Clinical trials in British medicine 1858 - 1948, with special reference to the development of the randomised controlled trial

Benjamin Toth

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Doctor of Philosophy in the Faculty of Medicine.

University of Bristol
Department of Social Medicine
November 1998

Word count 79,542

Abstract

This thesis arose from a desire to explore the reasons for two related phenomena: why randomisation wasn't introduced into clinical trials before the 1940s, and why in 1947 it was. In seeking to answer these questions, this study focuses on the development of clinical trials in Britain from the mid-nineteenth century to the MRC streptomycin trial published in 1948.

Control groups and quasi-random allocation and were known to British clinicians in the nineteenth century, well before they became formalised in statistical theory. Chapter two argues that the collective therapeutic enquiries of the 1860s were an attempt to reform the medical profession in the light of the deficiencies of the 1858 Medical Act. Reform in this context was an attempt to create an 'ideal practitioner', defined here as one who used a statistical style of knowledge to guide both medical practice and medical etiquette. However, methodological elements such as control groups were largely irrelevant to this enterprise.

Collective enquiries were overshadowed by the possibility of an exact knowledge of therapeutic action. Drugs such as diphtheria anti-toxin and Salvarsan, both products of German laboratories, mark the beginning of the modern era of therapeutics. Clinical trials played a secondary role in the development and testing of such drugs. Biological standardisation offered a powerful way for drug companies, research laboratories, and state authorities to promote and regulate the new laboratory based drugs.

At the British Medical Research Council, physiological research in a laboratory was regarded as the desirable way to test new drugs. As a result, clinical trials organised by the MRC in the 1930s are historically inconsequential. Nevertheless, they played a role in consolidating the position of the MRC as the leading medical research organisation in Britain at a time when pharmaceutical companies were beginning to produce a range of new therapies.

The organisational challenges facing the MRC in the immediate post-war period were changed considerably by the coming of the NHS. Adoption of a more strictly designed form of randomised controlled trial offered the MRC some specific advantages when dealing with a drug such as streptomycin. Using unpublished archival sources, the organisation of the streptomycin trial is reconstructed, and the organisational advantages of an RCT design highlighted.

Acknowledgments

I am greatly indebted to my supervisor, Professor Stephen Frankel, who has allowed this work to develop at its own pace, and who encouraged me at the outset to look more widely than I otherwise would have. At the Department of Health, Dr Russell Hamilton provided regular friendly motivation as only he can.

For any one in the field of medical history, the Wellcome Institute in London is indispensable. My friend and fellow student Alan Yoshioka has been very supportive, and kindly provided me with part of the transcript of his interview with Philip D'Arcy Hart. Also at the Wellcome, Dr Eileen Magnello convened an informal seminar on the history of clinical trials, and shared her work on Karl Pearson.

My thanks to the staff at the Public Record Office, especially in the reprographics department, who went to great lengths to help with copying material.

Closer to home, I am grateful to the staff of Library of the University of Bristol, especially Debbie Jones and Margaret Burke, who arranged for me to have extended access to archived volumes of the British Medical Journal.

Professor John Pickstone and Dr. Mark Jackson at the Wellcome Unit in Manchester enabled me to present work on the Therapeutic Trials Committee at their seminar in 1995. Professor Dennis Lindley kindly talked to me about RA Fisher, randomisation, and

Bayesian statistics. Correspondence with Professor Ian Hacking clarified the relationship between his work and that of Michel Foucault.

I owe the greatest debt to my family, especially during 1998 when I lived with this thesis as much as with them, and was finished a few days after my dear son Hugo was born.

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the Regulations of the University of Bristol. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree.
Any views expressed in the dissertation are those of the author and in no way represent those of the University of Bristol.
The dissertation has not been presented to any other University for examination either in the United Kingdom or overseas.

SIGNED **DATE** 20th November 1998

Table of Contents

List of Tables and Figures

Tables

Figures

Introduction

Purpose of this study

Several years ago the present author was among a group of health care researchers discussing the merits of different research methodologies. In the course of discussion it emerged that the first randomised controlled trial (RCT) in medicine was conducted only as recently as the 1940s.

The date, if correct, was surprising. The RCT is regarded as the *sine qua non* of clinical trial design. As a research method it is strikingly economical and powerful in its ability to evaluate therapies. Because of this it now seems, to a great majority of researchers and many clinicians, the intuitive way to evaluate the effectiveness of therapies. Alternative methods, if they exist, are almost unthinkable. Yet only a few generations ago the RCT was unknown to medicine.

Stuart Pocock's brief history in his textbook on clinical trials suggested that the date was correct.[1] According to Pocock the first rigorously organised randomised controlled trial in medicine was that of streptomycin as a treatment for tuberculosis of the lung, conducted under the auspices of the Medical Research Council (MRC), and published in the British Medical Journal in 1948.[2]

[1] Pocock 1983. Chapter 2.

[2] Pocock 1983 p17. The trial, described in more length in chapter 5 of this dissertation, published its report in the BMJ in 1948 (MRC 1948). The report is reproduced in Appendix 1.

Several other standard synoptic accounts of the development of the RCT in medical research are available.[3] Many of them refer to the streptomycin trial, although as with any 'first' there are other priority claims. A controlled of patulin as a treatment for the common cold was begun before the streptomycin trial and had many of its features.[4] Nevertheless, there is general agreement that the careful design of the streptomycin trial, especially the truly random allocation of patients to study and control groups, marks it out as especially significant.

The comparatively recent introduction raised a number of questions. I have chosen to focus on just two. Firstly, why wasn't the RCT introduced earlier in medicine, when we know from Hacking's important account that all the techniques were available to researchers in the early twentieth century?[5] Secondly, why did it suddenly become possible to randomly allocate patients in clinical trials the 1940s?

Like others who have looked at these questions, the answers proposed here are historical in character. The present study differs from most previous accounts of the history of RCTs in two principal ways however. Firstly, by and large previous studies took the form of wide-ranging but synoptic overviews. In contrast, the present study intentionally restricts itself to late nineteenth and early twentieth century Britain. To illustrate the effects of this restriction, the present study does not cover the development of clinical

[3] Bull 1952, Bull 1959; Lilienfeld 1982; Green 1954; Pocock 1983;. Spilker 1992; Lennard Jones 1993; Lock 1994, Gehan and Lemack 1995. Curiously, Bull's 1952 MD does not mention the streptomycin trial, even though Bull was a staff member of the MRC whilst completing his MD.

[4] MRC 1944. Proto-RCTs are reported in the eighteenth century and earlier. (see for example Lilienfeld 1982).

[5] The most important source on randomisation is Hacking (Hacking 1988) Standard accounts of the history of clinical trials emphasise the near emergence of RCTs in the work of Ambrose Pare in the sixteenth century and James Lind in the eighteenth century. Other aspects of clinical trial design have a long prehistory. For example Kaptchuk 1998 discusses the history of blinding and placebo controls in medical research.

trials in France in the early nineteenth century.[6] A second difference is that most synoptic accounts have been concerned purely with the lineage of ideas and therefore say little about the specific contexts in which clinical trials took place. Here I make positive use of the late nineteenth and early twentieth century medical world in seeking answers to why clinical trials became attractive and relevant to some parts of the medical profession.

To some readers this may suggest the sort of approach associated with relativism and the sociology of scientific knowledge. I do not entirely reject these categories but I take their main claim to be that ideas are the product of social settings and interests. This thesis does not support that view. The reality that binds together ideas, interests, ambitions, people, organisations, and societies is sufficiently complex to preclude any uni-directional mapping of interests onto knowledge.

Plan of this study

My original plan was to focus exclusively on the events in the decade or so before and after the 1946 streptomycin trial. By comparing the years between 1930 and 1939, when the RCT did not happen, with the years around 1948, when it did, I hoped to identify the changed conditions in which the RCT succeeded in impressing itself upon researchers.

Preliminary research, using the records of the MRC's Therapeutic Trials Committee held at the Public Record Office, Kew, suggested that this approach might be successful. Between the 1930s and the 1950s the standard methodology used in MRC sponsored trials did change significantly, from one that emphasised small case series to one which emphasised statistically designed controlled trials.

[6] Aspects of the history of nineteenth century clinical trials in France are discussed in Weisz 1995 and Rosser Matthews 1995. I am grateful to Dr Elsbeth Heaman for pointing out Weisz's study.

However, a focus on the immediate period around 1948, an approach which might be called 'before and after Bradford Hill', began to look less convincing, for two reasons.

Firstly, concurrent controls and random allocation were not new in the 1940s, even in medicine. More specifically, Philip D'Arcy Hart, a surviving member of the streptomycin trial organising committee, denies that random allocation was novel, or that the impetus for using it in the streptomycin trial came from Bradford Hill.[7] As Ian Hacking has demonstrated, the technique of randomisation was advocated in the nineteenth century, and by the 1920s had both powerful advocates and critics.[8]

Secondly, even allowing for the modesty of Bradford Hill, and his reluctance to emphasise statistical techniques in papers designed to be read by doctors, contemporaries do not appear to have regarded the design of the streptomycin trial as categorically different from what had taken place previously. The significance of the streptomycin trial therefore something of a modern construction, owing much to a subsequent generation of statisticians, clinical trialists and admirers of Sir Austin Bradford Hill.[9]

The story of how the streptomycin trial has come to assume a central position in the recent history of clinical trials is worth telling. But I have not tried to do so here. As I worked through the MRC archives it became clear that I should need to look back to a period before the 1930s if I was to understand the development of the RCT in Britain.

[7] D'Arcy Hart 1991 and D'Arcy Hart 1996 (interview transcript, available from author).

[8] Hacking 1988.

[9] Representative examples of this approach are Lock's account of the history of the RCT (Lock 1994) and Sutherland's account of the streptomycin trial (Sutherland 1998).

Methodology

The working assumption of this study is that at any one time questions of the design of clinical trials are also questions about how the medical profession should be organised. The most immediate methodological influence is that of Shapin and Shaffer, whose ground-breaking study of the Hobbes/Boyle dispute is a model of scholarship in the history of science.[10] More generally, I have drawn on Hacking's studies of the history of statistics. These provide a link to the figure of Michel Foucault, whose works are the presiding spirit of the house for historians and sociologists of science. Yet Foucault's works on methodology remain unused by most medical historians despite the fact that to study the relations that doctors establish among themselves and with society is also, inseparably, to study the creation and deployment of medical knowledge.

On this basis, I shall explore the ways clinical trials mediate between the clinicians who undertake them and other parties concerned with therapeutics, such as government, pharmaceutical manufacturers, and non-orthodox health care providers. Applying Shapin and Shaffer's maxim that 'solutions to the problem of knowledge are solutions to the problem of social order',[11] my approach is to look at how questions of method and questions of social and professional organisation are answered together. Looking at clinical trials may tell us something about medicine, and more precisely, about the way medicine should be organised and perceived.

The material I have used is drawn from two principle primary sources. The first is the British Medical Journal, used to describe the medical world into which a statistical form

[10] Shapin and Shaffer 1985.

[11] Shapin and Shaffer 1985 p332.

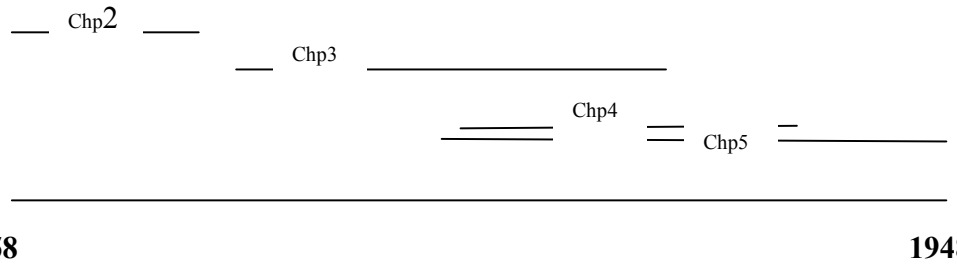
of knowledge was inserted in the nineteenth century. The second is the extensive archive of the Medical Research Council, now held at the Public Record Office, in Kew, London.

Chapter outline

The first chapter of this dissertation concerns the ways in which people have written about the history of statistics, since the recent history of clinical trials is in effect the history of the application of statistical method in medicine. The second chapter concerns the approach to therapeutic evaluation taken by the British medical profession in the late nineteenth century. The third chapter concerns the approach to therapeutic evaluation taken by the MRC in the years immediately before 1948. The fourth chapter looks at the early career of Bradford Hill, and discusses the MRC trial of anti-pneumonia serum. The fifth chapter concerns the streptomycin trial. The sixth chapter looks back over the others and draws out the conclusions of this study.

The chronology of the chapters tends to overlap, and there is a small gap between the end of the second chapter and the beginning of the third. The first chapter covers the modern historiography of statistics between the late-nineteenth and late-twentieth centuries. The second chapter focuses on the period between 1858 and the mid-1870s. The third chapter begins in the mid-1880s, but the bulk of the chapter concerns events in the 1930s. The fourth chapter starts in the early twentieth century and ends in the late 1930s. The fifth chapter is mainly concerned with the period 1946-48, but in describing the history of clinical trials in tuberculosis covers the period 1891 to 1948. I have set out the time periods covered by each chapter in Figure 1.

Figure 1: timeline of this study



What is a Randomised Controlled Trial?

The methodology employed in a RCT is straightforward and can easiest be described by discussing the three parts of its name in reverse order: trial, controlled, and randomised.

Wishing to know if a medical treatment is effective, we elect to carry out an investigation, or trial, to provide some empirical evidence with which to answer the question. The word trial has long usage in medicine, and in an analogous way to a legal trial refers to a formal attempt to establish the truth about the value of a therapy. In a clinical trial arrangements are made to allow a therapy to testify to its own efficacy, by studying its action on selected patients.

A set of patients is chosen for the trial. The members of this set are often called the subjects of the trial. It is customary today to obtain subjects' consent, but this was not an absolute requirement for most of the period being considered here. The set is divided in two lots or groups. One will receive the new treatment, the other an alternative, or an inactive treatment - the placebo. For example, if the new treatment were a drug given in the form of a yellow tablet, a suitable placebo would be a yellow tablet containing base but no active ingredient.

Like its legal counterpart, the purpose of a clinical trial is to bring evidence to bear. Unlike a legal trial, the RCT is comparative: effectiveness is decided by comparing the outcome in the group receiving the treatment with the outcome in the other group. The group receiving the new treatment is generally called the study group, while the group receiving the standard treatment or placebo is the control group. The control group provides a means for measuring the difference between the new therapy and a standard therapy or placebo. Clinical trials are controlled to the extent that they use control groups. The RCT is also concurrent. In an RCT the study and control groups are created at the same time. It is possible to compare a group of patients treated now with a group of patients in the past who were not treated. However good the matching, the use of so called historical controls is not encouraged because a variety of differences between the groups will exist, even if the patients in the two groups are carefully matched for diagnosis, age and prognosis.

Finally, a process of random allocation constructs the study and control groups. As subjects enter the trial they are assigned at random – randomised – to study or control group. The arguments for assigning patients at random, and some objections to it, are discussed in Appendix 3.

'The first RCT'

By present standards, the MRC trial of streptomycin as a treatment for tuberculosis, published in the British Medical Journal in 1948, was unusually fulsome in describing its methods and results. Conceived in 1946, the purpose of the trial was to reach a clearer understanding of the efficacy of a recently isolated anti-bacterial agent called streptomycin, which had shown promise as a treatment for pulmonary tuberculosis. The

trial was organised by a committee of the Medical Research Council (henceforward MRC), chaired by Dr Geoffrey Marshall, and received statistical advice from Professor Austin Bradford Hill, of the MRC's Statistical Research Unit.

Patients with confirmed pulmonary tuberculosis were recruited to a number of trial centres between January and September 1947. Patients were not told they were taking part in a research study. They were carefully chosen, according to the type and stage of their condition, and received the standard therapy for pulmonary tuberculosis, which at the time consisted of bed rest. On entering the study, patients were randomly allotted[12] to receive, or not to receive, 2 gms daily of streptomycin hydrochloride in addition to bed rest, given in the form of 4 intramuscular injections.

The results, presented as extensive tables and figures, showed clearly that streptomycin plus bed rest was a superior treatment to bed rest alone. (The report is reproduced in Appendix 1)

Recent studies of therapeutics and clinical trials

For many years therapeutics was a secondary topic within the history of medicine, even after Edwin Ackerknecht's call for more empirical research on the matter in 1967.[13] Similarly, the history of clinical trials attracted very little interest before the 1950s, after which time a series of synoptic histories were published.[14]

---

[12] The method of random allocation was fully described, unlike most present studies. It involved the use of sealed envelopes, opened in sequence, whose contents described whether the patient should be allotted to the study or control group. The envelopes were prepared by Bradford Hill, their contents determined by a table of random numbers.

[13] Ackerknecht 1967.

[14] For many years the standard work on the history of clinical trials was JP Bull's DM Thesis, submitted in 1950 (Bull 1951). There are several later studies, but they include little original material (Green 1954; Lilienfeld 1982; Lennard-Jones 1993; Armitage 1983; Vandenbroucke 1987; Gehan and Lemack 1995; Lancaster 1994 chapters 17 and 18).

The reputation of therapeutics as a suitable topic for serious interest has improved because of American scholarship. Firstly, Charles Rosenberg's work,[15] and subsequently John Harley Warner's study of the transformation of medical practice in America between 1820 and 1880. In Harley Warner's study:

> *Therapeutics is central to the professional image and legitimacy of physicians. Moreover, therapeutics, regarded as both a cognitive system and a set of social practices, is a useful indicator of the changing real and perceived roles of scientific knowledge in medicine. A study of therapeutic change, its determinants, and its meaning is thus a singularly productive means of assessing physicians' professional values and their perceptions of what constituted proper sources of knowledge*[16]

If one considers the themes that have interested many recent historians of medicine – knowledge, practice, professional identity – therapeutics would appear to offer a highly appropriate topic. Several recent studies have centred on questions of therapeutics with a view to exploring the relationship of knowledge, practice and identity in medicine. For example Armstrong's study of medical knowledge in twentieth century Britain[17]; Winter's study of mesmerism at University College Hospital[18]; Rusnock's study of James Jurin[19]; Weisz's study of the French Academy of Medicine;[20] and Oudshoorn's study of the commercialisation of sex hormone research[21]. The diversity of these studies precludes

[15] Rosenberg 1979.

[16] Harley Warner 1986 p1.

[17] Armstrong 1983.

[18] Winter 1991 and Winter 1994.

[19] Rusnock 1994.

[20] Weisz 1995.

[21] Oudshoorn 1995.

easy summary.[22] Nevertheless, as a collective sign of intellectual bearing they point to the revived fortunes of therapeutics as a topic for enquiry.

Even if the area of enquiry is restricted to the history of clinical trials, there are many recent scholarly works. Marks' recent study links reform in early twentieth century American medicine with the growth of co-operative clinical trials.[23] Tröhler's still unpublished thesis makes a strong case for regarding the numerical method as being established in British therapeutics much earlier, and much more extensively, than previously recognised.[24] Richards has discussed the politics of therapeutic evaluation by looking into the history of cancer clinical trials.[25] Further studies of the growth of quantification in medicine have been published by Cassedy[26] and Rosser Matthews.[27] Marcia Meldrum has completed a thesis on the development of RCTs after 1948.[28] AIDS clinical trials have been the subject of a major study by Epstein.[29] A study of the history of clinical trials of chemical contraceptives is currently in preparation by Lara Marks of Imperial College, London. Finally, the history of clinical trials is the topic of two recently

[22] Rusnock's study concerns seventeenth century England; Armstrong with twentieth century Britain. Weisz focuses on early nineteenth century France; while Oudshoorn's concerns twentieth century Holland.

[23] Marks 1997.

[24] Trohler 1978.

[25] Richards 1988.

[26] Cassedy 1984.

[27] Rosser Matthews 1992 and Rosser Matthews 1996.

[28] Meldrum's thesis, submitted to Stony Brook University in 1994, was unavailable at the time of completing this thesis. Two paper are available (Meldrum 1996 and Meldrum 1998).

[29] Epstein 1991, Epstein 1993, Epstein 1997.

completed theses, those of Alan Yoshioka at the Wellcome Institute in London, and Desiree Cox-Maksimov at Cambridge University.[30]

The importance of statistics

Much recent scholarship in the area of clinical trials is concerned with the issue of why, and how, evaluation couched in numerical terms came to the fore in the twentieth century. The explanations on offer, and the degree to which studies engage with the mathematics involved, varies.[31]

Nevertheless, the issue of quantification is central to any discussion of clinical trials. In particular, recent scholarship raises a question about the role of statistics. If the histories written by those involved in organising trials may be characterised as simply extolling the mathematical, quantifying virtues of statistics in medicine, historians and sociologists have tended to emphasise the importance of the social context in which statistics appear. Since statistics and mathematics are regarded as the disciplines least likely to be subject to the influence of their immediate context, the claim that statistics are in some way social requires special justification. Chapter one of this thesis is therefore an extended review of the history and philosophy of statistics.

Statistics play a central role in the history of clinical trials. Linked to the theme of quantification, the growth of statistical methods is used to organise a history that emphasises the slow separation of truth from error and the gradual turn towards more objective and rational ways to evaluate therapies. Statistics are however, arcane

[30] Thanks are due to Dr Eileen Magnello at the Wellcome Institute for convening an informal seminar in 1995 on the history of clinical trials.

[31] For example, Marks 1997 contains very little mathematics, while Rosser Matthews 1992 includes some of the mathematics associated with nineteenth century clinical trials.

mathematical constructions. Is it possible to write about clinical trials without writing about statistics? This approach has produced at least one notable study, that of the outcome of clinical trials of vitamin C as a treatment for cancer.[32] However tempting, to exclude statistics is to exclude from analysis the very feature that distinguished clinical trials. In the case of the Vitamin C controversy the ability of trials to produce objective knowledge was treated as a myth. Although this approach clears space for discussing the role of social forces in shaping the form of knowledge about therapeutics, it places significant elements out of reach.

Is it possible to account for therapeutic evaluation as 'inherently a social and political process'[33] while at the same time writing about at least some of the statistical aspects of clinical trials? The approach taken here is to treat statistical techniques as essentially a rhetorical resource. The bulk of this thesis examines the rhetorical effects of statistics in various settings between 1858 and 1948.

There is no single definition of rhetoric, but it can be broadly defined as the art and science of persuasion.[34] As such, rhetoric is something that both creates knowledge and in so doing sustains and produces advantage for the creator. Statistics are, arguably, the most transparent form of knowledge, the most modest in its assertions, the least able to conceal its workings. Statistics are concerned with objectivity, rigour and the removal of bias from argumentation. To characterise them as rhetoric therefore risks appearing unduly controversial, since in modern society the term rhetoric tends to be associated

[32] Richards 1988.

[33] Richards 1988 p 189.

[34] For a summary, see Barthes 1988.

with illegitimate forms of argument, or with verbal effectiveness as a mask of the true matters of the case.[35]

To defend the idea that statistics is a form of rhetoric, chapter one is a fairly extended review of the recent historiography of statistics. It argues that attempts to characterise statistical knowledge, particularly statistical inference, as an exact logical discipline have not been possible, and remain so. In the absence of a consensus about the logical basis of statistics, historical and sociological explanations for the authority of statistical method have developed. These have taken various forms, and three approaches are considered. Despite their differences, each finds a way of writing about statistics that makes it possible to appreciate both its logical construction and the role it played in specific historical and social contexts. The subsequent chapters explore the development of clinical trials in Britain since 1858, applying the methodology outlined in the introduction.

[35] An exception, a text written by a statistician that emphasises the rhetorical nature of statistical arguments, is Abelson 1995.

Chapter one

Approaches to the history and philosophy of statistics

Controversy in the theory and practice of statistics

Introduction

There is a large body of writing on the history and philosophy of statistics. At a conservative estimate, some 325 English language books and papers were published between 1910 and 1989 on the subject.[36] This figure is undoubtedly conservative. A more exhaustive search, and more permissive inclusion criteria,[37] might double that figure. My survey of this literature will be selective, and directed towards characterising only the main lines of its development.

Viewing the literature since 1910 as a whole, the increase in published output after 1950 is striking. For each published item before 1950, almost seven were published afterwards. While some of this growth will be a result of the general increase in published output in recent years, quickening interest in the basis of statistics after 1950 is due to a phenomenon described by Lancelot Hogben[38] in 1958 as 'the contemporary crisis in statistical theory'[39]. Hogben, although a Fellow of the Royal Society, and at the time Professor of Medical Statistics at the University of Birmingham, was not held in great

[36] See Appendix 2.

[37] I have tried to draw a distinction between two categories: literature which describes statistical methods and that which is about the underlying logic or theory or history of statistics. Only the latter is discussed here.

[38] Lancelot Hogben FRS (1895-1975). The major published sources on Hogben are GP (son of HG) Wells short biography (Wells 1978) and Werskey's study of left wing scientists in the 1930s (Werskey 1988). Hogben's papers, which include drafts of his recently published autobiography (Hogben 1998), are deposited at the University of Birmingham. He was a founding editor of the British Journal of Social Medicine, and published three papers on the methodology of randomised controlled trials in the 1950s. (Hogben and Wrighton 1952**a**; Hogben and Wrighton 1952**b**; Hogben 1954).

[39] On the title page of Hogben 1957.

esteem by professional statisticians contemporary with him[40] and there is little interest in him today outside specialist academic circles.[41] In retrospect though, Hogben can be credited with identifying the crisis and at least some of its causes.

Hogben's particular animosity lay with the unthinking acceptance of statistical methods by research workers. In the stylized language that was typical of his prose he described the condition of statistics in the 1950s:

> *It is not without reason that the professional philosopher and the plain man can now make common cause in a suspicious attitude towards statistics ...we witness on every side a feverish concern of biologists, sociologists and civil servants to exploit the newest and most sophisticated statistical devices with little concern for their mathematical credentials or for the formal assumptions inherent therein.*[42]

Hogben was a socialist as well as a scientist of high repute during much of his life. Linking his professional career with his political views, he wrote several popular science textbooks. In these he promoted the view that it was necessary for ordinary people to own, through understanding, the means of production of knowledge, if science were to serve socially useful ends. The pre-condition for social emancipation was therefore widespread scientific literacy.[43]

---

[40] Interview with Professor Dennis Lindley March 1996. For further light on Hogben, see Peter Medawar's autobiography. (Medawar 1986).

[41] For a recent study, which discusses Hogben, especially his ambiguous relationship to eugenics, see Mazumdar 1992, especially pp150-177.

[42] Hogben, 1957 p13.

[43] See for example his 1938 book 'Science for the citizen' (Hogben 1938), which he described as 'a self-educator based on the social background of scientific discovery'. Chapter 12 for example is entitled The Dark Satanic Mills – the superfluity of mere toil. It weaves together the invention of the steam engine, the theory of thermodynamics, and Marx's labour theory of value. It concludes by demonstrating the unscientific nature of the wage system and claims that science degenerates under capitalism. The epilogue makes the link between scientific progress and social emancipation particularly clear. Hogben's account owes much to Hessen's study of Newton presented to the International Congress of the History of Science and Technology in London in 1931 (Hessen 1971), with which he would have been familiar.

In the case of statistics, scientific literacy required research workers to understand the basis of the statistical tests they employed. On the professional statisticians, it placed a duty to be clear about the basis of the techniques they propounded. Hogben's major statistical work, published in 1957,[44] was based on two premises. Firstly, that most research workers were unaware of the theoretical derivation of techniques of statistical tests[45]. Hogben regarded this as evidence of the 'capitulation of the scientific spirit to the authoritarian temper of our time' caused by 'an increasingly widespread disposition of the younger generation of research workers to relinquish the obligation to examine the credentials of principles invoked in the day's work'.[46] Secondly, the absence, amongst the producing class of professional statisticians, of shared views about 'the fundamentals of their speciality at the most elementary level'[47] prevented real progress being made.

Two twentieth century controversies in statistical theory

Hogben's critique was motivated by his sense of the uncertainty about the intellectual foundations of statistical methods.[48] As a mathematically astute researcher in the 1950s he would have had no difficulty in identifying controversy among statisticians about the basis of their subject. Two disputes in particular will be referred to here. The more specific took the form of a dispute between RA Fisher,[49] the most eminent statistician of

[44] Hogben 1957.

[45] Hogben was reacting in print to the increasingly formulaic and unthinking (as he saw it) use of statistical tests of hypotheses in research. Danziger 1990 provides empirical evidence of this growth in the area of psychological research.

[46] Hogben 1957 p10.

[47] Hogben 1957 p13.

[48] Fisher did likewise at this time: 'It is no secret – it is a fact I have stressed particularly in a recent book of mine on scientific inference- that grave differences of opinion touching upon the nature of probability are at present current among mathematicians'. (Fisher 1958 p261).

[49] Ronald Aylmer Fisher (1890-1962) Born on 17th February 1890 in East Finchley. Youngest of eight children. Attended Gonville and Caius College. Wrangler in 1912. After Cambridge, he worked for the Mercantile and General Investment Company in London, then on a Canadian farm, before returning to England as a schoolteacher.

the day, and the almost equally eminent pair, Jerzy Neyman[50] and Egon Pearson[51], concerning the value of significance tests.[52] The dispute was highly personal.[53] It turned on an important issue concerning the relationship between statistical conclusions and scientific knowledge (see Appendix 3).

Confirming Hogben's critique of researchers, the dispute was not especially important to the users of statistical techniques, who drew, unknowingly, on a mixture of significance and decision testing, a hybrid which neither Fisher nor Neyman and Pearson approved.[54]

His interests in biology, genetics and eugenics resulted in early statistical publications, followed by an offer to work in Galton's statistical laboratory. He declined, in favour of a post at Rothamstead Agricultural Research Station, near Harpenden. Elected FRS in 1929. Became Galton Professor at UCL in 1933. Elected to Arthur Balfour Chair of Genetics at Cambridge in 1943. Retired in 1957, relocating eventually to Adelaide. For a biography, see Box 1978.

[50] Jerzy Neyman (1894 - 1981). Born in Bendry, Russia. Studied mathematics at the University of Kharkov, where he was subsequently a lecturer, before working as a statistician at the Bydgoszcz Agricultural Research Institute in Poland. Emigrated to England in 1933, where he worked with Egon Pearson. From 1938 Neyman lived in America, where he headed the Statistical Laboratory at the University of California, Berkeley. See Kendall 1984.

[51] Egon Sharpe Pearson (1895-1980). Son of Karl Pearson. Succeeded his father as Professor of Statistics at University College, London from the 1930s. For further information, see Bartlett 1981.

[52] The dispute concerned the value of Fisher's significance tests compared to Neyman and Pearson's' theory of decision testing. Most practicing statisticians adopted a hybrid of significance and decision testing, thereby avoiding the questions raised.(Gigerenzer 1989 p106) Seidenfeld argues that Neyman-Pearson theory has become the standard interpretation of statistical inference because it had a threefold advantage over Fisher's efforts to provide a theoretical justification. Firstly, the metaphor of repeated sampling proved popular; secondly it decisively rejected a subjectivist interpretation of probability; and thirdly, it could be expressed in clear mathematical language (Seidenfeld 1979) See Appendix 2 for a description of significance tests and decision tests.

[53] Fisher and Neyman clashed over several statistical topics. Fisher's biographer, his daughter Joan, follows Fisher's own explanation for the various Neyman-Fisher disputes: Fisher was an experimentalist who used statistics to improve the design of experiments; Neyman, according to Fisher, was a mathematician with little knowledge of experimental work. However, this explanation ignores Neyman's actual career, which was similar to Fisher's. Both had worked in applied research, in Neyman's case at the Bydgoszcz Agricultural Research Unit, Poland, in Fisher's at the Rothamstead Agricultural Research Station. Fisher Box's biography suggests a less auspicious reason why the two might have clashed: the failure to resolve the successorship to Karl Pearson at the department of applied statistics at University College London. Instead of appointing Fisher to run the department on Pearson's retirement, the department was split in two. Fisher was appointed Galton Professor of Eugenics; Pearson's son Egon, who had worked with his father for several years, headed a new Statistics Department. Neyman applied to both Fisher and Egon Pearson for a post. Fisher claimed he lacked funds to appoint him; in any event, Egon Pearson appointed Neyman as Lecturer in Statistics. (Box 1978, pp 257-266) In a late paper Fisher offered an explicitly ideological explanation for the opposing approaches. Whereas his own approach promoted disinterested science, Neyman's approach was associated with a soviet style utilitarianism that placed efficiency above truth (Fisher 1955).

[54] For Gigerenzer et al (Gigerenzer 1989) the debate about significance tests has been suppressed from 'the textbooks that have taught significance testing to the customer – the experimenter in the sciences.' Concerned with marketing an objective technique, textbooks have tended to censor the extent to which personal judgement necessarily informs statistical testing. Gigerenzer et al's conclusion may be too pessimistic. Viewed negatively, the absence of discussion may appear as an effort to produce an illusory sense of objectivity. Looked at more positively however, the exclusion

The proceeding generation of philosophically-minded statisticians such as Levi, Seidenfeld, Kyberg, and Hacking regarded the controversy between Fisher and his critics over statistical tests as highlighting basic deficiencies in the logical foundation of statistical inference.[55]

The dispute about statistical tests was quite specific. A more general controversy in the 1950s concerned the nature of probability. Two distinct schools of probability have emerged in the Twentieth century: one regarding probability as an objective part of the world[56], the other regarding it as a partly or wholly a subjective judgement.[57]

Something of the history of objective and subjective approaches to statistics is necessary at this point.[58] The subjective school is said to descend from the Reverend Thomas

---

of debate and controversy from standard textbooks may be no more than the outcome of a successful practice, in which a set of intellectual tools can be applied to situations in order to produce useful effects of quantification, precision and objectivity. Both interpretations can be applied to Neyman Pearson confidence intervals. To critics the interpretation of Neyman Pearson confidence intervals lacks rationale foundation, since it combines both frequentist and Bayesian elements . To supporters, this does little harm, given the psychological value of being able to say 'the true value lies between these values with 95% certainty' rather than the strictly accurate interpretation given by Neyman, which is obscure. For a discussion of the blurring of significance and decisions tests in epidemiology, see (Goodman 1993).

[55] Among the responses in Britain was that of Hacking, at the time a research fellow at the University of Cambridge (Hacking 1964; Hacking 1965; Hacking 1967; Hacking 1968) Seidenfeld 1979 provides a critical but supportive examination of Fisher's ideas. Kyberg has published several works on the foundations of statistics (Kyberg 1961; Kyberg 1974). Levi's main publication is Levi 1980.

[56] Often known as the frequentist school, because of the way it defines the probability of an event as its frequency in a long sequence. As Kyberg argues (Kyberg 1974), the 'frequentist school' consists of two parts. The logical part regards probability as determined by the application of a set of rules and assumptions to a set of data. For example the logical approach to the probability of throwing a six with one throw of a dice is as follows: there are six possible outcomes; assuming each is equally possible; the chance of a six is 1 in 6. The empirical part regards probability as determined by some empirical data. For a dice, the empirical argument would be: I do not know what the probability of a six is; I do not know if each outcome is equally possible. But if I roll the dice several times the frequency with which six appears will be [an estimate of] the probability of six. In simple examples the two approaches are indistinct. But in more complicated examples, and where the equi-possibility of outcomes can't be assumed, the approaches become distinct.

[57] The subjective school argues that the probability of an event is conditioned by prior knowledge about the event. This prior knowledge may vary from individual to individual, implying that the probability of an event may vary accordingly and cannot be wholly determined either by logic or empirical observation of outcomes.

[58] Discussions of the theory of statistics almost always include discussions of the history of statistics. In her study of Enlightenment statistical theory, Daston highlights the problem of doing so, and cautions statisticians who write histories not to overlook the empirical evidence of a diversity of approaches to defining probability in the

Bayes' posthumously published study of 1763.[59] The objective, or frequentist, approach became clearer after 1866.[60] However, distinct schools of probability did not exist in the eighteenth and nineteenth centuries. It is only in the twentieth century, in the wake of Fisher's rejection of Bayesian inverse inference (see Appendix 3), that the division into two schools becomes substantive enough to necessitate a choice, and for individuals to recognise the need to describe their basic alignment either as frequentist or Bayesian.

RA Fisher and controversy in statistics

Fisher proposed his theory of significance testing as an alternative to Bayesian inverse inference (described in Appendix 3). Since it did away with the need for prior subjective estimates of probability, significance testing created a powerful impression that a science of objective inference had been established,[61] heightening the apparent distinctiveness of the frequentist and subjectivist schools.

Despite his reputation as a leading statistician of the twentieth century, Fisher's work on statistical tests tended to produce controversy. Three reasons for this stand out.[62] Firstly,

---

Seventeenth century. Although convenient, the association of the subjective school with Thomas Bayes is unhelpful, since the connection between Bayesian statistics and Bayes' posthumously published 1763 paper is contentious. To further complicate the use of the name, Bayes is also associated with an entirely uncontentious theorem used to convert conditional probabilities, which was in fact first published by Laplace.

[59] Reverend Thomas Bayes (1702-61) Author of Essay towards solving a problem of chance, communicated by Dr Richard Price and published posthumously in the Transactions of the Royal Society in 1763. Bayes concern was to judge the chance that God existed on the basis of the state of the world. This approach amounted to judging (the probability of an unknown) cause from (observed) effects. Bayes solution required an estimate of the probability of the cause before the evidence was gathered. This, the so-called prior probability, places an element of subjectivity at the heart of judgements under uncertainty. To many statisticians in the early Twentieth century, including Karl Pearson, Ronald Fisher, and Jerzy Neyman , the value of statistical inference lay in the removal of subjectivity as far as possible from any role in judgements.

[60] The date of publication of John Venn's The Logic of Chance (Venn 1866). As with almost all concepts in statistics, a trail of precursors exist. Venn's definitive work was foreshadowed by that of Mill, Ellis, Cournot and Fries, who proposed frequentist interpretations of probability in the 1830s and 40s.

[61] The success of the objective approach is illustrated in Danziger's study of methods used in psychology (Danziger 1990).

[62] Joan Fisher Box brings out these strands in her biography (Box 1978).

as a synthesis, Fisherian significance testing represented a genuinely new approach, with immediate and lasting consequences. It was therefore subject to much attention from his colleagues. Secondly, Fisher tended to regard public controversy as a way to clarify lines of thought. He therefore courted controversy throughout his life.[63] Thirdly, although the techniques he introduced worked, and he was mathematically trained, Fisher tended to have solutions more readily to hand than their logical verification. Anecdotes reported by his daughter in her biography show Fisher happily solving complex problems for colleagues, but skipping their derivation.

Fisher's work was susceptible to challenge at several points. His colleague WG Gossett (who published under the pseudonym 'Student') challenged the need for randomisation in field trials.[64] Neyman and Pearson criticised significance tests for lacking a logical foundation.[65] Fisher's concept of fiducial probability, which was central to his thinking on statistical tests, has been described as 'merely a poorly sketched technique'.[66] And though much of his work was inspired by his wish to rid experimental design of Bayesian elements, one of the effects of Fisher's work was to create a new generation of subjectivist statisticians.[67] As a result, by the 1950s it was possible, as Hogben did, to

[63] For example, he challenged the view that smoking caused lung cancer by questioning the results produced by Bradford Hill and Doll in their epidemiological study. See Fisher 1959.

[64] The dispute is discussed in Egon Pearson's biography of Gossett (Pearson ES 1990) and Box's biography of Fisher (Box 1978).

[65] Neyman and Pearson 1928a and 1928b.

[66] Seidenfeld 1979 p106.

[67] Although seemingly eclipsed by Fisher, subjective probability survived. First with JM Keynes (1920s), then with Harold Jeffreys (1930s), and later with Rudolph Carnap, LJ Savage (1940s/50s), Dennis Lindley, Jack Good and de Finetti (1950s/60s/70s), a body of subjectivist or Bayesian statistics developed. By the 1950s there was a growing body of statistical theory which took the subjectivist position as its defining characteristic. It therefore offered an alternative interpretation of probability to the by then established view. The perceived challenge to the established view was all the greater because subjectivist statistics arose (so it was claimed) as an attempt to overcome the limitations of objectivist statistics.

characterise statistical theory as controversial and seriously divided along a fault line between frequentists and Bayesians.

Responding to the 1950s crisis in statistical theory

If we imagine philosophical discussion about statistics framed in a single question, the question would be 'how can reliable knowledge be secured from a limited empirical base?' This question was central to both the Neyman/Pearson–Fisher and the frequentist-Bayesian disputes: what is the underlying rationale and what are the prerequisites that permit a set of time-limited and local data to support or refute a statement about the world?

Judged by the growth in papers and books concerning the basis of statistics since the 1950s, the question of the reliability of statistical knowledge has become increasingly important. What the growth in output does not immediately reveal is the extent to which it contains new answers to the question of how reliability is achieved. Alongside continuing attempts to provide a rigorous proof for statistical methods, there has been a growth in historical studies. To illustrate, the next section follows the solution developed by Ian Hacking. I argue that a fully worked through justification for statistical methods seems unlikely. Consequently, historical explanations for the effectiveness of statistics have been proposed. Historical studies vary in their approach to incorporating historical factors. The approach that seems to work best is that which regards statistics as a 'style of knowledge formation', advocated by Hacking.

Ian Hacking's studies of the logic and history of statistics

The philosopher Ian Hacking is the leading British authority on both the logic and the history of statistics. His first substantial published work was a critical review of the logic

of statistical inference.[68] In this account, twentieth century theoretical statisticians have made progress in describing the nature of probability. Andrei Kolmogorov for example, provided a mathematical description of the basic rules of probability. Kolmogorov's axioms are in effect a set of rules for the formation of correct statements about probability. They are widely accepted as the theoretical basis of probability by statisticians who disagree about the nature of statistics, in conformance with Kolmogorov's assertion that his axioms supported 'an unlimited number of concrete interpretations' of probability.[69]

Hacking observes that Kolmogorov's axioms make no claim to providing a theoretical justification for statistical testing however. Nor do they suggest any way to construct rules for the formation of legitimate statements about statistical inference. Several authors, including Fisher, have written about these topics.[70] However, if one adopts the viewpoint of a philosopher none of the theories of probability or of the logic of statistical inference are wholly satisfying:

> *Neither frequentists nor subjectivists have been right about probability, but to discover exactly what there is to probability one needs a very different type of analysis than anything found in this essay…some say the word* [probability] *and its cognates are entirely vague, but our ability to use them regularly suggests they must be governed by some pretty stern regularities, which mark out the concept, or concepts of probability.*[71]

[68] Hacking 1965.

[69] Kolmogorov 1950 p1. Axiom IV for example states that the probability of a of a sure event, P(E), is 1. Kolmogorov's approach is regarded by Lucas as the paradigm of all axiomatic approaches. Lucas 1970 p 28.

[70] Hacking describes the problem as the connection between the frequency of an event and the truth of an assertion. Fisher used the term likelihood to describe his approach. See Hacking 1965 chapters 3-5.

[71] Hacking 1964 p227.

Hacking's views in 1965 can be summarised as follows: there is controversy about the underlying nature of statistical inference, and there appears to be little hope of resolving these issues. Despite uncertain foundations, statistical knowledge is not illusory. To understand why and how reliable knowledge can exist in a system whose rules for the formation of exact knowledge is incomplete requires some other form of explanation.[72]

The 'pretty stern regularities' referred to by Hacking in 1965 were set out in 1974, in a historical study of the origin of the modern concept of probability. In *The Emergence of Probability* Hacking argued that the single most important feature of the concept of probability is the simultaneous emergence of two forms of probability. Hacking calls the forms aleatory and epistemological; the aleatory being probability considered as stable frequencies of chance set-ups like a dice; the epistemological being the assessment of degrees of belief in propositions ("There is a good chance it will rain later tomorrow").

[72] Wittgenstein had a linguistic system in mind when he wrote: 'All testing, all confirmation and disconfirmation of a hypothesis takes place already within a system. And this system is not a more or less arbitrary and doubtful point of departure for all our arguments: no: it belongs to the essence of what we call an argument. **The system is not so much the point of departure, as the element in which arguments have their life**' Wittgenstein 1969 para 105. My emphasis.

The aleatory and epistemological views correspond to the modern frequentist and subjectivist schools of probability. The two schools are usually regarded as fundamentally at odds with one another, but Hacking argued that the opposition is illusory.[73] It will never be possible, decisively, to choose between aleatory and epistemological probability. They exist together or not at all, since they share the same fundamental assumptions about the nature of evidence:

> *It is better to expose the crudities of one's model at the start, than to conceal a methodology in banal phrases. I am inviting the reader to imagine, first of all, that there is a space of possible theories about probability that has been rather constant from 1660 to the present. Secondly, this space resulted from the transformation upon some quite different conceptual structure. Thirdly, some characteristics of that prior structure, themselves quite forgotten, have impressed themselves on our present scheme of thought. Fourth, perhaps an understanding of our space and its preconditions can liberate us from the cycle of probability theories that has trapped us for so long*[74]

The stern regularities, whose nature was only hinted at by Hacking in 1965, were in 1974 identified as a set of rules (called a conceptual structure) which outline a space of possible knowledge and determine the form and content of true statements.

His examination of the formation of the modern conceptual structure for statistics proceeds along the following lines. Before the seventeenth century probability was very much a second-rate sort of concept, belonging to the realm of opinion rather than knowledge. A fact had probability if it was approved by authority, if it was sanctioned by

[73] In one essay he argued that the opposition between frequentists and Bayesians masks 'the regimentation of reason that is so characteristic of … the calculus and language of probability Hacking 1992 p153.

[74] Hacking 1974 p16.

opinion. Hacking illustrates this with a nice quotation from Gibbons' *Decline and Fall*: 'such a fact is probable, but undoubtedly false'.[75]

A transformation in the late Renaissance concept of the 'sign' that made it possible for a thing (rather than a person or a book) to testify to the truth of something else, created the concept of empirical (rather than absolute or rational) knowledge.[76] A second transformation established that causality was no longer a property of absolute knowledge but of opinion.[77] This second transformation created the skeptical problem of knowledge, first set out by the Scottish philosopher David Hume. How do we acquire beliefs about things we are not currently experiencing? How do we know causes? We see a flame for instance, and conclude that it is hot. Hume notes that we start from a present impression – the sight of the flame – and suppose a causal relation – between flames and heat. But how do we come to believe in causal relations? Hume's claim was that it is not because of reasoning from first principles [i.e. absolute or rational knowledge]. But nor can empirical experience provide universal knowledge.

Hacking suggests that there emerged in the seventeenth century a new form of empirically based knowledge, the form we live with today, the form to which Hume's problem still applies. The transformations created the skeptical problem of knowledge, to which the modern concept of probability is a solution. The best that can be achieved in

[75] Hacking 1974 p19.

[76] Shapin has recently summarised the key events of this period in an overview of the scientific revolution. (Shapin 1996) Chapter 2 is especially relevant. It describes the clash between ancient and modern sources of knowledge, between the ancient texts and the book of nature as competing sources of knowledge. One of his illustrations is directly relevant. The Swiss magician and healer Paracelsus insisted that 'those who sought medical truths should put aside the ancient texts and take themselves directly to the study of herbs, minerals and stars … if I want to prove anything I do so not by quoting authorities but by experiments and reasoning' (Shapin 1996 p69).

[77] Hacking 1974 p180.

the circumstances first noticed by Hume are probable relations of causality, based on empirical evidence.

Probability, in its modern form, is the 'glue' that joins the evidence afforded by signs to knowledge. At a foundational level probability must always have two aspects if it is to function in this way: one concerning the frequency of signs (the aleatory); the other describing the degree of credence that exists between signs and knowledge (the epistemological). This, much condensed, was Hacking's answer to the question posed at the conclusion of *Logic of statistical inference*.

Responses to Hacking's studies

The line of argument in *The Emergence of Probability* was contentious among statisticians.[78] But overall, the effect of a publication by a leading logician that argued for the historical nature of fundamental statistical concepts must have suggested that a broadly social rather than mathematical frame of explanation for the history of statistics was plausible.[79] Without claiming that all subsequent discussion derives from it directly, recent scholarly work on the history and philosophy of statistics can be conveniently grouped into the themes arising from the studies Hacking published up to 1974.

[78] Garber and Zabell 1979. Although Hacking's periodisation matched that of the more conservative statistician Maurice Kendall, who judged the history of statistics to have begun in 1660 (Kendall 1960).

[79] Hacking was not the only person in the 1960s to suggest that reliable knowledge might be produced within structures it did not itself wholly determine. Publishing history of the 1960s points to a widespread desire to question the claims of scientific knowledge to be amenable to analysis only according to the hypothetico-deductive model. Thomas Kuhn published a general critique - *The Structure of Scientific Revolutions* - in America in 1962. (Kuhn 1962) It was originally published as part of the International Encyclopaedia of Unified Science, a series edited by Rudolph Carnap, who had written extensively on the theory of probability. Michel Foucault published a case study in the history of medicine - *The Birth of the Clinic* - in France in 1963 (Foucault 1973), and a general model - *The Order of Things* - in 1966 (Foucault 1970). Jurgen Habermas published a critique of positivism - *Knowledge and human interests* - in Germany in 1968, based on a lecture series of 1963/64 and his inaugural lecture of 1965. (Habermas 1972).

Is a logic of statistical inference possible?

Given that he is by training a logician and that he has made considerable efforts to derive a logic, Hacking's contention that statistical inference is not, in the strict sense, logical, must be regarded seriously. Recent literature appears to support this view. The dispute between the philosophy, if not the practice, of Bayesian and frequentist statistics remains as evident as ever was. More surprisingly, randomisation in experimental work, which is so secure as a practical technique, continues to lack a universally accepted logical basis.[80]

Continuing controversy about randomisation

Random allocation of subjects to study and control groups is the *sine qua non* of clinical trials, but is the element of trial design most often subject to controversy. For clinicians randomisation is particularly open to question on ethical grounds, since it involves a potential breach of the Hippocratic oath.[81] For some, randomisation epitomises the worst excesses of rationality, in which humans are rendered passive anonymous subjects of research.[82]

Despite a range of objections, random allocation in clinical trials is emphatically endorsed by researchers and textbooks as a technical device without which the results of research lack credibility. The justification for random allocation of subjects can be made on two levels. The first is more or less intuitive; the second is part of formal statistical theory.[83]

---

[80] For a survey on the debate and an assertion of the vitality of a Bayesian approach see Howson and Urbach 1993.

[81] Ethical arguments concerning RCTs are reviewed in Ashcroft 1997.

[82] Susan Lederer reviews the recent history of human experimentation from this perspective (Lederer 1995).

[83] The distinction may seem trivial; but it lies at the centre of the one of the major controversies in the field of clinical trials, the University Group Diabetes Program, a trial which was intended to make exemplary use of methodology. The interpretation of the relevance of differences in baseline risk factors found after the trial was stopped was

The intuitive explanation of the necessity for randomisation is that it is a way to ensure that the study and control groups are comparable. The following example illustrates the importance of comparability in clinical trials. Suppose the study group in a clinical trial were to be composed of subjects less ill than the subjects in the control group. A finding that the new treatment resulted in fewer deaths than the old treatment would be erroneous, since the treatment was being given to people with an *a priori* better chance of survival. We have biased our study against the old therapy.

We wish to conduct a fair trial. But there are many sources of this type of bias. If we allotted the first fifty subjects to the new treatment and the next fifty to the old treatment we might have arranged things so that some important characteristic were over-represented in one group. Or a researcher might consciously or unconsciously assign patients to study and control groups in favour of the newer treatment. One way to avoid the problems associated with establishing study and control groups is to base the allocation decision on the results of a process outside the control of the researchers: *random* allocation of subjects to study and control groups. A simple way to randomise subjects in a clinical trial would be to base the allocation decision on the results of a coin toss. Heads – treatment group. Tails – control group, or vice versa. There are practical objections to the use of simple randomisers, and more sophisticated methods are available. Tables of random numbers (that is, tables of numbers which have been selected according to a random process) can be used to determine the allocation of subjects to study and control groups (Figure 2). Another feature of modern trial design is the division

dependent on the view about the purpose of randomisation. For a fuller account see Marks 1997 chapter 7 especially pp211-213. For a summary of Alvan Feinstein's views see Feinstein 1977 p113.

of labour between who decide whether or not a patient is eligible to enter a study and those who assign random allocations.

Figure 2

Part of a Table of random numbers

The intuitive explanation – that random allocation ensures comparable groups –is only partly correct. While it is convenient to explain the need for randomisation in this way, random allocation cannot ensure that the characteristics of the study and control groups are the same. For example, it is rather unlikely that a random allocation process would divide a sample of 200 people composed of equal numbers of men and women into 2 groups each of which contains 50 women and 50 men. What random allocation does accomplish is the distribution of all characteristics of a sample between two or more groups in a chance way.

It is often seen upon inspection that two groups formed by random allocation are broadly balanced across a range of variables. For example if 2000 people were randomly allocated to two groups the average age of subjects in each group, the average height and weight, the average blood pressure, and the average IQ would probably be approximately the same in each group. Accounts sometimes imply that the purpose of randomisation is to achieve balanced samples.[84] It is expedient to imply this, but as is easy to demonstrate, it is only partly correct to do so. For any group of human subjects there will be hundreds of characteristics. It is rather likely that if one hundred characteristics were to be examined following random allocation, some of them would be distributed between the groups in an un-balanced way.[85] There are ways to ensure that characteristics that are known to be relevant to the outcome of a clinical trial can be balanced (through a

[84] Bradford Hill did so in the account of the streptomycin trial (MRC 1948) Oakley defends randomisation using the argument that 'the method allows researchers to evenly distribute both those factors which are known to be associated with different outcomes and those which may be, but are unknown' (Oakley 1990) p174. It would be more accurate to say that the purpose of random allocation is to allocate the factors randomly. But since this sounds like a tautology it isn't used as a justification.

[85] In their study of the efficacy of randomisation in clinical trials of acute myocardial infarction, Chalmers *et al* found that a relevant variable was mal-distributed in 14% of 'well-blinded' randomised controlled trials (Chalmers **T.C**. 1983).

technique called stratification). But stratification does not resolve the basic problem, since we cannot know beforehand all the characteristics that are relevant to outcome.

Turning to a more formal justification for the necessity of randomisation, one leading statistician has written:

> *Still deep within me, I have the feeling that the interpretation is clearer, the conclusions are stronger and the analysis has greater validity if treatments have actually been assigned at random. But why? I do not know, and none of the explanations which have been advanced are totally satisfactory to me.*[86]

For Fisher the act of randomisation assured the validity of the chain of inference based on significance tests carried out on samples drawn from a population. The level of significance *p* of the sample X in relation to a null hypothesis is, by Fisher's definition, the probability of drawing a sample from the reference set R which is as or more discrepant from the null hypothesis than X. (see Appendix 3 for a description of significance tests). The probability is only true over the reference set R provided there are no recognisable subsets of R to which X might belong. Since there must be subsets of R to which X will belong, the focus of effort to ensure epistemological validity must be on ensuring that it is not possible to recognise whether or not X is a member of R or a sub-set of R. Non-recognition can be achieved by being in a state of ignorance about the sample:

> *The necessary ignorance is specified by our inability to discriminate any of the sub-aggregates [sub-sets] having different frequency ratios, such as must always exist*[87]

[86] Folks 1984 p30.

[87] Fisher 1956 p35-36, quoted in Johnstone 1989.

Randomisation is the physical act which guarantees the ignorance required to legitimise inferences made using samples.

From a Bayesian point of view, Lindley has stated that randomisation is irrelevant to the design of experiments,[88] and Savage has written that 'randomisation is without value for statistics'[89] a view supported recently by Howson and Urbach.[90] In Kadane and Seidenfeld, randomisation is valuable, but only insofar as it acts as a rhetorical device for creating a sense of trust in the reader[91]

Even outside the subjectivist framework, doubt about the logical value of randomisation persists. In 1980, Levi set out his objection thus:

> *Fisher apparently believed that through randomization, information about the kinds of trial which could not be otherwise ignored could be ignored. And there is, indeed, a sense in which Fisher is right. Prior to finding out the results of randomization, the chance distribution of the t-statistic on the null hypothesis is perfectly determinate and if the experimenter could ascertain the value of that statistic without ever finding out which plots were selected for which treatment as a result of randomization, direct inference would justify a numerically definite likelihood for the null hypothesis… In practice, however, information as to which plots received which treatments is known and, indeed, used when computing the value of the t-statistic. And once it is known, the mere fact that the treatments were assigned at random contributes nothing to establishing the irrelevance of the information…. Thus it seems to me that the technique of artificial randomization has very little to recommend it, insofar as the rationale for it is the one which Fisher… offered.*[92] [emphasis added]

[88] Lindley 1992 p436. Howson and Urbach provide a re-statement of Bayesian objections to randomisation (Howson and Urbach 1993).

[89] LJ Savage, 1961, quoted in Kadane and Seidenfeld 1990.

[90] Howson and Urbach 1994 p 261.

[91] Kadane and Seidenfeld 1990.

[92] Levi 1980 p301-302.

In a review of the topic in 1989 Johnstone concluded that random samples are 'nice but not necessary'[93]. In fact, objections to randomisation are as old as the justification for randomisation. Fisher's colleague, the Guinness statistician W. G. Gossett, outlined one objection in 1937, when he argued that in the case of agricultural experiments balanced plots were more effective than randomised allocation.[94]

If logic cannot decide, can empirical study establish the value of randomisation? Analysis of trial results suggests that clinical trials using strict randomisation tend to find weaker treatment effects.[95] This has been taken as empirical support for the use of randomisation. However, the argument only has any force if one assumes what one intends to demonstrate, namely that the results of RCTs are closer to the true effect of the therapies under test than those of non-RCTs.[96] Empirical results cannot therefore provide a justification of equal rigour to logic concerning the necessity of randomisation.

In summary, even on a point so apparently settled as the need for randomisation, it is easy to find authors arguing strongly that insistence on its use is insufficiently warranted from a logical point of view. One leading biostatistician has argued that 'the idea of

[93] Johnstone 1989.

[94] Gossett 1937. For a summary of the position written by a leading medical statistician see Feinstein 1977 chapter 8.

[95] Chalmers **TC** 1983.

[96] In their recent overview of comparisons between randomised and non-randomised clinical trials, Kunz and Oxman assume that the true effect is that measured by the RCT. (Kunz and Oxman 1998) Hacking described the circularity of arguments used to justify statistics as themselves part of the statistical style of reasoning: 'the truth is what we find out in such and such a way. We recognise it as truth because of how we find it out. And how do we know that the method is good? Because it gets at the truth' (Hacking 1992 p135) Kunz and Oxman found that on the whole **non**-RCT studies tend to overestimate the effect of a therapy. Their explanation is that patients with a poorer prognosis tend to get selected for control groups in non RCTs. An alternative explanation is that where clinicians have some opportunity to consciously or unconsciously select patients likely to benefit from a new therapy, they are able to do so.

randomised allocation will retard therapeutic progress, however, if randomisation continues to be regarded as the principle ingredient in that progress'.[97]

At the risk of appearing to curtail further examination, from this point on I will assume that the claims of statisticians to provide objectively reliable knowledge cannot be grounded in formal logic. This does not mean that statistics are illogical. By widely accepted criteria, statistics provide the most reliable guide to the efficacy of therapies. However, the fact that these criteria are not timeless points to the possibility of studying the development and success of statistics without assuming that development and success are accounted for wholly within the terms of the self-understood logical superiority of statistical method.

How are statistics contingent upon circumstances? – A review of recent histories of statistics

Many statistical techniques are readily seen to have developed in response to practical problems. Fisher and Pearson developed their best known techniques – significance testing and the chi squared test respectively – whilst solving agricultural and biometric problems.[98] WG Gossett, better known as 'Student' spent much of his working life in the employment of the Guinness family, supporting work on crop yields. Some form of relationship between statistical techniques and material circumstances is therefore very plausible. Nevertheless it is still a question to decide what that relationship is.

History of statistics as the history of ideas

[97] Feinstein 1977 p120.

[98] Fisher's biography is again the best source. For many years he was the statistician at Rothamstead Agricultural Research Station, which undertook trials to increase the yield of crops through plant breeding and soil fertilizers. Pearson provided mathematical help to FR Weldon, who undertook experimental studies of natural selection. For details, see Magnello 1993.

The simplest possible history would be purely mathematical. It might trace the development of a technique like the computation of the average, by looking for the published occurrence of formulae which manifestly produce the arithmetic mean of several numbers.[99]

The first modern history of statistics, published by Isaac Todhunter in 1865, *was* largely a mathematical treatment of the development of statistics.[100] However, since Karl Pearson's lecture courses on the history of statistics (given at University College London between 1921 and 1933) it has been customary to emphasise the influence of practical and social issues on the development of statistical techniques.[101] This is done partly to enliven the subject, and partly to show the extent to which statistics is part of the real world. The social factors associated with the development of statistics are wide ranging, including gambling, jurisprudence, life assurance and concern for public health.[102]

Pearson's history made extensive use of these factors, but his lectures were organised around individual mathematicians, judging their contribution to the development of statistics against a background of contextual factors. This approach to the history of statistics creates two sorts of difficulty. The first concerns the precedence of ideas. Occasionally it is easy to see who first proposed an idea or a formula. More often though

[99] An account of this sort was published in the late 1950s. (Plackett 1958).

[100] Todhunter 1865.

[101] The lectures were eventually published in 1978, having been edited by Pearson's son Egon. (Pearson 1978) Pearson was a social thinker, as well as a leading theoretical statistician. He was also a biographer of Francis Galton. See Magnello 1996. The lectures are discussed in Hacking 1981.

[102] Bradford Hill, for example was fond of telling how it was the shortage of streptomycin that allowed him to impose a strict regime of random allocation in the 1948 streptomycin trial. In 1954 Fisher hinted at a greater role for social factors when he stated that 'this…overflow of statistical techniques from the quiet backwaters of theoretical methodology… into the working parts of going concerns of the largest size, suggest that hidden causes have been at work…, preparing men's minds, and shaping the institutions through which they work (Fisher 1954 quoted in Gigerenzer 1989 p70).

it is difficult to know how much of a new formulation is genuinely original and how much is the conscious or unconscious repetition of what has already been formulated.

For example, the normal curve, so named by Karl Pearson in the early twentieth century, is sometimes called the Gaussian distribution, in honour of Karl Friedrich Gauss. However the normal, or Gaussian, curve can be traced back to Laplace, or still earlier, to De Moivre.[103] Further examples of uncertain precedence in statistics abound. Random sampling may be a twentieth century idea, but Laplace anticipated it.[104] The American philosopher CS Pierce advocated the use of random allocation in experiments 50 years before Fisher;[105] Borel anticipated Ramsay's rational decision theory;[106] Radicke described a two sample means test 70 years before Fisher.[107] Bayes didn't formulate Bayes theorem.[108] In summary, almost every great idea in statistics has been anticipated or mis-attributed, sometimes so much so that the lineage of ideas is difficult to follow.

The second difficulty confronting a history of ideas approach is the transmission of statistical methods between disciplines. Why did statistical methods flourish in astronomy in the early nineteenth century but find no role in physics until the twentieth? Why were statistical methods applied to therapeutics in France in the early nineteenth century, but ignored by British medicine for 50 years or more? On this matter Pearson has little to say, since he built his account by stepping from one progressive individual to

[103] Pearson 1971 p 156-158.

[104] Stigler 1986 p164.

[105] Hacking 1988.

[106] Knobloch 1987 p215.

[107] Coleman 1989 p208.

[108] Stigler 1982; Stigler 1983.

another. In doing so, the differential rate with which statistical ideas developed in particular disciplines remains concealed in Pearson's history of statistics.

Resolving the difficulties of the history of ideas approach

Resolution to problems of precedence and transmission may come from taking greater account of the context in which statistical techniques developed. Contextual factors explain why, for example Galton, rather than Quetelet, formulated correlation, despite the mathematical techniques necessary being available to both, or why statistical approaches took hold in psychology before they did in sociology.

Context may be incorporated into the development of statistics in several ways. To compare the ways in which recent historical studies include contextual factors I will consider the approach to the development of correlation taken in three recent studies. These are Stephen Stigler's *The History of Statistics*[109], Theodore Porter's *The Rise of Statistical Thinking*,[110] and Ian Hacking's study of statistics *The Taming of Chance*.[111]

The history of correlation

Correlation is now seen as a useful if unremarkable concept in statistical analysis. It is based on the assumption that two population variables, for example height and weight, may be associated, so that as one variable changes so does the other. The extent to which variables change in line with each other can be measured using a simple formula, the result of which indicates the degree of correlation of the variables.

Despite the concept now seeming straightforward, and the mathematics needed to work out correlations between variables being readily available in the early nineteenth century,

[109] Stigler 1986.

[110] Porter 1986.

[111] Hacking 1990.

the mathematical technique of correlation did not appear till the late nineteenth century. At that point correlation was developed rapidly, first by the English polymath Francis Galton, then by Francis Edgeworth,[112] George Udny Yule[113] and Karl Pearson.

In his 1986 monograph *The History of Statistics*, Stigler rejects the idea that the delay in developing the theory of correlation might be due to intellectual inertia.[114] He argues that the key event was Galton's transformation of Quetelet's work on the normal distribution of population variables.[115] During the 1830s Quetelet published measurements of the chest circumference of 5,732 Scottish soldiers and found height to be distributed in what we would now call a normal curve.[116]

Quetelet, in Stigler's account was content to find normal curves and had no reason to develop his work beyond constructing 'normal man'. Galton was more ambitious. From his perspective, as a middle-class nineteenth century Englishman, he wanted to know how to distinguish racial groupings empirically. Quetelet's work showed him how this might be achieved. In *Hereditary Genius,* Galton argued that if a set of anthropomorphic data conform to a normal distribution, then it comes from persons of the same racial type. If it did not, there must be racial admixture.[117]

---

[112] Francis Ysidro Edgeworth (1845-1926). Edgeworth's biography is discussed briefly in MacKenzie 1981 p97-98. From 1891 Edgeworth was Drummond Professor of Political Economy at Oxford. In the early 1890s Edgeworth worked on the statistics of hereditary, but appears to have turned away from statistical work thereafter.

[113] George Udny Yule (1871-1951). Student of Karl Pearson. Later rejected Pearson's eugenic views.

[114] Stigler p 239.

[115] The argument forms chapter 8 of Stigler 1986. Galton is discussed at length in a recent study of fertility in Britain (Szreter 1996).

[116] The table is reproduced in Stigler 1986, p207 The data on chest circumferences were originally published in the Edinburgh Medical Journal of 1817 (see Hacking 1990 p109).

[117] Galton 1869, Quoted in Stigler 1986.

Having applied the normal curve to physical data, Galton extended his analysis to something that he was really interested in – individual talent. He argued that if 100 individuals from the same race were selected they could be measured for talent, and that a normal distribution would emerge.[118]

Galton next investigated how a normally distributed characteristic could be transmitted from generation to generation. In confirming that height was indeed passed from parents (Galton used the abstraction of mid-parent to represent the height of both parents) to offspring, Galton hit upon a mathematical formula for expressing the relationship between the height of a mid-parent and the height of their offspring.[119] The subsequent mathematization and generalization of Galton's work provided the human sciences with a powerful tool for measuring first regression and subsequently correlation between population variables.

Stigler's argument is that Galton's interest in heredity and the classification of racial characteristics allowed him to use Quetelet's work in a new way. The overall thesis of *The History of Statistics* is that individuals, working with materials that come to hand, using the context and prejudices of the day,[120] create methods that will eventually transcend the immediate context of their creation and the prejudices that informed them. In *The History of Statistics*, individuals predominate over social factors, which can never

[118] Stigler leaves unexamined the reasons why the normal distribution figured so largely in Nineteenth century thought. His approach can be contrasted with that of Foucault, who developed an analysis which explained why the normal curve became so potent in the nineteenth century. See Foucault 1981 pp135-145 especially p143, which introduces the idea of bio-power and the normalizing society. Foucault proposed, but never presented, a course on heredity, in 1969 (Foucault 1997 p7-10).

[119] Galton was helped by the Cambridge mathematician J Hamilton Dixon. (Stigler 1986 p285-289).

[120] Galton argued about the innate superiority of ancient Greece over nineteenth Century England, and of England over Africa.

be more than a 'climate'.[121] This is progressive history, one in which Quetelet 'failed to provide the real payoff he aimed for' although he 'helped create a climate of awareness'[122]

Although Porter's retrospective judgement of Quetelet in his 1986 monograph *The Rise of Statistical Thinking* is almost as harsh,[123] he ascribes a greater impact to nineteenth century liberal politics on the development of statistics. Porter characterises these as belief in the 'underlying stability of … society', the recovery of 'truths about mass phenomena even though the causes of each individual action were unknown', and 'the doctrine that order is to be found in large numbers'.[124] Porter's approach is similar to Stigler's. However, it places less emphasis on individuals, and carries less discussion of the mathematics they developed and more of the contextual factors involved. Porter argues that context is more important than mathematics because of the topic of statistics is essentially social.[125]

With Hacking's 1990 study *The Taming of Chance*, there is interest in individuals only insofar as they are 'convenient anchors for a particular organisation of sentences'[126] Hacking's account is about the development of an institutional and intellectual framework which supported a statistical style of reasoning. Individuals play a secondary role. To emphasise this anti-individualist approach, obscure individuals figure in *The*

---

[121] Accordingly, Stigler rejects MacKenzie's emphasis on social factors. See Stigler 1986 p267.

[122] The two quotations come from Stigler 1986 chapter 5 p 219-20.

[123] Porter 1986 p54-55.

[124] Porter 1986 p 5-6.

[125] Porter 1986 p11. The difference in approach is perhaps that of a historian turned statistician compared to a statistician turned historian.

[126] Hacking 1990 p8.

Taming of Chance as prominently as Quetelet and Galton do in Stigler and Porter. Quetelet is, if anything, given more respect than Galton in *The Taming of Chance*,[127] but such judgements are not of any real import. Hacking once again wishes to analyze the conditions that made statistics possible:

> *'My project is philosophical: to grasp the conditions that made possible our present organisation of concepts in two domains. One is that of physical indeterminism; the other is that of statistical information developed for purposes of social control'*[128]

In *The Taming of Chance*, the technique of correlation is barely described.[129] Galton and Pearson are invoked, not as originators of correlation, a claim which anyway Hacking shows to be problematic,[130] but as representatives of the ideas of partial-causation and non-causation. Correlation is important only insofar as its use requires the elimination of ordinary causality. Correlations replace causes as a way of thinking about the association between events. In *The Taming of Chance* correlation became possible only in a context in which causality has been effaced and statistical laws become real.

The three accounts offer competing solutions to why correlation appeared when it did. Put as a question, the answer from Stigler would be 'Galton's use of Quetelet, seasoned by his interest in heredity and the distribution of characteristics in society'. From Porter the answer would be 'the nineteenth century liberal view of society', but also Galton's

[127] In Stigler Galton transformed Quetelet's work. In Hacking, Quetelet makes the essential breakthrough. The reification of mean values for anthropomorphic data by Quetelet allowed Galton to see that the error curve could be though of as a distribution (Hacking 1990 p113).

[128] Hacking 1990 p5-6.

[129] Hacking 1990 p180-188.

[130] Hacking 1990 p187-88.

alignment towards eugenics[131] and social roots in Quaker idealism.[132] From Hacking, the answer would be 'the autonomy of statistical law; a hereditarian perspective; anthropometry; the normalisation of populations'. Stigler's explanation is individualistic; Porter invokes the social and political ideologies of nineteenth century Britain. Hacking's origin of statistics is harder to pin down. It is social, but not in the same way as Porter. In Hacking, correlation techniques arose because of two factors. Firstly, the growth of a social technology of statistical data collected for the purposes of social control.[133] Secondly, 'the autonomy of statistical law',[134] by which he meant a process whereby statistics of height, age, weight, death etc etc might be related to each other without having to refer them first to underlying cause. As Hacking puts it, the modern techniques of statistics orginate in 'the avalanche of printed numbers', 'bureaucracy', 'professional lust for precision', and 'the improvement of deviant sub-populations'. 'Statistical laws could only be noticed after the social phenomena had been enumerated, tabulated, and made public'.[135]

As the authors themselves suggest, the three studies are complementary.[136] Yet some comparison and judgement must be made because ultimately they offer competing frameworks for discussing the history of statistics.

The History of Statistics is a well-written and comprehensive history. In discussing the part played by individuals it also exposes the assumptions they had to make when

[131] Porter 1986 p129.

[132] Porter 1986 p131.

[133] Hacking 1990 p6.

[134] Hacking 1990 chapter 21.

[135] The phrases come from three sources: Hacking 1981; Hacking 1982; Hacking 1990.

[136] For Hacking's statement on this see Hacking 1990 p9.

developing ideas. But because the focus is on the individual, The *History of Statistics* cannot explain why a particular individual was able to do what he or she did. Why, for example, was Galton able to leap ahead of Quetelet, have the normal curve passed between generations, and thus produce regression theory and subsequently a theory of co-relation?[137] The technique of correlation, emerging from Galton's studies, assumes that each individual, and their characteristics, is an instantiation of a 'typical form of man',[138] distributed in a normal fashion. It required, at least at an early stage, a sustained belief in the ubiquity of the normal distribution. Without these premises, the technique of correlation would not have emerged.

By focussing on technical aspects of the history of statistics Stigler's approach systematically underplays the extent to which the social context made it possible for Galton to produce his re-working of Quetelet. Porter and Hacking show that what made social sciences amenable to probability statistics was not the transforming genius of Galton, but a social reality that was already being structured in a manner amenable to probabilistic interpretation. Porter's analysis can be seen as a transition between Stigler's approach and the fully worked through analysis by Hacking which emphasises the role of discursive structures in making reality possible.

The structural approach to the history of statistics

[137] Galton's philosophy – his commitment to the normal curve and his regard for classification of human characteristics – is brought out more clearly in Porter 1986 p140 and p144. Porter also argues that the genesis of correlation in Galton's thought is due to his attempt to find relations between scientific achievement and a rank scale of general ability (Porter 1986 p143).

[138] The phrase belongs to Dr Charles Roberts, a leading member of the Anthropomorphic Committee. For details on Roberts see footnote 16 on p133 of Szreter 1997. Hacking uses the phrase 'making up people' to describe the same thing (Hacking 1990 p6).

Despite the attractions there are several risks associated with a structural approach to the history of statistics. Taken to an extreme, historical events appear to be contingent upon structural elements. Foucault's case study, for example, of the transformation of French medicine at the end of the eighteenth century is a particularly valuable source for the history of medical statistics.[139] 'The visibility of the medical field assumes a statistical structure',[140] because of 'a spontaneous and deeply rooted convergence between the requirements of political ideology and those of medical technology'[141] Nevertheless, *Birth of the Clinic* is Foucault's most structuralist study. As Weiss has shown, the course of statistical methods in early nineteenth century France was not determined solely by the eighteenth century transformation in structure of its knowledge.[142] Gutting has raised the charge against *The Birth of the Clinic* that it leaves open the relationship between changes in knowledge and changes in practice. Did a statistical form of knowledge change the social structure of French medicine, or did reform in the structure produce a change in the epistemological framework of French medicine?[143]

Active elements in structuralist histories

The risks embodied in a structuralist approach are avoided in most histories by dividing the field of explanation into two parts and characterising one as a back-cloth against which an active element performs. Stigler's account suggests that the special cognitive ability of certain individuals is the active element against a backcloth of society. MacKenzie, and also Habermas, propose that 'cognitive interests' rather than individual

[139] Foucault 1973.

[140] Foucault 1973 p102.

[141] Foucault 1973 p38.

[142] Weiss 1995 chapter 7.

[143] This point is made by Gutting in his critical overview of Foucault's early works (Gutting 1989 p138).

genius can explain why specific techniques emerge when they do.[144] Buck introduces political considerations – the need to impose order on social and political life – into the quantification of Graunt and Petty.[145] Gould shows how Burt's extension of Spearman's work on factor analysis was driven by his desire to ground his bureaucratic activities concerning the ranking of schoolchildren in naturalistic categories.[146]

Yet the division into active and passive elements remains unsatisfactory if the connection between them remains unexamined. Yearley's critique of cognitive interest theory is important both because it offers one of the few critical readings of MacKenzie's study *Statistics in Britain*, and because it suggests a way to move beyond the limitations of active/passive elements.[147] The method he suggests is to look at 'the manner in which scientists present their arguments under various circumstances'[148].

A number of studies have begun to do this. Latour's analyses statistical techniques as methods to accumulate and deploy epistemological resources.[149] Hacking develops this idea and proposes an analysis based on statistics as a 'style of knowledge formation',

---

[144] MacKenzie 1981. Habermas 1978 pp308-9.

[145] Buck 1977.

[146] Gould. 1981 p285-296 For details of Burt's career in school inspection, his role in the creation of the 11+ exam, and his belief in the innateness of intelligence, see Hearnshaw 1981.

[147] MacKenzie's study of the statistics in Britain is widely acclaimed, because it seems to fulfill the requirements of the sociology science in a particularly difficult area, that of mathematics, which is generally considered to be the first discipline to have freed itself from a social history. Yearley shows that MacKenzie's study fails to substantiate its claims. One of the chief failings of the studies which have followed the cognitive interests approach is their tendency to construct controversies in a manner which suits the intended conclusions. On the basis of his review Yearley concludes: 'that the sociological theory of cognitive interests is vague and theoretically weak'. He is particularly critical of the use made of disputes between scientists, a construct he finds at the centre of MacKenzie's approach: 'scientists cognitive interests have been shown neither to determine the outcome of scientific disputes nor to prevent them from accommodating to alternative viewpoints …allusion to 'sides' seems to be a characteristic of the way in which participants present controversies in certain contexts' Yearley1982.

[148] Yearley 1982 p388.

[149] Latour 1987 p232-241 and Latour 1988 p90-91. Latour has always been wary of giving precedence to social explanations of science. His active elements construct the social just as much as the social determines the cognitive.

whose 'truth producing virtue' overcomes the limitations of realist and relativist approaches to the history of science.[150]

Statistics as a form of persuasion is the most recent proposal suggested for the study of statistics by Latour and Hacking, and it is the framework for analysing clinical trials adopted here. To understand the growth of statistics - to understand how it becomes the premier way of generating knowledge - it is necessary to follow the way in which it deploys its truth-producing virtues. The questions to ask whenever statistics come into play are: Why are they useful? How are they useful? What makes a technique plausible? I will argue that the adoption of techniques is as much a matter of utility as it is of logic. Plausibility, in the case of a statistical statement, is linked ultimately to processes of legitimation through usage. If statistics have a social as well as a mathematical identity it is not only because their creation is informed by social factors, but also because their application creates a series of tractable social objects.

Conclusion: statistics as rhetoric

There is support for the utilitarian view of statistics from statisticians themselves. Neyman was quite clear about the epistemic status of decision tests. He believed that the problem of inverse inference, coupled to a strictly frequentist interpretation of probability, precluded exact knowledge. Hence the title of his best known paper, published in 1933 with Egon Pearson, which concerned the most efficient tests of statistical hypothesis. They argued that in the absence of epistemological security, the best statistical tests were those that performed most efficiently.[151]

[150] Hacking 1992.

[151] Neyman and Pearson 1933.

Decision tests apply the deliberately diminished logic that is appropriate to approximate knowledge.[152] They are therefore part of approximative syllogistic logic,[153] a rhetoric of reasoning and positivist proof based on the assumption that what is frequent is in fact true and that what is infrequent is in fact false.[154]

The dual character of modern probability – aleatory and epistemic - links it to the old form of probability that preceded it. One therefore sees why modern statistics can be described in exactly the same terms Aristotle used to characterise rhetoric: 'the power to observe the persuasiveness of which any particular matter admits'.[155] Statistics are rhetoric; they are verbal tactics deployed to create epistemological and practical advantage. In the next chapter, this understanding of statistics is applied to the use of numerical reasoning about the value of therapeutics in British medicine in the late nineteenth century.

[152] Barthes 1988 p22.

[153] Syllogistic logic is family of arguments of the type 'If All As are Bs and all Bs are Cs, then all As are Cs'. The form of approximative syllogistic logic that applies to confidence intervals is set out in Hacking 1980 p152-153.

[154] Knowing the limitations of the model he was applying, Neyman made no claim about the truth or falsity of statements. He preferred to speak of the cost and benefit of behaving as if results were true or false. 'I shall behave as if I consider the result true' is an essential part of Neyman's theory. The standard for truth and falsity thereby becomes a function of the underlying frequentist statistical model.

[155] Aristotle 1991 p 74. Several rhetorical forms have an obvious and immediate parallelism with statistical techniques. Synecdoche for example, where the part stands for the whole, is the technique where inferences based on samples are extended to wider populations; metonymy is frequently used for example the statistical notion of population and the demographic notion of population do not necessarily refer to the same object but may be interchanged. For comments on the metaphor of population, see Cole 1994.

Chapter Two

The British Medical Journal and the creation of statistical medical knowledge in the nineteenth century

Introduction

The purpose of this chapter is to examine the ways in which statistical knowledge was promoted by the British Medical Journal (BMJ) in the 1850s and 60s. I will argue that a statistical form of knowledge played an important role in the efforts to reform the medical profession. Why and how this was done is examined in several settings, including the use of medical testimony in courts and in a series of clinical trials organised by the BMJ in 1862.

The introduction to this dissertation set out the claim, put most cogently by Shapin and Shaffer,[156] that problems in the organisation of social groups can be resolved, at least in part, through reform in their knowledge-making activities. Shapin and Shaffer wrote in the context of the late seventeenth century, and showed how the experimental methodology devised by Robert Boyle was linked to a particular solution to the recreation of social order after the restoration of Charles II. In this chapter I examine the idea that in a similar way a statistical form of knowledge was intended as the basis for both the creation of a scientific knowledge of therapies and the creation of social order in a renewed medical profession.

In the case of the BMJ this chapter considers some of the ways in which the journal sought to create a figure, which can be called the ideal practitioner. As I argue below, the figure of the ideal practitioner combined a way of knowing with a way of behaving. The

[156] Shapin and Shaffer 1985.

achievement of the BMJ in the second half of the nineteenth century was to show that a way of gentlemanly conduct would create rational knowledge, and that a rational epistemology would lead to honorable professional behaviour. However, it could not be simply a process of laying down by *fiat* that it was an aspect of gentlemanly behaviour to be scientific, or that a rational approach to knowledge would result in a gentlemanly code of conduct. If, in the case of medicine, gentlemanly conduct and scientific knowledge implied each other, it was because the principal characteristics of both originated in a set of rules which dictated that medical knowledge must have a probabilistic form if it was to be valid. These rules are explored below.

