Randomized Clinical Trial of Morphine in Acute Abdominal Pain

E. John Gallagher, MD
David Esses, MD
Conroy Lee, MD
Michael Lahn, MD
Polly E. Bijur, PhD

From the Department of Emergency Medicine, Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, NY.

Study objective: Administration of analgesia to patients with acute abdominal pain is controversial. We test the hypothesis that morphine given to emergency department (ED) patients with acute abdominal pain will reduce discomfort and improve clinically important diagnostic accuracy.

Methods: Pain was measured with a standard 0- to 100-mm visual analog scale. ED patients with acute abdominal pain were randomized in a double-blind fashion to 0.1 mg/kg intravenous morphine or placebo. The primary endpoint was the difference between the 2 study arms in clinically important diagnostic accuracy. Clinically important diagnostic accuracy was defined a priori by its complement, clinically important diagnostic error, using 2 independent, blinded investigators to identify any discordance between the provisional and final diagnoses that might adversely affect the patient's health status. The provisional diagnosis was provided by an ED attending physician, who examined the patient only once, 15 minutes after administration of the study agent. The final diagnosis was obtained through follow-up at least 6 weeks after the index ED visit.

Results: We randomized 160 patients, of whom 153 patients were available for analysis, 78 patients in the morphine group and 75 patients in the placebo group. Baseline features were similar in both groups, including initial median visual analog scale scores of 98 mm and 99 mm. The median decrease in visual analog scale score at 15 minutes was 33 mm in the morphine group and 2 mm in the placebo group. There were 11 instances of diagnostic discordance in each group, for a clinically important diagnostic accuracy of 86% (67/78) in the morphine group and 85% (64/75) in the placebo group. The difference in clinically important diagnostic accuracy between the 2 groups was 1% (95% confidence interval [CI] −11% to 12%). Analysis by efficacy and intention to treat yielded similar results. κ for interobserver concordance in classification of clinically important diagnostic accuracy was 0.94 (95% CI 0.79 to 1.00). No patients required naloxone.

Conclusion: Although administration of intravenous morphine to adult ED patients with acute abdominal pain could lead to as much as a 12% difference in diagnostic accuracy, equally favoring opioid or placebo, our data are most consistent with the inference that morphine safely provides analgesia without impairing clinically important diagnostic accuracy. [Ann Emerg Med. 2006;48:150-160.]

0196-0644/$-see front matter

doi:10.1016/j.annemergmed.2005.11.020

SEE EDITORIAL, P. 161.

INTRODUCTION

Background

The traditional reluctance of many physicians to administer analgesia to patients with acute abdominal pain can be traced to a recommendation made by Cope in 1921.[1] Dr. Cope, a widely respected surgeon of his day, admonished his colleagues to avoid administration of any agent that might mask physical findings in patients with acute abdominal pain. He feared that potent analgesics such as morphine might obscure the underlying diagnosis, resulting in delayed surgical intervention and adverse outcomes. In the context of the preantibiotic era, with the development of opioid antagonists and sophisticated imaging technology still many years in the future, Cope's caution may have been historically appropriate for the early 20th century.

Importance

Advances in medicine and surgery during the last 85 years, reflected in 21 successive editions of Cope's original text, have

Editor's Capsule Summary

What is already known on this topic

Despite 10 clinical trials that suggest that administration of analgesia to patients with acute abdominal pain decreases pain without compromising diagnostic evaluation, some physicians withhold analgesia from these patients.

What question this study addressed

This randomized, double-blind, placebo-controlled, clinical trial examined whether the administration of intravenous morphine to adult emergency department patients decreased diagnostic accuracy for clinically important conditions.

What this study adds to our knowledge

This study adds to a substantial body of evidence that indicates that appropriately dosed morphine does not adversely affect the evaluation of patients with abdominal pain. Morphine administration reduced discomfort without impairing diagnostic accuracy.

How this might change clinical practice

Physicians should provide analgesia for patients with abdominal pain. There is substantial evidence that this humane practice does not compromise diagnostic accuracy.

softened these initial proscriptions against analgesics, replacing them with tentative recommendations favoring their judicious use.[2,3] However, neither the practice of withholding nor administering analgesia to patients with acute abdominal pain enjoys the support of a sound evidence base. Although a critical appraisal of the aggregate weight of information contained in the 10 clinical trials published in the English literature suggests that administration of analgesia to patients with acute abdominal pain is safe,[4-13] there is by no means universal agreement.[14]

Goals of This Investigation

In an effort to explore this further, we designed a clinical trial to determine whether an association existed between administration of intravenous morphine to emergency department (ED) patients with acute abdominal pain and improvement in clinically important diagnostic accuracy.

MATERIALS AND METHODS

Study Design

We designed a prospective, randomized, double-blind, placebo-controlled trial to compare the effects of 0.1 mg/kg of intravenous morphine with placebo on clinically important diagnostic accuracy among patients presenting to the ED with acute abdominal pain. The study was approved by the institutional review board of the Montefiore Medical Center. All patients provided written, informed consent.

Setting

The study was conducted in the ED of the Montefiore Medical Center, the academic medical center for the Albert Einstein College of Medicine. This inner-city Bronx ED treats more than 80,000 adults annually.

Selection of Participants

Patients were eligible if they were 21 years or older, had atraumatic abdominal pain of less than 48 hours' duration, and were judged by the ED attending physician to warrant opioid analgesia for pain control. The ED attending physician was incorporated into the eligibility criteria because he or she was the physician ultimately responsible for the patient's treatment in the ED. Patients were excluded if they had isolated flank pain, were pregnant, were allergic to morphine, had self-medicated with analgesia before coming to the ED, received analgesic medication in the ED before screening for study entry, had a concurrent painful sickle cell crisis, had a systolic blood pressure less than 100 mm Hg, refused participation, or were unable to provide informed consent.

Interventions

Patients were randomized to receive 0.1 mg/kg morphine intravenously up to a maximum of 10 mg, or an equal volume of normal saline solution administered as a single intravenous bolus. The block allocation schedule was determined by an online randomization plan generator (available at: http://www.randomization.com), designed to allocate an equal number of patients to the morphine and placebo groups after every block of 10 subjects was randomized. The pharmacy used the schedule to determine the content of consecutive, identical-appearing, numbered vials. Labeled packets were provided to the ED, each containing a numbered vial and label to be attached to the patient's data collection instrument. An opaque, sealed envelope, which identified the contents of each vial (to be used in the event of an emergency) was also part of the prepared study packets. The master file linking the allocation assignment to the unique study identification number was maintained as hardcopy by the medical center's research pharmacist who prepared the study drugs. This information was sequestered by the pharmacist in a location unknown to the investigators.

Methods of Measurement

After enrollment, full-time, trained, fluently bilingual (English-Spanish) research associates gathered basic demographic information using a standardized data collection instrument, which was piloted on 10 patients and revised. The pilot data were then discarded. Before the study drug was administered, the research associate asked patients to rate their pain intensity on a standard horizontal 100-mm visual analog scale ranging from "no pain" to "worst possible pain." The visual analog scale has been shown to be a valid and reliable instrument for the measurement of acute abdominal pain in ED patients.[15] Fifteen minutes after administration of the study drug, the research associate again asked patients to rate their pain on a second visual analog scale without access to the initial visual analog scale. At that time, the research associate asked one of the full-time ED attending physicians not involved in the care of the patient to see the patient solely for study purposes. After obtaining a medical history and performing a physical examination, the study physician was asked by the research associate to select the single most likely diagnosis for that patient from a menu of options. The menu also contained an "other" category to be recorded as free text. The initial diagnosis provided by the study physician, which was made without access to any laboratory or radiographic information, constituted the provisional diagnosis. The research associate monitored vital signs at baseline, 15 minutes postbaseline, and then every half hour, for a total of 3 hours. After patients received a provisional diagnosis, they were eligible to receive additional analgesia, antiemetics, and any other agents deemed appropriate by the primary attending physician treating the patient. This physician had no role in the study.

