INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI®
800-521-0600

Acute Abdominal Pain in the Emergency Department: Physicians' Use of Opioid Analgesics and the Incidence of Serious Outcomes

Jacques Simon Lee

Department of Epidemiology and Biostatistics

McGill University

August, 1997

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of Master of Science

Abstract

Physicians have traditionally withheld opioid analgesics from patients with acute abdominal pain due to concerns of masking physical findings. No study has examined morbidity and mortality after narcotic administration. The purpose of this study was to determine: 1) frequency of abdominal pain requiring narcotic analgesics, and 2) rate of serious outcomes, (death, infection, perforation, obstruction or hemorrhage of abdominal organs), in order to assess the feasibility of a randomized clinical trial on the safety of narcotics. Of 860 patients with acute abdominal pain, 477(55%) completed a pain questionnaire, and 321 met study criteria for need of narcotic analgesia (37.3%). Of these, 36 (11.2%) experienced a serious outcome as assessed by telephone contact 2 to 3 weeks after initial visit. The overall rate of serious outcomes was 67 of 860 (7.8%). A clinical trial using serious outcomes as the primary endpoint is possible, but would need to randomize approximately 3200 patients.

Acknowledgments

I would like to express my gratitude to Tracey Makira, Mi-Lihn Trahn, and Amanda Sukdheo for their assistance in the collection of data as well as the often tedious task of data entry. As well, my heart-felt thanks go to Kathy Vandemheen and Teressa Cacciote for their practical guidance in the formation of data collection strategies as well as their expert advise on data form design. Sylvia Visentin deserves special recognition for almost three years of assistance with the preparation of innumerable revisions of this document, as well as for her impeccable professionalism during the many deadline crises that arose. I would also like to express my gratitude to Elaine Parker who has also been extremely helpful since inheriting Ms. Visentin's duties.

To my colleagues at the Ottawa Civic Hospital Division of Emergency Medicine, my undying gratitude for collection of data, in perhaps the most difficult research environments in Medicine. I also thank the clerical and nursing staff for their part in promoting data collection.

I wish to thank Georges Grenier, Luc Lalonde and especially Mrs. Lise Lee for the French translation of the abstract, whose timely assistance helped turn one of the most intimidating aspects of this undertaking into an enjoyable learning experience.

My thanks also go to Dr. George Wells and Dr. Jim Hanley, for innumerable and invaluable hallway statistical consultations over the past three years.

I take great pleasure in thanking my thesis supervisor, Dr. Stan Shapiro and co-supervisor, Dr. Ian Stiell, for being learned educators and accomplished role-models . They have been generous with their scarce time and demonstrated extraordinary patience, while at

the same time curbing my natural tendency to wander off topic. Beyond the wealth of knowledge and the practical understanding they have imparted to me, Dr. Shapiro and Dr. Stiell have kindled an enthusiasm for the research process within me: I can think of no better or more long-lived gift.

And finally, for her constant assistance in clerical, editorial, methodological, statistical, motivational, and innumerable other ways, let me express my eternal gratitude to Ms. Lauren Dade. Without her help and caring, this work would never have been completed.

1. Introduction

New onset (acute) abdominal pain is one of the most common complaints seen in emergency departments, representing 4-10% of all emergency department visits.[1,2,3,4,5,6] Many different diseases can cause acute abdominal pain, ranging in severity from benign (viral gastro enteritis) to life-threatening (ruptured abdominal aneurysm). The physical examination is still one of the most important tools used by physicians to decide if acute abdominal pain is caused by a serious disease, such as appendicitis, or not. Classic medical teaching has warned that giving strong pain relievers, such as opioid analgesics, to patients with acute abdominal pain may obscure the physical signs of serious diseases,[7,8] resulting in misdiagnosis and subsequent mismanagement. For this reason, many patients with acute abdominal pain today do not receive analgesics until a diagnosis has been made. Due to delays in receiving blood test results and specialist consultations, patients often suffer with abdominal pain for hours before receiving analgesia.

Recently, the practice of withholding narcotic analgesics in acute abdominal pain has been criticized, both for being uncompassionate as well as for being based on anecdotal experience rather than experimental evidence.[9,10,11,12] Unfortunately, no evidence exists to refute the theoretical concern that analgesics are dangerous in acute abdominal pain, either.[13,14] A literature review revealed no randomized clinical trials that could conclude with convincing power that opioid analgesics were or were not safe in acute abdominal pain. This lack of evidence to guide clinical practice prompted plans for a randomized trial to determine whether opioid analgesics could safely be given in acute abdominal pain. However, insufficient data exist on this patient population to assess the feasibility of a randomized

clinical trial. Therefore, this project was undertaken to describe the acute abdominal pain patient population in detail, and thus provide information necessary for the planning of future randomized clinical trials on the safety of opioid analgesics in acute abdominal pain.

2. Objectives

2.1 Principal Objectives

To assess the feasibility of a future randomized clinical trial, two questions need to be answered. 1) Do adverse outcomes occur frequently enough that the safety of opioid analgesics could realistically be assessed, or would the rarity of serious outcomes preclude a randomized clinical trial using this endpoint ? 2) Is the number of patients with abdominal pain who require opioid analgesics sufficient to justify such a potentially expensive randomized clinical trial ? These two questions that form the basis of the principal objectives of this thesis, and are formally defined as follows:

1.1 To determine the proportion of patients with acute abdominal pain who experience *serious outcomes*, (death, ischemic bowel, peritonitis, intra-abdominal abscess, septic syndrome, or the obstruction, perforation or hemorrhage of an abdominal viscus), under the current strategy of withholding analgesics.

1.2 To determine the frequency of patients who present with acute abdominal pain severe enough to require the use of an opioid analgesic.

2.2 Secondary Objectives

The data collection process for the principal objectives presented an opportunity to answer several other questions regarding the acute abdominal pain population. These secondary objectives are listed below, followed by a brief description of the underlying rationale.

2.1 To examine the agreement between physicians and patients regarding the need for opioid analgesics.

It has been shown in burn units and post-operative recovery units that physicians and patients disagree regarding pain severity.[15,16] Does the same hold true of emergency physicians and abdominal pain patients ? The frequency of patients requiring narcotic analgesics will vary depending on the definition of narcotic requirement used, therefore patient-physician agreement on need for analgesics was assessed.

2.2 To examine the inter-physician agreement on the physical examination of the abdomen.

At least three previous studies have used change in physical exam as an endpoint for assessing narcotic safety. However, the baseline intra and inter-observer variability of the abdominal exam is not well described. This information would be important in planning future randomized clinical trials if change in abdominal exam after narcotics was to be used as an outcome measure.

2.3 To examine the relationship between opioid analgesic use and the incidence of serious outcomes.

The data collection strategy provided a prospectively assembled cohort of patients with acute abdominal pain, some of whom would or would not receive narcotic analgesic. This allowed the preliminary investigation of the impact of narcotic analgesics on the acute abdominal pain population.

2.4 To estimate the baseline rate of positive and negative laparotomy among acute abdominal pain patients initially discharged from the ED, as well as among those admitted.

Negative findings at operation are another potential outcome measure, and the baseline rate among patients admitted from the Emergency Department is not well described

2.5 To estimate the baseline rate of return visits to the ED and subsequent hospitalization among acute abdominal pain patients initially discharged from the ED.

Return visits and hospitalization rates are also potential outcome measure for future randomized clinical trials.

2.6 To estimate the dose, route, and specific type of narcotic analgesics ordered by physicians in abdominal pain.

The current narcotic practice patterns of emergency physicians in acute abdominal pain is not described, and would be important in selecting a specific narcotic agent for the proposed clinical trial.

2.7 To estimate how often physicians withhold analgesia because of concerns that analgesics may mask the diagnosis in patients with acute abdominal pain.

This objective was included in order to assess the magnitude of the problem created by the current uncertainty regarding narcotic analgesic safety.

2.8 To determine the number of patients lost to follow-up during the study.

This objective would be used as an estimate of the loss-to-follow-up rate for future randomized clinical trials .

3. Literature Review

3.1 General Literature Search Strategy

The literature review was divided into five specific content areas: the frequency of acute abdominal pain, the safety of analgesic use in acute abdominal pain, the reliability of the abdominal exam, the measurement of pain, and the rate of serious outcomes in acute abdominal pain. A similar strategy was used to search the literature in each content area. Specific terms were searched for in the computer-based Medline database from 1966 to 1995 under the "Subject" fields for each literature search using OVID software.[17] The search was then repeated using the same terms under the "Title" field. This created an initial list of candidate articles which were manually screened and any pertinent articles where then retrieved. Subsequently, relevant secondary references cited in these articles were also retrieved. Finally, content experts at McGill University and the University of Ottawa were asked to provide any references pertinent to specific areas to verify that important articles had not been missed.

3.2 Frequency of Abdominal Pain: Potential magnitude of the problem

A literature search was conducted. The specific terms used to search for articles that referenced the frequency of patients with abdominal pain severe enough to require

abdominal analgesics were "Abdominal pain" and "Emergency Department" or "Emergency Health Services" or "Emergency Medicine". In addition, several articles found in searches of the other content areas made reference to the proportion of patients presenting to the Emergency Department with acute abdominal pain.

From this search, nine studies were found that made reference to the proportion of patients presenting to an emergency department with a principal complaint of abdominal pain.[1,2,3,4,5,6,18,19,20] No studies were found that examined the proportion of patients requiring analgesics. Acute abdominal pain was the principal complaint in 4 to 10% of patients presenting to an emergency department.[1,2,3,4,5,6,18,19,20]

Additionally, a pilot study conducted by Dr. J.V. Quinn at the Ottawa General Hospital, with an Emergency Department census of 50,000 visits per year, showed that 9.5% of all emergency department patients had a primary complaint of abdominal pain (personal communication). Of these, 50% were found to have pain requiring further investigation. It is unknown what proportion of these patients required narcotic analgesia.

Extrapolating crudely from U.S. figures of 90 million emergency department visits per year,[3] one could estimate 9 million emergency department visits per year in Canada (10% of the U.S. population). Therefore, patients with abdominal pain may total 360,000 and 900,000 per year in Canada (4-10% of 9 million). However, data exists that the annual rate of emergency department visits in Canada is higher than in the United States.[5,21] In 1990, there were 640 visits per 1000 population in Canada, versus 349/1000 visits in the U.S. .[21] Using these figures, the rate of Emergency Department visits may be as high as 14 million,

(640/1000 visits, based on a population of 22 million), with a consequent 560,000 to 1.4 million visit annually due to abdominal pain.

No data were found concerning the proportion of patients with abdominal pain who required narcotic analgesics. Informal polling of emergency physicians suggested that 30% was a reasonable estimate. Assuming 30% of the 360,00 to 900,000 patients with abdominal pain require narcotic analgesics, 108,000 to 270,000 Canadians per year currently may be suffering unnecessarily with severe pain while awaiting diagnosis. The results of the current study will more precisely define the magnitude of untreated abdominal pain.

3.3 Safety of Narcotics in Acute Abdominal Pain

Debate currently exists over the safety of narcotic analgesic use for acute abdominal pain prior to establishing a definitive diagnosis in the emergency department. It is now standard clinical practice to withhold analgesia.[7,8,22,23] The present uncertainty reflects two conflicting theoretical viewpoints. On the one hand, it is argued that analgesia may mask the physical findings of serious underlying pathology, and thus result in delayed or incorrect diagnosis. In the case of appendicitis, such delay can lead to perforation of the appendix, with a concomitant higher mortality rate. Although a consensus of expert opinion favouring the withholding of analgesia has existed,[7,8] it is based on anecdotal experience rather than research. On the other side of the debate is the position that "urgent relief of severe pain is good treatment, humane and unlikely, to say the least, to do harm nowadays by delaying diagnosis".[9] Unfortunately, there is no convincing body of literature to support this later argument either.[24,25]

3.3.1 Previous Studies on the Safety of Narcotics in Acute Abdominal Pain

A literature search of Medline references from 1966 to 1995 using the terms "Pain" and "Abdomen" and "Analgesia" under the "subject" field was conducted. Next, "Acute Abdomen" and "Analgesia", as well as "Appendicitis" and "Analgesia", were targeted. The above searches were then repeated under the "text" field. Only two randomized clinical trials,[24,25] one abstract[26] and one case-control study[27] were found that addressed the use of analgesia in abdominal pain.

All of these studies had methodologic weaknesses. The case-control study examined risk factors among a series of misdiagnosed appendicitis patients that had proceeded to litigation in the U.S.[27]. Administration of a narcotic occurred more often among patients initially misdiagnosed and discharged (37/66) compared to controls in whom the correct diagnosis was made (15/66) ($p < 0.001$). Selection bias may limit the generalizability of the result of this study, as analgesia use among patients involved in litigation is not likely to be representative of abdominal pain in general. Because of the case-control study design, the possibility of a reverse-causality bias can not be excluded: analgesics may have been given more often to patients who had already been misdiagnosed, rather than the analgesics causing such misdiagnosis. A randomized clinical trial would have been a more valid design to test whether analgesics *caused* the misdiagnosis.