The figure of the ideal practitioner as a bearer of probabilistic knowledge can be seen in a number of settings. The sections below describe some of these: the role of medical knowledge in court, especially the form of knowledge the ideal practitioner would bear and the codes of gentlemanly behaviour he would apply in his dealings with fellow practitioners; his relationship with irregular practitioners, particularly homeopaths; and finally, knowledge about the effectiveness of drugs. Although the latter is the focus of this research, I want to show that the knowledge of therapeutic effectiveness is not isolated from other aspects of knowledge making and practice, such as medical witnessing in court.

Medical reform in mid nineteenth century Britain

The middle years of the nineteenth century tend to be considered as a backwater for the use of statistics in British therapeutics. Between the excitement generated by Pierre Louis

and Jules Gavarret in the first third of the century,[157] and the English school of statisticians led by Karl Pearson at the turn of the twentieth century, there is commonly supposed to be little in the way of therapeutic evaluation. Progress in therapeutics itself is regarded as negligible until the second half of the nineteenth century.[158] It is the period covered by the expressive phrase 'therapeutic nihilism', when collectively practitioners had recognised the deficiencies of their old methods, exemplified by therapeutic blood-letting, but not yet secured a single effective treatment.

Lacking any remotely effective treatments, it might be argued that statistical analysis of therapeutics was unnecessary. Certainly, French clinical trials of the 1830s had little impact in Britain. Of Louis's leading English disciples, William Farr became involved in the descriptive statistics of populations. And although William Guy published on the numerical method, it had little impact on therapeutics. Shryock's characterisation is very largely the accepted view today:

> *The very progress physicians were making between 1830 and 1850 made it the more difficult for them to offer the public much encouragement. The first generation of critical clinicians had so much traditional trash to clear up, and such difficult foundations to lay, that it was never able to build a therapeutics that could impress the laity.*[159]

---

[157] PCA Louis is famed for his advocacy of numerical methods in medicine, though as Coleman notes, there is no comprehensive account of the numerical method. (Coleman 1982 p132 footnote 15) Louis is best remembered for organising a controlled clinical trial of therapeutic bloodletting to treat typhoid fever. This study is available in translation. (Louis 1836) 'Louis work … framed the issue of quantification in clinical medicine in the terms in which it would be debated in the Parisian Academy of Science and Medicine in the late 1830's' (Matthews 1992 p38). For some views on Louis, see Greenwood 1936, Bollet 1973 and Yankauer 1996. The debate at the Academy of Science and its outcome are discussed in Matthews 1992, Cole 1994, and Weisz 1995 chapter 7.

[158] Bynum 1994.

[159] Shryock 1979 p249 The impact of Farr and Guy on English epidemiology is discussed in Lilienfeld 1978. Major Greenwood bemoaned the lack of influence of Louis in Britain in his *Medical Dictator* (Greenwood 1936 p141) Louis' influence in America is better documented. See Steiner 1939, and Warner's studies, especially Warner 1985**b**.

Insofar as this depiction of therapeutic nihilism relates also to the development of formal statistical methods of evaluation, it is accurate enough. But it could hardly be otherwise, for the mature development of statistical tests really is an event of the twentieth rather than the nineteenth century. The depiction is less accurate, however, if one considers what might be called statistical thinking, rather than statistical testing.

By statistical thinking I mean any argument or method which uses aggregates of data – populations, samples, groups – to infer conclusions. Using the principal record book of British medical reform, the British Medical Journal, it becomes apparent that during the 1860s statistical thinking was used to evaluate therapies. The trials were not successful. However, they offer an early and clear example of the way in which statistics might work as rhetoric. In this chapter, rudimentary statistical analysis of therapies is portrayed as a style of knowledge formation, concerned with reform of medical knowledge and reform of medical organisation, hand in hand.

The Medical Act and the making of the British medical profession

On the 1st of October 1858 the Medical Act came into force.[160] It created a 'General Council of Medical Education and Registration of the United Kingdom'. For the first time all persons holding medical qualifications were required to register with the General Council.[161] The Act also called for the publication, under the direction of the General Council, of a book 'containing a list of medicines and compounds, and the manner of preparing them, together with the true weights and measures by which they are to be

[160] 21 and 22 Vict. c. 90.

[161] 11 separate qualifications were recognised, chiefly relating to those awarded by the Royal Colleges and Society of Apothecaries, but also included doctorates granted by the Archbishop of Canterbury.

prepared and mixed, and containing such other matter and things relating thereto as the General Council shall think fit, to be called "British Pharmacopoeia"'.[162]

The 1858 Act was based on the Medical Reform Bill introduced by the President of the Board of Health, W.F. Cowper, in December 1857.[163] The Act is of some importance, but as is increasingly recognised, it did not establish the modern medical profession to anything like the extent that synoptic accounts of medical progress suggest.[164] Cowper's Bill, the 16th medical reform bill to be put to Parliament between 1840 and 1858,[165] succeeded because it skillfully incorporated some of the perspectives of both the reforming (broadly speaking general practitioners and the majority of provincial hospital doctors) and conservative (the Royal Colleges and London élite) parts of the profession. For reformers, the Act offered some progress because it 'implied equality before the law of all registered medical practitioners',[166] while from the Colleges' perspective, the proposed General Council would extend their role in shaping the profession and certainly offered no threat of diminution to their status at the head of the profession.[167]

[162] BMJ August 14 1858 p688.

[163] It was drafted principally by John Simon, Medical Officer to the Board of Health. On 2nd June 1858 the Bill received its second reading; the third reading was passed on 29th July, and royal assent on 2nd August.

[164] Loudon's view is that in the long term the Act was of benefit to the public and the profession. But the Act provided little immediate benefit to either. His review is particularly critical of the idea that the Act unified the profession. (Loudon 1986 p297-301) For a recent historical review of the wider context of the Medical Act see Lawrence 1994 chapter 3.

[165] Petersen 1978 p34 A detailed account of the successive bills is given in Newman 1957.

[166] Petersen 1978 p33.

[167] During the Summer of 1858, the royal colleges successfully put pressure on the government to reduce the proposed powers of the new Council concerning the organisation and regulation of medical education and examination. The colleges were successful, since in its final version the bill gave the Council only the power to require information about education and examination.

Despite the skilful framing that permitted its enactment, the 1858 Medical Act fell some way short of meeting the aspirations of reforming practitioners.[168] It did nothing to alter the balance of power between the Colleges and the great mass of the profession. Nor did it do anything to protect general practitioners from the effects of unqualified practitioners, or poor law medical officers from their employers.[169]

In light of the difficulties of creating a powerful unitary profession through legislation, the newly formed British Medical Association[170] (BMA) took upon itself the task of doing what legislation seemingly could not:

> *What is the use of our calling to Jove to assist us – complaining of defective medical Acts of Parliament or an inefficient Medical Council – so long as this want of espirit de corps … exists amongst us. What Medical Council of Acts of Parliament can provide us with brotherly love?*[171]

To a very considerable extent, the BMA succeeded. As Petersen noted in her study of the years between the 1858 Act and the 1886 Medical Act Amendment Act, the legislative dates 'encompass a period when the profession's internal relations and its relationships with lay society were redefined in significant ways'.[172]

---

[168] Despite the largely negative assessment of the Act (see note 9 above) it is clear that the provisions of the Act were used to pursue reform. See BMJ 1859;ii:944-5. And when the Act failed to procure conviction of registered quacks, the BMJ wrote an encouraging editorial detailing recent successful prosecutions (BMJ 1860;i:400-401).

[169] A further flaw was the perpetuation of the partially qualified doctor (Loudon 1986 p298). Under the terms of the 1858 Act practitioners could enter the register on the basis of a single qualification. The need for qualification in medicine Surgery and midwifery was introduced in the 1886 Medical Act Amendment Act.

[170] Until 1856 the BMA was known as the Provincial Medical and Surgical Association (PMSA).

[171] BMJ 1863;I:272. The BMJ was commenting on recent cases in which practitioners had given evidence in open court about the wisdom of actions taken or therapies given by fellow practitioners.

[172] Petersen 1978 p3.

Much of the credit for the achievements between 1858 and 1886 must be given to the BMA, as Parry and Parry have argued.[173] However, it is not immediately obvious why the BMA should have succeeded, since the fledgling organisation of the 1850s had few resources at its disposal.[174] Its 1,500 or so members were mostly based in the provinces; London resident doctors having been admissible for membership only since 1853. It had a branch structure for its members, and several committees for addressing problems. As well, it had an annual meeting. Above all the BMA had a weekly journal, published in London, and distributed free to all members of the Association. The journal was the most visible, most important, and most obviously successful part of the BMA.[175] For many years it consumed the major part of the Associations resources.[176] But after mid-century it began to bring in substantial advertising revenues, and also to shepherd its readers towards membership of the BMA, since it was cheaper to join the BMA and receive the journal free than it was remain outside the Association and subscribe to the journal.[177]

The British Medical Journal (BMJ) of the nineteenth century was a reforming journal. Unlike its great rival The Lancet, which sought advance through direct conflict with the medical and social establishment, the BMJ sought to bring about reform through the

[173] Parry and Parry 1976 p147.

[174] Its success would have surprised Thomas Wakely, editor of the Lancet, who described the PMSA as 'a disgraceful abortion' comprising 'general practitioners who have allowed themselves to be made the dupes and puppets of the migratory showmen'. Quotation in Bartrip 1990 p 7. The early history of the PMSA/BMA is recounted in Vaughan 1959.

[175] It was for example, the envy of visiting members of the American Medical Association (AMA). The Journal of the American Medical Association, first published in 1883, was modeled on the BMJ, and was consciously intended to strengthen the AMA. (Knoll 1992).

[176] In 1858 for example, the total income to the BMA was £2436, of which £646 (26%) was advertising and other income from the BMJ. The total expenditure of the association was £2429, of which £1685 (70%) was printing costs and a further £200 (8%) at least is due to costs associated with the journal. (BMJ 1859;I:295).

[177] See Bartrip 1990 and Bartrip 1992.

creation of a collective consciousness for the BMA, and by implication, for the medical profession of the late nineteenth century as a whole.

This chapter concerns the methods by which the BMJ sought to create such a collective medical consciousness. Since it was no more than a journal, the BMJ could only reflect events taking place elsewhere, as the title of Bartrip's standard history suggests.[178] Yet it is apparent that the ambition of the BMJ was much greater than simply to reflect the life of the profession. More active than a mirror, it was able, through the manner of its publication and distribution, through the intimate relationship it enjoyed with the members of the BMA, and above all through the content and style of its pages, to draw its readers into a greater sense of collective identity. In the limited sense that will be explored below, it seems reasonable to suggest that the BMJ created the medical profession of the late nineteenth century. Week by week it set before its readers an image of the ideal medical profession. It laid down rules for the behaviour of one doctor to another, and suggested what the appropriate response of a fully formed profession would be to all the events of the day that had any bearing on medicine.

Among the means available for creating this textual image, the use of a highly rhetorical literary style figured prominently. As Bartrip has noted, the main function of the BMJ was, before the editorship of Ernest Hart in January 1867, to be 'a propagandist of the BMA in the struggle for professional advancement.'[179] Bartrip sees in this most rhetorical phase of the BMJ's existence an absence of original scientific work.[180] While it is true that the early BMJ published nothing that would be recognised as a scientific paper by

[178] Bartrip 1990.

[179] Bartrip 1990 p61.

[180] Bartrip 1990, introduction.

today's standards, Bartrip's view must be qualified in two ways. Firstly, to read through a few issues of the early BMJ will reveal that it does contain accounts of scientific research, reports from scientific societies, case studies and reviews of books. Secondly, the implication that a journal can be either rhetorical or scientific is not the case for the early BMJ. I will argue that, on the contrary, scientific reports and discussions played a vital rhetorical role in the debate about the constitution of legitimate knowledge. And equally, that concern for the legitimacy of knowledge was part of reform.

The sections below describe some of the situations in which science, rhetoric and reform acted most powerfully together. These include the role of medical knowledge in court, especially the form of knowledge the ideal practitioner would bear and the codes of gentlemanly behaviour he would apply in his dealings with fellow practitioners; his relationship with irregular practitioners, particularly homeopaths; and finally, knowledge about the effectiveness of drugs.

Medical testimony in court

In the early 1860s the BMJ returned again and again to the question of the relationship between medical and legal knowledge. For several reasons the legal system was a particularly compelling topic for the BMJ. Principally of course, the legal system offered the one significant practical distinction between registered and un-registered practitioners in the 1858 Act, whereby un-registered practitioners could be prosecuted for false claims to registration. It was also the arena of a distinguished profession whose status the BMJ saw in some ways as worth aspiring to. Finally, it was a public space, with its dealings regularly reported in the lay press.

These considerations made it likely that the BMA would follow court proceedings with some interest. The immediate reason, and the immediate motivation for the articles that appeared in the BMJ, was the humiliation of medical knowledge in the courts of the 1860s.[181]

1 The credibility of expert medical testimony in court.

The Journal reported several cases of conflicting expert testimony given by practitioners under oath:

> *We have already more than once called attention to the painful position in which the medical profession is placed, when, in a court of justice, it appears in division – cut in halves – and supporting two diametrically opposed conclusions.*[182]

Medical practitioners, and by implication the wider profession, were compromised by appearance in court as witnesses. The division of opinion exposed in public taught the world 'incredulity in medical science',[183] allowed 'the ridicule of a censorious and hostile public'[184] and highlighted the divisions in the profession. The position of medical witnesses by courts was therefore a serious matter for the BMJ.

The source of division was the role the medical man was called on to play. 'The truth is, that the evil which lies at the bottom of the mischief is this: that medical men, instead of giving a judicial, give an advocate's opinion, in the cases of the kind to which we are alluding. They are summoned to court to make the best of their client's case, not to

[181] Further information about medico-legal events in the nineteenth century is found in Smith R 1983, and in Smith R 1980 and 81.

[182] Editorial. BMJ 1862; i :15. Throughout this chapter I refer directly to the source of each quotation in the BMJ.

[183] BMJ 1862;i:15, see also BMJ 1862;ii:40.

[184] BMJ 1862;i:286.

declare the whole truth according to the light of scientific knowledge'.[185] The doctor should be a witness for the truth and not act as a partial advocate for a client. 'Medical men will never obtain for their evidence the respect it deserves, until they cease to appear as advocates in a court of law'.[186] Science – in the form of a neutral account of the 'barren facts of the case'[187] – provides a topic for the discourse of the medical witness in court, who may represent no interest except that of science itself.

Yet the circumstances in which evidence was given made it difficult for medical witnesses to appear as neutral experts. The BMJ identified two particular problems. Firstly, medical evidence was subject to critique by a lay jury. In the case of the suspected poisoner Smethurst, medical opinion as to the cause of death of his mistress Isabella Banks was thoroughly divided between arsenic poisoning (10 medical witnesses) and natural causes (7 medical witnesses).[188] 'The trial has, without doubt, left upon the public mind a very painful impression of the inadequacy of medical and scientific evidence'.[189] In the BMJ's view the problem lay not only with the divided nature of the evidence, but with the trial process, which artificially forced scientific evidence to come into conflict.[190]

The second problem identified by the BMJ was that the courtroom might recognise little distinction between medical evidence and testimony of non-medical witnesses. Called to give evidence about the sanity of a defendant, 'the opinions and conclusions of non-

[185] BMJ 1862;i: 258.

[186] BMJ 1862;i: 258.

[187] BMJ 1862;i:258.

[188] The trial is described fully in BMJ 1859;ii:707-11.

[189] BMJ 1859:ii:725.

[190] BMJ 1859:ii:725.

medical persons [were to] be held as sufficient'[191]. In general terms, complained one editorial, medical men occupied a lower status in court than judicial officials did, and the evidence they gave had little authority.[192]

As the BMJ observed, the court-room process was doubly prejudiced against medical knowledge, by artificially intensifying disputes and also by discounting the value of medical evidence. But the unpleasant truth was that medical opinion was divided. The Smethurst case showed that 17 medical witnesses might be evenly divided into two opposing views. In 1862, in the Windham case, and again in 1864, in the Townley case, several medical witnesses called on to pronounce on the sanity of a defendant gave diametrically opposed views.[193] An editorial quoted the views of the Lord Chancellor on the Windham case:

> *Here you have a medical man presented, who tells you that, according to his experience, the existence of cerebral disease is shown by certain bodily symptoms; while another medical man, or half a dozen, met his theory with a direct negative, and tell you that in their experience the particular symptoms relied on by the former witness as a criterion of mental disease may be fairly accounted for in another way, and present no certain indicia of its existence. Between these learned doctors, who is to determine?*[194]

The BMJ's response had two parts. Firstly, it called for a reduction in the variety of medical and expert opinion that could be put before a jury. This could be achieved quite easily through the elimination of the system of partisan medical advocacy, and its

[191] BMJ 1862;i:385. From the review of JG Davey's '*Medical evidence in our law courts*', Bristol, 1862

[192] BMJ 1862;i:15.

[193] The Townley trial, and several other similar cases, are discussed in Smith 1981 chapter 5.

[194] BMJ 1862;i: 258.

replacement by state sanctioned expert medical and scientific witnesses.[195] Secondly, the BMJ recognised that the solution lay in the hands of the profession as much as in the reform of court proceedings. There was a need for an internal discipline among medical witnesses. In part this was a matter for individual doctors called to give evidence: evidence given by a medical witness should consist of a plain narration, free from speculative hypothetical conjecture and the lofty and ornamented phraseology of science.[196] If possible, the evidence presented should take the form of an empirical demonstration of the facts in court.[197] Science in the court room required a suitable form of presentation: 'We have no desire to clip the wings of what has been called the scientific imagination; but we charge those who possess the gift to take care that they cultivate it within the boundaries of sound reasoning'.[198] In part also the internal discipline of individual medical witnesses must be the collective responsibility of the medical profession:

> *The performance of medical experts are becoming a libel on the profession; and it behoves the Association to organise some means of attaching a proper penalty to this kind of misconduct*[199]

In seeking to exert influence over the deployment of knowledge by its members, the BMA was emulating the Royal College of Physicians, which long held the right to

[195] The call for a 'state engine of this kind' was made in an editorial on the trial of Thomas Smethurst in 1859. BMJ 1859;ii:725.

[196] BMJ 1862;i:287.

[197] In the case of evidence about a defendant's state of mind, the BMJ recommended that 'historical' knowledge of the defendant's past character should be rejected in favour of an examination of his or her present state of mind. BMJ 1862:i:96-97. For further comments on the Windham case see the editorial BMJ 1862;i:146-7.

[198] BMJ 1862;i: 286.

[199] BMJ 1863;i:465 [letter from James Edmunds].

protect and develop its knowledge.[200] Yet the 1858 Act had not provided any of the authority needed to regulate its members in this way. Lacking power to control its members' activities, the BMA's approach in 1862 was to suggest that collective control over medical knowledge derived from the structure of medical knowledge itself. Proper management of medical knowledge could only be achieved by recognising the inherently collective nature of medical knowledge making.

In pursuit of this argument, the BMJ used events in courts to its advantage. It asserted that if medical opinion was divided it was because of the essentially probabilistic nature of medical knowledge. This extended quotation shows how a probabilistic interpretation of medical knowledge might explain why medical evidence foundered in court, and how a recognition of the impact of probabilism might contribute to a renewal of medical authority:

[200] A good guide to the early history of the Royal College of Physicians is Cook 1986. Chapter 2 describes the Royal College as a learned society which also 'possessed the juridical powers to make that learning the standard by which all other practices might be judged' Cook 1986 p71.

> *It is proverbial that medical men make bad witnesses … the root of the perplexity lies deeper than the surface … almost every fact we observe, or symptom that we notice, is a compound effect of sensation and inference, and the validity of our decisions mainly depends on the accuracy with which we weigh our inferences and the care we take to keep ourselves within the bounds of probability. We listen to the beating of a heart supposed to be diseased; and on hearing a particular sound, we infer the nature and position of the diseased structure, partly from the character of the sound and partly from the examination of other signs with which we have previously become acquainted. The object we have had in view, namely to understand the nature of the lesion has been revealed to our minds from the consideration of several separate facts; and our conclusions respecting that object may be relied on in proportion as our perceptions have been clear and our deductions unbiased and accurate. All legitimate inferences then are amenable to the laws of induction and if we could always confine ourselves to these without affirming anything in our descriptions which is not warranted by the limits of our knowledge, we might escape the impertinent interference of the lawyer, and do the state essential service.*[201]

If careful synthetic appraisal was needed when an individual reckoned his own knowledge, argued the BMJ, much the same was true for the profession as a group. Because medical knowledge was probabilistic, the individual could only have access to a part of the truth. Truth was therefore the property of the group.[202]

Medical knowledge must therefore be combined amongst individuals if it was to have validity: 'a fact in medicine … is the resultant of very numerous observations made by

[201] BMJ 1862;i:287. See also the editorial Law and Physic BMJ 1859;ii:743-5: 'the theory of probabilities, the elimination of chance, and the registration of empirical laws are forever forcing their importance upon his [the medical practitioner's] consideration… and the physician arrives at the truest conclusion in his practice who has the gift of bringing the largest number of facts before him at a glance, and, by accurately observing the elements of the facts, of best determining what are the 'middle results' of all'.

[202] Thomas Inman made the point that empiricism could be misleading if it was not recognised that experience was not 'invariable and certain (BMJ 1858;ii:945). The cause of differences in the empirical estimation of value of lime juice as a treatment for rheumatism was an insufficiency of experience, he claimed. The BMJ also began to publish lectures on the history of medicine. The purpose of Alexander Henry's lecture series of 1860 was to remind his audience that the remembrance of the few giants of medicine tended to obscure the true nature of medical progress, which came about when 'numbers of minds have been employed' (BMJ 1860;i:221).

fitting and capable enquirers'.[203] This explains the confidence of the BMJ in its own assertion in the Windham case that had the medical witnesses met before the trial to discuss the case 'we do not believe there would have been the smallest disagreement amongst them'[204] Medical knowledge becomes stable when its facts are amassed

> *According to M Fourier "there is in all statistical researches a general proposition to which too much cannot be paid-namely, that the indefinite repetition of events, generally denominated fortuitous, does away with the changeableness that may belong to them. In a series of an immense number of facts, there are none but constant and necessary relations, determined by the nature of things.'*[205]

From this time on, medical knowledge was to be collective.

2 Medical testimony by practitioners against practitioners

The collective structure of medical knowledge also implied a code of behavior for practitioners when called to give evidence in court about the practice of another practitioner. The code, never formalised but repeated week after week in the BMJ, suggested that a practitioner should not testify against 'a brother practitioner' in court. The rule, which might have appeared as simply an arbitrary injunction, was in fact dictated by epistemology. For if an individual's perceptions are but an element of the whole, then 'surely it is abundantly credible that another's may differ from ours often or always'.[206] The BMJ carefully avoided the view that a practitioner should never testify

[203] W.O. Markham. Address in medicine to the twenty ninth annual meeting of the BMA. BMJ 1862:I;31-33. The collective nature of knowledge making was repeatedly stressed by contributors to the BMJ. Robert Christison, Professor of Materia Medica, in his address to the twenty sixth annual meeting of the BMA, claimed that all genuine discoveries in therapeutics come about through the involvement of several persons (BMJ 1858;ii:675).

[204] BMJ 1862;I:15.

[205] BMJ 1860;I:383.

[206] P.M. Latham. Address in medicine. BMJ 1862;I:597.

against another, but in practice a sanction against doing so began to be the norm in the BMJ, which regularly highlighted transgressions.

The case of Dr Philbrick, in 1862, illustrates the way in which its collective structure provided a manner for presenting medical knowledge in court, and enjoined a modesty that was both behavioural and epistemological. At the request of a midwife, Dr Philbrick attended a woman who had safely delivered a child but who was now experiencing excessive haemorrhage. The cause was found to be a second child lodged in the birth canal, believed to be dead. Having making his diagnosis Dr Philbrick left his patient, but when called again declined to attend. A Dr Clark then attended. He delivered the second child, and remained with the woman. She recovered, but after a week complained of severe headaches. The headaches were followed by paralysis on one side of her body, and death ensued shortly thereafter.

At the inquest the Coroner called witnesses to establish the claim that the woman's death was caused or hastened by excessive haemorrhage, which in turn was due to the unskilful or neglectful treatment by Dr Philbrick. A number of witnesses give their opinion that excessive haemorrhage could have been the cause of the fatal event. In its column of December 13th 1862 BMJ asked the medical witnesses to re-cast their opinions in the light of aleatory rather than epistemological probability:

> *Now, we should like to ask these gentlemen, who have all had very large experience in midwifery, how many cases of apoplexy they have seen in their lives where the apoplexy has followed upon and could be fairly connected with excessive haemorrhage after delivery?… If they cannot justify their statements by facts, it follows that their statements were based merely on theoretical opinions…Surely a medical witness, in a case of this kind, might have most properly said: the haemorrhage might have occurred if the highest skill in the land had been employed. The apoplexy might have occurred – would very probably have occurred-if not one drop of blood had been lost.*[207]

The words that the BMJ put into the mouth of its ideal witness turned the source of variability on its head. Rather than reflecting the imperfect knowledge of practitioners, variability originated in the constitution of the patient. It followed therefore that medical knowledge must always be uncertain, and that combination and co-operation among practitioners was the best route to improving knowledge. Through its containment within a statistical framework, error and difference, used to harass the profession, were shown to be inevitable. Rather than the source of shame, they could be a source of knowledge.[208] [209]

Irregular therapeutics

Injustice at the continued existence of irregular practitioners after the 1858 Act was keenly felt in the BMJ. In terms of column inches, homeopathy presented the greatest challenge, but the threat was multi-faceted:

---

[207] BMJ 1862;ii:622.

[208] The point occurs in many places in the BMJ, but was made explicitly by Arthur Ransome in his address to the thirty second annual meeting of the BMA in Cambridge, when called for the creation of committees in each town to oversee the combination of medical observations. A. Ransome. The need of combined medical observation. BMJ 1864;ii:405-8. Also, the letter (BMJ 1865;i:207) in which he re-iterated the point: 'many important medical problems cannot be solved by detached individual observation; the weight of evidence necessary to true induction can only be amassed by associated labour.

[209] Much more could be written about the BMJ's coverage of medicine in the courts in the 1860s and 70s. Cases such as that of Dr Waters, (reported in BMJ 1863;i:377-379, 402-403, 431-432, 436-437 and 439-441) accused of impregnating a servant under his care while she was made insensible by him, shows how far the BMJ was willing to go to censure practitioners who gave testimony in court which might incriminate a fellow practitioner.

> *Every village has its bone-setter; rubbers are beginning to infest watering places…; hydropathy and homeopathy sweep away whole sections of the community.*[210]

The threat of irregular medicine was felt more keenly by regular practitioners in the nineteenth century than it is today. To understand the efforts by the BMJ to counter this threat around 1860 it is necessary very briefly to describe the position of irregular medicine at the time. As Lawrence[211] and Porter[212] have argued, quacks were often more modern, more scientific,[213] and more successful[214] than the regular practitioners who denounced them. And characteristics which might be predicted to differentiate regulars and quacks, for example, the use of secret remedies, were by no means distributed in the way which contemporary definitions projected backwards would forecast.[215]

Contrary to some assertions,[216] the Medical Act of 1858 did not attempt to exclude irregular practice; nor did it prevent regular practitioners from using the remedies and methods we now associate with unorthodox practice.[217] A glance at any of the *materia medica* in use around mid-century shows that the bulk of the pharmaceutical armamentum available to regular practitioners was similar to that used by medical

---

[210] BMJ 1860;ii:921.

[211] Lawrence 1994 p15.

[212] Porter 1986 and Porter 1989.

[213] Or at least claimed to be more scientific, as in the case of chiropractic (Martin 1994).

[214] In the treatment of the pox for example (Quetel 1990 p86-93). It may never be known if Dr Leake's Pills were effective, but they were enormously popular in Britain at the beginning of the nineteenth century.

[215] Alison Winter's study of the relationship between inhalation anaesthesia and mesmerism shows just how complicated the relationship between medicine and quackery can be, in this case between one of the most celebrated achievements of nineteenth century medicine, and a practice that was completely condemned (Winter 1991).

[216] Miley and Pickstone 1988 p152.

[217] The 1858 Bill was amended by the House of Lords during the summer of 1858 so that the Medical Act protected qualified medical practitioners who wished to treat patients homeopathically. See the medical news section, BMJ 1858;(August 7):670, for comment.

botanists and homeopaths.[218] In fact, as Cooter has argued, 'it is not difficult to extend discussion of the likenesses between orthodox and heterodox medicine in the early nineteenth century'.[219] In summary, when considering the mid-nineteenth century it is more accurate to regard the categories applied in the argument - 'regular' and 'quack' - as claims being asserted on behalf of, or applied to, groups of practitioners, rather than pre-existing explanatory categories into which groups could be consigned.[220]

In support of its attempts to create a distinctive boundary between regular and quack practice, the BMJ did not compare the value of the respective therapies on offer. Perhaps fearing the results that might be obtained, the profession simply excluded irregulars. Exclusion of homeopathy was difficult though because 'it is to a great extent practiced by legally qualified men'.[221] The BMJ followed the principles adopted by the BMA,[222] and sought to exclude homeopathic practitioners on the grounds that they brought the profession into disrepute.

---

[218] The seventh edition of the London Dispensatory, written by Anthony Todd Thompson, Professor of Materia Media and Therapeutics at University College London and Member of the Royal College of Physicians of London, includes: infusion of chamomile (used to treat dyspepsia and other stomach complaints); oil of marjoram (used to treat the pain of toothache); Beccabunga, an indigenous, perennial plant, common in rivulets and clear ditches (formerly considered a good antiscorbutic, but falling out of use in 1833); Arnica (used as a tonic and specifically to treat gout, amaurosis, rheumatism, and chlorosis). The British Pharmacopoeia, published under the auspices of the Medical Act of 1858 similarly lists many substances which would be used by medical botanists and homeopaths. The pages of the BMJ make clear the overlap between medical and homeopathic remedies. Nux Vomica, one of the great homeopathic remedies, was proposed for use, in small doses, as a treatment for epilepsy (BMJ 1869;ii:143). Walter Tyrrell, the author of the piece, claimed that strychnia was superior in some cases to potassium bromide, then the standard treatment, and implored his brother practitioners to 'give the remedy a fair trial'.

[219] Cooter 1988 p74. [my emphasis] Comments in the BMJ of the time confirm this view. It argued that the medical profession should not reject hydrotherapy out of hand, despite its apparently anti-medical outlook. Concerning homeopathy, the BMJ made much of the fact that a survey of homeopathic practitioners in Manchester showed that they were not averse to using allopathic treatments, nor did they use infinitesimal doses in the majority of prescriptions. Adopting this approach Nye argues that homeopathic remedies have a more involved history in the development of allopathic remedies than is often recognised (Nye 1990).

[220] Loudon, in his study of the origin of general practice, makes the point that in the absence of a regulatory framework the distinctions were bound to be claims (Loudon 1986 chapter 1).

[221] BMJ 1858; (August 14):689.

[222] outlined at the annual meeting of the Provincial Medical and Surgical Association of 1851. For details, see Nicholls 1988 p 137.

In the case of homeopathy, the behaviour pattern of the ideal practitioner was drawn directly from a gentlemanly code of honour.[223] The BMJ quoted at length, and approvingly, from the minutes of the 1858 annual meeting of the Reading Branch of the BMA:

> *Dr Cowan proposed the following resolution: 'believing that homeopathy is philosophically false, and in practice a dangerous delusion, this meeting resolves that encouragement to homeopaths is incompatible with the honour and interests of the medical profession; and that any member of this branch consulting or co-operating with them will justly forfeit the respect of his colleagues'... Dr Cowan observed that it was due to the profession itself to... cut off from association those who adopted a system which was at variance with all rationality and experience, and which in application was a dangerous delusion'*[224]

Despite the alarm at the incursion of irregular practice, the BMJ of the 1860s and 70s contains no substantive piece on the effectiveness of irregular therapies, preferring to treat the content of homeopathic practice with 'contemptuous silence'[225] while regularly disciplining any of its members who consulted with homeopaths. A directly applied code of gentlemanly conduct, unmediated by knowledge, was sufficient to police the boundary between regular and irregular practice.[226] We shall see below that in the main, statistically based evaluations of therapies were a way of disciplining knowledge within

---

[223] The same is true in the case of homeopathy in France, where a debate about the value of homeopathy in 1835 included no empirical evidence (Weisz 1995 p161).

[224] BMJ 1858; (August 14):689.

[225] BMJ 1862;I:285, in a review of W Roberts: *Homeopathy as practised in Manchester, contrasted with its alleged principles*. The editor of the London Medical Review was forced to defend his decision to publish 'one or two articles from the pen of a gentleman who has made himself somewhat conspicuous on the homeopathic side' (BMJ 1862;I:557) That the approach taken by the BMJ before 1850 to homeopathy and other quack treatments follows the model described here is made clear in Bartrip 1990 p41-46.

[226] The BMJ of 1861 provides several examples. Dr Ozane, a homeopath introduced to the Guernsey militia received 2 columns in March 1861 (BMJ 1861;I:281) The need to 'purge the medical body of alliance with homeopathy' forms the subject of an editorial of April 20th (BMJ 1861;I:422). Further excoriation against professional contact with homeopaths were made on June 8 (p614-615) June 20 (p685-686) July 20 (vol. ii:65).

the profession rather than establishing the boundary between the profession and its challengers.

Regular therapeutics

If professional unity could be achieved through simple exclusion of homeopaths and other irregular practitioners, some other approach was needed to bring coherence to therapeutic practices within the profession. Recalling though that the notion of a profession was still somewhat nebulous in the 1850s, the task faced by the BMJ was to simultaneously create a coherent knowledge of therapeutics and a coherent profession which could use that knowledge.

Before going on to examine the approach taken by the BMJ towards therapeutics, it is necessary to outline contemporary definitions and content of therapeutics in nineteenth century Britain.

This is Robert Christison's definition, addressed to the BMA annual conference in 1858:

> *Therapeutics – the doctrine of remedies - the theory and practice of cures – the conclusive part of the art of healing-that branch of medical learning to which every other is merely prefatory, subordinate, or fundamental-without which everything else in physic, anatomy, physiology, pathology, nosology, are all alike practically useless – the science which teaches the actions and mode of actions, the effects on health and on disease, and the special uses of the thousand articles, medicinal, dietetic, and regiminal, which have been established as the tools of the physician during an experience of two thousand years*[227]

A great variety of therapeutic substances were available in the nineteenth century. The London Dispensatory of 1833 lists several hundred preparations based on animal, vegetable and mineral substances. A few examples chosen at random:

[227] Christison R. The address in therapeutics to the twenty-sixth annual meeting of the BMA. BMJ 1858;August 14:671-675.

Table 1: examples of some nineteenth century therapeutic substances

Substance	Preparation	Usage	Page
Tansy	Infusion	Tonic, antihelminthic. Formerly a remedy for hysteria, but no longer since better remedies are available	628
Oyster shell	Powder	Antacid, but no longer needed since it has the same effect as lime	475
Iron	Solution	Tonic, antihelminthic	790
Tobacco	Infusion	Narcotic, diuretic. However, general effects are too violent for internal use	468
Oxalic acid	Solution	In very small doses, may be used as a beverage in febrile diseases. It is a virulent poison in large doses	692

Source Thomson 1833

In addition to therapeutic substances, treatment might consist of surgery, dietary regimes, rest, exercise, galvanism, and of course blood letting, which was widely used until mid-century, and still had adherents in the 1870s, thereafter falling into disrepute.

The spectrum of therapies was matched by the spectrum of people who might give the therapy. Medical botanists, homeopaths, mesmerists, bone-setters, clerics; all had access to the sorts of *materia medica* listed by Thomson, and indeed to the books on *materia medica*. Within the confines of the regular practitioners the provision of therapeutic substances was most likely to come from one of two sources: apothecaries, who were licensed to do so, and druggists, who had hitherto supplied apothecaries, but who increasingly treated the public directly in the nineteenth century. The apothecaries themselves existed towards the lower end of the medical hierarchy. In his position higher up the hierarchy a physician might be able to earn a living by giving advice and recommending therapies. But the physician class was itself divided. Licentiates of the Royal Colleges of the Physicians might need to dispense drugs in order to make a medical living. Fellows on the other hand would be able to charge sufficiently high fees

for their advice that they had no need to deal with drugs. An apothecary would certainly have to dispense drugs and charge for them rather than his advice. But a druggist would always charge less for drugs than an apothecary, and would be able to throw in some advice *gratis*.[228]

When it comes to describing what therapies were used and how they were given, the picture in Britain is somewhat unclear. The tripartite division of the medical world of the early nineteenth century – physicians, surgeons, and apothecaries – is only partly useful for describing the reality of therapeutic practice. Nevertheless, the sense of hierarchy that accompanies the tripartite division does mark therapeutics very deeply. Clearly, as Rosenberg and Warner have argued, therapeutics is central to medicine and offers a way of understanding several of its dimensions.[229] Their studies were set in the context of nineteenth century America. It is somewhat surprising that no comprehensive history of therapeutics in Britain exists.[230]

---

[228] For more detail on the role of drug dispensing in the lives of practitioners, see Digby 1994.

[229] Rosenberg 1979 and Warner 1987.

[230] The following is a sample from a wide ranging literature. GR Williams surveyed the history of therapies in the Eighteenth century (Williams 1975) and Mark Weatherall has provided an overview of the discovery of pharmaceutical compounds. (Weatherall 1990 and 1996). There are several accounts of specific remedies e.g. Peruvian Bark (Jarcho 1995), Rhubarb (Foust 1992). One of the few patient-centred accounts of therapeutics is Porter and Porter 1989. Various papers and chapters by John Parascandola cover the history of drug use (Parascandola 1977, 1985) Parascandola has also written specifically about pharmacology in Britain in the late nineteenth century (Parascandola 1976). Underwood 1963 and Holloway 1991 & 1995 discuss some of the organisations with an interest in therapeutics; Urdang 1963 p92-110 reviews the history of pharmacy in Britain.. Marland 1987b discusses the activities of chemists and druggists in mid-nineteenth century Britain. Company histories provide some insights into the business of drug development (Tweedale 1990). The early history of experimental pharmacology is discussed by Parascandola 1976. Because they form a contrast to orthodox medicine, there is always an interest in unorthodox and patent medicines. Campbell 1978 describes early nineteenth century attempts to disclose the composition of secret remedies. P.S. Brown uses contemporary material to describe the life of herbalists and medical botanists in Bristol (Brown 1982) and the nature of medical advertising (Brown 1987). There are several accounts of the development of specific remedies, and although these are out of period they provide some insight into the chronology of therapeutic discoveries (Lloyd 1961;Chance 1942). Tring 1977 describes one well known patent remedy, Holloways pills. Histories of more recent orthodox remedies tend to focus on their laboratory development and the problems of introducing scientific remedies, see for example Weindling 1992 Sturdy 1992. The clash between science and experience as a source of authority in therapeutics is discussed by Warner in several papers concerning blood-letting and the use of alcohol (Warner 1980). On blood-letting see also

By all accounts, the prevailing philosophy of health in early nineteenth century Britain was holistic, in the sense that illness was regarded as a problem affecting the whole system of the patient. And often monistic, in the sense that the various ills troubling a patient might be traced to a single cause. Therapies aimed to restore the patient to health by treating the excessive or deficient play of forces in the body.

Even when 'solidism' (the idea that diseases were located in specific organs) began to make its presence felt in the middle third of the century, the philosophy of treatment was to offer a patient a course of treatment - perhaps a tonic or a purgative – that looked beyond the proximate solidistic complaint. Therapy aimed to restore the patient back to health, and in doing so nullify the conditions that were causing the complaint. Implicit in the emphasis placed on the totality of the patient rather than the proximate complaint was the idea that no two patients were the same. Accordingly, two people presenting with the same symptoms might (should) receive a different therapy or combination of therapies.

A specific, that is a therapy used to treat a specific disease, was unusual, and regarded with some suspicion. They were redolent of quackery, since they lent themselves to advertising and required little of the art of physic. Paradoxically, the use of specifics also suggested submission to one of the rational systems of practice, which were held in disrepute since they appeared to encourage a mechanical approach to therapeutics.[231]

Niebyl 1979 and Risse 1982. The development of pharmacopoeias and formularies is discussed in Urdang 1942. In general terms there is much more material on American therapeutics, thanks to the work of Rosenberg (Rosenberg 1979) and Warner. To give just a few examples Dowling has written about the early period of chemotherapy (Dowling 1973); Marks has written about Cortisone (Marks 1992).

[231] For example, the physiological medicine of Broussais, in which diseases were considered within a framework of internal irritation, discussed in Canguilhem 1978.

The art and skill of the practitioner lay in their ability to match the constitution of the patient with a therapy or therapies, in order to restore the person to health and thereby moderate his or her illness. Any illness, even one such as cholera whose symptoms and course were so clear, might be treated with one of several therapies, and several therapies might be given over the course of any one illness. In a complex system the patient's view lay at the heart of the efforts to judge the efficacy of remedies. This is partly because the personalised nature of therapeutics precluded any other perspective. But it is also because the eighteenth century market for medicines was very open, forcing practitioners to compete among themselves for patients.[232] Overall, as Jewson[233] and Porter[234] have argued, patients in the eighteenth and early nineteenth centuries enjoyed an epistemological parity with their doctors, giving those able to choose a variety of options and the ability to make their own judgements of the therapies and practitioners on offer.[235]

Within a framework that differs so much from our own, statements about therapeutic confusion[236] and naiveté[237] in the nineteenth century must be treated with caution. If therapeutics are considered within a historical context, Harley Warner's judgement is definitive:

---

[232] Digby 1994.

[233] Jewson 1974, Jewson 1976.

[234] Porter and Porter 1989.

[235] Porter and Porter 1989 p206-207. The idea that the laity was fit to pass judgement on practitioners' claims was still alive in the minds of the BMA in the 1860s, when the BMJ criticised this aspect of court proceedings.

[236] Keele 1969 p9.

[237] Rawlins 1990.

> *Unmistakably, nineteenth-century therapeutics did work, though perhaps not when judged by criteria of efficacy satisfying to a twentieth-century pharmacologist. Physicians were not ordinarily simpleminded, passive or duplicitous, nor were they unobservant of the effects of their therapies'*[238]

Contemporary judgement has tended to characterise mid nineteenth century as a period of therapeutic nihilism.[239] This is at best a partial reading. It owes much to the pioneering work of RH Shryock,[240] who adopted the highly critical but equally highly motivated viewpoint of the physiologist Claude Bernard.[241]

If instead we adopt the perspective of a contemporary clinician's point of view, it would be just as accurate to highlight the vitality of therapeutics in the mid- nineteenth century:

> *...at any rate we can point to the introduction of some important new remedies ...among the best established of these I may mention iodine, chlorine, creosote, pyroloxic spirit, chloroform, cod-liver oil, glycerine, quina, morphia, strychnia, bebeerina, veratria, tannin, hydrocyanic acid, croton oil, pomegranate root, shield fern, aconite, buchu, cubebs, lobelia, Indian hemp, raw cotton, ergot, besides others of less note or more doubtful reception.*[242]

An increasing sense that regular therapeutics were of uncertain value is undoubtedly a feature of mid nineteenth century medicine, just as the sense of declining status of British medical science had been in the early nineteenth century.[243] But was it caused, as Sir William Jenner put it in 1869, by 'modern advances in science'?[244] Chronologically, the

---

[238] Warner 1986 p4.

[239] For example Paton 1979.

[240] Shryock 1936. In the University of Wisconsin Press 1979 edition, the relevant section is found on pages 248-262.

[241] Coleman 1985.

[242] Christison R Address in therapeutics. BMJ 1858; (August 14):675. Also W.O. Markham. Address in medicine. BMJ 1862;I:86.

[243] Warner 1991.

[244] W. Jenner. Address in medicine. BMJ 1869;ii:114-119. Also Shryock 1979: 'everyone appreciated, by 1850, the significance of the steam engine and of the electric telegraph. The former alone, in its many applications was

idea of a scientific therapeutics came after the appearance of uncertainty in therapeutics, so could not have caused it. Nevertheless in offering itself as the solution to uncertainty advocates of scientific medicine found it useful to present it as the cause.[245]

The greater cause of professional uncertainty was the changing social order of the medical profession.[246] The importance of the BMA lies in its efforts to provide a common solution to both the problem of the social order of the medical profession and the problem of medical knowledge.[247] Jenner, Bernard, and others merely reversed the chronology. In order to open therapeutics to the sorts of scientific inquiry they were advocating they made it appear that science, particularly physiology, had rendered the old therapeutics obsolete.

---

revolutionizing conditions of life throughout the Western world. Except for Jenner's mortal discovery, what did medicine have to offer which was comparable to these? (Shryock 1979 p248-9).

[245] In 1985 Warner (Warner 1985**a**) argued that the historiography of medicine went astray if it was based on the 'scientification' of medicine by laboratory science. The work presented here supports the need for caution when establishing the relationships between clinical medicine and medical science.

[246] Parry and Parry 1976 p134.

[247] The necessity for a common solution to the problem of medical knowledge and medical organisation was recognised at the time in the BMJ. Alfred Lochee, in his president's address to the BMA in 1861, made the following observation: 'Now, I am the more moved to start with this proposition, when I reflect upon the possible future of this Association, and what it is even probably destined to become in its relation to our profession and the public. It is not now simply a provincial society, but one, as its name denotes, desiring to embrace within it the entire number of those who practice medicine within the circuit of the British realms; and if it is ever to be deserving of its name-if it is ever to have an influence commensurate with it - it cannot be as the representative of a class - it cannot be as the avowed or presumed expositor of particular opinions; but it must be as the faithful interpreter and the equal advocate of the feelings and interests of all. Whatever, I say, may be the nature of the work it engages in, this must be the leading features of it; it must be the practical expression of some common sentiment among us; it must be the recognised instrument for achieving some general purpose…I must observe that, in an Association like ours, it is not possible that all the members who compose it should think exactly alike …[however] all questions of great, because of universal interest, are more likely to solved to the advantage of all after being submitted to the judgement of the many. (BMJ 1861;ii:142) The point was also made by W.H. Walshe, in his address in medicine to the Annual conference in 1862. Referring to the cause of advancement of the medical profession in recent times '…no, our advancement has sprung from the substitution of one true for many false systems of study - a one true system which is capable of being efficiently wielded by that multitude of men, endowed with well marked intellectual aptitude, honesty of purpose, zeal of character…' BMJ 1862:ii:141.

The BMJ and the promotion of rational empiricism

The BMJ's solution to therapeutic uncertainty was to promote rational empiricism. Before the BMJ's approach can be discussed it is necessary to say something about the contemporary meaning of these terms because they appear in the literature of nineteenth century medicine in a disconcerting series of guises. Harley Warner has identified a total of eight meanings for *rational* and *empirical* in the therapeutic discourse in the nineteenth century[248] Each term has four meanings. These relate to a positive and critical use of the term in a professional and a methodological context (Table 2). Harley Warner's analysis is based on the discourse of American therapeutics between 1820 and 1860, but the analysis appears to be relevant to the British context. Firstly, although the categories are arbitrary, and readily seen to be overlapping, they do provide some bearings when reading 'the singularly confusing body of rhetoric'[249] that constitutes mid-nineteenth century therapeutics. Secondly, what reads now as a muddled, inconsistent discourse is a reflection of the unstable structure of knowledge at the time rather than on the practitioners themselves. Thirdly, it is apparent that at least some part of the changing pattern of therapeutics was frankly linguistic, revealing itself in the struggle to settle the meaning of words.[250]

[248] Warner 1986 p41–46.

[249] Warner 1986 p41.

[250] For an earlier analysis of the struggle between empiricism and rationalism, see Aschner 1945.

Table 2: the meanings of rational and empirical in mid nineteenth century medicine

		Rational	**Empirical**
Professional	+	Grounded in medical knowledge and traditional wisdom. Synonymous with regular (i.e. part of the regular profession)	Rarely used because of the widely held view that quack medicine was always empirical
	-	Rigid adherence to dogma, excessive confidence in therapeutic systems	Quackish
Methodological	+	Prudent, judicious, based on induction, derived from general principles	Knowledge grounded in experience and observation; unclouded by theory
	-	Separated from experience, linked to therapeutic systems	Excessive reliance on rationalistic systems

Source: Harley Warner 1986 p41-46

What can be called the rational empiricism promoted by the BMJ took the positive implications of both terms. Therapeutics should be rational because professionally it needed to be prudent, judicious and grounded in medical knowledge. Therapeutics should be empirical because methodologically it needed to be firmly grounded in experience and observation, and should not rely on theoretical systems.

The pages of the BMJ around the 1850's and 60's show clearly how it achieved the fusion of two terms that were often regarded as antagonistic, and created an epistemology that was both rational and empirical.

The attack on individual experience

The idea that individual experience is insufficient occurs in several forms in the BMJ of the 1850's and 60's. In his speech about recent successful discoveries in therapeutics to the 26th annual meeting of the BMA, Robert Christison concluded:

> *Allow me to call your attention to only one circumstance connected with these examples of discovery in the more general doctrines of therapeutics. Scarcely one of them has been purely the discovery of one man. Each has required the independent research of several-often indeed of many-inquirers… let everyone, therefore, contribute his share to the general stock of knowledge…*[251]

Later that year, in his introductory lecture to the Liverpool Society, Thomas Inman set out the problems associated with experience.[252] He began his analysis by transferring the proverbial expression 'experience makes men grow wise' from the individual to the

[251] BMJ 1858;(August 14):675 The point that investigation must be a collective enterprise was made, the usual argument being that investigations were too big an undertaking for an individual (see for example BMJ 1862;I:686) The point was made repeatedly in the years around 1860, so that by 1862 the BMJ could say 'there is, however an infallible test by which the true value of the opinion which we hold respecting the value and effects of any particular agent; and that is, by admitting into our argument the opinions of others who have a perfect right to be considered as capable observers as ourselves' (BMJ 1862:ii:230).

[252] BMJ 1858;ii: 923-927 and 942-945.

group, and then observed that it does not appear to apply to the medical profession as a whole 'now three thousand years of age'. In an age of positivism, the argument continued, those in the profession who are using experience to take themselves forward, were in the vanguard. Yet there remained much confusion about the role of experience in medicine: 'a strange jumbling of ideas respecting experience in general, and especially in those departments of knowledge to which experience is commonly supposed to be necessary'.

If it is true that experience did not teach, and Inman used the example of fads in therapeutics to confirm that it did not, it was because of the 'indolence of mind [which] has ever existed in the human race, and will do so to the end of time'. Two of the major failings of medicine could be ascribed to this indolence. Firstly, it ensured 'repose with implicit and unvarying security upon conclusions once formed'. Secondly, it promoted the migration in herds between one new theory and another, or its opposite, the unwillingness to consider new theories.

However, while an openness to experience is a good thing Inman continued, raw or untutored experience was of no use: 'the shepherd knows little of the mountains over which he leads his flock'; likewise 'monthly nurses, midwives, and old hospital sisters, whose time is spent in the closest intimacy with sickness and death, are believed by the multitude to know more than the physician'. Experience was only useful, claimed Inman, to the prepared mind. It required 'steady powers of observation, habits of continued thought, and a power of calm judgement. It demands memory, discrimination, elasticity [of] mind, and self reliance.' Experience will only teach its lessons to a fine sensibility.

Inman then outlined some puzzles. How was it that experience has not perfected, for example, the knowledge of hysteria? Or of the treatment of pneumonia despite the recent aggregation of thirteen thousand cases? How was it that a fact so obvious as the circulation of blood was for seventeen centuries ignored? How could one avoid the *post hoc ergo propter hoc* fallacy? Why was it that two individuals, each guided by experience, may come to directly opposite conclusions about therapeutics?

Inman's main answer was that error occurs when experience is not sufficiently extensive. Given that elsewhere in his paper he observed that thirteen thousand observed cases were not sufficient to establish the proper treatment for pneumonia,[253] this is hardly convincing. Despite its incongruence, the idea that a large experience will improve medical knowledge supported the conclusion Inman wished to draw, which was that medicine must always be enlarging its empirical base:

> *The natural result of all this is that no one can feel that he has ever attained the ne plus ultra in medicine … of one thing we may feel certain, that where there are so many earnest minded individuals at work testing the true value of past theories, and cautiously framing new ones, medicine will ultimately attain as great a certainty as it is possible in the nature of things for her to do.*[254]

This conclusion readily integrates with Christison's - let everyone, therefore, contribute his share to the general stock of knowledge. The result is an epistemology of therapeutics that emphasises the need for the ceaseless, collective accumulation of reflective experience.

[253] BMJ 1858; ii: 925 col2.

[254] BMJ 1858 ii:945.

The attack on the validity of individual experience does not lessen the role of the individual. It re-situates individual experience, in two ways. Firstly, it places individual experience in relation to a real or possible collective series of observations. Secondly, because it stands in relation to other fragments of knowledge, every medical judgement takes on the character of an experiment:

> *From the very infancy of the medical art, one of the most powerful agents in the hands of the practitioner has been the abstraction of blood; but to the end of time, in every new case which comes under treatment, the question must still be ever new, ever unanswered, except by the judgement of the practitioner himself, how much good or how much ill may be done by the employment of this particular agent… and in forming this judgement, however difficult it might be to divest the mind of pre-conceived opinion, and of theory regarding the modus operandi of the practice, there can be no question that the most unbiased mind will come to the most correct conclusion; that the man whose theories are least hypothetical and are based on the largest generalisation of facts will be the least likely to be misled by appearances.*[255]

The effect is to abolish the isolated practitioner and replace him with a normative individual, whose knowledge is now by definition a sample of the totality of actual observers. A collective individual, or at least an individual whose identity is determined by a collective. In this way the variability of opinion exposed in the court might be resolved. The consequence is that medical knowledge becomes thoroughly statistical.

This epistemology creates a collective patient also, whose individuality is replaced by characteristics which can be measured as the degree of relative variation from the hypothetical aggregate-normal figure, and which may as a consequence be accounted for through aggregation. A new figure begins to appear in the BMJ – the table of cases.

[255] A. W. Barclay. On the real value of blood letting in acute diseases. BMJ 1859;I:187. [my emphasis] Also Latham: 'for in medicine nothing that deserves the name experience can be otherwise gained by what deserves the name experiment' and 'it is expedient that medical practice should in every case be conformed to the current idea of an experiment as far as the nature of that case will permit. (BMJ 1862;ii:59.

There were no rules for its construction, and no guarantee that it was meaningful. Nevertheless, for a few years, tables of cases sometimes filled page after page of the journal. Before discussing the importance of tabulated data, I will highlight the ways statistical knowledge was encouraged by the BMJ.

The encouragement of statistical knowledge

Within a structure of thought that is profoundly statistical, actual statistics are a sort of intensification that might crop-up anywhere in support of an argument.[256] In Inman's text for example, statistical arguments are used. They form only a small part of his text, but they carry the argument forward. The knowledge of hysteria is deficient, Inman argues, because not enough empirical knowledge has been brought to bear. Concerning marriage as a treatment for hysteria, and also the cause of hysteria:

> *Statistics recently collected have demonstrated that good has followed marriage in only twenty nine cases out of three hundred; and that in four hundred cases, only one hundred and thirty had any appreciable uterine disease.*[257]

Elsewhere in the BMJ, arguments employing a statistical form were used routinely in a variety of reports.[258] However, it was by no means universally agreed that statistics were an infallible guide to the success or failure of treatment in an individual case:

---

[256] The expectation of statistics was such that their absence might be noted, as in James Russell's account of the excision of the hip-joint, where the outcome of the operation: 'it is difficult to speak with absolute accuracy at present, the statistics of all the operations not being before the profession' (BMJ 1860;I:354).

[257] BMJ 1858;(Nov 6):925.

[258] The BMJ also campaigned for the collection of statistics of disease amongst the pauper population of workhouses. It highlighted the proposals made by BG Babbington to the Epidemiological Society in 1858 (BMJ 1860;I:54), and later that year highlighted the need for medical returns from workhouses in relation to raising the status and income of poor law medical officers (BMJ 1860;I:270-271).

> *To ransack the history of the past, or to collect the experience of the present, whether venesection has been followed by beneficial or baneful consequences in any one disease, can never settle the question beforehand, whether in the very next instance that occurs, the practice ought to be adopted or withheld. It may indeed give some general idea of the probabilities on either side… but in the end we must return to the study of the case we are about to treat, and our judgements must be formed on its own individual merits.*[259]

In view of the concerns expressed by Barclay and others about the relevance and validity of the statistics of therapeutic trials,[260] the fact that the BMJ regularly published the statistics of what it called clinical trials[261] throughout the 1850s and 60s, even after the method had been subject to severe criticism in France,[262] suggests that the meaning of such reports went beyond their immediate appearance.

At its simplest, a clinical trial would consist of a case series. The surgeon Henry Dove of Norwich reported a trial of essential oils for the treatment of puerperal fever as follows:

> *The oil of turpentine has for several years been used in this city [to treat] … puerperal fever with much advantage, and occasionally with almost magic effect…. I have seen the turpentine fail … and I have seen it add to the intensity of the disorder. Considering what a nauseous medicine turpentine is… I was induced to try in its stead, the essential oils, selecting that of peppermint…I have now used this oil in seven cases, and in another case, the oil of caraway, with all the advantages, and none of the disadvantages of the turpentine.*[263]

---

[259] Barclay AW. The real value of blood-letting in acute diseases. BMJ 1859;(March 5):187.

[260] The BMJ did not avoid publishing criticisms of statistics. It published Dr Fenner's critique of the statistics of tracheotomy in croup (BMJ 1861;I(Jan 26)); a critique of the statistics of iridectomy for glaucoma (BMJ 1861;I:280).

[261] The term clinical trial was widely used to describe a variety of inquiries. The word trial was invariably associated with the suggestion that there was some view in favour of the new treatment, or lack of evidence to support existing therapies: 'Aconite having been recently recommended by Mr. Page and Mr. Campbell de Morgan as worthy of fair trial in tetanus, and all old established methods of treatment having failed … I thought myself justified …' (BMJ 1860;I:68).

[262] Rosser Matthews 1992.

[263] Dove H. The essential oils in the treatment of puerperal fever. BMJ 1859;I:287-8 (Reproduced as Appendix 4).

By present standards, observation of seven cases by one clinician would not be regarded as a clinical trial. But clinical trial was the term used for such reports by their authors and by the publisher, and to ignore or dismiss them would be to overlook their contemporary importance. It should not be forgotten that these trials were undertaken by clinicians who were not sure if aggregated experience had any implications, either to their own practice or to the practice of others. The knowledge that the BMJ would likely publish their collected experiences would therefore have encouraged its members to be on the lookout for suitable material from their daily practice.[264]

By present standards the methodology of such trials seems incredibly crude. Read as the precursors of modern clinical trials they display little awareness of the risk of bias from not using a control group, and of not assigning subjects at random. Yet our sense of what they lack should not obscure their contemporary importance. They represent the attempt to demonstrate that a positive accumulation of knowledge from every day practice was possible, and that the measured rhythm of patient care might correspond with that of experimentation; resulting in an open mind, a fair trial, and a judgement passed under the auspices of the BMJ. Clinical trials of the eighteenth century were not unknown. But they tended to occur in special places or under unusual circumstances – on ships, in the army, in dispensaries.[265] The significance of the trials published by the BMJ in the mid-nineteenth century was the implication that the mentality of rational empiricism should be

[264] To further encourage publication, the BMJ would reproduce accounts of clinical trials from books under review, or from other medical journals. For example, on January 23rd 1858, in its reviews and notices section, the BMJ highlighted a trial of glycerine for consumption contained in the new edition of Richard Payne Cotton's 'On consumption: its nature, causes and treatment. London: Churchill, 1858. (p69).

[265] Ulrich Trohler's innovative study shows that the quantification which is usually supposed to have originated in early nineteenth century France can be found in late eighteenth century Britain. Lind's well known clinical trial of lime juice as a treatment for scurvy is shown by Trohler to be only one of many trials which took place in the eighteenth century. (Trohler 1978).

fused with that of general medical consciousness, and conversely, that the knowledge produced by aggregating experience might apply to every day practice. The point of reference for the practitioner as he prescribed should no longer be the individuality of the patient, nor even his own experience, but the collective experience which constituted the knowledge of a profession, which was available on the pages of the BMJ.

Presentation and tabulation of data

The presentation of data in a tabular form was used in the BMJ from its inception. A statistic such as a table of mortality[266] or the relative numbers of male and female deaths from diphtheria in 125 cases[267] could be presented as unexceptional and read without difficulty.[268]

In the case of clinical trials tabulation was not immediately exploited by the BMJ on any scale before the late 1850s. Single case reports continued to be a mainstay of publication.[269] Accounts that did aggregate data continued to use individual cases to add colour to their reports. (See Appendix 4)

As well as publishing the results of small-scale clinical trials, and in accordance with its commitment to accumulating evidence, large-scale statistical inquiries were also reported in the BMJ.[270] These presented the problem of how to deal with the mass of information.

---

[266] for example, the ten year mortality table for Irish districts facing p69 of the BMJ, January 23rd 1858.

[267] BMJ 1858;i:449.

[268] Tabulated data was an essential part of the statistical movement of the 1830s and 40s, which included many doctors. (Cullen 1975) Farr in particular made extensive use of tabulated data; according to Greenwood, Farr's predilection for tabulated data was a consequence of his rudimentary mathematical skills (Major Greenwood 1936). Coleman discusses tabulation as a tactic of the early French scientific hygiene movement (Coleman 1982 p145). For the wider history of the role of record keeping in the public health movement in Britain, see Szreter 1997, chapter 2.

[269] Leonard Sedgwick has no hesitation in calling his report of one patient treated for tetanus by aconite a fair trial. BMJ 1860;i:68-69.

[270] As well as original reports the BMJ began to publish overviews of statistical accounts. Graily Hewitt recommends episiotomy thus: 'a careful survey of the facts on record and the history of the subject generally have led me to the

The report of 276 cases of syphilis treated by calomel fumigation illustrates the difficulties and the interim solution adopted.[271] The data are briefly summarised: 164 cases occurred in women, 112 in men; 'of the total number of cases, 29 left the hospital without being properly discharged, 25 were made out-patients, and the remained were discharged as cured'.

The bulk of the report however consists of a simple tabulation of data. The table includes basic hospital management data, such as the date of admission and of discharge, suggesting that the source of data is the patient records used in the hospital. The inclusion of patient names gives further indication of the derivation of the data, as well as providing each row with an both an identity and a warrant of its authenticity. The eye tracks horizontally across the rows of the first five columns, but vertically down the 6th:

Table 3: selected tabulated data from the BMJ of 1858

Name	When admitted	Stage of disease	Fumigation commenced	When discharged	Result
Mary Ann C	February 27, 1856	Secondary affection	April 17, 1856	July 2, 1856	Cured
Maria C	March 13, 1856	Primary affection	March 13, 1856	April 4, 1856	Cured
Martha S	March 13, 1856	Secondary affection	March 18, 1856	May 13, 1856	Discharged, bad conduct
Sarah B	March 13, 1856	Secondary affection	April 17, 1856	May 29, 1856	Cured
Eliza B	April 10, 1856	Secondary affection	April 19, 1856	June 20, 1856	Cured

Source: BMJ 1858;(July 24):595

conclusion that, where laceration is actually threatened, the proper treatment would be to procure additional space by making a slight incision into the stretched and dilated perineal structures' (BMJ 1861;I:488).

[271] BMJ 1858;i:595-597. The history of therapies for syphilis is the subject of a paper by the social historian Owesi Temkin (Temkin 1955). Temkin, along with his fellow American historians Erwin Ackerknecht, Henry Sigerist and Richard Shryock, was attempting to establish the precedence of social factors in health. In the case of the history of therapeutics for syphilis, Temkin emphasised the influence of what he called the 'social background' in determining the course of syphilo-therapy. In highlighting the role of social factors he distinguished them from 'non-prejudiced clinical observation'. Temkin's approach has been influential. The study of social factors in medicine has tended to exclude by fiat a range of 'cognitive factors' and leave them unexamined. In his 1955 paper Temkin claims that: the enlightenment put an end to magical therapies for syphilis; the controversy about the use of mercury derived from the respective social positions of physicians and surgeons, and that the trend towards the use of guaicum in the eighteenth century was due to economic motives.

The sixth column is the most important. By tabulating cases a cross-sectional interpretation of the body of data suggests itself: a visual statistic of calomel as a treatment for syphilis. Not requiring any further mathematical interpretation, the table displays the weight of evidence in favour of calomel.

The BMJ of the time reports several large scale data collection exercises, some involving tables,[272] while others, because they were so extensive, had to further compress their data into what would now be called descriptive statistics.[273] It is important to recognise that the link between the assembled data in a table and the knowledge it promised did not necessarily involve quantification, since the table did not require further conversion into a numeric form or compression into a statistic. The term which best describes the link embodied in a table is enumeration. By this I mean a sort of assemblage of facts that does not carry with it the requirement to further transform the counted items.

The knowledge contained in a table was valuable because it had been assembled from fragments, each of which was a part of the whole. The table was an end in itself, an act of knowledge. It displayed the weight of evidence available, which might or might not yield a more precise answer. It is true that the data could be further compressed into numerical statistics, but chronologically and epistemologically the table precedes the numeric

[272] BMJ 1862;i:620 (primary syphilitic sores treated with mercury). BMJ 1863;ii:338-341 (excision of the knee joint). BMJ 1864;i:421 (use of skin flaps in amputation).

[273] Harrison for example published the statistics of 1000 cases in obstetrics (BMJ 1859;ii:869-71. AB Granville improved on this by reporting on 'upward of twelve thousand cases in his paper 'on certain phenomena, facts, and calculations incidental to or connected with, the power and act of propagation in females of the industrial classes in the metropolis' BMJ 1860;i:383-4.

summary, and for several years was the method of choice for presenting data in the BMJ.[274]

William Guy's 1860 lectures on numerical method

The BMJ provided further encouragement to the statistical approach by publishing the full text of William Guy's Croonian lectures on the numerical method in medicine.[275] Guy had studied under the French numerical clinician PCA Louis, and was at the time of the lecture series Professor of Forensic medicine at King's College, London.[276]

The substance of the lectures is discussed in Appendix 5. As an argument for the use of numerical method in medicine, Guy's Croonian lecture series has several points of interest. Firstly, he shows by analogy that the numeric method can exist alongside physiology and pathology. The numeric method is the applied art of medicine, while pathology and physiology are the applied sciences of medicine. Secondly, the numeric method offers a rational explanation of the source of variability in therapeutics, and offers a methodology for clinicians to deal with it. Third, in lecture 5 he describes a controlled clinical trial of belladonna as a treatment for scarlet fever. The trial allocated boys to study and control group on an alternate basis. Guy suggests that this sort of trial is useful where one doesn't trust either the sanity or the honesty of the person advocating a remedy. Third, he shows that useful results can be achieved from relatively small sample sizes. In doing so, Guy predates the later work of WG Gossett. Fourth, Guy skillfully

[274] Tables of the sort described here disappeared from the BMJ by the end of the 1860s. By then, tables consisted of rows of statistics rather than of individual patient data. A paper on lateral lithotomy from 1869 (BMJ 1869;i:253-255) contains 16 tables of statistics.

[275] Guy WA. The numeric method, and its application to the science and art of medicine. BMJ 1860; published in six papers:331-334;371-373;409-411;467-469;553-555;593-597.

[276] William Augustus Guy, 1810-1885. From 1838 Professor of Forensic Medicine, Kings College London. Honorary Secretary of the Statistical Society 1843-68. Editor, Journal of the Statistical Society 1852-56. Founding member of the Health of Towns Association. (DNB Vol 23 392-3).

avoids the criticism that the results produced by the numeric method do not apply to individuals. He does this by strongly agreeing with the criticism, and then proceeding to act as if it was irrelevant.

The view of the BMJ was expressed in an editorial:

> *Nowhere is there wanted a more rigid adherence to what Dr Guy in his Croonian lectures, now publishing in this journal, terms the science of numbers, than in medicine. That science thoroughly carried out is alone capable, we believe, of clearly defining the true pattern of disease, and of sweeping away the multitudinous errors which clog our art'*[277]

Therapeutic inquiries initiated by the BMJ

The BMJ also initiated a series of collective clinical enquiries into the efficacy of certain treatments. An epidemic of diphtheria beginning in 1856[278] led to the first therapeutic inquiry initiated by the BMJ. 'We propose to publish a collection of cases of this disease in a tabular form, embracing the principal points in the experience of the London and County hospitals'.[279]

The editorial sets out a list of what would become the columns of its table: age and sex; hygienic conditions; meteorological conditions; evidence of contagion; history of scarlatina; symptoms when first seen; time at which the leathery membrane appeared; presence of fungus in the membrane, as shown by microscopic examination; affection of internal organs; treatment and progress of the disease, with special attention to the

[277] BMJ 1860;i:564

[278] The epidemic was particularly troubling to medical practitioners. The unusual pattern of outbreak and virulence associated with diphtheria caused strong alarm in the public around 1859, who believed a new disease had appeared. For a review, see Hardy 1993 p80-109. The epidemic first appeared in several parts of England in the early autumn of 1856, and spread to London by early 1857. (Greenhow 1860 p70) (Newsholme 1898 p54) In addition to its complex epidemiology, which remained obscure even after the isolation of *Corynebacterium diptheriae* in 1883, patterns of care for diphtheria varied, and there was great controversy about the value of tracheotomy as a treatment in severe cases.

[279] BMJ 1859;I:87.

trachea; indications and success of tracheotomy; duration and event of the case; postmortem appearance in fatal cases.

Such a diverse and poorly defined data collection exercise indicates that the BMJ was seeking to answer several questions through the tabular method. The outbreak of diphtheria in England in 1856 was significant enough to prompt several inquiries and to expose the knowledge of the profession to scrutiny.[280] In response, the BMJ published some case studies and a review of the nature and treatment of diphtheria.[281] Predictably, these revealed the uncertain status of regular knowledge. For one author calomel was 'wholly contraindicated'[282] while another declared that many 'strongly advocate its early employment'[283] As a result, the BMJ initiated its own inquiry into diphtheria.

Publication of the table of diphtheria cases began on April 16th 1859. The result was not a success. The amount of detail supplied under some headings made the table difficult to read. The number of tabulated cases was only 74, but the table occupied 9 pages of the journal. The table was summarised on June 25th.[284] To the tabulated data could be added 55 summaries supplied by Mr. Ridgen of Canterbury, and 133 by Mr. Ellis of Woodhouse Eaves near Loughborough.

The summary of the treatment column shows that the common approach among practitioners was to give a tonic and stimulating regimen. 'The approach was so

[280] In particular, an inquiry into the causes, symptoms, and treatment of the disease initiated by the Privy Council in 1859 (Greenhow 1860 p vii).

[281] BMJ 1859;I:80-81 (case studies) and p81-83 (review).

[282] Laycock T. The parasitic nature of diphtheria. BMJ 1859;i:112-113.

[283] Bernard CE. Diptherite: its nature and treatment. BMJ 1859;I:81-83 Greenhow considered that since 'scarcely any two cases are precisely similar, it would be impossible to define rules of treatment applicable to every variety of diphtheria (Greenhow 1860 p261).

[284] BMJ 1859;I:497-498.

uniformly followed as to deprive us of the opportunity to compare it with any other'. Specific treatments included caustics, chlorate of potash and turpentine. However, no firm conclusions were drawn from the results.

The relative failure of its survey, and the greater failure of a survey under the direction of the Epidemiological Society,[285] led directly to the next phase of the BMJ's support for clinical investigations. The investigations into diphtheria had failed, according to the BMJ, because of the difficulties faced by medical observers when trying to record facts about disease and treatment. The want was not in the capability or aptitude of practitioners:

> *What is required is something directive, by which we could bring the facts and observations which constantly come under notice in every-day practice into scientific use, without interfering with the more immediately serious duties of that practice. This we believe may be brought about by the different medical societies… let them from time to time issue memoranda suggestive of the direction of inquiry, and of the mode in which that inquiry should be conducted.'*[286]

In 1862 the BMJ announced the formation of a committee to investigate the effects of remedies.[287] The committee appears to have been formed in response to a suggestion from Dr Handfield Jones.[288] It met for the first time on 7th August 1862, and was composed of 8 members of the Association, notably William Farr, John Hughes Bennett, and Handfield Jones. At the first meeting it was decided to investigate six subjects. The method of inquiry involved the committee, the members of the BMA, and the journal. A

[285] Reported in BMJ 1860;i:(January 14).

[286] BMJ 1860;i:(January 14).

[287] BMJ 1862;ii:175. It noted that a similar initiative by the metropolitan counties branch in 1857, had come to nothing (BMJ 1862;ii:284).

[288] Committee minutes, published in BMJ 1862:ii:177.

member of the committee would handle each subject. He would prepare a schedule for data collection and publish it in the BMJ along with an article directing attention to important points. Subsequently, he would receive the returns, and draw up a report to be presented to the next meeting of the Association.

The editorial and accompanying paper by Handfield Jones stressed the collective nature of the enterprise. This being so, Handfield Jones continued, the Association was in an ideal position to improve the status of therapeutics:

> *Assuredly we have the workers, and as certainly we have the materials. Our members are widely dispersed all over the Kingdom, in various and greatly different fields of labour;… who, I ask, have better opportunities than we, numbering in our ranks, as we do, hospital physicians and surgeons, experienced self-relying rural practitioners, exact observers of vital phenomenon, accurate chemical analysts, and practised experimental inquirers*[289]

Six therapeutic inquiries were initiated, and a seventh agreed subsequently (Table 4). Given the failure of the diphtheria inquiry to reach any conclusions, the new round of inquiries were structured in a way that might produce clear results. These were to be achieved by narrowing the focus of the inquiry in such a way as to guarantee that the data would be additive. The narrowing of focus affected all aspects of the inquiry. The illness under consideration was closely defined in the introductory essays[290]; the treatments under consideration were restricted to one or at most a few therapeutic substances.

Table 4: therapeutic inquiries initiated by the BMA in 1862

Inquiry	Condition	Therapeutic substances

[289] BMJ 1862;ii:187.

[290] 'It is of the utmost importance, therefore, at the onset of any inquiry into the effects of treatment, that those who take part in it should consent to the definition put forth by the member of the committee who takes charge of any special subject, whether they themselves think the definition strictly correct or not.' (BMJ 1862;ii:433).

1	Acute pneumonia	Antimony, moderate blood-letting, diet, or stimulants
2	Taenia	Oil of male fern, or kousso
3	Psoriasis	Arsenic, alkaline applications, pitch ointment
4	Jaundice	Mercurials, benzoic acid, podophyllon
5	Scarlatina	Chlorine mixture, carbonate of ammonia, quinine, and wet sheet
6	Epilepsy	Atropia
7	Progress of disease	

Source: BMJ 1862 (see text)

The introductory remarks made it clear that the results of the inquiry might be comparative.[291] The idea of an evaluation which compared the effects of remedies was new in relation to substances used by members of the BMA, although Guy had reported a comparative trial between a regular and a homeopathic treatment in his Croonian lectures. WD Farr, the most accomplished epidemiologist on the committee published a paper, introducing a new statistical measure for the efficacy of treatment.[292]

The proposed inquiries exhibited further innovation by associating an illness with at most three remedies. The limitation of the range of therapies under trial was prompted by a concern to make the data yield some definitive guidance, but such a course of action indicated therapeutic specificity, an idea that was not yet widely accepted.

The proposed therapeutic inquiries therefore carried some risk of failure. In his address to the thirtieth annual meeting of the BMA in 1862, Handfield Jones raised the idea that the focus on specific diseases and on specific remedies was justified by the new science of

[291] Farr 1862.

[292] Farr's interest in the design of clinical trials has hitherto been overshadowed by his interest in public health. It is clear though that Farr was applying the same sort of accounting techniques to questions both of public health and therapeutic efficacy. Farr used his actuarial work as a basis for an outcome measure of treatment that calculated the average length of time that a disease persisted at any moment during the course of illness. Applying this measure to different treatment regimens, Farr would be able to determine the amount of morbidity saved by various treatments, in an analogous way to life tables demonstrating the amount and value of life lost. In his 1862 paper (Farr 1862) he used data on small-pox in his possession to calculate the chances of recovery for various times after the disease had taken hold. He also calculated the average of how much longer the disease would last after any particular time. The distinguishing feature of these measures was their strictly statistical character. As such they had no meaning and no reference point to individual patients.

physiology. Before considering the outcome of the therapeutic inquiries therefore, I will briefly consider the role of physiology at this time in the argument for therapeutic inquiries.

Physiology

With Handfield Jones' BMJ paper in 1862, physiology became part of the effort to reform therapeutics. The role of physiology is quite specific in Handfield Jones' argument. It demonstrates that substances have a specific action on bodies, if those bodies are examined at the level of their tissue. Physiology therefore provided a scientific justification for the use of specifics. Physiology can also explain why drugs affect bodies differently. Opium, for example:

> *Is to be regarded chiefly as a toner or exciter of nerves; first, and in small doses, of the cerebro-spinal; and secondly, and in its more potent action, of the sympathetic. The second action is, in some of its results, counter to the first. It causes cerebral amnesia and sopor in precisely the same way as it stills an intestinal profluvium.*[293]

Handfield Jones shows how the seemingly contradictory actions of digitalis reported by empirical observation can be reconciled through experiments on living tissue. The old language – of tonics and depressants – is seen to be not wrong, but superficial compared to the deeper understanding afforded by physiology. Physiology also provided an explanation for why the effect of a drug varies on different types of personality, provided that personality is expressed in physiological terms:

[293] BMJ 1862;ii:191.

> *On the above view we can readily understand how opium so often disagrees with persons of irritable and weakly nervous systems, acting, perhaps, in the manner opposite to that intended. The cerebro-spinal nervous areas are more acted on than the vaso-motor nerves; and the result is therefore wakefulness and excitement*[294]

The role of physiology in therapeutics is captured in a phrase that occurs in an anonymous review of Marsh's '*Special therapeutics'* published in 1863. About half way through the review the skin is described as a 'physiological agency', the reviewer adding that it is surprising no one had noticed this before.[295] The relationship between therapeutics and the body was now to be mediated by physiology, which displaced the individuality of the body by breaking it down into an organic structure composed of common tissue surfaces.[296] The new relationship between the therapeutic substance and the body was established at the level of the tissue. It does not wholly do away with the old relationship between the patient and a therapy, but displaced it with a specific, generalised relationship between therapy and tissue.

In his study of the relationship between clinical medicine and physiology, Geison points to the highly complex relationship between the two.[297] As he notes, it is easy to find, at one level, frankly dismissive opinions held by each party of the others' claims. And at the same time, but on another level, the two disciplines appear to need each other: physiology needing the medical base to show its relevance, and medicine needing physiology to establish its scientific credentials. Geison raises the question of how, exactly, physiology

---

[294] BMJ 1862;ii:191.

[295] BMJ 1863;ii:345.

[296] The practical and the symbolic importance of animal experimentation is that is shows a continuity, beneath the level of individuals, of common tissues.

[297] Geison 1979.

benefited medicine.[298] In the case of therapeutics, the answer is clear. Edward Woakes, a provincial practitioner, places the hopes of the profession in physiology thus:

> *We want some recognisable parallelism between the symptoms to be assailed and the weapons brought to bear upon them. Let me not be misunderstood to aspire after the exactitude of the homeopath, who has his fixed remedy for every symptom… but rather to seek for a correspondence between the principle which the disease brings into play, and the principle upon which we make use of a remedy.*[299]

In Handfield Jones' 1862 paper physiology is fully integrated into the discourse of therapeutics. A physiological understanding of therapeutics underpins the association between remedies and illnesses. By operating at the level of tissues and cells it suggests a uniformity at the level of person which the therapeutic inquiries require if they are to be able to analyse their data in terms of aggregated individuals.

The fate of the Therapeutical Committee inquiries

The therapeutical inquiries were of considerable importance to the BMJ in the early 1860s. The requirement was to find some way of generating a scientific knowledge that could be created and shared among the members of the BMA. Physiological studies, 'carried out of course by younger men, who could devote more time to scientific research' would provide a sounder basis for understanding the effects of drugs on the body. Such work could not be left to private enterprise because it was too great a task for isolated researchers.

[298] Geison 1979 p68.

[299] BMJ 1862;ii:239.

It was also 'the duty as well as the interest of our medical corporations to take an honourable lead in placing our art on a basis of scientific truth it has not yet occupied.'[300] As to the prospect that collective inquiries might not be productive:

> *'In all labour there is profit. God denies nothing to well-directed diligence, as an old author says. It is impossible, if we work intelligently and earnestly, that we should fail to achieve some good results.*[301]

Diligence was not enough. In order to create the possibility of precise knowledge the committee had to construct trials that were artificial in relation to the prevailing therapeutic philosophy. The failure of the diphtheria inquiry had shown that simple aggregation of experience was not enough to resolve therapeutic questions. The lesson of the diphtheria inquiry was that one either had loosely defined trials which fully respected the variability of clinical judgement, but which could not produce consensus, or one had closely defined trials which might take knowledge forward, but which might also appear so artificial as to be irrelevant to daily practice.

The limitation of Handfield Jones' approach was seized upon immediately by the physician Horace Dobell.[302] The trials proposed had no value, he asserted. Firstly, because the tabular approach did not permit practitioners to record the specific details of the case in enough detail for others to profit from a horizontal reading of the table. Citing the use of alkalis in the treatment of rheumatic fever, in his experience, where alkalis had not worked, ' either the alkalis have come too late, or have not been properly used'.

[300] BMJ 1862;i:686. The BMA initially envisaged that it could meet the costs of medical research itself. Following the success of physiological researches into the cause of the cattle-plague in 1866 it was argued that the Government should subsidise medical research. In 1870, £2,000 was made available annually.

[301] BMJ 1862;i:687. Also BMJ 1862;ii:175 'here then is, is laid open before the Association, a great and most useful work… every dose of medicine prescribed may furnish a useful fact, if only the circumstances attending its administration and the results are carefully noted'.

[302] BMJ 1862;ii:238.

According to a perspective that was entirely conventional for the time, the effectiveness of a therapy was as much dependent on the skill and judgement of the clinician as any intrinsic effect of the remedy. Dobell's second argument concerns the uniqueness of each patient:

> *'What should we think of a farmer who, wishing to prove the value of a specimen of seed, should sow twenty bushels of it on twenty patches of land of the same geological constitution, and then estimate the value of the seed by the aggregate harvest? Is it not clear that he would be wrong; and that if he repeated the experiment fifty times with the same seed he might be no nearer the truth; unless he noticed not only the geological constitution of the soil, but its condition with relation to the succession of crops, the kind and amount of manure put upon it, the season of sowing, the birds and insects upon each patch etc?*[303]

Dobell's second argument again reflects the preeminence of individualism in therapeutics. It does so here by focussing on the uniqueness of the individual rather than the singularity of the disease/physician/treatment combination. Dobell's choice of agricultural metaphor is perhaps fortuitous, but it does relate closely to a problem solved by RA Fisher 50 years later.