We chose to obtain a provisional diagnosis from an attending physician who had not previously examined the patient in an effort to minimize diagnostic suspicion bias.[16] This form of detection bias could be introduced by use of a physician who had already examined the patient before administration of the study drug. In such a before-after setting, the clinical information obtained from an initial examination before administration of the study drug would naturally carry forward and might bias findings acquired from a second examination by the same individual after administration of the study drug.

Research associates obtained extensive contact information from patients to facilitate follow-up. Six weeks after the ED visit, all records were reviewed to obtain diagnostic test results, surgical findings, and pathology reports. Patients whose outcome was unclear from a review of the medical record were contacted by the research associates and, using a scripted telephone questionnaire, queried about their abdominal pain and whether they had visited an ED, been admitted to a hospital, undergone a diagnostic or therapeutic medical or surgical procedure, or consulted a physician or other health care provider since their index ED visit. Final diagnosis was determined from all available sources: medical record, patient's physician (permission for this had been obtained from the patient at enrollment in the study), and patient self-report.

In instances in which the final diagnosis could not be determined by the research associate at 6 weeks, follow-up was transferred to the investigators, who contacted patients with periodic telephone calls until either a definitive diagnosis was obtained or the presenting problem resolved. The definitive (final) diagnosis assigned to patients was determined by a convincing radiographic or tissue diagnosis, a clear response to medical or surgical intervention, or spontaneous resolution. In the latter instance, patients with no apparent cause for abdominal pain that had resolved and not recurred at follow-up were assigned the diagnosis of nonspecific abdominal pain. Patients were followed for up to 18 months.

Data Collection and Processing

Participants were enrolled between June 12, 2002, and May 27, 2004, when research associates were present in the ED. During this period, the research associate coverage increased from about 84 hours per week in 2002 to about 152 hours per week in 2004. Enrollment was based on the availability of research associates.

To minimize transcription error before analysis, all data were entered independently, in duplicate, by 2 individuals. All discordant fields were reconciled by referral to the hardcopy of the data collection instrument. SPSS Data Entry 4.0 (SPSS, Inc., Chicago, IL) was used to perform the double data entry.

Primary Data Analysis

Our primary target outcome variable was the difference in clinically important diagnostic accuracy in the morphine versus the placebo group. To develop a working definition of clinically important diagnostic accuracy, we first defined its complement, clinically important diagnostic error, as any disagreement between the initial provisional and final diagnosis that might reasonably be expected to have an adverse impact on the patient's health status. Clinically important diagnostic error was independently classified by 2 investigators through review of the data collection instrument and all available follow-up information. If either investigator classified an instance of diagnostic error as clinically important, it was coded as such for the final analysis. The proportion of patients whose provisional diagnosis was classified as accurate, ie, either identical to the final diagnosis or not different from it in a clinically important way, was calculated for each of the 2 arms of the study. Results are presented as the difference between the proportions of clinically important accurate diagnoses in the morphine versus placebo group, bounded by exact 95% confidence intervals (CIs). All classification of diagnostic accuracy was completed before unblinding.

Interobserver concordance in classifying the presence or absence of clinically important diagnostic accuracy was expressed as simple unadjusted agreement and with the κ statistic, both bounded by exact 95% CIs.

Analogous to diagnostic accuracy, we also calculated the accuracy of the provisional disposition decision as a secondary endpoint, reporting it as the difference between the morphine and placebo groups in sensitivity and specificity, bounded by 95% CIs. We defined sensitivity and specificity as follows: sensitivity=(provisional disposition of admit)/(actual disposition of admit) and specificity=(provisional disposition of discharge)/(actual disposition of discharge).

To ensure that the 0.1 mg/kg dose of morphine we had chosen provided meaningful analgesia, we calculated the median change in pain 15 minutes after administration of the study drug. Although measurements obtained from the visual analog scale are often analyzed parametrically as continuous data, we chose to use medians and interquartile ranges (IQR) with a box and whiskers plot rather than means and SDs because our changes in pain scores were nonnormally distributed.

Finally, we planned a priori to develop a simple logistic model to adjust any association between treatment group allocation and clinically important diagnostic accuracy for individual medication-induced changes in pain intensity. Results of the multivariate (adjusted) odds ratio (OR) derived from the logistic model were then compared with the bivariate (crude) OR. SPSS Version 13.2 (SPSS, Inc., Chicago, IL) was used for all data analyses.

We estimated that the proportion of clinically important accurate diagnoses in the morphine-treated group would be 80% and the proportion in the group treated with saline solution would be 60%. These values are based on the earlier work of Pace et al.[6] Using an estimate of a 20% difference in proportions for the effect size, with a 2-tailed α of 0.05 and power of 0.80, the number needed in each group was 80, for a total sample size of 160. nQuery Advisor release 3.0 (Los Angeles, CA) was used to calculate the sample size.

RESULTS

Of 578 ED patients who had acute abdominal pain and were screened for study eligibility between June 2002 and May 2004, 160 patients met entry criteria. Reasons for exclusion of the remaining 418 patients are listed in the CONSORT diagram shown in Figure 1. Of the 160 patients randomized, one individual allocated to the placebo group did not receive the intervention and was disqualified as a protocol violation. Five additional patients were excluded from the final analyses, either because the provisional diagnosis was missing or because more than 1 provisional diagnosis was selected. One patient in the placebo group was lost to follow-up, leaving 153 patients available for analysis.

As shown in Table 1, demographic characteristics of the study population were comparable in both groups. In particular, baseline pain intensity was similarly severe, with median visual analog scale scores of 99 mm (IQR 88 to 100 mm) and 98 mm (IQR 78 to 100 mm) in the morphine and control groups, respectively. As illustrated in Figure E1 (available online at http://www.annemergmed.com), reduction in pain was strongly associated with allocation group: median change in pain among patients randomized to morphine was −33 mm (IQR −8 to −73 mm) versus −2mm (IQR 1 to −16 mm) in the placebo controls.

Table E1 (available online at http://www.annemergmed.com) summarizes the final diagnoses by allocation group. The 153 patients occupied 27 unique diagnostic categories. A substantial majority (80%) of patients had an abdominopelvic computed tomography (CT) scan, 47% of patients had either abdominal or pelvic ultrasonography, and 33% of patients went to surgery. Except for 10 of the 14 renal colic patients (none of whom presented with typical renal colic with isolated flank pain, or they would have been excluded by protocol), who had noncontrast CTs, and 19 patients with an elevated creatinine level, who received oral contrast only, the remainder of study patients received intravenous and oral double-contrast CTs. By chance, the 2 largest diagnostic categories of biliary tract disease and nonspecific abdominal pain were imbalanced. More patients with biliary tract disease were allocated to morphine, and more patients with nonspecific abdominal pain were allocated to placebo.