Attard et al.[24] examined one hundred patients with acute abdominal pain "severe enough to warrant opiate analgesia" among patients referred to a surgical consultant at a teaching hospital in the United Kingdom by their general practitioner. Patients were randomized to receive an intramuscular injection of either a saline placebo or up to 20 mg of

papaveritum (equivalent to 12.5 mg of morphine).[28] Several outcome measures were reported: pain and tenderness before and after drug, surgeons confidence in diagnosis, and number of incorrect diagnoses. Inaccurate diagnosis were made in 2/50 of the papaveritum group and 9/50 in the saline group ($p < 0.05$). The authors stated that an "editorial in the BMJ 13 years ago recommended early pain relief in the management of acute abdominal pain. This study provides the scientific data to justify this recommendation." [24]

Methodologic concerns with the Attard study unfortunately make such a conclusion questionable. Their study design prevents the unqualified acceptance of the safety of analgesic use in acute abdominal pain, and the authors own data actually draw into question the safety of this practice. First, they use misdiagnosis rate as a proxy measure for demonstrating the safety of placebo versus analgesic. Adverse outcomes would have been a more convincing measure of relative safety than misdiagnosis, being a direct measure of safety. Although the authors noted no complications among their 100 patients, there is no comment made on the study's power to detect such complications. If complications occur at a rate of 1% of cases, this study would have less than 50% power to detect a 100% relative difference between the two treatment groups.[29]

Second, the manner in which subjects were classified as misdiagnosed raises concerns about the use of this outcome as a measure of safety. "Misdiagnosis" can be divided into two subgroups: those "under diagnosed" with a less serious diagnosis than was actually present, and those "over diagnosed" as having a less serious condition. The two patients classified as misdiagnosed in the analgesic groups were initially diagnosed as non-specific abdominal pain and observed. Both required a subsequent operation, where a final diagnosis of appendicitis

was made. Thus both the patients in the analgesic group were initially under-diagnosed. By contrast, in the placebo group none of the nine patients classified as misdiagnosed subsequently required surgery (i.e., none were *under* diagnosed). Six of the nine patients operated on in the placebo group turned out to have no intra-abdominal pathology (six patients were "over diagnosed"). The authors concluded from this that "analgesics seemed to facilitate accurate diagnosis and management".

However, it may be more informative to conclude that there was a higher rate of under-diagnosis in the analgesic group, and a higher rate of over-diagnosis in the placebo group, rather than stating that overall accuracy was better in either group. In attempting to determine if analgesic are safe in acute abdominal pain based on their data, the relevant question then is whether under-diagnosis is safer than over-diagnosis.

In acute abdominal pain, under diagnosis and subsequent delay of operative treatment carries much higher risk of complication and death than the risks from over diagnosis and over treatment with a negative exploratory surgery.[30] For example, failure to operate urgently in appendicitis can result in perforation, with consequently higher morbidity and mortality. Wound infections are three times more common in perforated appendicitis [33,34,35] with mortality rates 6 to 50 times higher after perforation.[30,36,35] Similarly, mortality from a strangulated bowel obstruction increases from 8% for early operation to 25% if delayed beyond 36 hours.[30] Conversely, the risk of morbidity from negative exploratory laparotomy is in the order of 10% with mortality of less than 1/5000.[37]

When viewed in this light, Attard et al's study shows a trend towards <u>increased</u> "under diagnosis" and thus <u>worse safety</u> in the analgesic group ($p = 0.051$). By combining

under and over diagnosis into the composite measure of misdiagnosis, important differences in the safety of analgesia and placebo may have been obscured. The authors do address the higher rate of under-treatment among the analgesic group in their discussion. They conclude that this is because the physical findings of appendicitis are delayed in certain patients, as opposed to "a significant problem that implies masking of symptoms by opiates". No explanation is given for the preponderance of such delayed presentations among the analgesic group.

A third problem with the conclusion of narcotic safety in acute abdominal pain is that all patients in the study were admitted to hospital. Of particular concern is what might have happened to the two patients with undiagnosed appendicitis if they had been discharged. Opiod safety among patients who might potentially be discharged from hospital is not address by this study, and thus the generalizability of these results to the practice of emergency medicine is limited.

Finally, the authors note that "although every attempt was made to maintain the double blind nature of the trial, clear differences in the patient's comfort resulted in an accurate assessment by the registrar of the treatment group each patient belonged to." Although a difficult problem to overcome, the lack of blinding may have led to differential treatment of patients".[24] Increased vigilance among the analgesic group may have resulted in fewer misdiagnoses, or more misdiagnoses among the placebo group.

The second pertinent clinical trial by Zoltie and Cust[25] was conducted on 288 consecutive patients admitted to the General Surgery department at a British teaching hospital during 1983. This trial compared two doses of a new sublingual analgesic,

buprenorphine, to placebo tablets. Unfortunately, there was no difference in pain scores between either dosage groups and the placebo groups, therefore it is difficult to assert that these patients received adequate analgesia. In addition, no inferential statistical tests were performed, nor did the published data permit a secondary analysis of the statistical significance of their results. Therefore, this study has added little to our understanding of analgesia in acute abdominal pain.

The third clinical trial by Burke et al.,[26] published in abstract form, reported 71 patients randomized to receive intravenous morphine or placebo. This study also used agreement between initial and final diagnosis as an endpoint, and found no difference between groups. The power of the study to conclude no difference was not reported in the abstract, nor were confidence intervals reported. In addition, the authors state that "Three diagnostic or mismanagement errors occurred in each group." (3/35 in the morphine group and 3/36 in the placebo group). However, these errors were not described in the abstract, nor was any comment made on the whether these errors resulted in adverse outcomes. It is unknown whether under and over diagnosis were grouped together as "errors" in this study.

3.4 Pain: Its Control, Theory, and Measurement

3.4.1 Pain Control in the Emergency Department: Increasing recognition

There has been a growing concern in recent years that pain has been under treated by physicians in general,[15,16,38] and in the Emergency Department specifically.[39,40,41,42,43,44] The primary duties of a physician are to prevent untimely death and alleviate suffering. And while much progress has been made in the ability to save lives, there are many who feel physicians have been much less successful in alleviating pain[45,46,47]: "numerous studies have shown that

pain and suffering are commonly ignored or under treated in most specialties of medicine"[41]. Of patients with acutely painful conditions treated at one university teaching center, 56% received no analgesia. Of those who were given analgesia, two thirds waited one hour while 42% waited two hours before receiving medications. [40]

Failure to adequately assess the patient's pain, fear of complications, and fear of being duped by "drug seeking" patients are given as a reasons for withholding analgesia [41]. In contrast, Wilson et al have noted "although drug abuse is a major concern in this country, we need to draw our attention to another type of drug abuse - the failure to treat patients in severe pain with adequate doses of narcotic analgesics.".[40]

Given the changing attitudes towards pain management, and despite the lack of evidence as to the safety of analgesics, it is not surprising that analgesic use in acute abdominal pain has been increasingly advocated.[41,45,48,49] Indeed, two recent editions of standard surgical texts now advocate the use of narcotic analgesia for acute abdominal pain[30,50]. Letters and editorials can be found in the literature favoring analgesia in undiagnosed abdominal pain,[48,49,50] even though only the two published trials discussed above have studied this issue.[24,25] Despite the methodologic difficulties described earlier, these articles are being used to support the use of analgesics as safe. A recent review article on the evaluation of abdominal pain cited the study of Zoltie et al. ,[25] as evidence that analgesia makes "no difference in diagnosis".[51]

3.4.2 Pain: Theory and Measurement Methodology

One of the main objectives of this thesis was to measure the incidence of abdominal pain severe enough to require an opioid analgesic. However, it is known that physicians and

patients often disagree regarding whether an analgesic is "required" or not.[15,16,38,39,40,41,42,43,44] The source of such disagreement is important and not well described .. Specifically, is physician-patient disagreement regarding the need for analgesics due to differences in perception of pain severity ? Or is disagreement due to physicians being more concerned about complications from narcotics, or masking the etiology of the acute abdominal pain. Thus, an assessment of pain severity was included in this protocol. On reviewing the literature, it was soon discovered that the measurement of pain presents unique challenges: a universally accepted gold standard does not exist. Furthermore, there is controversy regarding the fundamental nature of pain, with resultant controversy regarding its measurement. Thus an introductory discussion of the literature pertaining to pain and its measurement would be helpful in evaluating the methodology chosen in this thesis.

Webster's dictionary defines pain as "bodily suffering or distress, as due to injury, illness, etc."[52] This definition reflects an intuitive understanding of pain, and has been labeled the "sensory" definition of pain. Since 1965,[53] however, pain has been understood as a more complex interaction of sensory input modulated by nerve impulses descending from the brain and integrated at the level of the spinal cord. In this framework, originally termed the gate-control theory,[53] pain does not bear a one to one relationship to underlying tissue damage. Rather, pain is seen as varying according to the genetics, culture, psychology and experience of the individual.[54,55,56,57] For example, when having blood tests drawn certain individuals experience dramatic fear or panic, accompanied by profound changes in their physiologic state, and report experiencing intolerable pain if they permit the test at all. Other individuals may experience only a transient apprehension, show no physical manifestation, and report

feeling minimal pain on venipuncture from the exact same type of needle. Neurophysiologic differences between individuals are believed to generate different amounts of pain from a uniform stimulus.[55]

Melzack and Walls' theory of pain has important implications for pain measurement: assessing the intensity of a noxious stimulus can not be used as a proxy for the measurement of pain. Furthermore, any assumptions regarding the severity of pain being correlated to external signs of disease intensity are also tenuous. The pain experienced by two patients with appendicitis may differ significantly, due to genetics, culture or past experience. However, little data exists that would allow us to predict which individual would experience more or less pain. Consequently, patients' self-reports of pain have gained acceptance as the most valid method for pain measurement. It is recognized that such reports are subject to bias themselves: deliberate falsification by the respondent, and "reactivity" or the influence of the act of measurement on the resultant assessment of pain, for example. However, given the lack of a superior alternative, patients' self-reports are now generally accepted as the best available measurement[56] for pain.

Second, the concepts of acute and chronic pain should be differentiated. Whereas "acute pain is a transient, continuously changing state that differs radically from normal life; it is intimately related to intense emotional arousal, it is linked to tissue pathology, and it is usually characterized by clear, well focused sensory characteristics. Chronic pain, in contrast, is an enduring condition that has become a stable element in the daily life of the patient."[58] The focus of this study was on acute pain, as such pain is more likely to represent a rapidly

progressive pathological process, for which timely diagnosis and treatment are of more importance.

A third aspect of pain relevant to this discussion is its multi-dimensionality. Melzack states : "It is now evident that the word 'pain' refers to an endless variety of qualities that are categorized under a single linguistic label..." [55]. Melzack and Torgerson[59] originally proposed three main categories or dimensions to describe pain: an affective dimension, a pain intensity or evaluative dimension, and a sensory qualities dimension. Other domains have since been proposed including temporal course and the impact of pain on behaviour[60]. Many different scales have been developed to measure pain, both uni and multi-dimensional, and the choice of measurement tool is dependent on the specific pain dimension or dimensions of interest. In the present investigation the major focus was whether patients had pain of sufficient severity that a narcotic analgesic was required. It has been suggested that the distinction between the affective and evaluative dimensions becomes less significant in acute pain[61,62]. In addition, multi-dimensional pain scales take longer to administer than uni-dimensional scales, and require more training of the tester. Thus, the use of a uni-dimensional pain measurement tool of pain severity seemed appropriate.

3.4.3 Pain Measurement Scales

The two most widely used uni-dimensional pain intensity measurements are verbal rating scales,[59,63] and visual analog scales.[64,65,66,67,68] Both scale types have received some criticism in the literature; the relative strengths and weaknesses of each will be discussed below.

3.4.3.1 Verbal Rating Scale

The verbal rating scale consist of asking a patient to rate their pain as "none", "mild", "moderate", or "severe". Most verbal rating scales have decreased responsiveness to changes in pain severity when compared to the visual analog scale, due to the small number of categories available.[69,70,71,72] Increasing the number of categories confers additional problem, as additional descriptive categories may not necessarily increase responsiveness over the range of pain severity. For example, patients may not be able to differentiate "horrible" from "excruciating" pain. One of the major strengths of the verbal rating scale is its ease of use and familiarity for patients

3.4.3.2 The Visual Analog Scale

The visual analog scale is often described as a graphic representation of pain severity. In general, it consists of a line with descriptors at each end such as " No Pain" and "Unbearable pain" representing a spectrum of pain intensity. Patients are asked to indicate the severity of their pain by marking somewhere along this scale. The specific form of the visual analog scale chosen has been shown to have bearing on its measurement properties. Horizontal lines produce more uniform distribution of scores than vertical scales.[67,7273] Vertical stops at each extreme reduce marks made off of the scale. The line should not have intermediate marks or descriptors along its length as these cause artificial clustering of scores.[67,68,70] Compared to verbal rating scales, visual analog scales have increased sensitivity to change in pain state in response to treatment.[67,69]

3.5 Reliability of The Abdominal Examination

An issue fundamental to the perceived danger of narcotic use in acute abdominal pain is the concern that such analgesics will mask the physical findings of serious conditions and thus alter diagnostic accuracy. To this end, clinical trials reviewed in section 3.3.1 included analyses on the effect of narcotics on the reproducibility of the physical exam. However, as discussed in the following section, very little data exists on the baseline reliability of the abdominal exam: to what degree do physician reproduce the same clinical finding when they re-examine a patient. Such information is important in order to interpret the significance of any effect analgesics may have on the abdominal exam, and thus was included as a secondary study objective (see section 2.2). The literature pertinent to the reliability of the abdominal exam will be reviewed below.

3.5.1 Previous Studies

A Medline search strategy similar to that outlined above for the terms "acute abdomen" or "abdominal pain" and "physical examination", found one clinical trial[74] and one abstract.[75] The clinical trial by Bjerregaard was conducted at a Danish teaching hospital and studied 40 admitted patients, each examined by the same four physicians. The inter-observer agreement for physical signs was poor; for example, the kappa values for "tenderness, direct" was 0.31. In the abstract by Greene et al kappa's of 0.45 for superficial tenderness, 0.39 for percussion tenderness, and 0.38 for abdominal rigidity, were reported. The authors of the second study did however find excellent agreement for the physical finding of "tenderness to deep palpation" with a kappa of 1.0. Thus, agreement for most abdominal physical findings has been poor. Neither study reported intra-observer reliability, nor reliability in the presence of analgesia, nor was location of tenderness commented on.