At its first meeting the Committee on the Actions of Medicines resolved to report its findings to the BMA annual meeting, due in Bristol in August 1863. The schedules were published as shown in Table 5:

[303] BMJ 1862;ii:238.

Table 5: Schedules for the therapeutic inquiries		
Inquiry No.	**Condition**	**Schedule published in BMJ (1862)**
1	Acute pneumonia	October 25th
2	Taenia	November 22nd
3	Psoriasis	November 8th
4	Jaundice	December 20th
5	Scarlatina	November 29th
6	Epilepsy	Not published
7	Progress of disease	Not published

Source: BMJ 1862

Schedules were due to be returned by 1 July 1863,[304] to allow time for the committee to prepare a report on each topic. However, no reports were presented to the BMA Annual meeting in Bristol in 1863. Instead, a small notice recording the names of those who had made returns to the scarlatina and jaundice inquiries.[305]

And on the Friday, a report by Hughes Bennett outlining the sparseness of response (Table 6, col. 2):

Table 6: Schedules returned to the Committee		
Study	**Schedules returned (cases) by July 1863**	**By 1866** **ns = not stated**
Pneumonia	15(55)	21(152)
Taenia	18(ns)	18(100)
Jaundice	7(ns)	9(23)
Scarlatina	6(ns)	Ns
Psoriasis	Ns	3(3)

Sources Column 2:BMJ 1863(August 15):193 Column 3:BMJ 1866;(August 18):184

The failure of the therapeutic inquiries suggested by the results in Table 6 was made clear in Hughes Bennett's address to the BMA's 34th Annual conference held in Chester in 1866. His speech begins in the now familiar manner by outlining the virtues of the

[304] With the exception of the psoriasis study, which was held open until 1st January 1865.

[305] BMJ 1863;ii:132.

numeric method in medicine, and observing how the only good trial was one that encompassed varied clinical experience. However, the final number of returns to the therapeutic committee had been very low (Column 3 of Table 6). Considering that the BMA had a membership of over 2,000 during this period Hughes Bennett believed it would be possible for the membership to amass 1,500 cases of pneumonia and thereby resolve any uncertainty.

The end of therapeutic investigations and the beginning of medical research

Nevertheless, Hughes Bennett did not call for further returns of schedules. Instead he turned to the question of funding. 'The recent Government Report on the Cattle-Plague, for instance points out how the co-operation of various individuals may be so directed as to exhaust a medical inquiry'[306] Clearly, Hughes Bennett regarded the possibility of remuneration for the time spent on inquiries as essential if inquiries were to proceed to a conclusion.[307] His address ends with a plea for the BMA to set up a committee to secure 'chemical, histological and pathological research, combined with accurate, uniform and extended observation'.[308]

Summary and Conclusion

During the 1860s the BMA attempted to reform the circumstances of the medical profession by changing the way it organised and communicated medical knowledge.

[306] BMJ 1866;ii:184.

[307] The difficulty of raising funds for research is made clear in a reference to the unwillingness of the Medical Council to provide funds to Henry Acland, Professor of Physiology at Oxford University, to test the properties of drugs. The Council had the responsibility of publishing from time to time the standard pharmacopoeia for the medical profession, but refused to fund Acland on the basis that it was 'no part of its business to make such investigations'. (BMJ 1866;ii:185).

[308] BMJ 1866;ii:185.

These efforts included the first organised clinical trials, called therapeutic enquiries, whose aim was to establish the value of particular therapies in several conditions.

In reforming the epistemology of the profession, the BMJ promoted the use of simple statistical methods, which offered a way of accumulating knowledge among its members and of disciplining their beliefs.

The failure of the therapeutical enquiries can be ascribed to two factors. Firstly, Handfield Jones considered that a better response would have been obtained had some remuneration been available for filling in the forms. Secondly, among many members of the BMA there was resistance to the idea that data on effectiveness could be aggregated in a meaningful way.

In place of enquiries organised by the BMA, Handfield Jones pointed to the possibility of enquiries and researches being organised by the State. His reason for doing so was the success of the government-funded research included in the work of the Royal Commission on Cattle Plague, published in three reports between 1865 and 1866.[309] Among their work, the Commissioners arranged for various scientific investigations of the disease. One study, carried out on the farm of Robert Lowe, a Commissioner and future Chancellor of the Exchequer (1868-1873, as Lord Sherbrooke) was a clinical trial of sorts, in which a group of cattle were deprived of carbolic acid disinfection of their sheds, and rapidly succumbed to the plague. This research, paid for by the British Government, and providing dramatic and useful results, can be seen as the origin of state-

[309] Cattle plague, or Rinderpest, a highly contagious viral disease, struck first in Islington in 1865 and rapidly spread throughout Britain. By the end of 1865 the plague had resulted in 70,000 deaths, causing great financial loss. A Royal Commission was appointed in September 1865. During the course of the epidemic the commissioners presented three reports to Parliament. For details see Romano 1997.

supported biomedical research in Britain. In 1870, under Lowe's Chancellorship, Parliament approved the sum of £2,000 towards Auxiliary Scientific Investigations as a separate vote in the budget of the Education Committee of the Privy Council.[310] In the next chapter the role of state sponsored research into the effectiveness of therapies is explored, through the work of the Medical Research Council.

[310] For the events leading up to this, particularly the role of John Simon, see Sourkes 1993.

Chapter 3

Evaluating therapies in the 1930s: the evidence of the MRC Therapeutic Trials Committee

Introduction

This chapter describes the style of clinical trials promoted by the Medical Research Council in the 1930s. The chapter begins with a general introduction to the Medical Research Committee/Council (MRC)[311], emphasising those aspects of its early programme particularly concerned with therapeutic substances. It then describes the research sponsored by the MRC's Therapeutic Trials Committee (TTC), the main committee of the MRC responsible for the organisation of clinical trials in the period before the streptomycin trial. The chapter concludes with some provisional explanations for the methods adopted by the TTC.

The first section of this chapter draws extensively from the two principal published accounts of the early MRC. These are A. Landsborough Thomson's two volume semi-official history of the MRC,[312] and a more recent collection of essays edited by Joan Austoker and Linda Bryder.[313] In addition, WR Merrington's history of University

[311] Between 1913 and 1919 the supervisory body of the fund created by the 1911 Act was the 9 member Medical Research Committee of the National Insurance Commissioners, which first met on 24 July 1913, at the home of the Chairman, Lord Moulton, 57 Onslow Square, South Kensington. The Ministry of Health Act of 1919 transferred the powers and duties of the Insurance Commissioners to the Ministry of Health, with several exceptions, including the duties of the Medical Research Committee, which were to be transferred to a committee under the supervision of the Privy Council. An order of the Privy Council, 25 March 1920, set out a charter and the ways and means for the operation of a committee to be known as "The Medical Research Council". The constitutional differences between the Committee (1913-1920) and the Council (1920-) are less important than their similarities of purpose and de facto operation. The term MRC will therefore be used for both, bearing in mind that it refers to a committee of the National Insurance Commissioners before 1920, and The Medical Research Council after.

[312] Sir A Landsborough Thomson, Second Secretary to the MRC. Joined the MRC in 1919. Effectively second in command at the MRC until his retirement in 1957.

[313] Austoker and Bryder 1989a

College Hospital[314] has been used to give further detail on the MRC's first clinical research department.

The Medical Research Council

The origin of the MRC can be found in Section 16 of the National Insurance Act of 1911, where provision was made for 'one penny in respect of each person payable out of moneys provided by Parliament towards the expenses of sanatorium benefit to be retained by the Insurance Commissioners for the purposes of research.' [315]

The positioning of the provision within the section on sanatorium benefit raised the question about whether the funds for research should be devoted solely to the study of tuberculosis. Landsborough Thomson later justified the immediate broadening of the scope of the MRC's research programme by claiming the positioning of the sub-section on research as a matter of drafting convenience,[316] but in a subsequent paper accepted that the funds were originally intended expressly for research into tuberculosis.[317]

The facts are that the Departmental Committee on Tuberculosis, set up in 1912 by the Treasury to advise on the sanatorium benefit aspects of the 1911 Act, obtained differing interpretations of Section 16 from the Legal Adviser to the Insurance Commission and the Law Officer's Department. The Committee concluded that funds should be devoted to tuberculosis research, 'at least for a few years'.[318]

---

[314] Merrington 1976.

[315] 1&2 Geo. 5. Ch. 55.

[316] Landsborough Thomson 1973 p12.

[317] Landsborough Thomson 1973. The question of why research was included in the section of sanatorium benefit has been given a new direction by Worboys' discussion of why tuberculosis was singled out for relief in the 1911 Act. His suggestion is that the clause on research was included as a way of deflecting the criticisms of those who objected to the use of funds to support sanatoria (Worboys 1992 p65).

[318] Cited in Bryder 1989 p5.

In the event, the original programme[319] of the Medical Research Committee was completely overshadowed by the First World War. In 1914 the Treasury gave approval for expenditure on any research in connection with the war, and further debate about the intended limitation on the purposes of the fund did not take place thereafter.

During the First World War the efforts of the MRC focussed on supporting the Army Medical Service, through research on the treatment of wounded soldiers, investigation of infectious diseases at the frontline and the preparation of anti-typhoid and other vaccines for use by the armed forces.[320] Central to this work was the bacteriologist Almroth Wright, working out of St Mary's Hospital in London.[321] Approximately 10 million doses of the anti-typhoid vaccine developed in his laboratory were made available.[322] Wright also set up an army hospital laboratory in Boulogne to study the bacteriology of wound infection.

The general course of development of the MRC after the First World War owes much to the National Institute for Medical Research (discussed below), but also to the personal influence of Walter Morley Fletcher, appointed to the post of Secretary in 1914.[323] As Secretary, Fletcher's influence extended to every aspect of the scientific policy,

[319] Reproduced in Landsborough Thomson 1973 p28-29.

[320] Landsborough Thomson 1975 chapter 14.

[321] Sir Almroth Wright, 1861-1947, bacteriologist. His pupil and colleague, Leonard Colebrook recounted his life and work in Colebrook 1954. For a description of vaccine therapy, and a less partial assessment of Wright, see Worboys 1992. The fate of vaccine therapy is discussed in Keating 1988.

[322] The episode is recounted in Colebrook's biography. The statistics concerning the efficacy of Wright's vaccine were challenged by Greenwood and Yule in 1915 (Greenwood and Yule 1915), who were following up an earlier clash between Wright and Karl Pearson over the interpretation of statistics. The debate between Wright and Greenwood/Yule was an important episode in the developing relationship between clinicians, statisticians, and the State. See Appendix A of Colebrook 1954 and Rosser Matthews 1995.

[323] Sir Walter Morley Fletcher FRS 1873-1933. A graduate of Cambridge University, Fletcher was medically qualified, and a member of Foster's school of physiology, where his researches into the metabolism of muscle led to his election as a Fellow of the Royal Society in 1915.

organisation and administration of the MRC between his appointment in 1913 and his death in 1932. He was instrumental in ensuring the independence of the MRC from the newly created Ministry of Health. From 1920 onwards the MRC was responsible to the Privy Council rather than any department of state. Perhaps inevitably, the MRC and Ministry of Health came into conflict, and relations between the MRC and the Ministry during the 1920s have been described as tense and competitive.[324] Conflict came to a head in 1923 over the research initiated by the Ministry's Cancer Committee. The dispute about their respective roles led eventually to an official concordat between the MRC and Ministry of Health, which set out their respective spheres of interest. The Ministry was to be concerned with applied research relating to practical health problems, the MRC with basic bio-medical science. In her discussion of Fletcher's role at the MRC, Austoker has emphasised the importance of the concordat, which freed the MRC from any obligation to undertake a purely practical programme of research.[325] It was therefore able to pursue a programme of research in its new laboratories at the National Institute of Medical Research NIMR).

To a large extent the NIMR was the MRC in the early 1920s, and as such it fits almost too easily into sociological models which identify laboratories as a vital site for the construction of facts.[326] In broad terms, the rapid proliferation of specialised research institutions in all the major cities of Europe between 1888 and 1893 supports a sociological reading of laboratories as a concentration of power. But there remains the

[324] Austoker 1989 p25-26. The Ministry's research programme was small compared to that of the MRC. Nevertheless it produced 90 research reports between 1920 and 1939.

[325] Austoker 1989 p26.

[326] The classic statement is Bruno Latour 1987 chapter 2 in which the laboratory is portrayed as a site where facts are constructed, stabilized and mobilized, and through which resources are deployed and allies recruited. Latour's work is a theoretical extension of field research conducted by Latour and Woolgar at Jonas Salk's laboratory in 1976/77.

question of how laboratories were able to act as they did and effect 'the laboratory revolution', in medicine.[327] The answer, which Latour proposed in his study of Pasteur,[328] and to which the history of the MRC lends some support, is that laboratories extend their range beyond their physical limits by reproducing their techniques in the community. In the case of the MRC, although priority was given to basic sciences, the programme did not neglect more applied topics such as nutrition studies, wound management, and industrial medicine. A feature of the MRC's approach to applied research was the employment of laboratory techniques. Two examples will be cited in support of this characterisation. The first is HC Corry Mann's feeding experiment on boys undertaken during the early 1920s. Although the study was threatened with curtailment due to its excessive cost, Fletcher regarded it as an important contribution to public health because it extended, in his view, the precision of laboratory studies to human subjects. In so doing, it supported the MRC's position that improvements in micronutrients rather than housing or income were the key to improvements in public health.[329] The second, of which more later, is the attempt to model the spread of epidemic diseases in man using

[327] A phrase first given prominence in Arckerknecht 1967. For a summary of the changes that form part of the laboratory revolution, see Bynum 1994 chapter 4. Further detail and extensive references are contained in the various contributions to Cunningham and Williams 1992.

[328] Latour 1988, and also Latour 1992 p297-299.

[329] The Corry Mann feeding experiment is discussed at length in Petty 1987 and Petty 1989. The study took place at the Dr Barnardo's Home in Woodford, Essex, also known as the Boy's Garden City, one of several designed to remove sickly children from urban society and set them to live under country conditions. (Baker 1912 quoted in Smith FB 1988 p16) Corry Mann's study was published by the MRC in 1926 (Corry Mann 1926). The report includes a chi-squared statistical analysis by Greenwood and Newbold as an appendix. However, it is not clear if the boys were assigned at random to the study and control groups. Closer analysis of Corry Mann's data by Petty shows that the extra milk group contained an excess of stunted boys. But unlike the Lanarkshire milk experiment of 1930 (Leighton and McKinlay 1930), where it is known that the teachers intervened to ensure that the most needy children received extra rations, it is unclear whether or not the boys were assigned according to a stratified randomised allocation or a matched scheme. Corry Mann's description is 'as far as possible an equal number of the same age and rating [a categorised age-height-weight scale was used] were assigned to each of the three houses' (Corry Mann 1926 p9). The use of children in orphanages and homes as experimental subjects was a common practice. (Lederer 1995).

animal models, begun by William Topley[330] in the early 1920s, but developed by Greenwood and Bradford Hill in the 1930s.[331]

In summary, laboratory studies provided the MRC with an ideal for the organisation of every aspect of its research programme, helping to define its contribution in relation to that of the Ministry of Health. The NIMR was at the centre of its approach to therapeutic drugs. Before considering the implications of laboratory studies for the MRC's testing of therapeutic drugs, it will be helpful to outline the development of the NIMR. Doing so will highlight the relatively minor role played by the statistics department.

National Institute for Medical Research

Among the earliest decisions of the MRC was the creation of a 'central institute' somewhere in London.[332] The first scheme of research submitted by the MRC in November 1913 for ministerial approval contains the following section:

> *The organisation by which this research will be carried out should consist of the following departments:*
>
> *1 A competent body of investigators of the highest class in the permanent employ of the scheme and devoting their whole time to research under it. They would be supplied with proper laboratories, duly qualified assistants, etc., and would ordinarily carry on their researches in such laboratories.*
>
> *2 Skilled investigators in the permanent or temporary employment of the scheme….*
>
> *3 Individual investigators not in the employment of the scheme…*

[330] William Whiteman Carlton Topley (1886-1944). Pathologist and bacteriologist. Joined Greenwood as Professor of Bacteriology and Immunology at the London School of Hygiene and Tropical Medicine 1927.

[331] Greenwood 1936.

[332] Landsborough Thomson 1973 p26.

> *4 Statistical Department. This will mainly consist of persons in the permanent employment of the scheme who will be engaged in enquiries relating to diet, occupation, habits of life and other matters bearing upon the incidence of disease, the relative frequency of special types of lesions in diseases such as Tuberculosis, and in general with all statistical investigations useful either as preliminary to research or confirmatory of its results. It will possibly have to advise how the statistical material provided for under the act should be dealt with…*[333]

As Landsborough Thomson notes,[334] the NIMR, unlike the earliest biomedical research institutes,[335] was conceived as a truly multi-disciplinary research institute, oriented towards the advance of medical knowledge along a broad front rather than the solution of particular problems or the application of a particular research discipline.

The MRC acquired premises for its research centre in 1913, on the site of Mount Vernon hospital in Hampstead. However, plans to occupy the buildings were disrupted by the outbreak of war. The fledging MRC laboratories were handed to the Army Medical Service, which ran the site as a military hospital until 1920.

In the summer of 1914 the MRC appointed the heads of section for the departments of the Institute. They were: Sir Almroth Wright (Bacteriology); Henry Dale (Biochemistry and Pharmacology); Leonard Hill[336] (Applied physiology); and John Brownlee (Statistics). When the departments moved to Mount Vernon in 1920, Bacteriology came without Wright, who remained at St Mary's. Wright was considered for the post of Director of

[333] MRC First Scheme of Research 1913, in Landsborough Thomson 1973 p28.

[334] Landsborough Thomson 1973 p109.

[335] The Institut Pasteur in Paris (founded in 1888), the Institut Robert Koch in Berlin (1891) and the Lister Institute of Preventive Medicine in London (1891) The origins and early work of the Pasteur and Koch Institutes are described in Weindling 1992 Robert Kohler has pointed to one area of research more or less abandoned by the NIMR, that of bacteriological research, which was almost completely superceded by virology by 1930 (Kohler 1986 p73).

[336] Father of Austin Bradford Hill.

NIMR, but was passed over. For some time the Institute functioned without overall leadership, but eventually Henry Dale assumed the post of Director.

Under the idiosyncratic leadership of Brownlee, the statistical department did not flourish.[337] During the First World War, the statistical department was housed in rented accommodation in Bloomsbury. The MRC had offered to support the War Office by collecting statistical information about the sick and wounded. This enterprise expanded the department considerably. At one point over 100 clerks were employed, and the volume of records grew so large they required housing in the basement of the nearby British Museum. In addition, the costs grew out of hand, and, as Landsborough Thomson makes clear, very little of scientific value resulted from the undertaking.[338]

In 1920 Major Greenwood was seconded from the Lister Institute in order to provide more effective statistical advice within NIMR.[339] However, in 1927 he was appointed to direct the Department of Epidemiology and Vital Statistics at the London School of

[337] This view was widely held at the NIMR and MRC. Following Brownlee's untimely death in March 1927, rapid action was taken to break up Brownlee's massed data collections and meteorological instruments, and transfer the suite of rooms they occupied to use as laboratories. Only a few days before his death a confidential minute from Fletcher to the Council noted that 'the Council have already agreed on a past occasion that they do not attach importance to the retention of the Statistical Department within the National Institute itself… at present the Statistical Department costs about £3500 a year… it occupies three or four of the best ground floor rooms in the Institute, the liberation of which would give added laboratory and office room now much needed, and set free the Board room for Council and other meetings, which is at present temporarily used as an office' (FD1 7107 March 15th 1927. FD references are held at the MRC Archives at the Public Record Office, Kew) A year after his death, in a minute on the position of medical statistics, Greenwood's view was that 'poor Brownlee, a profound thinker, was constitutionally unfit to organise or effectively criticise' (FD1 7108 A memorandum on the present position and prospects of statistics and epidemiology. M Greenwood, February 11th 1928). Henry Dale's notes on the history of NIMR record that '[Brownlee] had proved to be entirely useless to members of other Departments needing statistical advice or co-operation' Source: Austoker and Bryder 1989 p50 fn. 69.

[338] After the war ended this section of the statistical department continued to compile statistics, but received an increasing number of individual enquiries, chiefly from the Ministry of Pensions concerning claims made by ex-soldiers with no army papers. The section was transferred to the Ministry of Pensions in February 1921. (Landsborough Thomson 1975 p277-280).

[339] For example, he provided a statistical appendix to Corry Mann's study of school-boy diet (Corry Mann 1926) There is yet no comprehensive biography of Major Greenwood, who can claim to have been the first medical statistician in Britain upon his appointment to the Lister Institute in 1909, and who did as much as anyone to found the discipline of medical statistics. For an account of his life see Hogben 1950. For a discussion of Greenwood's views on the role of statistics based on his correspondence with Raymond Pearl See Matthews 1992 chapter 5.

Hygiene and Tropical Medicine. When Brownlee died unexpectedly in 1927 the MRC took the opportunity to disband the Statistical Department. Thereafter, statistical advice to the NIMR came from Greenwood's unit, and from the MRC's Statistical Committee.[340]

The distinguishing feature of the NIMR was its orientation towards understanding health and disease in terms of chemical interactions and physiological mechanisms. Notable early work included Dale's research on the role of acetylcholine in the transmission of nervous impulses to voluntary muscles, for which he was awarded the Nobel Prize with Otto Lowy in 1936. Other examples include the research on insulin, and on the chemistry of steroid compounds. Such work resulted in the identification and synthesis of sex hormones undertaken by Alan Parkes at the NIMR laboratories from 1932; and on accessory food factors, later known as vitamins, especially the work of Edward Mellanby on rickets, and its prevention through the provision of vitamin D.[341]

Clinical Research at the NIMR

The earliest plans for the NIMR at Hampstead included a research hospital, which would provide suitable clinical material alongside the laboratory-based departments. At an early stage, when a plan for locating the NIMR within the Lister Institute was under consideration, the creation of a clinic with 50 beds adjacent to the Lister Institute seemed

[340] The Statistics Committee was formed from the Industrial Health Research Board Statistical Inquiries Committee, which found its remit widening in view of Brownlee's limited interest in providing practical advice. The Committee met 49 times between 1921 and 1950, and was finally disbanded at Bradford Hill's request in July 1961 (FD1 7115 Bradford Hill to RL Cohen 14th July 1961). It's first chairman was Major Greenwood, Karl Pearson having turned down an invitation to chair the Committee on account of his age and commitment to the editorship of Biometrika (see papers in FD1 7114). Initial membership included Leonard Hill (Bradford Hill's father) and G Udny Yule, the statistician from Cambridge University with whom Greenwood had worked on papers promoting the utility of statistical methods in adjudicating bacteriological disputes. Bradford Hill became vice-chairman in 1945 and chairman in 1948 (FD1 7113 extract from MRC minutes 15/10/48).

[341] Became Secretary to the MRC in 1933 following Fletcher's death.

likely. But when the plans to site the NIMR within the Lister Institute fell through, the offer from Lord Ivegh to fund the research hospital fell also. In any case, the MRC was by then worried that a special research hospital might acquire the reputation for being more interested in experimentation than medical care.[342]

In place of a department within NIMR, the Committee funded a research department at University College Hospital which was conveniently located near NIMR, led by Thomas Lewis, its first full time research clinician. The name 'Department of Clinical Research' was used only by the MRC. Within University College Hospital Lewis's department was known as the cardiographic department, reflecting the fact that Lewis's research at University College pre-dates his involvement with the MRC.[343]

At the same time as the MRC was establishing a Department of Clinical Research, the University of London was implementing the Haldane Commission's recommendation that academic clinical professorial units should be established in teaching hospitals. The evidence presented to the Haldane Commission by the Dean of University College Hospital set out the requirement for research facilities in hospitals:

[342] Landsborough Thomson 1975 p14.

[343] Thomas Lewis 1881-1945. Lewis's early career, described in Howell 1984, included the development of the electrocardiograph for clinical use, and other studies of the mechanism of the heart. He received a knighthood in 1921 for his MRC sponsored work on breathlessness of cardiac origin. Merrington 1976 discusses events at University College from the hospital's point of view.

> *'Increased facilities for research in clinical medicine, surgery, and obstetrics, both in their general and special departments, are urgently required. For this purpose, funds are needed in order to enable systematic research in these subjects to be carried out on similar lines to those possible in properly equipped physiological and pathological laboratories. University professors should be appointed in various subjects. Such University professors should be provided with properly paid assistants and adequately equipped departments…'*[344]

The Haldane Commission Report was published in 1913, and the University of London proposed the establishment of professorial units in medicine, surgery and obstetrics. The First World War delayed the implementation of the report. In 1919 the Government offered financial assistance for the formation of clinical professorial units at St Bartholomew's, St Thomas's, London and University College Hospital.[345]

The relationship between the professorial and MRC units at University College was very close. The endocrinologist TR Elliott had been appointed as a second senior researcher alongside Lewis at the MRC unit in 1919, but left that post shortly afterwards to head the professorial unit.

[344] Cited in Merrington 1976 p117.

[345] Funds were made available in connection with the plan for post war reconstruction, using funds from the Board of Education. In 1920 the professorial unit at University College Hospital was part beneficiary of a substantial bequest from the Rockefeller Foundation, amounting to over £1M. Why such an enormous bequest was made is explained in Fisher 1978a, which uses the unpublished 'History of the Rockefeller Foundation', by RM Pearce (Rockefeller Archive Centre MS XVI, 3996). This report makes clear the concern that the professorial units in London were judged likely to fail unless they received substantial extra funding. 'Lack of success in London would discourage trial of the plan in other parts of England and in other nations, and would also throw doubt on the success claimed for it in the United States. The situation in London therefore was critical, and to save the experiment it seemed essential that in one school at least the plan should be tried under the best possible circumstances as to staff, number of beds, laboratory equipment and financial backing'. (Hall 1976 p282) Biographical sketches of the first professors of medicine are contained in LJ Witts's reminiscences (Witts 1972).

The MRC did not establish further clinical research units until the early 1930s, when a change in funding arrangements allowed it to establish facilities at Guy's Hospital and the National Hospital for Nervous Diseases.[346]

MRC research on therapeutic drugs

Whilst the overall direction of early MRC policy owed much to the resources provided by NIMR and the person of Walter Fletcher, policy towards therapeutic drugs and the pharmaceutical manufacturers was shaped largely by Henry Dale, director of the NIMR Biochemistry and Pharmacology department. To describe the work of Dale and the MRC in relation to therapeutic drugs it is first important to outline the nature of the British pharmaceutical industry before the First World War, because it is so far removed in every aspect from the present pharmaceutical sector.

The British pharmaceutical industry during the first decade of the twentieth century consisted overwhelmingly of small to medium sized family-owned importers, finishers, and wholesalers of fine chemicals, media medica, and other preparations.[347] Moreover, the predominant pattern of business - described by Liebenau as the importation of raw materials, followed by rapid processing and then wholesale distribution or re-export of finished products - required little if any input from research and development facilities.[348] Accordingly therefore, and in contrast to the German, and to a lesser extent American,

[346] For a summary of further developments, see Landsborough Thomson 1973 p134-146 and 1975 p18-19.

[347] Slinn 1984 p100. Liebenau 1981. The main pre-war business of May & Baker was the preparation of bismuth, camphor, ether, and mercurials. (Slinn 1984 p80); Allen and Hanburys specialised in milk and malted foods, jujubes, pastilles and capsules, toilet soap, and the production of certain bulk galenicals such as cascara, sagrada, and liquorice (Tweedale 1990 p91). Glaxo was at the time principally a manufacturer of milk powder (Davenport Hines 1992 p41).

[348] Liebenau 1984. Despite the lack of innovation among British companies the strategy was very profitable. Glaxo increased turnover by 50% in the immediate pre-war years, and during the First World War demand from municipal authorities for dried milk caused turnover to rise from £50,000 in 1913 to £550,000 in 1918 (Davenport Hines 1992 p41-43). Allen & Hanburys reserves grew steadily and dividends on the firm's ordinary shares reached 50% in 1902. (Tweedale 1990 p115-116).

pharmaceutical industries, British manufacturers had little interest in new science-based products such as serum anti-toxins. They consequently lacked the capacity to undertake research on any scale, and had few links with universities.[349]

The MRC's first scheme of research, approved in December 1913, made no mention of research into therapeutic drugs.[350] Despite the lack of association between the earliest scientific objectives of the MRC and pharmaceutical manufacturers the position changed rapidly with the onset of the First World War.[351] The role played by the MRC, and particularly that of Henry Dale, in developing the UK pharmaceutical industry after 1914 is discussed below with reference to biological standardisation, and the MRC's Chemotherapy Research Committee, the forerunner of the TTC.

Henry Dale, Diphtheria anti-toxin, Salvarsan, and the development of biological standardisation

Henry Dale

Dr (later Sir) Henry Hallet Dale was appointed as Director of the NIMR Biochemistry and Pharmacology Department in June 1914. Before appointment to the MRC, Dale was Director of Henry Wellcome's research laboratories, the Wellcome Physiological Research Laboratory (WPRL), the leading commercial laboratory in Britain at the time,

[349] There are two exceptions to the general model. Firstly, the Wellcome Company, which funded an extensive physiological research laboratory, of which more in the section on HH Dale. But Henry Wellcome and his partner, Silas Burroughs, were Americans seeking to introduce an American style of pharmaceutical industry to Britain. The existence of the Wellcome research laboratory emphasises the difference between the American and British approach therefore. Secondly, the Evans Medical Company of Liverpool, which maintained close contact with the expanding Liverpool University Medical School, and is an exception to the predominant pattern of the British pharmaceutical sector. (Liebenau 1984).

[350] MRC First Scheme of Research. Reproduced in Landsborough Thomson 1973 p28-29 with minor omission.

[351] Note that Arthur Ewins, appointed to the NIMR Biochemistry and Pharmacology Department at the same time as Dale, became May & Baker's chief chemist in July 1917 (Slinn 1984).

and before that held the George Henry Lewes physiology studentship at University College London.[352]

At the WRPL, Dale had freedom to undertake research on any subject. His only work in support of the business at the laboratory was the standardisation of the extract of supra-renal glands.[353] Dale was made Director of WRPL in 1906. His first task was to clear up the contamination problem with serum processing at Brockwell Hall. His subsequent career at Wellcome was distinguished by further research success, and he was elected Fellow of the Royal Society in 1914, shortly before his appointment to the MRC.

Standardisation of diphtheria anti-toxin

In 1890 Emil Behring and Shibasaburo Kitsato observed that a blood serum extract from animals injected with sub-lethal doses of living or killed broth cultures of diphtheria bacillus had the ability to confer protection from the effects of lethal infection when

[352] His early life has been described by Tansey (Tansey 1990) Born 1875 into a strongly Methodist family, his father managed the London office of a Derbyshire stoneware firm. Dale won a scholarship to The Leys School Cambridge, and subsequently attended Trinity College Cambridge from 1894-1897, studying natural sciences, and specialising in physiology. He continued to attend the physiology laboratory after graduation, while also teaching biology at the Leys School. He left Cambridge in 1900, having won the Gedge Prize in Physiology (other winners include Walter Fletcher, AV Hill, EM Mellanby), for St Bartholomew's Medical School, where he qualified in 1902. In 1902 he won the GH Lewes Studentship, which allowed him to pursue research in physiology at University College London, under Ernest Starling. He worked at Paul Ehrlich's Laboratory in 1903. On return to London, and in need of a secure position to allow him to marry, he was offered a post at Henry Wellcome's Physiology Research Laboratory, upon the personal recommendation of Starling to Henry Wellcome.

[353] Dale's work on supra-renal glands led to publications that advanced his academic career but also brought him into conflict with commercial research. In 1905 he wanted to publish work on extract of supra-renal gland, using the word adrenaline. The uses of the word was objected to by staff from the Wellcome Chemical Research Lab, since it was similar to the term adrenalin, used by Parke Davis to market their extract of supra-renal gland. Wellcome ruled that Dale should use the term epinephrine. Dale objected on the basis that physiologists were not interested in the physical structure of a substance, but upon its action. He pointed out that the major paper published so far had been that of TR Elliott, entitled the action of adrenaline. (Elliott TR The action of adrenaline J Physiol 1904 31:20-21) To change the term would be to lose the intellectual connection between his and Elliot's work. Wellcome at first relented but then re-affirmed his original decision. In the ensuing conflict Dale asked for an outside expert, Langley from Cambridge, to adjudicate. Dale also received support from the Director of WRPL, Dowson, who pointing out that WRPL was meant to be entirely distinct from the business. Wellcome finally agreed to publication using the term adrenaline on March 8th 1906.

injected into other animals.[354] The first human treatment took place on Christmas Day 1891 in Berlin.[355] Despite early concerns about its utility, anti-toxic serum was produced on a commercial basis from 1892 by Hoechst. The first British cases were treated at the Eastern Hospital, in Hackney, in 1894.[356]

The principle concern of practitioners during the early stage of production was variability of the potency of batches of anti-toxin.[357] The response of Behring was to approach the well-known clinical scientist Paul Ehrlich, a colleague of his at the German Institute for Infectious Diseases.[358] Ehrlich's programme of standardisation of production had the added impact of improving the manufacturing process,[359] and created a system of control of the products. An attempt to measure and standardise the potency of diphtheria anti-toxin had been introduced by Behring, who calculated the amount of toxin needed to protect a guinea pig against 100 fatal doses (this amount being derived empirically from

[354] The concept of anti-toxin originates with Behring and Kitasato's experiments. Behring was awarded the Nobel Prize 1901.

[355] The patient recovered. For a summary of the introduction of diphtheria anti-toxin, see Weindling 1992. For technical details, see Parish 1965 especially p119-131.

[356] Eastern or Homerton Hospital, part of the Metropolitan Asylums Board (MAB). Opened in 1871 at Homerton Grove, Hackney. Success with the serum at Hackney led the MAB to establish facilities for the production of diphtheria anti-toxin in collaboration with the Royal Colleges of Physicians and Surgeons, using the laboratories on the Victoria Embankment and their stables in Balham. The MAB took over these facilities in 1904. In 1909 larger facilities were opened at Belmont in Surrey (Ayers 1971 p196-197).

[357] Baumler 1984 p55.

[358] Ehrlich is a key actor in the history of diphtheria anti-toxin and Salvarsan (see later in this chapter). He was a personal friend of Almroth Wright, and at one time employed Henry Dale for a brief spell. A good account of Ehrlich's life and work is found in Baumler's biography (Baumler 1984). The origin and organisation of the Institute of Infectious Diseases, better known as Koch's Institute, which was created to exploit the therapeutic potential of Tuberculin, an anti-toxin treatment for tuberculosis, is described in Weindling 1992. Further background material, on the growth of state support for science in Germany, is contained in Lenoir 1992.

[359] In order to improve the production of diphtheria anti-toxin at the WRPL, Dale introduced the standards of the academic laboratory to the commercial laboratory. Elsewhere, the requirements of anti-toxin production had the same effect. Before its re-organisation, the HK Mulford laboratory was situated above the stables used to house the serum horses. Diphtheria cultures became contaminated with tetanus spores from horse manure, and the horses injected with the toxin from these cultures acquired tetanus and died. (Gossel 1992).

experiments) of toxin.[360] As might be expected the results achieved by this method were variable. In 1897 Ehrlich introduced an indirect method, by which the potency of a test sample was compared to a preparation of anti-toxin of known Behring value. Samples of Ehrlich's original standard serum were kept in the dried state to maintain stability. This approach proved to be more successful, and for many years samples of standard serum were issued from Frankfurt all over the world.[361]

Following the demonstration that diphtheria antitoxin was effective,[362] the German government supported its development as a commercial product through standardisation and licensing.[363] As well as being a method of establishing the credentials of drugs, government licensing could be regarded as creating market advantage and was therefore

[360] Known as the Behring Unit. For details see Parish 1965 p125.

[361] Parish 1965 p127, also Baumler 1984 p62. For relations between Ehrlich and Behring concerning diphtheria, see Baumler 1984 p56-58.

[362] The first clinical trials took place in children's hospitals in Berlin, and subsequently at the Hospital for Sick Children in Paris, where Roux reported encouraging results with horse serum on a series of 300 children to the International Congress of Hygiene and Demography in Budapest (Parish 1965 p122). The results from the early trials have been described as disappointing until standard methods for preparation and treatment were developed. (Weindling 1992 p76. Baumler 1984 p55) Improved results were reported by Kossel, who obtained a 97% recovery rate in 78 children, where previous fatality rates had been as high as 50%, even in the best clinics. Note that one of the earliest controlled clinical trials is attributed to diphtheria anti-toxin. In 1898 Fibiger in Denmark was able to randomly allocate patients to anti-toxin or standard treatment, using date of admission as a randomiser. The fatality rate in the treated group was 3.3%, compared to 12.2% in the non toxin treated group. (Fibiger 1898) However insufficient the results from the non-controlled trials in Berlin and Paris, the researchers felt they had an efficacious remedy by 1894. In October the German Ministry of Culture held two conferences on diphtheria (Baumler 1984 p59), and a philanthropic committee for the distribution of the serum to poor children was established. Ehrlich and Wassermann began to promote the use of the serum through public lectures and a training programme for doctors. Official recognition had to wait until November 1895, when the Reich Health Office pronounced that diphtheria anti-toxin was safe and efficacious. (Weindling 1992 p79).

[363] The earliest practical outcome of state support for commercial pharmaceutical manufacture in Germany was the establishment of the Serum Institute (Institut für Serum Prufung und Forschung), in Berlin in 1896 with Ehrlich as Director. For details of state support, see Lenoir 1988 p79. Its purpose was the standardisation and testing of diphtheria antitoxin produced by Hoechst and Schering. Following the Lancet report on the relative potency of diphtheria antitoxins, Ehrlich concentrated on devising methods of standardisation. Ehrlich's move to new laboratories in Frankfurt in 1899 was prompted, according to Liebenau, by the proximity of Hoechst. Throughout this period Ehrlich worked closely with Hoechst, see for example the terms of agreement between Ehrlich and Hoechst concerning the production of diphtheria anti-toxin, reproduced in Baumler 1984 p58. The Berlin and Frankfurt organisations were in effect government sponsored laboratories for the endorsement of Hoechst products. By 1903, Hoechst estimated that it had produced almost 20,000 litres of diphtheria antitoxin, and generated income of over 4M Marks. (Lenoir 1988, Liebenau 1990).

readily accepted by German firms.[364] In America, several companies used scientific terminology in their advertising from the 1880s,[365] but the first state regulations were introduced only after deaths caused by contaminated batches of diphtheria antitoxin in 1902, and the contamination of smallpox vaccine with tetanus.[366] In contrast, the British market was largely un-regulated until 1925, when the Therapeutic Substances Act introduced a measure of control into the manufacture and use of biologically active chemicals.[367]

According to Burn, Dale's work on biological standardisation was amongst his most significant contributions to medical science.[368] The origin of Dale's work in this area can be traced quite precisely to his appointment to the WRPL in 1904.[369] Burroughs Wellcome, the company created by the American pharmacologists Henry Wellcome and Silas Burroughs in 1880, was the among the first British pharmaceutical manufacturers to exploit new advances and techniques, and to incorporate scientific terminology in its advertising material.[370] As early as 1891, an experimental laboratory was established at

[364] In contrast, British drugs could be excluded from foreign markets because they lacked any means of obtaining central authentication through license or standardisation. (Pfeffer N 1985) quoted in Austoker and Bryder 1989 p53.

[365] Liebenau 1985 highlights the role of scientific language by Lederle, Smith Kline and Wyeth. The case of Lederle is particularly interesting because it shows the links between state agencies and commercial organisations; in this case Ernest Joseph Lederle, who founded the Lederle Laboratories, was previously chief chemist at the New York City Health Department. Lederle ran the city's diphtheria. Antitoxin laboratory, the leading producer of anti-toxin until 1902.

[366] Liebenau 1984. For a general discussion see Marks 1987 chapter 3.

[367] Before 1925 the principal regulations related to the sale of poisons and the Pharmacy Acts of 1852 and 1868. For an overview of the development of pharmacy in Britain, see Kremers and Urdang 1963.

[368] Burn 1955.

[369] See Tansey 1990 for an account of Dale's research before he joined the WRPL.

[370] Burroughs Wellcome & Company was established in London in 1880 as a partnership between Silas Burroughs and Henry Wellcome. Burroughs, the senior partner, died of pneumonia in 1895. Examples of Burroughs Wellcome advertising are reproduced in the main history, that of Macdonald 1980. Burroughs Wellcome's earliest success was the preparation of drugs in compressed form. The term 'tabloid' was coined by Wellcome, and was protected by trademark for some time in the early twentieth century.

the company's commercial headquarters in Snowhill, though its use remain unclear. In 1894 facilities, known as the Wellcome Physiological Research Laboratories, were established in Charlotte Street for the manufacture of diphtheria anti-toxin. The facility was a success. Exports of diphtheria anti-toxin to America began in late 1894.[371] However, in 1896, the Lancet published a critical report on the quality of diphtheria anti-toxin produced by Burroughs Wellcome.[372] Since the company's advertising material emphasised the physiological tests made on its products, the Lancet report must have been potentially damaging.[373] In all events it led to the appointment of a new Director, and the removal to new premises in 1898, at Brockham Park, South London. Dale was responsible for all the work on standardisation at WRPL until his appointment to the MRC in 1914.

Salvarsan

Reference has already been made to the mass production of anti-typhoid vaccine on behalf of the British Government by Almroth Wright's laboratory during the First World War. A further effect of war was to highlight British dependence on foreign-produced drugs. Salvarsan, as well as being the first effective chemotherapeutic agent[374] has been described also as a turning point in the relationship between the pharmaceutical industry and the British state.[375]

---

[371] Parish 1965 p124.

[372] Martin 1896.

[373] See Tansey 1990 p159 footnote 78.

[374] There is a large literature on chemotherapy. Despite its age Galdston 1943 provides a reasonable introduction to the concept and history of chemotherapy. Parascandola describes Ehrlich's conception of chemotherapy in detail. (Parascandola 1981).

[375] Liebenau 1989 p164.

Salvarsan is the commercial name of a complex organic molecule, derived from arsanilic acid by Paul Ehrlich in 1909 as a treatment for syphilis.[376] Salvarsan is known also as '606' because it was the 606th preparation by Ehrlich of chemicals derived from arsanilic acid in his search for an agent that would kill organisms in the body.[377] Based on the principle of inner disinfection or chemotherapy, Ehrlich's concept of a chemotherapeutic agent, which he modified only towards the end of his life, was one which acted directly, physically affixing itself to the organisms it was to kill.[378] His ideal chemotherapeutic agent would destroy the parasite in one dose, whilst leaving the host organism unaffected.

Tests of the therapeutic potency of 606,[379] carried out first on birds infected with spirochetes, and then rats and mice, indicated that 606 was effective in treating relapsing fever. Further tests on rabbits indicated that Salvarsan effectively destroyed the spirochete organism *Treponema pallidum*, known since 1905 to be the cause of syphilis.[380]

---

[376] By this time, Ehrlich was based at a new research institute created for him through an endowment by Frau Franziska Speyer. The Georg Speyer Haus was the centre of Ehrlich's later researches and commercial relations with Hoechst. For a review of Ehrlich's work on Salvarsan, see part 3 of Baumler 1984. For a review of chemotherapy of venereal diseases see Dowling 1977 chapter 7.

[377] A photograph of Ehrlich's laboratory notebook showing the structure of compound 606 is included at p144 of Baumler 1984.

[378] Ehrlich divided the field of experimental therapeutics into three categories: organotherapy (use of organ extracts such as adrenaline); bacteriotherapy (use of anti-toxins and vaccines); and chemotherapy – which he described in an address at the opening of the Georg Speyer Haus in 1906 as the science of 'curing organisms infected with certain parasites in such a way that the parasites are exterminated within the living organism, so that the organism is disinfected'. Further discussion of Ehrlich's theoretical views on the basis of chemotherapy, and his work on the Salvarsan group are contained in Parascandola 1981.

[379] The tests were carried out at the Georg Speyer Haus by the Japanese bacteriologist Sahashiro Hata, beginning on June 2nd 1909 (Baumler 1984 p147).

[380] The organism was named as *Spirochaeta pallida* at the time. Initial work at the Institut Pasteur by Eli Metchnikoff suggested the existence of a transmissible agent in syphilis. The German scientists Fritz Schaudinn and Erich Hoffman isolated the responsible organism in 1905.

Clinical trials on human subjects had a demonstrative rather than confirmative role. Ehrlich's philosophy of therapeutic action,[381] his knowledge of therapeutic chemistry, and the evidence obtained from animal studies were enough to convince him that Salvarsan was a revolutionary treatment. Ehrlich set out his approach to the testing of drugs in a letter to Professor Ludwig Darmstaedter, of January 4 1905:

> *A therapeutic agent for a particular disease can be discovered only in an organism suffering from that disease. For many reasons, however (humanitarian considerations and the technique of scientific experimentation), the sick patient is a highly unsuitable subject to use in the discovery of medicines. Attention is not focussed on the patient until the drug has been recognised as effective in an extensive series of experiments in animals*[382]

The first test on human subjects was carried out at the lunatic asylum in Uchtspringe using a small number of late cases of syphilis, with dramatic effect.[383] The results of further tests were announced at the Congress on Internal Medicine, at Wiesbaden on April 18 1910.[384] In September 1910, at the congress of German Natural Scientists and Physicians, Ehrlich reviewed 10,000 cases treated with 606, which showed the drug to be effective.[385]

---

[381] In this respect, Ehrlich's side chain theory, a physiochemical explanation of immunity, suggested the existence of receptor molecules on cells, which facilitated the internal economy of cellular activity. Ehrlich realised that these receptors could be chemically blocked, thereby disrupting the cell. Ehrlich's willingness to test hundreds of chemicals testifies to his belief that specific chemical structure was the key to therapeutic activity. For a discussion of the role of side chain theory in the development of Ehrlich's work on chemotherapy, see Parascandola 1981. For the specific receptor for 606 see Parascandola 1977, which includes a discussion of Dale's criticism of Ehrlich's theory.

[382] Letter to Professor Ludwig Darmstaedter, January 4 1905, reproduced in Baumler 1984 p116.

[383] The trial took place in September 1909 (Baumler 1984 p153-154). The report of the trial is contained in a paper by Alt in the Munchener Medizinische Wochenschrift of March 15 1910 (Baumler 1984 p265). 27 cases of florid syphilis were tested "and perfectly amazing therapeutic results have been recorded".

[384] Baumler 1984 p162.

[385] Why then did Ehrlich continue testing during the autumn of 1910? The purpose of these tests, carried out by a network of carefully selected physicians, appears to have been two-fold. Firstly, as he later explained, extensive testing in humans was necessary to establish the nature and extent of harmful effects. Secondly, Hoechst had

Clinical trials in England confirmed the efficacy of Salvarsan in comparison to that of mercury based treatments.[386] One of the earliest reported trials was that of Blayney, in a paper to the Royal Academy of Medicine of Ireland in 1911.[387] Other early trials were also strongly positive,[388] and unlike later research found little evidence of side effects.[389]

Salvarsan, and its improved variant Neo-Salvarsan (known as '614') were produced exclusively in Germany under patent. With British stocks at the start of the First World War limited to what was already in the country, the Board of Trade used emergency powers to suspend the patents and trademarks of Salvarsan and Neo-Salvarsan and grant licenses to Burroughs Wellcome of England and Poulenc Freres of France (with May & Baker to be responsible for distribution in Britain) for the production of the drugs, which were marketed under the names Kharsivan and Neo-Kharsivan, and Arsenobenzol-billon and Nov-arsenobenzol-billon.[390]

---

difficulty producing Salvarsan in sufficient quantity to meet anticipated demand and of sufficient quality for release to the profession. Testing was a way of restricting distribution until greater quantities of drug could be manufactured.

[386] The reception of Salvarsan in England is discussed in Ross 1997. Unfortunately this paper contains a number of factual errors and no discussion of the effect of the First World War on supplies.

[387] BMJ 1911; ii:18.

[388] Browning and McKenzie tested 300 cases at the Western Asylum Research Laboratories in 1911; Gibbard and Harrison at the Military Hospital in Rochester Row were given supplies of Salvarsan by Ehrlich's laboratory, and found very positive results in 129 cases. (BMJ 1911;ii:654-55). At the International Medical Congress of 1913, Harrison estimated that in comparison to mercury, Salvarsan saved the British army 70-80,000 hospital in-patient days per year.

[389] The first significant report of side effects in the BMJ was in 1917, when an analysis revealed that over 30% of Salvarsan cases suffered moderate side effects such as headache and nausea, and about 10% suffered severe side effects (Lloyd Jones BMJ 1917; 1:152-4). It is possible that the number of side effects from arsphenamines produced under wartime license were greater than those produced by earlier German batches. The first report of the Salvarsan Committee reported that during the war the stringency of batch toxicity tests on rabbits was reduced, by lowering the experimental dose and shortening the period of observation (Medical Research Council 1919 p17).

[390] The announcement was made in the BMJ of April 10 1915 (p649). May & Baker's involvement with Salvarsan, and the eventual sale of its shares to Poulenc Freres is described in Slinn 1984 p89-99. It is interesting to note that one of the effects of the Salvarsan programme was the move from NIMR to May & Baker of Dr JA Ewins, who became chief chemist at the Wandsworth site of May & Baker in 1917, moving from Dale's department at NIMR. The research department grew steadily during the 1920s, and profits increased from £10,000 in 1922 to £44,000 in 1930 (Slinn 1984 p95). One source of profits were derivatives of arsanilic acid, including tryparsaminde, licensed to May & Baker in 1925 by the Rockefeller Institute of Medical Research.

The Board of Trade was empowered to attach conditions to the licenses granted under the temporary rules of 1914. In the case of Salvarsan and Neo-salvarsan, the conditions stipulated that that a sample of every batch produced should be submitted to the MRC's Department of Biochemistry and Pharmacology for biological tests. Although this condition was part of emergency wartime measures, it was in fact almost identical to one imposed by Ehrlich on Hoechst, who manufactured Salvarsan and submitted to his laboratory for testing.[391] Chemical analysis alone was insufficient, Ehrlich recognised, because the toxicity of Salvarsan could not be accounted for on the basis of the amount of residual hydroxyaminophenylarsinous oxide, a bi-product of the process of manufacture.[392] He therefore submitted each batch to biological testing for toxicity and efficacy, using a hypodermic injection of Salvarsan in alkaline solution in mice. Ehrlich's tests were repeated at the NIMR, initially using mice but in time using intravenous injection in rabbits.[393]

The whole issue was placed in the hands of a special Salvarsan Committee, established to consider the methods of manufacture, of biological testing, and of clinical administration of Salvarsan and its substitutes. Such was the promise of this work that in December

[391] Ehrlich conducted tests at the Georg Speyer Haus. Between September and December 1910, 375,395 ampoules of Salvarsan were tested, of which the vast majority were passed.

[392] Salvarsan Committee 1919 p7. BMJ 1915;i:649.

[393] Medical Research Council 1919 p15-16. It was found that hypodermic injections tended to result in variable amounts of the drug precipitating out of solution at the point of injection, leading to variable toxicity results. By a matter of their being more readily available, rabbits were used in Britain rather than mice. Variable results, thought to be due to both the rabbits themselves and the conditions in which the rabbits were kept, were soon noticed. Samples that had previously passed the toxicity test sometimes failed upon re-testing. With too many batches failing, the standard of the toxicity test was reduced: 'From time to time, therefore, it became necessary to relax the full rigidity of the standard originally demanded. During one period this was effected by lowering the test dose [by 33%]…at a later period, when shortage of supply again necessitated relaxation of the original standard, it was considered wiser to maintain the dose at 0.12 gm/kg but to shorten the period of probation, so that if the rabbit survived for three days after the injection without untoward symptoms, later death was not regarded as an adequate ground for rejection'.

1916 the MRC was urging the Government's Reconstruction Committee to consider the need for regulation on a wider basis than Salvarsan.[394] The memorandum noted that products such as anti-diphtheria and anti-tetanus sera were dependent on foreign methods of standardisation, and expressed the view that the existing absence of possibilities of control was "discreditable to our national position in the world of science and a source of grave danger to the community". It urged the establishment of a Government laboratory for biological standardisation of an analogous kind to the National Physical Laboratory.[395]

The exigencies of war had established a mentality in which the issue was not the clinical effectiveness of drugs, but cross-national comparisons of potency and manufacturing ability.[396] The outcome for the MRC was that biological standardisation became a special responsibility of the NIMR. In 1920 a Committee for Biological Standards and the Methods of Biological Assay was appointed, working under Henry Dale, then director of the Department of Biochemistry and Pharmacology. He represented the MRC at the international conference on diphtheria anti-toxin in 1921, and at conferences in 1925 and 1926 which established guidelines for the creation and maintenance of a wide range of

[394] Landsborough Thomson 1975 p244.

[395] The history of the National Physical Laboratory (NPL) reveal that it too originated in response to concern about German domination, in this instance of physical standards such as the ohm unit. The arguments were about the threat to British industry and the need for state finance of a physical laboratory. Following agreement by the Treasury, the establishment of the NPL was placed in the hands of a committee of the Royal Society, which included Francis Galton among its members. Responsibility for NPL passed to the Department of Scientific and Industrial Research upon its establishment in 1918. Work began on constructing the laboratory, at Bushey Park, in 1900. As Moseley 1978 makes clear, the establishment of the NPL was regarded as the first significant instance of Government support for science, and was achieved in the face of concerted opposition from The Times newspaper, and continued scepticism from the Treasury.

[396] Shortly after the MRC announced that Salvarsan was being manufactured under license in England the BMJ noted with satisfaction that those who had doubted the ability of British manufacturers to make Salvarsan had been proved wrong, on the basis of stringent scientific tests. (BMJ 1915 p689-690).

biological standards, including one for Salvarsan-type drugs.[397] Under Dale's leadership, the NIMR became a centre of international standardisation, under the auspices of the League of Nations.[398] In Britain, the position of the MRC at the centre of regulatory activities regarding therapeutics was consolidated by two events in 1925. Firstly, the passage of the Therapeutic Substances Act[399] and secondly, moves to reform the British Pharmacopoeia which saw responsibility taken from the General Medical Council and placed with a Pharmacopoeia Commission, whose members were selected by the MRC.

The terms of the Therapeutic Substances Act illustrate both the meaning of standardisation and the limits of state involvement in therapeutics in the 1920s. The Act made provision for 'the regulation of the manufacture, sale, and importation of vaccines, sera, and other therapeutic substances… the purity and potency of which cannot be adequately tested by chemical means.'[400]

As the Act suggests, the need for biological standardisation arose because the therapeutic and toxic effects of biologically active substances could not be adequately assessed by chemical analysis.[401] Why though did the Act not include provision for testing the effectiveness of treatments? At least two reasons can be suggested. The less important is

[397] Standardisation also covered the growing field of sex hormones. Dale's work in relation to the standardisation of sex hormones is discussed in Oudshoorn 1990.

[398] The other centre was the State Serum Institute in Copenhagen (LandsboroughThomson 1975 p247).

[399] Therapeutic Substances Act 1925 (15 & 16 Geo. 5. Ch. 60). Dale represented the MRC on the committees framing the Act.

[400] The Act covered immunological preparations; arsenical drugs of the Salvarsan type; insulin; and posterior pituitary hormone.

[401] The need for biological testing in the absence of known structure is highlighted by insulin, with which the MRC became involved in 1922 when it was offered the British patent rights by the University of Toronto. Until the mid-1950s, when Fredric Sanger finally deciphered the structure of insulin, no chemical tests for the activity of insulin were possible. There is some dispute about the earliest standard for insulin. Bliss suggests it was the amount necessary to reduce the blood sugar of a rabbit by 50% in 1-3 hours (Bliss 1988 p122), while Dale's unpublished notes suggest a cruder measure, the amount 'required … to throw 3 out of 5 rabbits into hypoglycaemic convulsions. (Murnaghan and Talalay 1992 p445).

that, as we have seen, the laboratories which produced the substances covered by the Act had little access to patients. More importantly, the new therapies were, at least nominally, based on physiological principles. The 'proof' of newer therapies did not therefore rely on empirical clinical testing, but on the demonstration that they behaved according to scientific laws and principles, which could be established under a microscope, in a test-tube, or in a laboratory animal.[402] Clinical trials on patients were therefore of secondary importance, as well as being difficult to organise. In the case of Salvarsan-type drugs, the trypanosome method of testing adopted in the 1920s involved comparison with a standard preparation of German Salvarsan or Neo-salvarsan:

> *Mice are infected, intraperitoneally, each with 0.5 cc of a suspension of* Treponema equiperdum *containing 7000 trypanosomes per cc. Two days later the blood of each mouse is examined and the number of trypanosomes per cc is estimated. They are then arranged in groups containing mice of similar degrees of infection. These are given doses of a test sample or standard Neoarsphenamine, intravenously. On the following day and on each subsequent day for a week, the blood of each mouse is examined for trypanosomes. A comparison is then made of the number of mice in which the blood is cleared of trypanosomes by means of the test sample and of standard Neoarsphenamine respectively.*[403]

In any case, the results of clinical trials were often difficult to interpret, costly, and late. In 1919, under the auspices of the Salvarsan Committee, Fildes[404] and Parnall published a report on the treatment of syphilis at the Royal Naval Hospital, at Haslar in

[402] In the philosophy of Claude Bernard, the laboratory formed the real world, in which the control of all extraneous factors allowed one to see the essential causal link, while clinical practice was merely the confused imitation of the laboratory. Bernard's rejection of statistics was not a rejection of quantification as such, but a rejection of a technique which tried to make up for the inherent limitations of clinical practice as a source of knowledge. The rejection of statistical knowledge was despite the fact that Radicke had introduced a statistical test of a self-controlled trial which allowed the rigorous testing of medical interventions using physiological outcome measures. Radicke's test was short lived and quickly forgotten, perhaps indicating the strength of institutional arrangements over cognitive possibilities (Coleman 1987).

[403] FD1 2508 Report on the therapeutic tests applied to Novostab N77 21/1/39.

[404] Paul Gordon Fildes 1882-1971. His supervisor at Cambridge had been Walter Fletcher. (Gladstone Bert and Knight 1973).

Portsmouth.[405] In 1929 LW Harrison,[406] Director of the Venereal Diseases Department there, prepared a second report, on all male cases treated at St Thomas's Hospital between 1920 and 1926.[407]

The studies did not attempt to compare the efficacy of Salvarsan with that of mercurial compounds, since it was widely accepted that Salvarsan was significantly superior to mercury. Fildes and Parnall showed that Neo-Salvarsan (a variant of Salvarsan which claimed less toxicity) was effective only in very early cases of syphilis. Harrison's analysis was an attempt to clarify the best course of treatment for different stages and presentations of syphilis, based on what he called the ultimate results of treatment. Harrison's approach was to draw up tables of all the possible combinations of symptoms and treatment regime, allocate cases to these tables, and record the outcome of treatment. It attempted too much, and therefore achieved nothing. For example, the table on page 11 of his 1929 report attempts to compare the effect of 5 brands of arsenobenzene compounds, in three patient groups, in the presence of two different adjuvant regimes, using an outcome measure split into 4 categories.[408] For good measure, he added summaries of the treatment of 598 patients as an appendix.

---

[405] Fildes and Parnall 1919. By an accident of history, James Lind, remembered for his prospective clinical trial of lemon juice as a treatment for scurvy, served as physician in chief at the Haslar Hospital, between 1758 to 1783. At the Haslar Hospital, Lind continued the controlled clinical experiments he had begun as a ship's surgeon. (Trohler 1981) Naval and Army medicine have formed an early site for evidence-based medicine. HJ Cook's work on armed forces medicine after 1688 is an important but neglected source of explanation (Cook 1990).

[406] Lawrence Whittaker Harrison, 1876-1964. First Ministry of Health Adviser on Venereal Disease. Founding editor of The British Journal of Venereal Diseases in 1925. His life and work are described in King 1974.

[407] Harrison 1929. VD treatment centres arose as a result of the 1913 Royal Commission on Venereal Disease. For a general discussion, see Evans 1992. The VD clinic at St Thomas's Hospital is described on p395-397 of King 1974.

[408] Harrison 1929 p11.

Lessons of Fildes' and Harrison's studies

While Fildes' study produced a clear result, it was no more than confirmation of the generally accepted position that the best response to Salvarsan was found in early cases of syphilis.[409] In Harrison's case, analysis of 6 years data had taken 3 years to complete. As he admitted, it had been 'much more difficult and time consuming than had been expected'.[410] Nor, in the view of the MRC, did it achieve its aim: 'it does not claim success for any particular treatment in any given stage of the disease'[411].

In place of the long drawn out, and rather insubstantial results of clinical studies, laboratory testing offered a means which was more precise. It also put the resources over which the Council had a high degree of control at the centre of efforts to increase knowledge and regulate the pharmaceutical industry. Pages 25-27 of the First Report of the Salvarsan Committee develop the argument for laboratory testing along the following lines. Biologically active chemicals such as Salvarsan require legislation to control the conditions under which they are supplied. Such legislation will be based around the compliance of batches with chemical and biological tests, and in general to impose

[409] Fildes involvement in the testing of Salvarsan and his methodology owed much to his position at the periphery of British clinical practice. On qualification in 1909 he took the most unusual step of working for a clinical scientist, William Bulloch, rather than taking up a clinical post. Bulloch himself had been ostracised by the clinicians at the London Hospital, who were in revolt against the appointment of clinical scientists. Fildes was able to gain early experience with Salvarsan through the bacteriologist James Macintosh, a personal friend of Ehrlich. Most of Fildes' publications between 1909 and 1915 concerned syphilis. His interest in clinical trials stemmed from the 'fog [that] began to descend as it does upon every new discovery in medicine. The drug became a staple article of medical practice; every one used it but no one studied it'. (Fildes and Parnall 1918) During the First World War Fildes took a post as chief of the laboratory at the Haslar Naval Hospital. From the laboratory he was able to promote standardised treatment regimes and monitoring, justifying these on the basis of organising and improving the laboratory service. Fildes was careful not to dictate the exact course of treatment but to improve and standardise data collection through the introduction of a card system, (illustrated on p 266 of Fildes and Parnall 1918) and a method of classifying treatment courses.

[410] Harrison 1929 preface first unnumbered page.

[411] Harrison 1929 Introduction.

conditions under which the substance may be sold. There are two possibilities for the arrangement of laboratory tests:

> *The system which has been in use during the war for the salvarsan group, by which each batch of product is tested by independent workers officially appointed, might be made permanent under the Government department appointed to administer the control. The Committee, after careful consideration of this possibility, are of the opinion that, in the long run, it would prove to be unworkable. As a war emergency, and while the products of a few manufacturers only were concerned, it has served its purpose. The extent to which the sale of remedies needing such control will expand in future cannot be foreseen. Unless it is proposed to restrict the sale to what is manufactured in this country, the number of the manufacturers who will desire to market their products here is quite unknown…the committee are therefore of the opinion that the alternative method, by which the testing is decentralized, will be the only practicable one to adopt.*[412]

The view of the MRC was that the role of Government in relation to drugs was to set standards and to act with other governments to enforce those standards. This line of action was developed through Dale's work on biological standardisation, and framed in the provisions of the 1925 Therapeutic Substances Act. In this way the MRC both involved and distanced itself from the producers of therapeutic substances, attempting to maintain authority in the area of therapeutic substances while not being responsible for the products themselves.

Nevertheless, given the example set in Germany, it was inevitable that the MRC would also seek ways to encourage the British pharmaceutical industry. The next sections describe the principal ways in which it set about doing so. It will be seen that the form of involvement was indirect at first, but became more direct during the late 1920s.

[412] Medical Research Council 1919 p25-26.

The Chemotherapy Committee

Work on standards shows the MRC acting at a distance to regulate the pharmaceutical industry. Less publicly, the MRC also took an active role in developing the industry, helping it to discover new therapeutic substances, and to improve methods of production.[413] The earliest recorded step taken by the MRC was the appointment of a chemical assistant to E J MacWeeny of Dublin, for the preparation of a series of synthetic dyes to be tested on the tubercle bacillus. The MRC also funded a study of the bactericidal activity of several compounds that involved co-operation between a bacteriologist and a chemist.[414]

A systematic attempt to promote the testing of promising chemicals and to co-ordinate the work of chemists and biologists began in 1926 with the creation of a joint exploratory committee of the MRC and Department for Scientific and Industrial Research,[415] with a view to increasing the productivity British pharmacology.

---

[413] Following the assignment of UK patent rights to the MRC, Dale and HW Dudley were able to increase the average yield of insulin from the pancreas by 800%. They also reduced the amount of alcohol needed in the process of extraction by 80% and introduced several other improvements. Fletcher to Charles Sherrington: 'if we had let the manufacturers begin two months ago, they would probably have made little progress and would now be scrapping their plant for changed methods. As it is we have five firms started very shortly, two are already beginning' Fletcher to Sherrington, January 1923. Quoted in Liebenau 1989 p171 footnote 22). Although Liebenau's claim that the insulin patent was crucial in determining the relationship between the MRC and he pharmaceutical industry, insulin was not especially important in framing MRC policy towards therapeutic substances. I have argued here that it was Dale's experience before his involvement in insulin - his work with diphtheria anti-toxin and the adrenaline controversy at WRPL and the MRC licensing of Salvarsan – that were of greatest importance, even if these substances lacked the lasting clinical importance of insulin. Dale's account of the standardisation of insulin (Murnaghan and Talady 1992) shows clearly that his ideas on standardisation presented to the International Congress of Physiology in Edinburgh in 1923 were based on those of Paul Ehrlich in relation to diphtheria antitoxin, which Dale had experienced first hand whilst visiting Ehrlich's laboratory in 1903. The MRC's involvement with insulin is not included in this thesis therefore. It is possible, as JP Swann suggests, that insulin 'marked the beginning of a widespread movement of collaborative medical research [between universities and industry] in America' (Swann 1986 p 3) In Britain however, 'Insulin AB' produced in partnership between Allen & Hanbury's and British Drug Houses, required little input from the MRC after initial work at NIMR (Tweedale 1990 p128-130).

[414] Landsborough Thomson 1975 p45.

[415] A body established during the First World War. The DSIR was part of the Government's response to the industrial crisis of the War, in which the deficiencies of British industry were exposed. During the 1920s the DSIR established co-operative research associations with several industries. The DSIR was disbanded in 1964. Events leading to the formation of the DSIR in 1916 and its subsequent history are described in Varcoe 1970 and Varcoe 1972.

In forming a Chemotherapy Committee (CC) in 1927 the MRC hoped to 'encourage co-operation between chemists, biologists and pathologists, and clinicians in the production of new compounds, in their experimental trials, and in the observations of their effects on human disease.'[416] It was anticipated that the CC would not organise trials itself but that a further specific body would be needed to organise the practical arrangements for clinical trials.

In practice, the majority of the CC's work consisted of arranging medium sized grants to individual chemists and pharmacologists in universities to allow them to employ research assistants. In terms of producing new substances for clinical trial the CC appears to have had very little impact during the 1920s at least. The minutes of the 5th meeting, of November 13th 1928 records:

> *9a 'Dr Dale stated that the biological workers were somewhat restless because the chemists were not providing them with nearly as many compounds to test as they had hoped for. Dr Keilin's department had had very little time occupied; and had been turning its energies in other directions. Dr Scott Macfie had tested a few compounds received from Prof. Robinson and elsewhere, and had then spent his activities in testing substances received from outside sources, not through the committee [including] May & Baker and there was some doubt as to whether these results would be available to the committee at all.*[417]

[416] Austoker & Bryder 1989b p45. The committee first met on November 7th 1927 (FD1 7205 Chemotherapy Committee Minute Book). Henry Dale chaired the Committee. Members were drawn from the chemical side and biological side. They included the chemists Harold King, then working at the NIMR on new chemicals for the treatment of trypanosomiasis (Landsborough Thomason 1975 p45) and G Barger, formerly head of the chemistry division at NIMR and then Professor of Chemistry at the University of Edinburgh. Members on the biological side included RT Leiper of the London School of Hygiene and Tropical Medicine (the Council's principal adviser on helminth infections), and Leonard Colebrook, a research clinician based at Queen Charlotte's maternity hospital, who later carried out work for the TTC, notably on prontosil as a treatment for puerperal fever. The biological secretary was the bacteriologist CH (later Sir Charles) Andrewes, best known for his work on a viral vaccine for dog distemper.

[417] FD1 7205 Chemotherapy Committee minutes.

A note from 1931 by Henry Dale for the Therapeutic Trials Committee suggests that no chemicals were put forward for clinical trial by DSIR.[418] While this recollection is not strictly accurate, it reflects the view that the intention of advancing British chemotherapy without working directly with the British pharmaceutical industry had not been fulfilled by the arrangements set in place during the 1920s.

Although the original remit of the CC had not formally included work with pharmaceutical manufacturers, the position appears to have changed dramatically after the 6th meeting of the CC, on December 18th 1928, when Dale reported that following an interview with Dr Ewins, who had worked with Dale at both WPRL and NIMR, and was now chief chemist at May & Baker, 'May & Bakers were now ready to give the committee, in confidence, a complete account of their research programme on chemotherapy'.[419] Given the tardiness of the 'house chemists' at NIMR, and the willingness of a leading pharmaceutical company to engage with the CC, the time appeared ripe for the closer relationship of the MRC and the British pharmaceutical industry. Events which led to the creation of the Therapeutic Trials Committee began towards the end of the 1920s when MRC entered into discussion with the Association of British Chemical Manufacturers (ABCM) to resolve 'the problem of securing trustworthy clinical trials of products produced by manufacturers'.[420] The Annual Report for 1930-31 sets out the case for closer co-operation between the MRC and industry:

[418] FD1 2498 summary of CC.

[419] FD1 7205 Chemotherapy Committee Minute Book. Minutes of 6th meeting, Item (3) December 18th 1928.

[420] FD1 2498.

> *In these intermediate stages of work* [between the discovery of a drug and its introduction to practice] *it is often necessary for rapid and economic progress that close cooperation should be established between the original scientific investigators and those whose work lies in finding the adaptations needed for large scale production and use. The general social and political reasons that make it obviously desirable for State-supported work like that of the Council to be brought into the field of co-operative effort with British manufacturing firms are reinforced by the intimate relations of this kind of manufacturing to health and life within the country.*[421]

On 16th February 1931 a group of senior figures from the MRC met with representatives of seven leading drug manufacturers (Boots Pure Drug Company; British Drug Houses; Graesser Monsanto; Allen and Hanbury's; Burroughs Wellcome; Evans Sons; and May and Baker) and the ABCM. The account of the meeting[422] records that Mr Pratt of the ABCM[423] felt that a committee was urgently needed. Dr Carr[424] believed that doctors in Britain were afraid of publishing clinical trials in case they should be suspected of pecuniary interest. Francis Fraser, Professor of Medicine at St. Bartholomew's Hospital, disagreed with this view. Thomas Elliott, Professor of Medicine at University College Hospital said that in any case such criticism could not be levelled at official trials under the auspices of a committee of the MRC. And so it was agreed that a committee should be formed to receive applications from manufacturers to have therapeutic substances subjected to clinical trial. On the 6th of March a memo was circulated within MRC, and on 13th March 1931 the Council ratified the formation of a committee. A number of titles for the committee were suggested, including the Clinical Trials Committee; the Clinical

[421] MRC Report for 1930-31.

[422] FD1 2498.

[423] J Donaldson Pratt OBE MA BSc FIC, Managing Director of the ABCM, 166 Piccadilly, London.

[424] Chief chemist at British Drug Houses.

Committee for New Remedies; the Therapeutic Committee, and the Therapeutic Trials Committee, the last of which was selected by Henry Dale.[425]

The Therapeutic Trials Committee

From their perspective, the pharmaceutical manufacturers needed to obtain testimonials from doctors for their products. From its perspective, the MRC needed to be seen supporting the manufacturers, but it also needed to maintain its position at the head of clinical research in Britain. Throughout the 1930s the Therapeutic Trials Committee (hereafter, TTC) played a central role in fulfilling these purposes.

The TTC met at the headquarters of the MRC in London 10 times between July 1931 and March 1939. It considered 67 applications for clinical trial of which 51 were supported.[426] During its existence two sub-committees were formed, concerned with the testing of sex hormones and anti-syphilitic remedies.[427] Whilst not formally disbanded, meetings of the TTC lapsed during the Second World War, and after the war it did not meet again as a committee. The committee's secretary FHK Green continued the work of the TTC without convening meetings, seeking advice from individual committee members as he thought necessary. Replying to an enquiry about the functioning of the TTC from the research superintendent of Monsanto in April 1947, Green wrote 'the Council's TTC has not been reconstituted as a committee since the end of the war, but the mechanism is still in action'.[428] Monsanto then submitted a benzene derivative for trial as

[425] FD1 2498.

[426] 67 applications were formally considered by the TTC. Green handled many more informal applications. Where he thought they were clearly unsuitable for the Committee he rejected them by return. See Appendix 6 for a summary of the applications considered by the TTC.

[427] FD1 5319 Minutes of TTC Meeting 15/2/32 Item 8 and meeting of 28/2/36 Item 34.

[428] FD1 2513. Letter from Green to Barrett 28 April 1947.

a scabies treatment and other companies submitted occasional requests for trials, but Green and his advisers were reluctant to reactivate the TTC.[429] Its moment had passed. The leading companies had formed their own organisation, The Therapeutic Research Corporation, to foster co-operation between manufacturers and represent their interests.[430] Green's account of the TTC was given in a Bradshaw Lecture to the Royal College of Physicians in November 1954.[431] The next sections describe the composition of the TTC, its relationship with the drug companies, and the methods of testing it promoted.

Committee membership

The initial membership was largely composed of MRC clinical scientists and those with close connection to the MRC, through Committee membership, or grant award. The one exception appears to be John Thompson-Walker, a senior figure from the Royal College of Surgeons (Table 7).

Table 7: initial membership of the MRC Therapeutic Trials Committee

Name	Position
TR Elliott	Clinical Research Department NIMR. At the time Director of the Medical Professorial Unit UCL. Member of Committee/Council 1920-26, 1927-31
EF Buzzard	Regius Professor of Medicine, University of Oxford
HH Dale	Director of NIMR
AWM Ellis	Director of Medical Professorial Unit, London Hospital
FR Fraser	Director of Medical Professorial Unit, St Bartholomew's Hospital
John Parsons	(Ophthalmologist) Member of Council 1928-32
John Ryle	(Clinician. Later Professor of Social

[429] FD1 2513. Correspondence between Green and LJ Witts, 19 August 1947 and 30 August 1947 indicates the extent to which they felt other avenues were open to manufacturers seeking trials and to clinicians seeking evidence.

[430] See footnote 152 on the TRC, in this chapter.

[431] Green 1954.

	Medicine, University of Oxford) Member of Council 1935-39.
JW Thomson-Walker	Hunterian Professor of Surgery, Royal College of Surgeons
Wilfred Trotter	(Surgeon) Member of Council 1929-33
DPD Wilkie	(Surgeon) Regius Professor of Surgery, University of Edinburgh. Member of Council 1933-37. Later Director of MRC Unit for Clinical Research in Surgery.
FHK Green	MRC Headquarters staff (Secretary to the Committee)

Source: FD1 5319 TTC Minute book

This was not to be a committee that shared its responsibility with industry, but one in which authority rested very firmly with the MRC. From the earliest stage the TTC used the network of clinical academic research units created by the MRC and London teaching hospitals in the wake of the Haldane report. The membership of the TTC did not change fundamentally over its lifetime. The membership was strengthened and broadened by the inclusion of Lord Dawson, President of the Royal College of Physicians, and Professor Edward Mellanby, Professor of Medicine in Sheffield joined from 1931. JA Gunn, Professor of Pharmacology at the University of Oxford, and Col. LW Harrison, head of the venereal diseases department at St Thomas's Hospital joined the Committee in 1937. Also in 1937, arrangements were made for Austin Bradford Hill to join the committee. The minutes of the MRC meeting of 18/3/38 record that 'it was agreed to appoint Dr A Bradford Hill as an additional member of the [Therapeutic Trials] Committee, in view of the fact that some of the trials organised were on a statistical scale'.[432]

[432] FD1 2505. Bradford Hill was at the time a member of Major Greenwood's staff at the MRC's Statistical Unit, based at the London School of Hygiene and Tropical Medicine. The letter of invitation records 'The Council have been glad to accept the recommendation that you should be invited to become a member of their Therapeutic Trials Committee, as you already have - with Greenwood's consent - kindly agreed to do.' (FD1 2505, 23/3/38 Landsborough Thomson to Bradford Hill).

Relationship with drug manufacturers

The drug manufacturers needed the TTC, at least at the outset, because they lacked ready access to clinicians who might provide a scientific endorsement of their products.[433] Following announcement of the TTC in the general and medical press in July 1931, the MRC received a steady supply of enquiries about testing. Allison Macbeth, British representative of the Dutch manufacturer Organon[434] wrote in 1933 to ask 'what procedure would you suggest, whereby the Organon series of concentrated liver extracts known as Pernaemon can be clinically tested?'.[435] In requesting a trial of an anti-helminthic, The British Drug Houses wrote in 1934 'we feel ourselves at a great disadvantage in that so far we have not received any official records suitable for publication describing its effectiveness'.[436]

For its part, the TTC needed the manufacturers because they were more active than government laboratories in developing new chemicals what might be valuable scientific substances. Despite the fact that this was the rationale for creating the TTC, the research

[433] During the 1920s, companies such as Glaxo and Allen and Hanbury's were largely concerned with dried milk and infant foods. The promotion of infant foods brought Glaxo representatives into contact with doctors, hospitals, charitable institutions, chemists and maternity and child welfare departments (Davenport Hines 1992 p94-95). For details of infant feeding schemes in Britain, see Dwork 1987 section 2. The development of one infant welfare scheme is described in detail by Marland 1993. See also Apple 1980 and Apple 1995 for a general discussion of the role of milk in the construction of scientific motherhood. However, by the early 1930's the infant food market was in decline, due, according to a survey carried out by Allen & Hanbury's (Tweedale 1990 p151), to the falling birth rate and the growing preference for whole milk, caused ironically by studies such as that of Corry Mann. Glaxo's diversification into pharmaceutical products stems from the 1930s, but food products still accounted for over half its income at the outbreak of the Second World War (Davenport Hines 1992 p136).

[434] NV Organon was founded in 1923 as a collaboration between the government funded Pharmaco-Therapeutic Laboratory at the University of Amsterdam and the Zwanenberg Slaughterhouse. The company, based in a laboratory at the Slaughterhouse, initially produced and marketed insulin, rapidly moved into the production of a female sex hormone, patented in 1925, called Menformon. Early clinical trials of Menformon took place in German gynaecological clinics, a reflection of their greater willingness to participate in studies. For material on Organon, see Oudshoorn 1990b and Oudshoorn 1994 p68-72 and chapter 5.

[435] FD1 2501 Macbeth to Green 22 April 1933.

[436] FD1 2502 Letter from British Drug Houses to TTC 23 January 1934.

output of the five leading British companies at this time was modest.[437] The TTC could not therefore restrict its work to British companies. During the main period of its activity roughly half of the applications to the TTC for clinical trials came from British companies, some of which were made by British companies acting as agent for a foreign manufacturer. The rest came mainly from German, Swiss and Dutch and American companies.

Despite the mutual need, both parties had reason to be wary of each other. The MRC would not want to be seen as working too closely with commercial manufacturers.[438] In creating the machinery of the TTC the MRC set up terms and conditions which emphasised both its independence from the manufacturers and its moral authority over the process of submission and selection.[439] Clinicians who had worked directly with manufacturers or commercial laboratories were viewed with suspicion by the TTC.[440] Nor would the MRC wish to be associated with all chemical manufacturers or any of the products of the members of the ABCM.[441]

However, at the outset the TTC was unprepared to receive applications. It was only after ABCM enquired about synthetic menthol on behalf of A. Boake Roberts that Green drew up an application form, intended to screen out the food products of drug companies and

[437] The combined publication output of BDH, Boots, Glaxo, May & Baker and Burroughs Wellcome between 1936 and 1941 was 307 papers. (Robson 1989).

[438] Publicly, at least, the MRC liked to distance itself from commercial organisations at this time. In response to a request from Sir Frederick Menzies for assistance with securing analysis of placentas at a commercial laboratory, Edward Mellanby replied 'the MRC will not co-operate so long as it involves assistance from a commercial firm'. FD1 2503 Letter to Menzies from Mellanby, 17 June 1935.

[439] The general tone of the early correspondence in the PRO files suggests that the MRC regarded itself as very much the senior partner. No representative of a commercial company ever had a place on the committee.

[440] FD1 2503 Mellanby to Menzies 18 June 1935.

[441] Amongst the early preparations summarily rejected by Green on behalf of the TTC were synthetic cough drops, patent foods, and vitaminised products.

products lacking in known active chemicals.[442] The application schedule included the following sections:

Purpose (of the substance)
Special advantages
Suggested mode of administration and dosage
Chemical formula and physical properties
Whether process is protected by patent
Pharmacological and toxicity tests
Particulars of any previous clinical trials
Whether substance is already on the market
References[443]

The terms and conditions drawn up by the TTC demarcated the type of scientific product the MRC wished to test. Applications were only accepted if they were for single active chemicals of known structure and which had shown biological activity in laboratory tests. Applications to have natural substances, food products, patent medicines and secret mixtures tested were invariably dismissed by Green.[444]

From their perspective, the manufacturers might well have preferred to work independently of the MRC. In submitting applications they had to reveal the structure of their new substances; they were encouraged to delay marketing activity until tests were complete,[445] and they had to allow MRC selected experts to test their products. Finally,

[442] FD1 2498 Pratt to Landsborough Thomson 12 May 1931.

[443] FD1 2498 Application Schedule. An example is shown in figure 3.

[444] Sometimes though, quack remedies caused far more work than regular applications. Sometimes applicants showed great persistence in putting their substance before the TTC. Umckaloabo, a secret remedy for the treatment of tuberculosis, had none of the qualities the TTC was looking for. However, its use was promoted by the Committee for the Investigation of Treatments of Tuberculosis, a charity with a network of upper class and parliamentary supporters. The TTC initially rejected application for a trial. However Lady Malcolm secured a meeting with Sir Edward Mellanby at which he agreed in principle to a trial provided the application form was completed and a specimen of Umckaloabo supplied. To the surprise of Mellanby, a sample of the root was supplied. The MRC was eventually able to reject the application because a herbarium specimen could not be supplied.