Table E2 (available online at http://www.annemergmed.com) lists all 22 of the clinically important instances of diagnostic discordance, each with its provisional diagnoses, final diagnoses, and allocation group. These were independently classified before unblinding the data by 2 investigators. There was interrater agreement on 151 patients' provisional and final diagnoses, 131 of which were classified by both raters as clinically accurate and 20 of which were classified by both raters as clinically erroneous. There was interrater disagreement on diagnostic accuracy in 2 patients. Rather than adjudicate the disagreement, the raters followed a conservative a priori protocol and classified the 2 remaining patients as members of the clinically important diagnostic discordance group. The unadjusted observed interrater agreement for classifying diagnostic accuracy was 99% (95% CI 95% to 100%), with an unweighted κ of .94 (95% CI 0.79 to 1.00).

Our primary endpoint, clinically important diagnostic accuracy, was 86% in the morphine group (67/78 provisional diagnoses correctly predicted the final diagnoses) versus 85% in the placebo group (64/75 provisional diagnoses correctly predicted the final diagnoses), for a difference in diagnostic accuracy of 1% (95% CI −11% to 12%). When we performed a logistic regression to adjust our primary endpoint for individual changes in visual analog scale pain scores, we found essentially no difference between the crude bivariate OR (OR 0.96; 95% CI 0.35 to 2.6) and its multivariate counterpart (OR 1.1; 95% CI 0.41 to 3.2).

The accuracy of the provisional disposition decision compared with the final disposition was similar in the placebo (55%) and morphine (58%) groups (95% CI for a difference of 3% −13% to 18%). The poor accuracy in both groups reflects a combination of low specificity (46% for morphine versus 30% for placebo; 95% CI for a difference of 16% −6% to 36%) and high sensitivity (98% for morphine versus 97% for placebo; 95% CI for a difference of 1% −10% to 11%) in the morphine and placebo arms, respectively. Of all the patients entered into the study, only 1 individual (allocated to morphine) had an unscheduled return visit to the ED 2 days later, resulting in admission. She was subsequently discharged from the hospital with the diagnosis of nonspecific abdominal pain (the same provisional diagnosis she had initially received), which resolved within the 6-week follow-up period without requiring intervention.

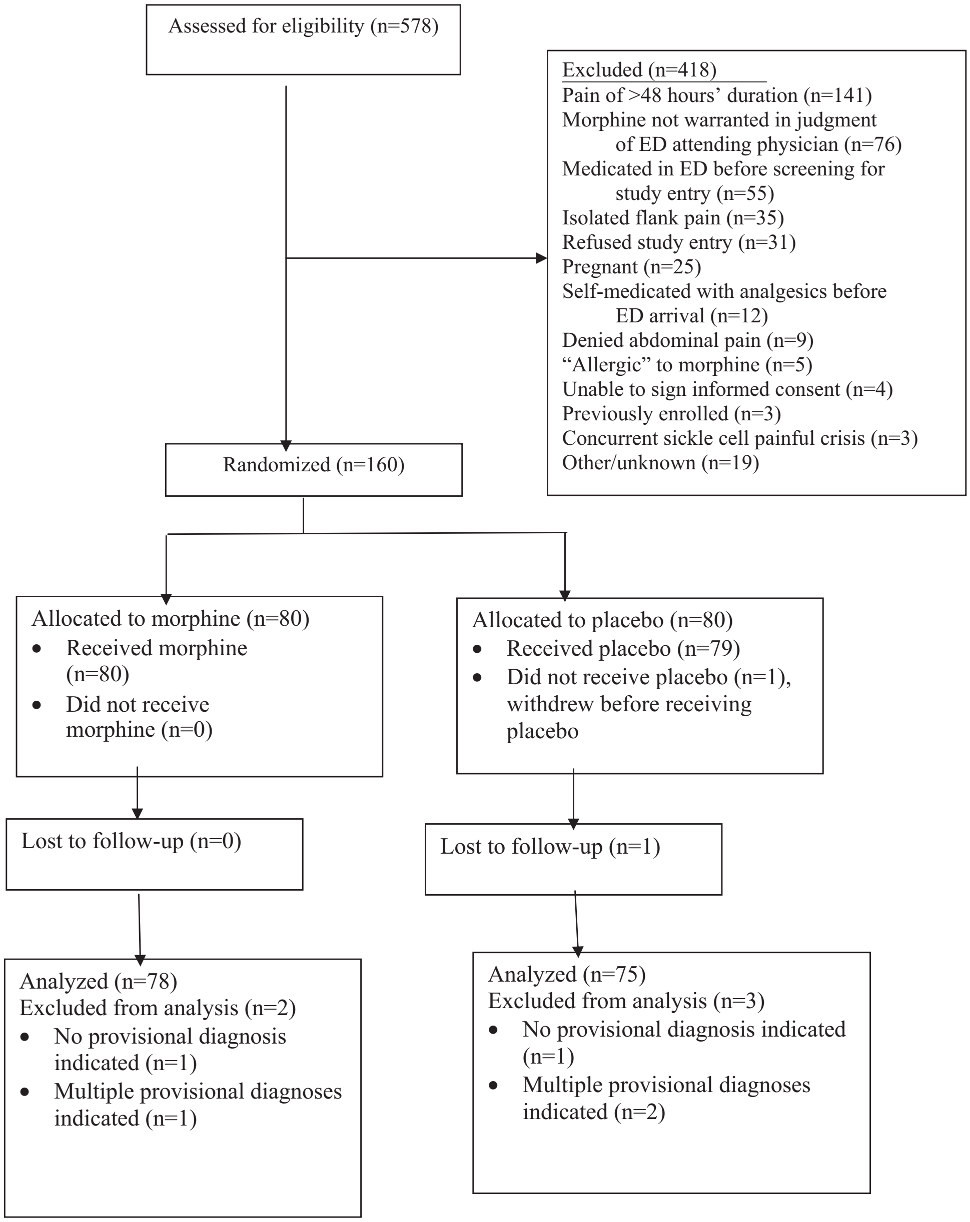

Figure 1. CONSORT diagram.

To assess blinding, research associates asked study physicians to indicate whether they believed the patient they had just treated had received opioid or placebo. The attending physicians correctly identified the study arm in 63% of cases.

Complications were similar in both groups: 13 patients had a systolic blood pressure less than 90 mm Hg recorded, 7 patients in the morphine group and 6 patients in the controls. All patients were easily treated with a saline solution bolus. Of the 37 patients complaining of nausea or vomiting, 16 patients were in the morphine group, and 21 patients were among the controls. Of these, 27 patients received antiemetics, 9 patients in the morphine group and 18 patients in the control group. There was no need to unblind providers or study personnel at any point during the trial. No patients required naloxone or any other opioid antagonist during their time in the ED.

Two patients died, one with biliary pancreatitis and a perforated gall bladder (allocated to placebo) and a second with mesenteric ischemia (allocated to morphine). The diagnosis in the first patient was made on CT and strongly suspected in the second patient, both within 90 minutes of ED arrival. Mesenteric ischemia with extensive bowel infarction was subsequently confirmed on invasive angiography.

LIMITATIONS

Inferences that can be drawn from these data are limited in several important respects. These limitations can be divided, in order of importance, into threats to internal and external validity. They are listed here and discussed in considerably greater detail in the Web version of the article (Appendix E1, available online at http://www. annemergmed.com).

