Review

Research on Balint groups: A literature review

Kaatje Van Roy*, Stijn Vanheule, Ruth Inslegers

Department of Psychoanalysis and Clinical Consulting, Ghent University, Ghent, Belgium

**ABSTRACT**

**Objective:** As the scientific literature on Balint groups (BGs) is scattered, this paper provides an overview of the literature on BGs published in peer-reviewed journals. Study characteristics are analyzed and the principal research topics are discussed.

**Methods:** ‘Web of Science’ and ‘Pubmed’ databases were searched and all English-language studies on BGs (empirical and non-empirical) were included.

**Results:** Of the 94 articles included, 35 are empirical studies adopting a qualitative, quantitative or mixed methodology. The research topics that emerged include outcome, characteristics of BG participants, themes addressed in BGs, BG processes, leadership and BG evaluations. The remaining articles were classified as historical articles, reports and reflective articles, for which the main discussion themes are presented.

**Conclusion:** Research on BGs proves to be diverse, scarce and often methodologically weak. However, indications of the value of BG work were found. Therefore, further research is strongly indicated.

**Practice Implications:** Points of interest that could to be further considered by BG workers and researchers are for instance long-term BG participation and ‘modified Balint groups’. Recommendations for future research on BGs are provided.

© 2015 Elsevier Ireland Ltd. All rights reserved.

**Contents**

1. Introduction .................................................................................................................. 000
2. Methods ....................................................................................................................... 000
   2.1. Search strategy ........................................................................................................ 000
   2.2. Data analysis .......................................................................................................... 000
3. Results ........................................................................................................................ 000
   3.1. Methodology of empirical articles ........................................................................ 000
   3.2. Variability in Balint group setting .......................................................................... 000
   3.3. Article topics .......................................................................................................... 000
      3.3.1. Outcome – effects of Balint group participation ............................................. 000
      3.3.2. Characteristics of Balint group participants .................................................. 000
      3.3.3. Themes addressed in Balint groups .............................................................. 000
      3.3.4. Balint group processes .................................................................................. 000
      3.3.5. Leadership in Balint groups ......................................................................... 000
      3.3.6. Evaluation of Balint groups .......................................................................... 000
      3.3.7. Historical-geographical articles .................................................................... 000
      3.3.8. Reports – anecdotal articles .......................................................................... 000
      3.3.9. Reflective articles ......................................................................................... 000
      3.3.10. Balint group observation as research data .................................................. 000
4. Discussion and conclusions ......................................................................................... 000
   4.1. Discussion .............................................................................................................. 000

* Corresponding author at: Tel.: +32 2646365.
E-mail addresses: Kaatje.Vanroy@UGent.be, kaatvanroy@yahoo.com (K. Van Roy), Stijn.Vanheule@UGent.be (S. Vanheule), Ruth.Inslegers@UGent.be (R. Inslegers).

http://dx.doi.org/10.1016/j.pec.2015.01.014
0738-3991/© 2015 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

In the 1950s, psychoanalyst Michael Balint introduced seminars for general practitioners (GPs) that were later called ‘Balint groups’ (BGs) [1,2]. These groups were set up in London and spread worldwide, though on a limited scale [3]. In Balint groups, GPs and/or other (para)medical professionals explore difficult interactions with patients through case presentations and discussions. Generally, BGs comprise six to twelve members and one or two leaders. Groups meet on a weekly to monthly basis over several years. In BG meetings, participants present cases that are subsequently commented on by the group members, expressing their thoughts, ideas and emotions. This process can help participants broaden their perspective on the initial difficulty they experienced, and can influence their overall perception of their practice and interactions with patients [4–7].

Activities of BG societies (such as registration of BGs, organization of workshops and conferences) [8] and associated literature demonstrate that BGs are still very much alive. Yet, research on the process and outcome of BGs is relatively scarce and sporadic, and therefore not always easy to find. This might explain why the short introductory literature reviews in some articles mention incomplete and sometimes contradictory findings. Therefore, the present study aims to gather the peer-reviewed literature on BGs in one article and to provide an up-to-date perspective on BG research.

2. Methods

2.1. Search strategy

Using ‘Balint group’ as a key word, we searched the ‘Web of Science’ and ‘Pubmed’ databases for publications up until March 2014. No restriction was set for year of publication. Abstracts were reviewed and all articles addressing BGs as a subject were included. Duplicates, non-English-language articles, meeting abstracts, book reviews, letters and editorials were excluded, as were articles that mentioned BGs only briefly in the context of another research topic. References from each article were checked for further peer-reviewed studies. In order to get a general overview, no further restrictions were imposed.