[445] This was not generally the rule when the companies were able to arrange their own tests. In these instances testing and advertising were part of the same activity of creating a market. A particularly clear example of the market-creating role of clinical trials concerns Metformon, which was marketed alongside tests designed to establish what the indications for its use might be. (Oudshoorn 1995 chapter 5).

they had to agree to the principle that results, favourable or not, might be published in a manner and place as the TTC saw fit.

In practice, the TTC worked hard to support British pharmaceutical manufacturers. This is particularly evident in the way that manufacturers' interests were recognised when under threat from foreign companies. In 1931 Green tried to speed up the completion of the trial of Harmol (a treatment for angina, see below) because an equivalent formulation was being tested for the German company Merck.[446] In 1932 Green reported to the TTC that he had persuaded JF Wilkinson, a haematologist working in Manchester, and financially supported by the MRC, not to publish results showing the comparatively poor activity of British preparations of hog stomach in the treatment of pernicious anaemia.[447] In 1935, Mellanby turned down a request for scientific co-operation between MRC and the American Council on Pharmacy and Chemistry because of sensitivity of British commercial interests.[448] In 1938, EW Assinder writing in the Birmingham Medical Review claimed that clinical trials showed the German preparation of Neoarsphenamine was superior to the

[446] On this matter, see the correspondence between Green and Gunn, and Green and Evans in December 1931, in FD1 2516.

[447] FD1 2499. Confidential supplementary item dated 12 January 1932 circulated before the second meeting of the TTC.

[448] The CPC was founded by the American Medical Association in 1906 to judge the claims made by drug manufacturers. The CPC is discussed by Marks 1997 chapter 1. See especially p24.

Figure 3: Application schedule to the TTC

British product. This caused a flurry of activity in the TTC, and testing of the British drug Novostab Batch n77, sold by Boots.[449]

The supportive approach towards the manufacturers is seen in also in the TTC's approach to publishing results. Although the policy of the TTC stressed freedom to publish, it did not always seek to publish the results of a trial when they were disappointing. Between 1931 and 1936 the trials of amylmetacresol, halarsol, nonylharmol hydrochloride and four other substances were not published because of poor results.[450] However, trial results were rarely definite, and the companies may have regarded any publication as useful to their marketing. The TTC meeting of 11 February 1937 discussed the possibility of 'a special form of publication for those reports of any anti-syphilitic remedy which might be found to be satisfactory, but which did not represent any real advance',[451] suggesting that members thought the TTC had set the threshold for publication too stringently.

Changing relationship between the TTC and drug manufacturers

The changing fortunes of the drug companies in securing testimonials and interest from clinicians can be described in several ways. Viewed in terms of sales, the 1930s were a period of considerable growth, suggesting that clinicians became more open to using commercial products. Despite the economic depression, May & Baker sales grew steadily, due in the main to sales of specialty drugs on behalf of its (by now) parent company Rhone-Poulenc.[452] Glaxo's income from pharmaceuticals doubled between

[449] FD1 2507. Correspondence between Sir Henry Dale and LT Harrison October 1938.FD1 2509 Papers on Novostab.

[450] See notes in FD1 2504. The general approach of the TTC appears to be that if the company was willing to accept the results of a poor clinical trial it was willing to forego publication.

[451] FD1 5319 TTC minute book. TTC minutes 11 February 1937, minute 41.

[452] Slinn 1984 p118.

1935/36 and 1938/39.[453] During this period the drug companies became more assertive in their relations with the TTC.

In terms of the direct relationship between the TTC and companies, the general tone of correspondence between the TTC and ABCM, and the rules drawn up by the TTC for handling trials, show that the ABCM adopted a submissive role in its dealings with the TTC in 1931. However, individual companies soon by-passed the ABCM and dealt with the TTC directly.[454] The application structure set up by Green began to break down after the first few meetings, making it difficult to determine whether or not substances had been formally considered.[455] By 1938, companies were beginning to dictate terms to the TTC. In a submission to the TTC, Ciba, proposed to make desoxycorticosterone acetate[456] available to the committee but stated they could not be bound by the usual undertaking not to issue it to independent workers also, as other firms were interested in marketing the product.[457]

Increasing success gave the companies greater freedom.[458] M&B 693[459] was an immediate worldwide success, following tests carried out by Lionel Whitby, Assistant

[453] Davenport Hines 1992 p88.

[454] This happened from an early stage. In August 1931, the General Manager of ABCM wrote to Green asking him to refer members of the ABCM back to him if they applied directly to the TTC. He enclosed a list of ABCM members. Group 6 included all the major drug manufacturers, and some 98 members in all. FD1 2499 Letter from General Manager of ABCM to Green 25 August 1931.

[455] For example, pseudo-ephidrene from Burroughs Wellcome was not considered in any way by the TTC but reported results at the 5th meeting on 5th March 1934.FD1 5319 TTC minute book.

[456] A synthetic version of suprarenal hormone.

[457] FD1 5319 TTC minute book. Minutes of 9th meeting 14 July 1938. Item 52.

[458] Tests on Prontosil were accepted at the TTC meeting of 28 February 1936. Results were rapidly produced and impressive, but left many questions about efficacy and toxicity un-resolved. However, when the question of further tests was discussed in 1938, it was not considered necessary for the TTC to sponsor further research because testing was so widespread FD1 5319 Minute book 8th meeting 7 February 1938 Item 45C para. (f).

[459] M&B 693 or Dagenan, a chemotherapeutic substance for streptococcal bacteria was first synthesised at the Wandsworth research laboratory of May & Baker in November 1937 (Slinn 1984 p124. The test book showing the entry for substance 693 is shown on p123).

Pathologist at the Bland-Sutton Institute of Pathology, Middlesex Hospital funded directly by May & Baker.[460] At its last meeting the TTC discussed a preliminary report on M&B 693, [461] despite the fact that it had not received an application from May & Baker, nor had Whitby's research had any connection with the MRC. The new relationship between TTC and drug companies was complete by 1941, when the leading pharmaceutical companies formed their own organisation, the Therapeutic Research Corporation, a reflection of their growing confidence and ability to organise their own research and clinical trials.[462]

Methods of testing

The most important fact about the methodology promoted by the TTC is, from the point of view of the development of clinical trial methodology, a negative one. During its existence it did not organise one rigorous comparative clinical trial, despite *prima facie* evidence of the problems of not doing so.[463] None of the factors that were later to be recognised as vital to producing meaningful evaluations of therapies were advocated by the TTC. By standards soon to be regarded as the norm, there was little attempt to frame research questions, little attempt to select patients, no random allocation, little systematic recording of results, and almost no attempt to measure results. In the case of multi-centre

[460] Whitby 1938. Whitby's study used experimental infection in mice to compare M&B 693 with sulphanilamide. He found that M&B 693 was as effective but less toxic. The paper acknowledges Dr Ewins at May & Baker for the preparation of compounds and for funding the studies at the Middlesex.

[461] FD1 5319 TTC minute book. Minutes of 10th meeting 28 March 1939. Item 59.

[462] The Therapeutic Research Corporation (TRC) was founded by Boots, May & Baker, British Drug Houses, Glaxo, and Burroughs Wellcome. ICI joined in 1942. The role of the TRC was to promote and integrate the work of drug manufacturers. It enabled them, at least in theory, to share and coordinate research in selected therapeutic areas. It also served as a means for presenting the views of the pharmaceutical industry. Although the TRC did not succeed in promoting co-operation between companies (this came mostly through acquisitions after World War II such as the merger of Glaxo with Allen & Hanbury's and Evans Medical) it did play a role in organising British penicillin production in the early 1940s (Davenport Hines 1992 p138-146).

[463] For example, the group trial of pneumonia anti-serum, discussed at more length in the next chapter.

studies, it was usual for Green to ask the clinicians to work to a standard schedule, but there was no enforcement of a schedule.

However, to categorise the TTC as a failure precludes an understanding of what it was trying to achieve with the resources at its disposal. I shall argue in this section that the 'methodological outlook' of the TTC embraced three conceptions of a clinical trial. The first is as an extension of a laboratory study, in which the patient is the *in vivo* substrate for physiological tests. The second conception of a clinical trial is a test on a series of cases, selected because they are likely to benefit from the drug. In this type of trial the synthetic judgement of an individual clinician forms the positive focus for evaluation of the utility of new drugs.[464] The third concept is that of a comparative clinical test. The latter was a special method, only advocated by the TTC when a British drug was being compared to a foreign produced variant.

Clinical trials as extensions of laboratory studies

Tests on the pure extracts of foxglove, digoxin and digitalinum verum, were agreed at the first TTC meeting on 8th July 1931.[465] EJ Wayne of the Department of Clinical Research, University College Hospital carried out tests.[466] Selected outpatients with auricular fibrillation were admitted to hospital.[467] The patient was rested in bed and ventricular

[464] One distinction between the case series approach and later conceptions of clinical trials is akin to the distinction between the versions of positivism held by Carnap and Popper. See Hacking 1983 p2-6. Carnap saw knowledge being built up by the successive verification of hypotheses, like a clinician testing a promising drug on a series of patients. Popper saw knowledge being built up by destruction of hypotheses.

[465] FD1 5319 TTC Minute Book Meeting 1 Item 3(e) and 3(f). Digoxin was isolated from the leaves of the foxglove *Digitalis lanata*. Digitalinum verum was extracted from the seed of *Digitalis purpurea*. (Wayne 1933) Pharmacologically, they belong to a group known as the cardiac glycosides, a group of chemicals which can be isolated in species of foxglove. Their common action is to slow the heartbeat and increase the force of contraction.

[466] The trial is reported in Wayne 1933.

[467] Patients with chronic bronchitis were excluded because coughing produced erratic variation in the ventricular rhythm.

rates were measured using the electrocardiograph. A low dose of drug was then given intravenously in dilute alcohol solution, with regular measurement of the ventricular rate. Dosage was successively increased over several days. Control patients were given intravenous dilute alcohol only. In three patients the effects of the active drug were compared with that of USPX ouabain, an American standardized preparation of the cardiac glycoside ouabain. In a further 13 cases, digoxin was given by mouth, and in 4 cases digitalinum verum was given by mouth. In these patients, the successive dose of digoxin was varied to stabilise the ventricular rate at between 60 and 70 beats per minute. Further tests were carried out to establish whether or not the drugs could be given by subcutaneous injection.

The results showed that the glycosides slowed the ventricular rate, while the alcohol control did not. Digoxin, but not digitalinum verum, was effective when given orally. Digoxin reduced the excess filling of the veins in the neck in 8/10 patients with mild congestive heart failure.

Wayne's glycoside study is characteristic of what can be termed the laboratory approach to drug testing. Reduced to its essentials, the method is close to human vivisection, because its principal aim is to observe the physiological effects of a drug in the human, with little attempt to determine the therapeutic value. The choice of patients who were suffering from a condition that the physiological action of the drug might affect should

not be overlooked however. By carrying out tests on patients who might benefit from the drug there was always a chance that therapeutic value would be demonstrated.[468]

In summary, the laboratory method used by the TTC involved a small sample of carefully selected outpatients admitted to research beds; given varying doses of a bio-active drug with effects relevant to the clinical condition of the patient. It employed careful measurement over a relatively short period of time of the effects of the drug, principally on some physiological variable. Alongside the primary aim there might be an attempt to standardise dosage and determine the best means of administration.

The case series approach

The commonest method used by the TTC for testing therapeutic substances would now be called the case series approach. As the modern name suggests the substance is given to a selected series of patients. Clinical impressions and other measurements are taken during the trial, and these are subsequently considered in order to assess the benefits of the drug.

The first example described below shows the case-series approach in its purest form.[469] In the second example, the experimenter added two statistical features. Firstly, a control group was used, to measure the rate of healing without calciferol. Secondly, the experimenter states that in order to reduce observer error, the same person carried out all

[468] There is a suggestion that Elliott and others were concerned about the charge of human experimentation being directed against the studies taking place in Lewis's laboratory, in which Wayne worked (Witts 1974) However there was no discussion of the ethics of such work at the TTC.

[469] There are several other examples from the TTC of extremely simple case series clinical trials. An example is the study on the value of amyl salicylate as a treatment for burns in which the drug was tested on three types of case at Wilkie's Department of Surgery in Edinburgh. Tests arranged by Green in September 1933. (FD1 2533 Green to Wilkie 4 September 1933).

observations.[470] This conception of observer variation is in direct opposition to the modern one, which would reduce intra-observer error by introducing several observers, and measure the inter-observer variation.

First example

Harmol hydrochloride, derived from an alkaloid found in the seeds of *Peganum harmala* (wild rue), was submitted for testing by Boots Pure Drug Ltd. as a coronary dilator to the first meeting of the TTC.[471] The drug was supplied to doctors at the professorial units at The London and Guy's Hospital, and to Crighton Bramwell, Assistant Physician at the Manchester Royal Infirmary. By the second meeting of the TTC, a full report had been submitted. It showed that Harmol was effective in giving short-term relief from angina, but was an irritant when given subcutaneously, and had produced renal colic in several patients. In the interim, Boots had offered for test the lactate salt of Harmol, which they claimed to be more soluble.[472] At the third meeting it was reported that Harmol had little value. The committee agreed to accept the application from Boots to have o-n-propylharmol lactate tested, and expressed the hope that the original researchers would test this substance also.[473] The results with Harmol and o-n-propylharmol were discussed at the fourth meeting. Harmol was agreed to be ineffective. Evans and Campbell

[470] 'To reduce the personal error to a minimum, all clinical examinations and taking of histories were carried out by myself week by week'. (Spence 1933 p911).

[471] The application was submitted on 20th May 1931. (FD1 2516 Submission from Boots Pure Drug Ltd. Accepted by TTC, FD1 5319 TTC Minute Book. Minutes of the first meeting, 8 July 1931. Item 3(c)) Harmol had been raised briefly by the CC. The Boots application was supported by a paper by JA Gunn, Professor of Pharmacology at the University of Oxford. The close relationship between some academic pharmacologists, drug companies and the MRC is illustrated by the detail of the Harmol submission. Gunn appears to have acted as an intermediary between Green and Boots, with messages about German tests and improved formulations reaching Green via Gunn. Following a note from Gunn saying that he had [in fact Boots had supplied him with] a better version of Harmol (FD1 2516 Gunn to Green 27/10/31), Green wrote back suggesting that this might be tested by the TTC (FD1 2516 Green to Gunn 24/11/31).

[472] FD1 5319 TTC Minute Book. Minutes of second meeting 15 January 1932. Item 6(b).

[473] FD1 5319 TTC Minute Book. Minutes of third meeting 8 July 1932 Item 14(d) and 18.

considered o-n-propylharmol useless, but Bramwell considered it to be useful in 'early' cases of coronary disease. Seeking to publish the results, the committee deferred the decision, proposing a small conference, to be attended by Professor Gunn, who had worked with Boots in submitting the original application.[474] The results were published in a short report in the Lancet in July 1933.[475]

The aim of the trial was to test both drugs in the following situations: (1) to cut short an attack (2) to forestall an attack (3) to reduce the frequency of attack through continuous treatment. The methodology consisted of giving the drug to selected patients and observing the clinical results. The results, published in Lewis's journal Clinical Science, showed that Harmol appeared to be effective in shortening an attack, but was less acceptable than trinitrin or amyl nitrite. Continuous testing of Harmol on 41 cases, and o-n-propylharmol on 30 cases, produced equivocal results. In seven Harmol cases, patients appeared to benefit. In 4 of the cases (the 'early' cases observed by Bramwell) o-n-propylharmol was effective, but was ineffective in others.

Example 2

The second case series concerns the clinical trial of calciferol as a treatment for rickets carried out on behalf of the TTC by JC Spence, Assistant Physician at the Royal Victoria Infirmary, Newcastle.[476] The purpose of the study was to determine the therapeutic value

[474] Delay was needed because the TTC wished to avoid separate, and probably conflicting, publications. The final combined report nicely manages to acknowledge differing valuations of o-n- propylharmol lactate without drawing attention to them as differences, but the occasion for further research. See FD1 2516 Letter from Green to Bramwell 3 April 1933.

[475] Bramwell 1933.

[476] The decision to test calciferol was made at the second meeting of the TTC (FD1 5319 TTC Minute Book. Minutes of second meeting, 15 January 1932, Item 9). The application evidently did not come from British Drug Houses, the manufacturer. It seems likely that the testing of calciferol was initiated by Edward Mellanby, who had pioneered work on the identification of anti-rachitic factors in a series of experiments on puppies sponsored by the MRC (see Landsborough Thomson 1975 p76-78). At this time Mellanby was deeply engaged in the politics of nutritional

of calciferol (pure vitamin D), as opposed to cod-liver oil and the International Standardised version of Vitamin D. The purpose was not therefore to confirm the anti-rachitic effect of vitamin D, but to test the value of its purified version.

Forty-four cases of uncomplicated rickets were chosen for the study, which began in February 1932.[477] Of these 19 were rejected because they had received gifts of food or because their fathers' had obtained employment. Of the remaining 25, 3 were chosen to act as standards for the optimum rate of cure. 'These were put under good hygienic conditions and given an adequate anti-rachitic diet containing milk, meat, liver, eggs, butter and vegetables, with one ounce of cod-liver oil daily in two cases, and four tablets of calciferol daily in the other.'[478] The remaining group of 22 was studied in various ways. The majority lived at home and received 3 cm^3 of an oily solution of calciferol. 8 cases went without treatment, partly to control for the known healing effects of sunlight. Two pairs of twins were observed, one of each pair acting as control. One boy, part of a family of five living in one room, confined to bed because he was unable to stand, was closely studied for 12 weeks.

Serial radiographs of the wrist were used to measure the extent of calcification, with the radiographs of the optimally treated children being used as a standard. Results showed that calciferol 'had an active curative effect on the rickets, and that it produced healing at

---

research, and more widely with the relationship between MRC and Ministry of Health. Shortly to become Secretary to the MRC, Mellanby would have welcomed further confirmatory results on the value of Vitamin D in rickets, since this would have supported his views on the place of nutritional research in supporting the public health, as opposed to the Ministry of Health view which emphasised the role of housing. On the politics of nutritional research, see Petty 1989, Smith and Nicolson 1989, Aronson 1982. The links between nutrition science and the food industry, especially the role of vitamins in shaping the advertising of specialised foods at this time are discussed in Horrocks 1995.

[477] The report of the study is in Spence 1933.

[478] Spence 1933 p911.

an optimum rate, acting as quickly and effectively as the usual therapeutic doses of cod-liver oil or irradiated ergosterol'.[479]

This case series extended the laboratory approach, as far as possible, into the community. The boy confined to bed was regarded as a boon because it allowed the researchers to closely monitor his conditions. The use of control patients gave the researchers some ability to account for the factors affecting the results, such as sunlight. The use of radiographs allowed a precise measurement of healing, in a manner analogous to a laboratory study.

Comparative clinical trials

The form of test used to determine the effectiveness of ergotoxine shows both the approach to comparative tests adopted by the TTC, and the reluctance of clinicians to adopt that approach. Ergotoxine ethanesulphonate, submitted by Burroughs Wellcome to the TTC in April 1931, was considered to be the active ingredient of ergot. Ergot was introduced into medicine in 1807 and had two main uses, both relating to its stimulant action on uterine muscles. It was used (mainly in the past) to intensify the contractions of a sluggish labour, and to stem haemorrhage after delivery by promoting the contraction of the uterus. Experience with ergot led nineteenth century textbook authors to conclude that it might be an effective treatment for haemorrhage, but was potentially fatal for mother and child when used to hasten a sluggish labour. The identification of the 'active element' of ergot may have been an interesting chemical problem, but in relation to the available body of therapeutic knowledge it is difficult to see what potential therapeutic

[479] Spence 1933 p915. The conclusion was precisely that desired by the advocates of purified vitamins. But was it justified? While the study showed clearly that children receiving no supplement healed more slowly, no direct comparison was made between heterogeneous preparations and the pure vitamin. Furthermore, in the one case studied most closely, healing slowed after the eighth week due to a deteriorating general diet.

advance it represented. Nevertheless, Burroughs Wellcome purified it, and the TTC accepted ergotoxine for trial in July 1931. Green wrote to Dr Aleck Bourne, a leading obstetrician at Queen Charlotte's Hospital asking him if he would be willing to test ergotoxine because of his special experience with ergot derivatives.[480] In referring to Bourne's earlier work with ergot derivatives Green appeared to Bourne to be implying that he was looking for a physiological experiment to be carried out on women in labour. On this basis Bourne refused the request. He offered Green some hope however:

> *It will be possible however, for us to use the drug after delivery of the child for cases of post partum haemorrhage, but as you can understand, the results obtained by injecting it as an ordinary clinical treatment can give no exact records but only clinical impressions. Unless it is employed for a very large number of cases of haemorrhage the impressions obtained by sisters, house surgeons and those in attendance cannot be very satisfactory evidence. However, if you will send me a supply of ergotoxine ethanesulphonate, I will have it used and careful records kept.*[481]

Green replied on the 27th

> *I note that you propose using the drugs in cases of post partum haemorrhage. The committee were anxious that its therapeutic effect should be tested against that of ergotamine in order to decide finally whether these two alkaloids, being similar pharmacologically, have a parallel clinical action. I take it that cases of post partum haemorrhage treated with ergotamine could conveniently be used for comparison with those treated with ergotoxine, and would make effective controls for the experiment?*[482]

Green's letter reveals that in accepting ergotoxine for trial the TTC was seeking to provide a British company with a scientific testimonial to allow it to compete with

[480] FD1 2517 Green to Bourne 23 July 1931.

[481] FD1 2517 Bourne to Green 25 July 1931.

[482] FD1 2517 Green to Bourne 27 July 1931.

foreign purifications of ergot such as Sandoz's ergotamine tartrate. Vials of ergotoxine were supplied to Bourne in September.

In November, Green wrote to Bourne to enquire about his results, and received a strikingly relaxed reply:

> *'I had a report yesterday, from those who have been using this drug at Queen Charlotte's Hospital, that it appears to produce the same clinical results as other preparations of ergotoxine. It must be clear to you that these remarks cannot have any scientific value whatever, as the observations have been chiefly made by the labour ward sisters and the house surgeons. No system of controls is possible, and in most cases the drug has been used in a routine way without, I fear, and intelligent appreciation that an investigation was being made. I explained, however, at the onset that I could not give you anything more than clinical impressions.*[483]

This reply prompted Green to immediately contact Sir Henry Dale, fearing the disappointment of Burroughs Wellcome if this was all the TTC could provide.[484] On Dale's advice he contacted Elliott at UCH who in turn contacted his colleague Professor Browne. Elliott wrote:

> *I regard the work as worth doing, in order to show that the British product of Burroughs Wellcome is as good as the foreign Sandoz. All that is needed is careful analysis... presumably in alternate cases, that are capable of ordinary clinical measurement. Would Moir care to undertake this? I fear there is no promise of an honorarium, but the work would bring your unit into closer association with the MRC*[485]

Browne arranged for Dr Chasser Moir to carry out tests. In accepting the task Moir asserted that a clinical trial comparing ergotoxine with ergotamine would be impossible because

---

[483] FD1 2517 Bourne to Green 26 November 1931.

[484] FD1 2517 Green to Dale 4 December 1931.

[485] FD1 2517 Elliott to Browne 12th December 1931.

> *' In the puerperium, for example, the rate of involution varies in accordance with so many conditions, e.g. anaemia, state of health of the patient, history, degree of post partum haemorrhage, presence and degree of sepsis and so on, that it would be impossible, I think, to say to what extent any drug influences it.*[486]

Given the confusing effects of patient characteristics and the difficulties of measuring involution the only reliable measure of the efficacy of ergot derivatives was a physiological one:

> *Then again involution is very difficult to measure.... the only useful test, to our minds, would be to put a bag inside the uterus, connect it with a manometer, inject the drug and note the effect on uterine infection (corrected by Elliott to contraction).* [487]

Moir carried out the experiment and published the results in the BMJ in 1932.[488] Even in this comparative trial the emphasis was on physiological experiment. A bag was inserted into the uterus of each woman in the study during labour. To the bag was attached a tube which was connected to a recording apparatus. The bag was left in pace several days after birth to make recordings. There was no suggestion that the women in the trial needed uterine stimulation, and no record of the clinical outcome of giving the drug.

Preference for laboratory tests

Thus although the companies needed testimonials of clinical trials on patients, for researchers associated with the TTC patient testing was scientifically problematic. Indeed, throughout this period the TTC and its researchers showed a preference for objective physiological tests as a way to determine the efficacy of therapeutic substances. The tendency to regard physiological experimentation as the standard pervaded much of

[486] In FD1 2517 Browne to Elliott 14th December 1931.

[487] FD1 2517 Browne to Elliott 14th December 1931.

[488] Moir 1932**a** and Moir 1932**b**.

its work. JF Wilkinson for example expressed the hope that one day a laboratory test for the effectiveness of liver extract treatments for pernicious anaemia would become available.[489]

An event that took place towards the end of its existence highlights the continuing importance of a physiological understanding of therapeutics to the MRC's approach to clinical trials. In the case study below, the MRC is seen to be on the defensive when challenged by EW Assinder's research on anti-syphilitic treatment, in which clinical trials on humans appeared to contradict MRC findings based on mice and rat tests.

EW Assinder and controversy over Novostab

The TTC's handling of the controversy surrounding Novostab provides a case study which illuminates the MRC approach to clinical trials. The episode began in 1938 when a venereologist, Dr EW Assinder, published a comparative trial of three treatments for syphilis. The substances involved were Novostab, a neo-arsphenamine compound produced by Boots Pure Drug Company, Mapharside from the American manufacturer Parke Davis and Company,[490] and Neo-salvarsan, presumably from a German manufacturer (presumably either Bayer or Hoechst).[491] The results showed that Neo-salvarsan was the most effective treatment, as measured by its ability to clear spirochetes from the exudate of syphilitic sores.

[489] FD1 2501 Wilkinson to Green May 1933.

[490] Assinder incorrectly labeled the American drug 'American Neoarsphenamine' in his original paper.

[491] Assinder 1938. Only the country of origin was identified in the original paper. Harrison wrote to Assinder on 14 October 1938 to enquire what the substances were; Assinder replied on 19th October. See FD1 2508.

A small-scale trial in a local medical magazine was to lead to considerable efforts at the MRC.[492] Why was this? Firstly, the conclusion, and the way it was presented. Assinder concluded that 'I am sorry to say that the best manufacturers are German…there is no doubt, I think, that No. 3 should be used as routine'.[493] Such a conclusion was undoubtedly unhelpful, and even unpatriotic in 1938, and was responsible for bringing an otherwise insignificant publication to the attention of the Ministry of Health. However, although the results clearly favoured the un-named German preparation, readers of Assinder's study would have been unlikely (and unable since the brands were not identified) to switch on the basis of a small-scale trial carried out in one hospital on a single batch of a drug.[494] The second reason for MRC interest was that in March 1939 the Ministry of Health began to enquire at the MRC about Assinder's paper, an action which prompted Green to write to Dale, seeking his advice on how to organise a response at the MRC.[495]

It turned out however that Dale was aware of Assinder's paper, and had begun his own investigation into the study. The reason was that Assinder's methodology challenged the basis of MRC authority over drugs controlled by the Therapeutic Substances Act. As has been discussed, the MRC was responsible, through the provisions of the 1925

[492] Correspondence between the MRC and Ministry of Health in FD1 2508 suggests that one further reason for the MRC's seemingly disproportionate interest was the impression it retained that the Ministry regarded Assinder's work as a serious challenge to the MRC's position. See FD1 2508 Green to Dale March 7th 1939: 'I am quite prepared, however, for you to tell me that I am wrong in venturing to minimise the practical significance of his claims, and it seems evident that the Ministry are inclined to take the situation seriously'.

[493] Assinder 1938 p16.

[494] Henry Dale, who was responsible for the interest shown at the MRC regarded the results as negligible. FD1 2508 Dale to Green March 8th 1939.

[495] FD1 2508 Green to Dale March 7th 1939: 'I am quite prepared, however, for you to tell me that I am wrong in venturing to minimise the practical significance of his claims, and it seems evident that the Ministry are inclined to take the situation seriously'

Therapeutic Substances Act, for the regulation of arsenic based anti-syphilitic drugs. The MRC's approach was to compare the potency of new arsenicals with standardised preparations of Salvarsan and Neosalvarsan using the trypanosome test in mice, described earlier in this chapter. In contrast, Assinder was undertaking a simple test on the blood of patients attending his VD clinic.[496] Novostab had been subject to tests by the MRC and Boots. If Assinder's results were correct they challenged at a stroke the validity of the MRC's use of animal rather than human subjects, and laboratory rather than clinical settings to test drugs. The principal table of results from Assinder's study are shown in Table 8 below:

Table 8: principal results from Assinder's comparative trial of anti-syphilitics

Drug	**Cases in which Sp. pallida were found**	**Serum examination after 24 hours or more**	**Toxic effects noted**	**Effect on sore**
English	9	SP present 7 SP absent 2	Very few 1 jaundice	Healing delayed often a month
American	7	SP present 1 SP absent 6	Marked often vomiting	Healing fair, generally about 14 days
German	8	SP present 0 SP absent 8	None noted	Healing rapid – generally a week to 10 days

Source Assinder 1938

Upon receipt of the paper in 1938, Dale sought further information about Assinder's work, using LW Harrison as an intermediary. Correspondence between Dale and Harrison at this time shows that Dale was particularly concerned about the implications of Assinder's paper for the workings of the Therapeutic Substances Act. To this end, Dale welcomed Harrison's suggestion that tests such as Assinder's might be carried out

[496] Ironically, the availability of laboratory facilities in VD treatment centres is due to LW Harrison, who specified a small lab area in his design of the VD centre at St Thomas's Hospital. The design is shown in King 1974.

on new batches of British neo-arsphenamine by VD clinics. This would allow the authorities to respond to Assinder's challenge without disturbing the Act:

> *I am glad you think that it would be worth while testing whether a subsidiary clinical trial of the therapeutic performance of the different batches of this product could be done regularly and systematically… I suggest that if this were done as an administrative matter under the VD Scheme, the difficulty would be avoided of having to make and submit to Parliament a new regulation under the Therapeutic Substances Act, involving trials on human patients.*[497]

Dale's first practical response to Assinder's paper was to have batches of Novostab re-tested, first by Boots in December 1938,[498] and then at NIMR in January 1939.[499] The Boots report showed that Novostab had a slightly lower potency than Neo-salvarsan; the NIMR test showed they had a slightly greater potency. The results of the repeat test provided some re-assurance. However, as Dale recognised, they did nothing to refute the charge that the results of laboratory tests on mice did not predict the efficacy of drugs in man.[500] Nonetheless, Dale took no further action until the Ministry of Health raised the matter in March 1939.

The enquiry from the Ministry evidently annoyed Dale. In response to Green's suggestion that the TTC should organise a trial of Novostab, Dale responded:

---

[497] FD1 2508 Dale to Harrison October 26th 1938. In 1937 there were 186 treatment centres in England and Wales. A survey by the Ministry of Health in 1936 showed that 88% of the arsenobenzene compounds provided under the Venereal Diseases Act of 1917 were used by VD centres. (Political and Economic Planning 1937 p291).

[498] FD1 2508 Report on re-examination of Novostab N77 for toxicity and therapeutic efficiency. 2nd January 1939. Also FD1 2508 Letter from Broom to Strangeways 30th December 1938.

[499] FD1 2508 Report on the therapeutic tests applied to Novostab N77 21/3/39. Department of Biological Standards, NIMR.

[500] FD1 2508 Dale to Harrison 30th January 1939.

> *I do not myself understand why they [the TTC] should be asked to advise the Ministry of Health on the use of information concerning individual batches of a well-known drug, obtained at the instigation of the Ministry's own V.D experts.*[501]

However, Dale had no option but to go along with the proposed TTC organised trial, because he doubted that Harrison or the network of VD clinics could arrange the trials Harrison had proposed.[502]. In place of testing under the auspices of the VD clinics, it was agreed on 16th May 1939 that the sub-committee should undertake a clinical trial of Novostab N77.[503] On 1st June Harrison sent a circular to 6 members of the Sub-committee (Anwyl Davies, Burke, Kemble, Lloyd, McElligott, Rorke) asking them to test ampoules of Novostab N 77, and outlining the methodology to be adopted:

> *'It is suggested that the batch be tested by each collaborator on ten early cases of syphilis, either primary or secondary, with discharge containing easily demonstrable S pallida. For the sake of uniformity it is suggested that all six collaborators adopt the following plan of investigation unless any modification is agreed upon in advance. The first dose for an adult to be 0.45 gm and the first examination of the discharge, 20 to 24 hours later'.*[504]

I have looked in the archives and in published accounts for the outcome of this trial, but nothing has come to light.[505] It may be that the proposed study was a casualty of war. Given the lack of enthusiasm for the study, its demise would not have troubled Dale,

---

[501] FD1 2508 Dale to Green 8th March 1939, marked confidential.

[502] FD1 2508 Dale to Green 8th March 1939, marked confidential. And Dale to Green 24th March 1939. This proved to be the case. Harrison admitted to Green that the enthusiasm for testing varied considerably between VD clinics, and there were 'serious practical difficulties' in sharing information between the clinics. FD1 2508 Green to Dale 14th March 1939, marked confidential.

[503] FD1 5320 Item (15), minutes of Sub-Committee on antisyphilitic remedies, 4th meeting 16th May 1939, held at the Ministry of Health. The matter was passed directly to the sub-committee without reference to the main TTC.

[504] FD1 2509 Circular from Green to members of the sub-committee on antisyphilitic remedies 1st June 1939.

[505] The MRC file on Novostab ends abruptly at this point. Searches for further information or publications include the MRC Annual reports for 1938-39 and 1939-45; Index Medicus for 1939-1942; the British Journal of Venereal Diseases and the Index of the BMJ between 1939 and 1945.

since any publication might have opened the question about the relationship between animal and human tests.

Summary and conclusion

Despite setting out to undertake basic biomedical research the MRC was drawn into the area of therapeutics at an early stage of its existence because of the need to produce a British version of Salvarsan. Success in this area proved to be the starting point for a programme of work which gave the MRC a leading role in the development of the British and International pharmaceutical manufacturing sector. This role, which included work on establishing biological standards for drugs, and the Therapeutic Substances Act of 1925, gave the MRC considerable authority over the area of therapeutic substances.

Biological standardisation was the MRC's first venture into the evaluation of the effectiveness of treatments. It offered the dominant model adopted by the MRC for interpreting the effectiveness of drugs in the 1920s. Standardisation offered a way of both comprehending and regulating the effects of biologically active substances. It had several benefits. It applied a physiological understanding of therapeutics and thus drew on the MRC's basic science remit. It required only the resources readily available to the MRC in its laboratories, and thus supported the MRC's central facility, the NIMR. Thirdly, it allowed the MRC to establish and maintain scientific and moral authority over the newer research-based drugs beginning to circulate between pharmaceutical manufacturers and clinicians in increasing amounts in the UK, while not entangling it with the day to day fortunes of the industry.

During the later 1920s the MRC, working with the DSIR, committed itself to developing British chemotherapy. At first it drew on the resources of chemists in the employ of the

British State, but when these proved to be insufficient a means was created for the MRC to test commercial products. This necessarily involved testing on patients, because the drug manufacturers had to have shown biological activity in the laboratory before approaching the TTC.

The MRC lacked immediate access to patients, but it was able to use a network of clinical scientists in the professorial units of London medical schools. It did not, because it could not, dictate the methods in which clinical trials should be undertaken. In some cases, trials were little more than physiological studies using human subjects. In some cases, trials consisted of a series of observations. Comparative trials were rarely undertaken.

Despite the ready availability of statistical advice, and the membership of Austin Bradford Hill from 1938, the TTC did not introduce methodological innovation. In the next chapter, I will show that there was a failure to innovate despite clear evidence as early as 1931 of the need for well-organised trials. The evidence is supplied by the trials of pneumonia antiserum whose tests came under the auspices of the TTC. The next chapter also introduces the work of Austin Bradford Hill, who advised on the pneumonia anti-serum trial.

Chapter four

Austin Bradford Hill and the MRC trial of anti-pneumococcus serum

Introduction

The purpose of this relatively short chapter is to discuss the case of an MRC sponsored trial in which a limited form of random allocation was used. The trial, a multi-centre trial of anti-pneumococcus serum, began in 1929 and concluded in 1933. It is important for several reasons. It illustrates the application of what was at the time an advanced technique of statistical design and analysis by clinicians. In doing so it highlights once more the question of why the TTC did not adopt more rigorous methods of trial design and analysis during the 1930s.

The MRC pneumonia anti-serum trial also introduces Austin Bradford Hill, who provided statistical advice on the trial, and who went on to become the trial statistician for the streptomycin trial, discussed more fully in the next chapter. Here I discuss Bradford Hill's early career, especially his work on experimental epidemiology, and the series of articles he published in the Lancet on medical statistics. These are interpreted as ways to improve his own position and the position of epidemiological and statistical methods within the MRC.

Serum treatment for lobar pneumonia

The pneumococcus associated with pneumonia was discovered in the late nineteenth century, the discovery being made by Pasteur in 1880 and independently by Sternberg in

1881. Within a few years Friedlander and Frankel had conclusively identified it as the cause of most pneumonias.[506]

Lobar pneumonia, the most common form of the disease, usually begins dramatically with a chill or sharp pain in the chest. A cough and blood in the sputum tend to occur. Following the chill, the temperature rises as high as 105°F, and in the absence of therapy remain high for several days. The patient experiences rapid and painful respiration, possible delirium, flushing and dry skin. After 5 to 10 days the fever subsides. Complications are frequent; especially serious are those caused by spreading infection, for example meningitis and endocarditis.[507]

Attempts to treat pneumonia with serum began shortly after the discovery of the bacterium.[508] The era of serum treatment for pneumonia began in earnest in 1910, when the German bacteriologists Neufeld and Haendel prepared sera for the different strains of pneumococcus. Unlike diphtheria serum treatment, pneumonia serum met with little success until the bacterium was recognised as existing in immunologically distinct types.[509] Amongst the earliest clinical trials was that reported by Cole in 1913, in which 19 patients were treated with either Type I or II serum.[510] Of the 11 patients with Type I pneumonia, only one died. As Marks notes, by 1917 Cole and his co-workers were

[506] The historical background given here is based on that of Dowling 1972 and Dowling 1977 p45-49. Pasteur identified the principal causative bacterium, now known as Diplococcus pneumoniae, in the saliva of healthy subjects. It was first associated with pneumonia by Carl Friedlander in 1882.

[507] Dowling 1977 p45.

[508] Early work leading to a serum treatment included Frankel's study of 1886 which conferred immunity on rabbits through injection of pneumococci.

[509] America led the work on typing and serum development, thanks to the philanthropy of John T Rockefeller, who funded an Institute – the Rockefeller – whose primary aim was to develop a treatment for pneumonia. The first Director was Rufus Cole.

[510] Cole 1913 JAMA;61:663-8.

reporting a mortality of 7.5% in serum treated patients, well below the 30% normally experienced in America.[511]

However, for at least a decade the use of serum as a treatment for pneumonia remained controversial, because of the variability of results obtained.[512] In the New York and Boston areas, progress was made, but elsewhere serum treatment was regarded with suspicion. The position changed somewhat in 1924, when Lloyd Felton found a method of concentrating and purifying serum.[513] More efficient extraction, biological standardisation, improved methods of typing, and the development of polyvalent serum[514] created the potential for commercial exploitation, with Lederle taking a leading role. Serum also obtained the backing of the Influenza Commission of the Metropolitan Life Insurance Company, who met the costs of serum for several years. The company also provided important statistical support for one of the trials discussed below.

American trials of serum were methodologically innovative. Lilienfeld cites a study by Maxwell Finland and colleagues at the Boston City Hospital, which used an alternate case design during part of their trial of serum treatment.[515] More innovative still was the

---

[511] Marks 1997 discusses the development of antipneumococcal serum as part of the efforts by therapeutic reformers to extend rational therapeutics from teaching hospitals and centres of excellence to the community. The methods employed in Bullowa's clinical trial lend support to Marks' claim that reformers regarded a culture of rational knowledge as necessary to the spread of rational therapeutics.

[512] Empirical evidence of variability of outcome is found in Bullowa 1929a, which shows death rates in treated and untreated cases higher in 1928 compared to 1927 (Bullowa 1929a p337 fig 3) In England, data from the Registrar General's Office showed geographical variation in the incidence and mortality of pneumonia in Britain, and also that the mortality rate varied between epidemics in a manner unrelated to the type of the causative organism, reflecting variability in the virulence of the organism itself (Langley 1931) In early editions of his Principles of Medical Statistics, Bradford Hill used the known variability of outcome in pneumonia as a way of introducing the difference between observed and expected values in a sample, and the effect of sample size on that difference. In later editions, the topic is discussed without specific reference to pneumonia (Bradford Hill 1971 chapter 10).

[513] Felton worked at the Department of Preventive Medicine and Hygiene at Harvard Medical School.

[514] Effective against both Type I and II pneumococcus.

[515] Finland 1930. The results showed a lower mortality and reduced duration of disease in Type I and II pneumonia when serum was used.

study of Jesse Bullowa and colleagues at the Harlem Hospital in New York.[516] Bullowa was acutely aware of the need for 'concurrent control', citing variability in the virulence of the organism and variability in patient response as reasons why this was so.[517]

Accordingly, as with the Boston study, Bullowa used the method of alternate allocation to form study and control groups. Bullowa's report also makes it clear that he recognised the need to control every aspect of the trial if he was to get a clear result:

> *'We have made the conditions of the experiment as similar as possible, by rating the cases, so that we know that equally severe cases shall be in each series, determined by rating them, by taking alternate cases, and by having a uniform standard treatment'*[518]

Bullowa's principal innovation was to use the services of Dr Leon Dublin, a statistician employed by the Metropolitan Life Insurance Company, New York. Under Dublin's guidance, Bullowa argued that the number of subjects needed in a controlled comparative clinical trial should be large enough to allow a two-fold difference in the ratio of the outcome to the standard error of the mean between groups. Although this does not constitute a pre-determined sample size calculation, it gave Bullowa an indicator of when a statistically significant result had been achieved.[519]. Applied to data from the Harlem

---

[516] From the trial report it is likely that the trial took place at two other hospitals also, Bellevue and The New York.

[517] See Bullowa 1929**a** p335 'It is impossible, at the outset, to determine clinically, what patients will recover…Patients may overcome the invading bacteria, and yet succumb as the result of loss of aerating surface, from cardiac inadequacy resulting from the illness, or from disturbance in vital equilibria. This serves to complicate our problem [of evaluation]' On p 337 Bullowa indicates how case severity varies annually, which he imputes to variation in the virulence of the bacterium.

[518] Bullowa 1929**a** p336.

[519] There is no suggestion that Bullowa carried on experimenting until he achieved a significant result. His paper to the symposium on the use of antipneumococcal serum in December 1927 suggests that he understood the role of the alternate case policy in relation to the statistical test applied. 'To evaluate the result of the treatment in pneumonia there are required adequate comparable series with and without treatment. We believe we have obtained such series by the devices adopted' (Bullowa 1928 p343).

Hospital, statistically significant results were found in Type I cases admitted to hospital within four days of onset:

> *'In fact, we have ceased, at Harlem Hospital, to alternate the use of serum in cases in which we find an invasion of the blood stream with Type I organisms, for it is felt that, for those cases statistical proof has been given, and it is unjust to withhold an available life-saving procedure; the ratio of the difference to the standard error in this type is 2.4 to 1.'*[520]

MRC trial of anti-pneumococcus serum

MRC involvement with anti-pneumococcus serum began in 1928, when RA O'Brien, Director of the Wellcome Physiology Research Laboratory (WPRL) wrote a private note to Walter Fletcher (secretary to the MRC), urging the MRC to undertake research in serum treatment for pneumonia. Fletcher's response was not helpful. He suggested that trials in Britain should be delayed until better American evidence was available.[521] With better results forthcoming,[522] and increasing pressure on him from other sources,[523] Fletcher wrote to O'Brien on April 9 1929 asking if serum might be made available for a trial. It was now O'Brien's turn to be cautious.[524] His response to Fletcher questioned the robustness of recent findings, and suggested that not enough cases might be forthcoming for inclusion in a trial.

O'Brien adopted a cautious approach because WRPL had previously invested two years of time and resources in developing a serum for pneumonia that had been ignored by

---

[520] Bullowa 1929**a** p339.

[521] FD1 2367 Fletcher to O'Brien May 23 1928.

[522] Principally, Park 1928.

[523] O'Brien had evidently found other ways to persuade Fletcher. O'Brien told Stanley Davidson, a bacteriologist from the University of Edinburgh, that the WPRL would be willing to produce serum if the MRC were willing to organise a trial. (see FD1 2367 Davidson to Fletcher 10 April 1929).

[524] His response is contained in two letters to Fletcher. The first dated 22 April 1928 and the second dated 26 September 1928, both contained in FD1 2367.

British clinicians, despite the efforts of Walter Topley to arrange clinical trials in Manchester. O'Brien argued that a large comparative trial in mice would be necessary. Results from large-scale trials were needed, O'Brien suggested, to avoid the criticism that results could have arisen by chance through the variability of serum potency. O'Brien also suggested that large trials were needed because mortality from pneumonia was rather lower in Britain than America:

> *'In view of these criticisms the only policy that seems to me to be possible is to adopt what the New York workers consider to be the best process of concentration, make a product, and by using large numbers of mice, try to get an answer to the question "is this material better than the current un-concentrated serum?"'*[525]

O'Brien suggested that the MRC arrange a small conference involving Topley, Francis Fraser, Stanley Griffith[526] and himself to discuss the possibility of a trial. It is not clear from the MRC archives whether or not this meeting took place, but the presumption must be that it did because by 1929 WPRL were attempting to produce concentrated serum.

Efforts to produce serum were not initially successful. On the 24th December 1929, O'Brien told Fletcher that WPRL serum was not yet available.[527] In the meantime however, Murray Lyon at Edinburgh Royal Infirmary had been offered a supply of serum for testing by the American pharmaceutical manufacturer Lederle. Murray Lyon wrote to

[525] FD1 2367 O'Brien to Fletcher, 26 September 1928.

[526] MRC bacteriologist who had previously worked at the Royal Commission on Tuberculosis (1901) research station at Stanstead, on the typing of tubercle bacilli.

[527] FD1 2367 O'Brien to Fletcher 24 December 1929.

Fletcher on 4th October 1929, requesting a grant of £200 from the MRC to cover the cost of serum, which was approved.[528]

Testing began at Edinburgh Royal Infirmary in December 1929, using alternate cases to form study and control groups.[529] Bullowa's clinical trial, begun in 1926 at the Harlem Hospital in New York, is likely to have been a model for Murray Lyons.[530] The use of a quasi-random method of treatment allocation does not appear to have been controversial to the readers of the BMJ. At least one British commentator regarded the use of 'no-selection case control' as vital to the estimation of the value of serum.[531]

However, the use of the alternate case method of allocation appeared to Green to be wasteful. In view of the high cost, Green suggested that treatment should be given only to selected patients in a serious condition.[532] To which Murray Lyon replied:

> *The suggestion that the remaining serum should be used in serious cases is noted - this will economise material, but may give less conclusive evidence*[533]

Green's preference at this point looks odd, because the MRC had sufficient stocks of serum to start discussion of a clinical trial with St Bartholomew's.[534] One interpretation is that Green was using concerns about price as a way of trying to direct Murray Lyons

[528] FD1 2367 Murray Lyon to Fletcher 4 October 1929. Landsborough Thomson placed an order on November 29th. In view of the MRC's standing, Lederle supplied a quantity without charge and sold the rest on favourable terms.

[529] FD1 2367 Murray Lyons to Green 22 January 1930.

[530] Bullowa was well known to the circle of pneumonia researchers in Scotland. Correspondence between Davidson (who undertook serum research at Aberdeen, see below) and Green in 1930 concerning the progress of the Edinburgh trial includes a copy of a letter from Bullowa to Davidson in which their personal friendship is clear. (FD1 2368).

[531] The phrase comes from JG Langley's paper summarised in BMJ 1931;ii:1139.

[532] FD1 2367 Green to Murray Lyon 25 February 1930.

[533] FD1 2368 Murray Lyons to Green 5 March 1930.

[534] FD1 2367 Note of January 30 1930.

towards what he regarded as a more useful methodology. The usual approach adopted by the MRC at this time, as discussed in the previous chapter, was the limited case series form of clinical trial, involving carefully selected patients thought likely to benefit. Even later, when the Wellcome sera was available, TR Elliott remained concerned about the use of non-selection.[535]

Events moved on, and in May 1930 it became clear that WPRL had succeeded in producing a stock of concentrated pneumonia serum.[536] Wellcome serum was available for testing at three centres in late 1930: Edinburgh,[537] St Bartholomew's, and Aberdeen.[538]

During 1931, responsibility for the trials of concentrated antipneumococcal serum passed, at Fletcher's instigation, to the newly formed TTC. At the first meeting, 8th July 1931, it was agreed to collate results from the trial centres and convene a conference to ensure that the results were comparable. The conference took place on 5th October 1931.[539] The main outcomes were: an agreement that more data were necessary; an attempt to include a research group in Glasgow; and the preparation of a standard scheme of inquiry.[540] A copy of the schedule is reproduced in Appendix 7. It shows that the MRC was now

[535] FD1 2368 Elliott to Green 27 November 1930.

[536] See correspondence between Landsborough Thomson and O'Brien in FD1 2368 May 1930. O'Brien was still concerned that the MRC should commit itself to further trials before WPRL scaled-up its production of serum.

[537] Despite Green's concerns about the shortage of serum, the supply of Lederle serum appears to have been more than sufficient for the trialists' needs. On 26 February 1931 Davidson reported to Green that he was still using Lederle serum, despite stocks of WPRL serum being available. (FD1 2368 Davidson to Green 26 February 1931).

[538] The trial at Aberdeen was led by Stanley Davidson, newly appointed Professor of Medicine. This appointment was part of the effort to spread the concept of academic medicine beyond London.

[539] FD1 2499 Green to Dawson of Penn 20th October 1931. After learning about the serum trials at his first Council meeting, Dawson, as President of the Royal College of Physicians, tried to instigate an RCP led trial, much to the annoyance of Fletcher. See FD1 2310 Fletcher to Tidy February 6th 1932.

[540] FD1 2370 Medical Research Council. Standard Scheme of Inquiry. 9/11/31 Reproduced at Appendix 7.

committed to an alternate scheme of patient allocation. Why had the position changed? The records offer no answer. In view of the fact that this was a multi-centre trial by virtue of events rather than planning, the decision to enforce alternate controls may have seemed to the MRC to be part of the overall process of establishing control over the research groups.[541]

In 1933 the MRC arranged a further conference on pneumonia serum.[542] Results from the three trial centres were called in before the meeting. When compared, the results were variable. Those at Aberdeen showed the influence of age on outcome.[543] Even taking this into account, the Aberdeen results were so favourable to serum that Elliott, who was drafting the paper, required further information on the severity of cases in each arm in order to confirm their accuracy.[544] Davidson's response, contained in correspondence to Landsborough Thompson and Green, indicated that the case severity in treatment and control groups was comparable. He considered it possible that the results were due to the play of chance in a small sample.[545]

---

[541] The decision to alternate was not welcomed by all the researchers. The Glasgow group of researchers, led by the bacteriologist Robert Cruikshank, remained outside the framework established by the MRC. Cruikshank maintained that the alternate method of patient allocation to study and control group could give misleading results. His views are set out in the paper describing the Glasgow experience with serum (Cowan 1932) and in correspondence with Green. (FD1 2373 Cruikshank to Green 15th July 1931) Matthews has suggested that bacteriologists tended to be skeptical about the value of the statistical design of experiments. They equated statistics with naïve empiricism, and thus with an unscientific approach (Matthews 1992 229-232). By 1933 however, Cruikshank was keen to have the Glasgow data pooled with that from Edinburgh and London. The proposal is contained in a letter from Cruikshank to Green, accepting Green's invitation for the Glasgow group to be included in the 1933 conference. (FD1 2372 Cruikshank to Green 3rd October 1933) The summary of British experience with antipneumococcal serum published in the Lancet includes Glasgow data. (Medical Research Council 1934).

[542] FD1 2372 Elliott to Green 4 October 1933. The conference took place on 10th November 1933.

[543] FD1 2372 Green to Elliott 27 October 1933.

[544] The request for further information was made after the conference. See FD1 2372 Green to Davidson 15 November 1933, where he described the results as 'the miracle in Aberdeen'.

[545] Davidson to Landsborough Thompson 24/11/33; Davidson to Green 28/11/33.

The Edinburgh data were also strongly in favour of serum, at least in Type I cases, in which mortality was 8.4% in cases treated with serum, compared to 20% in cases treated without. Again, Elliot felt he needed more information before he could accept the results.

The results of the trial were published in The Lancet in February 1934.[546] The report summarised the trials at each centre, including Glasgow, combining the data where it could. The combined result confirmed the widely held belief that patients under 40 were more likely to survive an episode of pneumonia whether or not they were treated with serum. The principal finding concerning the effectiveness of serum was that it reduced mortality in the 20-40 age group from 11.2% to 5.7% in Type I cases, and from 22.7% to 12.6% in Type II cases. The report also highlighted differences between the centres. Changes in mortality, based on Table III of the report, are shown in Table 9, the negative percentages being increases in mortality associated with serum treatment.

[546] Medical Research Council 1934.

Table 9: Effect of serum treatment on fatality in Type I pneumonia. Results from 4 centres in the MRC pneumonia antiserum trial

		Control		Serum				
	Age	**Cases**	**Deaths**	**Cases**	**Deaths**	**C mort.**	**S mort.**	**Change**
Aberdeen	20-40	22	3	25	1	13.64%	4.00%	9.64%
	40-60	13	6	10	0	46.15%	0.00%	46.15%
London	20-40	47	2	58	5	4.26%	8.62%	-4.37%
	40-60	23	6	14	6	26.09%	42.86%	-16.77%
Edinburgh	20-40	34	5	19	1	14.71%	5.26%	9.44%
	40-60	9	2	5	1	22.22%	20.00%	2.22%
Glasgow	20-40	121	15	38	1	12.40%	2.63%	9.77%
	40-60	32	6	15	3	18.75%	20.00%	-1.25%

Source Medical Research Council 1934 p293

The results of serum treatment in Type I pneumonia varied from a 46% decrease in fatality among 40-60 year olds in Aberdeen, to a 17% increase in fatality among 40-60 year olds in London. The paper explained the differences as being due to the chance inclusion of many more severe cases in the treatment arm at the London centre.

Austin Bradford Hill, at the time a staff member of Greenwood's Statistical Department at the London School of Hygiene and Tropical Medicine, commented on the adequacy of the statistical data in 1933. However, the papers at the Public Record Office no longer contain Bradford Hill's views on the results with serum.[547] His comments are therefore re-constructed from the accounts of Bradford Hill's given by Lock[548] and Austoker and

[547] Austoker, evidently using the files before they were transferred to the PRO, found the report in file 1487/VI. This is now listed as FD1 2372 but no longer contains any papers by Bradford Hill.

[548] Lock 1994.

Bryder.[549] According to Lock, Bradford Hill criticised the method of allocation as insufficiently robust:

> *'He showed, for example, that in the pneumonia trials there were two groups of patients, one of people aged between 20 and 39, and the other aged 40 to 60. Roughly 35% of the controls were aged 40 to 60 as opposed to 24% of the patients'.*[550]

Lock's point is that overloading the control group with older people would bias the result in favour of treatment, since it is known that younger patients have a better prognosis. Table 10 shows the number of serum cases and controls from all centres, calculated from the data in Table III of the Lancet report. The figures do not correspond exactly to those quoted by Lock. The probable reason is found in file FD1 2372 which contains a number of revised submissions from the centres, making it difficult to know how many centres, and which version of data Bradford Hill was using.

Table 10: age distribution in the serum treatment for pneumonia trial

Age Group	Number of serum Cases (%)	Number of controls (%)
20-40	251 (26)	418 (44)
40-60	97 (10)	188 (20)
Total	348	606

Source: Medical Research Council 1934

The re-calculated figures in Table 10 support Bradford Hill's critique. However, it was generally accepted that younger patients did form a different population to older patients when the prognosis for pneumonia was considered. The criticism that the groups were constructed separately in the trials therefore lacks force, since a separate analysis of each age group was needed to provide clinically useful information. The Lancet paper, while

[549] Austoker and Bryder 1989**b**.

[550] Lock 1994 p84.

acknowledging Bradford Hill's critique, was able to produce results comparing the death rates in the two groups separately:

> *It is unfortunate, from the point of view of statistical analysis, that the number of patients over the age of forty who were available for treatment was so much smaller than the number of younger patients. Nevertheless on this evidence* [i.e. the trial results] *it would seem that the life saving effects of serum are mainly restricted to the ages at which natural resistance is ordinarily high'*[551]

Austoker and Bryder quote Bradford Hill as criticising the trials for not using a rigorous system of random allocation.[552] This would probably have given a more balanced assignment across the ages, and removed any suspicion that the Aberdeen trialists were selecting patients with a good prognosis for the study group.

Impact of Bradford Hill's critique

However, judging by the trials organised by the TTC, the MRC appears to have paid little if any attention to Bradford Hill's proposals in 1933. If the TTC sought to control the conditions under which clinical trials were carried out, it did so by involving a network of researchers whose credentials as clinical scientists were known and who could be trusted to test therapeutic substances with only a minimum of central guidance.

The failure of Bradford Hill's criticism to have impact in 1933 can be ascribed to two factors. Firstly, given its disposition towards a physiological model of therapeutic trial it was not clear how or why the MRC should insist on the adoption of a formal scheme of randomisation by researchers. Researchers such as Bullowa and Davidson were if anything in advance of statistical theory when they adopted an alternating method of

[551] Medical Research Council 1934 p292.

[552] Austoker and Bryder 1989**b** p47.

selection to study and control groups.[553] In the early 1930s, RA Fisher was almost alone in promoting the virtues of randomisation.[554] Fisher did not fully set out methods for randomisation until 1935, with the publication of *The Design of Experiments*.[555]

Secondly, the position of statistical advice within the MRC lacked strength. At this time the majority of the work of the Statistical Department at the London School of Tropical Medicine and Hygiene was oriented towards applied descriptive epidemiology, for example the study of morbidity in the printing trades, carried out by Bradford Hill in 1927/28.[556] The Department was understaffed, and lacked medically qualified members.[557] Greenwood's view was that the advice given by the Department was unappreciated by the committees that received it:

[553] Even so, they were not the first to use alternation to form study and control groups. In 1910 Freeman used this method to study the efficacy of a whooping cough vaccine (reported in Matthews 1992 p231).

[554] See Hacking 1988 p429. Hacking notes that in 1932 Fisher was unable to secure an examiner for a thesis on randomisation by one of his students. Nonetheless, under certain circumstances the MRC accepted the value of random sampling from an early stage of its existence. In 1925, a request from the Eugenics Education Society for funds to carry out a large scale anthropometric survey was declined following guidance from the newly formed Statistical Committee, partly because the proposal did not suggest a random sampling method. By this time the vogue for eugenics in Britain was somewhat in decline, so the lack of adequate methods would have been a convenient way to sidestep the proposal. See FD1 7107 Application and correspondence May 1925.

[555] Fisher 1935.

[556] Referred to in FD1 7108 Report on the work of the Statistical Committee for the year 1927-28. Greenwood described the aims of the Department at the time as being the training of 'those who are endeavouring to acquire statistical method as a workable instrument in epidemiological and public health research work'. (FD1 7108 Memorandum on the present position and prospects of medical statistics February 1928 p4).

[557] See FD1 7108 Memorandum on the present position and prospects of medical statistics February 1928. At that time Greenwood regarded the key staff as himself, Elisabeth Newbold, and L Isserlis. 'It is certain that for some years to come the responsibility for this work must be on the shoulders of three people, Isserlis, Miss Newbold, and me, all of us over 40, only one of us a medical man. Our department is much stronger mathematically than medically' (FD1 7108 Memorandum p1).

> *This afternoon from 4 to 6 I spent on a committee at the MRC office… yesterday I was from 4.15 to 6.40 in the same place for a similar purpose… the effect of both those committees has been to save the Council from some fruitless expenditure… But it is not really pleasant to me to be regarded as a sort of devil's advocate steadily engaged in preventing young investigators having a show*[558]

To summarise, as of 1933 there were neither technical nor practical reasons for the MRC to consider formal schemes of randomisation as necessary to the design of clinical trials being undertaken by the TTC. The pneumonia anti-serum trial was inherited by the TTC. In all likelihood the problems associated with the trial did not suggest themselves as inherent to the trials authorised by the TTC. The solution offered by Bradford Hill did not therefore seem to apply to the TTC's programme. Even if randomisation had seemed necessary, the MRC had no way of controlling the methodological activities of researchers.

The next section introduces Austin (later Sir Austin) Bradford Hill. In particular it sets out Bradford Hill's attempts to make statistics more relevant within the MRC following the pneumonia anti-serum trial.

Austin Bradford Hill

Austin Bradford Hill, who made the comments about the statistics gathered by the pneumonia antiserum researchers, was to become the most eminent medical statistician of the twentieth century.[559] However, primary sources on Bradford Hill are meagre. He appears to have left no papers,[560] and two of the three named files relating to him at the

[558] FD1 7108 Memorandum on the present position and prospects of medical statistics February 1928.

[559] Sir Austin Bradford Hill 1897 – 1991. For his obituary see Doll 1993.

[560] The Cataloguing Unit of the Archive of Modern British Scientists at the University of Bath has no details of any papers left by Hill.

Public Record Office are closed until 2003.[561] The third PRO file consists principally of a draft of his paper on the philosophy of the clinical trial.[562] Finally, any papers left after his retirement as Director of the Statistical Research Unit have been lost as the Unit moved premises several times before reaching its current location as the MRC Biostatistics Unit at the University of Cambridge.[563]

Austin Bradford Hill was the third son of Leonard Hill, Director of the NIMR Department of Applied Physiology and member of the MRC Statistical Committee. Bradford Hill originally intended to train as a doctor, but at the age of nineteen contracted pulmonary tuberculosis. Having recovered, he studied economics at the suggestion of Major Greenwood, a friend of the Hill family.[564] Following in Greenwood's footsteps, Bradford Hill later attended Karl Pearson's statistics course at University College, London.[565]

Bradford Hill's first connection with the MRC was the receipt of a grant to study the epidemiology of tuberculosis in Essex, obtained with the assistance of Greenwood. He subsequently joined the staff of the Industrial Health Research Board, and then the Statistics Department at Hampstead. In 1933 he transferred to the London School of Hygiene and Tropical Medicine (LSHTM), where Greenwood was Professor. In 1945 he succeeded Major Greenwood as Professor of Medical Statistics and Director of the MRC's Statistical Research Unit. In that year also he became vice-chair of the Statistics

[561] Files FD1 2015 and FD1 2016.

[562] FD1 5081.

[563] Correspondence with Peter Armitage and Ian Sutherland suggests that many of the papers relating to the SRU were destroyed when it moved from London to Cambridge.

[564] Greenwood's first substantive post was as a demonstrator in Leonard Hill's physiological laboratory at the London Hospital Medical School (Hogben 1950**b**).

[565] Doll 1992; Doll 1993.

Committee, and chairman in 1948, until the Committee's dissolution in 1961. He retired in 1961 and died in 1991, aged 93.

Bradford Hill's work in the 1930s

Bradford Hill's work in the 1930s can be regarded as an attempt to consolidate both his personal position and the position of medical statistics within clinical science. In 1933 Greenwood took him out of MRC employment and appointed him as Reader in Medical Statistics at LSHTM. Bradford Hill's two major lines of work in the subsequent period were the completion of Greenwood's studies on mice, and a series of articles in the Lancet on the principles of medical statistics.

Experimental epidemiology

Experimental infection using colonies of mice was a bridgehead between epidemiology and experimentalism for Greenwood:

> *We have used up thousands of mice and in an experiment designed to test one point only and are using terrible numbers. Mice are dear and laboratory servants are dearer.* ***But this sort of thing must be done or people will chatter about epidemiology until the end of time****....I am sometimes, perhaps when I am tired and bored , tempted to doubt whether modern statistical developments have not done actually harm by tempting people to suppose that with such sharp tools they really can hack a way to truth through human statistics without leaving their studies or biometric laboratories.*[566] [my emphasis]

It was also a bridgehead between epidemiology and bacteriology:

[566] Greenwood to Raymond Pearl, June 9th 1925. Raymond Pearl Papers. Quoted in Matthews 1992 p263.

> *The experimental method which the report describes was devised to solve a particular type of problem that has proved refractory to other forms of study. On the one hand laboratory experiments, performed under strictly controlled conditions, have yielded a mass of information with regard to the response of the individual host to artificial infection, and as to methods by which that response may be modified. On the other hand the data collected by the epidemiologist have taught much in regard to the behaviour of naturally infected herds…The collection of statistical evidence under field conditions, however, is so beset with difficulties that the assessment of the relative importance of the interrelated factors, determining the course of events in infected communities or herds, has so far proved impossible.*[567]

The aim of the studies was to provide empirical data on the course of epidemic infectious diseases in populations, under controlled conditions. The general method was to use batches of 25 mice housed in cylindrical cages. Methods of establishing large herds of infection free mice, and of maintaining the cage environment, were laborious. Approximately 15,000 mice were needed each year, and cages had to be cleaned each day during experiments.

In this way it was possible to observe the course of experimentally introduced infections. The real aim of the work however was to test hypotheses regarding the impact of environmental and other influences on the course of infectious disease. What is the effect of regularly introducing small groups of uninfected mice into an infected herd? What is the effect of vaccination on herd mortality? How does diet affect the spread of infection?[568]

[567] Greenwood Bradford Hill and Topley 1936 Preface.

[568] The results of the mice experiments are summarised in Section XI, p 193-204 of Greenwood 1936.

Methodology in experimental epidemiology

It is possible that some form of random allocation of mice was used in the experiments, but the text does not say so. The following methodological statement is typical of the whole:

> *In this experiment 20 mice were taken, and fed on a mixture of a 24 hours' broth culture of Bact aertrycke [typhoid] and an active phage filtrate; 20 control mice were fed on the culture alone.*[569]

The introduction to the report provides the clearest statement on the need for random samples, while suggesting that a formal method of random allocation was not considered necessary:

> *An appropriate number of mice is taken and injected with a constant dose of a suspension of the organism under study. These mice are divided into batches of 25, and to each batch are added 100 normal mice…The method of interference under study – a special diet … and so on – is applied to one or more of the remaining herds… In experiments of this type the 100 mice at risk in each group are exposed, at the same moment, to contact with the same number of bacteria, and it would seem reasonable to assume that, provided the mice themselves are random samples and all other factors influencing the spread of infection are distributed between the groups in a purely random fashion, the mice in each group will behave in the same way, within the limits covered by random sampling errors. But the condition of randomness may not be fulfilled, and it can certainly not be assumed.*[570] [My emphasis]

The questions posed by Greenwood and colleagues were clearly aimed at answering questions relevant to the MRC using epidemiological and statistical methods in a laboratory setting. Bradford Hill's first research had been an attempt to resolve a question of excess mortality in the population of Essex, and to measure the impact of diet on

[569] Greenwood Bradford Hill and Topley 1936 p180.

[570] Greenwood Bradford Hill and Topley 1936 p21-22.

morbidity.[571] The laboratory work replicated surveys and natural experiments occurring in the field with those in controlled laboratory experiments. While the transfer of epidemiology to the laboratory improved upon the precision of field studies, laboratory work suffered the obvious shortcoming that mice in experimental conditions did not resemble human populations. Nor were the results of mice experiments conclusive. For example, the experiments on the effect of dietary variation on susceptibility and resistance to typhoid differed from previous results, but allowed no definitive conclusions about the role of diet.[572]

The Principles of Medical Statistics

The need for improved teaching in medical statistics was set out by Greenwood in 1927.[573] Bradford Hill's introduction to medical statistics *The Principles of Medical Statistics* was his greatest achievement of the 1930s.[574] First published as a series of articles in the Lancet and later as a book, *The Principles* was an attempt to recruit clinicians to a statistical perspective:

> *The worker in medical problems, in the field of clinical as well as preventive medicine, must himself know something of statistical technique, both in experimental arrangements and in the interpretation of figures. To enable him to acquire some knowledge of this technique I have tried to set down as simply as possible the statistical methods that experience has shown me to be most helpful in the problems with which medical workers are concerned.*[575]

[571] He concluded that the cause was probably due to selective emigration of the fittest to towns. Bradford Hill 1925**a** and 1925**b**.

[572] Greenwood Bradford Hill and Topley 1936 p178.

[573] FD1 7108 Instructions in medical statistics in England, by Dr M Greenwood. Undated manuscript, stamped 30 May 1927.

[574] Bradford Hill 1937.

[575] Bradford Hill 1937 preface.

In discussing the content of *The Principles,* I will emphasise the methods it employed to recruit readers as much as the contents themselves.[576] Bradford Hill's general approach is to use the facts of medicine to show the reader that medical knowledge is in essence numerical: 'The clinical assessment, or the clinical impression, must itself be numerical in the long run'.[577] Even when considering an individual, the knowledge used by a clinician is constituted from the 'population' of his or her experiences regarding the condition. To those who argue that each patient is different, Bradford Hill rejoins:

> *...yet if each patient is unique it is difficult to see how any basis for treatment can be sought in the past observations of other patients... in fact, of course, physicians must, and do, base their "treatment of choice" upon what they have seen before.*[578]

In their use of past experience doctors know that patients are variable. As a result, medicine is full of conflict, which statistics may help resolve. 'Far therefore, from arguing that the statistical approach is impossible in the face of human variability, we must realise that it is often essential.'[579] [my emphasis]

Having secured a sense that medicine inevitably concerns group effects, Bradford Hill describes the methods of dealing with populations. They include ways of obtaining samples and of describing them, ways of estimating whether or not two samples come from the same population, and whether or not a sample contains the expected value of a variable.

---

[576] His best known sleight of hand was the suggestion that the purpose of randomisation was to ensure that the treatment and control groups are initially equivalent. (Bradford Hill 1971 p255).

[577] Bradford Hill 1937 p266.

[578] Bradford Hill 1937 p6.

[579] Bradford Hill 1937 p6.

Each topic is contained in a short chapter, using examples drawn directly from clinical practice. At a deep level, Bradford Hill is suggesting that medical practice and medical statistics are concerned with the same thing – collective experience. The role of the statistician is to prevent the clinician from introducing error into the assessment of collective experience:

> *It seems that many people are not capable of using common sense in the handling and interpretation of numerical data until they have been instructed in quite elementary ideas and techniques. Mistakes which when pointed out look extremely foolish are quite frequently made by intelligent persons, ... there is often lacking what has been called a "statistical tact, which is rather more than good sense." That tact the majority of persons must acquire (with a minority it is undoubtedly innate) by a study of the basic principles of statistical thought and method.*[580]

In later editions, Chapter 20 of *The Principles of Medical Statistics* discusses the clinical trial. Chapter 20 first appeared as an article in the British Medical Bulletin in 1951,[581] and was incorporated into subsequent editions of *The Principles*. It is clear though that Bradford Hill's views on clinical trials in 1951 were based on several sources available to him in the 1930s. It seems reasonable to assume that success in 1948 encouraged Bradford Hill, and that the published views of 1951 were those he held in the mid 1930s.

Given Greenwood's considerable influence on Bradford Hill, the principal sources for Bradford Hill's views on clinical trials are most likely to be those of Greenwood and Yule, published in 1915,[582] and Greenwood's study *The Medical Dictator*, published in

[580] Bradford Hill 1937 p2.

[581] Bradford Hill 1951.

[582] Greenwood and Yule 1915. A frontal attack by two statisticians on the statistical approach used by the clinician Sir Almroth Wright. Pocock 1983 considers that this may have been the first paper to suggest the use of a randomised design in medicine. This overlooks Karl Pearson's suggestion of 1904 (see next chapter).

1936.[583] He also had available Fisher's *The Design of Experiments*,[584] articles by Mainland and Sutliffe,[585] along with his experience with trials like that of anti-pneumococcus serum in the early 1930s.

The criticism of alternating schemes of randomisation

Reading *The Principles* it becomes apparent that Bradford Hill sees with unusual clarity the methodological requirements necessary to the production of definitive results: 'The first step in the controlled trial is to decide precisely what it sets out to prove'.[586] In terms of the clinical trial the first step is to specify clearly the treatment and the group of patients who will be treated. Having decided what the treatment is, the first practical step is to define the criteria by which patients are admitted into the study. For example, one might decide to select for a particular trial only adults of either sex between the age of 40 and 60 with a raised blood pressure.

This seems to us a small and obvious step to take, but the disbenefits of prejudging which types of patients were to be studied would have figured more largely at the time. To do so means that the extent of clinical judgement is reduced to identifying whether or not a patient belongs to a pre-defined category. In requiring a pre-defined population, Bradford Hill is achieving two ends. Firstly, he is translating Fisher's requirement for a hypothetically infinite population into clinical terms. Secondly, he is placing statistical

[583] A study which is notable for attempting to revive the fortunes of PCA Louis. 'If only Louis had succeeded in really commanding the support of … *les superbes* as a contemporary called them, the great clinical teachers of Paris, … why we should have had to do something about it in England! I dare say that by now the Royal Colleges would be considering the desirability of establishing a diploma in clinical statistics… but this is mere day dreaming'. Greenwood 1936 p141.

[584] Fisher 1935.

[585] Mainland 1936, Sutliffe 1936.

[586] Bradford Hill 1971 p253.

concepts at the centre of trial design. If clinical medicine is to benefit, it must allow its research to be cast in terms of populations that are at once clinical and statistical:

> *It would of course be possible deliberately to incorporate more and different groups in a trial, but to start out without thought and with all and sundry included, with the hope that the results can somehow be sorted out statistically at the end, is to court disaster'*[587]

Once a population has been established, 'The next step in the setting up of the trial is the allocation of the specifically defined patients to be included in the treatment and non-treatment groups.'[588]

Fisher's criterion for judging whether or not a sample is random was that the samples should contain no recognisable characteristic. That criterion can be realised in the clinical trial only if the clinician does not allocate patients to trial and control group. In the anti-pneumococcus trial patient allocation was usually based on the sequence of admission. If strictly followed this method would meet Fisher's criterion.

> *Such a method may, however, be insufficiently random if the admission or non-admission of a case to the trial turns upon a difficult assessment of the patient and if the clinician involved knows whether the patient, if accepted, will pass to the treatment or control group. By such knowledge he may be biased, consciously or unconsciously, in his acceptance or rejection; or through fear of being biased, his judgement may be influenced*[589] [my emphasis]

The role of the statistician is to free the clinician from the responsibility of allocation. With the help of a table of random numbers, the statistician might prepare a series of

[587] Bradford Hill 1971 p254.

[588] Bradford Hill 1971 p254.

[589] Bradford Hill 1971 p255.

numbered envelopes each containing the direction either to the treatment of the control group.

> *After each patient has been brought into the trial the appropriately numbered envelope is opened and the group to which the patient is to go, treatment [T] or control [C], is given upon the slip inside.*[590]

It is here that Bradford Hill is at his most ambiguous. He knows that the purpose of randomisation, according to Fisher, is to ensure the validity of a parametric statistical test, but he presents randomisation as a way of freeing the clinician from the responsibility for constructing equivalent groups. The 'three great advantages' of randomisation claimed by Bradford Hill are entirely psychological:

> *(1) it ensures that our personal feelings… have not played a part…(2) it removes the very real danger that … believing our judgements may be biased, we endeavour to allow for that bias and in so doing may "lean over backwards" (3) having used such a random allocation we cannot be accused by critics of having set up personally biased groups for comparison.*[591]

The beauty of Bradford Hill's description of the design of clinical trials is the way in which it solves several problems simultaneously. It offers clinicians the promise of precise knowledge, provided they allow statisticians to help them. It translates Fisher's pioneering work into medical scenarios. It also suggests that the justification for human experimentation can be found within the precepts of trial design. For if one can circumscribe one's uncertainty by identifying those groups for whom one does not know whether or not a drug will work, then one *should* seek to find out whether or not the drug

[590] Bradford Hill 1971 p255.

[591] Bradford Hill 1971 p257.

works, and one *can* do so by entering these patients into a trial and allocating them at random to study and control groups.

Conclusion: the outcome of The Principles of Medical Statistics

In Chapter three I argued that the MRC had no reason to adopt statistical design in its clinical trials in the 1930s. Following the publishing success of *The Principles*, Bradford Hill was invited to join the TTC in 1938. In the short time before its dissolution he was unable to influence the work of the committee. During the Second World War he worked for the Royal Air Force. In 1945 he succeeded Greenwood as Professor of Medical Statistics at the London School, and also became Vice-chairman of the Statistics Committee.

In 1946 Bradford Hill was asked to provide statistical advice for an important MRC trial of a new treatment for tuberculosis called streptomycin, a chemotherapeutic agent first extracted from soil bacteria by American scientists in 1944. There were several parallels between streptomycin and anti-pneumococcus serum including its American origin, high price, and variability in the course of the disease treated. The next chapter describes the streptomycin trial and considers why the MRC was happy to introduce all the elements of the RCT.

Chapter five

Tuberculin, streptomycin, and the first published randomised controlled trial in medicine

Introduction

The purpose of this chapter is to discuss the clinical trial of streptomycin organised by the MRC in 1946.[592] The trial is widely regarded as the first clinical trial to fully meet the requirements of sound methodological design.[593] Whether or not this approbation is justified is open to question.[594] Nevertheless, when compared to the clinical trials organised by the TTC in the 1930s, the streptomycin trial clearly marks the beginning of a new phase in the MRC's approach to the design of clinical trials. Set against its earlier efforts, the streptomycin trial is distinctive and important in the context of the MRC and clinical trial design in Britain.

Nevertheless, as a trial of a therapy for tuberculosis, the streptomycin trial had precursors, which were known to the streptomycin committee.[595] The controlled trial of sanocrysin, in which subjects were paired and then allocated to treatment of control by the flip of a

[592] The trial is well documented. Papers occupying 10 files at the Public Record Office. The papers include the trial schedule, minutes of the organising committee, extensive correspondence relating to the procurement of streptomycin from America, and also a large number of enquiries concerning the availability of streptomycin for clinical use from clinicians and members of the public. I have attempted to locate the completed patient schedules from the study, but it appears that they have not been retained. The most likely explanation is that they were discarded when the Statistical Research Unit moved from London to Cambridge in the 1960s. Thanks are due to Dr Peter Armitage and Dr Ian Sutherland for the information they gave.

[593] Probably the strongest expression of this sentiment is Lock 1994, who argues that Bradford Hill deserved a Nobel prize for the introduction of randomised controlled trials into medicine.

[594] Like any priority claim, the situation is a good deal more complicated than any one perspective will allow. Reviewing the history of the clinical trial in 1982, the American epidemiologist-historian Abraham Lilienfeld emphasised the earlier American contribution. (Lilienfeld 1982). Oddly enough, the history of clinical trials written in the early 1950s by the MRC scientist John Prince Bull makes no mention of the streptomycin trial (Bull 1951 and Bull 1959).

[595] D'Arcy Hart 1946 and 1991.

coin, is well known.[596] Here I discuss clinical trials of tuberculin, which have been overlooked because tuberculin is regarded almost as a quack remedy.[597] However as will be shown, trials of tuberculin were probably more influential in the British context than sanocrysin.

This chapter attempts to do two things. Firstly, to highlight the methodological issues associated with trials of tuberculin. Secondly, by accounting for the origins and conduct of the streptomycin trial, to highlight the extent to which the trial design was innovative and discuss the role of the trial committee's statistician Austin Bradford Hill.

Tuberculosis

Although it is the world's largest killer, tuberculosis is now somewhat isolated from present day consciousness. Yet one hundred years ago it was part of the experience of thousands of families in Britain, and its course well known to many more.[598] The reason was that tuberculosis was the single largest cause of death. In cities the death rate might be from 200-400 deaths per 100,000 resident population per year.[599] Nationally, according to Worboys, the annual death toll from TB around the turn of the century was

[596] Amberson 1931. Referred to by most histories of the clinical trial.

[597] Citing FB Smith, one of the few authors to give any serious consideration to tuberculin, but whose account contains some errors, Lock writes: 'Until after World War Two, the history of tuberculosis was littered with abandoned false hopes. Just to take the most notable examples, in the 1890s there had been the fiasco of Ehrlich's tuberculin treatment…' (Lock 1994 p83) The MRC's official history is equally disparaging:. Referring to Camac Wilkinson: 'there was for example a vaccine introduced by a citizen of a medically advanced foreign country, in which he was evidently a prophet without honour' (Landsborough Thomson 1975) A recent, dismissive view is that of Leibowitz 1993, which against all evidence suggests that Tuberculin was little used after 1893.

[598] The principal studies of tuberculosis in Britain are Smith 1988, an interpretation of the decline of tuberculosis, and Bryder 1989, which discusses the anti-tuberculosis campaign in England between 1919 and 1942. In terms of publishing, interest in the history of tuberculosis in America has been prominent in recent years. For a representative selection, see: Ott 1996; Rothman 1995; Rosenkrantz 1994 (a compendium of extracts); Ellison 1994.

[599] Rates and other indices give a sense of the impact of tuberculosis. The historical epidemiology is well represented in Wilson 1990, and Bryder 1988.

50,000 in Britain.[600] Greenwood calculated that between 1848 and 1873 190,000 Englishmen between the ages of 20 and 25 died, of which 85,000 died of TB.[601] The scourge of tuberculosis persisted into the twentieth century. Introducing the 1911 National Insurance Bill to Parliament, Lloyd George claimed that one in every three among males aged between 15 and 55 who died did so from tuberculosis.[602]

A diagnosis of tuberculosis did not always mean a death sentence however. Nor was its course easily predictable. Surveys found tuberculosis to be highly prevalent amongst every population investigated. In 1900 the National Association for the Prevention of Tuberculosis[603] estimated that for every death from tuberculosis there were 5 sufferers, a factor that was increased to 10 in 1910. The calculations are necessarily inexact, but from the NAPT figures it may be estimated that in 1900 there were some 250,000 people with tuberculosis, and more who were pre-symptomatic. In 1913, school inspectors reported 15.4% of elementary school children to show obvious signs of tuberculosis.

Just as clinical tuberculosis did not always lead to death, so infection with *M tuberculosis* did not always lead to clinical disease. In Gloucestershire in 1931 50% of 563 children dying from all causes were found to have tuberculous lesions.[604] The results of the National Tuberculin Survey, conducted in 1949-50 by the MRC in 22 areas of England and Wales showed that by the age of 20 roughly 70% of the population showed signs of

[600] Worboys 1992**a** p48.

[601] Greenwood 1943 p19.

[602] Hansard 5th series 1911, xxv 626 quoted in Bryder 1988 p 37.