By choosing clinically important diagnostic accuracy as our primary target endpoint, rather than a simple determination of whether or not the provisional and final diagnoses were identical, we may have added unnecessary complexity and some ambiguity to the study. In the interest of exploring what our findings might have shown, had we chosen the more straightforward but, we believe, less clinically relevant approach, we performed a post hoc analysis examining whether or not the provisional and final diagnoses were the same. We found a 68% accuracy among the morphine group versus a 47% accuracy among the placebo group. The difference in diagnostic accuracy of 21% (95% CI 6% to 36%) is similar to the 19% difference favoring intravenous morphine reported by Pace and Burk,[6] on the basis of whose work we chose the Δ of 20% for the effect size in our sample size calculation.

A second threat to internal validity was our decision to perform an efficacy analysis of our data rather than an intention-to-treat analysis. However, a sensitivity analysis, which is available in the Web version of the article (available online at http://www. annemergmed.com), indicates that performing an intention-to-treat analysis under the assumptions of all 4 possible extremes of alternate findings would not have materially affected our conclusion that administration of morphine does not seem to interfere with clinically important diagnostic accuracy in acute abdominal pain.

A third threat to the internal validity of our findings is a consequence of the quirks of randomization. Our 2 largest final diagnostic categories of nonspecific abdominal pain and biliary tract disease were unevenly allocated to the 2 arms of the study.

Another limitation, linked to our sample size calculation, was a failure to adjust for the effect of clustering by physician provider. However, when we analyzed the degree of clustering by physician that actually occurred, there were so many physicians participating in the study that there was in fact very little clustering by provider.

Turning next to external validity, limitations that may have an adverse impact on generalizability include at least 4 additional considerations. First, patients were gathered by convenience sampling, determined by availability of research associates during a 2-year period. This nonconsecutive recruitment of subjects is likely to have caused some assembly bias. A second limitation related to external validity is spectrum bias,[17] reflected by the extremely high pain scores of our study cohort (median visual analog scale 99 and 98 mm in the morphine and placebo groups, respectively). A third constriction on external validity is the relatively young age of the patients enrolled in the study. Despite the planned exclusion of individuals younger than 21 years, our study cohort consisted of patients with a mean age only in their midforties (Table 1). A final potential threat to the generalizability of our data derives from the composition of our study population, which was composed principally of minorities (85% Latino and black). Interpretation of our findings may therefore need to be viewed, at least in part, in the context of previously described differences in pain perception[18] and treatment[19,20] associated with race and ethnicity.

Table 1. Baseline features by randomization group.

Baseline Features	Morphine	Control
	(n=80), %	(n=80), %
Sex		
Female	70%	61%
Male	30%	39%
Ethnicity		
Latino	60%	58%
Black	26%	26%
Other	14%	16%
Age (mean±SD)	46 ± 17	45 ± 17
Weight (kg) (mean±SD)	80 ± 22.0	80 ± 20.2
Baseline VAS, median (IQR)*	99 (88, 100)	98 (78, 100)

VAS, Visual analog scale.
*IQR 25th to 75th percentile values of baseline VAS pain scores in each group.

DISCUSSION

Abdominal pain is the most common problem bringing patients to the ED in the United States, accounting for more than 7 million visits per annum.[21] The most recent available data (published in 1999 and 2000) indicate that only a minority of individuals presenting with abdominal pain receive analgesia.[22,23] Wolfe et al[22] found that although 85% of emergency physicians believe judicious administration of pain medication does not alter important physical findings, 3 of 4 of these individuals deferred administration of opioids until a surgeon had seen the patient. Consistent with this, Graber et al[23] surveyed practicing surgeons and found that 2 of 3 believed that analgesia interfered with diagnostic accuracy.

Our main finding that morphine administration significantly reduces discomfort without impairment of clinically important diagnostic accuracy suggests that this widely available, inexpensive analgesic can be used safely and efficaciously in patients with acute, severe abdominal pain. This observation is consistent with all the previous clinical trials we could find in the English literature, as summarized in Figure 2,[5-8,10,11,13] and displayed in greater detail in Table 2.[4-13]

The 7 studies that provided data on diagnostic accuracy are summarized graphically in Figure 2, along with the present study.[5-8,10,11,13] The figure does not meet the methodologically rigorous Cochrane criteria for a formal meta-analysis and should not be viewed as such. Rather, it simply produces a pictorial summary of the findings of a heterogeneous group of studies, containing a mixture of inpatients and outpatients, adults and

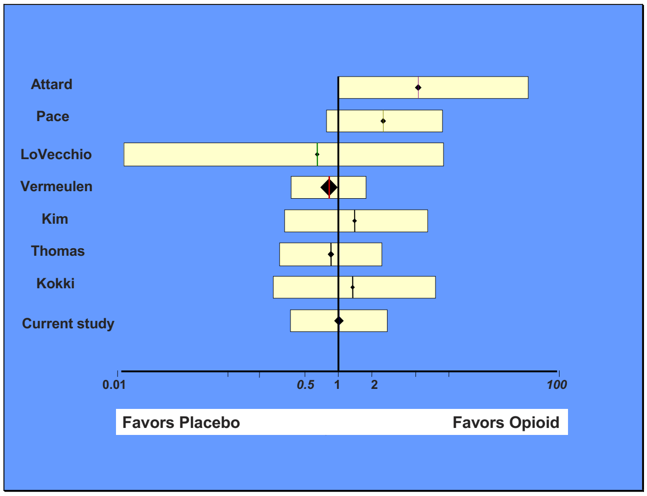

Figure 2. Summary of randomized clinical trials examining opioid analgesia and diagnostic accuracy in acute abdominal pain. Cochrane-style mini–meta-analysis, summarizing all work published in English on analgesia in acute abdominal pain targeting diagnostic accuracy. Vertical lines within each horizontal rectangle represent the OR for that study. Length of the rectangles reflects the width of the corresponding 95% CI.

children, and individuals with isolated right lower quadrant and undifferentiated abdominal pain. With the exception of the inpatient study by Attard et al,[5] the remainder of the clinical trials shown in Figure 2 have ORs whose point estimates are all reasonably close to the null point of 1 (ranging from 0.6 to 2.5), indicating little clinical or statistical difference in diagnostic accuracy between the opioid and placebo groups. Table 2 expands on the information displayed in Figure 2 by including the data extracted from all 10 previous studies in the English literature that have examined questions of opioid safety and efficacy in acute abdominal pain.[4-13] These are discussed chronologically in further detail below.

In 1986, Zoltie and Cust [4] randomized 288 hospitalized patients with acute abdominal pain to treatment with either the sublingual opioid buprenorphine or placebo. These authors reported that half the patients in the buprenorphine and placebo groups obtained pain relief, which did not impede arrival at the correct diagnosis.[4] Quantitative differences in diagnostic accuracy between the groups were not reported. The principal limitation of this study was the lack of an analgesic effect of buprenorphine that was superior to placebo, which may have accounted for the similar diagnostic outcomes among patients randomized to the 2 arms of the study.[4] In contrast, when we examined differences in improvement in visual analog scale pain scores in the 2 groups of patients randomized to intravenous morphine versus placebo in the present study, we found a large clinically and statistically significant difference in pain reduction (Figure E1, available online at http://www.annemergmed.com).

In 1992, Attard et al [5] randomized 100 hospitalized patients with acute abdominal pain to receive either 20 mg papaveretum (equivalent to about 12.5 mg morphine) or saline solution placebo intramuscularly. “Errors” in diagnosis or management occurred in 4% of patients randomized to papaveretum versus 18% of those randomized to placebo (95% CI for a difference of 14%, 2% to 26%).[5] As in the earlier trial by Zoltie and Cust,[4] this study was also conducted entirely on an inpatient service, thus limiting generalizability to the ED and outpatient settings where most abdominal pain is treated.