3.6 Rate of Serious Outcomes and the Feasibility of a Randomized Clinical Trial

One of the principal motivations for the current project was to determine whether a randomized clinical trial into the safety of narcotic analgesics would be feasible at the Ottawa Civic Hospital. Specifically, the baseline rate of serious outcomes was needed to perform sample size calculations for such a randomized clinical trial given that the serious outcome rate was to be used as the primary outcome. In addition, the frequency of patients presenting with acute abdominal pain requiring narcotics needed to be known to assess if such a study could be completed in a reasonable time frame Thus, the literature pertaining to the presentation of acute abdominal pain in the emergency department was reviewed and is discussed below.

Four articles were found which described the natural history of acute abdominal pain [1,2,18,19]. Stansiland et al [18] examined 600 patients both retrospectively and prospectively that were admitted to two hospitals in Leeds, England in 1972. The proportion collected retrospectively is not mentioned. Patients admitted from the emergency department with abdominal pain of less than 1 weeks duration were eligible, if records allowed abstraction of 35 variables. The authors comment that "the proportion of case notes which were not adequate on the grounds of incompleteness was surprisingly small".[18] No direct comment is made regarding the use of analgesics by the authors, but they did remark on pain severity. It is probable that the physician's impression of pain severity was reported, as this data was abstracted from medical charts. Unfortunately, the data were presented in bar graph format only, thus the following results are approximate: 39% of patients with appendicitis (n=100)

had severe pain, 30% of diverticulitis cases(n=100), 93% of perforated peptic ulcers patients (n=100), 18% of non-specific pain cases(n=100), 65% of cholecystitis (n=100), 60% of small bowel obstructions (n=50) and 65% of pancreatitis (n=50). Thus overall, approximately 31.5% of their patients overall had "severe pain". What proportion of these would require analgesics is unknown. From the final diagnoses reported, at least 150/600 (25%) of their patients had serious outcomes as defined in the current study (perforation or small bowel obstruction). However, since all their patients had been admitted, these results are not applicable to the patient population of interest for this project.

In a study conducted at an American teaching hospital emergency room, Brewer et al. 19 reviewed the charts of 1,000 consecutive patients with abdominal pain seen between 1971 and 1972. Abdominal pain represented 5% of their total emergency census, and 274 of these patients were admitted (27.4%). Of these, 150/274 patients (65%) underwent surgery, with a negative laparotomy rate of 20/150 (13%). Eleven patients (1.1%) initially discharged from the Emergency Department returned to their center with an acute surgical condition: 8 patients with a final diagnosis of appendicitis, and 3 patients with small bowel obstruction. One patient appeared to have suffered a complication from having been discharged: a patient who returned with a ruptured appendix required prolonged hospital admission. No other comment is made on serious outcome rate, nor do the authors comment on narcotic analgesic requirements.

A study by Bugliosi et al[20]of emergency patients at the Mayo clinic conducted from 1988-1989 focused on elderly patients. In this retrospective chart review, 127 patients from an annual census of 55,000 (0.2%) had abdominal pain for less than 1 week and were older

than 65 years of age. Overall 80/127 (62%) were admitted on initial presentation. Five out of the 127 (4%) patients were initially discharged and subsequently returned to the Emergency, and 2 of these underwent an operation. Overall, 53/127 had an operation (54 %), although the negative laparotomy rate was not reported. Final diagnosis revealed 9 patients with a perforated viscus and 13 patients with small bowel obstruction, 4 patients with an incarcerated hernia, 1 patient with a colonic obstruction due to cancer, and 1 patient with a sigmoid volvulus. Thus, overall 28/127 (21%) experienced serious outcomes. Due to the age restrictions, these results are also not applicable to our patient population. Again, no mention is made of analgesic use in this study.

In another retrospective review of emergency department patients,[1] 2401 of 44812 presented with abdominal pain (5.4%). The admission rate was 12.6% (304/2401), and 113 patients underwent appendectomy(4.7%), with a negative operative rate of 2% (2/113). Eleven patients initially discharged returned with a diagnosis of acute appendicitis (0.4% of all abdominal pain patients). Other diagnoses or complications were not reported, nor were analgesic use patterns commented on.

Finally, in the study which most parallels the patient population of interest to the current study, Lukens et al.[2] prospectively followed all adult patients discharged with a diagnosis of undifferentiated abdominal pain by telephone at day 2-3 and again after two weeks over a 12 month period. They were able to contact 307 of 403 eligible patients (76.1%). A total of 74 patients reported that they had subsequently seen another physician within 3 weeks (18%), and 14 patients were subsequently admitted (3.5%). Of these patients, two serious outcomes occurred (0.5%): one patient had a small bowel obstruction secondary

to an incarcerated hernia, and one patient had a ruptured ovarian cyst. This low rate of complications among discharged patients may not be generalizable to the current study as only patients with undifferentiated abdominal pain were included. Patients with a clinically based but unproved diagnoses such as gastroenteritis who returned with complications would not have been picked up by this study. As well, the loss to follow-up rate of 24% could be significant, if this group experienced a higher rate of complications. No mention of narcotic requirements is made in this study.

Thus, the literature review revealed insufficient data to establish the feasibility of a randomized clinical trial into the safety of narcotics in acute abdominal pain. Specifically, no information on the frequency of the requirement for analgesic was found, nor could a baseline rate for serious outcomes be established in the patient population of interest.

4. Methods

4.1 Project Funding

This project was funded via a peer reviewed competitive grant for pilot studies by the Emergency Health Services branch of the Ontario Ministry of Health.

4.2 General Study Design and Rationale

An observational cohort study design was employed. All patients with abdominal pain presenting to an Emergency Department were assessed for inclusion, and those meeting the eligibility criteria specified in section **4.5** were followed by a research assistant to see if they experienced outcome measures, defined in section 4.9.1, between 21 and 28 days after discharge from the Emergency Department. In addition, emergency physicians were asked to complete a questionnaire to verify patient eligibility, and assess pain severity and need for analgesia. Prospective data collection was chosen instead of retrospective to avoid over-estimating the serious outcome rate. As medical record documentation is more extensive for patients experiencing serious outcomes, retrospective data collection is more likely to find patients with serious outcomes and miss those without. As well, it was felt that physicians prospectively assessing eligibility would produce a more realistic estimate of eligible patient rates for any future trial compared to a retrospective chart review. A retrospective estimate would not have accounted for physicians' interpretation of the eligibility criteria, nor the extent to which physicians were willing to enroll patients. Prospective data collection was projected to be approximately equivalent in cost to a retrospective chart review. As well, this study was modeled after several previous successfully completed studies that had yielded good co-operation by emergency physicians and ancillary staff.[76,77,78]

4.3 Setting

This study was conducted at the Ottawa Civic hospital, an academic teaching center affiliated with the University of Ottawa Faculty of Medicine. The Ottawa Civic hospital is active in the Canadian College of Family Physicians residency training programs in emergency medicine. The Emergency Department is supported by a full complement of surgical, medical and intensive care facilities and sees an average of 55,000 patients per year. During the past five years, more than 2000 patients had already been enrolled at this Emergency Department in a series of clinical studies similar in design to the current study. The data collection steps of this study were closely modeled on these successfully completed investigations,[76,77,78] and required essentially the same degree of participation by the Emergency Department staff. As well as being a teaching hospital, the Ottawa Civic Emergency Department serves the community as a primary care facility: over 80% of patients are self-referred[79]. Thus, it was felt that the patient population seen at the Ottawa Civic Hospital was representative of those seen at other urban community and teaching hospitals within Canada and the results of this study are therefore expected to be generalizable to the practice of emergency medicine in Canada.

4.4 Ethical Considerations

The research ethics committee of the Ottawa Civic Hospital granted approval of this study protocol without requiring additional consent of the patient beyond that implied by the agreement to respond to questionnaire items. Normal patient management was not altered. Patients were not subjected to any new therapy, invasive procedure, undue risk nor

discomfort beyond those which would normally be experienced in the course of standard patient care. There was no randomization of patients in this study. Strict patient confidentiality was maintained at all times.

4.5 Study Population: Eligibility Criteria

4.5.1 Rationale for Inclusion Criteria

The eligibility criteria of the present study were designed to capture a sample from a population that could be eligible for a future randomized clinical trial on the safety of narcotic analgesics in acute abdominal pain. Thus, eligibility criteria were selected in anticipation of the requirements of such a study. Any patients in whom narcotics clearly represented an undue risk, independent of any diagnostic uncertainty introduced by narcotics were excluded ("unsafe" in **Table 1** below). Patients with allergies to narcotics or who were hemodynamically unstable are examples of this principle. In addition, those patients with a sufficiently clear diagnosis that physicians would not normally withhold analgesics were also excluded ("diagnosis clear"): it would be unethical to withhold narcotics from such patients in clinical practice or within a randomized clinical trial. The inclusion criteria were selected to capture a patient population in whom the controversy regarding the safety of narcotic analgesics was most acute; those patients in whom the need for opioid analgesics was the greatest, but in whom the risks of masking the diagnosis were also high. Thus, inclusion was restricted to those patients with a primary complaint of abdominal pain of less than 3 days duration. Emergency physicians have the greatest difficulty deciding whether to order analgesics for such patients; their pain tends to be more severe, and the possibility of a serious outcome is greater in acute abdominal pain.

Patients presenting to the participating emergency department with a primary complaint of acute abdominal pain and who were assessed by the emergency physician on duty were eligible for inclusion into this study. Patients referred directly to a specific consultant service were not included. The emergency physicians were asked to complete a check list of inclusion and exclusion criteria on all patients. In those cases where the emergency physician did not assess eligibility at the time of presentation, inclusion was assessed by the research assistant or principal investigator, based on the emergency department chart. Eligibility criteria included the following:

1. Abdominal pain as the primary complaint on presentation to the emergency department as assessed by the emergency physician.
2. Onset of pain within *3 days* of presentation to the emergency department.

Following is a table listing the exclusion criteria, their operational definitions, and the rationale behind each.

Table 1 : Definition of Exclusion Criteria and Rationale for Each

Exclusion Criteria	Definition	Rationale
Age < 16 years of age		Different spectrum of disease, Consent issues
Recurrent or chronic pain	Any patient with history of prior occurrence of same pain	Not target population
Diagnosis already established	Patients referred from another physician directly to an admitting service.	Diagnosis clear
Previous enrolment in study		Excess influence of repeated measures in same patient
Pregnant patients	LMP [1] >4 weeks by history or urine β-HCG [2] +	Unsafe
Traumatic Cause of Pain	Any MVA [3], assault or fall victim	Unsafe and Diagnosis clear
Other distracting pain	Any other concurrent pain reported by patient outside the abdomen	Not target population
Hemodynamic instability	SBP < 90, HR > 125 or in the judgment of the EP	Unsafe
Renal Colic	Typical history in the judgment of the EP with + hematuria	Diagnosis clear
Peritoneal Dialysis	Currently Receiving CAPD (Presumed peritonitis)	Diagnosis clear
Suspected abdominal aneurysm	In the judgment of the EP	Unsafe
Allergy to narcotics	According to patient or chart.	Unsafe

[1] LMP: Last Menstrual Period
[2] β-HCG: Beta Human Chorionic Gonadotropin Test
[3] MVA: Motor Vehicle Accident

4.6 Data Collection

Two main strategies were employed for data collection in this project: chart reviews on all patients presenting to the Emergency Department from October 2, 1995, to April 30, 1996, and questionnaire data collection on a sub-sample of patients for whom the physician was able to complete the questionnaire. Prior to initiation of this study, a research assistant and all attending physicians were familiarized with the study protocol, including all exclusion

and inclusion criteria, via two 30 minute formal training sessions. Several hypothetical case scenarios were presented to assess participant's understanding of their role in the study. In addition, posters with pertinent study details were mounted in prominent locations within the Emergency Department. Finally, physicians and clerical staff participated in a run-in period of two weeks. During this time, all aspects of the study protocol were tested for unforeseen difficulties.

4.6.1 Data Collection: Prospective Questionnaire data

All patients presenting to the emergency department routinely have their telephone number, address, next of kin phone number, and family physician's names verified by the registration clerk. A data sheet on brightly colored paper was attached by the clerks to the medical chart of any patient complaining of abdominal pain. This reminded physicians that the patient was a candidate for the study, and physicians then determined if the patient meet all inclusion and exclusion criteria by completing a check list (see Error! Reference source not found.). Blank study sheets were available through the department for physicians if no sheet had been attached by the clerk. If the patient was eligible, and the treating physician had sufficient time, they were to complete the "Physician pain assessment & Diagnosis" section of the questionnaire, which included a visual analog scale (VAS) assessing the Physician's perception of the severity of pain experienced by the patient, during initial history and physical exam (see Error! Reference source not found.).

The visual analog scales followed standard methodology. [64] All were 100 mm in length, ran horizontally, had end anchors, and were bounded by the terms "No pain" at the left-hand side and "Unbearable Pain" on the right-hand side. After the Physician rated the

patient's pain, the patient was asked to indicate their perception of pain severity on an identical VAS (see Error! Reference source not found.). To allow independent assessment, physicians were instructed to fold the data sheet prior to administering VAS to the patient. Physicians were instructed to re-examine all patients who remained in the emergency department after one hour, and especially those receiving narcotic analgesia (see **Figure 2**). The purpose of reassessment was to determine whether pain severity and diagnostic impression changed over time, and to compare this to the change in pain severity and diagnostic impression in those who received narcotics.

In a convenience sample of 44 patients who presented when two physicians were available, inter-observer agreement of the abdominal exam was assessed (see Appendix 1: Interobserver Reliability of Physical Exam and section **4.9.3** below). Each physician completed a standardized abdominal exam form, indicate the location of maximum tenderness, the presence or absence of key clinical signs, and whether they felt the patient had peritoneal irritation. Finally, each physician indicated how certain they were about the presence of peritoneal irritation on a Likert scale.