[603] Founded in 1898. For a discussion see Bryder 1988.

[604] These and other figures are found in chapter 1 of Smith 1988.

infection with *Mycobacterium tuberculosis*.[605] Present day estimates suggest that fewer than 10% of people infected will go on to develop clinical tuberculosis.[606] Against a generally positive picture of attempts to manage tuberculosis, evidence from health agencies suggests that tuberculosis remains a major threat to health.[607] Firstly, on a world scale tuberculosis remains the largest killer. While most deaths occur outside of the developed West, the decline in tuberculosis mortality has been halted in recent years in all the most developed countries. Secondly, in recent years there has been a growth in the number of multi-drug resistant strains of *M tuberculosis*.

Until the early twentieth century tuberculosis of the lungs was also known as consumption, or phthisis. Its general symptoms, described by Smith,[608] include lassitude, irregular appetite, flatulence, loss of weight, irritability, raised and unstable pulse rate, night sweats, facial pallor contrasted with flushed cheeks and wan eyes, emaciation, female amenhorrhoea and male impotence, running nose, frequent colds, harsh coughing, frequent spitting of foul sputum, in addition to the most dramatic sign of inner disease, that of haemoptysis – coughing up blood. The one common indication was fever in the late afternoon or night. When consumption was finally diagnosed it might advance and kill within weeks, or arrest, or spontaneously disappear.

The treatment of tuberculosis

Forms of treatment in the nineteenth century were numerous. Chief among them was cod liver oil, of which 1,500 gallons was consumed annually at the Hospital for Consumption,

[605] Daniels 1952 p347-348.

[606] For a review see Sutherland 1976.

[607] Klaudt 1994.

[608] Smith 1988 p2.

Brompton in the 1880s.[609] The efficacy of cod-liver oil was widely acknowledged at the time, and its superiority to other forms of treatment was on record. For example, in 1858 Richard Payne Cotton conducted a trial of glycerine at the Brompton:

> *'With a view to fairly testing its effects, I prescribed it, at the Consumption Hospital, for twenty-three of the inpatients in various stages of the disease… In only five instances did it seem to be of service, and even in these the improvement was but moderate, and might have been equally due to other causes, such as diet etc; and two of the patients afterwards progressed at a greatly increased rate under cod-liver oil.* [610]

Aside from sedatives and tonics, Payne Cotton commended fresh air, exercise and diet. Elsewhere, Latham noted the use of alkalis, bromides, morphia, iron, hypophosphites and calcium chloride.[611] At the end of the century, Ransome divided remedies into three types: specific remedies such as Tuberculin; anti-bacterial remedies such as creosote and guaicol; and constitutional remedies such as arsenic and iron.[612]

Institutionalising tuberculosis treatment

Although Villemin demonstrated the experimental transmission around 1865, tuberculosis was regarded as a hereditary rather than an infectious disease. Among sanitarians the role of environmental factors was pressed, especially around the time of the Public Health Act of 1866.[613] Evidence that infection played only a minor role was

---

[609] Williams 1887 p385.

[610] Payne Cotton 1858 p280.

[611] .Latham 1903 Chapter VII Quack and secret remedies such as Spahlinger treatment and Umckaloba had their enthusiastic advocates in the 1920s and 30s.

[612] Ransome 1896 chapter 10.

[613] Simon, in his Report to the Privy Council of 1867 noted the association between urban sanitary improvement and decline in tuberculosis. He regarded improvements which involved drying of the soil as particularly important. (Cited in Williams 1887 p81-82). FB Smith has suggested that tuberculosis attracted little interest among early public health reformers because of the widely held view that its origins were hereditary (Smith FB 1988 p37-9). Unlike his sanitarian colleagues such as Chadwick, who saw no role for contagious causes of disease, Simon had no difficulty in incorporating the infectious theory of tuberculosis into his thinking. For a discussion of the sanitarian perspective, see Ackerknecht 1948, Cooter 1982.

derived from hospital statistics. It was claimed that cross infection was unknown, and that staff at the Brompton were no more likely to contract tuberculosis than staff at non-specialist hospitals.[614]

However, the idea of transmission by infection found more favour among general practitioners. A collective inquiry instituted by the British Medical Association found that approximately one quarter of respondents knew of cases where transmission was by infection.[615]

The idea that tuberculosis was a communicable disease was beginning to be taken more seriously by the late 1870s. Hardy suggests that the manner in which tuberculosis is perceived as a disease entity was determined to a large extent by the resources one was able to muster.[616] Lacking any specific measure against tuberculosis, Medical Officers of Health continued to practice along sanitarian and hygienic lines after the discovery of *M tuberculosis* in 1882, integrating the newer knowledge concerning microorganisms into their existing lines of work.

The sanatorium – the segregation of tuberculous patients for the good of the patient and community – is a feature of twentieth century therapeutics.[617] Its rapid growth can perhaps be ascribed to the fact that as a medical intervention it did not require firm commitment to any particular theory of the cause of tuberculosis. Segregation appealed to

[614] Williams 1887 p87.

[615] BMA Collective Investigation Committee, cited in Williams 1887 p87.

[616] The examples she cites are firstly the association of tuberculosis with poverty by Medical Officers of Health around 1900 as a way of highlighting the potential of state action among the poor. (Hardy 1993 p262-263) Secondly, the concern about the role of housing conditions as a cause of tuberculosis among Victorians, because housing was regarded as falling within the remit of social intervention whereas a factor such as diet was not. (Hardy 1993 p265).

[617] Open-air treatment can be traced to Brehmer in 1859. Fee paying British patients attended alpine sanatoria in the 1890s.

those who believed in either the primacy of infectious transmission or hereditary causes. Environmentalists could point to the beneficial conditions provided by institutional settings. By 1900 there were 23 private sanatoria in Britain, offering more affordable alternative to German sanatoria. The first public sanatorium was opened in 1900, in Cumbria. Approximately 40 publicly funded sanatoria were opened in the following decade, including Frimley Park, in Hampshire, established as an outstation of the Brompton Hospital.[618] Whilst sanatoria offered a distinctive therapy, it was an expensive option, and most patients with clinical tuberculosis continued to be treated in Poor Law Infirmaries.[619]

Despite the growth in their provision, the effectiveness of sanatorium treatment was never fully accepted. Karl Pearson, in one of his early assaults on doctors' handling of statistical data, criticised one evaluation of sanatoria, and suggested that since tuberculosis was essentially a hereditary disease, it could only be properly treated if priority was given to eugenic measures.[620]

---

[618] For a fuller account, see Worboys 1992a. Also Bryder 1988 Chapter 2.

[619] Arthur Newsholme, in seeking an explanation for the historical decline of tuberculosis, attributed it to the isolation of patients in poor law infirmaries. In 1906 he calculated that between one quarter and one third of all the cases of tuberculosis in Brighton, Sheffield and Salford were treated in a poor law hospital (Newsholme 1906). For further details of Newsholme's campaign against the non-communicability thesis advanced by the physician Timbrell Bulstrode around the turn of the century, see Wilson 1990.

[620] Pearson 1907, quoted in Worboys 1992a. In fact most of the evidence accumulated by doctors showed that sanatoria had limited effectiveness. As early as 1903 Dr Lawrason Brown of the Trudeau Sanatorium in America showed that rest and diet had little effect on survival. (Brown 1903 quoted in Rosenblatt 1973 p171) Bryder 1988 p68-69 cites further critical evaluations of the results of sanatorium treatment. However, there were some positive results. At Frimley, 10 year survival rates improved for sputum positive cases (Brough 1949 p107), though this could be due as much to what happened after discharge as to improvements in sanatorium treatment. For an account of the history of the sanatorium movement in Glasgow, and a critical assessment of its effectiveness from the present perspective, see McFarlane 1989. The failure of sanatoria in McFarlane's account is shown by the lack of impact on incidence rates. While it may be fair to argue, as McFarlane does, that more could have been achieved if resources had been devoted to improving housing rather than sanatoria, he provides a rather weak argument that sanatoria in themselves were ineffective.

Whatever their effectiveness, sanatoria provided an institutional base for the development and assessment of therapy during the first decades of the twentieth century. Looking back over 35 years at the Midhurst Sanatorium, MC Brough of the Chest Service of the Oxford Regional Hospital Board described the development of therapies between 1906 and 1941. These included surgical techniques designed to rest affected lungs and collapse tuberculous cavities (pneumothorax, phrenic paralysis, and thoracoplasty) and drug therapies (tuberculin and sanocrysin or gold therapy).[621]

The concept of long-term settlements for treated consumptives was introduced in Britain in part as a response to concern about the effectiveness of sanatorium treatment. Sanatoria were effective, it was argued, but their value was masked by the return to society of treated sanatoria residents. Village settlements, where ex-sanatorium patients could continue to live apart from general society whilst being rehabilitated, were created in Papworth, Cambridgeshire, Preston Hall, Kent, and Sherwood, Nottingham.[622] Looking to the future, in 1945 Heaf and McDougall saw village settlements being incorporated into the new National Health Service, and the expansion of the scheme to all of the proposed Regions.[623] They did not know that within a decade the idea of village settlements would be gone forever, replaced by hospital and domicillary chemotherapy of tuberculosis.

---

[621] Detailed description of the methods and effectiveness surgical interventions is beyond the scope of this thesis. For more detail of some of these therapies see Rosenblatt 1973. For a British perspective on the effectiveness of collapse therapy after the introduction of streptomycin, see Ellman Johnson and Kagan 1955.

[622] Papworth Village Settlement, founded in 1917 by the Tuberculosis Officer for Cambridgeshire, PJ (later Sir Pendrill) Varrier Jones, with the support of Sims Woodhead, Professor of Pathology at Cambridge University. It consisted of a hospital, sanatorium, and permanent settlement where treated patients could live with their families and work in village industry. By 1938 Papworth Village Settlement had a population of over 100, including patients, their families, and staff, and a turnover in excess of £100,000 per annum. (Bryder 1984) Preston Hall is described in Heaf and McDougall 1945, which contains a bibliography on settlements.

[623] Heaf and McDougall 1945 p132.

Other aspects of the treatment of tuberculosis

As Bryder[624] notes, growth in the provision of sanatoria took place in the context of a general expansion of institutional provision, both charitably and state funded.[625] Whilst the main state provision in the first decade of the twentieth century continued to be the unloved and undervalued poor law infirmary, charitable work developed along novel lines, including tuberculosis dispensaries, which were in effect specialist out-patient clinics, providing treatment, but also health advice and home visits.[626]

The National Insurance Act of 1911 made special provision for tuberculosis.[627] Section 16 provided empowered local authorities to create sanatoria and other institutions for the treatment of tuberculosis. One shilling and 4d was to be provided for each person insured. Of this amount, known as sanatorium benefit, 6d was to be used to pay general practitioners for providing home treatment, and 1d was to set aside for research on tuberculosis. As has been discussed earlier, the 1d amounted to £57,000 and formed the basis of funds used by the MRC on medical research.[628]

---

[624] Bryder 1988.

[625] Major charities established around this time include the National Association for the Prevention of Tuberculosis and the King Edward VII Welsh National Memorial Association, to which the coal owner and Liberal MP David Davies of Llandinam contributed £150,000 on its establishment in 1910. By 1915 the Association provided 1586 beds for the treatment of tuberculosis. (Bryder 1988 p 27).

[626] The work of the dispensary in Street Somerset was described by its medical officer, Hilda Clark, possibly a member of the C & J Clark family, shoe manufacturers, who founded the Street Dispensary in 1904, and simultaneously funded two beds for their workers at the nearby Winsley Sanatorium. As Hilda Clark explained, tuberculin was particularly attractive to dispensaries because it was considered particularly useful in early and uncertain cases of tuberculosis. The routine for tuberculin injections was also favourable. It required regular but occasional attendance, and no in-patient care. (Clark 1915).

[627] National Insurance Act 1911 1&2 Geo. 5 c. 55.

[628] Bryder highlights the role of comparisons with Prussian State provision in the promotion of British legislation. (Bryder 1988 p37-8).

Causative organism and types of tuberculosis

The cause of tuberculosis is a bacterium, *Mycobacterium tuberculosis*, a slow-growing, rod shaped microorganism first identified by Koch in 1882.[629] The bacillus is strictly aerobic; its dependence on oxygen is a fundamental factor in the pathology of the disease. Tubercle bacilli are destroyed by heat, and by disinfectants. They are rapidly destroyed in bright sunlight and any form of ultra-violet radiation. They are more resistant to dry conditions and chemicals than most pathogenic organisms. The bacillus is pathogenic in a number of animals, but only the human and bovine forms are transmissible to man.[630]

Tubercle bacilli may be disseminated throughout the body via the blood stream or lymphatic circulation. No part of the body is immune from infection, but there are differences in the degree to which organs are susceptible. The lungs, liver, spleen, and bone marrow are most susceptible. The meninges of the brain are less susceptible, but infection in this site is life threatening.[631] The characteristic pathological sign is the tubercle, a grey or yellow focus in the infected organ consisting of an agglomeration of lymphocytes, epithelioid and other cells surrounding the bacillus. Tubercles vary in size and consistency. They may or may not contain dead cells; they may or may not heal spontaneously. Later stages of tubercle development in the lung produce the characteristic cavitation seen on X ray films.

The great majority of infections in man are due to the human type of *M tuberculosis*. Infection is spread by airborne droplets expelled when coughing, sneezing, or speaking.

---

[629] Koch's work on tuberculosis formed a central part of his work on identifying a set of rules for the association of microorganisms with disease. The genesis of Koch's postulates is discussed in King 1952 and Codell Carter 1985.

[630] The bovine form is transmitted in milk, and was considered responsible for 65% of abdominal tuberculosis (Daniels 1952).

[631] The likely cause of meningeal tuberculosis is discharge of bacillus from a focus in the brain.

Transmission may be by droplet, but may also be by the dried nuclei of droplets. Infection does not generally lead to clinical disease. 'Most individuals successfully overcome infection, and remain unaware of their encounter with *Mycobacterium tuberculosis*, although healed lesions in the lungs bear witness to the encounter'.[632] In the early twentieth century, when the mortality rate from tuberculosis was approximately 1%, it was estimated that 90% of the population had been infected at some time.[633]

Outstanding issues concerning tuberculosis

Two unanswered questions about the epidemiology of tuberculosis remain important. Firstly who, amongst the infected majority, will go on to become the clinically affected minority? Secondly, why did the death rate from tuberculosis decline from the 1860s onwards? The answers to who will succumb and why the overall rate declined are linked because the rate of mortality from tuberculosis has fallen steadily since the mid nineteenth century, well ahead of falling rates of infection with *M tuberculosis*.

Several theories have been advanced. The decline in tuberculosis mortality forms a central part of McKeown's thesis that improving standards of living, and especially nutritional standards, are responsible.[634] Challenging this view, Szreter,[635] Hardy,[636] and Wilson[637] emphasise the role of public health measures, especially, in the case of Wilson, the effects of segregation, the indirect effects of decline in other infectious diseases on the incidence of clinical tuberculosis. On one point, they agree with McKeown:

---

[632] Hardy 1993 p213.

[633] Delpine 1913. Cited in Bryder 1988 p4.

[634] McKeown 1962.

[635] Szreter 1989.

[636] Hardy 1993.

[637] Wilson 1990.

... none of the treatments in use in the nineteenth or early twentieth century had a significant influence on the course of disease[638]

The evaluation of tuberculosis therapies

Although the methods and substances used to treat tuberculosis prior to streptomycin may now appear to lack either logical or empirical foundation, they were subject to various sorts of evaluation, including clinical trials.

In relation to the history of clinical trials, tuberculin deserves special mention. First announced by Koch in 1890 as a specific cure for tuberculosis, tuberculin aroused intense controversy. It may have been that the formulation of tuberculin remained somewhat cloaked in secrecy,[639] or it may have been because tuberculin so clearly played a leading role in Koch's ambition to establish a research institute.[640] However, Virchow alluded to the over-riding reason in his characterisation of practitioners who used tuberculin as 'poisoners and murderers'.[641] Tuberculin, it emerged, consisted of fragments of the tuberculosis bacteria which, according to theory, stimulated the body's (then poorly understood) natural anti-bacterial resources. Tuberculin was therefore among the earliest vaccine therapies. Unlike prophylactic vaccination, vaccine therapies were administered as treatment after disease had affected the body. In contrast to serum therapy, which used the body's own products, vaccine therapies used the disease producing organism itself to stimulate the body's defences. Raised to a therapeutic principle, the idea of treating a disease with the organism that produced it was bound to be controversial. In Koch's

[638] McKeown 1976 p92.

[639] The secrecy was ascribed by the Lancet in 1890 to the Prussian Ministry of Health, which wanted to retain the monopoly over tuberculin. (Lancet 1890 quoted in Waksman 1965 p93).

[640] The Institut fur Infektions-krankheiten, better known as the Institute for Infectious Diseases, or simply the Koch Institute, was founded in Berlin in 1891.

[641] See Weindling 1992 p178.

hands, and as a cure for the major killing disease of the nineteenth century, tuberculin was a sensation.[642]

Early enthusiasm for tuberculin in England was soon replaced by skepticism as clinical trials on small series of patients revealed little effect on the progress of disease. In 8 cases at the City of London Chest Hospital, and 11 cases at Kings College Hospital, no improvement was seen.[643]

However, the novelty and potential consequence of tuberculin treatment attracted a number of advocates. Most importantly, Almroth Wright adopted tuberculin in his therapeutic research programme at St Mary's Hospital, London. Wright's support for tuberculin was part of his campaign to revolutionise medical practice through the widespread adoption of vaccine therapy.[644] Wright's first undertaking was to explain why it had not previously worked. Tuberculosis weakened patients' immune systems, he argued. The doses of tuberculin patients had hitherto received had been too large for them to cope with. Wright applied a method of measuring a consumptive's immunological strength (which he called their opsonic index) before treatment with a graduated series of tuberculin. With new tuberculin trials providing evidence of the value of tuberculin in selected patients,[645] tuberculin, vaccine therapy and Almroth Wright, were much in vogue. Wright was knighted in 1906 for his work on anti-typhoid vaccination. He had

[642] On the initial reaction to Koch's announcement see Smith 1988 p56-57.

[643] Cited in Smith 1988 p57.

[644] Wright transferred from the Army Medical School, at the Royal Victoria Hospital on Southampton Water to St Mary's in 1902. For details of his programme see Colebrook 1954.

[645] For example at the Mount Vernon Hospital (Worboys 1992a p91).

many supporters amongst London society, including the aristocracy, politicians and the playwright George Bernard Shaw.[646]

A further boost for tuberculin came with its adoption by the tuberculosis dispensary movement, and the associated championing of tuberculin by the Australian physician Camac Wilkinson.[647] Part of Wilkinson's argument was that dispensaries were cheaper and more effective than sanatoria. They were also more suited to the treatment of working class consumptives who could not afford to spend time resting. There were other advantages. As Hilda Clark, medical officer to the Street Dispensary in Somerset explained, tuberculin was attractive to dispensaries because it was considered particularly useful in early and uncertain cases of tuberculosis, precisely the sort that dispensaries hoped to deal with. In addition, the routine for tuberculin injections was a useful way of maintaining contact between dispensaries and their patients.[648] The arguments found some support. In Portsmouth, the local authorities opened a tuberculin dispensary in preference to a sanatorium.[649]

The Tuberculin Dispensary League campaigned for greater state support for dispensaries, arguing that they were more effective than sanatoria. Bernard Shaw, in his Preface to *The*

[646] Wright was the model for the leading character in Shaw's play The Doctor's Dilemma, Sir Colenso Rigeon. Their relationship is described in chapter 13 of Colebrook 1954, which suggests that at one point Shaw considered writing a biography of Wright. Amongst politicians, Arthur Balfour and Lord Moulton were keenly interested in Wright's work. Balfour particularly played a leading role in getting research included into the provisions for sanatorium benefit in the 1911 National Insurance Act. And Moulton, first chair of the Medical Research Committee, regarded Wright's work as a model for medical science in Britain.

[647] Robert Philip opened the earliest dispensary in Edinburgh, at 13 Bank Street, in 1887. The Victoria Dispensary in Edinburgh provided treatment but also home visits and education. Camac Wilkinson came to Britain from Australia in 1909, where he had been a lecturer in pathology and medicine at the University of Sydney. His championing of a single treatment, and his challenge to sanatoria advanced in his *Treatment of Consumption* (Camac Wilkinson 1908) gave him a marginal status. He persisted however, and led a long campaign on behalf of tuberculin.

[648] Clark 1915.

[649] Mearns Fraser 1912.

Doctor's Dilemma compared sanatoria unfavourably with tuberculin, arguing that dispensaries only needed proper organisation to be effective:

> *When it comes to prophylactic inoculation, the ... complete scientific process ... can only be brought down to reasonable cost by being very highly organised as a public service in a public institution.*[650]

By the time the Government was ready to legislate on state funding for tuberculosis treatment, the relationship between sanatoria and dispensaries was marked by a sense that they offered competing models of care. As has been noted, Section 16 of the National Insurance Act of 1911 made specific provision for state services to treat tuberculosis. The reports of the 1912 Departmental Committee formed to implement the sanatorium clauses defused the opposition by recommending that dispensaries become part of state provision.[651]

However, the dispensaries envisaged by the 1912 Committee were unlike those recommended by Wilkinson. They did not concentrate on the provision of tuberculin. Instead they were regarded as an outreach from, and a gateway to, the sanatorium, under the supervision of the local medical officers.[652]

---

[650] Shaw 1993 p365, originally published in 1911.

[651] The Committee released an interim report in 1912 and a final report a year later. For a discussion of the Departmental Committee see Bryder 1988 p39.

[652] The new role for sanatoria is described in the symposium on tuberculosis and the National Insurance Act reported in the British Journal of Tuberculosis 1912;6:133-143. The view of FB Smith, a leading authority on the matter, is that British policy towards tuberculosis services was overly influenced by the concept of sanatorium, despite there being little evidence of their effectiveness. (Smith FB 1993) He argues that the bias towards sanatoria increased after the First World War, during which a temporary moratorium on sanatorium building had been enforced, when the Ministry of Health came into being. Among other effects, rational discussion and clinical trial of the anti-tuberculosis vaccine BCG was delayed by a generation or more in Britain. This may be an exaggeration, as the lack of interest in BCG was multi-factorial. However, the brief review of the contemporary literature that has been possible here offers some support for Smith's view. Certainly, the policy of the Ministry of Health was directed towards sanatoria. Then as now policy was directed to the use of and availability of beds. MoH Circular 280 of 1921 called for the running of sanatoria 'at the lowest cost consistent with efficiency', a policy designed to maximise bed occupancy and minimise drug usage. But with waiting lists increasing, in 1924 the Government announced that

The voluntary tuberculin dispensary system was marginalised, but sanatorium treatment was not incompatible with the use of tuberculin. A survey for The Lancet in 1912 found it in routine use in over 200 institutions in Britain.[653] However, tuberculin became increasingly controversial after 1911, and was rejected at the 1912 International Congress on Tuberculosis.[654] In part this may be due to a decline in the fortunes of vaccine therapy.[655] It can also be seen that willingness to prioritise the use of tuberculin and to condemn sanatoria was the touchstone for the philosophy of the independent dispensary movement. With most tuberculosis experts now based in sanatoria the effectiveness of tuberculin *contra* the sanatorium system became the focus for controversy about how to organise tuberculosis services.

In the face of controversy, trials of tuberculin were undertaken. Noel Bardswell, Superintendent of the Midhurst sanatorium, tested 154 patients between 1911 and 1913, and found that in comparison to untreated cases a 'considerable' number of the patients treated with tuberculin showed little improvement. He concluded that tuberculin 'had no obvious influence on the lesions'.[656] In contrast Hilda Clark concluded from her experimental series at the Street Dispensary and a review of other studies that:

---

grants would be available to substantially increase the number of beds available. (The number of beds at the time was 18,245 and the number waiting for in-patient treatment was 4150. BMJ 1924;ii:80).

[653] Lancet 1912 cited in Bryder 1988 p26.

[654] Worboys ascribes this to a direct conflict between dispensaries and sanatoria. Worboys 1992a p65-66: 'The details of Sanatorium benefit and the related schemes were worked out by an expert committee headed by Lord Astor. This was packed with supporters of sanatoria and they recommended that local authorities organise their own institutions for residential care and develop dispensaries and after care institutions. After 1912, medical opinion, as at the International Congress on Tuberculosis, turned decisively against tuberculin and once again endorsed sanatoria; this was unsurprising, as by this time most tuberculosis experts worked in sanatoria'.

[655] Worboys suggests 1910 as the peak in interest in vaccine therapy. (Worboys 1992**b**).

[656] Bardswell 1913 p133.

'Tuberculin can be used with apparent good effect in nearly every case in which the condition, as judged by ordinary clinical methods, allows a reasonable probability of improvement'.[657]

The Tuberculin Dispensary League agitated for further trials of tuberculin. Two evaluations took place at this time, both organised by the MRC.[658] The first, at Midhurst, following on from Bardswell's study, was part of an MRC funded evaluation of the effect of sanatorium treatment. The study found little evidence of the effectiveness of either sanatoria or tuberculin use within sanatoria.[659] The second evaluation followed representations to Christopher Addison, the Minister of Health from the Tuberculin Dispensary League in July 1919.[660]

Addison asked the MRC to investigate. In comparison to its Midhurst study, the evaluation undertaken by Western, Burrell and MacNalty was small scale, and its dismissal of the claims of the League went unpublished until 1923, when the Lancet published the report.[661]

---

[657] Clark 1915 p7.

[658] In her discussion of the MRC's role in tuberculosis research Bryder ignores these trials. Her argument is that once the MRC was freed from the requirement to concentrate solely on tuberculosis research it found other topics more interesting. To be fair to the MRC, it probably devoted as many resources to tuberculosis research in the years after 1912 as was warranted by the possible gain from such research. In keeping with her line of argument Bryder discusses research on tuberculin as a diagnostic aid, but ignores the therapeutic use of tuberculin. (Bryder 1989 p8-10).

[659] This consisted of an inquiry at the Midhurst Sanatorium in Surrey between 1914 and 1919. The MRC paid for a clerk to record data on the outcome for patients discharged from Frimley since 1907. The inquiry took the form of a comparison of the survival of Midhurst patients compared to the general population. Not surprisingly, the sanatorium did not restore the tuberculous to the levels of health enjoyed by the general population. (Bardswell 1919).

[660] Anon 1923.

[661] Anon 1923. Comparison between the two evaluations is instructive. The MRC evaluation of sanatoria was a full scale evaluation funded by the MRC, and published as part of its Special Report Series in 1919, while the evaluation of tuberculin was unfunded and consisted mainly of an analysis of the often incomplete medical records submitted by the Tuberculin Dispensary League, which was only published in the Lancet after its more damming parts were leaked to the BMJ. Landsborough Thompson's only reference to tuberculin is in a section entitled 'the unorthodox fringe in the treatment of tuberculosis (Landsborough Thompson 1975 p9-10): 'More troublesome were some forms of treatment in a mid-way category [midway between recognised therapies and out-and-out quackery of the Umckaloba sort], promoted by unqualified people and supported by 'evidence' of testimonial quality. The promoters

Neither of the official evaluations of tuberculin found any evidence to support the inflated claims of the Tuberculin Dispensary League. Despite his efforts to promote tuberculin, Camac Wilkinson was increasingly isolated from medical opinion[662] Nevertheless, favourable evaluations of tuberculin appeared occasionally in the medical press in the 1920s. William Stobie, Medical Officer to the Oxfordshire Association for the Prevention of Tuberculosis compared patients treated with tuberculin to those not-treated. He explained that the untreated patients were those living further away from the dispensary, and they fortuitously formed a control series. His results showed that 7 year mortality was 38/83 (46%) in the tuberculin group and 311/346 (90%) in the non-tuberculin group.[663] In the face of continuing evidence that tuberculin might be effective, and reports that it was routinely use by the medical profession in some areas, the BMJ softened its line on tuberculin. In 1927 it was content to suggest in an editorial that further research was needed,[664] to which one correspondent suggested the BMJ should revive its tradition of collective inquiry.[665]

Although official interest in tuberculin as a treatment faded away, interest among practitioners did not disappear until other specific remedies became available. Throughout the 1930s between 13 and 16 evaluations or clinical trials of tuberculin as a

---

tended to stimulate pressure groups of patients and busybodies, anxious to champion the supposed pioneers against the alleged tyranny of the professional 'closed shop'. There was, for example, a vaccine introduced in Britain by a citizen of a medically advanced foreign country, in which he was evidently a prophet without honour.' For the record, it appears that Camac Wilkinson disowned himself from the entreaties of the Tuberculin Dispensary League (Western 1923 p984) He is also on record as stating that the only proper trial is one in which a large series of patients are followed up (Camac Wilkinson 1926a) For his account of the MRC enquiry, see Camac Wilkinson 1926**b**.

[662] According to Smith 1988 p60-61. One BMJ review of Camac Wilkinson states that after 18 years of proselytizing about the virtues of tuberculin 'a very different verdict has been pronounced and is widely accepted' BMJ 1926 i:1039.

[663] Stobie 1922. See also Gillespie 1926 for a favourable evaluation of tuberculin.

[664] Anon 1927.

[665] Carswell 1928.

treatment were published each year.[666] Nevertheless, tuberculin is now remembered if at all as a diagnostic test rather than a treatment.[667]. Writing in 1956, and reflecting on changing modes of treatment, Sir Robert Young regarded tuberculin treatment as an ephemeral vogue.[668] It is worth noting that the MRC were more amenable to tuberculin as a diagnostic test. In the late 1920s they undertook a trial of tuberculin as way of detecting tuberculosis in cattle herds. The unexpected result was that tuberculin not only detected tuberculosis, it appeared to clear it from the herds selected for trial.[669]

Methodological considerations in trials of tuberculin

In considering the methodological aspects of trials of tuberculin several factors about Almroth Wright, the first serious scientific advocate of its use in Britain, are important. Firstly, his criticism of the medial profession's claim to scientific status. Secondly, his rejection of statistical method, and thirdly his several clashes with Karl Pearson, the founder of modern statistical practice.

Wright privately spoke of the bankruptcy of medicine,[670] pointing to controversy among practitioners and failure to introduce effective therapies and preventive measures as evidence. At the heart of his critique was a criticism of the defectiveness of the empirical methods used by the medical profession. Wright had trained at Foster's pioneering school of physiology, so would have been aware of the similar criticisms made by Claude Bernard. In calling for the creation of special facilities and adequate resources to

[666] see Feldman 1946 Table 1.

[667] Response to tuberculin could be a sign of infection. Early results were considered of doubtful value. In 1923 McNeil, at the Royal Edinburgh Hospital for Sick Children described an improved method (McNeil 1923).

[668] Young 1956. Sir Robert Young, consultant physician to The Middlesex and Brompton hospitals.

[669] Buxton 1928.

[670] Colebrook 1954 p47.

undertake medical research in 1905 he claimed that the medical profession was incapable of carrying out research:

> *Is there then, in our midst any agency engaged upon the study of the problems of disease? It will, perhaps, suggest itself to the man in the street that the whole medical profession is just such an agency, and that our hospitals are institutions in which the study of the problems of disease is actively carried on. This is far from being the case. The medical profession is not an agency for medical research, nor are our hospitals instituted or administered for the purpose of solving the problems of medicine.*[671]

Clinical practice was no basis for medical science. In its place, Wright, like Bernard, Pasteur, Koch and Ehrlich, proposed the laboratory as the heart of the medico-scientific enterprise. And like Barnard he dismissed the pretensions of statistical methods as producers of scientific knowledge. In their place Wright proposed what he called the 'experiental method' – the diacritical judgement of empirical data by the individual scientist – as less prone to fallacy than any numerical method.[672]

However, Wright did not forego the use of statistics as a demonstrative or rhetorical device. In his *Short Treatise on Anti-typhoid Inoculation*[673] of 1904 he compared mortality in inoculated and un-inoculated groups of soldiers, and on the basis of improvement in the majority of groups concluded that inoculation was effective. When he subsequently made the case for anti-typhoid inoculation to the War Office his use of statistical concepts was at once sufficiently important to his argument and methodologically inexact for the Office to seek the views of the leading English statistician Karl Pearson:

---

[671] Wright A. The world's greatest problem. Liverpool Daily Post August 30th 1905, reprinted in Colebrook 1954 p264.

[672] Colebrook 1954 discusses Wright's philosophy.

[673] Wright 1904.

> *Your opinions ... are extremely valuable, not only as those of an expert, but as those of an unbiased critic, while those of us who have been working at the subject are more or less prejudiced one way or the other.*[674]

In a private report to the War Office Pearson suggested that Wright's results were imprecise when considered from the statistical viewpoint. Wright, who regarded statisticians and their techniques as pretty much irrelevant to medical science, dismissed the charge. The debate subsequently continued across several issues of the BMJ, which tended to side with Pearson against Wright.[675] In the event the Army did adopt inoculation against typhoid, and later published results showing that inoculation may have had some protective effect.[676]

Pearson's critique of Wright was one of the earliest practical applications of a statistical technique introduced by Pearson in 1904 to test the probability of the observed frequencies in an empirical distribution given the expected distribution based on theory.[677] It relied on the calculation of a test statistic, and is therefore based on the frequency interpretation of probability (see Appendix 3). The test assumed that the empirical data be regarded as random samples from a larger statistical population. With a view to applying his chi-squared test for independence in contingency tables, Pearson proposed to the War Office the design of a clinical trial that would meet the requirements of the statistical test. The design would involve the allocation of alternate soldiers to

---

[674] Karl Pearson Papers, University College London Box 159/1. RJS Simpson to Karl Pearson 24th May 1904. I am indebted to Matthews 1992 p 210 for the unpublished material from the Pearson archive. For a general discussion of Pearson, see Magnello 1993.

[675] The correspondence was initiated by Pearson's report and continued in BMJ 1904 p1344, 1432, 1490, 1667, 1727 and 1776.

[676] Leishman 1909, cited in Colebrook 1954 p40.

[677] Pearson called his method the square contingency coefficient. Fisher later called it the chi-squared statistic. (Fisher 1924 chapter 5).

treatment and control groups so that both could be considered random samples.[678] If this were done, the test statistic could be calculated, and, as Pearson put it 'we can exhibit your results in correlative form, showing a distinct relation between inoculation and immunity'.[679]

Set against the statistical approach advocated by Pearson, the conclusions of tuberculin enthusiasts such as Hilda Clark – tuberculin will work in patients in whom it is likely to work – sounds like special pleading. But it represents a persistent strand of thought in medicine that places the clinician's judgement at the centre of efforts to evaluate therapy. Among tuberculosis doctors the need to <u>select</u> patients for a treatment was a commonplace:

> *Excellent results have been obtained, but just as with sanatoria, so too with this agent [tuberculin]…it is only of value in carefully selected cases.*[680]

> *'Dr Heaf thought he was probably not alone in thinking that there was a certain amount of misuse of AP [artificial pneumothorax] treatment both clinically and administratively principally by faulty selection of cases…'*[681]

[678] Pearson, like Bradford Hill later, argued that the reason for allocation was to form groups of 'like risk'.

[679] Karl Pearson Papers, University College London Box 159/1. Karl Pearson to RJS Simpson May 24th 1904. Again, this quotation is from Matthews 1992. For the complex historiography of the chi-squared framework, which she judged to be Pearson's greatest single contribution to the development of statistical theory, see Magnello 1993 chapter 6.

[680] Sutherland 1912 p16.

[681] Tuberculosis Association 1933. 'There were, said Dr Heaf, certain types in which the AP treatment was unnecessary and sometimes harmful, such as the young adult with negative sputum and with a unilateral lesion shown in the radiograph' At the same meeting HC Toussaint described 7 causes of failure of AP. Accompanying the view that patients must be selected for treatment was an exquisite finesse in classifying the types of tuberculosis. (p1125) In 1948 one classification included 55 types, which were then rationalised into 32 types (Sekulich 1949).

> *[regarding the use of extrapleural pneumothorax in 1952] Despite serious postoperative complications and the disadvantage of protracted refills the results in this series indicate that there is still a limited place for the operation in carefully chosen cases'*[682]

At issue therefore was not the abstract notion of the efficacy of a treatment. The central issue was the clinician's ability to distinguish between those who may benefit and those who will not. In the case of tuberculin Clark was arguing that it was only by judging a patient's vitality (equivalent to Wright's Opsonic Index but without the need for test tubes) that one would know whether or not they were suitable for treatment.

The exclusion of drop-outs from the statistics presented by tuberculin enthusiasts can also be seen as part of this approach. By dropping out from therapy the patient was, as it were, highlighting the fact that they were un-suitable for treatment. To then include their results would be to unfairly contaminate the results of tuberculin. By excluding them the drug was able to show its full therapeutic potential. To a skeptic, and to the modern reader, all this sounds like no more than the unfair promotion of the therapy undergoing trial. It must be stressed though that at the time it violated no fundamental rule of medical thought.

From a statistical point of view, the selection of patients and the exclusion of dropouts invalidated statistical tests and were therefore forms of bias.[683] The offence against statistical propriety was greater still if the clinician claimed to be using a 'control' in his or her evaluation. The meaning of the word control varied considerably amongst those who used it. The only shared feature was the connotation that by controlling their results

[682] Baldry 1952.

[683] Fisher suggests that the conception of inadequate experimental design was common among experimental workers by 1935. (Fisher 1935 p2).

something precise and scientific was being offered. In this context, almost any form of comparison or effort at measurement could be described as a control.[684] Camac Wilkinson regarded the death (from tuberculosis) of a case he was prevented from treating as 'a striking control'.[685] Although not grounded in a statistical justification, control groups played a role in several evaluations of tuberculosis therapies because it was universally acknowledged that the course of tuberculosis was highly variable. By including a control group some sense of the efficacy of a method be judged. Bardswell used a control group.[686] In America, EL Trudeau used a control group.[687] So did Hartley and colleagues, in a trial of artificial pneumothorax at the Frimley sanatorium. They chose a group of patients for pneumothorax, and compared them to a group of age, sex and severity matched patients for whom pneumothorax was not recommended.[688]

If the use of controls, however formed, was regarded in the late nineteenth century as a relatively innocuous bid by clinicians to be scientific, by the early twentieth it could be construed as the less than competent adaption of statistical methods. Of considerable interest therefore is Karl Pearson's preface to Bardswell's sanatorium based clinical trial

[684] The concept of a control group was largely absent from medicine before the twentieth century, but increasingly used thereafter. In Appendix 5 of this thesis, the use of a control group in an experiment of 1860 is noted. The occasional use of control in medical research suggests that the idea of a control in nineteenth century medicine lacked connection to a body of theory. A control group could be used, or not. However, it is clear that control groups were used in medicine before they were part of the requirement of statistical design. Lancaster 1994 p223-4 Boring 1954 and Danziger 1984 survey give some nineteenth century examples the use of control groups in psychological research in the twentieth century. Both conclude that the use of control groups preceded the statistical framework which made them necessary.

[685] Wilkinson 1923 p676.

[686] Bardswell 1914.

[687] EL Trudeau, founder of the Saranac Lake sanatorium in the Adirondack Mountains of New York. His trial of tuberculin in 1906 compared 185 patients treated with tuberculin with a control group of 864 not so treated. Among patients with advanced tuberculosis the 27% were cured in the tuberculin group, compared to 6% in the non-tuberculin group. Trudeau noted that the results could be due to the fact that there was a tendency to give tuberculin to patients who showed signs of better nutrition. Cited in Dowling 1977 p74.

[688] Hartley, Wingfield and Burrows 1936 p41 and graph 17. The effect of comparing pneumothorax to control, when shown graphically, was to highlight the impact of pneumothorax.

of tuberculin. Pearson observed that the comparison of tuberculin treated patients with non-tuberculin treated patients was illegitimate so long as the groups were different. Pearson went on to describe the design for a randomised controlled trial in medicine. He suggested a way to randomise patients. He also proposed an ethical justification for human experimentation:

> *Is not the right attitude that which considers it very much as a treatment at the experimental stage, and when such an attitude is taken, are we not ethically justified in the only judicious* ***experimental*** *manner? We cannot tell a patient with any certitude that it will benefit him, and we are not bound therefore, to apply it. On the other hand we cannot definitely say at present that with selected classes of patients, cautiously treated, it is positively detrimental. If it were possible the scientific method would be to select patients suitable for tuberculin treatment, treat only those whose surnames begin with A-K, and then compare the results with simple sanatorium treatment of the remainder, L-Z, of these selected patients. Then in two or three years we should know exactly the value of the treatment."*[689] [the emphasis on experimental is Pearson's]

Pearson, the eugenic socialist, had little difficulty conceiving of experiments on populations. Hilda Clark described his proposed methodology shortly afterwards as 'impossible'.[690] And although there was some recognition among clinicians and researchers that Pearson's method of alphabetic randomisation was the ideal, it was also recognised that the scheme was ethically unacceptable.

[689] Bardswell 1913 pxiii.

[690] Clark 1915 p252 She did not say why. The reason is likely to be that tuberculin formed an essential link between the dispensary and its patients. A trial in which suitable patients were not given tuberculin would have threatened that link and raised suspicions that dispensaries were carrying out experiments on their patients.

> *Perhaps, scientifically, the best method would be to classify all patients on admission as being suitable or not suitable for treatment with tuberculin, and thereafter of all those regarded as suitable to give tuberculin to patients whose names begin with A-K and general measures only to those whose names began with L-Z ... It was considered however that this plan was too much in the nature of an experiment*[691]

As we have seen with trials of anti-pneumonia vaccine, from the late 1920s the ethical censure against no selection trials decreased. Among tuberculosis therapies, a form of randomised controlled trial was applied to sanocrysin in 1926 at the Detroit municipal sanatorium. This involved the selection of 24 patients considered suitable for treatment. The patients were then matched pair-wise, and by the flip of a coin assigned to treatment or non-treatment groups.[692] The methodology was repeated in 1944 at several Minnesota mental institutions, in a trial of a derivative of sulphonamide known as Promizole.[693]

The methods of evaluation proposed for tuberculin and hesitatingly used in trials of sanocrysin were subsequently adopted more fully in trials of the new anti-tuberculosis agent streptomycin. Methodological aspects of trials of streptomycin will be discussed in the next section.

The MRC randomised controlled trial of streptomycin

Streptomycin is widely regarded as the first effective treatment for tuberculosis. The award of the Nobel Prize for Physiology or Medicine to Selman Waksman in 1952 established his claim to have discovered that streptomycin halts the growth of *M tuberculosis.* As several authors have noted however, the honour should at least be shared

---

[691] Bardswell 1919 p41, footnote. Support for the use of 'proper' controls, and support for Pearson's alphabetic method came from the Bristol practitioner Ernest Weatherhead (Weatherhead 1928).

[692] Amberson 1931. Patients were asked if they wanted to be included in a trial, but not told of the experimental design which meant that half would not be treated.

[693] Hinshaw 1969. The Promizole trial extended to cover streptomycin. See next section of this chapter.

with his postdoctoral research assistant Albert Schatz, who carried out the earliest experiments with streptomycin.[694]

The discovery that an antibiotic substance produced by the soil living bacterium *Streptomyces griseus* was effective in stopping the growth of *M tuberculosis* was made in the soil laboratory at Rutgers University in the Summer of 1943.[695] The laboratory had included *M tuberculosis* in tests of potential antibacterial substances following a visit to his laboratory by two scientists from the Mayo Clinic, William Feldman and H Corwin Hinshaw. They were systematically searching for an effective antibiotic following a report that sulfanilamide had a limited suppressive effect in experimental tuberculosis.[696]

Following publication of the discovery of the bacteriostatic activity of streptomycin in January 1944,[697] Waksman wrote to Feldman on March 1st 1944 to ask if Feldman was prepared to test streptomycin on guinea pigs with experimental tuberculosis.[698] Between April and June a small sample of streptomycin was successfully tested on 4 tuberculous guinea pigs. Between July and September, a larger amount, this time supplied by Merck

[694] The controversy is described in Wainwright 1991, who argues that Waksman took little interest in antibiotics before the arrival of Schatz at his laboratory. For an account of Waksman's career see Comroe 1978. Waksman's own account is published in several places, including section III of Waksman 1965, chapter 15 of his autobiography (Waksman 1958), Waksman 1951, and Waksman 1954. Waksman earned a personal fortune from the royalties on streptomycin, (Wainwright's figures suggest that Waksman earned $500,000 in the period to mid 1950 alone) which he used in part to support his research at Rutgers University.

[695] For details of the microbiology of S griseus and the chemistry of streptomycin as understood in 1949, see the multi-author volume edited by Waksman (Waksman 1949).

[696] Feldman was a staff member of the Department of Comparative Pathology at the Mayo Graduate School. Hinshaw was a bacteriologist at the Department of Pulmonary Diseases, Mayo Clinic. For further details of their career, see Comroe 1978 and Feldman's own account (Feldman 1954).

[697] Schatz, Bugie and Waksman 1944.

[698] Comroe 1978 p960.

and Company, was tested on 24 control guinea pigs and 25 subjects, again successfully.[699] (See Figure 4).

Figure 4: Results of streptomycin on experimental tuberculosis in guinea pigs

[699] Report of the two studies is made in Feldman 1954.

Streptomycin was first used to treat human tuberculosis on a patient on November 20th 1944 at the Mineral Springs Sanatorium, Cannon Falls, Minnesota. The patient, a 21-year-old woman, had progressive late-stage tuberculosis of the right lung, and was therefore considered suitable for experimental administration of streptomycin. During the trial she was transferred from the Sanatorium to the Colonial Hospital, Rochester. Between November 1944 and April 1945 the patient (identified in the report as PT) received 5 courses of streptomycin. Over the summer of 1945 the patient continued to improve. She was discharged from the sanatorium in 1947 and was still alive in 1955. [700]

The involvement of Merck in the production of streptomycin meant that greater quantities were available from December 1944.[701] Karl Pfuetze and colleagues from the Mayo Clinic undertook further testing, as neither Feldman nor Hinshaw were in a position at the time to organise human trials. Testing of streptomycin began as an add on trial to the tail end of double blind matched pair randomised clinical trial of Promizole being carried out by Karl Pfuetze and Marjorie Pyle at mental institutions around Minnesota. According to Hinshaw, the Promizole trial was about two-thirds complete by the Summer of 1944, when the test drug was switched to streptomycin. This might have had some claim to be the first randomised controlled trial of streptomycin were it not for the fact that Hinshaw considered that it would have been unethical:

[700] The trial is described fully in Pfuetze et al 1954. Pfuetze states with confidence that the trial on PT was the first time streptomycin was administered to a human being. But Hinshaw also later recalled an early human trial (which he didn't date) in which the patient had developed tuberculosis due to a surgical accident. This manner was similar to the way he induced tuberculosis in guinea pigs. The patient was treated with streptomycin, but died. (Hinshaw 1969).

[701] Waksman had a long-standing relationship with Merck. Feldman suggests that Merck's involvement in streptomycin began formally at a meeting held on July 10 1944 (Feldman 1954 p863).

> *The matched pair, double blind controlled study was not extended to streptomycin because the therapeutic benefits were so obvious that we did not have the conscience to deny streptomycin to the controls.*[702]

On September 5th 1945, Hinshaw and Feldman published a report on 34 cases. By early 1946 the number of cases had grown to 75, and by late 1946 the number had grown to 100.[703] In May 1946 the results were regarded as sufficiently promising for a large scale multi-centre trial to be launched by the Veterans Administration (henceforward VA), in collaboration with the US Public Health Service, the National Research Council and the National Tuberculosis Association.[704] The study encompassed 7 centres, most of them belonging to the VA. Three study centres were in operation by July 1946, and by November 1946 each centre had enrolled at least 15 patients.[705]

The best known fact about the VA trial is that despite originally intending to randomly allocate patients to study and control groups, it did not do so. The official reason is that shortly after trial begun clinical material was in such short supply that to have done so would have reduced the number receiving streptomycin to an unacceptably low level. Harry Marks, who has studied the minutes of the committee, confirms that shortage of clinical material was the main motivating factor for <u>not</u> randomising. Marks' interpretation is that the committee was looking for quick results, and to have divided the

---

[702] Hinshaw 1969 p199.

[703] Respectively: Hinshaw and Feldman 1945; Hinshaw and Feldman 1946a; Hinshaw and Feldman 1946b. According to Hinshaw, following the first of these reports a series of articles appeared in the popular press.(Hinshaw 1954) This led to a stream of enquiries from patients and their relatives, which were handled by Dr Chester Keefer on behalf of a specially convened committee of the National Research Council.

[704] The earliest official trial of streptomycin was undertaken by the American Trudeau Society, following a large donation of streptomycin, presumably by Merck. See Hinshaw 1954 p13 and Marks 1987 p 131 footnote 2. The decision by the Veterans Administration followed the meeting of the National Tuberculosis Association in New York in May 1946, at which the results obtained by Hinshaw and colleagues were presented, along with findings from Cornell University reported by Walsh McDermott.

[705] Veterans Administration 1947.

limited clinical material between study and control would have hampered progress. [706] While this is the argument advanced in the VA's published paper, Marks also shows that there may be other factors involved. Carroll Edwards Palmer, the representative of the United States Public Health Service (USPHS) on the trial committee did wish to use an untreated control group, and was supported by the committee's statisticians.[707] When the USPHS wanted to fund a RCT, the VA did not participate.[708]

The VA trial published its preliminary report in 1947. Despite the methodological weakness, and accompanying reports from individual centres that made interpretation more difficult, the results confirmed the effectiveness of streptomycin. At 120 days, 85% of patients had some degree of clearing of their lesions, as shown by X-ray film. The partial clearing of exudative lesions was considered to be the principal benefit of streptomycin. Clearance was associated with weight gain in most patients. The study also confirmed that streptomycin damaged the vestibular apparatus of the inner ear, and that resistance to streptomycin developed in *M tuberculosis*. Guidelines for use accompanied the trial results, setting out the indications for therapeutic use of streptomycin.[709] For some presentations – tuberculous meningitis, acute miliary tuberculosis – streptomycin

[706] Marks 1987 footnote p 134 footnote13.

[707] Carroll Edwards Palmer 1903-1972. Leading advocate of random allocation and blinded observations in clinical research. Palmer had worked in the Biostatistics Department at Johns Hopkins University in the 1930s. Later became Professor of Biostatistics at the University of California School of Public Health. His career is discussed in Comstock 1972. I am grateful to Marks 1997 p122 footnote 107 for this reference.

[708] This trial may well have been that of Jenkins and colleagues, reported in 1947, which acknowledges the USPHS as a funding body and did divide patients into treatment and control groups, by a method not described in the paper. (Jenkins 1947) This trial appears to have been ignored in the standard accounts of the history of clinical trials of tuberculosis. It is not mentioned in D'Esposo 1982 or Hinshaw 1969. Jenkins' trial is of some interest because it did compare streptomycin treated patients with untreated controls. The trial design resembles the MRC's, and there are similarities in the graphical presentation of the results. There is no evidence that the MRC streptomycin trial was modelled on it, since it was underway by the time Jenkins published. It is likely that D'Arcy Hart at least was aware of the trial, as he read very widely during the 1940s, and had access to all the tuberculosis journals (D'Arcy Hart 1946).

[709] Committee on Therapy, American Trudeau Society 1947.

was recommended without reservation. It was less clear which types of pulmonary tuberculosis were affected by streptomycin, and what the role of streptomycin was in relation to other therapies. Given the possible side-effects and the ability of *M tuberculosis* to become resistant, the committee called for further and more adequately controlled trials 'to determine the possibilities and limitations of streptomycin therapy'.[710] The emergence of drugs such as PAS and Isoniazid in the 1950s meant that the VA research programme changed direction. It was clear that combinations of drugs were more effective, so further research on streptomycin alone was unnecessary.

Streptomycin production in America

By August of 1946 Merck and Co. had begun production of streptomycin on a significant scale. Production was increased by the use of a technique of submerged, or deep-vat, fermentation. By carefully controlling the conditions under which the cultures were maintained, far more streptomycin could be produced.[711] In 1946, American production amounted to 1,000kg. In 1947, when streptomycin was released onto the market, annual production amounted to 10,000kg. This level of production was sufficient, according to Max Tishler of Merck & Co., to meet demand in America.[712] In that year the combined sales of streptomycin and penicillin in America amounted to $112M, approximately half of all sales of synthetic drugs.[713]

[710] Committee on Therapy, American Trudeau Society 1947 p479.

[711] For details of the technique see Tishler 1949.

[712] Tishler 1949 In 1947 6 large US pharmaceutical manufacturers were producing streptomycin: Merck & Co; Charles Pfizer Co; Abbott Laboratories; Upjohn Company; Eli Lilly and Co; ES Squibb & Sons (Committee on Therapy, American Trudeau Society 1947 p480).

[713] Waksman provides some data on production in his 1951 paper (Waksman 1951 p346).

Year	Production (kg)	Exports (kg)	Price per gm($)
1945	*c* 300	None	-
1946	1,175	45	25, falling to 4.7

Streptomycin in Britain

The development of facilities for streptomycin production in Britain owes much to the commercial relations during the Second World War between the pharmaceutical manufacturers Glaxo and Merck.[714] Following several trips to America by Glaxo scientists, in January 1946 Glaxo committed itself to production of streptomycin under license from Merck.[715] Also in 1946, Boots had established a small plant for the experimental production of streptomycin.[716]

However, despite continuing collaboration with American manufacturers, British production was insignificant before 1948.[717] Importation of American streptomycin was impossible in 1945, and remained problematic until the middle of 1946 when the American authorities began to allocate small monthly quotas of streptomycin to Britain. Shortly after the quota scheme began, and for reasons that remain unclear, a large consignment of streptomycin was allocated to Britain. It was this consignment that made the British trial of streptomycin in 1947 possible.

1947	9,676	3,450	3.16
1948	37,709	26,500	1.57
1949	83,699	63,412	0.88
1950	92,446	66,419	0.60
1951 (1st qt.)	31,140	16,523	0.60

[714] Especially relevant is the transfer of American deep-vat fermentation technology to Glaxo in 1944/45. This was used to produce penicillin on a scale far greater than hitherto possible (Davenport Hines 1992 p147).

[715] Davenport Hines 1992 p180 and p243.

[716] FD1 6756 D'Arcy Hart to Mellanby 9th August 1946. Boots was one of the two companies authorised by the Ministry of Supply in 1944 to develop deep-vat fermentation technology using American expertise. (Davenport Hines 1992 p147).

[717] At the Streptomycin in Tuberculosis Committee meeting of 5th November 1947 it was announced that three firms – Boots, Glaxo and Distillers – were engaged in production. Of the three companies, only Boots had succeeded in producing any usable streptomycin, a batch of 28 gms. (FD1 6757 TCTC 38 Streptomycin in Tuberculosis Trials Committee. Minutes of 3rd meeting 5th November 1947) Enquires by Sir Edward Mellanby in early 1948 found that streptomycin was being manufactured on a commercial scale by Glaxo at its Barnard Castle factory in Durham in February 1948 at the rate of 0.5kg/month, rising to 20 kg/month, and that large-scale production at Ulverston, North Lancashire was anticipated in 1949, where the output was estimated to be in the order of 200 kg/month (FD1 6757 Sandercock to Mellanby 20th February 1948). At Boots, it was anticipated that 3kg/month would be supplied from April 1948 (FD1 6757 Drummond to Mellanby 18th February 1948).

The Public Record Office contains four files of requests and enquiries to the MRC concerning streptomycin.[718] The sheer volume of the enquiries and responses highlights the pressure the MRC came under from 1945 onward to release streptomycin for treatment and clinical trial, even at a time when it had no streptomycin at all. Doctors, who presumably had access to medical literature, made the earliest enquiries, on behalf of their patients but also their families. Later, following articles in the lay press and radio broadcasts, enquiries came from the public. The following excerpts are a just a sample from the enquiries received in the October 1946:

From the Winsley Sanatorium, near Bath:

> *Dear Sir, I am very anxious to obtain this antibiotic… could you kindly inform me if it is possible to obtain some for trial.*

From the Hawksmoor Sanatorium, Bovey Tracey, Devon:

> *Dear Sir, I should like to take part in any trials with STREPTOMYCIN.*

From a doctor in Lancaster

> *Dear Sir, I have under my care a patient suffering from tuberculosis of the lungs…I understand that supplies [of streptomycin] are limited and only available on a permit from you…*[719]

Origins of the British streptomycin trial

British policy in response to calls for streptomycin can be traced to the Spring of 1946. At the time the Ministry of Health had good reason to contain an issue which, if mishandled, could cast doubt on its ability to manage policy for the rational and equitable

[718] The PRO contains three folders of enquiries FD1 6760 – 6763.

[719] FD 1 6760 correspondence with the MRC between October 25th and 29th 1946.

provision of health care resources.[720] There may be no *prima facie* evidence that the NHS Bill had a direct influence on policy towards streptomycin. However, left to the free market, streptomycin would violate the principles of the Beveridge Report,[721] in that it would neither be universally available nor free. Nor would it be directed at those who needed it most. It would go to those who could afford it rather than those who needed it. Streptomycin, it can be argued, represented a threat to the Ministry's ability to organise and direct resources along lines consistent with the aspirations of the Beveridge Report.

Similarly, the streptomycin issue raised the question not only of what contribution the MRC would make to the NHS, but also how the MRC would operate within the structures of the emerging NHS. Whether or not the NHS Bill was influential, it was clear from the interest generated by streptomycin that some form of official response was necessary.[722]

---

[720] Webster 1988 p94-103. MoH officials were deeply involved in the final stages of the National Health Service Bill during 1946. It was published on 20 March 1946 after many months of detailed negotiation and drafting. The Bill secured a second reading on 30th April and over the summer passed through standing committee, committee and report stages, received a third reading on 31st of October, and Royal Assent on 6th November. According to Webster, the fate of Bevan's new scheme for a comprehensive health service was far from assured in 1946, largely because of the opposition of general practitioners. (Webster 1998 p25).

[721] For the Beveridge Report see Abel-Smith 1994.

[722] This is one example of the way in which new organisational structures became part of the streptomycin story: as will be argued below, the MRC wrested the initiative to 'do something'' about streptomycin from the MoH in mid 1946. Thereafter, most streptomycin supplies were assigned to the MRC, who distributed it to selected individual researchers and hospitals. However, in mid 1947 the Chief Medical Officer at the Ministry instituted a scheme for the distribution of streptomycin to teaching hospitals, a type of hospital particularly favoured in the NHS Bill, having a direct relationship with the Ministry rather than one mediated by Regional Hospital Boards. Streptomycin was thereby used to strengthen the Ministry-Teaching hospital relationship, and to assert the role of the Ministry in dealing with streptomycin. The response of the MRC was to attempt to block extensions of this scheme. The MRC succeeded in transferring Ministry streptomycin to itself in early 1948, so that it could run its own limited scheme for the treatment of tuberculous tracheo-bronchial ulceration. The amount transferred was sufficient for treatment until the end of 1948. In September 1948, shortly after the inauguration of the NHS, the Ministry announced its scheme for the distribution of streptomycin, creating in each new Region one distribution centre and a number of hospitals where treatment could be provided 'staffed and equipped to ensure the necessary scientific control of the treatment'. (Ministry of Health 1948) The Ministry centres included all the ones used by the MRC for its trials. One small piece of evidence points to MRC displeasure at the Ministry's initiative. In August 1948 the head of a commercial laboratory in Northumberland wrote to GS Wilson. The letter stated that he had been asked by the Walker Gate Infectious Diseases Hospital to undertake streptomycin assays on treated patients. 'I do not understand how this venture by the Ministry is related, if at all, to the investigation by the MRC… I should be glad to know whether you

The idea that a British clinical trial of streptomycin was needed originated in July 1946, when the initiative for policy making on streptomycin slipped from the Ministry of Health to MRC.[723] From this time on the form of policy response concentrated on the need for clinical trials. As I argue below, clinical trials offered a way to contain the several facets of the 'issue' of streptomycin. The trial that was subsequently arranged has been hailed for its methodological innovation. In the conclusion to this chapter it will be argued that the methodology of the trial also contributed to the overall policy aims of containing the issue of streptomycin.

However useful, the need for a clinical trial in Britain itself needed justification.[724] At its core therefore, official policy towards streptomycin in Britain involved emphasising the uncertain benefits of its use. Uncertainty had several practical advantages. Firstly, it allowed the MRC to respond to enquiries with a message about the uncertainty of streptomycin's efficacy. Secondly, it self evidently justified the need for research before making streptomycin more widely available.

Beginning in 1946 the Ministry, and subsequently the MRC, began to consciously build policy around the uncertainty of the therapeutic effect of streptomycin. British understanding of the effectiveness of streptomycin was established by Philip D'Arcy Hart, who delivered the Mitchell Lecture at the Royal College of Physicians on July 9th

---

would like us to try to meet the request'. (FD1 6757 Messer to Wilson 19.8.48) Wilson took the trouble to visit Messer, and advised him not to comply, citing building work and staff changes as the preventing reasons. (FD1 6757 file note of Wilsons's visit on 14.9.48).

[723] FD1 6756 File note by Edward Mellanby following a meeting with Professor H Raistrick, then working for the Ministry of Supply, and who later sat on the Trial Committee. The note reads 'Raistrick came here to discuss the question of streptomycin. He had persuaded Dalrymple-Champneys [Sir Weldon Dalrymple-Champneys, Deputy Chief Medical Officer] to put off the meeting which he had called at the Ministry of Health, and to which I had promised to go, so that we could see what the Medical Research Council could do in this matter'.

[724] An alternative policy response would have been for the Ministry to have made small quantities available at teaching hospitals, using American information to guide usage.

1946.[725] By setting his discussion in a historical framework, D'Arcy Hart emphasised the many false dawns in tuberculosis chemotherapy. He quoted Waksman's formulaic assertion of the need for more research:

> *The effect on human tuberculosis justifies cautious optimism for certain forms of the disease, but, subject to any very recent information, Waksman's own words, of November 1945, still hold: 'Prolonged treatment and studies of many cases are absolute prerequisites for any serious consideration of the efficacy of streptomycin in the treatment of tuberculosis. To date sufficient information has not yet been accumulated' The need for caution has been learned from bitter experience of past failures with gold, copper and tuberculin, and from the false promise given by animal experiments with the sulphones.*[726]

In response to enquiries to the MRC from individuals, Landsborough Thomson was more forthright. In May 1946 he responded to a plea for streptomycin from a Medical Officer for Health:

> *It is perhaps poor consolation, but I may mention that the first reports of the value of streptomycin in tuberculosis seem to have been rather too optimistic, and the indications are that its chief uses may be in other conditions.*[727]

While the need for caution expressed in D'Arcy Hart's Mitchell Lecture was at least defensible, MRC correspondence to individuals is hardly based on fact at all. Despite its contrivance, Landsborough Thomson's form of reply became the norm when it was

[725] Philip D'Arcy Hart, later secretary to the Streptomycin in Tuberculosis Trials Committee and Director of the Tuberculosis Research Unit. At the time of the Mitchell Lecture he a staff member of the NIMR, having joined MRC in 1936 to undertake research on pneumoconiosis among coal miners in South Wales. Dr D'Arcy Hart is one of two surviving member of the Streptomycin Committee, the other being Guy Scadding.

[726] D'Arcy Hart 1946 p852. The published version appears to be a transcript of the lecture, as D'Arcy Hart introduced new information via footnotes rather than change the text itself. Waksman later considered that the therapeutic value of streptomycin was 'well established' by 1945 (Waksman 1951 p347).

[727] FD1 6760. Landsborough Thomson to X 29 May 1946.

included in a standard response used to fend off enquiries.[728] A broadly pessimistic assessment about the role of streptomycin was repeated in an article written for the BMJ in December 1946.[729] The article stressed the potential hazards of using streptomycin in strong language. However, so negative was the article that it resulted in a cable from the British Supplies Council Office in Washington, stating that the American medical authorities were seriously concerned about the toxicity statement.[730] Landsborough Thomson responded privately that 'the MRC do not feel repentant as the trouble is due to their playing it up'.[731]

In other contexts, and for a different audience, the MRC's assessment of the potential of streptomycin was markedly more optimistic. Asked by the Ministry of Supply for advice in April 1947, the Streptomycin Committee resolved that:

> *We feel that the evidence already available,* ***mainly from work in the USA****, is sufficient to justify the view that in this country alone at least 50 to 100 kg per month could be absorbed during the next 2 to 3 years for more general trial'*[732] [my emphasis]

[728] FD1 6760 MRC Standard statement on streptomycin 8 October 1946, paragraph 6. The statement is marked restricted. Since the statement was included verbatim in correspondence this classification was probably intended to prevent its dissemination to other departments of state.

[729] FD1 6756 FHK Green to Graham Murphy 4th December 1946. 'Streptomycin: the present position'. Green subsequently made it clear that the MRC wished the piece to appear as if it came from the BMJ. (FD 1 6756 Green to Murphy 10th December 1946. The published report appears as Streptomycin: the present position. BMJ 1946;ii:906.

[730] FD1 6769 Jan 30 1947.

[731] FD1 6769 Landsborough Thomson to Miss EMR Russell-Smith MoH 5/2/47. The propaganda issued by the MRC appears to have been successful. In a statement issued by the Ministry of Health in 1950 it was noted that 'the claim that streptomycin does no more than … produce recovery as a physical and mental wreck. This is not true.' (Ministry of Health 1950) The phrase repeats that of the MRC in 1946, suggesting that it had become current among clinicians.

[732] FD1 6756 Streptomycin Trials Committee. Minutes of second meeting 18th April 1947. The likely reference is Keefer's report on 1000 cases treated with streptomycin (Keefer et al 1946). But this was published in September 1946. The only subsequent data likely to be sufficiently compelling was the VA trial. But its preliminary results were not published until late 1947. (Veterans Administration 1947).

It would appear that the expression of uncertainty about the value of streptomycin was dependent on the audience.

Organising the MRC streptomycin trial

Therapeutic uncertainty created a framework in which policy options were clarified. Once the MRC took the initiative on streptomycin in July 1946 it made quick progress in developing the methodology of a trial. The day after meeting Raistrick, Mellanby discussed streptomycin with his deputy, Harold Himsworth. At this meeting Mellanby tested the idea that a trial of streptomycin should be limited to tuberculosis. The following day, 26th July 1946, Mellanby contacted a small group of specialists to arrange a conference on tuberculosis research. Among those invited was Austin Bradford Hill.[733] The conference took place on 29th July at the MRC Offices in Old Queen Street, at 5.30 p.m. Apologies were received from Heaf and Raistrick. Others were represented by deputies. Bradford Hill did not attend, but sent his assistant, WJ Martin.

At the conference, Mellanby informed the group that a limited amount of streptomycin would shortly be available from British manufacturers.[734] The conference accepted Mellanby's view that the trial should be limited to tuberculosis in the first instance. It also agreed 4 types of tuberculosis suitable for trial, of which the main group was to be acute, rapidly progressing tuberculous broncho-pneumonia at ages 15-25 (hereafter this group is

[733] The full list is S Roodhouse Gloyne (London Chest Hospital); HM Macaulay (Middlesex County Council); FRG Heaf, (Senior Medical Officer to London County Council and Consultant to Papworth Village Settlement); Philip D'Arcy Hart (NIMR); Geoffrey Marshall (Brompton Hospital); JG Scadding (Brompton Hospital); JC Hoyle (Brompton Hospital); and Geoffrey Todd, (Medical Superintendent, King Edward VII Sanatorium, Midhurst). Professor H Raistrick, (Royal Infirmary Glasgow); Professor GS Wilson, (Director of the MRC Emergency Public Health Laboratory Service). (FD 1 6756 Letter of invitation 26th July 1946).

[734] Despite assurance from Raistrick on 24th July this turned out not to be the case. Correspondence between the Ministry and MRC on 29th August (FD1 6756 Marchbank to Mellanby) included a note from the Ministry of Supply saying that streptomycin in the quantities needed would not be available until the Spring of 1947.

referred to as the main study group).[735] It was agreed to hold the trial in several hospitals, most of which were represented at the conference. The need for a central laboratory to test the sensitivity to streptomycin of *M tuberculosis* in each case was agreed. It was agreed to appoint a field-worker, Marc Daniels, lately MRC Prophit Scholar.

The following Sunday, a small group met at the home of Geoffrey Marshall to consider the methodology of the trial.[736] It was agreed that control cases were essential for the main group in the study, but unnecessary in the meningitis group.[737] The view was that a double blind method was needed for the main group and that shortage of supplies would make it easier to refuse streptomycin to control cases.

Thus, most of the elements of the subsequent trial were in place by the Sunday. The study groups and study centres had been chosen; and the need for a double blind control group methodology agreed. Up to this point there is no evidence that Bradford Hill had played any role in the discussions.

The conference on streptomycin continued on August 27th 1946, this time attended by Bradford Hill. The main decision of the meeting was agreement of provisional study centres. The choice was made difficult by the issue of transferring patients from hospitals under London County Council authority to those outside its control. The provisional list for the main group was the Brompton Hospital (voluntary sector); the County Sanatorium

[735] Other groups were tuberculous meningitis in children; acute tuberculous broncho-pneumonia in children under 5; acute miliary tuberculosis at ages 15-25. (FD1 6756 MRC Clinical Trials of streptomycin TCTC1 paragraph 3).

[736] FD1 6756 MRC Clinical Trials of streptomycin TCTC2. The meeting consisted of Hoyle, Marshall, and Scadding from the streptomycin conference, and Professor RV Christie, Secretary of the MRC Penicillin Clinical Trials Committee.

[737] During 1947, as the meningitis trial progressed, the clinicians instituted an alternate case scheme of randomisation to determine if both intrathecal and intramuscular injections of streptomycin were necessary in cases of tuberculous meningitis. (FD1 6756 Streptomycin in Tuberculosis Trial. Working Sub-committee TCTC 30. Minutes of the fourth meeting of pathologists, 23rd June 1947 p3).

at Harefield (Middlesex County Council) and Colindale Hospital (London County Council).

Concerning British supplies, it appeared that Boots would be able to produce a small amount of streptomycin starting in the Autumn of 1946. The amount would be sufficient to treat 6 patients in the main study group in the first month of supply, and 12 patients thereafter.[738] The conference considered this amount as suitable for a pilot study. D'Arcy Hart's recollection of the meeting, in a letter to Landsborough Thomson was that:

> *The most important outcome of last night's meeting (at which the Ministry of Supply people were there) was that in all probability there will be a pilot trial in the Autumn followed by the main trials in the new year. This procedure is dictated by the supply position but may have advantages in that we shall learn a lot of lessons from the pilot trial"*[739]

At the meeting, a Dr Madigan gave a brief report on current trials of streptomycin in Kent. Madigan's presentation highlights the fact that the MRC was not the only organisation seeking to test streptomycin, and that there was, by mid 1946, a small unofficial market for British and American streptomycin.

No further meetings of the conference took place. The arrangement of supplies had now placed the MRC, rather than the Ministry, at the forefront of dealings with the Ministry of Supply. Correspondence forwarded by the Ministry to the MRC shows that the Ministry was losing touch with the supplies position.[740] Two items of correspondence at the PRO

---

[738] Using dosage figures supplied by Corwin Hinshaw which suggested 2-3 gms daily for periods of 3-6 months. (FD1 6756 Hinshaw to Hart August 26th 1946).

[739] FD1 6756 D'Arcy Hart to Landsborough Thomson 28th August 1946 D'Arcy Hart later predicted that the pilot study would be able to begin in November. (FD 1 6756 D'Arcy Hart to Landsborough Thomson 4th September 1946).

[740] FD1 6756 Marchbank to Mellanby 29th August 1946. See also FD1 6764 Landsborough Thomson to Hale 11th November 1946, a minute concerning the purchase of a large consignment of American streptomycin. In paragraph (8) where the suggestion is made that the Ministry of Health pay for the consignment, Landsborough Thomson notes that 'The Ministry of Health have not so far been consulted on the point'.

suggest that the policy of organising a trial was succeeding in putting the MRC at the centre of streptomycin work. A request from the Welsh National School of Medicine to be included in the trials was sent to the Ministry, who forwarded it to Landsborough Thomson. An offer to undertake trials was received directly from the Institute of Child Health. Both requests show that a series of formal and informal contacts were all pointing to the MRC as the focus for work on streptomycin.[741]

The formation of a trial committee was discussed towards at end of September. Despite D'Arcy Hart's wish to avoid publicity by not including the word streptomycin, it was called the Streptomycin in Tuberculosis Trials Committee, and met for the first time on 21st November 1946.

During October 1946, it emerged that the Ministry of Supply had secured 50kgs of American streptomycin at cut-price for the MRC. As far as the record shows, the offer was totally unexpected. As a result, Landsborough Thomson had to secure the resources for the purchase from the Treasury.

In the minute to the Treasury, Landsborough Thomson made no reference to any research other than the proposed MRC trial. In his argument the proposed clinical trial plays several roles. The trial will provide the medical profession with 'soundly based knowledge of its value in different conditions'. It will absorb supplies and thereby put off 'the question of permitting their purchase by agents for sale in this country. Equally, the question of releasing British supplies for sale will arise if these are not being wholly

[741] FD1 6756 Jameson (MoH) to Landsborough Thomson 18th October 1946; Moncrieff (Director, Institute of Child Health) to Mellanby 9th November 1946.

taken up for clinical trial'.[742] Thompson offered the trial to the Treasury as a way of dealing with the more embarrassing financial aspects of streptomycin in return for Treasury funds to make the purchase of streptomycin..

The first meeting of the Streptomycin in Tuberculosis Trials Committee must have been a pleasant affair. As a result of Landsborough Thomson's minute to the Treasury, provisional approval for the purchase had been granted. Considerable progress was made at the meeting. The entry criteria to the main study group was changed, so that it now consisted of 'acute, rapidly progressive, bilateral pulmonary tuberculosis of recent development, unsuitable for collapse therapy, bacteriologically proven, age limits 15-25'. The study centres were confirmed as Brompton, Harefield and Colindale. The idea of a pilot study was dropped, and it was decided to wait until sufficient supplies for the main study arrived in January. The use of controls in the main study group was confirmed as essential. It was agreed to convene a working sub-group of the pathologists involved in the trial.[743] It was decided that a press notice should be issued, indicating that trials were to happen, and that no supplies for private use could be considered.[744] By the 10th of December Daniels had drawn up a set of case record sheets for the trial.[745] In early January 1947, the American streptomycin arrived safely.[746]

---

[742] The quotes are taken from FD1 6764 'Streptomycin' Minute from Landsborough Thomson to E Hale, Assistant Secretary, Supplies Finance, HM Treasury 12th November 1946.

[743] The group met first on January 2nd 1947 at the Central Public Health Laboratory at Colindale. (FD1 6756 TCTC 11. Streptomycin in Tuberculosis Trials, Working Sub-committee).

[744] FD1 6756 Streptomycin in Tuberculosis Trials Committee. Minutes of the first meeting TCTC 8 21st November 1946.

[745] FD1 6756 Daniels to Wilson 10th December 1946.

[746] FD1 6756 D'Arcy Hart to committee members 6th January 1947. The forms were printed for the MRC by G Pulman & Sons, 24 Thayer Street, London W1.

The second meeting of the Committee took place at the LSHTM on 18th April 1947.[747] It was reported that the age limit in the main study group had been raised to 30. Daniels reported that admission of cases had begun on 20th January 1947. To date 45 subjects had been admitted to the main study group. A preliminary analysis showed that the condition of 5 members of the control group had deteriorated, while none of the study group had. It was agreed that further supplies of American streptomycin should be bought if possible, as it appeared that British streptomycin production was slower than anticipated.

The meeting also agreed to extend the trial, following 'considerable correspondence suggesting various extensions of the categories under investigation'. A small group met on 22nd April and agreed to extend the trial to cases of miliary tuberculosis without meningitis, tracheo-bronchitis, and laryngitis complicating pulmonary tuberculosis. None of these would have control groups. Four further proposed extensions were deferred, pending the next American consignment. The extension into eye cases was deferred pending further enquiries. Only one proposed extension, that of genito-urinary tuberculosis, was rejected.[748]

The only reference at the PRO to the method used to allocate patients to study and control groups, the now famous 'method of sealed envelopes' occurred in a discussion about how to respond to an appeal on behalf of a girl to be included in trials.[749] The girl's father was able to obtain streptomycin through contacts in America and wished her to be treated at

[747] FD1 6756 Streptomycin in Tuberculosis Trials Committee. Minutes of the second meeting TCTC 25 18th April 1947.

[748] The conclusions of the sub-group on extensions were included in the minutes of the main committee.

[749] The relevant correspondence is a handwritten letter from the girl's doctor to the MRC; internal correspondence at the MRC, and the resultant letter from the MRC to the doctor. With the exception of FHK Green, who handled the enquiry, names and other details are omitted from this reference because of the possibility of identification. The letters are however available at the PRO. The correspondence took place in the early Summer of 1947.

one of the study centres. Looking for a way to turn down the application, a colleague of D'Arcy Hart pointed out to Green that:

> *While you may not be able to pass this on to Dr A--- the strongest point against any possible acceptance of his case is that with the control system we dare not take isolated cases of this kind – we don't decide whether the case is to be a treated one or a control case. Professor Bradford Hill has worked out a sealed envelope system for us and we take our sealed instructions in rotation after the case has been accepted for admission to the trials. Dr A---'s case, were it otherwise quite suitable, might be a "control" case which would mean a bed alone in whichever centre was decided upon and I scarcely think that would improve the position for his patient!*[750]

The author added 'it will be nice when we do not have to answer negatively these many enquiries'. Green used the longstanding argument that the value of streptomycin was unclear, which as can be seen from the quotation, was by now untrue.[751]

During the Summer of 1947 the Committee agreed to shorten the length of treatment from 6 months to 4 months, following a discussion between D'Arcy Hart, Scadding, Houghton and Daniels with Hinshaw at which it became clear that the 6 month scheme had been abandoned some time ago in America.[752] Admission to the trial ended in September 1947.

The third meeting of the Committee took place on 5th November 1947. Several points of interest came from the meeting. Firstly, that despite extending the trial to several other hospitals, recruitment to the trial was only half the number initially considered necessary. As a result, 17 kgs of the original consignment of 50 kgs was unused, in addition to the

---

[750] FD1 6756.

[751] D'Arcy Hart later sent details of how to use streptomycin to the doctor. (FD1 6756).

[752] FD1 6756 D'Arcy Hart wrote to the committee members to ask their approval on 23rd July 1947, at the point when some patients were just past the four month stage.

further supply of 30kgs of American streptomycin the MRC had received in late 1947. With a large quantity of streptomycin in hand, the committee were able to initiate several extension studies. In the area of pulmonary tuberculosis alone these included a comparison of different dosage regimes of streptomycin in pulmonary tuberculosis, a trial of streptomycin as an adjuvant to chest surgery, and a trial on acute spreads of tuberculosis in previously treated cases.

Useful as these trials might be, the primary purpose for undertaking them was to retain first call on the streptomycin under its control. In the case of case of streptomycin as an adjuvant in chest surgery, correspondence between George Mason and Edward Mellanby in early 1948 shows that both men regarded streptomycin supplies and streptomycin trials as controlled by the MRC.[753] Further evidence of the importance of streptomycin to the MRC arises out of its handling of tuberculous tracheo-broncial ulceration. This study was the first extension granted by the Streptomycin in Tuberculosis Trials Committee in response to pressure from clinicians, and had quickly produced outstanding results.[754] In February 1948 D'Arcy Hart argued that although the trial had given an answer, it would be unfavourable to the MRC if it ceased treating patients with ulceration and allowed the Ministry of Health to become responsible for the release of streptomycin for this indication:

[753] The correspondence began in January 1948. (FD1 6757 Ogilvie to Mason 22 November 1947; Mason to Mellanby 3 January 1948; Mellanby to Hart 6 January 1948; Mellanby to Mason 6 January 1948Mason to Mellanby 14th January 1948).

[754] No control group was used. The result of treatment for two months or more with streptomycin was that 15 of the 17 study cases showed either improvement or complete healing. (FD 1 6757 Agnew to Thomson 10 February 1948).

> *A better plan would, I think be for the Ministry to supply us with extra streptomycin to carry on this abridged "public service" for them during the next few months, until we have our report on the wider pulmonary groups*[755]

In a carefully drafted letter of 19th February 1948, Green put this suggestion to Dalrymple-Champneys, to which he received a positive response.[756] At its next meeting the Committee noted the Ministry's gesture with approval, and halved the daily dose, presumably to conserve its supply and thereby extend the period over which it could run the tracheo-bronchitic ulceration treatment programme.[757]

At the fourth meeting of the Committee, a draft report on the treatment of tuberculous meningitis was discussed. It was decided, after long debate both at the meeting and afterwards that the meningitis report should be offered to the Lancet and the pulmonary report to the BMJ.[758] It was reported that Dr Daniels had called in the case-notes, with a view to writing the report.[759] It was agreed that 'there should be an independent panel of specialists to examine (separately) the X-ray films, with a view to their statistical

[755] FD1 6757 D'Arcy Hart to Thomson 3rd February 1948.

[756] FD1 6757 Green to Dalrymple-Champneys 19th February 1948. The copy of the letter retained by Green is annotated 'D-C answered this verbally "yes" 11/3/48. The Ministry delayed the transfer until June 1948. In June, Green suggested a further extension to the arrangement, to which Dalrymple Champneys agreed. (FD1 6757 Green to Dalrymple Champneys 3rd June 1948, and response Dalrymple Champneys to Green 17th June 1948).

[757] FD1 6757 Streptomycin in Tuberculosis Trials Committee. Minutes of the fourth meeting TCTC 52. 12th March 1948 paragraph (5).

[758] It is frustrating that so little of Bradford Hill's professional life is known. He is a shadowy presence, even at the meetings he attended. In this instance it may have been Bradford Hill's opposition to publishing the pulmonary report in the Lancet that was decisive. Lock reports that Bradford Hill quarrelled with Robbie Fox, editor of the Lancet, and afterwards transferred his allegiance to the BMJ (Lock 1994 p85) Bradford Hill's bibliography neither confirms or rejects this since he continued to publish in both journals throughout the 1940s and 50s. (Doll 1993).