In 1996, Pace and Burke [6] conducted the first study of opioid use among ED patients with acute abdominal pain. Of 75 patients randomized to intravenous morphine (0.1 mg/kg, titrated to analgesic effect, maximum dose 20 mg) versus saline solution placebo, 71 (35+36) patients completed the trial. Although the authors did not claim diagnostic superiority for the opioid group, the emergency physician’s provisional diagnosis agreed with the final diagnosis in 80% of the patients randomized to morphine versus 61% to placebo (95% CI for a difference in diagnostic accuracy of 19%, −2% to 40%, favoring morphine). These authors concluded that use of opioids did not impair the physician’s ability to accurately evaluate patients with acute abdominal pain.[6] Although their findings favoring morphine achieved clinical significance and very nearly attained statistical significance as well, the comparatively small number of patients enrolled in the trial, as reflected in the wide CIs, limited the precision with which inferences might be drawn about the utility of opioid analgesia in acute abdominal pain.[6]

In 1997, LoVecchio et al[7] randomized ED patients with acute abdominal pain and peritoneal findings to intravenous opioid (19 patients received 10 mg morphine; 13 patients received 5 mg morphine) versus placebo (16 patients). Half of the 32 patients who received morphine had a change in their physical examination results. Of these, 29 of 32 (91%) patients had an ED diagnosis that agreed with their final diagnosis versus 15 of 16 (94%) patients in the placebo group. No adverse events or diagnostic delays were attributed to opioid administration.[7] Because of the sample size, which was smaller than Pace and Burke’s,[6] the 95% CI surrounding the estimate of 3% difference in diagnostic accuracy ranged from as much as 13% favoring placebo to 19% favoring morphine. The authors noted that substantially larger numbers of patients would be needed to determine whether a statistical or clinical difference existed that favored a strategy of either administering or withholding morphine.

In 1999, Vermeulen et al[8] randomized 340 ED patients with acute right lower quadrant pain to either 0.1 mg/kg intravenous morphine or placebo. The primary goal of the study, which was published in the radiology literature, was to determine whether opioid analgesia would alter the accuracy of ultrasonography in patients with suspected appendicitis. They found that opioids modified neither the sensitivity nor specificity of ultrasonography in this cohort, nor did opioid-induced pain

Table 2. Randomized clinical trials of opioid administration to patients with acute abdominal pain.

Study	Population and Setting	Opioid Route and Dose(s)	Evaluator	Diagnostic Accuracy Opioid Group N_1/N_2 (%)	Diagnostic Accuracy Placebo Group N_1/N_2 (%)	Difference in Diagnostic Accuracy (Opioid Group – Placebo Group) (95% CI)
Zoltie[4]	Adult inpatient, UK	Buprenorphine sublingual 200 μg and 400 μg	"Surgeon," not otherwise specified	Not reported	Not reported	Not calculable
Attard[5]	Adult inpatient, UK	Papaveretum intramuscular 20 mg	"Registrar" (surgical resident/fellow)	48/50 (96)	41/50 (82)	14 (2–26)
Pace[6]	Adult ED (military), US	Morphine intravenous 0.1 mg/kg (20 mg max)	Emergency medicine attending physicians and senior residents	28/35 (80)	22/36 (61)	19 (−2, 40)
LoVecchio[7]	Adult ED, US	Morphine intravenous 5 mg or 10 mg	Emergency medicine attending physicians and senior residents	29/32 (91)	15/16 (94)	−3 (−19, 13)
Vermeulen[8]	Adult ED, Switzerland, suspected appendicitis patients undergoing ultrasonography	Morphine intravenous 0.1 mg/kg	ED and radiology residents	156/175* (89)	150/165* (91)	−2 (−8, 5)
Mahaddevan[9]	Adult ED, Singapore, suspected appendicitis	Tramadol intravenous 1 mg/kg	Emergency medicine residents	Not reported	Not reported	Not calculable
Kim[10]	Pediatric ED, US	Morphine intravenous 0.1 mg/kg	Pediatric emergency medicine attending physicians and fellows	24/29 (83)	24/31 (77)	5 (−15, 26)
Thomas[11]	Adult ED, US	Morphine intravenous (15 mg max)	Not specified	24/38 (64)	24/36 (67)	−3 (−18, 25)
Wolfe[12]	Adult ED, US, suspected appendicitis patients scheduled for appendectomy	Morphine intravenous 0.075 mg/kg	Emergency medicine attending physicians and surgical residents	Not reported	Not reported	Not calculable
Kokki[13]	Pediatric ED, Finland	Oxycodone, buccal transmucosal 0.1 mg/kg	"Surgeons," not otherwise specified	28/32 (88)	26/31 (84)	4 (−13, 21)
Current study	Adult ED, US	Morphine intravenous 0.1 mg/kg	Emergency medicine attending physicians	67/78 (86)	64/75 (85)	1 (−11, 12)

*Accuracy of decision to perform or not perform appendectomy used as proxy for diagnostic accuracy.

relief alter the appropriateness of the decision to operate on the patient. Using the latter as a proxy for diagnostic accuracy in suspected appendicitis, the correct decision was made in 89% of the morphine group (156/175) and 91% of the placebo group (150/165) (95% CI for a difference of −2%, −8% to 5%).[8] As the authors pointed out, the validity of their findings was necessarily limited to suspected appendicitis presenting with right lower quadrant pain and could not be generalized to other locations and causes of abdominal pain.

In 2000, Mahadevan and Graff[9] randomized ED patients with acute right lower quadrant pain to tramadol, 1 mg/kg intravenously versus placebo. Seven abdominal signs were assessed before and after administration of the study agent. The tramadol (N=33) and placebo groups (N=33) appeared to have equivalent normalization of abdominal signs. Ten patients in each group had appendicitis, but diagnostic accuracy was not reported. As with the limitations of the earlier Zoltie and Cust[4] study, interpretation of these findings is impaired by the lack of a clinically significant difference in analgesic effect between the tramadol and placebo groups.

In 2002, in the first study of pediatric abdominal pain, Kim et al[10] randomized 60 children (aged 5 to 18 years) presenting to the ED with acute abdominal pain to either 0.1 mg/kg morphine or saline solution intravenously. These investigators found that, despite a reduction in pain, there was no significant difference in diagnostic accuracy or delay in surgical intervention between the 2 groups. The diagnostic accuracy of the pediatric fellow or attending physician was 83% (24/29) in the morphine group and 77% (24/31) in the placebo group (95% CI for a difference of 5%, −15% to 26%). Unfortunately, as the authors acknowledge, inferences of no true difference in diagnostic accuracy are substantially limited by low statistical power (about 18%) because of the small sample size.[10]

In 2003, Thomas et al[11] published the first clinical trial of opioid use in abdominal pain to appear in the American surgical literature. These investigators examined the use of morphine in ED patients presenting with acute abdominal pain. Of 74 individuals randomized to titrated doses of morphine (maximum 15 mg) versus placebo, no differences were detected in masking of physical findings or diagnostic accuracy. Indeed diagnostic accuracy was similar to that found in the current study, although, because of the smaller sample size, the difference of 3% was bounded by substantially less precise CIs (−18% to 25%) than in our study. The authors, who included the current editor of the most recent edition of Cope's updated text on the acute abdomen,[3] concluded that "there was no evidence supporting the contention that MS (morphine) administration was deleterious in any way."[11]

In 2004, Wolfe et al[12] administered morphine 0.075 mg/kg intravenously versus placebo in a randomized, double-blind, crossover design to 22 patients scheduled for emergency appendectomy. Each individual was randomized either to morphine followed by placebo (N=11) or placebo followed by morphine (N=11). The aim of the study was not diagnostic accuracy but rather changes in physical findings after administration of morphine and placebo. Three patients in each group had a change in their physical findings after medication. The authors concluded that morphine could provide significant pain relief without alteration of the physical examination among patients with clinical signs of appendicitis.[12] As in the 2 earlier studies that focused solely on suspected appendicitis,[8,9] any conclusions that can be drawn from this study are limited by a lack of generalizability to undifferentiated acute abdominal pain.