4.6.2 Data Collection: Retrospective Chart Review

Patients with abdominal pain presenting during the study period were also independently identified by twice weekly review of a computerized log of all patients presenting to the Emergency Department. Hospital regulation require that all patients be registered into such a log when receiving medical attention: consequently, compliance with registration is very high. The log included patients' presenting complaints: the charts of patients with any abdominal or pelvic complaints were then reviewed in detail to assess

eligibility. Regardless of whether physician had completed the data form, all eligible patients were contacted between two and three weeks after their visit to the Emergency Department. A structured phone interview was used to determine if patients had been hospitalized within two weeks of discharge from the emergency department, and if so whether it was because of abdominal pain (see Appendix 2: Telephone Interview). For all patients admitted directly from the Emergency Department or reporting re-admission at phone contact, the inpatient hospital record was reviewed for evidence of the specified serious outcomes.

4.7 Losses to Follow-up

If a patient could not be reached by telephone, the next of kin was contacted to ascertain the patient's location. If it was impossible to contact the patient, the next of kin was used as proxy respondents for outcome. If this was unsuccessful, patients were classified as lost to follow-up.

4.8 Rationale for Length of Follow-up

Due to the rapid progression of acute abdominal pain, it was anticipated that all serious outcomes related to the initial visit would present within 2 to 3 weeks after discharge. In one case series of misdiagnosed appendicitis, the average time to return to hospital was 39 hours[27]. It was recognized that neoplasm might present in a more insidious fashion and not be detected within this follow-up period. However, the consequences of delaying diagnosis in this disease group are also less significant. As the main interest of this study was to investigate the rate of *acute* serious outcomes, a two week follow-up period was chosen.

Figure 1 : Study Sheet

SPARED PILOT STUDY

ALL HOUSESTAFF: Please consult with staff physician when completing this form

Last Name: ________________ **Unique #:** ____________ **Clerk:** ______

Date: _____/_____/_____ ❑ Cubicles ❑ Observation ❑ Resuscitation

Physician's Initials: ____________ ❑ Housestaff ❑ Staff Physician ❑Clin. Assoc.

EXCLUSION CRITERIA: (Please check all that apply)

❑ **Abdominal Pain duration ≥ 3 days** (if ptnt had previous episode of pain which resolved: = **eligible**)
❑ **Identical pain to a previously diagnosed condition** (e.g. previously proven biliary colic or known pancreatitis)
❑ **Hemodynamically Unstable:** (SBP < 90 mm Hg, HR > 125 or in physician's opinion)
❑ **Suspicion of Abd Aortic Aneurysm / Aortic Dissection**
❑ **Most likely diagnosis Renal Colic**
❑ **Patient Less than 16 years of Age**
❑ **Diagnosis previously established**
❑ **Allergy to Demerol, Morphine or Fentanyl**
❑ **Pregnant Patient**
❑ **Peritoneal Dialysis**
❑ **Traumatic cause of pain**
❑ **Main Complaint NOT Abdominal Pain**

IF ANY BOX CHECKED, do not complete the remainder of questionnaire. Thank you for your help.

PHYSICIAN PAIN ASSESSMENT & DIAGNOSIS

Time of Initial pain assessment: ____________ **Duration of Pain (hours) :** ____________

What is your presumptive diagnosis after initial exam: ______________________
How certain are you of this? ❑Very Certain ❑Certain ❑Neutral ❑Uncertain ❑Very Uncertain

Does the patient have peritoneal irritation ? : ❑No ❑Yes
How certain are you of this? ❑Very Certain ❑Certain ❑Neutral ❑Uncertain ❑Very Uncertain

Please rate how severe the patient's pain is now: **(Mark line with one vertical slash)**

No pain |——————————————————| Unbearable pain

Will you order parenteral narcotic medications for this patient ?: ❑No ❑Yes

If No, indicate the **single most important reason** why not: **(Choose one only please)**

❑ Pain not severe enough to require narcotic
❑ Narcotic analgesic may mask the diagnosis
❑ Other (please specify): ____________________
❑ Consultant may object
❑ Other treatment may relieve pain

Please fold along dotted line and administer following section to patient after completing above

PATIENT PAIN ASSESSMENT:

"We would like to call you in two weeks to make sure you have recovered from your abdominal pain. Please give:

Daytime Telephone Number: __________ **Evening Telephone Number:** ____________
as well as the name and telephone number of a **contact person** in case you are admitted to hospital:"

Contact Person's Name: __________ **Relation to ptnt:** ________ **Telephone:** ____________

Ask patient: "Please rate how severe your pain is now by marking line with one vertical slash":

No pain |——————————————————| **Unbearable pain**

Please ask patient : **"Do you need any kind of pain killer ?"** ❑No ❑Yes
Please ask patient: **"Do you need a strong pain killer ?"** ❑No ❑Yes

2 Staff MD's Available?: ❑No ❑Yes **If "Yes" Please complete INTER-OBSERVER form now** (attached)

Figure 2: Reassessment Sheet

SPARED PILOT STUDY: PAIN & DIAGNOSIS REASSESSMENT

Please complete on: 1) All patients GIVEN NARCOTIC ANALGESICS
2) All patients NOT DISCHARGED 1 HOUR after initial exam

PHYSICIAN PAIN & DIAGNOSIS REASSESSMENT

Please complete within 3 hours of initial exam

Time of Repeat Exam: _____

What was your final diagnosis after repeat exam: _____

How certain are you of this ? ❑Very Certain ❑Certain ❑Neutral ❑Uncertain ❑Very Uncertain

Does the patient have peritoneal irritation ?: ❑No ❑Yes

How certain are you of this? ❑Very Certain ❑Certain ❑Neutral ❑Uncertain ❑Very Uncertain

Please rate the severity of the patient's pain now : (Please mark line with one vertical slash)

No pain |——————————————————————| Unbearable pain

Please fold along dotted and administer to patient after completing above section

PATIENT PAIN REASSESSMENT:

How much pain do you have now ? (Please ask patient to mark line with one vertical slash)

No pain |——————————————————————| **Unbearable pain**

Physician Comments:

_____

Thank you very much for your participation !

4.9 Outcome Assessment Criteria

4.9.1 Primary Outcomes: Serious Outcomes

The occurrence of serious outcomes was determined by review of the medical charts of all admitted patients by the principal investigator. Operative, pathologic, or radiographic reports, as well as laboratory test were used to confirm the presence of any serious outcome noted in the medical chart. Serious outcomes were defined as:

- Perforation of an abdominal viscus
- Bowel Obstruction: Complete or partial
- Gastro-Intestinal Hemorrhage
- Intra-Peritoneal Hemorrhage
- Intra-Abdominal Abscess formation
- Ischemia of abdominal viscus
- Peritonitis or intra-abdominal abscess
- Sepsis syndrome, using American College of Chest Physicians definition[80] of two or more of: i) temp >38 or < 36°C, ii) heart rate > 90, iii) respirations > 20 or PaCO2 < 32, iv)WBC > 12,000 or <4000, or > 10% band, **and** the presence of end organ dysfunction with elevated lactate, oliguria or PaO2/FiO2 < 280.

4.9.2 Primary Outcomes: Narcotic Requirement

As discussed under **3.4**, "Pain: Its Control, Theory, and Measurement", the assessment of pain severity is problematic. Even given that pain severity has been meaningfully quantified, there is no agreement as to a threshold amount of pain at which

opioid analgesics are "required". Therefore, four measures of narcotic requirement were used to assess the frequency of patients with abdominal pain requiring narcotic analgesics : 1) the patient's assessment of need for analgesic, 2)whether analgesics were actually given in each case, 3) whether physicians intended to order narcotics, and 4) the physician's assessment of narcotic requirement including those patients in whom narcotics were withheld for any reason other than insufficient pain severity. The last definition of narcotic requirement was intended to capture patients who are currently denied narcotic analgesics, but who would be eligible for a future randomized clinical trial on the safety of narcotics in acute abdominal pain. The four measure are described below.

4.9.2.1 ***Patients' assessment of Narcotic Requirement***

Patients assessment of narcotic need was measured by their response to the question: "Do you need a strong pain killer" (see Appendix 1: Interobserver Reliability of Physical Exam). The phrasing of this question attempted to parallel the question for physicians regarding analgesia need. Identical wording could not be used as the terms "narcotic analgesic" would not have been comprehensible to the majority of patients. Instead, the phrase "strong pain-killer" was chosen as an equivalent, and more recognizable, phrase. In addition, patients were asked if they needed. "any kind of pain killer" in order to tap into patients who wanted relief from their pain, but would not go so far as to say a "strong pain killer" was necessary.

4.9.2.2 Narcotic Requirement as Measured by Actual Analgesic Use

Secondly, the frequency of patients requiring opioid analgesics was measured by the number of patients who actually received narcotics in the emergency department. This data

was abstracted directly from nursing records of narcotic administration, which are highly accurate due to administrative policies regarding controlled substances. This data was available for all patients whether a questionnaire had been completed or not.

4.9.2.3 ***Physicians' Assessment of Narcotic Requirement***

The frequency of patient requiring narcotics was assessed by, physicians' response to the question : "Will you order a narcotic analgesic for this patient". Physicians could respond "Yes" or "No" only to this questions, with "yes" responses defined as indicating need for analgesia for measure 3. If they responded "no", physicians were asked to chose the main reason why not, from a list of 6 options: pain not severe enough to require narcotic, consultant may object, narcotic analgesic may mask the diagnosis, other treatment may relieve pain or other reason (with space provided for the physician to elaborate, see Appendix 1: Interobserver Reliability of Physical Exam). If "Pain not severe enough" and another reason applied, physicians were instructed to choose "pain not severe enough". In this manner, it could be determined whether physicians' decisions were based primarily on a perception that the pain was not severe enough to warrant a narcotic analgesic, rather than fears regarding the safety of narcotics or other reasons. The wording of the question deliberately avoided asking physicians to make hypothetical judgments such as "would you order a narcotic"; responses to hypothetical questions may not reflect accurately actual behaviour. Instead, the focus was on what physicians had actually decided to do.

4.9.2.4 Proportion of Patients Requiring Analgesics: Definition for Determining Feasibility

None of the previous three measures are ideal to determine the proportion of patients with acute abdominal pain who "require" narcotics and thus would be eligible for a randomized clinical trial. The physicians assessment, (measure 1), has the advantage of reflecting the fact that the final decision regarding whether narcotics are given or withheld is ultimately made by the treating Physician, with varying degrees of input from nurses and the patient themselves. Also, in a future randomized clinical trial, decisions on patient eligibility would likely be made by emergency physician, thus the physician's assessment of narcotic requirement is the most accurate predictor of eligibility for future trials.

However, it is recognized that at present, some physicians may withhold narcotic analgesics out of concerns of masking the diagnosis, consultant objections, or one of the other reasons listed above. Such patients in fact have abdominal pain severe enough to require analgesics according to the physician, and thus would be eligible for the proposed randomized clinical trial. Thus, for the purposes of assessing potential patient enrollment of a future randomized clinical trial, "need for analgesic" should include all patients who are currently given narcotic analgesics, plus those patients in whom narcotics are withheld for any reason other than "Pain not severe enough".

4.9.3 Physician Inter-Observer Reliability of the Abdominal Exam

In a convenience sample of patients presenting when there was double physician coverage in the Emergency Department, both physicians were asked to independently complete a standardized abdominal exam form (see Appendix 1: Interobserver Reliability of Physical Exam). This form was modified from an ongoing world-wide abdominal pain

database which had enrolled over 100,000 patients to date[81,82,83] and listed 9 signs pertinent to the examination of the abdomen, each of which were marked as present or absent. The location of maximum tenderness was indicated on a three by three grid overlying a diagram of the abdomen. Finally, physicians commented on whether the patient had peritoneal irritation on exam, and rated their degree of certainty on a 5 point Likert scale.

4.9.4 Secondary Outcomes:

In addition to the primary objectives, 19 other end-points were collected to better describe the acute abdominal pain population. Two main sources were used: review of the emergency department or inpatient medical records, and questionnaires completed by the physician and patient. The specific methods used to collect secondary data are described below:

4.9.4.1 Telephone Follow-up and Medical Chart Review

1. Return visits within 2 weeks of study enrollment among patients initially discharged from the emergency department were determined by telephone interview and verified against the medical chart. Return visits were further classified as planned or unplanned.
2. The **rate of subsequent unplanned hospitalization** among eligible patients initially discharged from the emergency department was assessed by telephone interview and verified against the medical chart.
3. The **rate of positive and negative laparotomies** among eligible patients initially discharged from the emergency department as reported by patients and verified by operative records was recorded. Laparotomies were classified as "positive" if any of the

serious conditions listed in section 4.9.1 above were found at operation. The pathology report was used to verify the operative report when applicable (e.g., appendicitis or intra-abdominal neoplasm). b) The rate of positive or negative laparotomies within 2 weeks of study enrollment among patients initially admitted from the emergency department was determined in a similar manner.

4. The **all cause mortality rate** within 2 weeks of study enrollment was determined for patients initially discharged from the emergency department by medical chart review for patients who died at the Ottawa Civic Hospital. Telephone contact of next of kin and/or family physicians was used to determine if patients had died elsewhere. b) The all cause mortality rate for patients initially admitted from the emergency department was also determined by medical chart review.
5. The **number of abdominal pain patients who were excluded, and reason for exclusion** was determined for all patients presenting with an abdominal or pelvic complaint during the study period by daily review of the emergency treatment record of treatment. This data was collected for planning of any future randomized clinical trials on the safety of narcotics in this patient population.
6. Patients eligible for the study and discharged from the Emergency Department were classified as **lost to follow-up** if telephone contact could not be made after seven attempts within two weeks, including attempting to contact next of kin and family physicians.
7. If a narcotic was ordered by emergency physicians, the **specific drug, dosage, and route** was abstracted from the medical chart.