2.2. Data analysis

After reading through all included articles, we composed a list of variables to be completed for each of the articles. Three broad categories of variables were used: (1) general article information (year of publication, country where research was done); (2) information provided about the BG (the author’s relationship to BGs, length and frequency of sessions, group composition, information on leaders, specifications on terminology used, description or definition of BG); (3) type of paper, i.e., ‘empirical articles’ (using a quantitative, qualitative, or mixed quantitative-qualitative methodology), ‘historical-geographical articles’, ‘reports—anecdotal articles’ (with or without case examples), or ‘reflective articles’. Each article was critically appraised. Given our purpose of mapping out the range of articles on BGs, all articles were retained. Articles using qualitative data were scored according to the NICE methodology checklist for qualitative studies [9]. These studies were rated independently by two researchers and disagreements were discussed. Articles failing to meet standards of quality for qualitative research were classified as ‘reports—anecdotal articles’. For the quantitative articles, potential remarks with regard to the interpretation of the results are provided below. As a number of the remaining papers (n = 59) provided critical reflections, rich reports on personal experiences, or instructive information about the context of BGs that may be of interest for future research, these papers were also included in this review. Finally, for each article, the topic, topic variables and results or findings were summarized. As the overall body of empirical literature was too diverse to make any meaningful quantitative synthesis, we chose to qualitatively synthesize the article topics and to present the results of the articles in a schematic way.

3. Results

In Fig. 1, the numbers of articles included and excluded throughout the search process are presented. The database search yielded 362 articles. Screening the abstracts led to the exclusion of 32 articles that were not related to BGs. After excluding duplicates (n = 60), non-English-language articles (n = 149), meeting abstracts, book reviews, letters and editorials (n = 24), the number of included articles scaled down to 97. Hand searches and bibliographic reviews yielded an additional 22 papers. Finally, 25 papers were excluded since they only marginally mentioned BGs. This resulted in a total of 94 articles included in this study.

3.1. Methodology of empirical articles

Of the 94 included articles, 35 (37%) were empirical papers. Among these articles, 21 used a quantitative methodology, 10 used a qualitative methodology, and four applied a mixed quantitative-qualitative methodology. Almost all quantitative studies made use of self-report questionnaires, measuring for instance work satisfaction, burnout, attitudes, empathy, personality, psychosocial self-efficacy and BG evaluation (see Table 1). In the qualitative studies, researchers mainly used semi-structured interviews, field notes, video-taped sessions, audio-taped sessions (with or without transcriptions) and open questionnaires (see Table 2).

3.2. Variability in Balint group setting

Although BGs were initially set up for GPs, some papers address BGs for other professionals. These include BGs for nurses [10–13], ‘industrial physicians’ [14], specialists [15], physiotherapists [16] and community health workers [17]. A relatively high number of papers report on BGs for medical residents [5,18–27], medical fellows [28] or medical students [29–34], working in family medicine [5,21,23–27,29], obstetrics and gynecology [18,22], psychiatry [4,31,32] or oncology [19,28] departments. In some countries (mostly US) a limited time of BG participation (often 6 months) is mandatory for residents [4,5,21–23,35,36]. Some BGs are mixed, welcoming professionals from various backgrounds, such as GPs, medical specialists and/or counselors [4,37,38]. Generally, BG participants do not cooperate with each other in their everyday work, yet some BGs are organized for professionals
working in the same unit [13,38–41]. Reports on other types of ‘modified Balint groups’ indicate the use of different proceedings, such as case preparation [19], presenting cases in rotation [4,42], taking the most recent consultation as a case [17], working on questions [10], position related difficulties [14] and professional role conflict [28], giving homework assignments [43], combining meetings with theoretical teaching [4,44], rotating leadership [42], or modifying the BG according to a mindfulness technique [45]. Some modified BGs have different focuses such as a family systems approach [37], cognitive therapy [43] or an additional focus on diagnostics [38,40]. Often these modified groups have different names such as ‘Balint-style group’, ‘Balint clinical reflection group’ or ‘Balint-like group’. Generally, the number of participants in a BG is between 6 and 12, with extremes of 4 [34] to 15 [19,24,46] and 17 participants [47]. Meeting frequency is often once per week or once every fortnight, sometimes once per month. Meetings generally last between one and two hours over a period of one or two years. However, the period of group meetings is variable, ranging from approximately 6 to 12 weeks [4,32,34,43,48] up to 12 [38] and 17 years [49].

3.3. Article topics

In this section, we briefly discuss the main findings of the empirical articles and the chief topics addressed in the other papers. Specific information on the study designs is provided in Tables 1 and 2. Due to space limitations, only summaries of the findings are presented; for more information we refer to the articles themselves.

3.3.1. Outcome – effects of Balint group participation

Several quantitative and qualitative studies reported on outcome or effects of BG participation. Quantitative studies categorized under this topic have a research design of minimally two assessment moments. Results on item level are not presented here. The following outcome variables of BG participation were addressed:

Psychosocial self-efficacy [11,12,18,22,27,50]. All six articles addressing this topic made use of the Psychological Medicine Inventory (PMI). Three studies [11,12,27] found an increase in psychosocial self-efficacy while the other three [18,22,50] reported no significant increase. Rabinowitz [11,12] reported significant changes only after long-term participation (i.e., 10 to 12 months) but not after short-term participation (i.e., 6 months).