[759] Daniels was at the time based at the BMA Headquarters, Tavistock North, Tavistock Square, London. (FD1 6757 handwritten note from Daniels to Green, received by MRC 15th June 1948, and D'Arcy Hart 17th September 1948, minute to Committee).

analysis'.[760] The committee also discussed a proposal to undertake trials of PAS. The minutes report that 'the Committee was definitely opposed to any scheme which would interfere with or replace the research now proceeding in a number of centres'.[761] On the face of it this reluctance is odd, since the efficacy of PAS was no more or less uncertain than streptomycin.[762] The explanation is found in the minutes. PAS was already in use in Britain. There was therefore no particular benefit to the MRC in organising trials. It was decided to let others, namely the Tuberculosis Association Research Committee, pursue research on PAS.[763]

The fifth meeting of the Committee took place on 6th September 1948. By that time, a slot in the BMJ had been arranged for October 30th 1948, and the main business of the meeting was a discussion of the draft paper, written by Daniels and D'Arcy Hart. The draft was discussed and amended 'page by page'.[764] The amendments are not recorded. At that time the discussion and conclusion had not been written. This was written and circulated after the meeting. The only surviving comment on the conclusion and summary on file is from GS Wilson, who wanted the strength with which

[760] FD1 6757 Streptomycin in Tuberculosis Trials Committee. Minutes of the fourth meeting TCTC 52. 12th March 1948. Paragraph (9).

[761] FD1 6757 Streptomycin in Tuberculosis Trials Committee. Minutes of the fourth meeting TCTC 52. 12th March 1948. Paragraph (8).

[762] For the use of PAS in the treatment of tuberculosis see Watkin Edwards 1950.

[763] FD1 6757 Streptomycin in Tuberculosis Trials Committee. Minutes of the fourth meeting TCTC 52. 12th March 1948. Paragraph (9).

[764] FD1 6757 Streptomycin in Tuberculosis Trials Committee. Minutes of the fifth meeting TCTC 63. 6th September 1948 Paragraph (2).

recommendations were made to be made more tentative , for example by qualifying 'it has a place in the ….' With 'it may have a place …'[765]

The results of the trial were published in the BMJ on October 30th 1948, and is reproduced in Appendix 1. The group effects of treatment with streptomycin were clear (Table I in the paper). Six months after entering the study, patients who had been treated with streptomycin had fared much better than those given bed-rest alone. There were significantly fewer deaths in the treatment group (referred to as the s group), and among those who survived 51% had shown considerable improvement, compared to only 8% in the control group, as measured by changes in X-ray appearance.

However, the purpose of the trial had been to learn more about which types of tuberculous patients benefited, rather than simply to confirm what was by then a widely held view.[766] As a result, the bulk of the paper is filled with what is now called sub-group analysis, and with description of individual cases where streptomycin had failed, or bed rest succeeded. Sub-group analysis is best described by way of an example. The paper showed that streptomycin was effective for the study population. However, further analysis showed that most of the improvement occurred in patients who entered the study with a normal evening body temperature. Those who entered the study with a raised temperature were as likely to recover from bed rest as they were from treatment with streptomycin.

[765] FD1 6757 Wilson to Daniels 23rd September 1948. The conclusions in the final report were made more tentative, and the paper concluded 'these conclusions are of necessity lacking in precision; much organized work is yet required to determine the precise indications of streptomycin'.

[766] The paper makes little reference to existing research on streptomycin. It ignores a controlled trial, possibly randomised on streptomycin versus bed-rest, published a year before, which showed clearly that streptomycin was effective in treating tuberculosis. (Jenkins et al 1947).

Modern reports of clinical trials rarely include details of individuals. And sub-group analysis is usually regarded as a post hoc attempt to extract further results that had not been hypothesised at the outset.[767] Had the streptomycin trial simply reported that streptomycin was better than bed-rest, the report might well have been regarded as inconsequential. The case notes and sub-group analyses were essential to the streptomycin trial because they added a greater sense of clinical relevance to the results.

Guidance for the profession

Also on the agenda at the fifth meeting was an invitation to send representatives to a sub-committee on streptomycin of the Ministry of Health Standing Advisory Committee. With the results of several trials now available to the Committee, it was able to draw up a draft pamphlet for its fifth meeting, advising doctors on the appropriate use of streptomycin. It is this pamphlet, more than anything else, which highlights the achievements and limitations of the trial on pulmonary tuberculosis.

The section on indications for pulmonary tuberculosis begins:

> *The facts given indicate the importance of resisting pressure to try the effect of this new drug in all the many forms of pulmonary tuberculosis. In the present state of knowledge, the simplest general criteria recommended for selection for streptomycin might be as follows: 'pulmonary tuberculosis in which the lesions requiring treatment are of recent development, progressive, and unlikely to benefit from conventional methods (e.g. bed-rest and/or collapse therapy) alone'*[768] [emphasis in original]

[767] One of the problems with the post-hoc identification of sub-groups in which differences are apparent is that they break Fisher's rule about relevance in the reference set. Statistical tests do not therefore carry the same conviction as they do on the main groups in the study.

[768] FD1 6757 Streptomycin in the treatment of tuberculosis. Draft pamphlet TCTC 61 3.9.48 p4.

While this offered a reasonably clear line on which patients with pulmonary tuberculosis might benefit from streptomycin, it in fact added very little to what was known about the indications for streptomycin. The advice repeated the inclusion criteria to the streptomycin trial, but these criteria had been chosen precisely because it was already known in 1946 that this group was likely to benefit.[769]

The guidance concludes with an attempt - unwarranted by the design or conclusions of the trial - to set out a plausible scenario in which streptomycin might by integrated with existing regimes:

> *Streptomycin should not be used as the only therapeutic measure in pulmonary tuberculosis. Indeed its major role may be to make possible the use of collapse procedures which, in its absence, would have had to be delayed or never performed.*[770]

After the trial

What evidence is there that the streptomycin trial reduced uncertainty about its effectiveness, or in other ways contributed to the rational use of streptomycin? While the main result of the trial was evident, the results of the sub-group analysis, if given credence, tended to make the overall message from the trial far less clear. Body temperature, sedimentation rate, presence of tubercle bacilli in the sputum, development of resistance – all appeared to be relevant to the outcome. To a skeptical clinician the inclusion of accounts of individual cases within the report added to the sense of endless individual variation in the response to streptomycin.

[769] The fact that the trial produced a clear, statistically significant result on a sample size half that judged necessary is testimony to the care with which the entry criteria to the trial were constructed.

[770] FD1 6757 Streptomycin in the treatment of tuberculosis. Draft pamphlet TCTC 61 3.9.48 p4.

Restriction of the extent to which trial results could be applied to individual patients was highlighted almost immediately. A letter from Dr Bernard Freedman of the Dulwich Hospital concerned the rational use of streptomycin to treat tracheobronchitis in advanced pulmonary tuberculosis. Although an uncontrolled MRC series had shown streptomycin to be effective in healing tracheal ulceration in younger, earlier cases, Freedman's point was that the value of streptomycin in later cases remained to be established. The MRC passed the case on to the Ministry 'as it is now the concern of that Department, rather than of the Council.'[771]

The subsequent use of streptomycin in the UK has not been investigated here. But several facts suggest that the use of streptomycin as a treatment for tuberculosis may not have been extensive. Firstly, American and British studies had revealed side-effects to streptomycin treatment and the tendency for resistant strains of *M tuberculosis* to develop. In his review of the history of treatment for tuberculosis, Sir Robert Young, consultant physician to the Middlesex and Brompton regarded these as a considerable brake on the effectiveness of streptomycin:

> *[Following its discovery] disappointment was felt when it was found that this substance had toxic effects, particularly on the auditory functions, and that the tubercle bacilli frequently developed resistance to the drug after comparatively short periods of treatment.*[772]

In a similar vein, a 1950 review of the streptomycin trial summarised its findings as follows:

[771] FD1 6757 Green to Freedman 28th September 1948.

[772] Young 1956.

> *There was a striking difference between the treated and control cases after three months, but after twelve months there was much less to choose between them.*[773]

Secondly, the rapid introduction of more effective drugs for the prevention and treatment of tuberculosis.[774] Thirdly, a continuing belief among tuberculosis specialists that treatment was multi-faceted:

> *Streptomycin turns out to be one tiny shade of colour in the palette of the physician and not the great spectrum we had hoped.*[775]

In summary therefore, a brief survey of the literature on streptomycin post-1948 suggests that its impact in the British specialist journals was not extensive. Referring to streptomycin and PAS, an editorial in the journal Tubercle in 1950 stated that: 'The truth is that we do not know which patients with active disease should be denied these drugs.'[776]

Conclusion

From the perspective of the history of ideas, the trial of streptomycin for pulmonary tuberculosis has come to be regarded as a turning point in the struggle to find rational ways to evaluate the effectiveness of therapies. After this trial, it is argued, the proper way to conduct clinical trials of therapies was clear to all but the most refractory.

In this chapter I have considered the streptomycin trial from another perspective, that of its day to day organisation as revealed by the MRC archives. My purpose has not been to

---

[773] Mullard 1950.

[774] Principally, Isoniazid, introduced in 1952. For a review of the major trials of Isoniazid a chemoprophylactic agent, see Ferebee 1970.

[775] Harley Williams 1950, quoted in Watkins Edwards 1950

[776] Anon 1950 (Tubercle).

challenge the special status of the trial, or its claims to be the first rigorously controlled clinical trial, although I hope that by now it is clear that these claims are less certain than the secondary literature suggests.

Although I have discussed some of the non-scientific factors associated with the trial, it is not because I wanted to explain the adoption of randomisation in clinical trials in terms of social or cognitive factors. MacKenzie has tried such an analysis, when he made an explicit link between the adoption of particular statistical techniques and the 'social interests of the British professional middle class'.[777] In a very limited way, Bradford Hill undertook this sort of analysis when he claimed that the shortage of streptomycin created the circumstances in which he was able to introduce random allocation. MacKenzie's study of the history of statistics was an enormously ambitious attempt to disturb the order of scientific knowledge at its strongest point. It was probably bound to fail, and it did fail, as reviews and subsequent corrections have shown.[778]

Rather than identify the sociological factors that shaped scientific knowledge, I have described the clinical trial of streptomycin unfolding within a domestic drama, whose leading players were the MRC, the MoH, the public, the British medical profession, and the emerging National Health Service. I have tried to portray the streptomycin trial as a drama that shaped and redefined the relationship between the various actors. The trial methodology does not so much reflect the interests of the groups involved as help to shape the relationships between the groups.

[777] MacKenzie 1981 p221.

[778] Notably, Magnello 1993, who set the maths and the historical record straight, and Yearley 1982, who raised some uncomfortable questions about the claims from a sociological point of view.

The MRC trial was a way of generating knowledge about the efficacy of streptomycin. However, the trials were at the same time a means to solve a set of organisational problems concerning the relationship between the MRC and Ministry of Health; MRC and the public; and possibly, the MRC and the emerging NHS.

Bradford Hill was right, I believe, to claim that this was the first strictly controlled clinical trial. He was right to claim that he devised a method of random allocation that had not been used before in clinical trials in Britain, and perhaps anywhere. The method of sealed envelopes immediately commended itself to clinical trialists, as a way of controlling the operation of clinical trials in hospitals

He was wrong though to suggest that he was responsible for imposing methodological rigour on the Committee, and wrong to identify the shortage of supplies as the reason why the methodology was acceptable. In fact, all the evidence suggests that Mellanby, D'Arcy Hart and the other members of the Streptomycin Conference worked out the main elements of the methodology at a very early stage. Bradford Hill was also wrong to suggest that production shortages and a crisis in British dollar holdings were responsible for methodological innovation. The evidence shows that there was no streptomycin at all when the MRC first conceived the trial. And later, following the American consignment, the Committee had more than it needed, so much so that the main study could not absorb the supply allotted to it.

The impetus for organising a trial in the middle of 1946 was the will of the MRC to be central to the roll-out of a powerful new drug in Britain. The trial served to confirm that the MRC was the proper body to generate knowledge about the effectiveness of drugs for

the NHS. It showed the MRC how to create a network of trial centres within the emerging NHS, and how it could draw on NHS patients while retaining control for itself.

Particular elements of the methodology can be interpreted as further and more specific ways to regulate the access of hospitals and doctors to streptomycin, and to draw together patients in diverse hospital settings while bypassing organisational and medical authority. Randomisation was a way to construct study and control groups in such a way that statistical tests could be applied to patient outcomes. It was also a way to extend the reach of the MRC into the hospitals participating in the trials, and maintain control over the distribution of streptomycin.

Bradford Hill later re-organised the events that shaped the streptomycin trial to make it appear that shortage of streptomycin and shortage of dollars made it possible to overcome objections to randomisation. The truth is much more prosaic. The 'method of sealed envelopes' was acceptable to the Committee only in so far as it provided extra control over the clinicians conducting the study.

Thus there are two types of explanation for the MRC streptomycin trial. The better known places the trial at the end of a series of developments in trial methodology, and focuses on the trial as a new solution to the problem of how knowledge can be secured. The other explanation concerns the solution to a set of material problems concerning the position of the MRC, in which the method of securing knowledge is part of the tactics of securing a territory.

Chapter six

Conclusion

Two questions were posed at the outset of this study. Why did randomisation not take place in medical research before the 1940s? And, why did it become possible to allocate patients at random around 1948? In the conclusion, I would like to draw together some answers to these questions on the basis of the findings of this study.

Why wasn't randomisation used before the 1940s?

To begin with, it is clear that random allocation of patients to treatment and control groups could have taken place before the 1940s, and indeed did. As Hacking has shown, the earliest instances of random allocation in experimental research took place in the late nineteenth century. And although RA Fisher never described an RCT of medical therapy, Karl Pearson did, when he suggested the use of schemes of random allocation for trials of typhoid anti-toxin and tuberculin in the early twentieth century. Intellectually therefore, randomisation was possible in clinical trials from early in the twentieth century.

A possible explanation for why randomisation was not systematically adopted is that doctors were not sufficiently interested in clinical trials. This explanation is unconvincing however, since there is a rich history of clinical trials in the late nineteenth and twentieth centuries. The Index Catalogue of the Surgeon General's Office, the forerunner to Index Medicus, contains 10 pages of small-print references to clinical evaluations of diphtheria anti-toxin.[779]

[779] Surgeon General's Office 1899.

Chapter two of this study has suggested that doctors were intensely interested in the possibility of reliable knowledge about therapeutics in the 1860s.[780] At that time clinical trials, albeit of a simple type, were instigated by the British Medical Association in order to achieve the twin goals of improving knowledge and uniting the fractious body of medical practitioners. The BMA used, and thereby consolidated, its dispersed membership to accumulate data that would resolve a series of problems related to the subjective independence and variability of practitioners' judgements. In creating 'the view from nowhere' or 'observations without an observing subject'[781] the BMA was not seeking to disavow subjectivity altogether, although some members clearly thought it was trying to. Rather it sought to create a new type of doctor whose point of reference was no longer the subjective judgement of the isolated practitioner, but the well-stocked collective consciousness of a profession.

At about the same time, William Guy described the design for a quasi-randomised controlled clinical trial. However, in the view of Guy such designs were applicable only when the practitioner or the therapy was untrustworthy. In organising its therapeutic enquiries, the BMA emphasised the collective nature of the enterprise. Data were to be aggregated and shared among trusted members of the BMA. It would have made no sense to introduce the element of ignorance entailed by assigning patients to treatments without first matching patient to treatment. The BMA sought to combine the considered experience of clinicians. Just as stable averages emerge from combined data, the act of

[780] I argued also that the interest in objective knowledge among doctors in a sense preceded the entry of laboratory science into medicine. As Foucault (Foucault 1972 p53 and p181) and Harley Warner (Warner 1995 Osiris) have warned in different ways, the conventional historiography that regards the laboratory as undermining the legitimacy and then revolutionising clinical medicine risks underplaying the extent to which clinical medicine established its own relations with the laboratory.

[781] The phrases come from Porter 1995 and Swijtink 1989 respectively.

aggregation was considered to be sufficient to arrive at stable knowledge about therapeutics.

By contrast, the laboratory based therapies created by Ehrlich, Koch and Pasteur in the late nineteenth century, and by Wright in the twentieth, had little need for statistics in creating their knowledge. Better than any statistics, the laboratory created specific scientific therapies based on physiological therapeutic principles such as inner disinfection or chemotherapy. Much of the testing of drugs could be done on experimentally induced lesions in animals. For the laboratories, clinical trials were connected not so much with establishing the therapeutic value of a drug as with promoting its use and distributing it in a controlled manner.

When laboratories sought to extend their reach they did so by creating links with pharmaceutical manufacturers. The career of Henry Dale illustrates the close relationship between medical science, state and industry that was possible around the turn of the twentieth century. At the heart of the relationship between laboratories, state, and drug companies was biological standardisation. Standardisation suited the companies (it gave them a reliable standard scientific product to market), the laboratories (they could use their facilities to generate income), and governments (it gave them regulatory control over the pharmaceutical sector).

The MRC played a leading international role in the drive to standardise key pharmaceuticals in the 1920s. However, the orientation towards biological standardisation left MRC in a weak position when it came to supporting the British pharmaceutical industry. In an effort to mediate between pharmaceutical companies and British clinicians, the MRC organised clinical trials throughout the 1930s. Despite being

encouraged to use random allocation in comparative clinical trials, the MRC adopted a style of research that best suited its organisational structure. A central committee handled applications to have substances tested. If accepted, trusted researchers were allowed to do pretty much as they pleased by way of clinical tests.

The answer to why randomisation wasn't used before the 1940s is that it suited no one's purpose to allocate patients at random in clinical trials.

Why was randomisation possible by 1948?

Nothing that came from the Therapeutic Trials Committee suggested that the MRC would be innovative in the area of clinical trials. Yet in 1946 the MRC implemented the first rigorously designed RCT. Credit for achieving this is usually given to Austin Bradford Hill. Greenwood and Bradford Hill had certainly done much in the 1930s to make statistics more relevant within the MRC, following the failure of Brownlee to do so in the 1920s.[782] However, the unpublished records of the committee that organised the 1946 trial of streptomycin strongly suggest that Bradford Hill had little direct influence over the research design. As D'Arcy Hart has suggested on several occasions, the methodology of the 1946 streptomycin trial is less innovative than most commentators have claimed, and that Bradford Hill's role has been over-stated.[783] All of the methodological elements of the trial were suggested before 1946 in situations or publications that would have been readily accessible to clinical scientists associated with the MRC. D'Arcy Hart's account receives further support from the absence of any

[782] They did it in two ways Firstly, they completed a series of studies in which epidemiological techniques were put into laboratory settings. Secondly, Bradford Hill published an acclaimed series of articles on medical statistics in the Lancet.

[783] D'Arcy Hart 1991 and interview, 1996.

evidence that members of the committee had difficulty accepting random allocation to study and control groups.

In contrast, an administrative perspective on the streptomycin trial suggests why the MRC required little convincing to adopt a randomised controlled study design in 1946. It allowed the MRC to solve a series of problems associated with streptomycin once it had seized the initiative for the handling of streptomycin from the Ministry of Health in 1946. Firstly, the notion of a clinical trial in which no one knew who would receive the drug gave the MRC a ready response to enquirers wanting to be entered into the trial. Secondly, and most importantly, Bradford Hill's method of sealed envelopes offered a way of placing the MRC at the centre of the network of trial hospitals and clinicians who were otherwise outside its direct control. The scheme of randomisation gave the MRC control over the distribution of the drug in these settings.

The standard account is that given considerable uncertainty about the value of streptomycin, and great shortage of supply, Bradford Hill was able to overcome any resistance and introduce the methodology that he had for so long wanted to. Is the standard account wrong? The argument rests on three assumptions. Firstly, that there was resistance to randomisation and control in clinical trials. Secondly, that there was a large degree of uncertainty about the value of streptomycin. And thirdly, that only limited supplies of the drug were available.

The evidence presented here suggests that by 1947 there was little resistance to the idea of random allocation to treatment and control groups in clinical trials. It was a technique that had been used on several occasions. The novelty introduced by the streptomycin trial was the degree of central co-ordination that accompanied the conduct of a clinical trial.

Concerning what was known about the effectiveness of streptomycin, there is no doubt that the MRC over-stated the degree of uncertainty concerning the value of streptomycin, so much so that the American government protested.

The claim that streptomycin was in short supply is almost certainly true, until late 1946. By the time the trial began it would have been possible to obtain greater supplies from America. And had the Ministry of Health been in control of streptomycin at that point, distribution to specialist centres on the basis of known American experience might have been possible. Had there been the political will it is possible also that American technology and experience with deep-vat production could have been imported to Britain in 1947. This did not happen of course, and British production did not reach American levels of efficiency until 1949.

The evidence concerning shortage of supply, therapeutic uncertainty, and resistance to innovation provides little support for Bradford Hill's subsequent account of the genesis of the trial methodology. It is possible that the historical record does not adequately reflect Bradford Hill's contribution in July 1946. Nonetheless, it is now time to accept that the methodology of the streptomycin trial was also the product of administrative contest and the solution to a series of practical organisational difficulties faced by the MRC in the 1940s.

Further research

Against a background of growing interest in the history of clinical trials, this study has mapped out some of the territory of the British experience between 1858 and 1948. Following Marks' study of certain clinical trials in twentieth century America there has been a tendency to regard the advancement of clinical trials and reform of the profession

as almost synonymous.[784] While the present study provides some support for this view in the case of the BMA in the nineteenth century, factors other than reform come into play in the twentieth century.

The study of therapeutics, clinical trials, and medical knowledge making in general will be restricted if it is portrayed largely in relation to medical reform. I have portrayed clinical trials as a way of mediating relationships between the MRC, the British State and the pharmaceutical industry in the 1930s; and between the MRC, MoH and the public in the 1940s. Foucault, and more recently Latour and Shapin & Shaffer, have suggested ways in which discursive practices operate. This framework might be applied to a series of more detailed studies on clinical trials. One topic that appears intriguing is the changing relationship between State and pharmaceutical industry during and after the Second World War. Other topics I would have liked to pursue but have been unable to include: Major Greenwood, who surely deserves a biography; the critique of RCTs made by Lancelot Hogben in the 1950s; the attitude of pharmaceutical companies towards RCTs in the 1950s; the relationship between pharmaceutical companies and clinicians in the 1930s; the impact of the 1925 Therapeutic Substances Act; and the history of tuberculin dispensaries.

[784] For example Porter 1995 p204.

Appendix 1

The MRC streptomycin trial

Appendix 2

Publications on the history and philosophy of statistics 1910-1989

The list of publication below comes from several sources, including: the principle bibliographic tool for the history and philosophy of science, the Isis cumulative bibliography and supplements,[785] the catalogues of major libraries; and the bibliographies contained in theses and books.

The creation of this list was made difficult because judgements had to be made in some instances concerning whether a text was about the underlying principles of statistics rather than being a 'tutor text'. I have included one (Folks 1981) since it uses the history of statistics to introduce statistics to undergraduate students. The contributions to one multi-authored text have been listed individually (Daston et al 1987); while those in another have not (Owen 1976), reflecting the differential importance of the collections. Where an item was originally published in a non-English language I have used their date of publication in English.

Inevitably, this list offers only a partial listing of the relevant material. Nevertheless, it offers a fair picture of the growing scholarly interest in the basis of statistics. Table 1 shows the number of publications in each decade.

[785] Whitrow 1971-84; Neu 1980-85; Neu 1990.

Table 2.1:publications in English on the history and philosophy of statistics 1910 – 1989

Decade of publication	Number of publications
1910-19	2
1920-29	9
1930-39	17
1940-49	13
1950-59	32
1960-69	44
1970-79	82
1980-89	126
TOTAL	**325**

Source: various (see Appendix 2 para 1)

1911
Yule GU. An introduction to the theory of statistics. London: Griffin.

1912
Edgeworth FY. On the use of the theory of probabilities in statistics relating to society. J Roy Statist Soc 76:165-193.

1921
Keynes,JM. A treatise on probability. Macmillan, London.
Pearson,K. The history of statistics in the 17th and 18th centuries against the changing background of intellectual, scientific and religious thought. Lectures by Karl Pearson given at University College London during the academic sessions 1921-1933. Charles Griffin (published in 1978 but listed here).

1922
Fisher,RA. On the mathematical foundation of theoretical statistics. Phil. Trans. of the Royal Society A 221, 309-368.

1924
Pearson,K. Historical note on the origin of the normal curve of errors. Biometrika 16, 402-404.

1925
Fisher,RA. Statistical methods for research workers. Edinburgh, Oliver and Boyd.

1926
Archibald,RC. A rare pamphlet of Moivre and some of his discoveries. Isis 8, 671-684.

1928
Campbell,NR. An account of the principles of measurement and calculation. Longman Green, London.
Greenwood,M. Graunt and Petty. Journal of the Royal Statistical Society 91, 79-85.

1929
Walker,HM. Studies in the history of statistical method. With special reference to certain educational problems. Williams and Wilkins, Baltimore.

1931
Ramsey,FP. The foundations of mathematics and other logical essays. RKP, London.

1932

Comp,BH. Definitions of probability. American Mathematical Monthly 39, 285-288.
Westergaard,HL. Contributions to the history of statistics. King, London.

1933
Neyman,J; Pearson,ES. The testing of statistical hypotheses in relation to probabilities a priori. Proceedings of the Cambridge Philosophical Society 29, 492-510.

1934
Irwin,JO. Some aspects of the development of modern statistical method. Mathematical Gazette, 18-34.
Struick,DJ. On the foundation of the theory of probability. Philosophy of Science 1, 50-70.

1935
Pearson,K. Statistical tests. Nature. 136, 296-297,500.
Willcox WF. Definitions of statistics. Rev Int Inst Statist;3:388-99

1936
Pearson,K. Method of moments and method of maximum likelihood. Biometrika 28, 34-59.

1937
Anon. Mathematics and medicine. Lancet. i, 31.
Gossett WG. (Student) Comparison between balanced and random arrangements in field plots. Biometrika, 29:363-79.
Neyman,J. Outline of a theory of statistical estimation based on the classical theory of probability. Phil. Trans. of the Royal Society A ccxxxvi, 333-380.

1939
Frechet M. The diverse definitions of probability. J Unified Sci 8:7-23
Nagel,E. Principles of the theory of probability. International Encyclopaedia of Unified Science ed. Vol. Volume 1. University of Chicago Press, Chicago.
Steiner,WR. Some distinguished American medical students of Pierre-Charles-Alexander Louis of Paris. Bulletin of the History of Medicine 7, 783-793.
Von Mises,R. Probability, statistics and truth. Hodge, London.

1940
Greenwood,M. Medical statistics from Graunt to Farr. Biometrika 32, 101-127.

1942
Greenwood,M. Medical statistics from Graunt to Farr. Biometrika 32, 203-225.
Kendall MG. On the future of statistics. J Roy Stat Soc;105:69-80.

1943
Greenwood,M. Medical statistics from Graunt to Farr. Biometrika 33, 1-24.

1945
Braithwaite, RB. The two concepts of probability. Philosophy and Phenomenological Research;5:513-32

1946
Bradford Hill,A. Statistics in medicine. Transactions of the Manchester Statistical Society 1946-47.

1947
Barnard,GA. The meaning of a significance level. Biometrika 34, 169-182.
Bradford Hill,A. Statistics in the medical curriculum. BMJ ii, 366.

1948
Greenwood,M. Medical statistics from Graunt to Farr. Cambridge University Press, Cambridge.

1949
Kneale,W. Probability and induction. Clarendon Press, Oxford.
On the reconciliation of theories of probability. Biometrika;36:101-16.
Reichenbach,H. The theory of probability. University of California Press, London.

1950
Anon. 50 years of statistics. BMJ i, 68.
Kolmogorov,A. Foundations of the theory of probability. Chelsea Publishing Company, New York.

Neyman,J. Probability and statistics. Holt, New York.

1951
Hotelling H. The impact of RA Fisher on statistics. JASA;46:35-46
Underwood,EA. The history of the quantitative approach in medicine. British Medical Bulletin 7, 265-274.
Youden,WJ. The Fisherian revolution in methods of experimentation. Journal of the American Statistical Association 46, 47-50.

1953
Bradford Hill,A. The philosophy of the clinical trial. Public Health Service Publications, Number 388.US Government Printing Office, Washington.

1954
Boring,EG. The nature and history of experimental control. American Journal of Psychology 67, 573-589.
Dodd,SC. The scientific measurement of fitness for self-government. The Scientific Monthly 78, 94-99.
Fisher,RA. Expansion of statistics. American Scientist 42, 275-282.
Savage LJ. The foundations of statistics. New York Wiley.

1955
Braithwaite,RB, Scientific explanation. Cambridge University Press, Cambridge.
Goodman N. Fact fiction and forecast. Harvard University Press, Cambridge Mass.
Nagel,E, Principles of the theory of probability. In: International encyclopaedia of unified sciences. Vol. Vol1 Part 2. University of Chicago Press, Chicago.
Neyman,J, Statistics- servant of all sciences. Science 122, 401-406.
Rosen,G Problems in the application of statistical analysis to questions of health. Bulletin of the History of Medicine 24, 27-45.

1956
Kendall,MG. The beginnings of a probability calculus. Biometrika 43, 1-14.

1957
Barker,S. Induction and hypothesis. Cornell University Press, Cornell.
Hogben,L. Statistical theory. The relationship between probability, credibility, and error. Allen and Unwin, London.
Mises,R von. Probability, statistics and truth. 2nd rev ed. ed. Macmillan, New York.
Spencer-Brown,G. Probability and statistical inference. Longmans, London.

1958
Cox,DR. Some problems connected with statistical inference. Annals of Mathematical Statistics 29, 357-363.
Ehrenfest-Afanassjewa,T. On the use of the notion 'probability' in physics. American Journal of Physics 26, 388-392.
Gnedenko BV. The main stages of the theory of probability; Paris: Herman
Plackett,RL. The principle of the arithmetic mean. Biometrika 45, 30-35.
Toulmin,SE. The uses of argument. Cambridge University Press, Cambridge.

1959
Fisher RA. Mathematical probability in the natural sciences. Technometrics,1:21-29.
Grendaner U. Probability and statistics: the Harald Cremer volume. Stockholm: Ahlmquist and Wicksell.
Kendall,MG. Where shall the history of statistics begin? Biometrika 47, 447-449.
Leblanc,H. On so-called degrees of confirmation. Br. J. Phil. Sci. 10, 513.
Popper,K. The propensity interpretation of probability. Br. J. Phil. Sci. 10, 25-42.

1960
Mainland,D. The use and misuse of statistics in medical publications. Clinical Pharmacology and Therapeutics 1, 411-422.
Nunnally,J. The place of statistics in psychology. Educational and Psychological Measurement 20, 641-650.

1961
Jeffreys,H. Theory of probability. Clarendon Press, London.
Kyberg,HE. Probability and the logic of rational belief. Wesleyan University Press, Middletown.
Shryock,RH. The history of quantification in medical science. Isis 52, 215-237.
Woolf,H. Quantification - a history of the meaning of measurement in the natural and social sciences. Bobbs-Merrill, Indianapolis.

1962
Barnard,GA; Cox,DR. The foundations of statistical inference: a discussion. Methuen, London.
Birnbaum,A. On the foundation of statistical inference. Journal of the American Statistical Association 57, 269-306.
Carnap,R. Logical foundations of probability. 2nd ed. University of Chicago Press, Chicago.

David,FN. Games, gods and gambling. The origins and history of probability and statistical ideas from the earliest times to the Newtonian era. Hafner, New York.
Pearson,ES. Some thoughts on statistical inference. Annals of Mathematical Statistics 33, 394-403.
Savage,LJ. The foundations of statistical inference. Methuen, London.

1963

Gillespie,C. Intellectual factors in the background of analysis by probabilities. In: Scientific change. ed.: Crombie,AC, London, 431-453.
Glass DV. John Graunt and his Natural and Political Observations. Proc Roy Soc B 159:2-37
Hacking I. Guessing by frequency. Proc Aristol Soc. LXIV:62
King AC. Pathways to probability. New York: Holt, Rinehart, and Winstone.
van Leeuwen,HG. The problem of certainty in English thought 1630-1690. Nijhoff, The Hague.

1964

Coleman,JS. Introduction to mathematical sociology. Collier Macmillan, New York.
Eisenhart,C. The meaning of 'least' in least squares. Journal of the Washington Academy of Sciences 54, 23-33.
Hacking,I. On the foundations of statistics. Br. J. Phil. Sci. 15, 1-27.
Kyberg,HE; Smokler,HE. Studies in subjective probability. John Wiley, New York
Yates,F. Sir Ronald Fisher and the design of experiments. Biometrics 20, 316.

1965

Cooper,N. The concept of probability. Br. J. Phil. Sci. 14, 226-238.
Pearson,ES. Some incidents in the early history of biometry and statistics. Biometrika 52, 3-18.

1966

Bakan D. The test of significance in psychological research. Psychological Bulletin;66:423-37
Laudan,L. The clock metaphor and probabilism: the impact of Descartes on English methodological thought, 1650-1655. Annals of Science 22, 73-104.
Rankin,B. The history of probability and the changing concept of the individual. Journal of the History of Ideas 27, 483-504.
Sheynin,OB. Origin of the theory of errors. Nature. 27 August, 1003-1004.

1967

Hacking I. Possibility. Philosophical Review;76:143-68.
Hilts,VL. Statist and statistician: three studies in the history of 19th century English statistical thought. Ph.D. Dissertation, Harvard University.
Jeffreys,H. Theory of probability. 3rd ed. Clarendon Press, Oxford.
Levi,I. Gambling with truth. Alfred Knopf, New York.
Skipper JM. The sacredness of .05. American Sociologist, 2:16-18.

1968

Byrne,EF. Probability and opinion: a study of the medieval pre-suppositions of post-medieval theories of probability. Martinus Nijhoff, The Hague.
Good IJ. The estimation of probabilities. MIT Press.
Hacking I. Possibility. Australian Journal of Philosophy;46:98-112.
Hodgkinson,RG. Social medicine and the growth of statistical medicine. In: Medicine and science in the 1860s. Ed: Poynter,FNL Wellcome Institute of the History of Medicine, London, 183-198.

1969

Birnbaum,A. Concepts of statistical evidence. In: Philosophy, science and method. Ed: Morgenbesser,P St Martins Press, New York.
Cassedy JH. Demography in early America. Harvard: Harvard U Press.
Lissitz,RW. A longitudinal study of the research methodology in the Journal of Abnormal and Social Psychology, the Journal of Nervous and Mental Disease and the American Journal of Psychiatry. Journal of the History of the Behavioural Sciences 5, 248-255.
Morrison DE. Significance tests re-considered. American Sociologist, 4:131-40.
Quine,WV. Existence and quantification. In: Ontological relativity and other essays.

1970

Lucas,JR. The concept of probability. Clarendon Press, Oxford.
Morrison,DE. The significance test controversy. Aldine Publishing, Chicago.
Pearson,ES. Studies in the history of statistics and probability. Charles Griffin and Co., London.

1971

Forman,P. Weimar culture, causality, and quantum theory, 1918 - 1927: adaptation by German physicists and mathematicians to a hostile intellectual environment. Historical Studies in the Physical Sciences 3, 1-115.
Hacking,I. The Leibniz-Carnap program for inductive logic. Journal of Philosophy 68, 597-610.
Hacking,I. Equipossibility theories of probability. Br. J. Phil. Sci. 22, 339-355.
Provine,W. Origins of theoretical population genetics. Chicago University Press, Chicago.

1972
Ayer,AJ. Probability and evidence. Macmillan, London.
Cowan,RS. Francis Galton's statistical ideas: the influence of eugenics. Isis 63, 509-528.
DeFinetti,B. Probability induction and statistics. John Wiley, London.
Edwards AW. Likelihood. Cambridge University Press.
Gillespie,C. Probability and politics: Laplace, Condorcet, and Turgot. Proceedings of the American Philosophical Society 16, 1-20.
Hacking,I. The logic of Pascal's wager. American Philosophical Quarterly 9, 186-192.
Kahneman D. Subjective probability: a judgement of representativeness. Cognitive Psychology;3:430-54
Lancaster HO. Development of the notion of statistical dependence. Math Chronicle (New Zealand);2:1-16
Plackett,RL. The discovery of the method of least squares. Biometrika 59, 239-251.
Popper,KR. Objective knowledge. Oxford University Press, Oxford.
Sheynin OB. On the mathematical treatment observations in L Euler. Archive for the History of Exact Science;9:45-56

1973
Blackburn,S. Reason and prediction. Cambridge University Press, Cambridge.
Fine,TL. Theories of probability: an examination of the foundations., London.
Foucault,M. The birth of the clinic. Tavistock, London. (includes a section on the development of a statistical form of knowledge)
Laudan,L. Induction and probability in the 19th century. In: Proceedings of the 4th international congress for the logic, methodology and philosophy of science. North-Holland, Amsterdam, 429-438.
Norton,BJ. The biometric defence of Darwinism. Journal of the History of Biology 6, 283-316.
Rabinovitch NL. Probability and statistical inference in ancient and medieval Jewish literature. Toronto U Press.
Stove,DC. Probability and Hume's inductive scepticism. OUP, Oxford.

1974
Barndorff-Nielsen OE. et al Proceedings of the conference on foundational questions in statistical inference. Aarhus: University of Aarhus, Department of Theoretical Statistics.
Cochran,WG. The vital role of randomization in experiments and surveys. In: The heritage of Copernicus: theories 'pleasing to the mind. Ed: Neyman,J MIT Press, Cambridge, Mass.
Edwards,AWF. History of likelihood. International Statistical Review 42, 9-15.
Gottinger,HW. Review of concepts and theories of probability. Scientia 109, 83-110.
Harter L. The method of least squares and some alternatives. International Statistical Review;42:147-74
Kyberg,HE. The logical foundations of statistical inference. Synthese Library, v65. D Reidel, Dordrecht.
Maistrov,LE, Probability theory: a historical sketch. Academic Press, London.
Redhead,MLG. On Neyman's Paradox and the theory of statistical tests. Br. J. Phil. Sci. 25, 265-271.
Sheynin,OB. On the prehistory of the theory of probability. Archives of the History of the Exact Sciences 12, 97-141.

1975
Baker,KM. Condorcet: from natural philosophy to social mathematics. University of Chicago Press, Chicago.
Cullen,MJ. The statistical movement in early Victorian Britain: the foundations of empirical social research. Barnes and Noble, New York.
Hacking,I. The emergence of probability. Cambridge University Press, Cambridge.
Hacking I. Possibility. Philosophical Review;83:249-56.
Meier,P. Statistics and medical experimentation. Biometrics 31, 511-529.

1976
Brakel,J van. Some remarks on the prehistory of the concept of statistical probability. Archives of the History of the Exact Sciences 16, 119-136.
Owen,DB Ed. On the history of statistics and probability. Proceedings of a symposium, May 27-29, 1974. Marcel Dekker, New York.
Kempthorne,O. Of what use are tests of significance and tests of hypothesis? Communications in Statistics - Theory and Methods A5, 763-777.
Sheynin,OB. PS Laplace's work on probability. Archives of the History of the Exact Sciences 16, 137-187.
Harper WL Foundations of probability theory. Proceedings of an international research conference. Dordrecht: Reidel (3 vols).
Selvin,HC. Durkheim, Booth and Yule: the non-diffusion of an intellectual innovation. Archives of European Sociology 17, 39-51.
Van Brakel,J. Some remarks on the prehistory of the concept of statistical probability. Archives of the History of the Exact Sciences 16, 95-112.
Schneider I The introduction of probability into mathematics. Hist Math; 3:135-40.

1977
Buck,P. Seventeenth century political arithmetic: civil strife and vital statistics. Isis 68, 67-84.
Cowgill GL. The trouble with significance tests and what we can do about it. American Antiquity;42:350-68
Cox DR. The role of significance testing. Scan J Statistics;4:49-70

Gore,SM. Misuse of statistical methods: critical assessment of articles in the British Medical Journal from January to March 1976. BMJ i, 85-87.
Heyde,CC; Seneta,IJ, Bienaym: statistical theory anticipated., New York.
Runnalls,AR. Bayesianism and a subjectivist philosophy of induction. Ph.D. Dissertation, Cambridge University Cambridge.
Seal HL Studies in the history of probability and statistics. Biometrika;64:429-39.
Thomas L. Biostatistics in medicine. Science;198:675.

1978

Acree,MC. Theories of statistical inference in psychological research: a historico-critical study. Ph.D. Dissertation, Clark University.
Barone J. A history of the axiomatic formulation of probability. Arch Hist Exact Sci;18:123-90
Barone,J; Novikoff,A. A history of the axiomatic formulation of probability from Borel to Kolmogorov. Part 1. Archives of the History of the Exact Sciences 18, 123-190.
Box,JF. RA Fisher: the life of a scientist. John Wiley, New York.
Canguilhem,G. On the normal and the pathological. Reidel, Dordrecht.
Eggleston R. Evidence, proof and probability. Weidenfeld and Nicholson.
Hilts,VL. Aliis Externundum, or, the origin of the Statistical Society of London. Isis 69, 21-43.
MacKenzie,D. The development of statistical theory in Britain 1865-1925: a historical and sociological perspective. Ph.D. Dissertation, Edinburgh University Edinburgh.
MacKenzie,D. Statistical theory and social interests. Social Studies of Science 8, 35-83.
Menard,C. Three forms of resistance to statistics: Say, Cournot, Walras. History of Political Economy 12, 524-541.
Norton,BJ. Karl Pearson and statistics: the social origins of scientific innovation. Social Studies of Science 8, 3-34.
Norton,BJ. Karl Pearson and the Galtonian tradition. Ph.D. Dissertation, University College London
Rescher N. Peirce's philosophy of science: critical studies in his theory of induction and scientific method. University of Notre Dame Press.
Shafer G. Non-additive probabilities in the work of Bernoulli and Lambert. Archive for the History of Exact Science;19:309-70
Sheynin,OB. C.F Gauss and the theory of errors. Archives of the History of the Exact Sciences 18, 245-300.
Simonton,DK. Independent discovery in science and technology. A closer look at the Poisson distribution. Social Studies of Science 8, 521-532.
Sprott,DA. Gauss's contribution to statistics. History of Mathematics 5, 183-203.

1979

Anding,JE, Peirce's defense of scientific method. Ph.D. Dissertation, University of Waterloo Ottawa.
Eyler,JM, Victorian social medicine: the ideas and methods of William Farr. Johns Hopkins University Press, Baltimore.
Giere,RN, Foundations of probability and statistical inference. In: Current research in philosophy of science. Proceedings of the PSA critical research problems conference. Philosophy of Science Association, East Lancing,
Harre,R Social being. Blackwell, Oxford.
Lovie,AD, The analysis of variance in experimental psychology: 1934-1945. British Journal of Mathematical and Statistical Psychology 32, 151-178.
MacKenzie,D Karl Pearson and the professional middle class. Annals of Science 36, 125-143.
Norton,BJ, Charles Spearman and the General Factor in Intelligence: genesis and interpretation in the light of sociopersonal considerations. Journal of the History of the Behavioural Sciences 15, 142-154.
O'Donnell,JO The crisis of experimentalism in the 1920s: EG Boring and the use of history. American Psychologist 34, 289-295.
Seidenfeld,T, Philosophical problems of statistical inference: learning from RA Fisher. Reidel, Dordrecht.
Young,RM, Why are figures so significant? The role and the critique of quantification. In: Demystifying social statistics. Eds.: Miles,I; Evans,J Pluto Press, London, 63-74.

1980

Basu D. Randomization analysis of experimental data. J Am Stat Assoc, 75:575-82.
Box,JF RA Fisher and the design of experiments. American Statistician 34, 1-7.
Cohen,LJ, Some historical remarks on the Baconian conception of probability. Journal of the History of Ideas 41, 219-231.
Daston,LJ, Probabilistic expectation and rationality in classical probability theory. Historia Mathematica 7, 234-260.
Eyler,JM, The conceptual origins of William Farr's epidemiology: numerical methods and social thought in the 1830s. In: Times, places and persons: aspects of the history of epidemiology. Ed: Lilienfeld,AM Johns Hopkins Press, Baltimore, 1-22.
Fienberg,SE; Hinkley,DV, RA Fisher: an appreciation. Springer Verlag, New York.
Freudenthal H Huygens foundations of probability. Hist Math 7:113-117
Hacking,I, The theory of probable inference: Neyman, Peirce and Braithwaite. In: Science, belief and behaviour: essays in honour of RB Braithwaite. Ed: Mellor,DH Cambridge University Press, Cambridge,
Hilts,VL, Statist and statistician: three studies in the history of 19th century English statistical thought. Arno, New York.
Kruskal,WH Representative sampling. IV. The history of the concept in statistics. International Statistical Review 48169-195
Levi,I, The enterprise of knowledge: an essay on knowledge, credal probability, and choice. MIT Press, Cambridge, Mass.
Lindley DV. LJ Savage: his work in probability and statistics. Annals of Statistics;8:1-24

Samuelson,PA, Fishers 'reproductive value' as an economic specimen in Merton's zoo. Transactions of the New York Academy of Science 39, .
Schneider I Christian Huygens contribution to the development of a calculus of probabilities Janus;67:269-79
Sheynin,OB, On the history of the statistical method of biology. Archives of the History of the Exact Sciences 22, 323-371.
Stigler,SM, American contributions to mathematical statistics in the 19th century. Arno, New York.
Stigler,SM, R.H. Smith: a Victorian interested in robustness. Biometrika 67, 217-221.
Vickers,M, Truth, consensus and probability: on Peirce's definition of scientific truth. Pacific Philosophy Quarterly 61, 183-203.

1981

Cohen PC Statistics and the state. William and Mary Quarterly;38:35-55
Conference on HPS Probability theory and probabilistic thinking in the classical modern period. Dordrecht: Reidel (8 references)
Daston,LJ, Mathematics and the moral sciences: the rise and fall of the probability of judgements. In: Epistemological and social problems of the sciences in the early nineteenth century. Eds.: Jahnke,HN; Otte,M Reidel, Dordrecht, .
Diamond M. Nightingale on Quetelet. J Roy Stat Series A 144:66-79
Dobbertin,GF, Statistics and social science: the introduction of inferential statistics into higher education in America from 1890-1930. Ph.D. Dissertation, Michigan State University.
Evans B. IQ and mental testing: an unnatural science and its social history. Macmillan
Folks,JL. Ideas of statistics. John Wiley, New York.
Hacking,I. How should we do the history of statistics? Ideology & Consciousness 8, 15-26.
Hacking,I. Karl Pearson's History of Statistics. Br. J. Phil. Sci. 32, 177-182.
Laudan,L. Science and hypothesis: historical essays on scientific methodology. Reidel, Dordrecht.
MacKenzie,D. Statistics in Britain, 1865-1930: the social construction of scientific knowledge. Edinburgh University Press, Edinburgh.
Murphy,TD. Medical knowledge and statistical methods in early nineteenth century France. Medical History 25, 301-319.
Porter,TM. The calculus of liberalism: the development of statistical thinking in the social and natural science of the nineteenth century. Ph.D. Dissertation, Princeton University.

1982

Buck,P. People who counted: political arithmetic in the eighteenth century. Isis 73, 28-45.
Cowles,M. On the origin of the .05 level of statistical significance. American Psychologist 37, 553-558.
Dale, AI. Bayes or Laplace Arch Hist Exact Sci 27:23-47
Downton F. Legal probability and statistics. J Roy Stat Soc A;145:395-402
Eddy DM. Probabilistic reasoning in clinical medicine: problems and opportunities. In Kahnman et al eds. Judgements under uncertainty. Cambridge: Cambridge University Press.
Hacking,I. Biopower and the avalanche of printed numbers. Humanities in Society 5279-295.
Horwich,P. Probability and evidence. Cambridge: Cambridge University Press.
Niiniluoto,I. Statistical explanation. In: Contemporary philosophy: a new survey. Ed: Guttorm,F Nijhoff, The Hague.
Sheynin,OB. On the history of medical statistics. Archives of the History of the Exact Sciences 26, 241-286.
Stigler,SM. Thomas Bayes Bayesian inference. Journal of the Royal Statistical Society A 145, 250-258.
Weatherford R. Philosophical foundations of probability theory. London: RKP.
Wynn,HP. Controlled versus random experimentation. The Statistician 31, 237-244.
Yearley S. The relationship between epistemological and sociological cognitive interests. Studies in the Philosophy and History of Science 13, 353-88 (discussion of MacKenzie's study on the history of statistics in Britain).

1983

Armitage,P. Trials and errors: the emergence of clinical statistics [Presidential address delivered to the Royal Statistical Society June 22 1983]. Journal of the Royal Statistical Society A 146, 321-334.
Baird,D. The Fisher/Pearson chi squared controversy: a turning point for inductive inference. Br. J. Phil. Sci. 34, 105-118.
Bennett,JH. Natural selection, hereditary, and eugenics. Including selected correspondence of RA Fisher with Leonard Darwin and others. Clarendon Press, Oxford.
Hacking,I. Nineteenth century cracks in the concept of determinism. Journal of the History of Ideas 44, 455-475.
Heidelberger M. et al Interdisciplinary studies of scientific development. Bielefeld: University of Bielefeld (23 references).
Lancaster HO. Development of the notion of statistical independence. Math Chron;2:1-16.
Lindley DV. The role of randomization in inference. Proceedings of the Philosophy of Science Association. East Lansing. Vol 2 431-446.
Shapiro,BJ. Probability and certainty in seventeenth century England. Princeton University Press, Princeton.
Shoesmith E. Expectation and the early probabilists. Hist Math: 10:78-80.
Stigler,SM. Who discovered Bayes theorem? American Statistician 37, 296-325.

1984

Cassedy,JH American medicine and statistical thinking, 1800 - 1860. Harvard University Press, Cambridge, Mass.
Cohen,IB. Florence Nightingale Scientific American 250(3):128-136.
Duncan,OD. Notes on social measurement: historical and critical. New York: Russell Sage.
Gould,SJ. The mismeasure of man. Harmondsworth: Pelican Books.
Levi,I. Decisions and revisions. CUP, Cambridge.
Lindley,DV. A Bayesian lady tasting tea. In: Statistics: an appraisal. Eds.: David,HA; David,HT Iowa State University

Press, Iowa, 455-477.
Mainland,D. Statistical ritual in clinical journals: is there a cure? - II. BMJ 288, 920-922.
Rosenthal,SB. Mathematical necessity, scientific fallibilism, and pragmatic verificationism: the Peircean synthesis. International Philosophical Quarterly 24, 1-19.
Suppes,P. Probabilistic metaphysics. Blackwell, Oxford.

1985
Atkinson AC (ed.) A celebration of statistics. New York: Springer.
Porter,TM, The mathematics of society: variation and error in Quetelet's statistics. British Journal of the History of Science 18, 51-69.
Urbach P. Randomization and the design of experiments. Phil Sci, 52: 256-273.

1986
Hacking,I Making up people. In: Reconstructing individualism. Ed: Heller,T Stanford University Press, Stanford.
Oakes,M, Statistical inference: a commentary for the social and behavioural sciences. John Wiley, Chichester.
Porter,TM The rise of statistical thinking 1820-1900. Princeton University Press, Princeton.
Stigler,SM The history of statistics: the measurement of uncertainty before 1900. Harvard University Press, Cambridge, Mass.
Swijtink,ZG, D'Alembert and the maturity of chances. Studies in History and Philosophy of Science 17, 327-349.

1987
Berger JO. Testing a null point hypothesis. JASA;82:112-39.
Coleman,WD Experimental physiology and statistical inference: the therapeutic trial in nineteenth century Germany. In: The probabilistic revolution. Vol. 2.Eds: Kruger,L; Gigerenzer,G; Morgan,MS MIT Press, Cambridge, Mass, 201-227.
Danziger K. Statistical method and the historical development of research practice in American psychology. In The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge.
Daston,LJ Rational individuals versus laws of society: from probability to statistics. In: The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge
Daston,LJ The domestication of risk: mathematical probability and insurance 1650-1830. In: The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge, 237-260.
Dear,P Jesuit mathematical science and the reconstruction of experience in the early 17th century. Studies in History and Philosophy of Science 18, 133-175.
Gigerenzer,G Probabilistic thinking and the fight against subjectivity. In: The probabilistic revolution. Volume 2: ideas in the sciences. Eds.: Kruger,L; Gigerenzer,G; Morgan,MS MIT Press, Cambridge Mass., 11-34.
Gigerenzer,G; Murray,DJ, Cognition as intuitive statistics. Lawrence Erlbaum, Hillsdale, NJ.
Hacking,I Prussian numbers, 1860-1882. In: The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge, 378-394.
Hacking,I Was there a probabilistic revolution 1800-1930? In: The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge, 45-55.
Heidelberger,M Fechner's indeterminism: from freedom to laws of chance. In: The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge, 117-156.
Kruger,L The slow rise of probabilism: philosophical arguments in the nineteenth century. In: The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge, 59-89.
Luyer,B-P Probability in vital and social statistics: Quetelet, Farr, and the Bertillons. In: The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge, 317-335.
Metz,KH Paupers and numbers: the statistical argument for social reform in Britain during the period of industrialisation. In: The probabilistic revolution. Volume 1: ideas in history.
Porter,TM Lawless society: social science and the reinterpretation of statistics in Germany, 1850-1880. In: The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge, 351-375.
Stigler,SM The measurement of uncertainty in nineteenth century social science. In: The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge, 287-292.
Swijtink,ZG The objectification of observation: measurement and statistical methods in the nineteenth century. In: The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge, 261-285.
Wise,MN How do sums count? On the cultural origins of statistical causality. In: The probabilistic revolution. Volume 1: ideas in history. Eds.: Kruger,L; Daston,LJ; Heidelberger,M MIT Press, Cambridge, 395-425.

1988
Daston L. Classical probability in the enlightenment. Princeton University Press.
Hacking,I. Telepathy: origins of randomisation in experimental design. Isis 79, 427-451.
Plackett,RL. Data analysis before 1750. International Statistical Review 56, 181-195.

1989
Desrosieres,A The part in relation to the whole: how to generalise? The prehistory of representative sampling. In: The social survey in historical perspective. Ed: Bulmer,M Cambridge University Press, Cambridge.
Gigerenzer,G The empire of chance: how probability changed science and everyday life. Cambridge University Press,

Cambridge.
Johnson DJ. On the necessity of random sampling. Br J Phil Sci, 40:443-457.
Signorile V. Buridan's Ass: the statistical rhetoric of science and the problem of equiprobability. In Simons HW. Rhetoric in the human sciences. New York: Sage, 69-88.

Listings of this sort are inevitably only partially complete. Were it more extensive I believe the difference between the pre- and post- Second World War periods would be greater. The creation of this list was made more difficult because judgements had to be made in some instances concerning whether a text was about the underlying principles of statistics rather than being a 'tutor text'. I have included one introductory text (Folks 1981) since it uses the history of statistics to introduce statistics to undergraduate students. The contributions to one multi-authored text have been listed individually (Daston et al 1987); while those in another have not (Owen 1976), reflecting the differential importance of the contributions. Some of the contributions were not originally published in English. I have used their date of publication in English here.

Appendix 3

Statistical models of hypothesis testing and estimation

Introduction

The purpose of this appendix is threefold. Firstly, to highlight the major logical problem confronting statistical testing, namely that of direct and indirect inference. Secondly, to describe the principles underlying statistical hypothesis testing. Thirdly, through the analysis of two competing models for statistical hypothesis tests, to argue that the logical foundations of statistical tests are less established than might be expected. The same criticism applies to two theories of interval estimation, described very briefly at the end of the appendix. The appendix concludes by describing Hacking's resolution to the disjunction between the utility of statistics and their basis in formal logic. Without denying that statistics produce verifiable knowledge, Hacking argues that the processes of verification have a stylistic or rhetorical basis as much as they have a basis in formal logic.

Significance tests and decision tests

Two approaches to statistical hypothesis testing form the main body of the appendix. The first, known as significance testing, is described in some detail, since it readily demonstrates the basic design of various statistical tests, including the one used in the streptomycin trial. The second, known as decision testing or hypothesis testing,[786] the latter name reflecting the view that it is the accepted standard approach to hypothesis testing,[787] is then introduced.

[786] Seidenfeld 1979 p3.

[787] Kyberg 1974 p 22.

Significance tests are attributed to RA Fisher. Decision tests (these names are adopted from now on), are attributed to Jerzy Neyman/Egon Pearson. Both are a product of a eruption of interest in statistical tests which took place in the first half of the twentieth century.[788] Chronologically, significance testing preceded decision testing.[789] Neyman and Pearson argued that their approach was based on a general model, of which Fisher's was a particular instance. However, Fisher had already introduced what he regarded as a general framework for statistical tests in 1922,[790] and regarded decision tests as unscientific and even inimical to science.

The approaches are alike in many ways. Both concern reasoning about the truth of hypotheses given some empirical data. The logic of the two approaches is superficially the same, and both lead to similar sorts of phraseology about the truth and rejection of hypothesis. However, their manner of proceeding from statistical data to inferences is radically different.

In formal terms, both are solutions to the problem of inverse inference and are therefore solutions to the epistemological problem of inverse inference. Although the concept of inverse inference appears remote from statistical testing it may be usefully described here, since it forms the general problematic to which many statistical tests are solutions,

[788] Depending on how casually the term is applied, hypothesis testing was anticipated in the early eighteenth century, when John Arbuthnot used a test to establish to reality of divine intervention on the basis of birth statistics (Arbuthnot 1710 discussed in Gigerenzer 1989 p135). Much of the framework for statistical tests was established by the American philosopher CS Peirce in the second half of the nineteenth century. For further examples of anticipation and, in the case of Fisher, co-discovery, see Gigerenzer 1989 p 84-90. Hacking argues that Neyman-Pearson confidence interval theory was anticipated by the Harvard mathematician Edward B Wilson (Hacking 1980 p 143).

[789] Fisher worked out his method of significance testing in the early1920s. Neyman and Pearson's earliest papers on decision testing appeared in Biometrika in 1928.

[790] Fisher 1922.

and because the approaches taken by Fisher and Neyman-Pearson reflect different ways of overcoming the limitations on knowledge it imposes.

Direct and indirect inference

Direct and indirect inference will be contrasted by means of a simple example which will be used subsequently to illustrate the method of significance testing. If we know that a coin is 'fair' we also know that the chance of a flip landing heads is the same as landing tails. Fairness is in this instance synonymous with equi-probability. We therefore know using either the law of equiprobability or intuition that the probability of heads = 0.5. Its 'fairness' can be labeled a characteristic, or parameter of the coin,[791] which might also be called its 'bias', with the value, for this coin of 0.5.

Knowing the value of this parameter we can proceed, by direct inference, to assert the probability of outcomes for events involving this coin which incorporate a chance ingredient (e.g. tossing the coin). Table 3.1 shows the probability of several outcomes in a chance set-up, for two bias values, a fair coin, and a coin heavily weighted to land tails. In each case the probability is calculated using direct inference.

Table 3.1: The chance of various event for two values of θ

Bias	**θ = 0.5 (i.e. fair coin)**	**θ = 0.1 (e.g. a coin biased towards landing tails)**
Outcome/number of trials	**Probability**	
one head/one toss,	0.5	0.1
one head/two tosses,	0.75	0.19
one head/three tosses,	0.875	0.271
two heads/three tosses,	0.375	0.027
10 heads/20 tosses,	0.1762	$6.44204 \times E^{-06}$

[791] Many of the terms used today, such as parameter, population, and sample originate in Fisher's 1922 paper *On the mathematical foundations of theoretical statistics* (Fisher 1922).

While the probability of 1 head from 2 tosses is 0.75, the chance of 10 heads from 20 tosses is not 0.75. There are many possible heads-outcomes from 20 tosses, ranging from 0 to 20. The general method for calculating the probability of getting x heads from y tosses is the binomial distribution, a mathematical formulae[792]. The probability of obtaining 0 to 20 heads from a series of 20 tosses of a fair or un-biased coin is given in Table 3.2.

[792] The binomial distribution is sometimes associated with James Bernoulli, who published a geometric device for generating the binomial distribution in his great text on probability, Ars Conjectandi, 1713. His work was anticipated by Pascal, Stifel, and has been located in a Chinese text of 330BC (Walker 1931).

Table 3.2: the probability of obtaining r heads in 20 flips of a fair coin

r =	Probability	r =	Probability
0	9×10^{-7}	11	0.1602
1	1.9×10^{-5}	12	0.1201
2	2×10^{-4}	13	0.0739
3	0.0011	14	0.0370
4	0.0046	15	0.0148
5	0.0148	16	0.0046
6	0.0370	17	0.0011
7	0.0739	18	2×10^{-4}
8	0.1201	19	1.9×10^{-5}
9	0.1602	20	9×10^{-7}
10	0.1762		

Suppose now that the bias of a coin is unknown, but that some data produced by flipping the coin is available. Inverse inference is the process of asserting the value of the bias of the coin on the basis of the known data. The problem when applying inverse inference is that of knowing what value to place on a parameter. If I toss a coin 3 times and all three tosses return heads, should I conclude that the coin is biased? It might be a fair coin, which has by chance given an unusual result. What is the value of the coin's bias?

The obvious answer in this example is to toss the coin some more times. In many situations it is either not possible or too costly to expand the empirical data on which to base inverse inferences. And in any case, while increasing the data is helpful, the problem persists: any particular result supports several values of the unknown parameter. Thus 10 heads in twenty tosses provides good evidence to support the theory that the coin is fair. However, it also supports the theory that the coin is biased. Support offered by the data for this conclusion is less strong, but there is no guarantee that more strongly supported conclusions are in fact the correct ones.

In one school of probability the problem of inverse inference is overcome by allowing individuals to assign a value to an unknown parameter before data is gathered. Having gathered some empirical data, the individual re-calculates their estimate of the value of the unknown value. This is the subjectivist or Bayesian School of statistics.[793]

In the other school, the concept of individualistic assertions about the value of parameters (sometimes called personal probabilities or betting rates) is regarded as illogical, since it can lead to divergent results on the basis of the same empirical data.[794] The frequentist or objective school of probability finds work-arounds for not pre-judging the value of a parameter, as discussed below.

Statistical hypothesis tests

Statistical hypothesis tests employ inverse inference – inferring from empirical data the likely value of an unknown parameter - to test hypotheses framed as probability distributions of the unknown parameters.

In historical terms there are two types of test, attributed to Fisher on the one hand and Neyman/Pearson on the other. This is not how contemporary users see statistical testing. From the perspective of the modern user, statistical testing is a toolkit, consisting of an amalgam of procedures, derived in part from Fisher, in part from Neyman-Pearson, with additions from several other sources, including Neyman's theory of interval estimation, used to estimate values rather than test hypotheses, and non-parametric techniques which do not make assumptions about the nature of unknown parameters.

---

[793] For a modern re-statement of Bayesian statistics, see Howson and Urbach 1992.

[794] Fisher's objection to this approach is that by adopting it we 'seem forced to regard mathematical probability, not as an objective quantity measured by observed frequencies, but as measuring merely psychological tendencies, theorems respecting which are useless for scientific purposes'. (Fisher 1935**a** p6-7).

The adequacy of such an amalgam is rarely considered. The pragmatic decision has been taken that the toolkit is robust, or at least offers the most robust way of proceeding in an uncertain world. Adequation is therefore mainly a practical matter, based circularly on success and usage.[795] The irrelevance of a theory of epistemology to statistics, in the sense implied by Habermas,[796] can be sensed in the following quotation:

> *'Despite basic philosophical differences, in their main practical aspects the two theories are complementary rather than contradictory, and ...a unified approach is possible that combines the best features of both'*[797]

The two approaches to statistical hypothesis testing are set out below.

Significance tests

Significance tests allow data to refute hypotheses. They do not therefore lead to statements about the extent to which hypotheses are true, but are able to indicate the un-truth of particular hypotheses. This approach is close to that recommended by Karl Popper for constructing scientific knowledge,[798] except that Popper tended to favour

[795] Hacking described the circularity of arguments used to justify statistics as themselves part of the statistical style of reasoning: 'the truth is what we find out in such and such a way. We recognise it as truth because of how we find it out. And how do we know that the method is good? Because it gets at the truth' (Hacking 1992 p135).

[796] 'Hence I should like to put forth the thesis that since Kant science has no longer been seriously comprehended by philosophy. Science can only be comprehended epistemologically, which means as one category of possible knowledge, as long as knowledge is not equated effusively with the absolute knowledge of a great philosophy or blindly with the scientistic self-understanding of the actual business of research' Habermas 1978 p4.

[797] Lehman 1993 p1242.

[798] Popper started by assuming that Hume was correct. That is, empirical data can never confirm the truth of a proposition. 'But there is a further negative result; there are logically valid negative arguments leading in the inductive direction: a counterinstance may disprove a law' (Popper 1983 p111). 'My solution of the logical problem of induction was that we may have preferences for certain of the competing conjectures; that is, for those which are highly informative and which so far have stood up to eliminative criticism. These preferred conjectures are the result of selection, of the struggle for survival of the hypotheses under the strain of criticism [i.e. hypothesis testing], which is artificially intensified selection pressure' (Popper 1983 p112-3).

experiments that were capable of definitively rejecting hypothesis, rather than statistical experiments that must attach probabilities to statements about refutations.[799]

However, the conditions Popper set out as necessary for the formation of scientific knowledge are rarely achieved in reality[800]. More often, experiments appear to offer provisional findings at best, suggesting that the epistemological modesty inherent in probabilistic statements reflects the aspirations of research workers better than Popper's refutationalist concept of decisive experiments. *Statistical* hypothesis testing is therefore a powerful practical means of establishing an empirical way of scientific knowing. It requires hypotheses and data of a particular type, and part of the art of research design is the selection and framing of research questions in terms of hypotheses that are amenable to significance testing.

It is not always easy to show the framework of significance testing using examples that are of most relevance to health care. Instead, a straightforward research question and experimental set-up will be described in which the framework is very clear. The following example concerns the researchable question 'is this coin biased?'. The question can easily be transformed into an assertion capable of empirical test: 'this coin is not biased' becomes the hypothesis that is to be tested experimentally. The test will be based on data produced by tossing the coin. The analogy between the coin and an example from health care will be drawn subsequently.

---

[799] Popper was critical of statistical hypothesis testing because he considered the distributional aspect of the test to be based on a theoretically repeated experimental arrangement, not the decisive test of a hypothesis (Popper 1983 p202). Popper, like Fisher, argued that a hypothesis was like a single event, and required a theory of testing which gave a probability to a single event, not a hypothetical sequence. Popper's propensity definition of probability is very similar to that of Hacking in chapter 1 of *Logic of Statistical Inference*.

[800] The limitations of Popper's philosophy are reviewed by Chalmers (Chalmers 1982).

We have a coin, and we wish to know something about a property of the coin called its 'bias'. Although bias is a property of the coin I cannot inspect it directly, unlike say, its colour. In fact very few properties can be inspected directly. The weight of a coin requires me to use a device which supply me with a weight. Even colour requires me to shine a light on the coin.

I must therefore further define bias including the conditions under which a value can be supplied to it. I will interpret the bias of the coin as its tendency to land heads or tails, and measure this by putting the coin in a machine which tosses the coin twenty times, just as I might evaluate the colour of a coin by putting it in a spectrometer.

I will get data from the coin-tossing machine. In this case the data are:

H T H T H H T T T T H T H H T H H H H T.[801]

Given this data, what can be said about the coin's bias? Bias has already been interpreted as the tendency to land heads or tails. Does this mean that I should achieve any particular sequence of heads and tails? No. Part of the set up here is that each flip of the coin is independent, so could land heads or tails, a result that is unaffected by the result of the previous trial. So that if 49 tosses of a fair coin produce 49 heads the chance of getting a head on the next toss is the same as the chance of getting a tail on the next toss.

In the coin toss example set out above, the particular sequence of results tells me little about the bias of the coin. One way of representing the data is to summarise it. The summary number is called a statistic.[802] There are many statistics that can be derived

[801] Thanks to Hyat Robbins-Toth for tossing a coin to provide these data.

[802] The theory set out here originated in Fisher 1922.

from the dataset. For example, 1 is the value of the statistic 'number of runs of three or more heads'. To come to the point, valuable statistics are those whose probability distribution is known. The statistic here is: 'the number of heads'. It is possible to derive the probability of any number of heads in a trial, which consists of twenty flips of an unbiased coin. Table 3.2 shows the probabilities of each value of heads calculated from the binomial distribution. Note that implicit in the data of Table 3.2 is the assumption that the coin is fair. The simplest way of defining a statistical hypothesis test is that it a mechanism for allowing empirical data to 'testify' to the value of the assumptions that generate probability distributions. If we toss a coin 10 times and get 9 heads, the probability distribution based on a fair coin is not well supported by the data, and hence the assumption that the coin is fair may not be true.[803]

The data in Table 3.2 can be shown graphically, revealing a symmetrical curve (Figure 3.1)

In this case there were 11 heads. Does this mean the coin is fair? It might. It might mean that the coin was slightly more likely to give heads (i.e. that it was slightly biased to heads).

[803] Fisher: 'Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis' (Fisher 1935**a** p16).

Figure 3.1 Probability curve (also known as probability density function) for the number of heads produced by tossing a fair coin 20 times

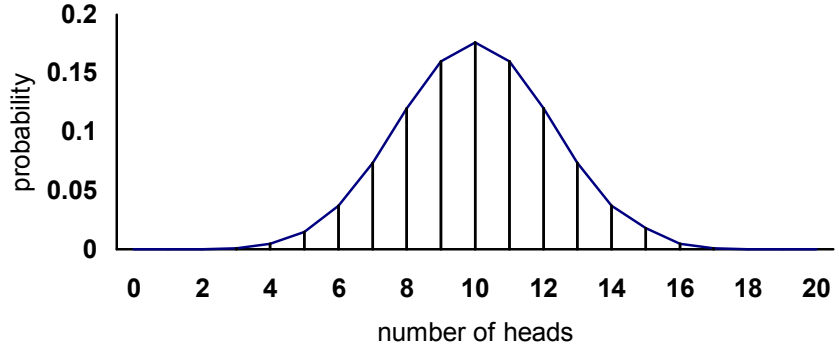


Table 3.2 and Figure 3.1 show the probability of any particular result from a series of 20 flips of a fair coin. Tossing 11 heads would appear to be an unremarkable result, even though the probability of such a result is 0.162. Tossing 4 heads would be an unusual event, with a probability of 0.0046. Should a fair coin be tossed 20 times and this result occur, the frequency interpretation of probability says that an event has occurred which should only be seen once in every 217 sequences of 20 flips.

The next step in the argument is to make a connection between the results obtained and the hypothesis's underlying distribution. Fisher's logic at this point is convoluted but elegant. It runs as follows:

1. If the coin is fair the probability distribution function is as shown (in Figure 3.1)

2. If the distribution is as shown and the data (4 heads) have been obtained, then

3. 'Either an exceptionally rare chance has occurred or the [probability curve] is not true'[804].

[804] Fisher 1956 p39. A third possibility is that the underlying statistic is not normally distributed. Part of Fisher's approach was to select statistical parameters that are known to be distributed normally, such as Student's t

If we accept that rare events are (by definition) unlikely to occur, then:

4. The probability distribution imputed to the bias parameter of the coin on the basis of the hypothesis does not seem plausible, and hence

5. The hypothesis that this coin is fair is effectively refuted.

In this example the probability of 4 heads is 0.0046. The convention has grown up that in judging the plausibility of an event the probability **P** of all events with the same or less probability as the index event should be calculated. In this example:

> ***P*** *= prob. (4 heads) + prob.(3 heads) + prob.(2 heads) + prob.(1 head) + prob.(0 heads) + prob.(16 heads) + prob.(17 heads) + prob.(18 heads) + prob.(19 heads) + prob.(20 heads)* ***= 0.006***

A further convention is that if **P** is numerically less than 0.05, then the null hypothesis may be rejected. In this case, the null hypothesis is that the coin is unbiased. On the result of 4 heads we therefore reject the hypothesis that the coin is unbiased. The result is said to be significant at the 0.05 level. A result of 6 heads would have a P of 0.0575, and would not be significant at the 0.05 level. Other levels of P are sometimes used, for example 0.01.

This form of testing does not prove hypotheses. We have not proved that the coin is biased. But, in a Popperian style, we have refuted, within the framework of the significance test, the claim that the coin is not un-biased.

---

distribution. Where an underlying statistic is not normally distributed, such as the mean in the case of two samples not drawn from the same population, so-called non-parametric statistical tests are used. They use test statistics which are independent of any assumption about underlying populations. Non-parametric tests were introduced in the 1940s, by Wilcoxon (1945) and Mann and Whitney (1947).

Limitations in the logic of significance tests

Fisher never set out the logical framework for significance tests more robustly than has been presented here.[805] Fisher asks us to accept that if the probability curve is true then events with a low probability are unlikely to occur. It does not follow that if a low probability event does occur the curve does not apply. Logically, the occurrence of an event defined as rare says nothing about the truth or falsehood of the premises that derive the probability of the event. Although elegant therefore, the logic of significance tests lacks force.[806]

The further limitation of significance testing is the degree of arbitrariness it includes. Three aspects of significance testing that include an arbitrary element:

The choice of P level. In the example here 4 heads refuted the null hypothesis at the 0.05 level, but did not at the 0.01 level. The choice of significance level, while not entirely arbitrary, is not associated with any objective rules to determine its appropriate level.

The choice of statistic. Recalling that a statistic generally involves a contraction of the available data, it is apparent that any one set of data can support several statistics. It can be shown that particular empirical data can both refute and not-refute a hypothesis depending on the test statistic derived from it and used in a significance test.[807]

---

[805] According to Hacking 1965 p82.

[806] Fisher was aware of this, as the following quotation shows: 'I have assumed, as the experimenter always does assume, that it is possible to draw valid inferences from the result of experimentation…it is however, certain that many mathematicians, if pressed on the point, would say that it is not possible rigorously to argue from the particular to the general….We may at once admit that any inference from the particular to the general must be attended with some degree of uncertainty, but this is not the same as to admit that such inferences cannot be absolutely rigorous, for the nature and degree of the uncertainty may itself be capable of rigorous expression.' (Fisher 1935**a** p3-4). Against this view can be set Habermas's criticism of the circularity of scientific justifications of scientific knowledge. (Habermas 1974).

[807] See Howson and Urbach 1993 p181. Howson and Urbach argue that this is an especially difficult problem for the chi-squared test.

The choice of rejection region. In significance testing we are looking for outcomes which have a low probability of occurring. By tradition only, these regions are concentrated at the tail of a distribution. It would be valid, using Fisher's logic of significance tests to create a rejection region consisting of 3, 14, 16 and 17 heads. The combined probability of these outcomes is 0.0438. To use this rejection region, toss the coin 20 times. If 3, 14, 16 or 17 heads is obtained, reject the null hypothesis.

In summary, despite their appearance as an objective method for testing hypotheses, significance tests make use of subjective judgement and arbitrary decisions at several points in their methodology.

Decision tests

Decision test theory aimed to solve the problem of inverse inference without reliance on subjective prior probabilities. The intention of decision tests is therefore the same as significance tests. It was developed by Jerzy Neyman and Egon Pearson, beginning in the late 1920s. They regarded decision test theory as setting significance tests on a logical basis. Despite the superficial similarity of significance and decision tests, the underlying structure of their arguments are different. Fisher regarded them as incompatible, and regarded decision tests as un-scientific.

The following discussion introduces decision testing by way of an example originally set out by Kyberg and reproduced in Howson and Urbach.[808] The discussion that follows concerns three key differences between significance and decision tests: the epistemological intent of the test; the nature of the test; and the interpretation of the result.

[808] Howson and Urbach 1992.

Whereas significance testing concerns a single hypothesis, decision testing requires 2 competing hypothesis. The purpose of a decision test is to calculate which of the hypotheses is the best. Kyberg's example concerns a bulb grower who receives a shipment composed of a mixture of red and yellow bulbs, and who has forgotten if the proportion of bulbs is 40/60 or 60/40 red/yellow. He undertakes an experiment in order to supply some evidence about the true proportion. The experiment consists of randomly selecting 10 bulbs from the shipment. On the basis of a statistic (the number of red bulbs in the sample) he will decide which of the two competing hypotheses (labelled H1 and H2) is true. Table 3.3 shows the probability of obtaining any particular result for H1 and H2

Table 3.3. The probability of getting red bulbs in a sample of 10 from a shipment of mixed red/yellow bulbs

Number of reds in sample	H1 40% Red	H2 60% Red
	Probability	
0	0.0060	0.0001
1	0.0403	0.0016
2	0.1209	0.0106
3	0.2150	0.0425
4	0.2508	0.1115
5	0.2006	0.2006
6	0.1115	0.2508
7	0.0425	0.2150

8	0.0106	0.1209
9	0.0016	0.0403
10	0.0001	0.0060

The epistemological intent of Neyman-Pearson tests

Significance testing supports a three-step model of scientific progress in which:

- A theory (hypothesis) is proposed
- Data is collected in an attempt to reject the theory
- Either the theory is rejected or it is not

Fisher maintained that the statistical testing of hypotheses could only be thought of in this way.[809]

By contrast, Neyman-Pearson decision testing offers a rule of inductive behaviour:

> *The problem of testing a statistical hypothesis occurs when circumstances force us to make a choice between two courses of action: either take step A or take step B, with no other course of action contemplated. Moreover in order to speak of a test of a statistical hypothesis, it is necessary that the desirability of actions A and B depend on the frequency function p(e) of some observable random variables and that p(e) be uncertain*[810]

Significance testing is about testing hypotheses, seeking to reject them. Decision testing is less ambitious. It seeks to indicate the comparative merits of divergent courses of action, on the basis of some relevant empirical evidence. In the bulb example, the choice for the bulb grower is whether to regard the shipment as being composed of 40/60

[809] Fisher's views are set out in Fisher 1956.

[810] Neyman 1950 sect 5.1.2.

red/yellow bulbs or 60/40. With a little ingenuity, most situations can be regarded as involving a choice between two courses of action. In agricultural research for example, an experiment can be undertaken which will help us to decide if it will be more productive to use variety X or variety Y of corn. A clinical trial can be construed similarly, substituting the notion of the efficacy of a drug for the productivity of a seed variety.

In summary, significance tests involve one decision, and are in a sense open. The decision is whether or not to reject a hypothesis. It is open because non-rejection does not imply acceptance. Decision tests involve a binary decision, and are in a sense closed. The decision is whether to accept H_1 or H_2. The decision is closed because decision test theory assumes that one of the hypotheses is better, or more efficient, than the other.

The nature of the test

Decision tests in their simplest form always concern alternative hypotheses, allowing the data to recommend one or the other hypothesis. The decisions carry with them the risk of making the wrong choice. Two risks, or errors, are focussed on: that of taking course B when course A is correct (i.e. regarding a hypothesis as false when it is in fact true (often called Type I error)); and that of taking course A when course B is correct (i.e. regarding a hypothesis as true when it is in fact false (Type II error)).

Somewhat like significance testing, decision tests require a rule to choose between the hypothesis. In this case the grower decides that if the sample contains 6 or more reds he will reject H1. It can be seen that if this rule is applied, the grower is accepting, with a probability of 0.1663, the chance that he will reject H1 when it is true (the sum of the darker shaded cells in Table 3.3). This is the Type I or alpha error. He is also accepting,

with probability 0.3664, that he will reject H2 when it is true to the extent of the lighter shaded cells.

Much of the theory of decision tests is concerned with the management of Type I and II error. It is desirable to minimise the risk of both error. This is not possible however, since for any experiment, as the risk of Type I error decreases, the risk of Type II error increases, and vice versa. The practical routine recommended in textbooks is to fix the Type I risk at some figure (typically 0.05), and then adjust the size of data collection to achieve a desired level of Type II error. This is usually expressed in terms of 1-(Type II error), which is also known as the power of the experiment. As the amount of data collected increases the Type II error decreases and the power increases.

The interpretation of the test result

The frequency of Type I and II errors are calculated on the basis of repeat sampling of a population, and therefore have a strict frequentist interpretation: 'if I make choice A 100 times then on 95 occasions I will be right'. In contrast Fisher conceived of the result of a significance test applying in this instance to a single member of the hypothetical population, provided there are no recognisable subsets of the population to which the member might belong. Where significance tests offer a strongly typed scientific judgement (the rejection of a hypothesis); decision tests offer a choice of behaviour (either A or B) without requiring the experimenter to decide about the truth or falsity of either.

Advantages and validity of decision tests

The use of competing hypotheses creates a framework in which much of the arbitrariness of significance testing is removed.[811] Seidenfeld argues that there are three reasons why decision test theory is preferred to significance test theory: an explicit thoroughgoing frequentist framework; a clear mathematical formulation; and ease of application to many practical cases.[812] Nevertheless, Seidenfeld considers decision testing to be theoretically flawed, and Howson and Urbach have questioned the validity of frequency interpretations embodied in the outcome of decision tests.[813] The next section looks briefly at Fisher's critique of decision tests, which he judged to be radically unsound.

Fisher's critique of decision tests

In order for the mathematics of decision tests to work it is necessary to assume that the empirical evidence is a random sample, taken by some repeatable method, from a larger population, called by Neyman the fundamental probability set. In the result of a decision test this rule (called the repeat sample rule) lends itself to the interpretation of the result in the familiar form: 'given the data if we do X then 95 times out of 100 we will be doing the right/best/correct/true thing'.

Fisher regarded the repeat sample rule as contrary to the enterprise of science. His point was that in science the repeat sample rule is a fiction, since each repetition of an experiment will draw use a sample from what is in effect a new population:

[811] Howson and Urbach 1993 p200-202, and Gigerenzer 1989 p99-102.

[812] Seidenfeld 1979 p30.

[813] Howson and Urbach 1993 p203-208.

> *Pearson and Neyman have laid it down axiomatically that the level of significance of a test must be equated to the frequency of a wrong decision "in repeated samples from the same population". This idea was foreign to the development of tests of significance given by the author [Fisher] in 1925, for the experimenter's experience does not consist in repeated samples from the same population.*[814]

Yet if the repeat sample idea was problematic, it had the virtue of providing a strictly frequentist interpretation of empirical evidence. Fisher's own concept, which he claimed offered a more robust confirmational link between a set of results and a hypothesis, required a 'hypothetical infinite population', a concept just as difficult to map onto the experimenter's experience as fundamental probability sets.

A simple clinical trial using either significance or decision tests

There is a new drug, which cures people of tuberculosis. I want to test this hypothesis. My general approach is to assume that the drug has no effect and to test that hypothesis. If I disprove it I may say, with caution, that the hypothesis that the drug is ineffective has been disproved.

Taking a population, I will subject the members to the drug. I will measure the number of cures by looking at x-rays six months after a course of treatment. Unlike my coin–tossing experiment, I have no frequency distribution for the number of x-ray clearances for the whole population. Instead, I will manipulate the situation to create a parameter 'differences between groups'. I do this by dividing the patients into two groups, and treating only one of the groups. I will compare the number of clearances in one group with that of the other. The parameter 'differences between groups' can take the value 0,

[814] Fisher 1935**b.**

1, 2, 3, etc and I can work out a probability distribution for each of the values. I can then apply a statistical test to the data produced by the trial.

Estimation

Estimation theory tries to supply what hypothesis testing cannot. It tries to supply a value to a parameter, based on some data, and the assumption that the parameter is unknown. Typically, estimates will be set out as the interval between an upper and lower bound, with an attached probability. For example, the mean height of a population lies between 183 cms and 191 cms, with 95% probability.