In 2005, Kokki et al[13] randomized children 4 to 15 years old who presented to a pediatric ED with acute abdominal pain to 0.1 mg/kg oxycodone (N=32) or placebo (N=31), delivered by the buccal mucosa.[13] Oxycodone was associated with significantly greater pain relief and similar diagnostic accuracy (88%) when compared to saline solution placebo (84%). Despite improving pain, opioid administration did not mask clinical evidence of peritonitis. The negative appendectomy rate was similar in the oxycodone (23%) and control (27%) arms, which is consistent with reported negative appendectomy rates in children.[24] The principal limitation of this study was that, similar to several of its antecedents, it was underpowered to identify a meaningful difference in diagnostic accuracy. This limitation was underscored in an accompanying commentary by Bowen et al.[25]

When we reviewed our 22 instances of diagnostic discordance, we found that more than 1 in 3 of them (36%) was attributable to incorrect assignment of the provisional diagnosis of nonspecific abdominal pain to patients who subsequently were found to have a more serious problem. Although nonspecific abdominal pain was our most common final diagnosis overall, comprising nearly 1 in 4 members (23%) of the study cohort, it appears inadvisable to select this as an early working diagnosis among ED patients presenting with sufficient abdominal pain to require morphine.

The poor accuracy of the disposition decision, which had no impact on the care of any patient, can be attributed largely to the study design. Because the study attending physician who was recruited to see the patient was asked to record a provisional diagnosis and disposition shortly after the patient's arrival in the ED, almost all the physicians chose admission as the safest, most sensible option. As the patient's abdominal pain evolved and diagnostic test results returned, the clinical picture became clearer, and a more appropriate disposition decision could be made. Because this naturally drove any changes in disposition from admission toward discharge, most of the discrepancies in provisional disposition took the form of false positives (provisional disposition of admit changed to final disposition of discharge) rather than false negatives (provisional disposition of discharge changed to final disposition of admit). Because false positives tend to erode specificity and false negatives reduce sensitivity, this is reflected in disposition specificities of only

46% and 30% compared to disposition sensitivities of 98% and 97% in the morphine and placebo groups, respectively.

Although the present study is the largest clinical trial of opioid use in ED patients with undifferentiated acute abdominal pain of which we are aware, the imprecision of our estimate, reflected by the width of the CIs, still leaves room for as much as an 11% to 12% superiority in clinically important diagnostic accuracy associated either with administration or withholding of morphine. Despite this, a difference of 1% in diagnostic discordance between the 2 groups is the point estimate and is therefore the best approximation of the true difference in diagnostic accuracy associated with morphine use in acute abdominal pain.

Indeed, the only difference between the 2 study arms that we identified was a clinically and statistically significant difference in pain reduction among those patients randomized to morphine compared with those who received placebo. The difference in median improvement in visual analog scale pain scores was 31 mm, which is roughly twice the value of 13 to 16 mm, which has been shown to represent the minimal clinically significant change in the intensity of acute abdominal pain.[15]

The management of acute abdominal pain varies enormously among institutions, depending largely on the quality and availability of diagnostic imaging, accessibility of interpretative expertise, and institutional custom. In our medical center, unless adult patients with abdominal pain are highly unstable, they are rarely transported directly from the ED to the operating room without first undergoing a CT scan. The results of the CT scan then heavily influence subsequent management, particularly the decision to operate. As this trend persists and images obtained by increasingly sophisticated technology continue to trump the clinical examination in identifying candidates for surgical intervention, the logic of the century-old argument that abdominal pain ought not to be altered by analgesic administration because of its putative diagnostic importance is commensurately weakened.

Although administration of intravenous morphine to adult ED patients with acute abdominal pain could lead to as much as a 12% difference in diagnostic accuracy, equally favoring opioid or placebo, our findings are most consistent with the inference that morphine safely provides analgesia without eroding clinically important diagnostic accuracy. These data are congruent with the aggregate weight of evidence accumulated from previous studies examining this question during the past 20 years.[4-13]

Supervising editor: Robert K. Knopp, MD

Author contributions: EJG and ML conceived of the study. CL and ML took primary responsibility for conduct of the trial. EJG, DE, and PEB analyzed the data. EJG wrote the manuscript and all authors contributed substantially to its multiple revisions. EJG takes responsibility for the paper as a whole.

Funding and support: The authors report this study did not receive any outside funding or support.

Publication dates: Received for publication September 14, 2005. Revisions received October 11, 2005, and November 1, 2005. Accepted for publication November 9, 2005. Available online February 13, 2006.

Reprints not available from the authors.

Address for correspondence: E. John Gallagher, MD, Department of Emergency Medicine, Albert Einstein College of Medicine, Montefiore Medical Center, 111 E. 210th St, Bronx, NY 10463; 718-920-7459, fax 718-798-0730; E-mail jgallagh@montefiore.org

REFERENCES

1. Cope Z. *The Early Diagnosis of the Acute Abdomen.* New York, NY: Oxford University Press; 1921.
2. Silen W. *Cope's Early Diagnosis of the Acute Abdomen.* 20th ed. New York, NY: Oxford University Press; 2000.
3. Silen W. *Cope's Early Diagnosis of the Acute Abdomen.* 21st ed. New York, NY: Oxford University Press; 2005.
4. Zoltie N, Cust MP. Analgesia in the acute abdomen. *Ann R Coll Surg Engl.* 1986;68:209-210.
5. Attard AR, Corlett MJ, Kidner NJ, et al. Safety of early pain relief for acute abdominal pain. *BMJ.* 1992;305:1020-1021.
6. Pace S, Burke TF. Intravenous morphine for early pain relief in patients with acute abdominal pain. *Acad Emerg Med.* 1996;3: 1086-1092.
7. LoVecchio F, Oster N, Sturmann K, et al. The use of analgesics in patients with acute abdominal pain. *J Emerg Med.* 1997;15:775-779.
8. Vermeulen B, Morablia A, Unger PF, et al. Acute appendicitis: influence of early pain relief on the accuracy of clinical and US findings in the decision to operate: a randomized trial. *Radiology.* 1999;210:639-643.
9. Mahadevan M, Graff L. Prospective randomized study of analgesic use for ED patients with right lower quadrant abdominal pain. *Am J Emerg Med.* 2000;18:753-756.
10. Kim MK, Strait RT, Sato TT, et al. A randomized clinical trial of analgesia in children with acute abdominal pain. *Acad Emerg Med.* 2002;9:281-287.
11. Thomas SH, Silen W, Cheema F, et al. Effects of morphine analgesia on diagnostic accuracy in emergency department patients with abdominal pain: a prospective randomized trial. *J Am Coll Surg.* 2003;196:18-31.
12. Wolfe JM, Smithline HA, Phipen S, et al. Does morphine change the physical examination in patients with acute appendicitis? *Am J Emerg Med.* 2004;22:280-285.
13. Kokki H, Lintula H, Vanamo K, et al. Oxycodone versus placebo in children with undifferentiated abdominal pain. *Arch Pediatr Adolesc Med.* 2005;159:320-325.
14. Nissman SA, Kaplan LJ, Mann BD. Critically appraising the literature-driven practice of analgesia administration for acute abdominal pain in the emergency room prior to surgical evaluation. *Am J Surg.* 2003;185:291-296.
15. Gallagher EJ, Bijur PE, Latimer C, et al. Reliability and validity of a visual analog scale for acute abdominal pain in the ED. *Am J Emerg Med.* 2002;20:287-290.
16. Feinstein AR: *Clinical Epidemiology: The Architecture of Clinical Research.* Philadelphia, PA: WB Saunders; 1985:618–621.
17. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978; 299:926-930.