Burnout/satisfaction [12,19,22,45]. Using the Maslach Burnout Inventory (MBI), one study [22] found no significant effect on burnout and a second [19] failed to report on statistical tests and therefore remained inconclusive. A third study [12] using two other burnout questionnaires found a significant decrease in burnout levels after 10 months of participation, but not after 6 months. A fourth study [45] did not find any significant effect of BG participation on subjective satisfaction.

Attitudes [18,22,28,43,45,51,52]. Seven articles made use of various questionnaires focusing on different aspects of participants’ attitudes. Brock and Stock [51] presented leaders’ perceptions of attitudes or skills that are attainable through BG seminars; Dokter, Duivenvoorden and Verhage [52] reported individual changes in perception of patients but did not report on statistical tests; Adams et al. [18] found no significant effect of BG participation on professionalism; Ghetti et al. [22] found unchanged scores in participants’ empathy; Sekeres et al. [28] reported no significant effect on participant’s overall attitudes (only in domain “view of oneself as a physician”) and Abeni et al. [45] reported a general maturation in participants’ defense mechanisms; finally, Hartmann’s [43] pilot study of participants’ attitudes towards somatizing patients showed no significant overall change.

Specific expertise/knowledge [44,47,50]. Amiel et al. [47] found no effect of BG participation on breaking bad news; Rabin et al. [50] reported increased self-efficacy cognitions related to the treatment of drug addicts, although significant at final assessment only (30 months); finally, a third study [44] was inconclusive on the effect on knowledge of pharmacotherapy and psychotherapy of anxiety.