There are two frequentist theories of how to produce interval estimates: Fisher's fiducial intervals and Neyman/Pearson's confidence intervals, which were both introduced around 1930. Although they appeared to be very similar[815] it is now clear that they offer radically different interpretations. Taking an example, the 95% interval for the mean of a population, calculated on the basis of a known mean of a sample of that population. The interval can be calculated using a formula:

$x \pm 1.96\sigma / \sqrt{n}$

where x = sample mean

σ = standard deviation of population

n = sample size

If a sample of 9 measurements has a mean of 5, and it is known that the standard deviation of the population is 0.5, then the 95% interval:

[815] Fisher for example in 1935: 'It has been, naturally, of great interest to me to follow the attempts which Drs Neyman and Pearson have made to develop a theory of interval estimation independently of some of the concepts I have used. That, whenever unequivocal results have been obtained by both methods they have been identical, is of course, a gratifying confirmation of the hope that we are working along sound lines' Fisher 1935**c.**

= 5 +/- 1.96x0.5 / □9

= 5 +/- 0.327

i.e. the 95% interval for the population mean is 4.673 – 5.327

The fiducial interpretation of this interval is that the true value of the population mean lies between the upper and lower bound of the interval with 95% probability. The confidence interpretation is equally clear but it is **not** that the true mean lies between the upper and lower bound. The interpretation warranted by theory is that this may be (with 95% probability) an interval that contains the true value, rather than the true interval may be contained in this interval. The confidence interpretation is a statement to the effect that, before any sample is drawn, there is a 0.95 relative frequency that an interval calculated using the formula will contain the population mean. Since this is rather abstract from the point of view of practical statistics, Neyman argued that while it was not logically true that any value can be attached to the probability of a parameter's value occurring in a particular interval, the researcher should calculate the interval and then state that the true value lies in the interval with a probability of 0.95. This is reasonable because 'in the long run he will be correct in about 95% of all cases'.[816] While it is generally reckoned that fiducial interval theory is more problematic than the confidence interval theory,[817] the fiducial interpretation of the meaning of an interval is the more intuitive, and is applied to confidence intervals.[818]

[816] Neyman 1937 p263.

[817] Seidenfeld 1979 p107.

[818] Confidence intervals have also been subject to criticism. See Seidenfeld 1978 Howson and Urbach 1993 chapter 10.

Hacking has defended a fiducial interpretation of confidence intervals. He does so by firstly denying that inference is subject to the laws of logic. His argument, hedged about with caution lest it be misinterpreted were 'there is no such thing as a logic of statistical inference'.[819] He goes on to argue that if this is the case, the fault lies with the theory of logic rather than the practice of statistics. As a justification for the theory of estimation he suggests a deliberately diminished syllogistic logic, where the step from calculating an interval to interpreting its meaning has no warrant except what is acceptable:

> *Critics of Neyman and Pearson sometimes say that confidence intervals perpetuate a confidence trick on the innocent research worker. The routine technician conducts an experiment and obtains a 95% confidence interval. But even when this person has been taught somewhere along the line that you cannot attach a 95 per cent probability to the statement, 'θ is in f(x)', what the interval means to the researcher is just, 'the probability that θ is in f(x) is 95 per cent'. That is what the confidence interval feels like to the research worker, and that is how it is often used, or so the critics say. I think it is only the logicist instincts of the critics, and their false view of language, that leads them to impugn the research worker'*[820]

In place of a logic, he later suggested a statistical style of reasoning, within which procedures which lacked a full logical justification were nevertheless capable of producing truthful statements.[821]

Conclusion

Significance testing and decision testing arose in the twentieth century as ways to provide a means of using empirical evidence to support inductive inference without requiring prior subjective estimates of the value of parameters. Fisher created significance tests to give evidence (empirical data) the opportunity to contradict hypotheses. The framework

[819] Hacking 1980 p145.

[820] Hacking 1980 p 152-153. For a recent re-statement of the logical fallacy of a frequentist interpretation of confidence intervals, see Burton 1998.

[821] Hacking 1992.

supporting significance testing, while apparently robust, on closer inspection appears to be logically rather weak and cannot exclude some of the arbitrary subjective elements Fisher set out to eradicate. Decision testing, the theory of Neyman and Pearson, solves many of the unanswered questions posed by significance testing, but offers a less robust form of knowledge, based on what Neyman called inductive *behaviour*. Ultimately, neither significance testing nor decision testing can claim to succeed in offering a complete logic model of reasoning based on inverse inference.

Today neither of the theories of hypothesis testing predominates in the scientific literature. In their place, techniques of estimation are used. Estimation attempts to provide what testing cannot: that is, a reasoned guess at the true value of a parameter, based on empirical data. Using estimates it becomes possible to gauge the magnitude of difference between two hypotheses rather than simply state that the data supports one hypothesis over another.

Several authors consider estimation theory to be as flawed as hypothesis testing from a logical point of view. Yet hypothesis testing and interval estimation continue to be effective ways of generating knowledge. Efficiency and utility are therefore not determined entirely by the internal logical characteristics of hypothesis tests and interval estimation, but by a broader framework. In the case of interval estimation Hacking has defended the intuitive interpretation of confidence intervals by appeal to what he called a stylistic interpretation of statistical knowledge.

Appendix 4

THE ESSENTIAL OILS IN THE TREATMENT OF PUERPERAL FEVER

By H Dove, M.R.C.S. Eng., Norwich

The oil of turpentine has for several years been used in this city and neighbourhood in a great variety of forms of puerperal fever with much advantage, and occasionally with almost magic effect.[822] It is usual to commence the treatment with half an ounce of turpentine and an equal amount of castor oil, repeating a drachm of the former every four hours. I have seen the turpentine fail in cases well suited for its peculiar action, and I have also seen it add to the intensity of the disorder, and hurry on its fatal result.

Considering what a nauseous medicine turpentine is, that it irritates the kidneys, suffuses the eyes and produces more or less head symptoms, I was induced to try, in its stead, the essential oils, selecting that of peppermint, and giving 30 or 40 minims in divided doses during the twenty-four hours. I have now used this oil in seven cases, and in another case, the oil of carraway, with all the advantages and none of the disadvantages of the turpentine. The dull colour of the complexion, oedematous condition of the surface, and offensive evacuations, usually observed in puerperal fever, point out the necessity of commencing the treatment with at least one stimulating dose of aperient.

[822] The use of turpentine to treat puerperal fever is not mentioned in the London Dispensatory. (see Thomson 1833 p501-510).

For illustration, I will detail the most severe of the eight cases. Mrs. G, a delicate woman aged 20, primipara, attended by a midwife, had an easy labour, and did well for four days; but on the fifth she complained of chills, profuse perspirations, headache, intense thirst, vomiting and purging of offensive matters, and pain and distension of the abdomen. She was restless, her countenance was anxious, breathing short and hurried, tongue covered with a white fur, pulse 160; the lochia and urine were scanty; the skin was of dull colour and oedematous. A dose composed of tincture of rhubarb and castor oil, of each half an ounce, with five minims of the oil of peppermint in a little water, was immediately administered, and thirty minims of the oil of peppermint were given in divided doses, during the twenty-four hours. A spirit lotion was applied to the head, and mustard poultices to the abdomen. On the following day, the vomiting had ceased; the headache was relieved, and the pulse was considerably reduced. The purging, thirst, and perspirations continued for a few days and gradually ceased. In this case convalescence was slow, differing from the others, in which convalescence was remarkably rapid. Instead of the loathing usually expressed where turpentine has been used, there was an evident desire to take this oil, and, indeed to continue it, when the necessity for it had ceased. My belief is, that almost all the essential oils would do just as much good as the turpentine, and I do not think that I shall have recourse to the turpentine again.

BMJ 1859(April 9th):287-288

Appendix 5

William Guy's lectures on statistics, 1860

Guy delivered the three Croonian Lectures for 1860 at the Royal College of Physicians. The lectures were printed in six parts in the BMJ, beginning on Saturday May 5th 1860.

Lecture one: In his first lecture Guy establishes mathematics at the head of all knowledge because of its formal perfection. Using the example of a sea-journey conducted across the oceans with unerring accuracy, Guy places astronomy at the head of what he calls the applied sciences, and stellar navigation at the head of the applied arts, because of their 'accuracy, certainty, and precision'. By analogy, any discipline that is able to use numeric devices is capable of the perfect precision of a long sea journey.

Guy next divides applied arts and sciences into those which deal with properties and those which deal with relations. In general, applied sciences deal with properties and applied arts with relations. However, the applied arts each have a corresponding applied science. The applied art of architecture requires applied sciences that deal with the properties of materials; similarly the arts of gardening and farming require the science of botany.

A peculiarity of the applied sciences 'is worth noticing. I mean the wide variations in quality and consistence to which portions of matter bearing the same name, and even found in the same place, are liable.'[823] Accordingly, empirical experiments are needed to test the strength of an iron bar, and the judicious architect or engineer will use average

[823] BMJ 1860;I:333.

values rather than the value of a single bar. Variation is a more marked feature in the applied arts and sciences, which deal with animate objects. Guy uses an example that later formed the basis of RA Fisher's work most notable work on experimental design:

> *'Take again, the case of the farmer. The seed he sows is of variable quality; the land he sows it in consists of a variable composition of soils; the mechanical preparation of the soil is more or less complete; no two specimens of the manure he applies have exactly the same strength…from the same land, with the same preparation, the same quality of the same manure, and, as nearly as he can judge, the same seed, he one year obtains an abundant, another year a scanty, crop.'*[824]

The science and art of medicine take for themselves the most compound and complex object possible – the human body. The behaviour of this body is difficult to predict because any response is determined by the myriad of elements that make up the individual body. 'The sense of difficulty and perplexity grows as we pass from individual to species',[825] since to the complexity of the individual must be added the variability that exists between individuals. Age and sex are only the most immediately obvious of the characters that vary between individuals. But even two individuals of the same age and sex will vary considerably. Variability, in fact, is at the heart of the medical enterprise. It explains why the ordinary results of medical practice and inquiry are variable. Consequently, a method is needed to collect, arrange, classify and analyse facts if they are to advance medical knowledge. The numerical method is the most appropriate one in the arts of agriculture, government, and medicine, where facts and events are the result of 'the combined action of a great number of forces and causes.'[826] The numerical method

[824] BMJ 1860;I:334.

[825] BMJ 1860; I:371.

[826] BMJ 1860:I:373.

has the virtues of 'terseness of expression, clearness of exposition by tabular forms, and precision and truthfulness in setting forth the amount of our experience'.[827] 'Facts are only useful when brought together in masses.'[828] The reason is that individual facts, being the product of a multitude of invisible operations, are variable, and it is only by bringing facts together 'that we can strike an average which shall truly represent the state of the case both absolutely and for the purpose of comparison'.

Guy's first lecture is ingenious. By the time he discusses medicine he has set up a series of examples to which the audience can readily respond. He has created a necessary link between applied sciences and applied arts which respects both. In doing so he resolves the tension between the science of medicine and the art of medicine. He has shown also that variability is an inescapable feature of the applied arts and sciences. His chief message is a warning of the dangers of ignoring the influence of variability. A physician who does so will take credit where he does not deserve it, but more importantly lose it where he is not to blame. The application of the numerical method in medicine form the subject of his next lecture.

Lecture two: 'The first and most obvious principle of the numerical method is, that the individual facts which we bring together, and from which we obtain our averages, should be in everything but the inseparable incident of variable numerical values, counterparts of

[827] BMJ 1860:I:373.

[828] BMJ 1860:I:373.

each other'.[829] Secondly, data should not be lost from a class of persons 'in consequence of their having changed their names.'[830]

The principal statistic used by the numerical method is the average. When properly arrived at, the average is a figure from which conclusions may be drawn. But Guy urges caution. Averages are not always the most appropriate number from which to draw conclusions. Extreme values may be the most valuable, for example in the case of an engineer who will use the minimum strength of some material, and make calculations on the basis of that rather than the average. Similarly, in criminal trials of infanticide, the weight of the lung (used to determine if any breath was drawn) should be compared to the range of possible values of new born and still-born lungs, not their average.

Guy adds a further caution. Results derived from the numerical method do not apply to either individuals or real groups. For example, knowing that deaths in England and Wales amount to one in every 45 of the population does not allow us to assemble 45 persons and know that one will be dead within a year:

> *It must be obvious ... that this large truth, expressed in numbers, has no practical application whatever to any individual man, woman or child; and it could not even be safely applied to any single group of forty-five persons, though closely resembling ... the population of England and Wales of which it forms a part'.*[831]

A multitude of individual facts is needed to form a proper value for an average or an extreme result. Consideration of how life tables or insurance tables are constructed, Guy

[829] BMJ 1860:I:409.

[830] BMJ 1860:I:409.

[831] BMJ 1860;i:411.

argues, shows the necessity for taking into account many individual facts if values for the group are to be reliable.

Guy next discusses how many individual facts are needed. His first answer is that the number of observations should equal the range of values possible. If a sample of 25 persons, as similar as it is possible to make them, have a range of pulse from 46 to 92 beats per minute, it follows that 46 (rather than 25) individual facts are needed to obtain an accurate average. Guy next shows that as the number of facts decreases, variation in the summary figure increases. He uses the example of the average age at death of the English aristocracy, which is 60½, calculated from 1,600 deaths recorded in the Annual Register. If the class is divided into two groups of 800, two averages can be obtained. On inspection these differ from the whole class average by little more than a year. Further divisions into groups of 400, 200, 100, 50, and 25 facts shows the range in average obtained increasing on each successive division.

Guy's second answer to the question of how many facts are needed to create a true average is that such calculations are difficult, despite the work of Gavarret.[832] It is important to find a way of dealing with small numbers of facts if the numerical method is to be applicable to medicine.

Lecture three: Guy calculates the incidence with which the sample average corresponds with the population average. Using an example involving 6400 deaths, on no less than 1

[832] For a discussion of Gavarret, see Rosser Matthews 1992 (Thesis) p62-78. Gavarret devised an equation for calculating the limits of oscillation of a statistical average. The equation was a formalised version of Guy's demonstration that the variation between extremes increased as sample size decreased. Gavarret in effect devised a confidence interval for averages. When applied to Louis' study on typhoid, it was clear that Louis' sample size of 140 patients was too small to give confidence it the validity of the results obtained. Gavarret is perhaps the father of meta-analysis, since he recommended that several hundred trials were necessary before a result could be established by the numerical method (Rosser Matthews 1992 p65-66).

in 6 occasions does a sample of 50 deaths have the same average as the population average. From this Guy concludes:

> *'See that there are so many instances of coincidence between the averages derived from small numbers of fact and the true average, and so many other instances in which the averages of small numbers of facts differ but little ... we shall certainly be justified in making use of these ... provided that we speak of the evidence they afford with due reserve'*[833]

Guy next discusses real applications of the numerical method in medicine. There are three groups into which they can be divided. In the first group are topics whose statistics are so plain that a small amount of data is sufficient to provide conclusive evidence. Guy cites his work on the health of compositors included in the First Report of the Health of Towns Commission, 1844. In the second group are topics that require laborious, subtle and skilful treatment. This group comprises the bulk of epidemiological and therapeutic inquiries. The third group are those topics which concern unlikely hypotheses – 'cases in which we distrust either the sanity or the honesty of the person who proposes a remedy or preventive.'[834] In this group Guy includes the example of a quasi-randomised controlled trial of belladonna as a treatment for Scarlet fever:

[833] BMJ 1860;(July 21):553.

[834] BMJ 1860;I:554 (col. 1).

> *An experiment on a very small scale is all that the supporters of Hahnemann's hypothesis ... have a right to expect at our hands. Such an experiment, devised with true logical acumen, and carried out with corresponding care, was made by my friend Dr Balfour, at the Royal Military Asylum at Chelsea. Scarlet fever having broken out in the institution, Dr Balfour took 151 boys ... and divided them into two sections, taking them alternately from the list, to prevent the imputation of selection. To the first section (76) he gave belladonna; to the second (75) he gave none; the result was that two in each section was attacked with the disease.*[835]

The middle part of Guy's third lecture consists of arguments in support of his assertion that small-scale statistics can yield useful results. To the examples of the compositors and belladonna, Guy added a survey of the health of journeyman bakers, and the mortality of workers associated with the brewing industry compared to other labourers. In all cases statistics derived from a small number of observations were sufficient to confirm the hypothesis under test.

In the final part of his third lecture Guy begins by observing the stability in the frequencies of certain phenomena – marriage, suicide. He considers Quetelet's inference – that such results drain social phenomena of free will – as dangerous. Medical statistics show the fallacy of Quetelet's conclusion. Firstly, as he has already demonstrated, statistical stability is the inevitable result when large numbers of individual facts are brought together. Secondly, in the causes of death that are clearly independent of human volition, there is more variation in rates of death than in those that are clearly caused by human volition. For example, 'the highest rate of fluctuation belong to epidemic and contagious maladies',[836] while deaths by violence and privation show relatively little

[835] BMJ 1860;i:554.

[836] BMJ 1860;i:595.

variation. In summary, Quetelet and Buckle are wrong to deny the role of human volition in statistical work.

Guy's conclusions are that:

1. The numerical method admits of most useful application where we can apply the results obtained from a considerable body of facts to a similar large body of facts of the same order.
2. The average results which we obtain from large numbers of facts, as well as the extreme values, are as useful, for all purposes of comparison and reasoning , as any other ascertained truths, whatever the shape which they assume.
3. The numerical method admits of very limited use in actual practice.
4. The numerical method is to be commended as contributing to precision, accuracy, and truthfulness.
5. Average results obtained from small bodies of facts may be employed with advantage to test and refute opinions carelessly adopted and rashly propounded; to establish new truths when the figures exhibit themselves in great uniformity, or when compared with similar small collections of facts, if the numerical results are widely divergent; and lastly, to indicate possibilities to be afterwards confirmed or invalidated by more extensive induction.

The series concludes with a description of a clinical trial carried out in 1536 by Ambrose Pare, concerning the use of different dressings for gunshot wounds.

Appendix 6

Substances submitted to the TTC for clinical trial

Mtg.	Date	Application	Company	Indication	Outcome
1	08/07/1931	Nonyl-harmol hydrochloride	Boots Pure Drug	Amoebic disease	Accepted
1	08/07/1931	Amyl-meta-cresol	Boots Pure Drug	Urinary antiseptic	Accepted
1	08/07/1931	Propyl guaiacol	Boots Pure Drug	Anti-helminthic	Declined
1	08/07/1931	Harmol hydrochloride	Boots Pure Drug	Coronary disease	Accepted
1	08/07/1931	Ergotoxine ethanesulphonate	Burroughs Wellcome	Uterine stimulation	Accepted
1	08/07/1931	Digoxin	Burroughs Wellcome	Heart disease – cases of auricular fibrillation	Accepted
1	08/07/1931	Digitalinum verum	Burroughs Wellcome	Heart disease – cases of auricular fibrillation	Accepted
1	08/07/1931	Halarsol	May & Baker Ltd	Syphilis	Provisional
1	08/07/1931	Parosan	May & Baker Ltd		Declined
1	08/07/1931	antipneumococcal serum	MRC sponsored	lobar pneumonia	Accepted
2	15/01/1932	Acetyl b-oxynapthoic acid	BDH and ICI	pain relief in inoperable cancer	Accepted
2	15/01/1932	Ephedrine vs. pseudo ephedrine	Burroughs Wellcome	asthma	Accepted
2	15/01/1932	Ephedrine vs. pseudo ephedrine	Burroughs Wellcome	asthma	Accepted
2	15/01/1932	Calciferol	British Drug Houses	rickets	Accepted
2	15/01/1932	Amyl salicylate	A Boake, Roberts	topical treatment of burns	Accepted
2	15/01/1932	hog's stomach	referred by Council	pernicious anaemia	Declined
2	15/01/1932	Oestrin	Sex Hormones Committee	ovarian dysfunction	Accepted
3	08/07/1932	Proviron	Schering-Kahlbaum	bilateral orchidectomy, and possibly cases of premature senility and impotence, possibly prostatic enlargement	Accepted
3	08/07/1932	o-n-propylharmol lactate	Boots Pure Drug Co	angina	Accepted
3	08/07/1932	Quinanil	Chemotherapy Committee	antiseptic	Provisional
4	27/03/1933	Khellavis	Upsher-Smith	ureteric calculi	Accepted
4	27/03/1933	Sodium phenanthridine-9-carboxylate	Prof. GT Morgan	cinchonphen substitute	Declined
4	27/03/1933	Pentnucleotide	Smith Kline & French and Manley & James	principally, for induction of artificial leucocytosis in preparation for surgical operations	Accepted

4	27/03/1933	Preparine	British Drug Houses	spasmolytic	Deferred
5	05/03/1934	Hepamult	HR Napp Ltd		Declined
5	05/03/1934	Prostigmin	Hoffman La Roche	postoperative atony of the intestine	Accepted
5	05/03/1934	Progestin	Organon Laboratories	various gynaecological disorders	Accepted
5	05/03/1934	Cortin	Organon Laboratories	suprarenal deficiency	Accepted
5	05/03/1934	Avenyl Cream	Burroughs Wellcome	syphilis	Accepted
5	05/03/1934	Solu-salvarsan	Bayer Products Ltd	syphilis	Accepted
5	05/03/1934	Concentrate of Vitamin A and D	Glaxo Laboratories	deficiency disease	Rejected
5	05/03/1934	pseudo-ephedrine	Burroughs Wellcome	myasthenia gravis	Accepted!
5	05/03/1934	Synotropan	Hoffman La Roche	Atropine like action. Suggested use for bladder pain	Accepted
5	05/03/1934	Staphylococcus Toxoid	Burroughs Wellcome	Chronic staphylococcal skin infections	Accepted
5	05/03/1934	Profundol	HR Napp Ltd		Declined
5	05/03/1934	Hombreol	Organon Laboratories	Prostatic enlargement	Deferred
5	05/03/1934	Staphylococcus Antitoxin	Lister Institute	Acute osteomyelitis; staphylococcal septicaemia	Accepted
6	28/02/1936	Merthiolate	Captain DP Lambert	Tuberculosis	Accepted
6	28/02/1936	Doryl	Dr J Chasser Moir	Postoperative retention of urine	Accepted
6	28/02/1936	Neocryl	ABCM	Syphilis	Accepted
6	28/02/1936	Testosterone	Ciba	Prostatic enlargement	Accepted
6	28/02/1936	Tetra-n-amyldiaminodecane	Boots Pure Drug	Amoebicide	Accepted
6	28/02/1936	Streptozon S	Bayer Products Ltd	Streptococcal infections, esp. puerperal fever	Accepted
6	28/02/1936	Bismutrat	Wilcox, Jozeau & Co	Syphilis	Accepted
6	28/02/1936	Eustab	Boots Pure Drug	Syphilis	Accepted
6	28/02/1936	Mapharsan	Parke Davis & Co	Syphilis	Accepted
6	28/02/1936	Prontosil	Bayer Products Ltd	Streptococcal infections, esp. puerperal fever	Accepted
7	11/02/1937	Helborsid	Roche Products Ltd		Deferred
7	11/02/1937	Preparation 2020	Ciba	Asthma	Accepted
7	11/02/1937	Quindoline methochloride	ICI Ltd	External antiseptic	Accepted
7	11/02/1937	Tussipect	Beiersdorf Ltd		Declined
7	11/02/1937	Adovern	Roche Products Ltd		Declined
8	07/02/1938	Jensen Diphtheria Prophylaxis	British Drug Houses	Diphtheria	Declined

8	07/02/1938	Vitamin D3	Vitamins Committee	Rickets	Accepted
8	07/02/1938	Eupaverin	E. Merck		Declined
8	07/02/1938	Methyl Isomyn	Burroughs Wellcome	Benzedrine like effects	Accepted
8	07/02/1938	Sodium thioethamyl	Parke Davis & Co	Anaesthetic	Provisional
8	07/02/1938	Preparation 3259	Ciba	Vascular disease esp. Raynaud's syndrome	Accepted
9	14/07/1938	Morpholine nicotinamide	Chase Laboratories		Declined
9	14/07/1938	Sulphanilic-acid-4-acetanilide	E. Merck	Streptococcal infection	Provisional
9	14/07/1938	Preparation 2834/35	Ciba	Histamine like effect	Provisional
9	14/07/1938	Trichlorethanol	E. Merck	Anaesthetic	Accepted
9	14/07/1938	Desoxycorticosterone acetate	Ciba	Supra-renal deficiency	Accepted
9	14/07/1938	Diethylstilboestrol	British Drug Houses	Amenorrhoea, menopause, abortion	Accepted
10	28/03/1939	Hexoestrol	Boots Pure Drug	Oestrogen deficiency	Accepted
10	28/03/1939	Pyridacil	Cilag Chemisches		Declined
10	28/03/1939	M&B 693 Dagenan	May & Baker Ltd	Puerperal fever	Unclear

Appendix 7

9/11/31

Medical Research Council

SERUM TREATMENT OF LOBAR PNEUMONIA

Standard scheme of inquiry to be used by the different investigators

(1) Typing of organisms: All cases to be typed. When possible the 'rapid method' should be used before treatment is begun. Failing this, typing should be done as soon as possible after the first dose of serum has been given.

(2) Choice of cases: Age limit 20 to 60, classified separately by decades. Patients over 60 not to be included in statistical series of treated cases, or controls. Cases admitted after the 5th day of the disease and patients admitted 'moribund' to be excluded from published series of both patients and controls.

(3) Controls: With above reservations, alternate cases are to be treated with serum, the remainder classed as controls.

(4) Full clinical records to be kept of every case, including:-

Date and mode of onset

Evidence of general physical condition, nutrition, alcoholism. Condition on examination

(5) Blood cultures: When possible, and always in seriously ill cases. Cases with positive blood-culture to be classified separately in publications.

(6) Publications: Reports on the results of work to be submitted to the Council and, when approved, may be published by the authors as 'Reports to the Therapeutic Trials Committee of the Medical Research Council'.

Glossary

ABCM	Association of British Chemical Manufacturers
BDH	British Drug Houses Ltd.
BMJ	British Medical Journal
CC	Chemotherapeutic Committee of the MRC
DNB	Dictionary of National Biography
DSIR	Department of Scientific and Industrial Research
LSHTM	London School of Hygiene and Tropical Medicine
MAB	Metropolitan Asylums Board
MoH	Ministry of Health
MRC	Between 1913 and 1920, refers to the Medical Research Committee. From 1920, refers to The Medical Research Council.
NIMR	National Institute for Medical Research
NPL	National Physical Laboratory
PAS	Para amino salicylic Acid
PMSA	Provincial Medical and Surgical Association
PRO	Public Record Office, Kew
SRU	Statistics Research Unit of the MRC
TRC	Therapeutic Research Corporation
TTC	Therapeutic Trials Committee of the MRC
USPHS	United States Public Health Service
VA	Veterans Administration
WRPL	Wellcome Physiological Research Laboratories

Bibliography of works consulted

Archival sources

Hogben papers, University of Birmingham Library.

MRC papers, Public Record Office, Kew.

Interview transcripts

Lindley D. V. 1995. Interview by B Toth.

D'Arcy Hart, P. A. 1996. Interview by A Yoshioka.

Bibliographies and reference works

Bynum, W. F. and R. Porter. 1993. Companion encyclopaedia to the history of medicine (2 vols). London: Routledge.

Harper, P. 1993. *Guide, 1973-93.* Bath: National Cataloguing Unit for the Archives of Contemporary Scientists.

Higby, G. J., and E. C. Stroud. 1995. *The history of pharmacy: a selected annotated bibliography.* New York: Garland Press.

Jordanova, L. 1983. The social sciences and the history of science and medicine. In *Information sources in the history of science and medicine*, ed. P. Corsi and P. Weindling, 81-96. London: Butterworth.

Matthews, L. G. 1962. *History of pharmacy in Great Britain.* Edinburgh: Churchill Livingstone.

Neu J ed. 1980-85 *Isis cumulative bibliography, 1966-75*. London: Mansell.

Neu J. ed. 1990. *Isis cumulative bibliography, 1976-85*. London: Mansell.

Stephen, L. and S. Lee eds. 1885-1900. *Dictionary of National Biography*. London Macmillan.

Surgeon General's Office 1899. *Index catalogue of the library of the Surgeon-General's Office*. Second Series. Washington: Government Printing Office.

Waring, E. J. 1878. *Bibliotheca therapeutica or, bibliography of therapeutics.* London: New Sydenham Society.

Whitrow M ed. 1971-84 *Isis cumulative bibliography*. *A bibliography of the history of science formed from Isis critical bibliographies 1-90, 1913-65*. London: Mansell.

Theses

Barr, B. C. 1992. Spare children, 1900-1945: inmates of orphanages as subjects of research in medicine and in the social sciences in America. Ph.D. Dissertation, Stanford University. (a copy is available at the Wellcome Library in London)

Bull, J. P. 1951. A study of the history and principles of clinical therapeutic trials. MD Thesis, University of Cambridge.

Epstein S. 1993. Impure science: AIDS, activism and the politics of knowledge. Ph.D. Dissertation, Berkeley: University of California.

Liebenau, J. 1981. Medial science and medical industry, 1890-1929. Ph.D. Dissertation, Pennsylvania: University of Pennsylvania.

Magnello, M. E. 1993. Karl Pearson: evolutionary biology and the emergence of a modern theory of statistics. D Phil Thesis, Oxford: University of Oxford.

Marks, H. M. 1987. Ideas as reforms: therapeutic experiments and medical practice, 1900 - 1980. Ph.D. Dissertation, Cambridge, Massachusetts: Massachusetts Institute of Technology.

Petty, E. C. 1987. The impact of the newer knowledge on nutrition: nutrition science and nutrition policy 1900-1939 Ph.D. Dissertation, London: University of London.

Robson, M. 1993. The pharmaceutical industry in Britain and France 1919-1939. Ph.D. Dissertation, London: University of London.

Rosser Matthews, J. 1992. Mathematics and the quest for medical certainty: the emergence of the clinical trial 1800-1950. Ph.D. Dissertation, Emory: Duke University. Available from UMI Microfilms. A copy is held at the Wellcome Library, London.

Swann, J. P. 1985. The emergence of cooperative research between American universities and the pharmaceutical industry, 1920-1940. Ph.D. Dissertation, Madison: University of Wisconsin.

Tansey, E. M. 1990. The early scientific career of Sir Henry Dale FRS (1875-1968). Ph.D. Dissertation, London: University of London.

Tröhler, U. 1979. Quantification in British medicine and surgery, 1750-1830, with special reference to its introduction into therapeutics. Ph.D. Dissertation, London: University of London.

Varcoe I. M. 1972. The DSIR [Department of Scientific and Industrial Research]: a study in the growth of organised knowledge. Ph.D. Dissertation, Oxford: Oxford University.

Journals read

American Review of Tuberculosis 1930 - 1950

British Journal of Tuberculosis 1930 – 1950

(From 1943 known as British Journal of Tuberculosis and Diseases of the Chest)

British Medical Journal 1858 – 1882

Tubercle 1930 - 1950

Published sources

Abel-Smith, B. 1964. *The hospitals, 1800-1948.* London: Heinemann.

---. 1994. The Beveridge Report: its origin and outcomes. In *Beveridge and social security: an international perspective*, ed. J. Hills, J. Ditch, and H. Glennerster, 10-22. Oxford: Clarendon Press.

Abelson, R. P. 1995. *Statistics as principled argument.* New York: Lawrence Erlbaum.

Abraham, J. 1994. Bias in science and medical knowledge: the Opren controversy. *Sociology* 28:717-36.

---. 1995. The production and reception of scientific papers in the academic industrial complex: the clinical evaluation of a new medicine. *British Journal of Sociology* 46:167-90.

---. 1995. Science, politics and the pharmaceutical industry. London: UCL Press.

Ackerknecht, E. H. 1946. Incubator and taboo. *Journal of the History of Medicine* 1:144-48.

---. 1948. Anticontagionism between 1821 and 1867. *Bulletin of the Institute of the History of Medicine* 22:562-93.

---. 1950. Elisha Bartlett and the philosophy of the Paris clinical school. *Bulletin of the History of Medicine* 24:60.

---. 1962. Aspects of the history of therapeutics. *Bulletin of the History of Medicine* 36:389-419.

---. 1967. A plea for a behaviourist approach to the history of medicine. *Journal of the History of Medicine* 22:211-14.

Alborn, T. L. 1996. The business of induction: industry and genius in the language of British scientific reform, 1820-1840. *History of Science* xxxiv:91-121.

Allen, G. 1976. Genetics, eugenics, and society: internalists and externalists in contemporary history of science. *Social Studies of Science* 6:105-22.

Alleyne, J. 1733. *A new English dispensatory, in four parts.* London: Thomas Astley and S Austen.

Allison, S. T., and J. M. N. Nilsson. 1947. Treatment of pulmonary tuberculosis with streptomycin. *American Review of Tuberculosis* 56:579-88.

Almeder, R. 1980. *The philosophy of Charles S. Peirce: a critical introduction.* Oxford: Basil Blackwell.

Altman, D. G., and C. J. Doré. 1990. Randomisation and baseline comparisons in clinical trials. *Lancet.* 335:149-53.

Alvarez-Dardet, C., and M. T. Ruiz. 1993. Thomas McKeown and Archibald Cochrane: a journey through the diffusion of their ideas. *BMJ* 306:1252-54.

Amberson, J. B., B. T. McMahon, and M. Pinner. 1931. A clinical trial of sanocrysin in pulmonary tuberculosis. *American Review of Tuberculosis* 24:401-35.

American Trudeau Society. 1947. Report of the Committee on Therapy and its Subcommittee on Streptomycin. *American Review of Tuberculosis* 56:477-81.

---. 1947. Report of the Laboratory Subcommittee on Streptomycin. *American Review of Tuberculosis* 56:481-83.

Amsterdamska, O. 1987. Medical and biological constraints: early research on variation in bacteriology. *Social Studies of Science* 17:567-687.

---. 1991. What is a virus? The case of Tobacco Mosaic Disease. *Studies in History and Philosophy of Science* 22:557-88.

---. 1993. From pneumonia to DNA: the research career of Oswald T. Avery. *Historical Studies in the Physical and Biological Sciences*:1-39.

Anderson, W. 1992. The reasoning of the strongest: the polemics of skill and science in medical diagnosis. *Social Studies of Science* 22:653-84.

Anon 1842. Medical statistics? *Lancet.* i:

---. 1881. Editorial. *Lancet.* ii:424-25.

---. 1923. The work of the Tuberculin Dispensary League. *Lancet.* ii:984-88.

---. 1925. Therapeutic Substance Act. 15&16 Geo 5 Chapter 60.

---. 1927. Treatment of tuberculosis by tuberculin [editorial]. *BMJ* ii:69-70.

---. 1931. Clinical trials of new remedies. *Lancet.* ii:304.

---. 1936. An earlier British Medical Journal - a prospectus issued in 1798. *BMJ* ii:777.

---. 1937. Mathematics and medicine. *Lancet.* i:31.

---. 1950. 50 years of statistics. *BMJ* i:68.

---. 1950. Streptomycin and PAS [editorial]. *Tubercle* xii

---. 1954. Obituary: Thomas Carnwath. *BMJ* i:937.

---. 1964. Obituary of Francis Fraser. *BMJ* ii:950.

---. 1977. Randomized clinical trials [editorial]. *BMJ* i:1238-39.

---. 1984. A fair trial [editorial]. *BMJ* 289:336-37.

---. 1987. The controlled clinical trial: consensus? [editorial]. *Lancet.* 1:547.

---. 1992. On the falsification of ideas [editorial]. *Lancet.* 340:344-45.

---. 1992. Blind in the face of evidence [noticeboard]. *Lancet.* 340:1344-45.

---. 1993. Clinical trials and clinical practice [editorial]. *Lancet.* 342:877-78.

---. 1994. Blinded by science [editorial]. *Lancet.* 343:553-54.

Apple, R. D. 1980. To be used only under the direction of a physician: commercial infant feeding and medical practice. *Bulletin of the History of Medicine* 54:402-17.

---. 1988. 'They need it now': science, advertising and vitamins, 1925-1940. *Journal of Popular Culture* 22:65-83.

---. 1993. Science in the marketplace: Acnotabs and the Food and Drug Administration. *Public Understanding of Science* 2:59-70.

Armitage, P. 1960. The construction of comparable groups. In *Controlled clinical trials. Papers delivered at the conference convened by The Council for International Organizations of Medical Sciences*, ed. A. Bradford Hill, 14-18. Oxford: Blackwell Scientific.

---. 1971. *Statistical methods in medical research.* Oxford: Blackwell.

---. 1982. The role of randomization in clinical trials. *Statistics in Medicine* 1:345-52.

---. 1983. Trials and errors: the emergence of clinical statistics [Presidential address delivered to the Royal Statistical Society June 22 1983]. *Journal of the Royal Statistical Society A* 146:321-34.

---. 1995. Before and after Bradford Hill: some trends in medical statistics. *Journal of the Royal Statistical Society A* 158:143-53.

Armstrong, D. 1977. Clinical sense and clinical science. *Social Science and Medicine* 11:599-601.

---. 1983. *Political anatomy of the body: medical knowledge in Britain in the twentieth century.* Cambridge: Cambridge University Press.

Aronson, N. 1982. Nutrition as a social problem: a case study of entrepreneurial strategy in science. *Social Problems* 29:474-87.

Aschner, B. 1945. Empiricism and rationalism in past and present medicine. *Bulletin of the History of Medicine* 17:269-86.

Assinder, E. W. 1938. Treatment of syphilis. *Birmingham Medical Review* 13:16-23.

Ashcroft, R. E. et al 1997 Implications of socio-cultural contexts for ethics of clinical trials. *Health Technology Assessment* 1(9).

Atherton, C. R. 1993. From science to evidence - the testimony on causation in the Benedictin cases. *Stanford Law Review* 46:1-86.

Austoker, J. 1989 Walter Morley Fletcher and the origins of a basic biomedical research policy. In Austoker, J., and L Bryder. Eds. *Historical perspectives on the role of the MRC: essays on the history of the Medical Research Council of the United Kingdom and its predecessor, the Medical Research Committee, 1913-53* Oxford: Oxford University Press. p23-33.

Austoker, J., and L. Bryder, eds. 1989**a** *Historical perspectives on the role of the MRC: essays on the history of the Medical Research Council of the United Kingdom and its predecessor, the Medical Research Committee, 1913-53.* Oxford: Oxford University Press.

Austoker, J., and L. Bryder. 1989**b** The National Institute for Medical Research and related activities of the MRC. In Austoker, J., and L. Bryder, eds. *Historical perspectives on the role of the MRC: essays on the history of the Medical Research Council of the United Kingdom and its predecessor, the Medical Research Committee, 1913-53.* Oxford: Oxford University Press. p35-57.

Ayer, A. J. 1972. *Probability and evidence.* London: Macmillan.

Ayers, G. M. 1971. *England's first state hospitals and the Metropolitan Asylums Board 1867-1930.* London: Wellcome Institute of the History of Medicine.

Bäulmer, E. 1984. *Paul Ehrlich: scientist for life.* London: Holmes & Meier.

Bailar, J. C. 1983. Introduction. In *Clinical trials: issues and approaches*, ed. S. H. Shapiro and T. A. Louis, 1. New York: Marcel Drekker.

Baird, D. 1983. The Fisher/Pearson chi squared controversy: a turning point for inductive inference. *British Journal of the Philosophy of Science* 34:105-18.

Bakan, D. 1966. The test of significance in psychological research. *Psychological Bulletin* 66:423-37.

Baker, W. 1912. The Boys' Garden City. *The Child*:60-64.

Balas, E. A., S. M. Austin, and B. G. Ewigman. 1995. Methods of randomized controlled clinical trials in health services research. *Medical Care* 33:687-99.

Baldry, P. E., J. J. Sumner, and R. L. Ward. 1952. The place of extrapleural pneumothorax in the modern treatment of pulmonary tuberculosis. *Tubercle* 43:336-41.

Baldwin, M. 1995. The snakestone experiment: an early modern medical debate. *Isis* 86:394-418.

Bardswell, N. D. 1914. *Preliminary report of the treatment of pulmonary tuberculosis with tuberculin. With a prefaratory note by Professor Karl Pearson.* London: HK Lewis.

Bardswell, N. D., and J. H. R. Thompson. 1919. *Pulmonary tuberculosis: mortality after sanatorium treatment.* MRC Special Report Series, 33. London: HMSO.

Barnard, C. 1865. *An introduction to the study of experimental medicine.* New York: Dover.

Barnard, G. A. 1947. The meaning of a significance level. *Biometrika* 34:169-82.

Barnard, G. A., and D. R. Cox. 1962. *The foundations of statistical inference: a discussion.* London: Methuen.

Barnes, B., ed. 1969. *Sociology of science.* Harmondsworth: Penguin.

---. 1977. *Interests and the growth of knowledge.* London: Routledge and Kegan Paul.

---. 1982. *T.S. Kuhn and social science.* Basingstoke: Macmillan.

---. 1985. *About science.* Oxford: Basil Blackwell.

Barnes, B., and S. Shapin. 1979. *Natural order: historical studies of scientific culture.* London: Sage.

Barnes, D. S. 1992. The rise or fall of tuberculosis in Belle-Epoque France: a reply to Allan Mitchell. *Social History of Medicine* 5:279-90.

---. 1995. *The making of a social disease: tuberculosis in nineteenth century France.* Berkeley: University of California Press.

Baron, J. H. 1982. James Lind would not have approved [letter]. *Lancet.* i:1313.

Barone, J., and A. Novikoff. 1978. A history of the axiomatic formulation of probability from Borel to Kolmogorov. Part 1. *Archives of the History of the Exact Sciences* 18:123-90.

Barrow, L. 1991. Why were most medical heretics at their most confident around the 1840s? In *British medicine in an age of reform*, ed. R. French and A. Wear, 165-85. London: Routledge.

Barry, J. 1987. Publicity and the public good: presenting medicine in eighteenth-century Bristol. In *Medical fringe & medical orthodoxy*, ed. W. F. Bynum and R. Porter, 29-39. London: Croom Helm.

Barthes, R. 1988 The old rhetoric: an aide-memoire. In Barthes, R. *The semiotic challenge*. Oxford: Blackwell.

Bartlett, M. S. 1981. Pearson, Egon Sharpe, 1895-1980. *Biometrika* 68:1-7.

Bartley, M. 1990. Do we need a strong programme in medical sociology. *Sociology of Health and Illness* 12:371-90.

---. 1992. *Authorities and partisans: the debate on unemployment and health.* Edinburgh: Edinburgh University Press.

Bartrip, P. J. 1990. *Mirror of medicine: a history of the British Medical Journal*. Oxford: Clarendon.

---. 1992. The British Medical Journal: a retrospect. In Bynum W. F., Lock, S. and R. Porter eds. *Medical journals and medical knowledge*. London: Routledge.

Basu, D. 1980. Randomization analysis of experimental data: the Fisher randomization test. *Journal of the American Statistical Association* 75:575-82.

Bean, W. B. 1977. Walter Read and the ordeal of human experiments. *Bulletin of the History of Medicine* 51:75-92.

Beecher, H. K. 1955. The powerful placebo. *JAMA* 24 Dec:1602-6.

---. 1966. Ethics and clinical research. *New England Journal of Medicine* 274:1354-60.

---. 1970. *Research and the individual: human studies.* Boston, MA: Little Brown.

Beer, J. J. 1958. Coal tar dye manufacture and the origins of the modern industrial research laboratory. *Isis* 49:123-31.

Bell, J., and T. Redwood. 1880. *Historical sketch of the progress of pharmacy in Great Britain.* London: Pharmaceutical Society of Great Britain.

Bennett, J. H., ed. 1972-74. *Collected papers of RA Fisher (5 vols.).* Adelaide: University of Adelaide. Individual papers consulted are listed below under Fisher, RA.

---. 1972. Bibliography of RA Fisher. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 11-22. Vol. 1. Melbourne: University of Melbourne Press.

Berliner, H. 1985. *A system of scientific medicine. Philanthropic foundations in the Flexner era.* London: Tavistock Press.

Berman, A. 1963. Conflict and anomaly in the scientific orientation of French pharmacy 1800-1873. *Bulletin of the History of Medicine* 37:440-62.

Birnbaum, A. 1962. On the foundation of statistical inference. *Journal of the American Statistical Association* 57:269-306.

---. 1969. Concepts of statistical evidence. In *Philosophy, science and method*, ed. P. Morgenbesser. New York: St Martins Press.

Bland, J. M. 1985. Is the clinical trial evidence about new drugs statistically adequate? *British Journal of Clinical Pharmacology* 19:15-60.

Bliss, M. 1988. *The discovery of insulin.* rev ed. London: Faber.

Bloom, B. S. 1986. Controlled studies in measuring the efficacy of medical care: a historical perspective. *International Journal of Technology Assessment in Health Care* 2:299-310.

Boissel, J. P. 1989. Impact of randomized clinical trials on medical practice. *Controlled Clinical Trials* 10 (Suppl):120S-134S.

Bollet, A. J. 1973. Pierre Louis: the numerical method and the foundation of modern medicine. *American Journal of Medical Science* 266:92-101.

Bonchek, L. I. 1979. Are randomized trials appropriate for evaluating new operations? *New England Journal of Medicine* 301:44-45.

Boring, E. G. 1954. The nature and history of experimental control. *American Journal of Psychology* 67:573-89.

Borrell, M. 1976. Organotherapy, British physiology, and discovery of internal secretions. *Journal of the History of Biology* 50:235-68.

---. 1985. Organotherapy and the emergence of reproductive endocrinology. *Journal of the History of Biology* 18:1-30.

Box, J. F. 1978. *RA Fisher: the life of a scientist.* New York: John Wiley.

---. 1980. RA Fisher and the design of experiments. *American Statistician* 34:1-7.

Boyd Orr, J. 1928. Milk consumption and the growth of schoolchildren. *Lancet.* i:202-3.

Bradford Hill, A. 1925. *Internal migration and its effects upon the death rates: with special reference to the county of Essex.* Medical Research Council Special Report Series, 95. London: HMSO.

---. 1937. *Principles of medical statistics.* London: The Lancet Ltd.

---. 1947. Statistics in the medical curriculum. *BMJ* ii:366.

---. 1951. The clinical trial. *British Medical Bulletin* 7:278-82.

---. 1952. The clinical trial. *New England Journal of Medicine* 247:113-.

---. 1952. Chemotherapy of pulmonary tuberculosis in young adults. An analysis of the combined results of three Medical Research Council trials. *BMJ* i:1162.

---. 1953. Observation and experiment. The Cutter lecture in preventive medicine. *New England Journal of Medicine* 248:995.

---. 1953. *The philosophy of the clinical trial.* Public Health Service Publications, Number 388. Washington: US Government Printing Office. (Contained in Bradford Hill 1962)

---. 1954. Cortisone and aspirin in rheumatoid arthritis [letter]. *BMJ* i:1437.

---. 1958. The Harben Lectures, 1957. The experimental approach in preventive medicine. *Journal of the Royal Institute of Public Health* 21:177, 185,209.

---, ed. 1960. *Controlled clinical trials. Papers delivered at the conference convened by The Council for International Organizations of Medical Sciences [Vienna 1959]*. Oxford: Blackwell Scientific Publications.

---. 1960. Aims and ethics. In *Controlled clinical trials. Papers delivered at the conference convened by The Council for International Organizations of Medical Sciences*, ed. A. Bradford Hill, 3-7. Oxford: Blackwell Scientific.

---. 1960. Conclusion: the statistician. In *Controlled clinical trials. Papers delivered at the conference convened by The Council for International Organizations of Medical Sciences*, ed. A. Bradford Hill, 168-71. Oxford: Blackwell Scientific.

---. 1962. *Statistical methods in clinical and preventive medicine.* Edinburgh: Livingstone.

---. 1963. Medical ethics and controlled trials. *British Medical Journal* i:1043-49.

---. 1969. The power and limitations of statistical methods. In *Two meetings: 'People and Cancer' and 'The problem of relevance in cancer research'*, ed. The British Cancer Council, 79-82. London: British Cancer Council.

---. 1971. *Principles of medical statistics.* 9th edition. London: The Lancet Ltd.

---. 1990. Suspended judgement. Memories of the British streptomycin trial in tuberculosis. The first randomized clinical trial. *Controlled Clinical Trials* 11:77-79.

Bradford Hill, A., and R. Doll. 1954. The mortality of doctors in relation to their smoking habits. A preliminary report. *BMJ* i:1451.

Braithwaite, R. B. 1955. *Scientific explanation.* Cambridge: Cambridge University Press.

Brakel, J. van. 1976. Some remarks on the prehistory of the concept of statistical probability. *Archives of the History of the Exact Sciences* 16:119-36.

Bramwell, J. C., M. Campbell, and W. Evans. 1933. Harmol hydrochloride and o-n-propylharmol lactate in angina pectoris. *Lancet.* ii:69.

Brand, J. L. 1965. *Doctors and the state. The British medical profession and government action in public health, 1870-1912.* Baltimore: Johns Hopkins University Press.

Brandt, A. M. 1986. *No magic bullet: a social history of venereal disease in the United States since 1880.* New York: Oxford University Press.

Brent, J. 1993. *Charles Sanders Peirce: a life.* Bloomington: Indiana University Press.

Brewer, L., and E. Bogen. 1947. Streptomycin in tuberculous tracheobronchitis. *American Review of Tuberculosis* 56:408-14.

Brieger, G. 1967. Therapeutic conflict and the American medical profession in the 1860s. *Bulletin of the History of Medicine* 41:215-23.

Brieger, G. H. 1981. The history of medicine and the history of science. *Isis* 72:537-40.

Broome, J. 1984. Selecting people randomly. *Ethics* 95:38-55.

Brough, M. C. 1949. Thirty-five years at Midhurst: survey of progress in treatment of pulmonary tuberculosis based on Midhurst Annual Reports for 1906-1941. *Tubercle* 30:103-9.

Brown, B. W. 1980. Statistical controversies in the design of clinical trials. *Controlled Clinical Trials* 1:13-27.

Brown, P. S. 1982. Herbalists and medical botanists in mid-nineteenth century Britain with special reference to Bristol. *Medical History* 26:405-20.

---. 1985. The venders of medicines advertised in eighteenth century Bath newspapers. *Medical History*:352-69.

---. 1986. Medicines advertised in eighteenth century Bath newspapers. *Medical History*:152-68.

---. 1987. Social context and medical theory in the demarcation of nineteenth-century boundaries. In *Medical fringe & medical orthodoxy*, ed. W. F. Bynum and R. Porter, 216-33. London: Croom Helm.

---. 1991. Medically qualified naturopaths and the General Medical Council. *Medical History* 35:50-77.

Brownlee, J. 1916. Certain considerations concerning the epidemiology of phthisis pulmonalis. *Public Health* 29:130-45.

Bryder, L. 1984. Papworth village settlement - a unique experiment in the treatment and care of the tuberculous? *Medical History* 28:372-90.

---. 1985. Tuberculosis and the Medical Research Council. *Society for the Social History of Medicine Bulletin* 37:72-75.

---. 1988. *Below the magic mountain: a social history of tuberculosis in twentieth century Britain.* Oxford: OUP.

---. 1989. Tuberculosis and the MRC. In Austoker, J., and L Bryder. Eds. *Historical perspectives on the role of the MRC*. Oxford: Oxford University Press. p1-21.

Buck, P. 1977. Seventeenth century political arithmetic: civil strife and vital statistics. *Isis* 68:67-84.

---. 1982. People who counted: political arithmetic in the eighteenth century. *Isis* 73:28-45.

Bud, R. 1998 . Penicillin and the new Elizabethans. *British Journal of the History of Science* 31:305-33.

Bull, J. P. 1959. The historical development of clinical therapeutic trials. *Journal of Chronic Diseases* 10:218-46.

Bullowa, J. G. M. 1928. The control. *Bulletin of the New York Academy of Medicine* 4:339-43.

---. 1929**a**. The serum treatment and its evaluation in lobar pneumonia. *Bulletin of the New York Academy of Medicine* 5:328-61.

---. 1929**b**. Serum treatment of the pneumococcus pneumonias. *BMJ* i:943-46.

Bulpitt, C. J. 1983. *Developments in biostatistics and epidemiology Volume 1. Randomised controlled clinical trials.* Amsterdam: Martinus Nijhoff.

Burks, A. W., ed. 1958. *Collected papers of Charles Sanders Peirce.* Belknap ed. Vol. Volume VII. Cambridge, Mass.: Harvard University Press.

Burn, J. H. 1955. Sir Henry Dale's contribution to therapeutics. *BMJ* i:1357-59.

Burnham, J. C. 1996. How the concept of profession evolved in the work of historians of medicine. *Bulletin of the History of Medicine* 70:1-24.

Burton, P., L. C. Gurrin, and M. J. Campbell. 1998. Clinical significance not statistical significance: a simple Bayesian alternative to p values. *Journal of Epidemiology and Community Health.* 52:318-23.

Buss, A. R. 1976. Galton and the birth of differential psychology and eugenics: social, political and economic forces. *Journal of the History of the Behavioural Sciences* 12:47-58.

Buxton, JB, and A. MacNalty. 1928. *The intradermal tuberculin test in cattle: collected results of experience.* MRC Special Report Series, 122. London: HMSO.

Bynum, W. F. 1970. Chemical structures and pharmacological action: a chapter in the history of nineteenth century molecular pharmacology. *Bulletin of the History of Medicine* 44:518-38.

---. 1994. *Science and the practice of medicine in the nineteenth century*. Cambridge: Cambridge University Press.

Bynum, W. F., and J. Wilson. 1991. Periodical knowledge: medical journals and their editors in nineteenth-century Britain. In *Medical journals and medical knowledge: historical essays*, ed. W. F. Bynum, S. Lock, and R. Porter, 6-28. London: RKP.

Camac Wilkinson, W. 1908, The treatment of consumption. London: Macmillan.

---.1923. Therapeutic use of tuberculin. *BMJ* i:674-76.

---. 1926**a**. The therapeutic use of tuberculin. *BMJ* i:1036-39.

---. 1926**b**. The tuberculin dispensary. *BMJ* i:674-76.

---. 1926**c**. Tuberculin treatment of tuberculosis. *BMJ* ii:274-75.

Campbell, R. 1992. *Truth and historicity.* Oxford: Clarendon Press.

Campbell, W. A. 1978. Some early chemical analyses of proprietary medicines. *Isis* 69:226-33.

Canada, R. O. 1947. Streptomycin therapy in progressive pulmonary tuberculosis. *American Review of Tuberculosis* 56:508-33.

Canguilhem, G. 1978. *On the normal and the pathological.* Dordrecht: Reidel.

---. 1981. What is scientific ideology? *Radical Philosophy* 29:20-25.

Cantor, D. 1992. Cortisone and the politics of drama, 1949-1955. Chap. 2 In *Medical innovations in historical perspective*. Science Technology and Medicine in Modern History, ed. J. V. Pickstone, ed. J. V. Pickstone, 165-84. Basingstoke: Macmillan.

Cantor, G. 1989. The rhetoric of experiment. In *The uses of experiment*, ed. D. Gooding, T. Pinch, and S. Schaffer, 159-80. Cambridge: CUP.

Carlson, D. G. 1977. African fever, prophylactic quinine, and statistical analysis: factors in the European penetration of a hostile West African environment. *Bulletin of the History of Medicine* 51:386-96.

Carnap, R. 1962. *Logical foundations of probability.* 2nd ed. Chicago: University of Chicago Press.

Carpenter, K. J. 1986. *The history of scurvy and vitamin C.* Cambridge: Cambridge University Press.

Carswell, R. 1928. Collective investigation and tuberculosis. *BMJ* i:375.

Cassedy, J. H. 1984. *American medicine and statistical thinking, 1800 - 1860.* Cambridge, Mass.: Harvard University Press.

Cassidy, W. A., and E. Dunner. 1947. Pulmonary tuberculosis treated with streptomycin. *American Review of Tuberculosis* 56:552-64.

Cecil, R. L. 1939. Present status of serum therapy in pneumonia. *Bulletin of the New York Academy of Medicine* 74:341-54.

Chalmers, A. 1982. *What is this thing called science?* 2nd revised ed. Milton Keynes: Open University Press.

---. 1990. *Science and its fabrication*. Milton Keynes: Open University Press.

Chalmers, I. 1983. Scientific inquiry and authoritarianism in perinatal care and education. *Birth* 10:151-64.

---. 1997. Assembling comparison groups to assess the effects of health care. *Journal of the Royal Society of Medicine* 90:379-86.

Chalmers, T. C. 1981. The clinical trial. *Milbank Memorial Fund Quarterly* 59:324-29.

---. 1983. The control of bias in clinical trials. In *Clinical trials: issues and approaches*, ed. S. H. Shapiro and T. A. Louis. New York: Marcel Drekker.

Chalmers, T. C., P. Celano, H. S. Sacks, and H. Smith. 1983. Bias in treatment assignment in controlled clinical trials. *New England Journal of Medicine* 309:1358-61.

Chance, B. 1942. Bishop Berkeley and his use of tar-water. *Annals of Medical History* 4:453-67.

Charlton, J., Murphy, M. 1997. *The health of adult Britain 1841-1994*. Volume 1. London: the Stationary Office. p220-221.

Chen, W. 1992. The laboratory as business: Sir Almroth Wright's vaccination programme and the construction of penicillin. In *The laboratory revolution in medicine*, ed. A. Cunningham and P. Williams, 245-92. Cambridge: Cambridge University Press.

Church, A. 1940. On the concept of a random sequence. *Bulletin of the American Mathematical Society* XLVI:130-35.

Clark, H. 1915. *The dispensary treatment of tuberculosis.* London: Ballière, Tindall and Cox.

Cochrane, A. L. 1972. *Effectiveness and efficiency: random reflections on health services*. London: Nuffield Provincial Hospitals Trust.

Cockburn, W. C. 1991. The international contribution to the standardization of biological substances. I. Biological standards and the League of Nations 1921-1946. *Biologicals* 19:161-69.

Codell Carter, K. 1985. Koch's postulates in relation to the work of Jacob Henle and Edwin Klebs. *Medical History* 29:353-74.

Cole, J. 1994. The chaos of particular facts: statistics, medicine and the social body in early nineteenth century France. *History of the Human Sciences* 7:1-27.

Cole, R. 1913 Treatment of pneumonia by means of specific serums *Journal of the American Medical Association* 61:663-6.

Colebrook, L., and M. Kenny. 1936. Treatment with prontosil of puerperal infection. *Lancet.* ii:1319-26.

---. 1936. Treatment of human puerperal infections. *Lancet.* i:1279-86.

Colebrook, L., and A. W. Purdie. 1937. Treatment of 106 cases of puerperal fever by sulphonilamide. *Lancet.* ii:1237-42 and 1291-94.

Coleman, W. 1982. *Death is a social disease. Public health and political economy in early Industrial France*. Madison, Wisconsin: University of Wisconsin Press.

---. 1985. The cognitive basis of the discipline: Claude Bernard on physiology. *Isis* 76:49-70.

---. 1989. Experimental physiology and statistical inference: the therapeutic trials of nineteenth century Germany. In Kruger, L., G Gigerenzer, and M. S. Morgan eds. *The probabilistic revolution. Volume 2*, Cambridge, Mass.: MIT Press. 201-225

Coleman, W., and F.L. Holmes. 1988. *The investigative enterprise: experimental physiology in nineteenth century medicine*. Berkeley: University of California Press.

Committee on Therapy, American Trudeau Society. 1947. Report. *American Review of Tuberculosis* 56:477-81.

Comroe, J. H. 1978. Pay dirt, part 1. *American Review of Respiratory Diseases* 117:773-81.

---. 1978. Pay dirt, part II. *American Review of Respiratory Diseases* 117:957-968.

Comstock, G. W. 1972. In memoriam: Carroll Edwards Palmer, 1903-1972. *American Journal of Epidemiology* 95:305-7.

Cook, H. J. 1988 *The decline of the old medical regime in Stuart London*. Ithaca: Cornell University Press.

---. 1990. Practical medicine and the British armed forces after the glorious revolution. *Medical History* 34:1-26.

---. 1994. Good advice and little medicine: the professional authority of early modern English physicians. *Journal of British Studies* 33:1-31.

Cooter, R. 1982. Anticontagionism and history's medical record. In Wright, P. and A Treacher eds. *The problem of medical knowledge*. Edinburgh: Edinburgh University Press.

---. 1984. *The cultural meaning of popular science: phrenology and the organisation of consent in nineteenth century Britain.* Cambridge: Cambridge University Press.

---. 1984. Open-air therapy and the rise of open-air hospitals. *Society for the Social History of Medicine Bulletin* 35:44-46.

---, ed. 1988. *Studies in the history of alternative medicine.* London: Macmillan.

---. 1993. *Surgery and society in peace and war: orthopaedics and the organisation of modern medicine 1880-1948.* Basingstoke: Macmillan.

Cooter, R., and S. Pumfrey. 1994. Separate spheres and public places: reflections on the history of science popularization and science in popular culture. *History of Science* xxxii:237-67.

Copeman, W. S. C. 1967. *The Worshipful Society of Apothecaries of London 1617-1967.* London: Pergamon.

Cornfield, J. 1976. Recent methodological contributions to clinical trials. *American Journal of Epidemiology* 104:408-21.

Corry Mann, H. C. 1926. *Diets for boys during the school age.* MRC Special Report Series No 105. London: HMSO.

Cowan, J. M. et al. 1932. Treatment of lobar pneumonia by Felton's serum. *Lancet* ii:8-11.

Cowan, R. S. 1972. Francis Galton's statistical ideas: the influence of eugenics. *Isis* 63:509-28.

Cowen, D. L. 1957. The Edinburgh pharmacopoeia 1. Historical development and significance. *Medical History* 1:123-39.

---. 1957. The Edinburgh pharmacopoeia 2. Bibliography. *Medical History* 1:340-51.

---. 1982. The influence of the Edinburgh Pharmacopoeia and the Edinburgh dispensatories. *Pharm Hist (GB)* 12:2-4.

Cowles, M., and C. Davis. 1982. On the origin of the .05 level of statistical significance. *American Psychologist* 37:553-58.

Cranberg, L. 1979. Do retrospective controls make clinical trials inherently fallacious? *BMJ*:1265-66.

Crawford, C. 1991. A scientific profession: medical reform and forensic medicine in British periodicals in the early nineteenth century. In *British medicine in an age of reform*, ed. R. French and A. Wear, 203-30. London: Routledge.

Crellin, J. K. 1981. Internal antisepsis or the dawn of chemotherapy? *Journal of the History of Medicine* 36:9-18.

Crombie, A. C., ed. 1963. *Scientific change.* London: Heinemann.

Crosby, A. W. 1997. *The measure of reality: quantification and western society, 1250-1600.* Cambridge: Cambridge University Press.

Cullen, M. J. 1975. *The statistical movement in early Victorian Britain: the foundations of empirical social research.* New York: Barnes and Noble.

Dale, A. I. 1991. *A history of inverse problems.* Berlin: Springer-Verlag.

Daniel, T. M. 1997. *Captain of death: the story of tuberculosis.* Rochester: University of Rochester Press.

Daniels, M. 1952. Tuberculosis. In *The British Encyclopaedia of Medical Practice*. 2nd ed., ed. Rt. Hon Lord Horder, 344-57. Vol. 12. London: Butterworth & Co. p344-357.

Danziger, K. 1990. *Constructing the subject: historical origins of psychological research.* Cambridge: Cambridge University Press.

D'Arcy Hart, P. 1946. Chemotherapy of tuberculosis: research during the past 100 years. *BMJ* ii:805-10, and 849-55.

---. 1954. The role of the host in the chemotherapy of tuberculosis. *BMJ* ii:767-74.

---. 1960. The organization of controlled clinical trials. In *Controlled clinical trials. Papers delivered at the conference convened by The Council for International Organizations of Medical Sciences*, ed. A. Bradford Hill, 145-50. Oxford: Blackwell Scientific.

---. 1991. Randomized controlled clinical trials [letter]. *BMJ* 302:1271-72.

Daston, L. J. 1980. Probabilistic expectation and rationality in classical probability theory. *Historia Mathematica* 7:234-60.

---. 1981. Mathematics and the moral sciences: the rise and fall of the probability of judgements. In *Epistemological and social problems of the sciences in the early nineteenth century*, ed. H. N. Jahnke and M. Otte. Dordrecht: Reidel.

---. 1987. The domestication of risk: mathematical probability and insurance 1650-1830. In *The probabilistic revolution. Volume 1: ideas in history*, ed. L. Krüger, L. J. Daston, and M. Heidelberger, 237-60. Cambridge: MIT Press.

---. 1987. Rational individuals versus laws of society: from probability to statistics. In *The probabilistic revolution. Volume 1: ideas in history*, ed. L. Krüger, L. J. Daston, and M. Heidelberger, 295-304. Cambridge: MIT Press.

---. 1992. Objectivity and the escape from perspective. *Social Studies of Science* 22:597-618.

---. 1995. The moral economy of science. *Osiris 2nd Series* 10:3-24.

Daston, L. J. and P. Galison. 1992. The image of objectivity. *Representations* 40:81-128.

Davenport-Hines, R. T. P. 1986. Glaxo as a multinational before 1963. In *British multinationals: origins, management and performance*, ed. G. Jones, 137-63. Aldershot: Gower.

Davenport-Hines, R. T. P., and J. Slinn. 1992. *Glaxo: a history to 1962.* Cambridge: Cambridge University Press.

David, F. N. 1962. *Games, gods and gambling. The origins and history of probability and statistical ideas from the earliest times to the Newtonian era.* New York: Hafner.

Davies, C. 1988. The health visitor as mother's friend: a woman's place in public health, 1900-1914. *Social History of Medicine* 1:39-59.

Dean, M. 1994. *Critical and effective histories. Foucault's methods and historical sociology.* London: Routledge.

Dehue, T. 1998. Deception, efficiency, and random groups. *Isis*:653-73.

Dempster, A. P. 1966. Review of Logic of Statistical Inference. *Journal of the American Statistical Association* 61:1233-35.

D'Esopo, N. D. 1982. Clinical trials in pulmonary tuberculosis. *American Review of Respiratory Diseases* 125 (Suppl. 3 pt 2):85-93.

D'Esopo, N. D., and J. Steinhaus. 1947. Streptomycin therapy with special reference to tuberculosis. *American Review of Tuberculosis* 56:589-605.

Desrosieres, A. 1989. The part in relation to the whole: how to generalise? The prehistory of representative sampling. In *The social survey in historical perspective*, ed. M. Bulmer. Cambridge: Cambridge University Press.

Digby, A. 1994. *Making a medical living: doctors and patients in the English market for medicine, 1720-1911*. Cambridge: Cambridge University Press.

Doll, R. 1982. Clinical trials: retrospect and prospect. *Statistics in Medicine* 1:337-44.

---. 1992. Sir Austin Bradford Hill and the progress of medical science. *British Medical Journal* 305:1521-26.

---. 1993. Sir Austin Bradford Hill, 1897-1991. *Statistics in Medicine* 12:795-808.

---. 1994. Austin Bradford Hill, 8 July 1897 - 18 April 1991. *Biographical Memoirs of the Royal Society*:129-40.

---. 1995. Sir Austin Bradford Hill: a personal view of his contribution to epidemiology. *Journal of the Royal Statistical Society A* 158:155-63.

Dowling, H. 1972. Frustration and foundation: the management of pneumonia before antibiotics. *JAMA* 220:1341-45.

---. 1973. Comparisons and contrasts between the early arsphenamine and early antibiotic periods. *Bulletin of the Institute of the History of Medicine* 47:236-49.

---. 1977. *Fighting infection: conquests of the Twentieth Century*. Cambridge, Mass: Harvard University Press.

Dwork, D. 1987. *War is good for babies and other young children: a history of the infant and child welfare movement in England 1898-1919*. London: Tavistock.

Dykstra, D. L. 1955. The medical profession and proprietary medicines during the nineteenth century. *Bulletin of the History of Medicine* 29:401-19.

Ellison, D. L. 1994. *Healing tuberculosis in the woods: medicine and science at the end of the nineteenth century*. New York: Greenwood Publishing.

Ellman, P., J. H. P. Johnson, and A. Kagan. 1955. The changing place of artificial pneumothorax in the treatment of pulmonary tuberculosis. *British Journal of Tuberculosis and Diseases of the Chest* L:130-51.

Emerson, J. D., E. Burdick, D. C. Hoaglin, F. Mosteller, and T. C. Chalmers. 1990. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. *Controlled Clinical Trials* 11:339-52.

Epstein, S. 1991. Democratic science? AIDS activism and the contested construction of knowledge. *Socialist Review* 21:35-64.

---. 1996 *Impure science: AIDS, activism, and the politics of knowledge*. Berkeley: University of California Press.

---. 1997. Activism, drug regulation, and the politics of therapeutic evaluation in the AIDS era: a case study of ddC and the 'surrogate marker' debate. *Social Studies of Science* 27:691-726.

Epstein, S., and B. Williams. 1946. *Miracles from microbes: the road to streptomycin.* New Brunswick: Rutgers University Press.

Evans, D. 1992. Tackling the "hideous scourge": the creation of venereal disease centres in early twentieth century Britain. *Social History of Medicine* 5:413-33.

Evans Medical Supplies Ltd. 1959. *The story of Evans Medical Supplies, 1809-1959.* Liverpool: Evans Medical Supplies.

Eyler, J. M. 1979. *Victorian social medicine: the ideas and methods of William Farr.* Baltimore: Johns Hopkins University Press.

---. 1980. The conceptual origins of William Farr's epidemiology: numerical methods and social thought in the 1830s. In *Times, places and persons: aspects of the history of epidemiology*, ed. A. M. Lilienfeld, 1-22. Baltimore: Johns Hopkins Press.

Farr, W. 1862. A method of determining the effects of systems of treatment in certain diseases. *BMJ* ii:193-195.

Feinstein, A. R. 1970. Clinical Biostatistics II. Statistics versus science in the design of experiments. *Clinical Pharmacology and Therapeutics* 11:282.

---. 1977. The Haze of Bayes, the Aerial palaces of decision analysis, and the computerized ouija board. *Clinical Pharmacology and Therapeutics* 21:482-96.

---. 1983. An additional basic science for clinical medicine: 1. The constraining fundamental paradigms. *Annals of Internal Medicine* 99:393-97.

---. 1983. An additional basic science for clinical medicine: the limitations of randomized trials. *Annals of Internal Medicine* 99:544-50.

---. 1987. The intellectual crisis in clinical science. *Perspectives in Biology and Medicine* 30:215-30.

---. 1995. The crisis in clinical research. *Bulletin of the History of Medicine* 69:288-91.

---. 1995. Meta-analysis: statistical alchemy for the 21st century. *Journal of Clinical Epidemiology* 48:71-79.

Feldberg, G. D. *Disease and class: tuberculosis and the shaping of modern North American society*. New Jersey: Rutgers University Press.

Feldman, W. H. 1946. The chemotherapy of tuberculosis, including the use of streptomycin [Harben Lectures 1946]. *Journal of the Royal Institute of Public Health* 9:267-88,297-324,343-63.

---. 1954. Streptomycin: some historical aspects of its development as a chemotherapeutic agent in tuberculosis. *American Review of Tuberculosis* 69:859-68.

Feldman, W. H., and H. C. Hinshaw. 1944. Effects of streptomycin on experimental tuberculosis in guinea pigs. A preliminary report. *Proceedings of Staff Meetings of the Mayo Clinic* 19:593-99.

Feldman, W. H., A. G. Karlson, and H. C. Hinshaw. 1947. Streptomycin in experimental tuberculosis. *American Review of Tuberculosis* 56:346-59.

Ferebee, S. H. 1970. Controlled chemophrophylaxis trials in tuberculosis: a general review. *Advances in Tuberculosis Research* 17:28-106.

Fibiger, J. 1898. Om serumbehandling af difteri [on the treatment of diphtheria with serum]. *Hospitalstidende* 6:309-25. (Summary consulted in Cochrane Library)

Figlio, K. 1976. The metaphor of organization: an historiographical perspective on the bio-medical sciences of the early 19th century. *History of Science* 14:17-53.

---. 1979. Sinister medicine? A critique of left approaches to medicine. *Radical Science Journal* 9:14-68.

---. 1982. How does illness mediate social relations? Workmen's compensation and medico-legal practice 1890-1940. In Wright, P., and A. Treacher, eds., *The problem of medical knowledge*. Edinburgh: Edinburgh University Press.

---. 1986. The historiography of scientific medicine: an invitation to the human sciences. :262-86.

Fildes, P., and R. J. G. Parnall. 1919. *An investigation into the ultimate results of the treatment of syphilis with arsenical compounds.* MRC Special Report Series 41. London: HMSO.

Findlen, P. 1993. Controlling the experiment: rhetoric, court patronage and the experimental method of Francesco Redi. *History of Science* 31:35-64.

Fine, P. E. M. 1979. John Brownlee and the measurement of infectiousness. *Journal of the Royal Statistical Society A* 142:347-62.

Finland, M. 1930. The serum treatment of lobar pneumonia. *New England Journal of Medicine* 202:1244-47.

Finland, M. et al 1946. Development of streptomycin resistance during treatment. *Journal of the American Medical Association* 132:16-20.

Finlay, M. 1992. Quackery and cookery: Justus von Liebig's extract of meat and the theory of nutrition in the Victorian age. *Bulletin of the History of Medicine* 66:404-18.

Fisher, D. 1978**a**. The Rockefeller Foundation and the development of scientific medicine in Britain. *Minerva* 16:20-41.

---. 1978**b**. Rockefeller philanthropy and the British empire: the creation of the London School of Hygiene and Tropical Medicine. *History of Education* 7:129-43.

Fisher, M. W., G. W. Fishburn, and J. B. Wallace. 1947. Streptomycin and bed rest in the treatment of pulmonary tuberculosis. *American Review of Tuberculosis* 56:534-39.

Fisher, R. A. 1922. On the mathematical foundation of theoretical statistics. *Philosophical Transactions of the Royal Society A* 221:309-68.

---. 1925. *Statistical methods for research workers.* Oliver and Boyd: Edinburgh.

---. 1926. The arrangement of field experiments. *Journal of the Ministry of Agriculture* 33:503-13.

---. 1935**a**. *The design of experiments.* London: Oliver and Boyd.

---. 1935**b.** Contribution to a discussion of J Neyman's paper on statistical problems in agricultural experimentation. *Journal of the Royal Statistical Society, Supplement*, 2:173

---. 1935**c**. The logic of inductive inference. *Journal of the Royal Statistical Society* 98:39-54.

---. 1954. Expansion of statistics. *American Scientist* 42:275-82.

---. 1955. Statistical methods and scientific induction. *Journal of the Royal Statistical Society B* 17:69-78.

---. 1958. The nature of probability. *Centennial Review* 2:261-174.

---. 1959. *Smoking; the cancer controversy: some attempts to assess the evidence.* Edinburgh: Oliver and Boyd.

---. 1972-74. On the "probable error" of a coefficient of correlation deduced from a small sample. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 205-35. Vol. 1. Melbourne: University of Melbourne Press.

---. 1972-74. Studies in crop variation 1. An examination of the yield of dressed grain from Broadbalk. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 236-65. Vol. 1. Melbourne: University of Melbourne Press.

---. 1972-74. On the interpretation of chi-squared from contingency tables, and the calculation of p. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 336-45. Vol. 1. Melbourne: University of Melbourne Press.

---. 1972-74. Statistical tests of agreement between observation and hypothesis. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 460-68. Vol. 1. Melbourne: University of Melbourne Press.

---. 1972-74. The conditions under which chi-squared measures the discrepancy between observation and hypothesis. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 480-88. Vol. 1. Melbourne: University of Melbourne Press.

---. 1972-74. A method for scoring coincidences in tests with playing cards. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 557-61. Vol. 1. Melbourne: University of Melbourne Press.

---. 1972-74. The elimination of mental defect. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 580-82. Vol. 1. Melbourne: University of Melbourne Press.

---. 1972-74. The biometrical study of heredity. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 583-604. Vol. 1. Melbourne: University of Melbourne Press.

---. 1972-74. The arrangement of field experiments. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 83-94. Vol. 2. Melbourne: University of Melbourne Press.

---. 1972-74. Eugenics: can it solve the problem of decay of civilisations? In *Collected papers of RA Fisher*, ed. J. H. Bennett, 113-21. Vol. 2. Melbourne: University of Melbourne Press.

---. 1972-74. Modern eugenics. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 122-27. Vol. 2. Melbourne: University of Melbourne Press.

---. 1972-74. Periodical health surveys. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 128-31. Vol. 2. Melbourne: University of Melbourne Press.

---. 1972-74. Studies in crop variation. VI Experiments on the response of the potato to Potash and nitrogen. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 375-87. Vol. 2. Melbourne: University of Melbourne Press.

---. 1972-74. The statistical method in psychical research. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 388-91. Vol. 2. Melbourne: University of Melbourne Press.

---. 1972-74. Statistics and biological research [letter]. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 392-93. Vol. 2. Melbourne: University of Melbourne Press.

---. 1972-74. The arrangement of field experiments and the statistical reduction of the results. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 437-55. Vol. 2. Melbourne: University of Melbourne Press.

---. 1972-74. Randomisation, and an old enigma of card play. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 162-65. Vol. 3. Melbourne: University of Melbourne Press.

---. 1972-74. The logic of inductive inference. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 272-315. Vol. 3. Melbourne: University of Melbourne Press.

---. 1972-74. Eugenics, academic and practical. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 444-49. Vol. 3. Melbourne: University of Melbourne Press.

---. 1972-74. The half-drill strip system agricultural experiments [letter]. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 482-83. Vol. 3. Melbourne: University of Melbourne Press.

---. 1972-74. Has Mendel's work been rediscovered? In *Collected papers of RA Fisher*, ed. J. H. Bennett, 514-36. Vol. 3. Melbourne: University of Melbourne Press.

---. 1972-74. Professor Karl Pearson and the method of moments. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 29-45. Vol. 4. Melbourne: University of Melbourne Press.

---. 1972-74. "Student". In *Collected papers of RA Fisher*, ed. J. H. Bennett, 218-26. Vol. 4. Melbourne: University of Melbourne Press.

---. 1972-74. Mathematical probability in the natural sciences. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 398-406. Vol. 5. Melbourne: University of Melbourne Press.

---. 1972-74. Lung cancer and cigarettes? In *Collected papers of RA Fisher*, ed. J. H. Bennett, 423-24. Vol. 5. Melbourne: University of Melbourne Press.

---. 1972-74. Cancer and smoking. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 425-31. Vol. 5. Melbourne: University of Melbourne Press.

Fisher, R. A., and W. A. Mackenzie. 1972-74. Studies in crop variation. II. The manurial response of different potato varieties. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 469-78. Vol. 1. Melbourne: University of Melbourne Press.

Fissell, M. E. 1991. The disappearance of the patient's narrative and the invention of hospital medicine. In *British medicine in an age of reform*, ed. R. French and A. Wear, 92-109. London: Routledge.

---. 1991. *Patients, power and the poor in eighteenth century Bristol.* Cambridge: Cambridge University Press.

Fitz-Gibbon, C. T. 1986 Apr Vol 21(1). In defence of randomised controlled trials, with suggestions about the possible use of meta-analysis. Special Issue: Adult aphasia. *British Journal of Disorders of Communication*:117-24.

Fleck, L. 1979. *The genesis and development of a scientific fact.* Chicago: University of Chicago Press.

Fleiss, J. L. 1982. Multicentre clinical trials: Bradford Hill's contributions and some subsequent developments. *Statistics in Medicine* 1:361-68.

Fleming, A. 1946. *Chemotherapy: yesterday, today, and tomorrow.* Cambridge: Cambridge University Press.

Folks, J. L. 1981. *Ideas of statistics.* New York: John Wiley.

---. 1984. Use of randomization in experimental research. In *Experimental design, statistical models, and genetic statistics*, ed. K. Hinkelmann, 17-32. New York: Marcel Drekker.

Foucault, M. 1970. *The order of things.* London: Tavistock.

---. 1972. *The archaeology of knowledge.* New York: Pantheon Books.

---. 1973. *The birth of the clinic.* London: Tavistock.

---. 1981. *The history of sexuality. Volume 1: an introduction*. London: Pelican Books.

---. 1991. Introduction. In *The Normal and the pathological*, G. Canguilhem, 7-24. New York: Zone Books.

--- 1997. *Ethics: subjectivity and truth*. New York: The New Press (The essential works of Michel Foucault 1954-1984, Volume 1)

Foust, C. M. 1992. *Rhubarb: the wonder drug.* Princeton: Princeton University Press.

Fox, N. J. 1988. Scientific theory choice and social structure: the case of Joseph Lister. *History of Science* 26:367-97.

Freedlander, B. L., and F. A. French. 1947. Constitution of chemotherapeutic agents in experimental tuberculosis. *American Review of Tuberculosis* 56:373-75.

Freidson, E. 1970. *Profession of medicine: a study of the sociology of applied knowledge.* New York: Dodd Mead.

Freiman, J. A., T. C. Chalmers, H. Smith, and R. R. Kuebler. 1978. The importance of beta, the Type II error and sample size in the design and interpretation of the randomized controlled trial: survey of 71 "negative" trials. *New England Journal of Medicine* 299:690-94.

French, R. D. 1975. *Antivivisection and medical science in Victorian society.* Princeton, NJ: Princeton University Press.

Fujimara, J. H. 1987. Constructing 'do-able' problems in cancer research: articulating alignment. *Social Studies of Science* 17:257-93.

Fussell, G. E. 1935. The technique of early field experiments. *Journal of the Royal Agricultural Society of England* 96:78-88.

Fye, W. B. 1990. Vasodilator therapy for angina pectoris: the intersection of homeopathy and scientific medicine. *Journal of the History of Medicine and Allied Sciences* 45:317-40.

Gage, T. B. 1993. The decline of mortality in England and Wales 1861 to 1964: decomposition by cause of death and component of mortality. *Population Studies* 47:47-66.

Galdston, I. 1924. Tuberculosis in folk medicine. *Medical Life* 31:393-96.

---. 1939. The ideological basis of discovery. *Bulletin of the History of Medicine* 7:729-35.

---. 1940. Some notes on the early history of chemotherapy. *Bulletin of the History of Medicine* 8:806-18.

---. 1943. *Behind the sulfa drugs: a short history of chemotherapy.* New York: D. Appleton-Century Company.

Galton, F. 1869. *Hereditary genius.* London: Macmillan.

---. 1872. Statistical inquiries into the efficacy of prayer. *Fortnightly Review* ns:125-35.

---. 1875. Statistics by intercomparison with remarks on the law of frequency of error. *Philosophical Magazine* 49:33-46.