8. The **proportion of study eligible patients referred for surgical consultation** was abstracted.

4.9.4.2 Patient and Physician Questionnaire

9. The **number of patients for whom physicians indicated they would NOT order narcotic analgesics** was recorded on the study questionnaire(see **Figure 1**).
10. The **primary reason that narcotics were withheld** was recorded.
11. The **physicians' impression of the patient's pain severity** was assessed via standardized 100 mm VAS.
12. The **patient's perception of their pain severity** on an identical scale.
13. Physicians indicated if the patient had **signs of peritoneal irritation,** recorded as yes or no.
14. Physicians also recorded their **confidence in the presence of peritoneal irritation** on a 5 point Likert scale (see Appendix 1: Interobserver Reliability of Physical Exam).
15. Physicians **repeat pain severity rating** on the VAS scale was repeated on those patients who stayed in the emergency department for longer than one hour (see **Figure 2**).
16. The **patients repeat pain severity rating** on VAS was reassessed after 1 hour on those not discharged by that time.
17. The **repeat exam for presence of peritoneal signs** was conducted after 1 hour.
18. The **confidence in the presence of peritoneal irritation** was reassessed after 1 hour
19. Any difficulties arising in the implementation of this study were recorded, for possible modification of future randomized clinical trial protocols.

4.10 Sample Size Requirements for Equivalence Trials

Prior to initiating this study, sample size calculations were undertaken to determine how precise an estimate of the primary endpoint, serious outcome rate, could be achieved with the resources and time available.

The purpose of the proposed clinical trial is to establish the equivalence of narcotics and placebo in the rate of serious outcomes in acute abdominal pain. This has implications for the planning of the trial, particularly the formulation of the test hypothesis and the sample size calculations.[84,85] When determining the effectiveness of a new therapy, a null hypothesis of no difference between therapies is tested: rejection of the null hypothesis allows acceptance of the superiority of the new therapy. However, when the intent is to establish equivalence of two therapies, the null hypothesis should not be formulated in the same manner.[85] Rather, a null hypothesis of superior efficacy of the new therapy to control by some clinically important difference should be used in equivalence trials. Rejection of this null hypothesis permits acceptance of the alternate that the two treatments are not different by any clinically important amount. Donner (p. 202) gives sample size calculations based on such a hypothesis test in Equation 1.[84]

Equation 1: Sample Size for Equivalency Trials

Null and Alternative Hypotheses: $h_o: P_C \geq P_E + \delta$

$h_a: P_C < P_E + \delta$

$$n = 2\left(\frac{Z_\alpha\sqrt{2\bar{P}(1-\bar{P})} + Z_\beta\sqrt{P_E(1-P_E)+P_C(1-P_C)}}{P_E - P_C - \delta}\right)^2$$

where n = Total sample size

Z_α= Quantile of the normal standard deviation for the probability of type I error

Z_β = Quantile of the normal standard deviation for the probability of type II error

P_E = Proportion of experimental group with outcome

P_C = Proportion of control group with outcome

$\bar{P}$ = Pooled proportion $(P_e + P_c)/2$

δ = Minimal clinically important difference in proportion with outcome

It was known that on average 50,000 patients are seen annually at the participating Emergency Department; thus 25,000 patients were expected to be seen within 6 months of data collection. A pilot study by Dr. James Quinn done at an emergency department serving a similar population in the same city demonstrated 9.5% of all emergency department patients had abdominal pain. It was unknown what proportion of patients with abdominal pain would be eligible for the study (i.e., had acute abdominal pain, an uncertain diagnosis, and no contra-indication to opioids). Therefore sample size projections were made using two different assumptions of the proportion of abdominal pain patients who would be eligible . 1) Assuming 50% of the 9.5% of patients with abdominal pain would be eligible for the study, or 4.75% of all patient seen, then 1187 patients could be enrolled in 6 months. 2) If 30% of abdominal pain patients were eligible, or 2.85% of all emergency department patients, then 750 patients would be entered in 6 months. The following table presents the expected

precision attainable with 6 months of data collection using different rates of patient eligibility and rates of serious outcome:

Figure 3: Sample Size Calculations

% of Total Patients Eligible	# Eligible Patients	% of Eligible Patients w. Serious Outcomes	Upper CI	Lower CI
2.85% *	**750**	**1.00%**	**2.09**	**0.45**
2.85%	**750**	**2.00%**	**3.36**	**1.17**
2.85%	**750**	**4.00%**	**5.73**	**2.76**
4.75% †	1187.5	1.00%	1.80	0.54
4.75%	1187.5	2.00%	3.01	1.31
4.75%	1187.5	4.00%	5.32	2.99

* Assuming 30% of all abdominal pain patients eligible

† Assuming 50% of all abdominal pain patients eligible

Thus, over a wide range of possibilities, it seemed reasonable that 6 months of data collection would allow acceptable precision in estimating the serious outcome rate.

4.11 Data Analysis

Due to the observational nature of this study, the majority of statistics reported are descriptive.

4.11.1 Discrete Outcomes

When appropriate, simple proportions and percentages were reported for each discrete outcome. Between group proportions were compared against the null hypothesis of no difference using the Chi square statistic. For example, the rate of serious outcomes was compared between those patients with and without returned questionnaires. For

polychotomous outcomes such as final ED diagnosis, groups were combined if there were fewer than five subjects in each cell. A 5% significance level was used in all cases.

For primary outcomes, 95% upper and lower confidence bounds were also reported, calculated using the Score Method with continuity correction.[86] (Equation 2) rather than the commonly used Wald method (Equation 3). Vollset has shown that the Wald method performs poorly in producing confidence intervals that contained the true parameter value in computer generated simulated trials when the underlying proportion, p, was close to 0 or 1.[86] In contrast, 95% confidence intervals produced by the Score method with continuity correction covered the true estimate at least 95% of the time, even when p was less than 1%. However, the Score method was found to be conservative, in that the lower boundary slightly exceeded 95% in all cases. However, of 13 methods investigated, Vollset recommended the Score method due to its consistent coverage and conservativeness.

Equation 2: Score Method with Continuity Correction

$$\frac{\left(x \pm \frac{1}{2}\right) + \frac{Z_\alpha^2}{2}}{n + Z_\alpha^2} \pm \frac{Z_\alpha^2 \sqrt{\left(x \pm \frac{1}{2}\right) - \frac{x \pm \frac{1}{2}}{n} + \frac{Z_\alpha^2}{4}}}{n + Z_\alpha^2}$$

where x= number of success
n= number of trials
Z_α=1-α quantile of the standard normal distribution

Equation 3: Wald Method

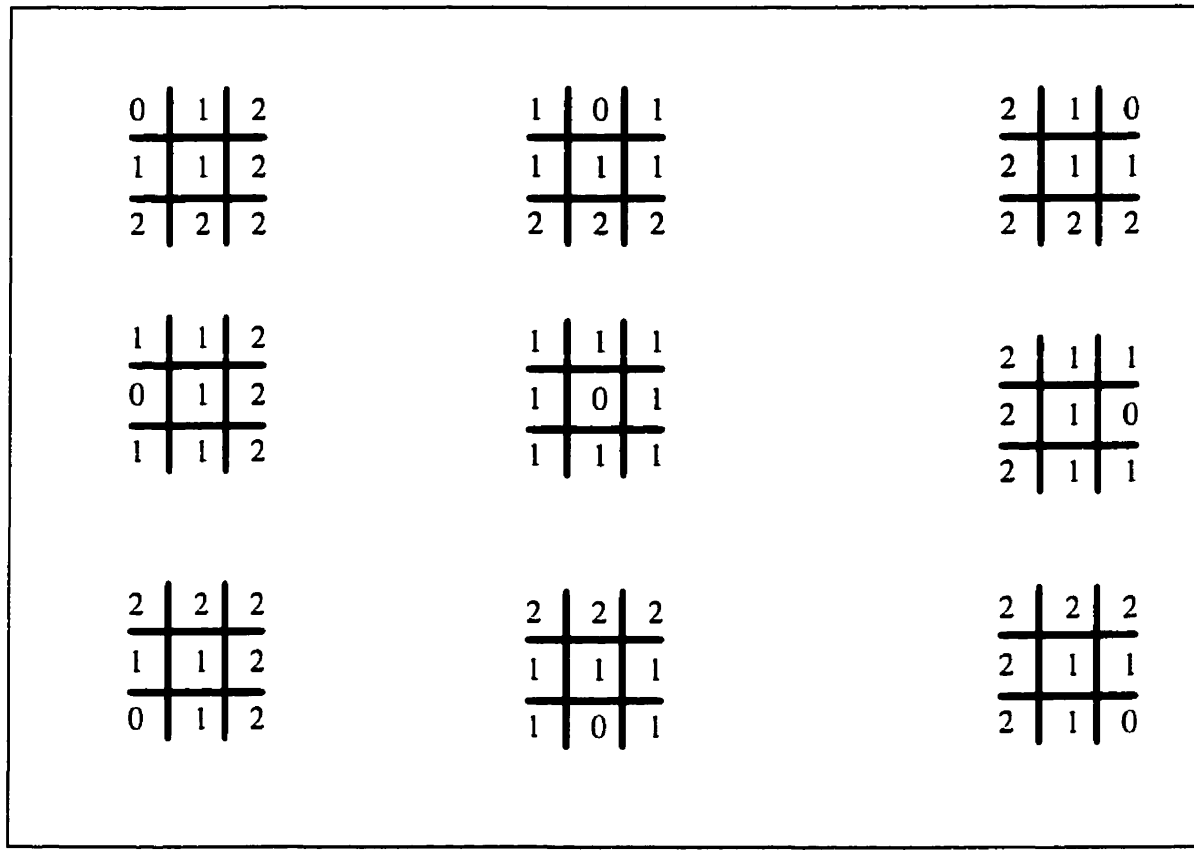
$$p \pm z_\alpha \sqrt{\frac{p(1-p)}{n}}$$

where p= proportion trials positive
z= 1-α quantile of the standard normal distribution
n= number of trials

4.11.2 Physician Agreement on the Abdominal Exam

For dichotomous categorical data, such as physician agreement for the presence of peritonitis, the kappa statistic was used.[87] For agreement on the location of maximal abdominal tenderness, the weighted kappa statistic was employed to allow credit for varying degrees of agreement[87]. For ease of calculation, the weighted kappa is calculated in terms of the proportion of disagreement, and weights are assigned corresponding to the degree of disagreement. The weights for degree of disagreement on location for abdominal pain were assigned as indicated in the following diagrams, with "0" representing perfect agreement and "2" representing maximal disagreement. Thus, for right upper quadrant location of maximum tenderness, if the second physician also observed right upper quadrant tenderness, then the kappa weight assigned would be "0". However, if the second physician chose left upper quadrant, a disagreement weight of "2" would be assigned.

Figure 4: Assignments of Weights for Agreement on Location of Maximum Tenderness



The weighted kappa, κ_w, is given by [87]:

$$\kappa_w = 1 - \frac{q_O'}{q_E'} \qquad q_O' = \Sigma \frac{wf_o}{N} \qquad q_e' = \Sigma \frac{wf_e}{N} \qquad f_e = \frac{r_i c_i}{N}$$

Where W= Weight

f_o= Observed Frequency in each cell of r by c table

f_e= Expected Frequency in each cell of r by c table

r_i= Sum of observed frequencies in the *i*th Row in r by c table

c_i= Sum of observed frequencies in the *i*th Column in r by c table

4.11.3 Continuous Outcomes

Means and standard deviations were reported for all continuous outcome measures. Student's t test was used to compare group means against the null hypothesis that both means were drawn from the same population. Visual analog pain scores were treated as continuous measures. Patient-physician agreement on pain severity scores were assessed using the random effects Intraclass Correlation Coefficient, as given by Fleiss:[88]

Equation 4: Intraclass Correlation Coefficient

$$\hat{R} = \frac{N(S-E)}{N \cdot S + k \cdot R + (Nk - N - k)E}$$

where N= number of subjects

k= number of raters

S= mean sum of squares due to between subject variation

R= mean sum of squares due to between rater variation

E= mean sum of squares due to error

5. Results

5.1 Demographics of the Patient Population

Between October 1st, 1995 and April 30th 1996, 31,772 patients presented to the Emergency Department. Of these, 860 patients (2.7%) were found to be eligible. The flow of patients into the study is presented in **Figure 5** and **Figure 6**.

Figure 5: Flow diagram of patients into the study.