Besides quantitative measures of pre-defined outcome variables, a number of qualitative studies investigated the effects of BG participation. One pilot study [53] outlined two criteria for defining the type of change that BG participation might induce: ‘knowledge of one’s own limits’ and ‘minimal interference of one’s...
<table>
<thead>
<tr>
<th>Article</th>
<th>Instruments</th>
<th>Assessment moments</th>
<th>Participants (+control) modified BG</th>
<th>Time in BG</th>
<th>Topic + variables</th>
<th>Results (remarks)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quantitative papers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abeni et al. [45]</td>
<td>REM-71 + SAT-P + GCQ</td>
<td>2</td>
<td>30 (8 Caregivers + 10 physicians + 12 nurses) modified BG</td>
<td>30 h (7–1 hr–30 sessions)</td>
<td>Outcome: defense mechanisms, subjective satisfaction; process: group climate: ↑ engagement ↓ conflict (only in group of caregivers)</td>
<td>No effect on psychosocial self-efficacy; no effect on professionalism</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adams et al. [18]</td>
<td>PMI + Musikc 360-degree evaluation (only professionalism items)</td>
<td>2</td>
<td>7 Residents (+6 control)</td>
<td>7 (1x/2wk–7 sessions)</td>
<td>Outcome: psychological medicine skills; professionalism</td>
<td>No effect on BBN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amiel et al. [47]</td>
<td>2 questionnaires (7 + 3 + 4 items) evaluating OSCE (8 cases)</td>
<td>2</td>
<td>17 GPs (+17 control)</td>
<td>21 h (7–1.5h–14 sessions)</td>
<td>Outcome: breaking bad news (BBN)</td>
<td>No effect on BBN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bar-Sela et al. [19]</td>
<td>MSI + expectations questionnaire</td>
<td>2</td>
<td>15 Residents (8 Jr.–7 sr. residents comparison) modified BG</td>
<td>18 h (1x/month–1.5h–1 yr)</td>
<td>Outcome: burnout; evaluation: topics + group contribution</td>
<td>No effect on expected group contribution</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cataldo et al. [21]</td>
<td>JSPE + Work Satisfaction Survey (3 items)</td>
<td>2</td>
<td>4 GPs (+4 control) 6 month 104 h (1x/wk–2h–1 yr)</td>
<td>Part. characteristics: empathy; work satisfaction</td>
<td>No effect in empathy; no ↓ in work satisfaction between ‘attendees’ and ‘non-attendees’</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dokter et al. [52]</td>
<td>Questionnaire (own design): demographic–Balint characteristics–personality characteristics–Leary’s interaction rose</td>
<td>3</td>
<td>12 GPs (to–14 and 8 GPs) (+22 control)</td>
<td>78 h (1x/2wk–1.5h–2 yrs)</td>
<td>Outcome: attitude, patient perception; part. characteristics: attitude, personality, patient perception</td>
<td>No effect on BBN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ghetti et al. [22]</td>
<td>MSI + PMI + JSPE</td>
<td>2</td>
<td>17 Residents</td>
<td>12 h (1x/month–1h–1 yr)</td>
<td>Outcome: burnout; psychological medicine skills; empathy</td>
<td>No effect on burnout; no effect on psychosocial self-efficacy; no effect on empathy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hartmann et al. [43]</td>
<td>Attitude questionnaire (own design; 10 items)</td>
<td>2</td>
<td>4 GPs (+4 control) modified BG</td>
<td>19.5 h (1x/wk–1.5h–13 sessions)</td>
<td>Outcome: attitudes</td>
<td>Results only on item level</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson et al. [23]</td>
<td>MBI + Rokeach score + WEP + IE + FRR + POI</td>
<td>2</td>
<td>132 Residents (+74 control) + 6 month mandatory part.</td>
<td>104 h (1x/week–2h–1 yr)</td>
<td>Part. characteristics: personality</td>
<td>Proportion ‘non-attendees’ = 35% ‘non-attendees’ less intuitive than ‘attendees’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joukamaa et al. [56]</td>
<td>Patients: SCL-25; GPs: assessment scale of patient’s mental health</td>
<td>1</td>
<td>10 GPs (+2 control)</td>
<td>1</td>
<td>Part. characteristics: workload; control; satisfaction; quality of work; cooperation; training; health; attitudes to psychosomatic patients</td>
<td>Experienced BG part.: overall higher scores (except for ‘workload’)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kjeldmand et al. [55]</td>
<td>Questionnaire (own design; 49 items)</td>
<td>1</td>
<td>20 GPs (5 BG &lt; 1.5 yrs; 12 BG &gt; 1.5 yrs) (+21 control)</td>
<td>7 part. from diff groups</td>
<td>Part. characteristics: workload; control; satisfaction; quality of work; cooperation; training; health; attitudes to psychosomatic patients</td>
<td>Evaluation BG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parker and Leggett [31]</td>
<td>2 evaluation questionnaires (own design; 5 + 6 items)</td>
<td>2</td>
<td>20 students</td>
<td>64 h (1x/wk–1 hr–6h–8 wks)</td>
<td>Outcome: professional medicine skills; drug-treatment self-efficacy</td>
<td>Sessions rated positive; contribution of BG participation to educational needs rated medium (only descriptive statistics)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rabin et al. [50]</td>
<td>PMI + questionnaire on drug treatment self-efficacy (own design; 1 items)</td>
<td>4</td>
<td>22 Physicians</td>
<td>144 h (1x/2 wks–2h–2.5 yrs)</td>
<td>Outcome: psychological medicine skills; drug-treatment self-efficacy</td>
<td>No effect on psychosocial self-efficacy; ↑ in self-efficacy cognitions related to the treatment of drug addicts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rabinowitz et al. [11]</td>
<td>PMI + part. listing important mental health topics</td>
<td>3</td>
<td>13 Nurses</td>
<td>48 h (1x/2 wks–2h–1 yr)</td>
<td>Outcome: psychological medicine skills; psychosocial repertoire</td>
<td>↑ in psychosocial self-efficacy (long-term but not short-term); no effect on psychosocial repertoire</td>
</tr>
</tbody>
</table>
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Article</th>
<th>Instruments</th>
<th>Assessment moments</th>
<th>Participants (+control) modified BG</th>
<th>Time in BG</th>
<th>Topic + variables</th>
<th>Results (remarks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rabinowitz et al. [12]</td>
<td>PMI + burnout questionnaire (Shirom + Melamed)</td>
<td>3</td>
<td>10 Nurses</td>
<td>40 h (1x/2wks–2h–10 months)</td>
<td>Outcome: psychological medicine skills; burnout</td>
<td>[ In psychosocial self-efficacy; ]</td>
</tr>
<tr>
<td>Sekeres et al. [28]</td>
<td>Attitudes questionnaire (own design; 32 items)+ evaluation questionnaire (own design; 21 items)</td>
<td>3</td>
<td>28 Fellows modified BG</td>
<td>18 h (1x/2wks–1.5–2h–6 months (=10 sessions))</td>
<td>Outcome: attitudes; evaluation BG</td>
<td>Outcome: no effect on attitudes (only in domain “view of oneself as a physician”); evaluation: safe group, decompress, social activity (Results mainly comparing GPs &amp; specialists in primary care)</td>
</tr>
<tr>
<td>Stojanovic-Spehar et al. [44]</td>
<td>Questionnaire on knowledge of pharmacotherapy (own design; 5 items) and use of psychotherapy (own design; 3 items)</td>
<td>2</td>
<td>111 GPs &amp; specialists in primary care modified BG</td>
<td>? (4 weekends)</td>
<td>Outcome</td>
<td>Outcome</td>
</tr>
<tr>
<td>Turner and Malm [27]</td>
<td>PMI</td>
<td>2</td>
<td>6 Residents (+8 control)</td>
<td>18 h (1x/2wks–1 hr–9 months)</td>
<td>Outcome: psychological medicine skills</td>
<td>Preference to discuss terminally ill, female patients of same age; [ verbal activity participants; ] reflection on patient; [ reflection on self ]</td>
</tr>
<tr>
<td>von Klitzing [13]</td>
<td>Session transcriptions (word counts)`</td>
<td>1</td>
<td>7 Nurses</td>
<td>? (?–1.5h–1 yr)</td>
<td>Process: verbal reflective activity; themes</td>
<td>Inconclusive (no significance tests)</td>
</tr>
</tbody>
</table>