18. Edwards CL, Fillingim RB, Keefe F. Race, ethnicity and pain. *Pain*. 2001;94:133-137.
19. Todd KH, Samaroo N, Hoffman JR. Ethnicity as a risk factor for inadequate emergency department analgesia. *JAMA*. 1993;269: 1537-1539.
20. Todd KH, Lee T, Hoffman JR. The effect of ethnicity on physician estimates of pain severity in patients with isolated extremity trauma. *JAMA*. 1994;271:925-928.
21. McCaig LF, Burt CW. *National Hospital Ambulatory Medical Care Survey: 2002 Emergency Department Summary: Advance Data from Vital and Health Statistics; No. 340*. Hyattsville, MD: National Center for Health Statistics; 2004.
22. Wolfe JM, Lein DY, Lenkoski K, et al. Analgesic administration to patients with an acute abdomen: a survey of emergency medicine physicians. *Am J Emerg Med*. 2000;18:250-253.
23. Graber MA, Ely JW, Clarke S, et al. Informed consent and general surgeons' attitudes toward the use of pain medication in the acute abdomen. *Am J Emerg Med*. 1999;17:113-116.
24. Lintula H, Kokki H, Vanamo K. Single-blind randomized clinical trial of laparoscopic versus open appendicectomy in children. *Br J Surg*. 2001;88:510-514.
25. Bowen ME, Poehling KA, Gunn VL, et al. Does the use of oxycodone affect diagnostic accuracy in children with acute abdominal pain? *Arch Pediatr Adolesc Med*. 2005;159:326-328.

IMAGES IN EMERGENCY MEDICINE

(continued from p. 150)

DIAGNOSIS:

Hemopneumopericardium. This process is recognized radiographically by an air-fluid level within the pericardium. Hemopneumopericardium is rare and most commonly observed after blunt trauma and in association with pneumothorax or pneumomediastinum. Pneumothorax may lead to pneumopericardium by a pleuropericardial connection through a pericardial tear. Interstitial air may also track along the perivascular planes of pulmonary vessels into the pericardium.

A nontension hemopneumopericardium may initially appear asymptomatic but can rapidly progress to a life-threatening cardiac tamponade. The presentation of a tension hemopneumopericardium may be delayed or clouded by hemorrhage, cardiac contusion, or tension pneumothorax.

This patient received a subxiphoid pericardial window and pericardial drainage in the operating room.

APPENDIX E1.

Limitations

Inferences that can be drawn from these data are limited in several important respects. These limitations can be divided, in order of importance, into threats to internal and external validity. We believe the greatest potential threat to the internal validity of our findings is our choice of clinically important diagnostic accuracy as the primary target endpoint. By choosing this outcome rather than a simple determination of whether or not the provisional and final diagnoses were identical, we may have added unnecessary complexity and some ambiguity to the study. Nevertheless, we elected to do this because it seemed clinically important to determine whether diagnostic discordance was simply an artifact of taxonomy that had no impact on patient outcome or was a reflection of significant diagnostic error, with implications for the patient's health status. For example, there were several patients with the provisional diagnosis of biliary tract disease, which was subsequently changed to biliary pancreatitis as laboratory tests, sonography, and CT results returned. Coding these cases as instances of diagnostic discordance did not seem clinically sensible. Thus, we independently determined whether a discordance in diagnostic nomenclature was clinically important or not and coded each case accordingly before unblinding the study.

In the interest of exploring what our findings might have shown, had we chosen the more straightforward but, we believe, less clinically relevant approach, we performed a post hoc analysis examining simply whether or not the provisional and final diagnoses were the same. We found a 68% accuracy among the morphine group versus a 47% accuracy among the placebo group. The difference in diagnostic accuracy of 21% (95% CI 6% to 36%) is similar to the 19% difference favoring intravenous morphine reported by Pace and Burk,[6] on the basis of whose work we chose the Δ of 20% for the effect size in our sample size calculation. Using the strategy we chose a priori, we not only failed to find a difference in diagnostic accuracy of 20% favoring morphine but also, because we did not design an equivalence or noninferiority trial, we cannot make unequivocal claims of comparable diagnostic accuracy under morphine and placebo. Rather, we can only note that our data are most consistent with this interpretation.

A second threat to internal validity was our decision to perform an efficacy analysis of our data rather than an intention-to-treat analysis. To examine the difference this choice might have made in our findings, we performed 4 post hoc intention-to-treat analyses under conditions of maximally and minimally divergent outcomes. This approach also serves to test the robustness of our findings by combining a sensitivity analysis with an intention-to-treat analysis. We accomplished this by assigning, in turn, combinations of accurate and erroneous diagnoses to those patients excluded from our efficacy analysis because of protocol violations or lost to follow-up. If we assume that all the missing patients in the morphine arm were correctly diagnosed and all the missing patients in the placebo arm were misdiagnosed, then the difference in diagnostic accuracy between the 2 groups would be 6%, which is only a 5% improvement over the 1% difference we found. If we assume the opposite, ie, that all the missing patients in the morphine arm were misdiagnosed and all the missing patients in the placebo arm were correctly diagnosed, the difference in diagnostic accuracy increases to only 2%. If one then assumes that all the missing patients in both groups were misdiagnosed, the difference in diagnostic accuracy increases from 1% to 4%. Finally, if one again assumes the opposite, ie, that all missing patients in both arms were correctly diagnosed, the difference in diagnostic accuracy drops to 0%. All 4 of these differences are bounded by 95% CIs that embrace 0%, indicating that, similar to our reported difference of 1%, none were statistically significant. Thus, it does not appear that performing an intention-to-treat analysis under the assumptions of all 4 possible extremes of alternate findings would have materially affected our conclusion that administration of morphine does not seem to interfere with clinically important diagnostic accuracy in acute abdominal pain.

A third threat to the internal validity of our findings is a consequence of the quirks of randomization. Our 2 largest final diagnostic categories of nonspecific abdominal pain and biliary tract disease were unevenly allocated to the 2 arms of the study (24/35 patients with nonspecific abdominal pain received placebo, and 17/20 individuals with biliary tract disease received morphine). We have no evidence that this represents selection bias caused by tampering with study packets but speculate that this constitutes an example of how large study cohorts must be to ensure a consistently balanced dichotomy of random allocation across multiple strata, particularly when the number of strata is large compared to the number of patients occupying the cells that compose each stratum.

Another limitation related to the above is that our sample size, which is quite small when placed in the context of the large number of causes of abdominal pain, precluded any meaningful planned subset analyses designed to explore an effect on diagnostic performance that might be associated with the patient's underlying condition.