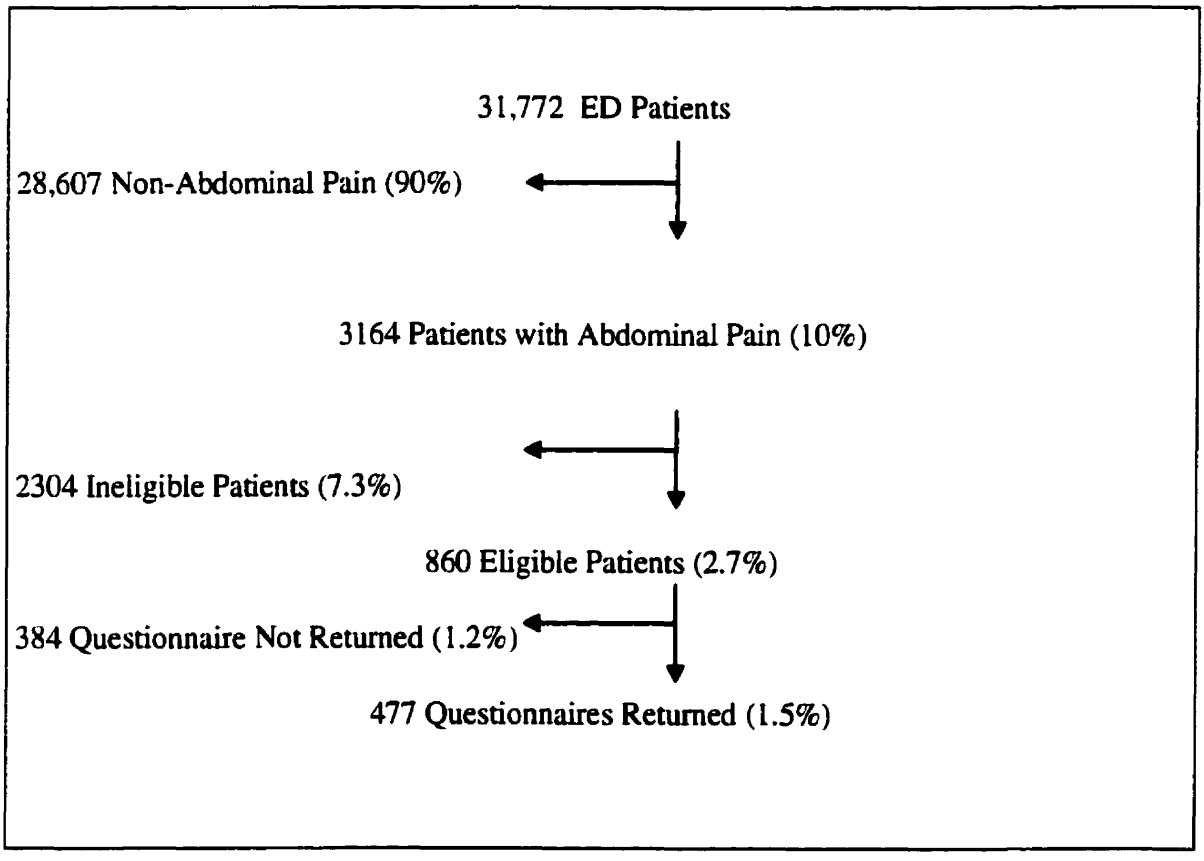
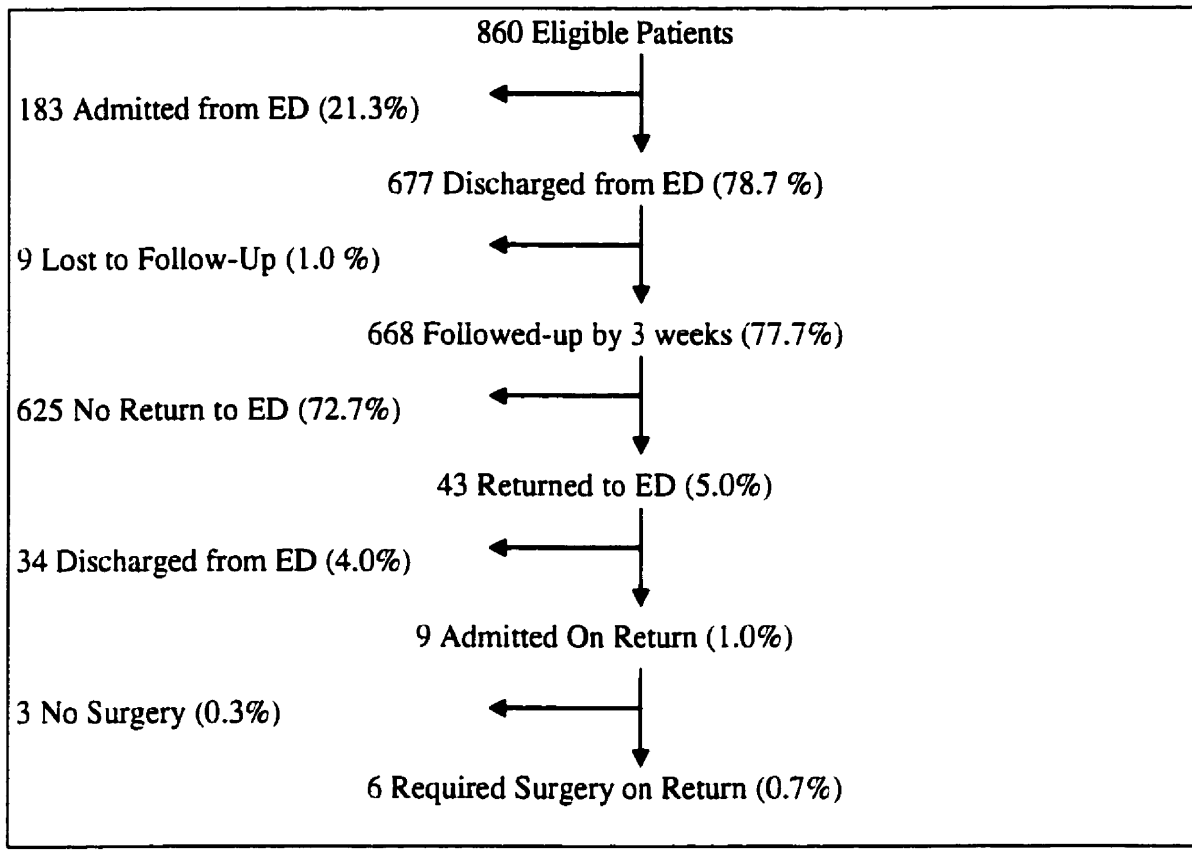


Figure 6: Diagram of Patient Flow Within the Study



Of 3164 patients who presented to the emergency department with a chief complaint of abdominal pain, 2304 were excluded and thus the 860 eligible patients represented 27 % of all abdominal pain. The breakdown of the reasons for exclusion of the 2304 patients is presented in **Table 2.** Pain duration of greater than 48 hours was the most common exclusion criteria, followed by patients in whom abdominal pain was not judged to be the primary complaint by the emergency physician.

Table 2: Frequency of Individual Exclusion Criteria

Reason For Exclusion	Frequency N=2304	%
Pain > 48 Hours	752	32.6
Pain Not Primary Complaint	524	22.7
Suspected Renal Colic	260	11.3
Recurrence of Pain with Established Diagnosis	189	8.2
Pregnancy	124	5.4
Direct to Specialist	90	3.9
Less than 16 years old	87	3.8
Previously Enrolled	59	2.6
Referred with a Diagnosis	57	2.5
Recurrence of Non-Specific Abdominal Pain	55	2.4
Allergy to Demerol	29	1.3
Suspected Aneurysm	20	0.9
Traumatic Abdominal Pain	15	0.7
Confusion or Language Barrier	14	0.6
No Emergency Chart	11	0.5
Allergy to Morphine	9	0.4
Peritoneal Dialysis	6	0.3
Hemodynamically Unstable	3	0.1

The demographic characteristics of the study population are outlined in **Table 3**. The acute abdominal pain population seen at the Ottawa Civic Hospital had a mean age of 41.0 years, age range of 16-95 years, and 61.1% were female. Although the mean age was higher in the groups of patients with a completed questionnaires ($z=2.09$, $p=0.014$), the admission rate, surgical rate, and rate of serious outcomes were similar.

Table 3 : Demographic and Clinical Characteristic of All Eligible Patients, and Patients With and Without Returned Questionnaires

Characteristic	All Patients		Questionnaire Returned		No Questionnaire Returned	
	N=860	%	N=477	%	N=384	%
Mean Age, years (S.D.)	41.0 (18.4)	-	39.8† (17.1)	-	42.4† (19.9)	-
Female	526	61.1	285	59.7	241	62.8
Admitted	188	21.9	108	22.6	80	20.8
Surgery	87	10.1	50	10.5	37	9.6
Serious Outcome	66	7.7	34	7.1	32	8.3

† $z=2.09$, $p=0.014$

The emergency physician final diagnosis is presented in **Table 4**. Serious outcome included all diagnoses corresponding to the definition given in section **4.9.1, "Primary Outcomes: Serious Outcomes"**. Non-specific abdominal pain and gastro-enteritis were the two most common diagnoses in all groups, but there was a statistically significant difference between patients with and without returned questionnaires in the type of diagnoses ($\chi^2_{12}=33.8$, $p=0.001$). Specifically, biliary colic and appendicitis were nearly twice as common in the questionnaire returned group, and urinary tract infection was the diagnosis more often among those patients without a returned questionnaire.

Table 4: Comparison of Final Diagnosis, Patients With and Without Questionnaires

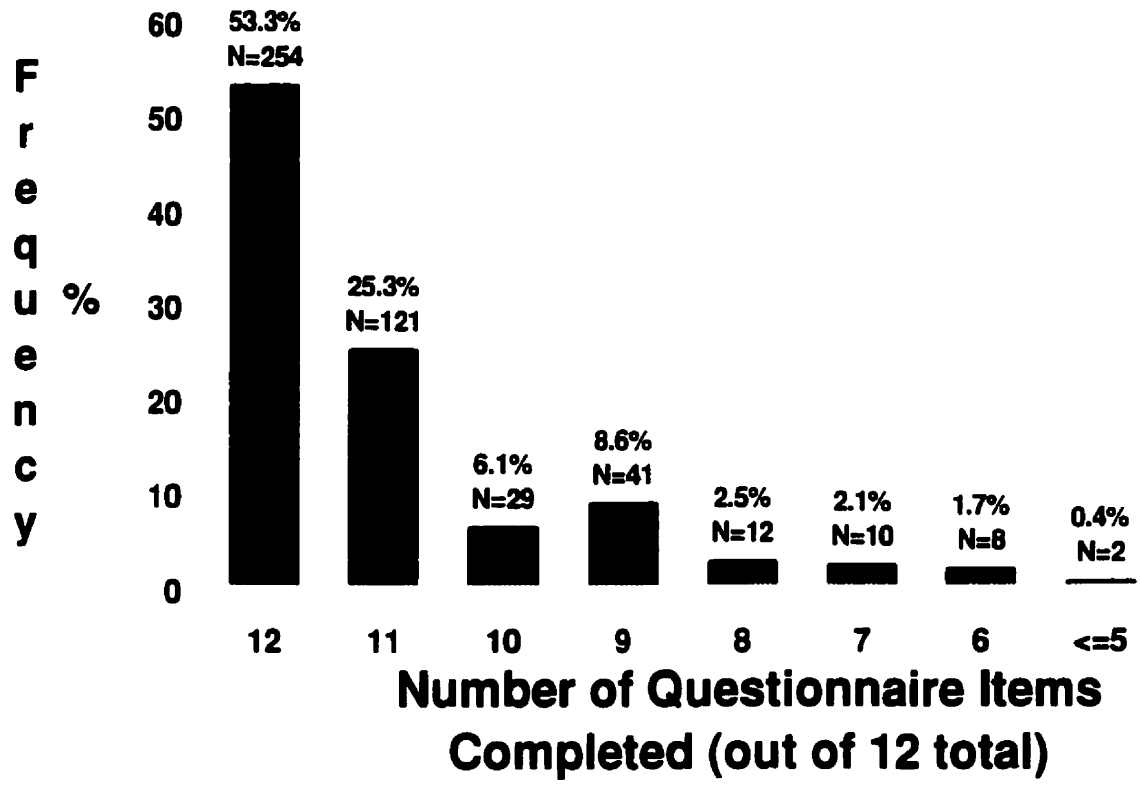
Diagnosis	All Patients		Questionnaire Returned		Questionnaire Not Returned	
	N=860	%	N=476	%	N=384	%
Abdominal Pain, Not Yet Diagnosed	220	25.6	127	26.7	93	24.2
Gastro-enteritis	139	16.2	75	15.8	64	16.7
Gastric / Esophageal	74	8.6	42	8.8	32	8.3
Pelvic Pain	73	8.5	34	7.1	39	10.2
Biliary Colic	62	7.2	44	9.2	18	4.7
Serious Outcomes	59	6.9	28	5.9	31	8.1
Appendicitis	54	6.3	38	8.0	16	4.2
Urinary Tract	52	6.0	17	3.6	35	9.1
Constipation	33	3.8	15	3.2	18	4.7
Abdominal Wall	25	2.9	13	2.7	12	3.1
Biliary / Hepatic, Other	24	2.8	16	3.4	8	2.1
Irritable Bowel/Spasm	18	2.1	10	2.1	8	2.1
Systemic	14	1.6	7	1.5	7	1.8
Diverticulitis & Other Intestinal	13	1.2	10	1.7	3	0.5

$\chi^2_{13} = 31.6$, $p=0.003$

5.2 Rate of Returned Questionnaires

Of the 860 eligible patients, physicians returned questionnaires for 477 patients (55.4%). However, not all questions were answered on returned questionnaires. The degree of completion of questionnaires is described in **Figure 7.** Note that 78.6% of returned questionnaires were either completed or missing a response to only one item.

Figure 7: Frequency Histogram for Number of Items Responded to on 477 Returned Questionnaires



The response rates to each item on the questionnaire are listed in **Table 5**. In total, 402 out of 4608 total items were incomplete (8.7%). In general, items addressed to the patient were most frequently incomplete. Among the items addressed to physicians, the pain severity visual analog scale was the item they failed to complete most often.

Table 5: Frequency of Questionnaire Items not Answered

Questionnaire Item	Frequency No Response N=402	Percentage
Physician's assessment of pain severity	52	12.9
If no narcotics, reason why not	39	9.7
Certainty of peritoneal irritation	28	7.0
Will your order narcotic analgesics	26	6.5
Certainty of Presumptive Diagnosis	23	5.7
Does the patient have peritoneal irritation	22	5.5
Time of Initial Assessment	19	4.7
Presumptive Diagnosis	10	2.5
Duration of Pain	1	0.2
Patient: "Do you need a strong pain killer ?"	67	16.7
Patient: "Do you need any pain killer ?"	63	15.7
Patient assessment of pain severity	52	12.9

5.3 Principal Outcomes

5.3.1 Objective 1.1: Incidence of Serious Outcomes

At least one serious outcome occurred in 67 of the 860 eligible patients with acute abdominal pain (7.8 %). **Table 6** presents overall serious outcome rate, broken down into individual types of serious outcomes. Nine patients experienced more than one serious outcome: six subjects with both perforation and peritonitis were classified as peritonitis, two patients with both ischemic and obstructed bowels were classified as ischemic bowel, and one patient with a bowel obstruction and gastro-intestinal hemorrhage was classified as a bowel obstruction. Partial bowel obstruction was the most common serious outcome, followed by complete bowel obstruction and peritonitis.