Quantitative parts in mixed method papers

| MUSHAM and BROCK [5] | MBTI | 1 | 16 Residents (9 freq vs 7 infreq attenders) | ? (1x/wk–1 hr–7?) | Part. characteristics: personality Evaluation BG | Tendency to positive group ratings (only descriptive statistics) |
| PARKER and LEGGITT [32] | 2 evaluation questionnaires (own design; 5 + 6 items) | 1 | 42 Students | 6–8 h (7–1 hr–6–8 wks) | Facts BG; leadership; process; themes; objectives; outcome: attitudes Facts BG; leadership; possible to present them here | Detailed results (not possible to present here) |

Historical papers

| BROCK and STOCK [51] | Questionnaire (own design) | n/a | 354 Family practice residency directors | n/a | Facts BG; leadership; process; themes; objectives; outcome: attitudes Facts BG; leadership; possible to present them here | Detailed results (not possible to present here) |
| JOHNSON et al. [35] | Questionnaire (own design) | n/a | 298 Family practice residency directors | n/a | Facts BG; leadership; process; themes; objectives; outcome: attitudes Facts BG; leadership; possible to present them here | Detailed results (not possible to present here) |

Abbreviations: BG; Balint group; part.: participants; REM-71: Response Evaluation Measures-71; SAT-P: Satisfaction Profile; GCQ: Group Climate Questionnaire—short version; OSCE: objective structured clinical examination; MB: Maslach Burnout Inventory; PMI: Psychological Medicine Inventory; JSPE: Jefferson Scale of Physician Empathy; MBTI: Myers-Briggs Inventory; WEPS: Work Environmental Preference Schedule; IE: Rotter’s Internal-External Locus of Control; FIRO-B: Schutz’s Fundamental Interpersonal Relationship Orientation Behavior Test; POI: Personal Orientation Inventory.

\[ Time in BG: \] approximate number of hours calculated by multiplying mentioned session frequency (x/week or x/month), session length (hours) and overall duration of Balint group (weeks, months or years) (possible holiday breaks could not be taken into account, thus for the longer lasting groups the calculated numbers may be slightly overestimated).

\[ a \] If instrument is not self-report.

own psychopathology’. Four studies [4–6,16] used semi-structured interviews to describe participants’ perception of the effect of BG participation. Among the effects we found: understanding case dynamics, awareness of one’s own and patients’ feelings, using a new perspective/conceptual framework [4], competence in the physician–patient encounter, recognizing different aspects of professional identity [6], increased self-awareness and interacting with patients differently [5,16]. Finally, Samuel [54] observed individual changes in some participants’ approach towards the group and their patients, as well as a maturation of their defenses.

3.3.2. Characteristics of Balint group participants

Five quantitative studies [5,23,52,55,56] compared characteristics of BG participants and professionals with no or only limited experience with BGs. Dokter et al. [52] compared ‘Balint characteristics’, personality traits and perceptions of patients in a group of Balint participants and a control group, but failed to report on statistical tests. Comparing a group of residents participating in a BG for two years (labeled ‘attendees’) with residents who left after the obligatory 6 month participation (labeled ‘non-attendees’), Cataldo et al. [21] found no significant differences in empathy or overall work satisfaction. Moreover, Johnson et al. [23] found that their group of ‘non-attendees’ was less intuitive. Kjeldmand et al. [55] found that experienced BG participants (>1.5 years) had significantly higher scores on self-reported control, satisfaction, quality of work, co-operation, training, health and attitudes towards psychosomatic patients than GPs with no BG experience. Finally, although relying on a very small sample, Joukamaa, Lehtinen and Karlsson [56] noticed that BG participants showed lower ability to detect patients’ mental disorders than non-BG-participants.
### Table 2
Overview of included qualitative papers.