An additional limitation, also linked to our sample size calculation, was a failure to adjust for the effect of clustering by physician provider. When we analyzed the degree of clustering by physician that actually occurred, there were so many physicians participating in the study that there was in fact very little clustering by provider (median number of patients per provider=2 [IQR 1 to 3]). Had significant clustering occurred, the effect of this would have been to decrease the precision of our 1% estimate of the difference in diagnostic discordance between the opioid and placebo groups by widening the CIs bounding this estimate (−11% to 12%).

Turning next to external validity, limitations that may have an adverse impact on generalizability include at least 4 additional considerations. First, patients were gathered by convenience sampling, determined by availability of research associates during a 2-year period. This, combined with substantial initial resistance to the administration of opioids in abdominal pain patients among a wide range of providers, resulted in a lengthy enrollment period. The study duration was then extended further by the necessity of following up many patients longer than the planned period of 6 weeks because the cause of abdominal pain in many individuals could not be determined with reasonable certainty until more time had elapsed. This nonconsecutive recruitment of subjects is likely to have caused some assembly bias, which threatens generalizability.

A second limitation related to external validity is spectrum bias,[17] reflected by the extremely high pain scores of our study cohort (median visual analog scale 99 and 98 mm in the morphine and placebo groups, respectively). Because patients in severe pain may be less likely to have important physical findings reduced below the threshold of examiner detection by opioid administration, generalization of our results to individuals with less severe abdominal pain may not be appropriate.

A third constriction on external validity is the relatively young age of the patients enrolled in the study. Despite the planned exclusion of individuals younger than 21 years, our study cohort consisted of patients with a mean age only in their midforties (Table 1). Because this is considerably younger than the age of our overall patient population and those presenting with abdominal pain, we suspect this may reflect a relative reluctance of physicians to administer opioids in the doses specified by study protocol (0.1 mg/kg, up to maximum of 10 mg intravenously) to older patients, perhaps because of a fear that they may be more likely to hypoventilate than younger individuals. Consequently, our findings cannot be generalized to the pediatric age group and should be applied only with caution to the geriatric age group.

A fourth potential threat to the generalizability of our data derives from the composition of our study population, which was composed principally of minorities (85% Latino and black). Interpretation of our findings may therefore need to be viewed, at least in part, in the context of previously described differences in pain perception[18] and treatment[19,20] associated with race and ethnicity.

Table E1. Frequency of final diagnoses, diagnostic imaging, and surgical intervention by randomization group.

	Morphine (N=78)				Control (N=75)			
Final Diagnosis	**Dx**	**CT**	**Ultrasonography**	**Op**	**Dx**	**CT**	**Ultrasonography**	**Op**
Acid-peptic disease*	2	2	1	0	4	1	1	0
Adnexal torsion*	1	0	1	1	1	1	1	1
Appendicitis*	7	7	1	7	5	5	2	5
Biliary tract disease*	17	12	17	16	3	1	3	3
Constipation*	1	1	0	0	0	0	0	0
Degenerating fibroid*	3	2	2	1	1	1	1	1
Diabetic dysautonomia†	0	0	0	0	1	1	0	0
Diverticulitis*	4	4	1	2	2	2	1	1
Endometriosis*	0	0	0	0	1	1	1	0
Gastroenteritis*	5	3	0	0	2	1	0	0
Hepatitis†	1	1	1	0	0	0	0	0
Herpes zoster†	1	0	0	0	0	0	0	0
Inguinal hernia (incarcerated)*	0	0	0	0	3	3	0	3
Lower lobe pneumonia†	1	1	0	0	0	0	0	0
Mesenteric ischemia*	1	1	0	0	0	0	0	0
Nonspecific abdominal pain*	11	8	4	0	24	18	10	3
Ovarian cyst (with/without torsion; ruptured/unruptured)*	4	2	3	1	6	6	6	3
Pancreatitis*	0	0	0	0	4	4	2	0
Pelvic inflammatory disease*	3	2	3	0	3	2	3	0
Polycystic spleen†	1	1	1	0	0	0	0	0
Regional enteritis*	0	0	0	0	1	1	0	0
Renal colic*	6	6	0	0	8	8	0	0
Retroperitoneal bleed (with Coumadin)†	0	0	0	0	1	1	0	0
Small bowel obstruction*	2	2	0	1	3	3	2	2
Spontaneous bacterial peritonitis*	1	1	1	0	0	0	0	0
Tubo-ovarian abscess*	1	1	1	0	0	0	0	0
Urinary tract infection*	6	6	0	0	1	0	0	0

CT, Abdominopelvic CT; *Dx*, diagnosis; *Op*, operative intervention.
*Initial menu of diagnostic choices included these entities plus abdominal aortic aneurysm, ischemic colitis, malignancy, and perforated viscus.
†Initial diagnostic choices entered as free text under "other."

Table E2. Clinically important instances of diagnostic error by randomization group.

Final Diagnosis	Provisional Diagnosis	Allocation
Acid peptic disease	Nonspecific abdominal pain	Morphine
Acid peptic disease	Nonspecific abdominal pain	Control
Appendicitis	Biliary tract disease	Control
Appendicitis	Nonspecific abdominal pain	Control
Biliary tract disease	Nonspecific abdominal pain	Morphine
Biliary tract disease	Nonspecific abdominal pain	Control
Degenerating fibroid	Diverticulitis	Morphine
Lower lobe pneumonia	Small bowel obstruction	Morphine
Ovarian cyst	Appendicitis	Morphine
Ovarian cyst	Nonspecific abdominal pain	Control
Ovarian cyst	Diverticulitis	Control
Ovarian cyst	Degenerating fibroid	Control
Pancreatitis	Acid peptic disease	Control
Pancreatitis	Constipation	Control
Pelvic inflammatory disease	Incarcerated inguinal hernia	Morphine
Pelvic inflammatory disease	Appendicitis	Morphine
Retroperitoneal bleed (with Coumadin)	Diverticulitis	Control
Small bowel obstruction	Gastroenteritis	Morphine
Small bowel obstruction	Nonspecific abdominal pain	Control
Tubo-ovarian abscess	Diverticulitis	Morphine
Urinary tract infection	Appendicitis	Morphine
Urinary tract infection	Nonspecific abdominal pain	Morphine

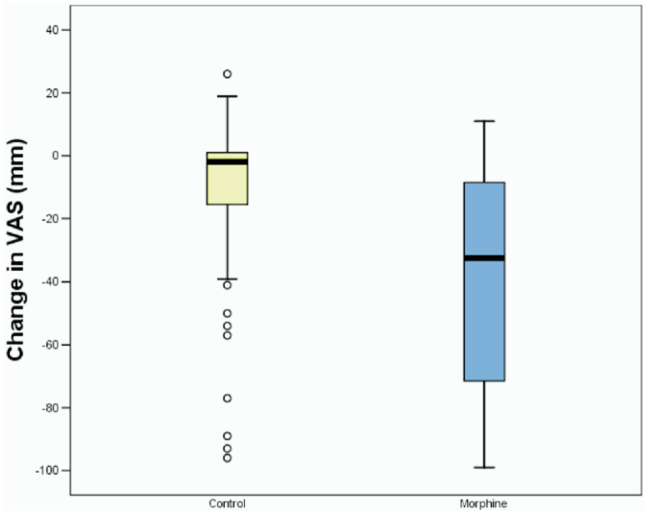

Figure E1. Box and whisker plot comparing changes in VAS pain scores in the morphine and placebo groups. Shaded boxes represent interquartile ranges (IQR). Dark horizontal line within each box is the median change in pain for each group (–2mm for the control group vs. –33mm for the morphine group).