Table 6: Objective 1.1, Frequency of Serious Outcomes for All Patients

Serious Outcome	Frequency N=860	Percentage	95% Confidence Interval (of %)[1]
Total, at least 1 Complications	67	7.8	6.1-9.8
Death	3	0.3	0.1-1.1
Peritonitis or Abscess	11	1.3	0.7-2.3
Perforation	10	1.2	0.6-2.2
Ischemic Bowel	4	0.5	0.2-1.3
Bowel Obstruction, Complete	13	1.4	0.8-2.5
Bowel Obstruction, Partial	19	2.2	1.4-3.5
Hemorrhage	5	0.6	0.3-1.6
Sepsis	2	0.2	0.0-0.9

[1] Score Method, Continuity Corrected[86]

Serious outcomes were then compared between the group of patients with returned questionnaires versus those without. No statistically significant difference occurred between the two groups in rates of serious outcome (**Table 7**).

Table 7: Frequency of Serious Outcome , With and Without Returned Questionnaire

Serious Outcome	Frequency, Questionnaire N=477	Frequency, No Questionnaire N=384
Yes	34	33
No	443	351

$\chi^2_1 = 0.437$, $p=0.51$

5.3.1.1 Deaths

Three patients died during the course of this study. The first patient was an 85 year old woman who presented comatose, with severe necrotizing pancreatitis and hypotension. In consultation with the patient's family, it was decided not to continue aggressive resuscitation, in view of her poor prognosis. The second patient was a 94 year old man with a final emergency department diagnosis of small bowel obstruction. This patient died of an acute myocardial infarction five days after presentation. The third patient was a 70 year old male, who had been referred to the surgical service after having received narcotic analgesics. The patient was later discharged by the surgical service with a diagnosis of non-specific abdominal pain, but re-admitted four days later with acute cholecystitis. The patient suffered from severe concomitant COPD, and was too ill to undergo invasive surgery. Therefore, a percutaneous cholecystostomy was performed which the patient tolerated well. The patient died of acute respiratory failure 37 days after study enrollment. All of these patients had received narcotics; however, it is unlikely that this contributed to the deaths of the first two patients, as both patients had diagnoses that ultimately turned out to be correct at the time of their deaths. It is less clear that narcotics did not the interfere in the initial diagnosis of the third patient, and whether failure to make the correct diagnosis on the initial visit contributed to the patient's ultimate demise.

5.3.2 Objective 1.2: Narcotic Requirements

Four different measures were used to determine how often patients, presenting to the Emergency Department with abdominal pain, required opiod narcotics. The number of

patients that actually received narcotic analgesics in the Emergency Department, as determined by chart review, was the first measure. In all 250/860 or 29% of abdominal pain patients received narcotic analgesics (95% confidence interval, 26.1-32.2).

Next, in the sub-sample of patients for whom a questionnaire was returned, physicians were asked to indicate whether they would order a narcotic analgesic (**Table 8**). If physicians withheld narcotics, they were asked to indicate their reason for doing so. In the 450 cases where this item was answered, physicians stated they would order narcotics 28.7 % of the time, similar to the proportion of patients that in the end actually received narcotics.

Table 8: Objective 1.2, Physicians' Assessment of Narcotic Requirement

Physicians Response to question: "Will you order a narcotic analgesic. If not, please give reason why not"	Frequency N=450	(%)	95% CI (%)
"Yes"	129	28.7	24.5-33.1
"No", Reason: "Pain not Severe "	213	47.3	41.5-50.8
"No", All Other Reasons	95	21.1	18.7-26.6
"No", No Reason Given	13	2.9	1.6-5.0

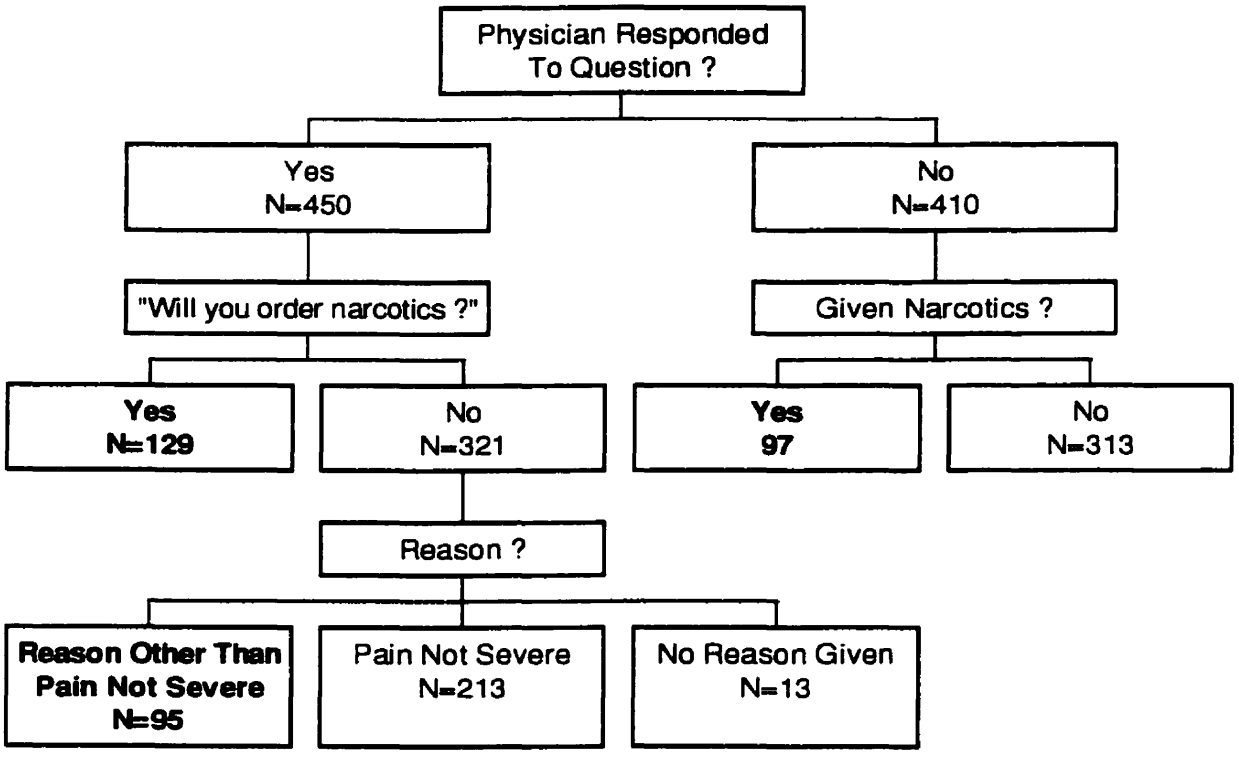
Patients response to the question "Do you need any kind of pain killer" and "Do you need a strong pain killer?" were the third and fourth items used to assess narcotic requirement, and are presented in **Table 9.**

Table 9: Objective 1.2, Patients' Assessment of Narcotic Requirement

	"Do you need any pain killer" N=413 (%)	"Do you need a strong pain killer" N=409 (%)
Yes (%)	237 (57.4)	154 (37.7)
95% CI (%)	52.5-62.2	32.8-42.6

Narcotic requirement, for the purpose of determining eligibility for the proposed randomized trial, included cases in whom physicians decided to give narcotics, as well as those where narcotics were withheld for any other reason than "pain not severe enough" (see section **4.9.2**, "**Primary Outcomes: Narcotic Requirement**"). For the 410 patients for whom physicians did not answer the question "will you order a narcotic analgesic", need for narcotic was determined by whether the physician actually ordered an analgesic or not. In total, 321/860 (37.3 %) of subjects met this definition of narcotic requirement and are presented in **Figure 8.**

Figure 8: Frequency of Patients Requiring Narcotics for Entry into Future Clinical Trial



Total Patients Requiring Narcotics (Eligible for Future Clinical Trial)

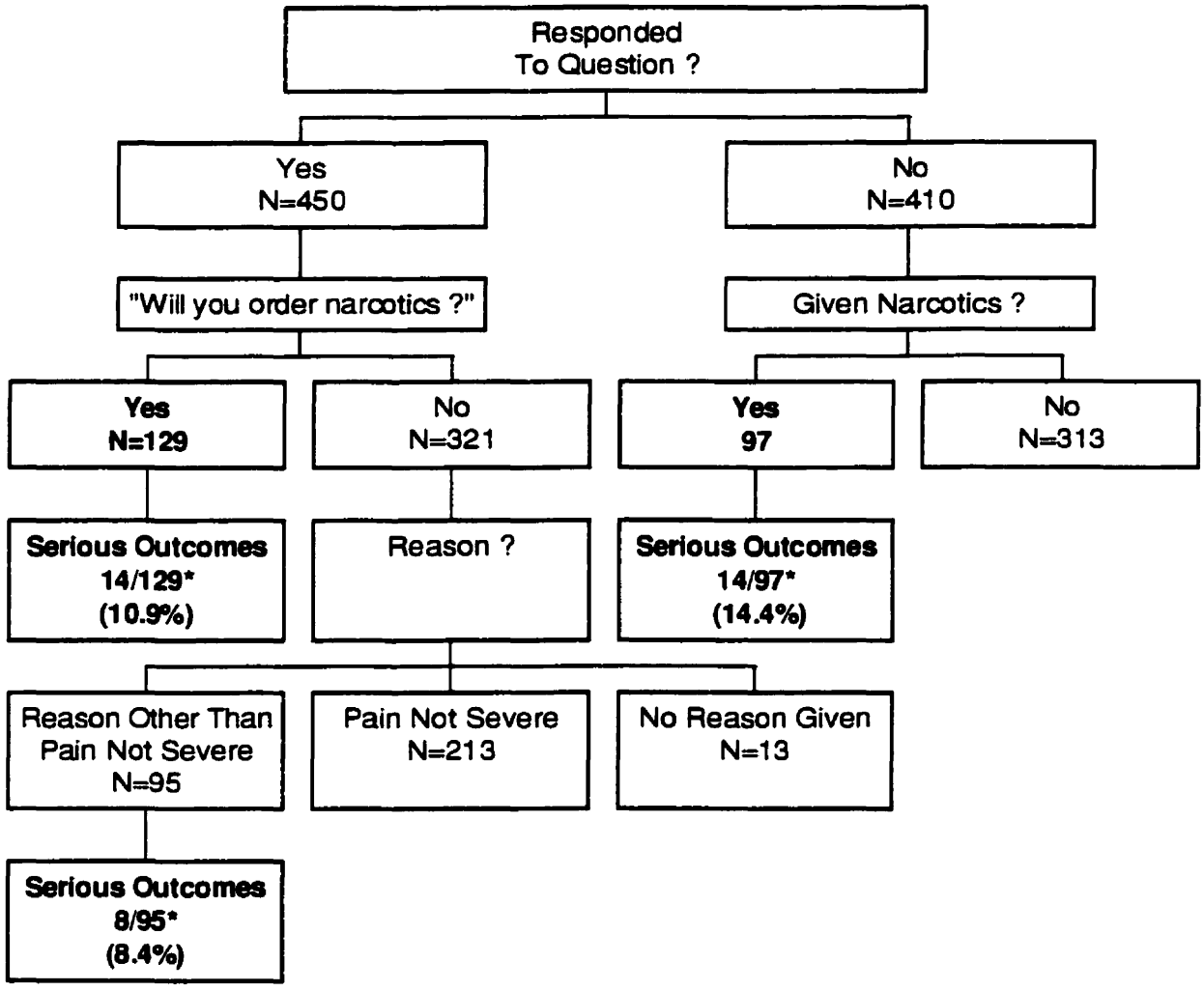
=129+95+97

=321

5.3.3 Serious Outcomes Among Patients Requiring Narcotic Analgesics

Among the 321 subjects defined as requiring narcotics above, 36 serious outcomes were observed (11.2%), as shown in the **Figure 9.**

Figure 9: Frequency of Serious Outcomes Among Patients Requiring Narcotic



Estimated overall serious outcome rate =14* + 14* +8* = 36/321

= 11.2 %

(95% CI: 8.1-15.3 %)

5.4 Secondary Outcomes

5.4.1 Objective 2.1:Patient-Physician Agreement on Need for Analgesics

Physicians and patients agreement regarding the need for analgesics was measured directly, by comparing the patient and physician categorical assessments of need for analgesics, as well as indirectly, by comparing patient and physician pain severity ratings.

5.4.1.1 Patient-Physician agreement on Need for Analgesics: Categorical Measures

Patient-physician agreement on need for narcotic is presented in **Table 10**. Answers were provided for both questions in 390 cases. The overall agreement between patients and physicians was only moderate, with a Kappa value of 0.53 (95% CI: 0.43 to 0.63). Of note, in 57/147 (38.8%) of patients that stated they needed a strong pain killer, physicians felt narcotic analgesics were not required.

Table 10: Objective 2.1, Patient-Physician Agreement on Need for Analgesics

"Do you need a strong pain killer?" (Patient)	"Will you order a narcotic analgesic for this patient ?" (Physician)		
	No	Yes	Total
No	220	23	243
Yes	57	90	147
Total	277	113	390

As discussed in section **5.3.2, Objective 1.2: Narcotic Requirements,** physicians occasionally withheld narcotics for reasons other than that they felt the patients pain was not severe enough to require narcotics. Therefore, patient and physician agreement regarding need for analgesics was re-analyzed, considering physicians reasons for withholding narcotics. Cases where the physician withheld narcotics for reasons other than "Pain not severe enough"

were included as requiring narcotics according to the physician (Table 11). Again, however, agreement was only moderate, with a Kappa value of 0.46 (95% CI: 0.37 to 0. 54).