<table>
<thead>
<tr>
<th>Article</th>
<th>Data + participants</th>
<th>Analysis</th>
<th>Time in BG</th>
<th>Topics</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Qualitative papers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brock and Johnson [96]</td>
<td>Process notes of 66 BG sessions with GPs</td>
<td>Description</td>
<td>? (1x/wk–7–7)</td>
<td>BG as research method</td>
<td>Typology of 5 potential harmful GP roles: description + example</td>
</tr>
<tr>
<td>Dahlgren et al. [16]</td>
<td>Semi-structured interviews with 3 BG part. (physiotherapists)</td>
<td>Interpretative phenomenological approach (metaphors)</td>
<td>? (1x/month–7–18 months)</td>
<td>Process; effects</td>
<td>8 process elements grouped into 4 phases (e.g. expression of difficulties, meeting other perspectives, applying insights to practice)</td>
</tr>
<tr>
<td>Graham et al. [4]</td>
<td>Semi-structured interviews with 17 BG part. (psychiatry residents and counselors)</td>
<td>Description</td>
<td>16 h (1x/wk–1 1/4h–12 wks)</td>
<td>Evaluation; process; effects</td>
<td>Evaluation: groups were anxiety provoking; process: e.g. group container, self-reflection; effects: e.g. understanding case dynamics, awareness of own and patient’s feelings, new perspective/ conceptual framework</td>
</tr>
<tr>
<td>Kjeldmand and Holmström [60]</td>
<td>Semi-structured interviews with 8 BG leaders</td>
<td>Systematic text-condensation method</td>
<td>? (part. from diff groups)</td>
<td>Leadership; process</td>
<td>3 categories of difficulties in BGs: (1) related to individual member, (2) related to group/leader, (3) related to group surroundings</td>
</tr>
<tr>
<td>Merenstein and Chillag [36]</td>
<td>Observation of 14 BG sessions (field notes); interviews with 10 BG leaders; 7 focusgroups with BG part.</td>
<td>Editing style</td>
<td>? (part. from diff groups)</td>
<td>Leadership</td>
<td>Comparison of different BGs in terms of format, themes discussed, dynamics, leadership</td>
</tr>
<tr>
<td>Pinder et al. [61]</td>
<td>Observation of 6 BG &amp; 2 non-BG meetings (field notes; interviews with 13 BG part. (GPs - registrars); discussion with leaders)</td>
<td>Ethnographic approach/case studies</td>
<td>? (part. from diff groups)</td>
<td>Process; evaluation</td>
<td>Process: group dynamics; evaluation: positive and negative experiences</td>
</tr>
<tr>
<td>Samuel [54]</td>
<td>Tape records; leader’s notes; report by 11 BG part. (pre: expectations; post: evaluation, change of others); group attitude questionnaire by 11 BG part. (pre + post)</td>
<td>Description</td>
<td>90 h (1x/2wks–1.5h–2.5 yrs)</td>
<td>Themes; process; effects</td>
<td>Themes: often personal themes; process: identification with cases, use of group for immediate help in daily work; effects: maturation of defenses, some change in attitudes towards group and patients, little sensitivity towards other members’ change</td>
</tr>
<tr>
<td>Torppa et al. [34]</td>
<td>Leaders’ notes on 2 BGs (medical students)</td>
<td>Grounded theory</td>
<td>15 and 7.5 h (1x/2wks and 1x/wk–1.5h–10 and 5 sessions)</td>
<td>Themes</td>
<td>Themes: e.g. feelings related to patients, building professional identity, negative role models, cooperation with other medical professionals</td>
</tr>
<tr>
<td>Van Roy et al. [7]</td>
<td>Observation notes; tape records + transcripts of 2 case discussions in 2 BGs (GPs + mixed)</td>
<td>Description</td>
<td>? (part. from diff groups)</td>
<td>Process</td>
<td>Characterisation of change in participants in 2 case discussions</td>
</tr>
<tr>
<td><strong>Qualitative parts in mixed method papers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Musham and Brock [5]</td>
<td>Semi-structured interviews with 16 BG part. (residents)</td>
<td>Not mentioned</td>
<td>&gt;24 h (1x/wk–1 hr– &gt;6 months)</td>
<td>Evaluation; effects</td>
<td>Evaluation: little initial understanding of BGs, reasons for infrequent attendance; effects: heightened self-awareness, interacting with patients differently</td>
</tr>
<tr>
<td>Parker and Leggett [32]</td>
<td>Unstructured written feedback from 16 BG part. (medical students)</td>
<td>Thematic analysis–grounded theory</td>
<td>6.8 h (7–1 hr–6.8 wks)</td>
<td>Evaluation</td>
<td>Reflections on value of BG, limitations in relevance for students, advice for adaptation</td>
</tr>
<tr>
<td><strong>Qualitative and quantitative parts not separable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson et al. [62]</td>
<td>Qual: open evaluation forms + focus groups with 21 BG leaders; quant: evaluation forms</td>
<td>Qual: grounded hermeneutic editing approach + quant: PCA</td>
<td>? (part. from diff groups)</td>
<td>Leadership</td>
<td>5 essential leadership skills: creating safe climate of safety, guarding over group norms, encouraging group movement, understanding group process, personality/style of leader</td>
</tr>
</tbody>
</table>

Abbreviations: BG part.: Balint group participants; PCA: principal component analysis.

* Time in BG: approximate number of hours calculated by multiplying mentioned session frequency (x/week or x/month), session length (hours) and overall duration of Balint group (weeks, months or years) (possible holiday breaks could not be taken into account, thus for the longer lasting groups the calculated numbers may be slightly overestimated).


3.3.3. Themes addressed in Balint groups

Several studies reported on the themes that were addressed during BG sessions. Nevertheless, Torppa et al. [34] were the only authors to present a systematic overview of the themes addressed in (student) BGs, which are illustrated with examples. The majority of papers provide only a brief non-systematic selection of the themes that were addressed [11,13,15,19,25,28,36,37,43,48,57–59]. Brock and Stock [51] provide an overview of the frequency with which specific themes are addressed in BGs. As general trends, Samuel [54] observed that themes often represented a personal involvement with particular kinds of problems, and von Klitzing [13] observed a tendency for participants to present terminally ill patients that were similar to themselves with regard to gender and age.