Garber, D., and S. Zabell. 1979. On the emergence of probability. *Archives of the History of the Exact Sciences* 21:33-53.

Garrison, F. H. 1933. On quackery as a reversion to primitive medicine. *Bulletin of the New York Academy of Medicine* 9:601-12.

Gehan, E. A. and N. A. Lemack. 1995 *Statistics in medical research: developments in clinical trials*. New York: Plenum Publishing Corporation.

Geison, G. L. 1978. *Michael Foster and the Cambridge School of Physiology. The scientific enterprise in late Victorian society.* Princeton: Princeton University Press.

---. 1979. Divided we stand: physiologists and clinicians in the American context. In *The therapeutic revolution: essays in the social history of American medicine*, ed. M. J. Vogel and C. E. Rosenberg, 67-90. Philadelphia: University of Pennsylvania Press.

---. 1990. Pasteur, Roux and rabies: scientific versus clinical mentalities. *Journal of the History of Medicine* 45:341-65.

Gelfand, T. 1970. Empiricism and eighteenth-century French surgery. *Bulletin of the History of Medicine* 44:40-53.

---. 1977. A clinical ideal: Paris 1789. *Bulletin of the History of Medicine* 51:397-411.

---. 1981. The decline of the ordinary practitioner and the rise of the modern medical profession. In *Doctors, patients, and society*, ed. M. S. Staum and D. E. Larsen, 105-29. Hamilton, Ontario: Wilfred Laurier University Press.

Giere, R. N. 1979. Foundations of probability and statistical inference. In *Current research in philosophy of science. Proceedings of the PSA critical research problems conference*, ed. P. D. Asquith and H. Kyberg. East Lancing: Philosophy of Science Association.

Gigerenzer, G. 1987. Probabilistic thinking and the fight against subjectivity. In *The probabilistic revolution. Volume 2: ideas in the sciences*, ed. G. Kruger, G. Gigerenzer, and M. S. Morgan, 11-34. Cambridge Mass.: MIT Press.

---. 1989. *The empire of chance: how probability changed science and everyday life.* Cambridge: Cambridge University Press.

Gigerenzer, G., and D. J. Murray. 1987. *Cognition as intuitive statistics.* Hillsdale, NJ: Lawrence Erlbaum.

Gilbert, N., and M. Mulkay. 1984. Experiments are the key: participants histories and historians histories of science. *Isis* 75:105-25.

Gilles, D. A. 1987. Was Bayes a Bayesian? *Historia Mathematica* 14:325-46.

Gillespie, C. 1960. *The edge of objectivity.* Oxford: Oxford University Press.

---. 1963. Intellectual factors in the background of analysis by probabilities. In *Scientific change*, ed. A. C. Crombie, 431-53. London: .

---. 1972. Probability and politics: Laplace, Condorcet, and Turgot. *Proceedings of the American Philosophical Society* 16:1-20.

Gillespie, J. R. 1926. Tuberculin in the treatment of tuberculosis. *BMJ* ii:35.

Gillespie, R. 1991. *Manufacturing knowledge: a history of the Hawthorne experiments.* Cambridge: CUP.

Gladstone, G. P., C. J. G. Bert, and G. W. Knight. 1973. Paul Gordon Fildes, 1882-1971. *Biographical Memoirs of the Fellows of the Royal Society* 19:317-47.

Goldberger, J., C. H. Waring, and W. F. Tanner. 1923. Pellagra prevention by diet among institutional inmates. *Public Health Reports* 38:2361-68.

Goldberger, J., C. H. Waring, and D. G. Willets. 1915. The prevention of pellagra: a test of diet among institutional inmates. *Public Health Reports* 30:3117-31.

Goldman, L. 1991. Statistics and the science of society in early Victorian Britain; an intellectual context for the General Register Office. *Social History of Medicine* 4:415-34.

Goldsmith, M. 1946. *The road to penicillin: a history of chemotherapy.* London: Lindsay, Drummond.

Golinski, J. 1990. The theory of practice and the practice of theory. *Isis* 81:492-505.

---. 1993. The rhetorical maelstrom [essay review]. *Isis* 84:746-49.

Good, I. J. 1974. Random thoughts about randomness. In *Proceedings of the 1972 biennial meeting of the Philosophy of Science Association*, ed. K. F. Schaffner and R. S. Cohen, 117-35. Dordrecht: Reidel.

Gooding, D., T. Pinch, and S. Schaffer, eds. 1990. *The uses of experiment.* Cambridge: Cambridge University Press.

Goodman, S. N. 1989. Meta-analysis and evidence. *Controlled Clinical Trials* 10:188-204.

---. 1993. P values, hypothesis tests and likelihood - implications for epidemiology of a neglected historical debate. *American Journal of Epidemiology* 137:485-96.

Goodman, S. N., and R. Royall. 1988. Evidence and scientific research. *American Journal of Public Health* 78:1568-74.

Gore, S. M. 1981. Assessing clinical trials - why randomize? *BMJ* 282:1558-1960.

Gossel, P. P. 1992. A need for standard methods: the case of American bacteriology. In *The right tools for the right job: at work in twentieth century life sciences*, ed. A. E. Clarke and Fujimara.JH, 287-311. Princeton: Princeton University Press.

Gossett, W. S. [‘Student’]. 1908 The probable error of a mean. *Biometrika* 6:1-8.

---. 1923. On testing varieties of cereals. *Biometrika* 15:271-93.

---. 1931. The Lanarkshire milk experiment. *Biometrika* 23:398-407.

---. 1937. Comparison between balanced and random arrangements in field plots. *Biometrika* 29:363-79.

Gottinger, H. W. 1974. Review of concepts and theories of probability. *Scientia* 109:83-110.

Gotzsche, P. C. 1989. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal anti-inflammatory drugs in rheumatoid arthritis. *Controlled Clinical Trials* 10:31-56.

Gould, S. J. 1981 *The mismeasure of man*. London: Penguin.

Green, D. P., and I. Shapiro. 1994. *Pathologies of rational choice theory: a critique of applications in political sciences.* Yale: Yale UP.

Green, F. H. K. 1944. Clinical evaluations of new remedies in Britain. *British Medical Bulletin* 2:58-60.

---. 1954. The clinical evaluations of remedies. *Lancet.* ii:1085-90.

Greenhow, E. H. 1860. *On diphtheria.* London: John W Parker.

Greenwood, M. 1913. On methods of research. *Lancet.* i:158-65.

---. 1932. *Epidemiology, historical and experimental [The Herter Lecture for 1931].* Baltimore MD: Johns Hopkins University Press.

---. 1935. *Epidemics and crowd diseases.* London: Williams and Northgate.

---. 1936. *The medical dictator and other biographical studies.* London: Williams and Norgate.

---. 1940. A statistical Mare's nest. *Journal of the Royal Statistical Society* 103:246-48.

---. 1943. *Authority in medicine: old and new [Linacre Lecture 6 May 1943].* Cambridge: Cambridge University Press.

---. 1944. Mr. Shaw on doctors. *BMJ* ii:570-71.

---. 1948. *Some British pioneers of social medicine.* Oxford: Oxford University Press.

---. 1948. *Medical statistics from Graunt to Farr.* Cambridge: Cambridge University Press.

Greenwood, M., A. Bradford Hill, W. W. C. Topley, and J. Wilson. 1936. *Experimental epidemiology.* MRC Special Report Series, 209. London: HMSO.

Greenwood, M., and G. U. Yule. 1915. The statistics of anti-typhoid and anti-cholera inoculation and the interpretation of such statistics in general. *Proceedings of the Royal Society of Medicine. Section on Epidemiology and State Medicine* 8:113-94.

Gross, A. G. 1990. *The rhetoric of science.* Cambridge, Mass.: Harvard University Press.

Gunn, J. A. 1932. Remarks on the outlook of research in therapeutics. *British Medical Journal* ii:389-92.

Gutting, G. *Michel Foucault's archaeology of scientific reason*. Cambridge: Cambridge University Press.

Guy, W. A. 1839. On the value of the numerical method as applied to science, but especially to physiology and medicine. *Journal of the Royal Statistical Society* 2:25-47.

Habermas, J. 1972. *Knowledge and human interests.* London: Heinemann.

Hacking, I. 1964. On the foundations of statistics. *British Journal of the Philosophy of Science* 15:1-27.

---. 1965. *Logic of statistical inference.* Cambridge: Cambridge University Press.

---. 1967. Possibility. *Philosophical Review* 76:143-68.

---. 1968. Possibility. *Australian Journal of Philosophy* 46:98-112.

---. 1971. Equipossibility theories of probability. *British Journal of the Philosophy of Science* 22:339-55.

---. 1971. The Leibniz-Carnap program for inductive logic. *Journal of Philosophy* 68:597-610.

---. 1972. The logic of Pascal's wager. *American Philosophical Quarterly* 9:186-92.

---. 1973. *Proof and eternal truths.* Oxford: Oxford University Press.

---. 1975. *The emergence of probability.* Cambridge: Cambridge University Press.

---. 1975. *Why does language matter to philosophy?* Cambridge: Cambridge University Press.

---. 1975. Possibility. *Philosophical Review* 83:249-56.

---. 1980. The theory of probable inference: Neyman, Peirce and Braithwaite. In D. H. Mellor ed. *Science, belief and behaviour: essays in honour of RB Braithwaite*. Cambridge: Cambridge University Press. p141-160.

---. 1981. How should we do the history of statistics? *Ideology & Consciousness* 8:15-26.

---. 1981. Karl Pearson's History of Statistics. *British Journal of the Philosophy of Science* 32:177-82.

---. 1982. Biopower and the avalanche of printed numbers. *Humanities in Society* 5:279-95.

---. 1983. *Representing and intervening.* Cambridge: Cambridge University Press.

---. 1983. Nineteenth century cracks in the concept of determinism. *Journal of the History of Ideas* 44:455-75.

---. 1984. Trial by number. *Science* (Nov):69-70.

---. 1986. Making up people. In *Reconstructing individualism*, ed. T. Heller. Stanford: Stanford University Press.

---. 1987. Was there a probabilistic revolution 1800-1930? In *The probabilistic revolution. Volume 1: ideas in history*, ed. L. Krüger, L. J. Daston, and M. Heidelberger, 45-55. Cambridge: MIT Press.

---. 1987. Prussian numbers, 1860-1882. In *The probabilistic revolution. Volume 1: ideas in history*, ed. L. Krüger, L. J. Daston, and M. Heidelberger, 378-94. Cambridge: MIT Press.

---. 1988. Telepathy: origins of randomisation in experimental design. *Isis* 79:427-51.

---. 1990. *The taming of chance.* Cambridge: Cambridge University Press.

---. 1992. Statistical language, statistical truth and statistical reason: the self-authentication of a style of scientific reasoning. In McMullin, E., ed. *The social dimensions of science.* Notre Dame: Baton Rouge. p130-157

Hald, A. 1990. *History of probability and statistics and their application before 1750.* New York: John Wiley.

Hall, D. L. 1974. Biology, sex hormones and sexism in the 1920s. *Philosophical Forum* 5:81-96.

---. 1976. The critic and the advocate: contrasting British views on the state of endocrinology in the early 1920s. *Journal of the History of Biology* 9:269-85.

Haller, J. S. 1984. Aconite: a case study in doctrinal conflict and the meaning of scientific medicine. *Bulletin of the New York Academy of Medicine* 60:888-904.

Hamlin, C. 1992. Predisposing causes and public health in early nineteenth century medical thought. *Social History of Medicine* 5:43-70.

Hansen, E. H., and L. Launso. 1989. Is the controlled clinical trial sufficient as a drug technology assessment? *Journal of Social and Administrative Pharmacy* 6:117-26.

Hardy, A. 1993. *The epidemic streets. Infectious disease and the rise of preventive medicine, 1856-1900*. Oxford: Clarendon Press.

Harrison, L. W. 1929. *The treatment of syphilis: a survey of records from St Thomas's Hospital*. London: HMSO. (MRC Special Reports Series 132)

Hartley, P. H. S., R. C. Wingfield, and V. A. Burrows. 1936. *The expectation of survival in pulmonary tuberculosis. An analysis of 8766 cases treated at the Brompton Hospital Sanatorium, Frimley.* Aldershot: Gale and Polden.

Harvey, A. M. 1981. *Science at the bedside: clinical research in American medicine 1905-1945.* Baltimore MD: Johns Hopkins University Press.

Harville, D. A. 1975. Experimental randomization: who needs it? *American Statistician* 29:27-31.

Heaf, F. R. G., and J. B. McDougall. 1945. *Rehabilitating the tuberculous.* London: Faber and Faber.

Hearnshaw, L. S. 1981. *Cyril Burt, psychologist.* New York: Vintage Books

Hellman, S. 1991. Of mice but not men. *New England Journal of Medicine* 324:1585-89.

Herman, J. 1994. Experiment and observation. *Lancet.* 344:1209-11.

---. 1995. The demise of the randomized controlled trial. *Journal of Clinical Epidemiology* 48:985-88.

Hessen, B. 1971. The social and economic roots of Newton's 'Principia'. In *Science at the cross roads*. 2nd ed., ed. N. I. Bukharin, 149-212. London: Frank Cass.

Hill, G. B. 1983. Controlled clinical trials - the emergence of a paradigm. *Clinical and Investigative Medicine* 6:25-32.

Hilts, V. L. 1978. Aliis Externundum, or, the origin of the Statistical Society of London. *Isis* 69:21-43.

---. 1980. *Statist and statistician: three studies in the history of 19th century English statistical thought.* New York: Arno.

Hinshaw, H. C. 1954. Historical note on the earliest use of streptomycin in clinical tuberculosis. *American Review of Tuberculosis* 70:9-14.

---. 1969. Tuberculosis chemotherapy: reminiscences of early trials. *Scandinavian Journal of Respiratory Diseases* 50:197-203.

Hinshaw, H. C., and W. H. Feldman. 1944. Evaluation of chemotherapeutic agents in clinical trials: a suggested procedure. *American Review of Tuberculosis* 50:202-13.

---. 1945. Streptomycin in treatment of clinical tuberculosis: a preliminary report. *Proceedings of Staff Meetings of the Mayo Clinic* 20:313.

---. 1946**a**. Streptomycin in treatment of clinical tuberculosis. *American Review of Tuberculosis* 54:191-4.

---. 1946**b**. Treatment of tuberculosis with streptomycin; a summary of observations on one hundred cases. *JAMA* 132:778-82.

Hodgkinson, R. G. 1968. Social medicine and the growth of statistical medicine. In *Medicine and science in the 1860s*, ed. F. N. L. Poynter, 183-98. London: Wellcome Institute of the History of Medicine.

Hogben, L. 1948. Book review: Principles of Medical Statistics. By A Bradford Hill. Fourth Edition. London: Lancet, 1948. *British Journal of Social Medicine* II:43.

---. 1950**a**. *Chance and choice by cardpack and chessboard.* London: Max Parrish.

---. 1950**b**. 'Major Greenwood, 1880-1949'. *Obituary Notices of the Fellows of the Royal Society*:138-54.

---. 1954. The assessment of remedies. *The Medical Press* iii:331-38.

---. 1957. *Statistical theory. The relationship of probability, credibility and error. An examination of the contemporary crisis in statistical theory from a behaviourist viewpoint.* New York: WW Norton & Company.

---. 1963. *Science in authority.* London: Allen and Unwin.

---. 1998. *Lancelot Hogben: scientific humanist. An unauthorised autobiography edited by Adrian and Anne Hogben.* Woodbridge: Merlin Press.

Hogben, L., and M. Sim. 1953. The self controlled and self recorded clinical trial for low grade morbidity. *British Journal of Preventive and Social Medicine* 7:163-69.

Hogben, L., and R. Wrighton. 1952a. Statistical theory of prophylactic and therapeutic trials. I: limitations of the unique null hypothesis. *British Journal of Preventive and Social Medicine* 6:89-117

---. 1952b. Statistical theory of prophylactic and therapeutic trials II Methods of operational advantage. *British Journal of Social Medicine* 6:205-25.

Hollman, A. 1997. *Sir Thomas Lewis: pioneer cardiologist and clinical scientist*. New York: Springer Verlag.

Holloway, S.W.F. 1966. The Apothecaries Act 1815: a re-interpretation. *Medical History* 10:107-29.

---. 1991. *Royal Pharmaceutical Society of Great Britain, 1841-1991: a political and social history.* London: The Pharmaceutical Press.

---. 1995. Producing experts, constructing expertise. The School of Pharmacy of the Pharmaceutical Society of Great Britain, 1842-1896. *Clio Medica* 30:116-40.

Honigsbaum, F. 1989. *Health, happiness, and security: the creation of the National Health Service.* London: Routledge.

Horrocks, S. M. 1995. The business of vitamins: nutrition science and the food industry in inter-war Britain. *Clio Medica* 32:235-58.

Horton, R. 1995. The rhetoric of research. *BMJ* 310:985-87.

Horwich, P. 1982. *Probability and evidence.* Cambridge: Cambridge University Press.

Hostettler, J. 1993. *Thomas Wakely - an improbable radical.* Chichester: Barry Rose Law Publishers.

Howard-Jones, N. 1982. Human experimentation in historical and ethical perspectives. *Social Science and Medicine* 16:1429-48.

Howson, C., and P. Urbach. 1989. *Scientific reasoning: the Bayesian approach.* La Salle, Ill: Open Court.

---. 1993. *Scientific reasoning: the Bayesian approach.* 2nd ed. La Salle, Ill: Open Court Publishing.

Hughes, R. E. 1985. James Lind and the cure of scurvy: an experimental approach. *Medical History*:342-51.

Inman, H. F. 1994. Karl Pearson and RA Fisher on statistical tests: a 1935 exchange from Nature. *American Statistician* 48:2-11.

Irwin, J. O. 1934. Some aspects of the development of modern statistical method. *Mathematical Gazette*:18-34.

Jacobs, M. S. 1940. Paul Ehrlich and his relation to modern chemotherapy. *Bulletin of the History of Medicine* 8:956-64.

James, R. R. 1994. *Henry Wellcome.* London: Hodder and Stoughton.

Jarcho, S. 1995. *Quinine's predecessor: Francesco Torti and the early history of Cinchona.* Baltimore: The Johns Hopkins University Press.

Jardine, N. 1992. The laboratory revolution in medicine as rhetorical and aesthetic accomplishment. In *The laboratory revolution in medicine*, ed. A. Cunningham and P. Williams, 304-23. Cambridge: Cambridge University Press.

Jeffreys, H. 1961. *Theory of probability.* London: Clarendon Press.

Jenkins, D. E., W. M. Peck, J. J. Rowland Reid, and H. S. Willis. 1947. Effect of streptomycin on early tuberculous pulmonary lesions. A preliminary report. *American Review of Tuberculosis* LVI:387-95.

Jewson, N. D. 1974. Medical knowledge and the patronage system in eighteenth century England. *Sociology* 1974:369-85.

---. 1976. The disappearance of the sick man from medical cosmology, 1770-1870. *Sociology* 10:225-44.

Johnstone, D. J. 1989. On the necessity for random sampling. *British Journal of the Philosophy of Science* 40:443-57.

Jordanova, L 1993. Has the social history of medicine come of age? *Historical Journal* xxxvi:437-39.

---. 1995. The social construction of medical knowledge. *Social History of Medicine* 8:361-81.

Kadane, J. B. 1996. *Bayesian methods and ethics in a clinical trials design.* New York: John Wiley.

Kadane, J. B., Seienfeld, T. 1990. Randomization in a Bayesian perspective. *Journal of Statistical Planning and Inference* 25:329-45.

Kaplan, B. B. 1993. '*Divulging of useful truths in physik': the medical agenda of Robert Boyle.* Baltimore: Johns Hopkins University Press.

Kaptchuk, T. K. 1998. Intentional ignorance: a history of blind assessment and placebo controls in medicine. *Bulletin of the History of Medicine* 72:389-433.

Kaye, D. H. 1987. Apples and oranges: confidence coefficients and the burden of persuasion. *Cornell Law Review* 73:54-77.

Keating, P. 1988. Vaccine therapy and the problems of opsonins. *Journal of the History of Medicine* 43:275-96.

Keefer, C. S. et at 1946. Streptomycin in the treatment of infections: a report of one thousand cases. *Journal of the American Medical Association* 132:4-11, and 70-77.

Keele, K. D. 1968. Clinical medicine in the 1860s. In *Medicine and science in the 1860s*, ed. F. N. L. Poynter, 1-11. London: Wellcome Institute of the History of Medicine.

Kelves, D. J., and G. L. Geison. 1995. The experimental life sciences in the twentieth century. *Osiris (Second Series)* 10:97-121.

Kempthorne, O. 1966. Some aspects of experimental inference. *Journal of the American Statistical Association* 16:11-34.

---. 1977. Why randomize? *Journal of Statistical Planning and Inference* 1:1-25.

Kendall, D. G., M. S. Bartlett, and T. L. Page. 1984. Neyman, Jerzy 16 April 1894 - 5 August 1981 - obituary. *Bulletin of the London Mathematical Society* 16:160.

Kendall, M. G. 1956. The beginnings of a probability calculus. *Biometrika* 43:1-14.

---. 1960 Studies in probability and statistics. X. Where shall the history of statistics begin? *Biometrika* 47:447-9.

Kim, K-M. 1994. *Explaining scientific consensus: the case of Mendelian genetics*. New York: Guilford Press.

King, A. 1974. The life and times of Colonel Harrison. *British Journal of Venereal Diseases* 50:391-403.

King, L. S. 1952. Dr Koch's postulates. *Journal of the History of Medicine and Allied Sciences* 7:350-61.

King, M. D. 1971. Reason, tradition and the progressiveness of science. *History and Theory* 10:3-32.

Kitcher, P. 1993. *Science without legend, objectivity without illusions.* Oxford: OUP.

Klaudt, K ed. 1994. *TB, a global emergency*. Geneva: WHO.

Knobloch, E. 1987. Emile Borel as a probabilist. In Kruger, L et al eds. *The probabilistic revolution. Volume 1: ideas in history*. p 215-235 Cambridge, Mass,: MIT Press.

Knoll, E. 1992. The American Medical Association and its journal. In *Medical journals and medical knowledge*, ed. W. F. Bynum, S. Lock, and R. Porter, 146-64. London: Routledge.

Kohler, R. E. 1978. Walter Fletcher, FG Hopkins, and the Dunn Institute of Biochemistry: a case study in the patronage of science. *Isis* 69:331-55.

---. 1978. A policy for the advancement of science: the Rockefeller Foundation. *Minerva* 16:480-515.

---. 1979. Medical reform and biomedical science: biochemistry - a case study. In *The therapeutic revolution: essays in the social history of American medicine*, ed. M. J. Vogel and C. E. Rosenberg, 27-66. Philadelphia: University of Pennsylvania Press.

---. 1986. Bacterial physiology: the medical context. *Bulletin of the History of Medicine* 60:55-74.

Kolmogorov, A. 1956. *Foundations of the theory of probability.* 2nd English Edition ed. New York: Chelsea Publishing Company.

Kremers, E. and G. Urdang. 1963. *History of pharmacy*. 3rd ed. revised by Glenn Sonnedecker. Philadelphia: J. B. Lippincott.

Krüger, L., L. J. Daston, and M. Heidelberger, eds. 1987. *The probabilistic revolution (2 vols.).* Cambridge, Mass.: MIT Press.

Kuhn, T. S. 1961. The function of measurement in modern physical science. *Isis* 52:162-76.

---. 1962. *The structure of scientific revolutions*. Chicago: University Press.

Kunz, R., and A.D. Oxman. 1998. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ* 317:1185-1190.

Kuriyama, S. 1995. Interpreting the history of bloodletting. *Journal of the History of Medicine and Allied Sciences* 50:11-46.

Kyberg, H. E. 1961. *Probability and the logic of rational belief.* Middletown: Wesleyan University Press.

---. 1974. *The logical foundations of statistical inference.* Synthese Library, v65. Dordrecht: D Reidel.

---. 1983. The reference class. *Philosophy of Science* 50:374-97.

Lancaster, H. O. 1994. *Quantitative methods in biological and medical sciences.* New York: Springer-Verlag.

Landsborough Thomson, A. 1973. *Half a century of medical research, Volume 1, origins and policy of the Medical Research Council (UK).* London: HMSO.

---. 1973. Origin of the British legislative provision for medical research. *Journal of Social Policy* 2:43-47.

---. 1975. *Half a century of medical research, Volume II: the programme of the Medical Research Council (UK).* London: HMSO.

Langley, G. J. 1931. Serum treatment of pneumonia. *BMJ* ii:1139.

Lasagna, L. 1955. The controlled clinical trial: theory and practice. *Journal of Chronic Diseases* 1:357-58.

Latham, A. 1903. *The diagnosis and modern treatment of pulmonary consumption.* London: Bailliere, Tindall and Cox.

Latour, B. 1983. Give me a laboratory and I will raise the world. In *Science observed: perspectives on the social study of science*, ed. K. Knorr-Cetina and M. Mulkay. London: Sage.

---. 1987. *Science in action.* Milton Keynes: Open University Press.

---. 1988. *The pasteurization of France.* Cambridge, Mass.: Harvard University Press.

---. 1988. The politics of explanation: an alternative. In *Knowledge and reflexivity: new frontiers in the sociology of knowledge*, ed. S. Woolgar, 155-76. London: Sage.

---. 1992. The costly ghastly kitchen. In *The laboratory revolution in medicine*, ed. A. Cunningham and P. Williams, 295-303. Cambridge: Cambridge University Press.

---. 1993. *We have never been modern.* London: Harvester Wheatsheaf.

Latour, B., and S. Woolgar. 1979. *Laboratory life: the social construction of scientific facts.* London: Sage Publications.

Lawrence, C. J. 1985. Incommunicable knowledge: science, technology and the clinical art in Britain 1850 - 1914. *Journal of Contemporary History* 20:503-20.

---. 1985. Moderns and ancients: the "new cardiology" in Britain 1880-1930. *Medical History* Suppl 5:1-33.

---. 1992. *Medical theory, surgical practice.* London: RKP.

---. 1994. *Medicine in the making of modern Britain 1700-1920.* London: Routledge.

Lazarsfeld, P. 1961. Notes on the history of quantification in sociology - trends, sources, and problems. *Isis* 52:164-81.

Lécuyer, B.-P. 1987. Probability in vital and social statistics: Quetelet, Farr, and the Bertillons. In *The probabilistic revolution. Volume 1: ideas in history*, ed. L. Krüger, L. J. Daston, and M. Heidelberger, 317-35. Cambridge: MIT Press.

Lederer, S. E. 1984. "The right and wrong of making experiments on human beings": Udo J. Wile and syphilis. *Bulletin of the History of Medicine* 58:380-97.

---. 1992. Orphans as guinea pigs: American children and experimentation, 1890-1930. In *In the name of the child: health and welfare, 1880-1940*, ed. R. Cooter, 96-123. London: Routledge

---. 1994. *Subjected to science: human experimentation in America before the Second World War.* Baltimore: Johns Hopkins University Press.

Lehmann, R. S. 1995. Statistics in the behavioural sciences: a conceptual introduction. London: Brooks/Cole.

Leibowitz, D. 1993. Scientific failure in an age of optimism: public reaction to Robert Koch's tuberculin cure. *New York State Journal of Medicine* 93:41-48.

Leighton, G., and M. L. Clark. 1929. Milk consumption and the growth of schoolchildren: second preliminary report on tests to the Scottish Board of Health. *Lancet.* i:40-43.

Leighton, G., and P. L. McKinlay. 1930. *Milk consumption and the growth of school children: report on an investigation in Lanarkshire schools.* Edinburgh: HMSO.

Lennard-Jones, J. E. 1993. Evolution of the controlled clinical trial. *Mt. Sinai. J. Med.* 60:180-85.

Lenoir, T. 1988. A magic bullet: research for profit and the growth of knowledge in Germany around 1900. *Minerva* 26:66-88.

---. 1992. Laboratories, medicine and public life in Germany 1830-1849. In *The laboratory revolution in medicine*, ed. A. Cunningham and P. Williams, 14-71. Cambridge: Cambridge University Press.

Levi, I. 1967. *Gambling with truth.* New York: Alfred Knopf.

---. 1980. *The enterprise of knowledge: an essay on knowledge, credal probability, and choice.* Cambridge, Mass.: MIT Press.

---. 1983. Direct inference and randomization. In *Proceedings of the 1982 biennial meeting of the Philosophy of Science Association*, ed. P. D. Asquith and T. Nickles, 447-63. Vol. Vol 2. East Lansing, Mi.: Philosophy of Science Association.

Lewis, T. 1930. Research in medicine: its position and its needs. *British Medical Journal* i:479-83.

Liebenau, J. 1984. Industrial R&D in pharmaceutical firms in the early 20th century. *Business History* 26:329-46.

---. 1985. Scientific ambitions: the pharmaceutical industry, 1900-1920. *Pharmacy in History* 27:3-11.

---. 1987. *Medical science and medical industry.* London: Macmillan.

---. 1989 The MRC and the pharmaceutical: the model of insulin. In Austoker, J., and L. Bryder, eds. *Historical perspectives on the role of the MRC: essays on the history of the Medical Research Council of the United Kingdom and its predecessor, the Medical Research Committee, 1913-53.* Oxford: Oxford University Press. p163-180.

---. 1990. The rise of the British pharmaceutical industry. *British Medical Journal* 301:724-28.

---. 1990. Paul Ehrlich as a commercial scientist and research administrator. *Medical History* 34:65-78.

Lilienfeld, A. M., ed. 1980. *Times, places and persons: aspects of the history of epidemiology.* Baltimore: Johns Hopkins University Press.

---. 1982. Ceteris Paribus: the evolution of the clinical trial [Fielding H Garrison Lecture]. *Bulletin of the History of Medicine* 56:1-18.

Lilienfeld, D. E. 1978. "The greening of epidemiology": sanitary physicians and the London Epidemiological Society (1830-1870). *Bulletin of the Institute of the History of Medicine* 52:503-28.

Lindley, D. V. 1983. The role of randomisation in inference. In *Proceedings of the 1982 biennial meeting of the Philosophy of Science Association*, ed. P. D. Asquith and T. Nickles, 431-46. Vol. 2. East Lansing, Mi.: Philosophy of Science Association.

---. 1984. A Bayesian lady tasting tea. In *Statistics: an appraisal*, ed. H. A. David and H. T. David, 455-77. Iowa: Iowa State University Press.

Lloyd, C. 1961. The introduction of lemon juice as a cure for scurvy. *Bulletin of the History of Medicine* 35:123-32.

Lock, S. 1994. The randomised controlled trial - a British invention. In *Technologies of modern medicine*, ed. G. Lawrence, 81-87. London: Science Museum.

Logan, W. P. D. 1950. Mortality in England and Wales from 1848 to 1947. *Population Studies* 4:132-78.

Long, E. R. 1940. The decline of tuberculosis with special reference to its generalised form. *Bulletin of the History of Medicine* 8:819-43.

Loudon, I. 1986. *Medical care and the general practitioner.* Oxford: Clarendon Press.

Loudon, J., and I. Loudon. 1992. Medicine, politics and the medical periodical 1800-50. In *Medical Journals and medical knowledge*, ed. W. F. Bynum, S. Lock, and R. Porter, 49-69. London: Routledge.

---. 1992 Medicine, politics and the medical periodical 1800-1850. In *Medical journals and medical knowledge*, ed. W. F. Bynum, S. Lock, and R. Porter, 49-69. London: Routledge.

Louis, P. C. A. 1836. *Researches on the effects of blood letting in some inflammatory diseases and on the influence of tartarized antimony and vesication in pneumonitis.* Boston: Hillard, Grey.

---. 1836. *Pathological researches on phthisis.* Boston: Hilliard, Gray.

Löwy I. 1994. Recent historiography of biomedical research. In Lawrence G ed. *Technologies of modern medicine*. London: Science Museum.

Löwy, I., O. Amsterdamska, J. V. Pickstone, and P. Pinnell, eds. 1993. *Medicine and change: historical and sociological studies of medical innovation.* Paris: Mouton.

Lucas, J. R. 1970. *The concept of probability.* Oxford: Clarendon Press.

Macdonald, G. 1980. *In pursuit of excellence: one hundred years of Wellcome, 1880-1980.* London: Wellcome Foundation.

Mackenzie, D. A. 1976. Eugenics in Britain. *Social Studies of Science* 6:499-532.

---. 1978. Statistical theory and social interests. *Social Studies of Science* 8:35-83.

---. 1979. Karl Pearson and the professional middle class. *Annals of Science* 36:125-43.

---. 1981. *Statistics in Britain, 1865-1930. The social construction of scientific knowledge.* Edinburgh: Edinburgh University Press.

---. 1981. Sociobiologies in competition: the Biometrician-Mendelian debate. In *Biology, medicine and society 1840-1940*, ed. C. Webster. Cambridge: Cambridge University Press.

MacNalty, A. 1951. The control of pulmonary tuberculosis in England. *British Journal of Tuberculosis and Diseases of the Chest* XLV:90-96.

Madison, D. L. 1996. Preserving individualism in the organizational society: 'cooperation' and American medical practice, 1900-1920. *Bulletin of the History of Medicine* 70:442-83.

Mainland, D. 1934. Chance and the blood count [editorial]. *Canadian Medical Association Journal* 31:656-58.

---. 1936. Problems of chance in clinical work. *British Medical Journal* ii:221-24.

---. 1938. *The treatment of clinical and laboratory data: an introduction to statistical ideas and methods for medical and dental workers.* London: Oliver & Boyd.

---. 1952. *Elementary medical statistics: the principles of quantitative medicine.* Philadelphia: WB Saunders.

---. 1960. The clinical trial - some difficulties and suggestions. *Journal of Chronic Diseases* 11:484-96.

---. 1960. The use and misuse of statistics in medical publications. *Clinical Pharmacology and Therapeutics* 1:411-22.

---. 1984. Statistical ritual in clinical journals: is there a cure? - II. *BMJ* 288:920-22.

Mann, H. B. and D. R. Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics* 18:50-60.

Margolis, J. 1989. *Texts without referents: reconciling science and narrative.* Oxford: Blackwell.

Marks H. M. 1988. Notes from the underground: the social organisation of therapeutic research. In *Grand rounds: one hundred years of internal medicine*, ed. R. C. Maulitz and D. E. Long. Philadelphia: University of Pennsylvania Press.

---. 1992. Cortisone, 1949: a year in the political life of a drug. *Bulletin of the History of Medicine* 66:419-39.

---. 1997. *The progress of experiment: science and therapeutic reform in the United States, 1900-1990.* New York: Cambridge University Press.

Marland, H. 1987**a**. *Medicine and society in Wakefield and Huddersfield, 1780-1870.* Cambridge: CUP.

---. 1987**b**. The medical activities of mid-nineteenth-century chemists and druggists, with special reference to Wakefield and Huddersfield. *Medical History* 31:415-39.

---. 1993. A pioneer in infant welfare: the Huddersfield Scheme 1903-1920. *Social History of Medicine* 6:25-50.

Martin, S. 1896. An address on the serum treatment of diphtheria. *Lancet* ii:1059-62.

Martin, S. C. 1994. "The only truly scientific method of healing". Chiropractic and American science, 1895-1990. *Isis* 85:207-27.

Martin-Lof, P. 1970. On the notion of randomness. In *Intuitionism and proof theory*, ed. A. Kino, J. Myhill, and R. E. Vesley, 73-78. Amsterdam: North-Holland.

Maulitz, R. C. 1979. 'Physician versus bacteriologist': the ideology of science in clinical medicine. In *The Therapeutic revolution: essays in the social history of American medicine*, ed. M. J. Vogel and C. E. Rosenberg, 91-107. Philadelphia: University of Philadelphia Press.

---. 1987. *Morbid appearances: the anatomy of pathology in the early Nineteenth century.* Cambridge: CUP.

Mayo, D. G. 1981. In defence of the Neyman-Pearson theory of confidence intervals. *Philosophy of Science* 48:269-80.

Mayo, O. 1987. Comments on "randomisation and the design of experiments" by P Urbach. *Philosophy of Science* 54:592-96.

Mazumdar, P. 1992. *Eugenics, human genetics and human failings. The Eugenics Society, its sources and its critics in Britain.* London: Routledge.

McFarlane, N. 1989. Hospitals, housing, and tuberculosis in Glasgow, 1911-51. *Social History of Medicine* 2:59-85.

McKeown, T. 1976. *The modern rise of population.* London: Edward Arnold.

McKeown, T., and R. G. Record. 1962. Reasons for the decline in mortality in England and Wales during the Nineteenth century. *Population Studies* 16:94-122.

McKeown, T., R. G. Record, and R. D. Turner. 1975. An interpretation of the decline of mortality in England and Wales during the Twentieth century. *Population Studies* 29:391-422.

McNeil, C. 1923. The cutaneous tuberculin test (Pirquet). *BMJ* i:673-75.

Meachen, G. N. 1936. *A short history of tuberculosis.* London: Bale, Sons and Danielsson.

Means, J. H., and A. L. Barach. 1921. The symptomatic treatment of pneumonia. *JAMA* 77:1217.

Mearns Fraser, A. 1912. National Insurance and the municipal control of tuberculosis. *British Journal of Tuberculosis and Diseases of the Chest* 6:86-92.

Medawar, P. 1964. Is the scientific paper a fraud? *Saturday Review* August 1:43-44.

---. 1986. *Memoir of a thinking radish*. Oxford: Oxford University Press

MRC (Medical Research Committee/Council) – committee publications. Publications with named authors are listed under the name of the lead author.

---. 1919. *Reports of the Special Committee upon the Manufacture, biological testing, and clinical administration of Salvarsan and its substitutes. No 1.* Special Report Series 44. London: HMSO. (Salvarsan Committee).

---. 1934. The serum treatment of lobar pneumonia. A report of the Therapeutic Trials Committee of the Medical Research Council. *Lancet.* i:290-95.

---. 1944. Clinical trial of Patulin for the common cold. *Lancet.* ii:373-5. (Patulin Trial Committee).

---. 1948. Streptomycin treatment of pulmonary tuberculosis. *BMJ* ii:769-82. (Streptomycin in Tuberculosis Trials Committee).

---. 1954. A five year assessment of patients in a controlled trial of streptomycin in pulmonary tuberculosis. *Quarterly Journal of Medicine* 23:347. (Streptomycin in Tuberculosis Trials Committee).

Meldrum, M. 1996. "Simple methods" and "determined contraceptors": the statistical evaluation of fertility control, 1957-1968. *Bulletin of the History of Medicine* 70:266-95.

---. 1998. “A calculated risk”: the Salk polio vaccine field trials of 1954. *BMJ* 317:1233-6.

Merrington, W. R. 1976. *University College Hospital and its medical school: a history*. London: Heinemann.

Miley, U. and J. Pickstone. 1988. Medical botany around 1850: American medicine in industrial Britain. In Cooter, E. ed. *Studies in the history of alternative medicine*. Basingstoke: Macmillan.

Ministry of Health. 1948. Streptomycin treatment of tuberculosis. *BMJ* ii:527-28.

---. 1950. Streptomycin in the treatment of tuberculous meningitis. *Tubercle* 41:214-15.

Mises, R. von. 1957. *Probability, statistics and truth.* 2nd rev ed. New York: Macmillan.

Mitchell, A. 1990. An inexact science: the statistics of tuberculosis in late nineteenth century France. *Social History of Medicine* 3:387-403.

Moir, C. 1932**a**. Clinical comparison of ergotoxine and ergotamine. *British Medical Journal* i:1022-24.

---. 1932**b**. The action of ergot preparations on the uterus, with a note by Sir HH Dale. *British Medical Journal* i:1119-1122.

Morgan, N. 1986. A note on the proposed amalgamation of the Lister Institute of preventive medicine and the Medical Research Committee. *Annals of Science* 43:287-89.

Morrison, D. E., and R. E. Henkel, eds. 1970. *The significance test controversy.* Chicago: Aldine Publishing.

Moscucci, O. 1990. *The science of woman: gynaecology and gender in England, 1800-1929*. Cambridge: Cambridge University Press.

Moseley, R. 1978. The origins and early years of the National Physical Laboratory: a chapter in the pre-history of British science policy. *Minerva* 16:222-50.

Mullard, K. S. 1950. The use of streptomycin and PAS in the surgical treatment of pulmonary tuberculosis. *Tubercle* 41:236-40.

Murnaghan, J. H., and P. Talalay. 1992. Dale's account of the standardization of insulin. *Bulletin of the History of Medicine* 66:440-50.

Murphy, T. D. 1981. Medical knowledge and statistical methods in early nineteenth century France. *Medical History* 25:301-19.

Nelson, R. B. 1979. Are clinical trials pseudoscience? *Forum Med* 2:594-600.

Nester, M. R. 1996. An applied statisticians creed. *Journal of the Royal Statistical Society - Series C* 45:401-10.

Neuberg, L. G. 1989. *Conceptual anomalies in economics and statistics: lessons from the social experiment.* Cambridge: Cambridge University Press.

Neuberger, M. 1932. Resumption of discredited methods of treatment. *Medical Life* 39:694-98.

Newman, M. 1994. Randomised Controlled Trials [letter]. *Journal of Epidemiology and Community Medicine* 48:423-24.

Newsholme, A. 1898. *Epidemic diphtheria.* London: Swan Sonnenschein.

---. 1906. An inquiry into the principal causes of the reduction of the death rate from phthisis during the last forty years with special reference to the segregation of phthisical patients in general institutions. *Journal of Hygiene* 6:304-84.

Neyman, J. 1937. Outline of a theory of statistical estimation based on the classical theory of probability. *Philosophical Transactions of the Royal Society A* ccxxxvi:333-80.

---. 1950. *A first course in probability and statistics.* New York: Holt.

---. 1955. Statistics - servant of all sciences. *Science* 122:401-6.

---. 1967. R. A. Fisher: an appreciation. *Science* 156:1456-60.

Neyman, J., and E. S. Pearson. 1928**a**. On the use and interpretation of certain test criteria for purposes of statistical inference. Part 1. *Biometrika* 20A:175-240.

---. 1928**b**. On the use and interpretation of certain test criteria for purposes of statistical inference. Part 2. *Biometrika* 20A:263-94.

---. 1933. The testing of statistical hypotheses in relation to probabilities a priori. *Proceedings of the Cambridge Philosophical Society* 29:492-510.

Nicholls, P. A. 1988. *Homeopathy and the medical profession.* London: Croom Helm.

Niebyl, P. H. 1977. The English bloodletting revolution, or modern medicine before 1850. *Bulletin of the History of Medicine* 51:464-83.

Oakes, M. 1986. *Statistical inference: a commentary for the social and behavioural sciences.* Chichester: John Wiley.

Oakley, A. 1989. Who's afraid of the Randomized Controlled Trial? Some dilemmas of the scientific method and "good" research practice. *Women and Health* 15:25-59.

Ott, K. 1996. *Fevered lives: tuberculosis in America since 1870.* Cambridge, Mass.: Harvard University Press.

Oudshoorn, N. 1990**a**. On measuring sex hormones: the role of biological assays in sexualizing chemical substances. *Bulletin of the History of Medicine* 64:243-61.

---. 1990**b** On the making of sex hormones. *Social Studies of Science* 20:5-33

---. 1995. *Beyond the natural body. An archaeology of sex hormones.* London: Routledge.

Owen, D. B., ed. 1976. *On the history of statistics and probability. Proceedings of a symposium, May 27-29, 1974.* New York: Marcel Drekker.

Papineau, D. 1994. The virtues of randomization. *British Journal of the Philosophy of Science* 45:437-50.

Pappworth, M. H. 1967. *Human guinea pigs. Experimentation on man.* London: RKP.

---. 1990. 'Human guinea pigs' - a history. *BMJ* 301:1456-60.

Parascandola, J. 1974. The controversy over structure-activity relationships in the early twentieth century. *Pharmacy in History (US)* 16:54-63.

---. 1976. Chemistry and pharmacology in Britain, 1840-1920. *Pharm Hist (GB)* 6:1-3.

---. 1977. Carl Voegtlin and the 'arsenic receptor' in chemotherapy. *Journal of the History of Medicine* 32:151-71.

---. 1981. The theoretical basis of Paul Ehrlich's chemotherapy. *Journal of the History of Medicine* 36:19-43.

---. 1985. Patent medicines in Nineteenth Century America. *Caduceus* 1:1-41.

---. 1985. Industrial research comes of age: the American pharmaceutical industry 1920-1940. *Pharm Hist (US)* 27:12-21.

---. 1988. The beginnings of pharmaceutical work in the Federal government. *Pharmacy in History (US)* 30:179-87.

---. 1992. *The development of American pharmacology: John J Abel and the shaping of a discipline.* Baltimore: Johns Hopkins University Press.

Parascandola, J., and E. Keeney. 1983. *Sources in the history of American pharmacology.* Madison Wisconsin: AIHP.

Parish, H. J. 1965. *A history of immunization.* Edinburgh: E&S Livingstone.

Park, W. H. and G Cooper. 1928. Antipneumococcus serum in lobar pneumonia: method of administration and dosage. *Journal of the American Medical Association* 90:1349-51.

Park, W.H., Bullowa, J. G. M. and M. B. Rosenbluth. 1928. Treatment of lobar pneumonia with refined specific antibacterial serum. *Journal of the American Medical Association* 91:1503-8.

Parnall, C., B. L. Brock, and R. E. Moyer. 1947. Streptomycin in the treatment of pulmonary tuberculosis. *American Review of Tuberculosis* 56:565-78.

Parry, N., and J. Parry. 1976. *The rise of the medical profession. A study in collective social mobility.* London: Croom Helm.

Paton, W. 1979. The evolution of therapeutics. *Journal of the Royal College of Physicians (London)* 13:74-83.

Payer, H. C. 1996. *Roche: a company history.* Basel: Editiones Roche.

Payne Cotton, R. 1858. *On consumption: its nature, causes and treatment.* 2nd ed. London: Churchill.

Pearson, E. S. 1962. Some thoughts on statistical inference. *Annals of Mathematical Statistics* 33:394-403.

---. 1965. Some incidents in the early history of biometry and statistics. *Biometrika* 52:3-18.

---. 1990. *"Student": a statistical biography of William Sealy Gossett.* Oxford: Clarendon Press.

Pearson, E. S., and M. G. Kendall, eds. 1970. *Studies in the history of statistics and probability.* London: Charles Griffin and Co.

Pearson, K. 1894. Dilettantism and statistics. *Nature.* 51:145-46.

---. 1924. Historical note on the origin of the normal curve of errors. *Biometrika* 16:402-4.

---. 1935. Statistical tests. *Nature.* 136:296-97,500.

---. 1936. Method of moments and method of maximum likelihood. *Biometrika* 28:34-59.

---. 1978. *The history of statistics in the 17th and 18th centuries against the changing background of intellectual, scientific and religious thought. Lectures by Karl Pearson given at University College London during the academic sessions 1921-1933.* London: Charles Griffin.

Petersen, M. J. 1978. *The medical profession in mid-Victorian London.* Berkeley: University of California Press.

Pfeffer, N. 1985. Biological standardisation and sex hormones. BSA Medical Sociology Group September 1985.

---. 1994. *The stork and the syringe: a political history of reproductive medicine.* Oxford: Polity Press.

Pfuetze, K. H., M. M. Pyle, H. C. Hinshaw, and W. H. Feldman. 1954. The first clinical trial of streptomycin in human tuberculosis. *American Review of Tuberculosis and Pulmonary Diseases* 71:752-54.

Philip, R. 1924. The effect of the anti-tuberculosis campaign on the diminution of the mortality from tuberculosis. *Edinburgh Medical Journal* 31:482-526.

Pickstone, J. V. ed. 1992. *Medical innovations in historical perspective*. London: Macmillan.

---. 1993. The biographical and the analytical: towards a historical model of science and practice in modern medicine. In *Medicine and change. Proceedings of the symposium ISERM Paris 21-23 April 1992*, ed. I. Löwy, 23-47. London: John Libby and Co.

---. 1993. Ways of knowing: towards a historical sociology of science technology and medicine. *British Journal of the History of Science* 26:433-58.

---. 1994. Museological science? The place of the analytic/comparative in nineteenth century science, technology and medicine. *History of Science* 32:111-38.

---. 1994. Objects and objectives: notes on the material culture of medicine. In *Technologies of modern medicine*, ed. G. Lawrence, 13-24. London: Science Museum.

Pickstone, J. V., R. Cooter, and C. S. Murphy. 1985. Exploring 'clinical research'. Academic medicine and clinicians in early Twentieth century Britain. *Society for the Social History of Medicine Bulletin* 37:79-81.

Plackett, R. L. 1958. The principle of the arithmetic mean. *Biometrika* 45:30-35.

---. 1972. The discovery of the method of least squares. *Biometrika* 59:239-51.

Political and Economic Planning. 1937. *Report on the British health services.* London: PEP.

Pocock, S. J. 1983. *Clinical trials: a practical approach.* Chichester: John Wiley.

---. 1985. Current issues in the design and interpretation of trials. *BMJ* 290:39-42.

---. 1996. Clinical trials: a statistician's perspective. In *Advances in biometry*, ed. P. Armitage and H. David, 405-21. Chichester: Wiley.

Popper, K. R. 1959. The propensity interpretation of probability. *British Journal of the Philosophy of Science* 10:25-42.

---. 1968. *Conjectures and refutations: the growth of scientific knowledge.* New York: Harper and Row.

---. 1972. *Objective knowledge.* Oxford: Oxford University Press.

---. 1983. *A pocket Popper*. London: Fontana.

Porter, R. 1985. Laymen, doctors, and medical knowledge in the Eighteenth century: the evidence of the Gentleman's magazine. In *Patients and practitioners: lay perceptions of medicine in pre-industrial society*, ed. R. Porter, 283-314. Cambridge: CUP.

---. 1985. The patient's view: doing medical history from below. *Theory and Society* 14:175-98.

---. 1985. Lay medical knowledge in the eighteenth century: the evidence of the Gentleman's Magazine. *Medical History* 29:138-68.

---. 1986. The language of quackery in England 1660-1800. In *Language and society*, ed. P. Burke and R. Porter. Cambridge: Cambridge University Press.

---. 1988. Quackery and the eighteenth century medical market. In *Studies in the history of alternative medicine*, ed. R. Cooter, 7-9. London: Macmillan.

---. 1989. *Health for sale: quackery in England 1650-1850.* Manchester: Manchester University Press.

---. 1989. The early Royal Society and the spread of medical knowledge. In *Medical Revolution of the Seventeenth century*, ed. R. French and A. Wear, 280-81. Cambridge: CUP.

---. 1991. The rise of medical journalism in Britain to 1800. In *Medical journals and medical knowledge: historical essays*, ed. W. F. Bynum, S. Lock, and R. Porter, 6-28. London: RKP.

Porter, T. M. 1985. The mathematics of society: variation and error in Quetelet's statistics. *British Journal of the History of Science* 18:51-69.

---. 1986. *The rise of statistical thinking 1820-1900.* Princeton: Princeton University Press.

---. 1987. Lawless society: social science and the reinterpretation of statistics in Germany, 1850-1880. In *The probabilistic revolution. Volume 1: ideas in history*, ed. L. Krüger, L. J. Daston, and M. Heidelberger, 351-75. Cambridge: MIT Press.

---. 1992. Objectivity as standardization: the rhetoric of impersonality in measurement, statistics, and cost-benefit analysis. *Annals of Scholarship* 9:19-60.

---. 1994. Making things quantitative. *Science in Context* 7:389-407.

---. 1995. *Trust in numbers: the pursuit of objectivity in science and public life.* Princeton: Princeton University Press.

Poynter, F. N. L., ed. 1965. *The evolution of pharmacy in Britain.* London: Pitman Medical Publishing.

Quen, J. M. 1975. Case studies in nineteenth century scientific rejection: mesmerism, perkinism and acupuncture. *Journal of the History of the Behavioural Sciences* 11:149-56.

Quetel, C. 1990. *History of syphilis.* London: Polity Press.

Ramsey, M. 1977. Medical power and popular medicine: illegal healers in 19th century France. *Journal of Social History* 10:560-87.

---. 1982. The repression of unauthorized medical practice. *Eighteenth Century Life* 7:118-35.

---. 1982. Traditional medicine and medical enlightenment: the regulation of secret remedies in the Ancien Régime. *Historical Reflections* 9.

---. 1984. The politics of professional monopoly in nineteenth century medicine: the French model and its rivals. In *Professions and the French state 1700-1900*, ed. G. L. Geison. : University of North Carolina Press.

---. 1988. *Professional and popular medicine in France 1770-1830.* Cambridge: Cambridge University Press.

Rankin, B. 1966. The history of probability and the changing concept of the individual. *Journal of the History of Ideas* 27:483-504.

Ransome, A. 1868. On epidemics, studied by means of statistics of disease. *BMJ* ii:386-88.

---. 1896. *The treatment of phthisis.* London: Smith Elder & Co.

Rawlins, M. D. 1990. Development of a rational practice of therapeutics. *British Medical Journal* 301:729-33.

Rescher, N. 1958. Can there be random individuals? *Analysis* 18:114-16.

---. 1978. *Peirce's philosophy of science: critical studies in his theory of induction and scientific method.* Notre Dame: University of Notre Dame Press.

Richards, E. 1988. The politics of therapeutic evaluation: the vitamin C and cancer controversy. *Social Studies of Science* 18:653-701.

---. 1991. *Vitamin C and cancer controversy: medicine or politics?* London: Macmillan.

Risse, G. B. 1971. The quest for certainty in medicine: John Brown's system of medicine in France. *Bulletin of the History of Medicine* 45:1-12.

---. 1973. Calomel and the American medical sects during the Nineteenth century. *Mayo Clinic Proceedings* 48:57-64.

---. 1979. The renaissance of bloodletting: a chapter in modern therapeutics. *Journal of the History of Medicine* 34:3-22.

---. 1985. Typhus fever in Eighteenth century hospitals: new approaches to medical treatment. *Bulletin of the History of Medicine* 59:176-95.

---. 1991. The history of therapeutics. In *Essays in the history of therapeutics*, ed. W. F. Bynum and V. Nutton, 3-8. Vol. Clio Medica 22. Amsterdam: Editions Rodolphi.

Robbins, J. M. 1997. Class struggles in the tubercular world: nurses patients and physicians, 1903-1915. *Bulletin of the History of Medicine* 71:412-34.

Robson, M. 1985. The British pharmaceutical industry in World War One. *Society for the Social History of Medicine Bulletin* 37:60-62.

Romano, T, M. 1997 The cattle plague of 1865 and the reception of "The Germ Theory" in mid-Victorian Britain. *Journal of the History of Medicine and Allied Sciences* 52:51-80.

Rosen, G. 1946. Mesmerism and surgery: a strange chapter in the history of anaesthesia. *Journal of the History of Medicine* 1:527-50.

---. 1951. The new history of medicine: a review. *Journal of the History of Medicine* 6:516-22.

---. 1955. Problems in the application of statistical analysis to questions of health. *Bulletin of the History of Medicine* 24:27-45.

Rosenberg, C. E. 1979. The therapeutic revolution: medicine, meaning, and social change in Nineteenth century America. In *The therapeutic revolution: essays in the social history of American medicine*, ed. M. J. Vogel and C. E. Rosenberg, 3-25. Philadelphia: University of Pennsylvania Press.

Rosenblatt, M. B. 1973. Pulmonary tuberculosis: evolution of modern therapy. *Bulletin of the New York Academy of Medicine* 49:163-96.

Rosenkrantz, B., ed. 1994. *From consumption to tuberculosis.* New York: Garland Publishers.

Ross, J. E., and M. Tomkins. 1997. The British reception of Salvarsan. *Journal of the History of Medicine and Allied Sciences* 52:398-423.

Rosser Matthews, J 1995. *Quantification and the quest for medical certainty.* Princeton, NJ: Princeton University Press.

---. 1995. Major Greenwood versus Almroth Wright: contrasting visions of scientific medicine in Edwardian Britain. *Bulletin of the History of Medicine* 69:30-43.

Rothman, D. J., and H. Edgar. 1992. Scientific rigor and medical realities: placebo trials in cancer and AIDS research. In *AIDS: the making of a chronic disease*, ed. E. Fee and D. M. Fox, 194-206. Berkeley: University of California Press.

Rothman, S. M. 1994. *Living in the shadow of death: tuberculosis and the social experience of illness in American history.* New York: Basic Books.

Rothstein, W. 1972. *American physicians in the Nineteenth century: from sects to science.* Baltimore: Johns Hopkins University Press.

Royal College of Physicians of London. 1787. *Pharmacopoeiae Londinensis specimen alterum.* London: Royal College of Physicians.

Rubin, D. B. 1978. Bayesian inference for causal effects: the role of randomization. *Annals of Statistics* 9:45-58.

Rusnock, A. 1994. The weight of evidence and the burden of authority: case histories, medical statistics and smallpox inoculation. *Clio Medica* 29:289-315.

Russell, C. A. 1978. Edward Frankland and the cheapside chemists of Lancaster. *Annals of Science* 35:253-73.

Sadler, J. 1978. Ideologies of art and science in medicine: the transition from medical care to the application of technique in the British medical profession. In *The dynamics of science and technology: social values, technical norms and scientific criteria in the development of knowledge*, ed. W. Krohn, W. T. Layton, and P. Weingart, 117-215. Dordrecht: D. Reiedel.

Scadding, J. G. 1960. Clinical aspects of controlled trials in pulmonary tuberculosis. In *Controlled clinical trials. Papers delivered at the conference convened by The Council for International Organizations of Medical Sciences*, ed. A. Bradford Hill, 52-56. Oxford: Blackwell Scientific.

Schatz, A., E. Bugie, and S. A. Waksman. 1944. Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. *Proc. Soc. Exp. Biol. Med.* 55:66-69.

Schulz, K. F., I. Chalmers, R. J. Hayes, and D. G. Altman. 1995. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 273:408-12.

Seidenfeld, T. 1978. Direct inference and inverse inference. *Journal of Philosophy* 75:709-730.

---. 1979. *Philosophical problems of statistical inference: learning from R. A. Fisher.* Theory and Decision Library, v22. Dordrecht: Reidel.

---. 1981. Levi on the dogma of randomization in experiments. In *'Henry E. Kyburg, Jr. & Isaac Levi'*, ed. R. J. Bogdan, 263-91. Amsterdam: D. Reidel.

Sekulich, M. 1949. A key to the classification of pulmonary tuberculosis. *British Journal of Tuberculosis and Diseases of the Chest* XLIII:12-17.

Shamaskin, A., and L. C. Morris. 1947. Streptomycin in the treatment of pulmonary tuberculosis. *American Review of Tuberculosis* 56:540-51.

Shapin, S. 1982. History of science and its sociological reconstructions. *History of Science* 20:157-211.

---. 1988. Understanding the Merton thesis. *Isis* 79:594-605.

---. 1991. A scholar and a gentleman: the problematic identity of the scientific practitioner in early modern England. *History of Science* 29:279-327.

---. 1992. Discipline and bounding: the history and sociology of science as seen through the externalism-internalism debate. *History of Science* 30:333-69.

---. 1994. *A social history of truth: civility and science in Seventeenth century England.* Chicago: University of Chicago Press.

---. 1996. *The scientific revolution.* Chicago: University of Chicago Press.

Shapin, S., and S. Schaffer. 1985. *Leviathan and the air pump: Hobbes, Boyle and the experimental life.* Princeton: Princeton University Press.

Shaw, G. B. 1931. *Doctors' delusions, crude criminology and sham education.* London: Constable.

---. 1993. *The complete prefaces. Volume 1: 1889-1913.* London: Allen Lane. (edited by Dan Laurence and Daniel Leary)

Sheynin, O. B. 1966. Origin of the theory of errors. *Nature.* (27 August):1003-4.

---. 1974. On the prehistory of the theory of probability. *Archives of the History of the Exact Sciences* 12:97-141.

---. 1976. PS Laplace's work on probability. *Archives of the History of the Exact Sciences* 16:137-87.

---. 1978. C.F Gauss and the theory of errors. *Archives of the History of the Exact Sciences* 18:245-300.

---. 1980. On the history of the statistical method of biology. *Archives of the History of the Exact Sciences* 22:323-71.

---. 1982. On the history of medical statistics. *Archives of the History of the Exact Sciences* 26:241-86.

Shortt, S. E. D. 1983. Physicians, science and status: issues in the professionalization of Anglo-American medicine in the nineteenth century. *Medical History* 27:51-86.

Shryock, R. H. 1930. Public relations of the medical profession in Great Britain and the United States 1600-1870. *Annals of Medical History* 2:308-39.

---. 1937. *The development of modern medicine: an interpretation of the social and scientific factors involved.* New York: Alfred Knopf.

---. 1953. The interplay of social and internal factors in the history of modern medicine. *The Scientific Monthly* 76:221-30.

---. 1961. The history of quantification in medical science. *Isis* 52:215-37.

---. 1961. The history of quantification in medical science. In *Quantification: a history of the meaning of measurement in the natural and social sciences*, ed. H. Woolf, 31-63. Indianapolis: Bobbs-Merrill.

---. 1979. *The development of modern medicine.* Madison: University of Wisconsin Press.

Signorille, V. 1989. Buridan's Ass: the statistical rhetoric of science and the problem of equiprobability. In Simons, H. W. ed. *Rhetoric in the human sciences*. London: Sage. P69-88.

Silverman, W. A. 1980. *Retrolental fibroplasia - a modern parable.* New York: Grune and Stratton.

---. 1983. The public character of scientific medicine. *Perspectives in Biology and Medicine* 26:343-53.

---. 1991. Memories of the 1953-54 Oxygen Trial and its aftermath. The failure of success. *Controlled Clinical Trials* 12:355-58.

Silverman, W. A., and D. G. Altman. 1996. Patients' preferences and randomised trials. *Lancet.* 347:171-74.

Silverman, W. A., and I. Chalmers. 1992. Sir Austin Bradford Hill: an appreciation. *Controlled Clinical Trials* 13:100-105.

Slinn, J. 1984. *A history of May & Baker, 1834-1984.* Cambridge: Hobsons Ltd.

Smith, D. 1995. Henry Dale, skepticism, theory, luck, strategy and tactics. *Clio Medica* 33:195-98.

---, ed. 1997. *Nutrition in Britain: science, scientists and politics in the twentieth century.* London: Routledge.

Smith, D., and M. Nicolson. 1989. The 'Glasgow School' of Paton, Findlay and Cathcart: conservative thought in chemical pathology, nutrition and public health. *Social Studies of Science* 19:195-238.

---. 1993. Health and ignorance: past and present. In *Locating health: sociological and historical explorations*, ed. S. Platt, H. Thomas, S. Scott, and G. Williams, 221-44. Aldershot: Avebury.

Smith, F. B. 1985. Gullible's travails: tuberculosis and quackery 1890-1930. *Journal of Contemporary History* 20:733-56.

---. 1988. *The retreat of tuberculosis: 1850 - 1950.* London: Croom Helm.

---. 1993. Tuberculosis and bureaucracy. Bacille Calmette et Guérin: its troubled path to acceptance in Britain and Australia. *Medical Journal of Australia* 159:408-11.

Smith, R. 1980. Scientific thought and the boundary of insanity and criminal responsibility. *Psychological Medicine* 10:15-23.

---. 1981. *Trial by medicine: insanity and responsibility in Victorian trials.* Edinburgh: Edinburgh University Press.

---. 1983. Defining murder and madness: an introduction to medico-legal belief in the case of Mary Ann Bough. *Knowledge and Society* 4:173-225.

Solomon, S. M., and E. J. Hackett. 1996. Setting boundaries between science and law: lessons from Daubert v Merrell Dow Pharmaceuticals, Inc. *Science, Technology and Human Values* 21:131-56.

Sourkes, T. L. 1993. John Simon, Robert Lowe, and the origin of state supported biomedical research in Nineteenth-Century England. *Journal of the History of Medicine and Allied Sciences* 48:436-53.

Spence, J. C. 1933. Clinical tests of the antirachitic activity of calciferol. *Lancet.* ii:911-15.

Spiegelhalter, D. J., L. S. Freedman, and M. K. Parmar. 1994. Bayesian approaches to randomized trials. *Journal of the Royal Statistical Society A* 157:357-416.

Spilker, B. 1992 *A guide to clinical trials*. New York: Raven Press.

Steiner, W. R. 1939. Some distinguished American medical students of Pierre-Charles-Alexander Louis of Paris. *Bulletin of the History of Medicine* 7:783-93.

Stepan, N. 1978. The interplay between socio-economic factors and medical science: Yellow Fever research in Cuba and the United States. *Social Studies of Science* 8:397-429.

Stigler, S. M. 1980. R.H. Smith: a Victorian interested in robustness. *Biometrika* 67:217-21.

---. 1982. Thomas Bayes Bayesian inference. *Journal of the Royal Statistical Society A* 145:250-58.

---. 1983. Who discovered Bayes theorem? *American Statistician* 37:296-325.

---. 1986. *The history of statistics: the measurement of uncertainty before 1900.* Cambridge, Mass.: Harvard University Press.

---. 1987. The measurement of uncertainty in nineteenth century social science. In *The probabilistic revolution. Volume 1: ideas in history*, ed. L. Krüger, L. J. Daston, and M. Heidelberger, 287-92. Cambridge: MIT Press.

Stobie, W. 1922. Treatment of tuberculosis by tuberculin. *BMJ* ii:473.

Stove, D. C. 1973. *Probability and Hume's inductive skepticism.* Oxford: OUP.

Struhl, P. R. 1975. Peirce's defense of scientific method. *Journal of the History of Philosophy* 13:481-90.

Sturdy, S. 1992. The political economy of scientific medicine: science, education and the transformation of medical practice in Sheffield, 1890 - 1922. *Medical History* 36:125-59.

Sullivan, M. D. 1993. Placebo controls and epistemic control in orthodox medicine. *Journal of Medicine and Philosophy* 18:213-31.

Sullivan, R. B. 1994. Sanguine practices: a historical and historiographic reconstruction of heroic therapy in the age of Rush. *Bulletin of the History of Medicine* 68:211-34.

Suppes, P. 1983. Arguments for randomization. In *Proceedings of the 1982 biennial meeting of the Philosophy of Science Association*, ed. P. D. Asquith and T. Nickles, 464-75. Vol. Vol 2. East Lansing, Mi.: Philosophy of Science Association.

---. 1984. *Probabilistic metaphysics.* Oxford: Blackwell.

Sutherland, H. G. 1912. The work of the St Marylebone dispensary for the prevention of consumption. *British Journal of Tuberculosis and Diseases of the Chest* 6:12-18.

Sutherland, I. 1976. Recent studies in the epidemiology of tuberculosis. *Advances in Tuberculosis Research* 19:1-63.

Sutliffe, W. D. 1936. Adequate tests of curative therapy in man. *Annals of Internal Medicine* 10:89-96.

Swann, J. P. 1985. Arthur Tatum, Parke-Davis, and the discovery of Mapharsan as an antisyphilitic agent. *Journal of the History of Medicine and Allied Sciences* 40:167-87.

---. 1986. Insulin: a case study in the emergence of collaborative pharmaceutical research. *Pharmacy in History* 28:3-13, 65-74.

---. 1988. *Academic scientists and the pharmaceutical industry: cooperative research in Twentieth-Century America.* Baltimore: Johns Hopkins University Press.

---. 1989. Biomedical research and government support: the case of drug development. *Pharm Hist (US)* 31:103-16.

---. 1995. The evolution of the American pharmaceutical industry. *Pharm Hist (US)* 37:76-86.

Swijtink, Z. 1989. The objectification of observation: measurement and statistical methods in the nineteenth century. In Kruger, L., L. Daston, and M Heidelberger eds. *The probabilistic revolution. Volume 1*, Cambridge, Mass.: MIT Press. 261-285

Symposium on Clinical Trials. 1962. *Clinical trials: report of a symposium organised by the Department of Pharmaceutical Sciences of the Pharmaceutical Society of Great Britain at the School of Pharmacy, University of London, April 5, 1962.* London: Pharmaceutical Press.

Szreter, S. 1988. The importance of social intervention in Britain's mortality decline c1850-1914. *Social History of Medicine* 1:1-37.

---. 1997 *Fertility, class and gender in Britain, 1860-1940*. Cambridge: Cambridge University Press.

Tansey E. M. 1994. The funding of medical research before the Medical Research Council. *Journal of the Royal Society of Medicine* 87:546-48.

Tansey, E. M., and R. C. E. Milligan. 1990. The early history of the Wellcome Research Laboratories. In *Pill peddlers: essays in the history of the pharmaceutical industry*, ed. J. Liebenau, G. J. Higley, and E. Stroud, 91-106. Madison, Wisc.: American Institute for the History of Pharmacy.

Temkin, O. 1955. Therapeutic trends and the treatment of syphilis before 1900. *Bulletin of the History of Medicine* 29:309-16.

Thomas, D. P. 1969. Experiment versus authority: James Lind and Benjamin Rush. *New England Journal of Medicine* 281:932-34.

Throm, C. 1995. *Das diphtherie-serum: ein neuses therapieprinzip, seine entwicklung und markteinfuhrung*. Stuttgart: Wissenschaftliche Verlageseinscheft.

Thompson, A. T. 1833. *London Dispensatory.* 7th ed. London: Longman Rees.

Thomson, H. C. 1935. *The story of the Middlesex Hospital Medical School*. London: Murray.

Tishler, M. 1949. Production and isolation of streptomycin. In *Streptomycin: nature and practical applications*, ed. S. Waksman, 32-54. London: Balliere, Tindall & Cox.

Todhunter, I. 1865. *A history of the mathematical theory of probability from the time of Pascal to that of Laplace.* London: Macmillan and Co.]

Topley, W. W. C. 1940. *Authority, observation and experiment in medicine.* Cambridge: Cambridge University Press.

Tring, F. C. 1977. Thomas Holloway and his patent pills. *Pharm Hist (GB)* 7:6-7.

Tröhler, U. 1981. Towards clinical investigation on a numerical basis: James Lind at Haslar Hospital 1758-1783. In *XXVII Congreso Internacional de Historia de la Medicina: Actas*, 414-19. Barcelona: Academia de Ciencies Mediques de Catalunya i Balears.

---. 1988. 'To improve the evidence of medicine': arithmetical observations in clinical medicine in the Eighteenth and early Nineteenth centuries'. *History and Philosophy of the Life Sciences* 10:31-40.

Tuberculosis Association. 1933. Misapplication of artificial pneumothorax. *Lancet* i:1124-6.

Tweedale, G. 1990. *At the sign of the plough. Allen and Hanburys and the British pharmaceutical industry 1715-1990.* London: John Murray.

Underwood, E. A. 1951. The history of the quantitative approach in medicine. *British Medical Bulletin* 7:265-74.

Urbach, P. 1985. Randomization and the design of experiments. *Philosophy of Science* 52:256-73.

---. 1987. Clinical trial and random error. *New Scientist* 116:52-55.

---. 1993. The value of randomization and control in clinical trials. *Statistics in Medicine* 12:1421-31.

---. 1994. Reply to David Papineau. *British Journal of the Philosophy of Science* 45:712-15.

Urdang, G. 1942. The first century of the Pharmaceutical Society of Great Britain. *Journal of the American Pharmacy Association* 3:420-26.

---. 1942. The mystery about the first English (London) Pharmacopoeia (1618). *Bulletin of the Institute of the History of Medicine* 12:304-13.

---. 1944. Retail pharmacy as the nucleus of the pharmaceutical industry. In *Essays in the history of medicine, presented to Professor Arturo Castiglioni*, 325-46. Baltimore: Johns Hopkins University Press.

---. 1944. How chemicals entered the official pharmacopoeias. *Pharm Arch* 15:17-37.

---. 1951. The antibiotics and pharmacy. *Journal of the History of Medicine* 6:388-405.

Usadi, M. M. 1993. Baconian science, nineteenth century alternative practice, and clinical medicine. *Pharos* 56:25-29.

Van Brakel, J. 1976. Some remarks on the prehistory of the concept of statistical probability. *Archives of the History of the Exact Sciences* 16:95-112.

Vandenbroucke, J. P. 1987. A short note on the history of the randomized controlled trial. *Journal of Chronic Diseases* 40:985-87.

van Leeuwen, H. G. 1963. *The problem of certainty in English thought 1630-1690.* The Hague: Nijhoff.

Van Zwanenberg, D. 1983. The training and career of those apprenticed to apothecaries in Suffolk, 1815-1858. *Medical History* 27:139-50.

Varcoe, I. M. 1970. Scientists, government and organised research: the early history of the DSIR, 1914-1916. *Minerva* 8:192-216.

Vaughan, P. 1959 *Doctors Commons: a short history of the British Medical Association* London: Heinemann.

Venn, J. 1866 *The logic of chance*. London: Macmillan

Veterans Administration. 1947. The effect of streptomycin on tuberculosis: preliminary report of a cooperative study of 223 patients by the Army Navy and Veterans Administration. *American Review of Tuberculosis* 56:485-507.

Veterans Administration, and W. B. Tucker. 1949. Evaluation of streptomycin regimens in the treatment of tuberculosis. *American Review of Tuberculosis* 60:715-54.

Wainwright, M. 1991. Streptomycin: discovery and resultant controversy. *History and Philosophy of the Life Sciences* 13:97-124.

Waksman, S. A., ed. 1949. *Streptomycin: nature and practical applications.* London: Bailliere, Tindall & Cox.

---. 1951. Streptomycin: isolation, properties and utilization. *Journal of the History of Medicine* 6:318-47.

---. 1954. *My life with microbes*. New York: Simon and Schuster.

---. 1965. *The conquest of tuberculosis.* London: Robert Hale.

Walker, H.M. 1931. *Studies in the history of statistical method. With special reference to certain educational problems*. Baltimore: Williams and Wilkins.

Ward, P. S. 1981. The American reception of Salvarsan. *Journal of the History of Medicine and Allied Sciences* January:44-62.

Warner, J. H. 1980. Therapeutic explanation and the Edinburgh blood-letting controversy: two perspectives on the medical meaning of science in the mid-nineteenth century. *Medical History* 24:241-58.

---. 1980. Physiological theory and therapeutic explanation in the 1860s: the British debate on the medical use of alcohol. *Bulletin of the History of Medicine* 54:235-57.

---. 1985**a**. Science in medicine. *Osiris 2nd Series* 1:37-58.

---. 1985**b**. The selective transport of medical knowledge: antebellum American physicians and Parisian medical therapeutics. *Bulletin of the History of Medicine* 59:213-31.

---. 1986. *The therapeutic perspective. Medical practice, knowledge and identity in America 1820 - 1885.* Cambridge, Mass: Harvard University Press.

---. 1987. Medical sectarianism, therapeutic conflict, and the shaping of orthodox professional identity in antebellum American medicine. In *Medical fringe and medical orthodoxy*, ed. W. F. Bynum and R. Porter, 234-60. London: Croom Helm.

---. 1991. Science, healing and the physician's identity: a problem of professional character in nineteenth-century America. In *Essays in the history of therapeutics*, ed. W. F. Bynum and V. Nutton, 65-88.(Clio Medica Vol 22). Amsterdam: Editions Rodolphi.

---. 1991. The idea of science in English medicine: the "decline of science" and the rhetoric of reform, 1815-1845. In *British medicine in an age of reform*, ed. R. French and A. Wear, 136-64. London: Routledge.

---. 1991. Ideals of science and their discontents in late nineteenth century American medicine. *Isis* 82:454-78.

---. 1992. The fall and rise of professional mystery. Epistemology, authority and the emergence of laboratory medicine in nineteenth century America. In *The laboratory revolution in medicine*, ed. A. Cunningham and P. Williams, 110-41. Cambridge: Cambridge University Press.

---. 1994. The history of science and the sciences of medicine. *Osiris 2nd Series* 10:164-93.

Watkin Edwards, T. A. 1950. Some aspects of the use of streptomycin and PAS. *Tubercle* 41:250-55.

Wayne, E. J. 1933. Clinical observations on two pure glycosides of digitalis, digoxin and digitalinum verum. *Clinical Science* 1:63-76.

Weatherall, M. 1990. *In search of a cure: a history of pharmaceutical discovery.* Oxford: OUP.

---. 1996. Drug therapies. In *Companion encyclopaedia of the history of medicine*, ed. W. F. Bynum and R. Porter, 915-38. Vol. Vol 2. London: Routledge.

Weatherhead, E. 1928. Efficacy of tuberculin. *BMJ* i:611-12.

Webster, C. 1988. *The health services since the war.* Vol. 1. London: HMSO.

---. 1998 *The National Health Service: a political history*. Oxford: Oxford University Press

Weindling, P. 1992. From medical research to clinical practice: serum therapy for diphtheria in the 1890s. In *Medical innovations in historical perspective*, ed. J. V. Pickstone, 72-83. Basingstoke: Macmillan.

---. 1992. Scientific élites and laboratory organisation in fin de siècle Paris and Berlin. In *The laboratory revolution in medicine*, ed. A. Cunningham and P. Williams, 170-88. Cambridge: Cambridge University Press.

Weisz, G. 1978-9. The politics of medical professionalization in France, 1845-1848. *Journal of Social History* 12:3-30.

---. 1995. *The medical mandarins: the French academy of medicine in the nineteenth century and early twentieth centuries.* Oxford: Oxford University Press.

Wells, G. P. 1978. Lancelot Thomas Hogben, 9 December 1895 - 22 August 1975. *Biographical Memoirs of the Royal Society* 24:183-221.

Werskey, G. 1988. *The visible college: a collective biography of British scientists and socialists of the 1930s*. London: Free Association Books.

Whitby, L. E. H. 1938. Treatment with 2(p-aminobenzenesulphonamidopyridine). *Lancet.* ii:1210-12.

White, M. J. 1988. The unimportance of being random. *Synthese* 76:171-78.

Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics* 1:80-83.

Williams, C. J. B. 1887. *Pulmonary consumption. Its etiology, pathology and treatment, with an analysis of 1000 cases to exemplify its duration and modes of arrest.* London: Longmans Green and Co.

Williams, G. R. 1975. *The age of agony: the art of healing c1700-1800.* London: Constable.

Wilson, L. G. 1990. The historical decline of tuberculosis in Europe and America: its causes and significance. *Journal of the History of Medicine and Allied Sciences* 45:366-96.

Winter, A. 1991. Ethereal epidemic: mesmerism and the introduction of inhalation anaesthesia to early Victorian London. *Social History of Medicine* 4:1-27.

---. 1994. Mesmerism and popular culture in early Victorian England. *History of Science* xxxii:317-43.

Wise, M. N. 1987. How do sums count? On the cultural origins of statistical causality. In *The probabilistic revolution. Volume 1: ideas in history*, ed. L. Krüger, L. J. Daston, and M. Heidelberger, 395-425. Cambridge: MIT Press.

---, ed. 1995. *The values of precision.* Princeton: Princeton University Press.

Witkop, B. 1982. Paul Ehrlich: his ideas and his legacy. In *Science, technology and society in the time of Alfred Nobel*, 146-66. Oxford: Pergamon Press.

Wittgenstein, L. 1969. *On certainty*. Oxford: Blackwell.

Witts, L. J. 1960. The ethics of controlled clinical trials. In *Controlled clinical trials. Papers delivered at the conference convened by The Council for International Organizations of Medical Sciences*, ed. A. Bradford Hill, 8-13. Oxford: Blackwell Scientific.

---. 1964. *Medical surveys and clinical trials.* 2nd ed. Oxford: Oxford University Press.

---. 1972. The first whole-time professors of medicine in the University of London. *History of Medicine* 4:13-17.

Wood, M. T. 1979. Random Assignment to Treatment Groups: A Strategy for Judicial Research. *Criminology* 17:230-41.

Woodward, J., and D. Richards, eds. 1977. *Health care and popular medicine in nineteenth-century England: essays in the social history of medicine.* London: Croom Helm.

Woolf, H. 1961. *Quantification - a history of the meaning of measurement in the natural and social sciences.* Indianapolis: Bobbs-Merrill.

Worboys, M. 1992**a**. The sanatorium treatment for consumption in Britain 1890-1914. Chap. 3 In *Medical innovations in historical perspective*. Science Technology and Medicine in Modern History, ed. J. V. Pickstone, ed. J. V. Pickstone, 47-71. Basingstoke: Macmillan.

---. 1992**b**. Vaccine therapy and laboratory medicine in Edwardian Britain. In *Medical innovations in historical perspective*. Science Technology and Medicine in Modern History, ed. J. V. Pickstone, ed. J. V. Pickstone, 84-103. Basingstoke: Macmillan.

Wright, A. 1904. *A short treatise on anti-typhoid inoculation*. London: Constable.

Wright, P., and A. Treacher, eds. 1982. *The problem of medical knowledge*. Edinburgh: Edinburgh University Press.

Wrighton, R. 1953. The theoretical basis of the therapeutic trial. *Acta Genetica et Statistica Medica* 4:312-43.

---. 1954. Auxiliary randomization. *Acta Genetica et Statistica Medica* 5:39-63.

---. 1955. Statistical experimentation in therapeutics. *British Journal of Preventive and Social Medicine* 9:135-39.

---. 1957. Random selection and the therapeutic trial. *British Journal of Preventive and Social Medicine* 11:61-62.

Wynn, H. P. 1982. Controlled versus random experimentation. *The Statistician* 31:237-44.

Yankauer, A. 1996. Pierre Charles-Alexandre Louis: the impact of a clinical trial. *Pharos* 59:15-19.

Yates, F., and K. Mather. 1972. Ronald Aylmer Fisher. In *Collected papers of RA Fisher*, ed. J. H. Bennett, 23-52. Vol. 1. Melbourne: University of Melbourne Press.

Yearley, S. 1982. The relationship between epistemological and sociological cognitive interests: some ambiguities underlying the use of interest theory in the study of scientific knowledge. *Studies in the Philosophy and History of Science* 13:353-88.

Yoshioka, A. 1995. *The social roots of Lancelot Hogben's Mathematics for the Million*. Paper presented to the British Society for the History of Mathematics, 3-5 April 1995, Gregynog, Wales.

Young, J. H. 1960. Patent medicines, an early example of competitive marketing. *Journal of Economic History* 20:648-56.

---. 1961. *The toadstool millionaires: a social history of patent medicines in America before regulation.* Princeton, NJ: Princeton University Press.

Young, R. A. 1956. Clinical aspects of tuberculosis and diseases of the chest 1906-1956. Then and now. *British Journal of Tuberculosis and Diseases of the Chest* 50:23-32.

Young, R. M. 1976. Science is social relations. *Radical Science Journal* 5:65-129.

---. 1979. Why are figures so significant? The role and the critique of quantification. In *Demystifying social statistics*, ed. I. Miles and J. Evans, 63-74. London: Pluto Press.

Youngson, A. J. 1979. *The scientific revolution in Victorian medicine.* London: Croom Helm.