3.3.4. Balint group processes

Seven qualitative articles [4,6,7,16,54,60,61] studied the process of BG meetings or BG participation (i.e. how BG participation leads to change). Dahlgren et al. [16] investigated participants’ descriptions of the BG process and deduced eight elements grouped into four phases (e.g. expression of difficulties, meeting other perspectives, applying insights to practice). Gratamini et al. [4] described pathways through which change in BG participants occurred, pointing at aspects such as the group’s container function or self-reflection. Kjeldmand and Holmström [6] indicated parallel processes, a sense of security, and the recognition of participants’ professional identity as some of the group processes at work. Samuel [54] noted that participants often identified with each other and their patients in the case discussions, and that they often used the group for immediate help in their daily work. Pinder et al. [61] pointed out helpful as well as limiting group dynamics by making use of detailed case examples. Kjeldmand and Holmström [60] indicated potentially negative group processes such as scapegoating. Reading BG work from a Lacanian theoretical framework, Van Roy et al. [7] described two participants’ process of change over one session. Furthermore, two quantitative studies [13,45] investigated BG processes. Over the course of several sessions, Abeni et al. [45] found increased group engagement and decreased group conflict, but this was only the case in a group of caregivers. Finally, using word counts in session transcripts, von Klitzing [13] observed that participants’ verbal activity and reflections about their patients increased, whereas reflection about themselves decreased over the sessions.

3.3.5. Leadership in Balint groups

Five articles focused on various aspects of BG leadership by making use of either a qualitative methodology [36,60], a mixed qualitative-quantitative methodology [62] or a quantitative methodology [35,51]. Kjeldmand and Holmström [60] focused on leaders’ experiences of difficulties in their groups, while Johnson et al. [62] extracted five essential leadership skills from BG leaders’ evaluation forms and focus groups (creating a safe climate, guarding over group norms, encouraging group movement, understanding the group process, personality/style of leader). Relying principally on observations of several BGs, Merenstein and Chillag [36] discussed several leadership-related issues, e.g. personality of leaders, degree of hierarchy and degree of control. Brock and Stock [51] quantitatively investigated leaders’ perceptions of group objectives, format, issues, attitudes or skills attainable through BG participation and leaders’ professional backgrounds. Group objectives and leaders’ professional backgrounds were later reassessed by Johnson et al. [35], who also included data on leaders’ training.

3.3.6. Evaluation of Balint groups

Several papers focused on participants’ evaluation of their participation in a BG by using qualitative interviews [45,61], qualitative written reports [32] or quantitative questionnaires [19,28,31,32]. Some described rather positive group evaluations. For instance, Sekeres et al. [28] reported that the residents evaluated the groups as safe, as an opportunity to decompress, and as a social outlet. Other authors outlined participants’ negative experiences, especially in mandatory groups: Graham et al. [4] reported that residents participating in a BG experienced the groups as anxiety provoking and that some struggled to adapt to the learning process. Similarly, Musham and Brock [5] observed that residents initially poorly understood the purpose of the groups; moreover, participants indicated factors such as time, discomfort and not being convinced of the relevance of BG work to their clinical work as reasons for infrequent attendance [5]. Finally, some papers reported mixed findings. Although somewhat tentatively, Parker and Leggett [31,32] mentioned participants’ rather positive evaluation of individual group sessions, whereas participants were more hesitant about the relevance of the groups to their clinical practice. Pinder et al. [61] provided detailed group evaluations (including both positive and negative aspects) by interviewing the presenters after the group meetings.

3.3.7. Historical-geographical articles

Two studies presented a number of facts about US BGs. Brock and Stock [51] conducted a survey study offering data about existence, leadership, meeting frequency, objectives and composition of BGs in US family practice residencies, with a follow up study ten years later in 2000 [35]. Other, non-empirical articles provide historical information about the introduction of BGs (the initial groups by Michael Balint as well as other groups) and about Balint societies [2,3,15,49,63–68].

3.3.8. Reports – anecdotal articles

A large part of the non-empirical papers are reports about BGs. Frequently, they describe (co-)leaders’ or participants’ BG experiences. They often comprise detailed information about BG meetings (e.g. [17,42]), initiatives of setting up BGs (e.g. [66]), difficulties encountered (e.g. [39]), issues addressed (see 3.3.3), interventions applied (e.g. [26,48]), instruments used (e.g. ‘initial interview card’ [69]), a group’s evolution (e.g. [15,20,24,25,39]) or group evaluations (e.g. [14,18,25,29,30,70]). Sometimes the reports describe specificities of BGs for special target groups or specificities about ‘modified Balint groups’ (see Section 3.2). Some papers also offer case examples, which are further analyzed in the paper (e.g. [14,17,24,38,40,41,69,71–76]), either serve as a mere illustration (e.g. [29,57,77]). One paper consisted of an excerpt of a transcript of a BG meeting [78].

3.3.9. Reflective articles

In a substantial number of the non-empirical papers, the authors reflect on diverse BG related topics. The depth of reflection varies: some articles mainly present different aspects of what BG work is, whereas others provide a critical reflection about specific Balint-related issues. The most frequently discussed topics include: the need for Balint training, the place for such training in (continuing) medical education (e.g. [71,77,79,80]), the role of mandatory groups (e.g. [81]) and the future of BGs (e.g. [49]). Several papers focus on the specificity of BGs (e.g. [82–86]), comparing them to other forms of group discussions (e.g. [87–89]), discussing the possibility of BGs for other professions (e.g. [74]) or the necessity to adapt BGs to the participants’ needs (e.g. [26,58]). Some authors reflect on the change that BG participation might facilitate (e.g. [69,77,90]) or on leadership issues (e.g. [33,41,66,74,85,91,92]). In certain papers, the authors use
theoretical concepts as a framework for understanding BG processes [72,86,91,93,94]).

3.3.10. Balint group observation as research data

Michael Balint introduced his seminars (later called ‘Balint groups’) as ‘training-cum-research’ groups [82]. This means that these seminars not only aimed to ‘train’ GPs, but also to investigate, as a group, aspects of general practice. In line with this last aim, some papers discuss the use of BGs as a research method [69,71,84,95]. Some studies actually used BG observations as research data to study for instance GPs’ defenses [54,96]) or the phenomenon of ‘third party in general practice consultations’ [46]. Bourne and Lewis [64] reflected upon the involvement of BGs in such research projects. The scientific value of BGs was questioned by Sowersby [97].

4. Discussion and conclusions

4.1. Discussion

The main aim of the present study was to provide an overview of the peer-reviewed articles on Balint groups. We included empirical (n = 35) as well as non-empirical (n = 59) papers. We reviewed the variations in BG setting, diverse outcome variables (such as psychosocial self-efficacy, burnout and change in attitudes), participants’ characteristics, themes, processes, leadership issues, group evaluations, historical information and the main topics addressed in reports and reflective articles. As we demonstrated, Balint group research includes very diverse research topics, with low numbers of studies focusing on the same topic. Moreover, several articles appeared to be methodologically weak. With regard to the quantitative studies, these shortcomings include the use of (too) small samples, the omission of a control group, the lack of information about the reliability and validity of the instruments used, the incorrect use of statistics (e.g. providing no information on the significance of results) or the misleading presentation of the results (e.g. in the abstract or title); for the qualitative studies, the shortcomings concern for instance a lack of systematic approach or information about the data-analysis. However, despite these shortcomings, we found indications of the value BGs may still have today. This became evident from some qualitative studies on BG effects and participants’ evaluations as well as from personal reports and reflective articles. In order to gain more insight into this value and into aspects of potential improvement of BG work, more solid and systematic findings are needed.

4.2. Limitations

There are limitations to this literature review. As mentioned above, only English-language articles were included, though a large number of articles in other languages are available. Reviewing these articles and contrasting them with the current study surely could be worthwhile. Moreover, books (e.g. [1,98,99]), conference proceedings and articles in national Balint society journals were not taken into account, although they may contain interesting views. Finally, given the occasionally flexible distinction between what is a BG and what not, papers using different names for their groups may have been overlooked.

4.3. Practice implications

This review study highlights specific points of interest for both professionals involved in BG work and (future) BG researchers. First, since some papers reported effects (e.g. psychosocial self-efficacy, burnout) only after long-term BG participation [11,12,55]), BGs should be organized for a sufficient length of time (1 or 1.5 year at least) to allow for change. Next, the topic ‘modified Balint groups’ was addressed in various papers, indicating that this is an issue that is alive among BG professionals. This applies for instance to BGs for students/residents who may have particular needs (e.g. [26,31]). This leads to the broader issue of what the core of BG work is and what may be fruitful adaptations. In order to allow for meaningful comparison and discussions on this topic, articles should supply information about proceedings, goals, group composition, leaders’ profession and authors’ relationship to the BG. Further considerations for researchers include setting up well-considered study designs. This implies for instance learning from the findings and shortcomings of previous studies; this task could be facilitated by the present study. As already noted, defining and selecting appropriate (outcome) variables is an important and difficult question to address. Grasping the core of BG work and defining the ‘limited but considerable change’ as referred to by Michael Balint [1] appears to be a tough but necessary task. In this respect, we argue that well-designed qualitative studies are needed, as they allow for a more explorative focus. Moreover, qualitative studies may counter the difficulty of finding sufficient numbers of participants, which is often not evident due to the low number of Balint groups. For instance, studies investigating BG participants’ experiences through a critical incident technique [100] could shed light on crucial aspects of BG work and potential effects of BG participation. Valid research findings may not only help BG workers to enhance their practice, they could also help policy makers to make more informed and appropriate decisions.

Acknowledgments

We want to thank our colleagues Joachim Cauwe, Virginie Debaere and Femke Truijens for their critical and useful help in revising the manuscript. All authors are employed by Ghent University who funded this research.

The authors declare no conflicts of interest related to this work.

References

The student training

Lipsitt Abeni

mental professionalism.

Pract

A

D.

Med

S.

Leggett K. Behavioral

and

S.


Kulenovic M. Balint groups in the postwar period in Croatia (with regard to medical and social conditions). Coll Antropol 1999;23:293–8.