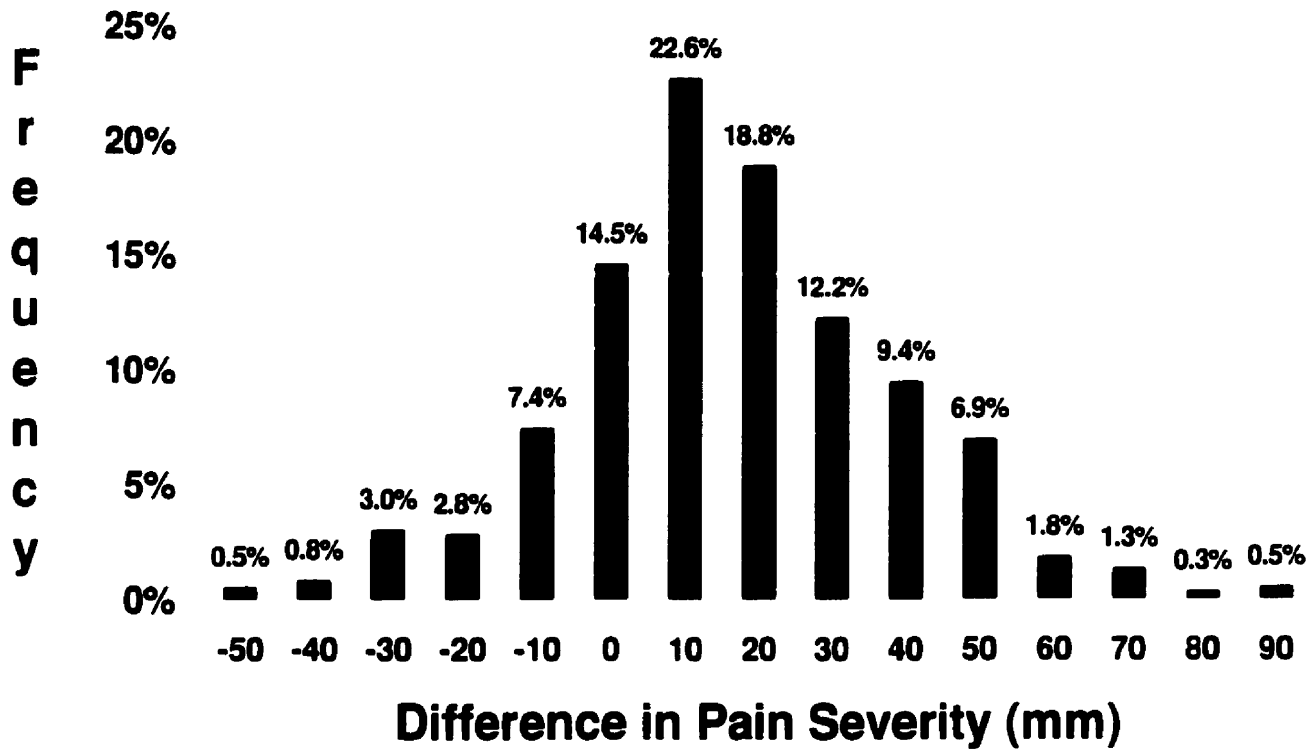
Table 11: Patient-Physician Agreement on Narcotic Requirement, accounting for Physician Reasons for Withholding Narcotics

Patient: "Do you need a strong pain killer ?"	"Will you order a narcotic analgesic for this patient ?" (Physician)		
	No	Yes	Total
No	182	36	218
Yes	73	118	191
Total	255	154	409

5.4.1.2 Patient-Physician Agreement on Need for Analgesics: Pain Severity

Patient-physician agreement on pain severity was examined by plotting the frequency distribution of difference between patients' and physician' assessments of pain severity on the 100mm visual analog scale. Physician pain assessments were subtracted from the patient's pain score and are presented in **Figure 10.** Positive values in the horizontal axis of the figure represent the physician's under-estimation of pain severity relative to the patient's assessment. From the histogram, we see that in nearly 50% of cases, physician's pain severity assessments were 20 mm or less severe than the patient's self assessment. In nearly 30% of case, the discrepancy was as large as 30mm.

Figure 10: Frequency Histogram of Patient-Physician Difference in Pain Severity



To examine whether disagreement regarding need for analgesia was related to perceptions of pain severity, the patient-physician differences in pain scores were calculated by subtracting physician scores from patient scores. The mean differences in pain scores are presented in **Table 12**, broken down by patient and physician categorical assessments of need for analgesics. Again, we see that the patients' self-assessments of pain severity were higher than the physicians' assessment, as demonstrated by the positive mean difference in pain scores seen in all cells. Differences in pain severity assessments were largest (25.9 mm) when the physician stated she or he would not order a narcotic, but the patient felt a strong pain killer was needed, suggesting that disagreement regarding need for analgesics may be related to divergent perceptions of pain severity.

Table 12: Mean Difference in Patient-Physician Pain Severity Scores, by Categorical Assessments of Need for Narcotics

	"Will you order a narcotic analgesic ?" (Physician) Mean Difference, VAS Score in mm (SD)		
"Do you need a strong pain killer ?" (Patient)	No	Yes	Total
No	10.0 (20.5) N=218	4.3 (13.8) N=23	9.4 (20.0) N=241
Yes	25.9 (27.1) N=56	12.4 (17.6) N=88	17.7 (22.7) N=144
Total N=391	13.3 (22.8) N=274	10.7 (17.1) N=111	12.5 (21.3) N=385

To further examine the relationship between perception of pain severity and assessment of narcotic requirement, the mean differences in patient and physician pain severity assessments on 100 mm visual analog scale were compared in **Table 13**, taking physicians reasons for withholding narcotics into consideration. The difference in pain severity was greatest when patients felt a narcotic was indicated, and physicians reason for withholding narcotics was "pain not severe enough" (mean difference, 32.5 mm).

Table 13: Differences in Pain Severity, According To Patient and Physician Assessment of Narcotic Requirement, and Physicians' Reasons for Withholding Narcotics

"Do you need a strong pain killer ?" (Patient)	Physician Reasons for Withholding Narcotics Mean Difference, VAS Score in mm (SD)		
	Pain Not Severe Enough	Other Reason	Total
No	11.6 (19.8) N=161	4.6 (22.0) N=50	9.9 (20.5) N=211
Yes	32.5 (26.3) n=28	18.6 (26.8) N=27	25.6 (27.2) N=55
Total N=	14.7 (22.1) N=189	9.5 (24.6) N=77	13.2 (22.9) N=266

5.4.1.3 Agreement of Pain Severity Scores: Intraclass Correlation Coefficient

Agreement between patients and physicians pain ratings in terms of the intraclass correlation coefficient (ICC), was found moderate at 0.60 ($F_{1, 410, 0.05}$=4.12, p<0.0001). The intraclass correlation coefficient was then re-calculated, stratified according to patient and physician assessment of need for narcotics, in **Table 14**. Again, poor agreement was observed in cases where patients felt a narcotic was required but physicians disagreed.

Table 14: Intraclass Correlation Coefficient, According to Patient and Physician assessment of Narcotic Requirement

"Do you need a strong pain killer ?" (Patient)	"Will you order a narcotic analgesic ?" (Physician)	
	No	Yes
No	0.73	0.51
Yes	0.14	0.56

5.4.2 Objective 2.2: Inter-Physician Agreement on the Abdominal Exam

The following table presents the inter-physician Kappa statistic for agreement of several physical signs considered to be clinically important in the diagnosis of acute abdominal pain. Due to logistical problems in arranging repeated examinations by two staff physicians, the agreement statistics are based on 44 comparisons, which is reflected by the wide confidence intervals.

Table 15: Objective 2.2, Physician Agreement on Abdominal Exam

	Kappa	95% Confidence Interval
Location of Max. Tenderness	0.52	0.38 to 0.65
Murphy's Sign	0.49	0.23 to 0.75
Presence of Peritoneal Irritation	0.47	0.03 to 0.90
Movement Pain	0.30	0.01 to 0.60
Guarding	0.26	-0.02 to 0.55
Rebound	0.24	-0.04 to 0.53
Percussion Sign	0.11	-0.19 to 0.40
Distention	-0.04	-0.24 to 0.16
Bowel Sounds	0	0
Rigid Abdomen	0	0

Overall agreement on the location of maximum tenderness improved when partial credit was given for near-agreement, as assessed by the weighted kappa using weight scheme outlined in section 4.11.2. The overall weighted kappa was **0.70**. Simple kappa statistics for agreement on each location of abdominal pain are presented in **Table 16**. The right lower quadrant, right upper quadrant and epigastric locations showed moderate to good reliability.

Table 16: Objective 2.1, Agreement for Location of Pain

	Kappa	95% Confidence Interval
RLQ	0.74	0.53 to 0.95
RUQ	0.65	0.20 to 1.00
Epigastrum	0.64	0.37 to 0.90
Supra-Pubic	0.45	0.01 to 0.90
L Flank	0.37	-0.19 to 0.93
LLQ	0.33	-0.08 to 0.74
Peri-Umbillical	-0.03	-0.08 to 0.01
R Flank	-0.04	-0.09 to 0.02
LUQ	0	0

5.4.3 Objective 2.3: Relationship between Opiod Analgesic use and Serious Outcome

The occurrence of serious outcomes was compared between patients who received narcotic analgesics and those who did not in **Table 17**. A significant difference between those receiving narcotics and those who did not in terms of eventual serious outcome rates was found by χ^2 analysis. Due to the observational nature of this study, however, no causal inference can be made.

Table 17: Objective 2.3, Serious Outcome by Analgesic Status

Serious Outcome	No Analgesic		Received Analgesic		Total	
	N	%	N	%	N	%
No	576	94.3	219	87.6	795	92.3
Yes	35	5.7	31	12.4	66	7.7
Total	611	-	250	-	860	-

χ^2_1=15.7 p<=0.001

5.4.4 Objective 2.4: Rate of Positive and Negative Laparotomy

The baseline rate of laparotomies with or without surgical pathology (positive or negative) among all patients is presented in **Table 18**. The negative laparotomy rate was low, with surgical pathology absent in 5.8 %.

Table 18: Objective 2.4, The rate of positive and negative Procedures

Total that underwent Surgery	87/860 (10.1 %)
Positive Procedures	82/87 (94.2%)
Negative Procedures	5/87 (5.8%)
Surgery after discharge from ED	6/666 (1.3 %)

The type of surgery performed is outlined in **Table 19**. One patient underwent multiple procedures. At laparotomy, the patient was discovered to have metastatic cancer, and received a bowel resection and an incidental appendectomy. This patient was classified as a bowel resection, as this was the operations most directly related to the primary underlying disorder. All laparoscopic procedures (laparascopic cholecystectomy, appendectomy) were classified as "laparoscopy".

Table 19: Description of Procedures

	Positive	Negative	Total
Exploratory Laparotomy	13	0	13
Appendectomy	34	2	36
Cholecystectomy	4	0	4
Bowel Resection	5	0	5
Laparoscopy	8	2	9
Dilatation & Curettage	4	0	4
Colonoscopy	3	0	3
Upper Endoscopy	3	1	4
ERCP	4	0	4
Hernia Repair	1	0	1
Nephrostomy	1	0	1
Exam Under Anesthetic	1	0	1
Angiogram	1	0	1
Total	82	5	87

5.4.5 Objective 2.5: Rate of Subsequent Return Visits and Hospitalization

The total number of patients who returned to the Emergency Department after discharge from hospital, as determined by telephone follow-up at 3 weeks, is presented in the following table. There was a trend towards higher re-admission rate among those who had received narcotics, however the small number of occurrences in each group was small.

Table 20: Objective 2.5, Rate of Subsequent Return Visits and Hospitalization among Discharged Patients

	Frequency, Overall N=677	Frequency, Given Narcotics N=677	Frequency, No Narcotics N=677
Return Visit to ED[1]	43	10	33
Subsequently Hospitalized[2]	9	4	5
Loss-to Follow-up Rate[3]	9	3	3

[1] χ^2_1=0.12, p=0.91
[2] χ^2_1=3.40 continuity corrected, p= 0.06
[3] χ^2_1=1.04 continuity corrected, p=0.31

5.4.6 Objective 2.6: Type of Narcotic Analgesic Used

A variety of narcotic analgesics were used at the study center, including demerol, fentanyl, codeine and morphine. The frequency, dose range, and median dose of each narcotic given are presented in **Table 21**. Meperidine was by far the narcotic most often given at the study center, and usually at an initial dose of 25 mg.

Table 21: Objective 2.7, Narcotic Choice, and Dose

Narcotic Type	Frequency N=249	%	Dose Range	Median Initial Dose
Meperidine	206	82.7	12.5-100 mg	25 mg
Morphine	8	3.2	2-5 mg	3.75 mg
Fentanyl	24	9.7	25-150 μg	87.5 μg
Codeine	10	4.0	60 mg	60 mg
Leritine	1	0.4	50 mg	50mg

5.4.7 Objective 2.8: Reason for Withholding Narcotic Analgesic

The specific reasons physicians gave for withholding narcotics are examined in **Table 22.** Physicians gave "Pain not severe enough" as the reason for withholding narcotics most frequently, in 48.7% of cases. In 13.7% of cases, physicians stated they would try another treatment for abdominal pain before giving narcotic analgesics.

Table 22: Objective 2.8, Physicians Reasons for Withholding Narcotics, and Mean Pain Severity

Reason	Frequency N=437	%	MD Mean Pain Score (mm)	Patient Mean Pain Score (mm)
Will Order Narcotic	129	29.5		
Pain not Severe enough	213	48.7		
Other Treatment	60	13.7	42.5	60.6
Other Reason	13	3.0	62.2	57.8
Consultant May Object	11	2.5	72.0	77.3
May Mask Diagnosis	11	2.5	45.8	73.2

The type of "other treatments" attempted are presented in **Table 23.** Antacids were used most often, followed by anti-spasmodic treatments.

Table 23: List of " Other Treatments" Attempted

Other Treatments Attempted	Frequency N=60	%
Antacids +/- Viscous Lidocaine	25	41.6
Antispasmodic	8	13.3
Anti-Emetic	5	8.3
Rehydration	4	6.6
Antibiotics	3	5.0
No other treatment given	3	5.0
Ketorolac	2	3.3
Laxatives, Enemas	2	3.3
Catheterization	1	1.6
Miscellaneous	7	11.6

Finally, four "other" reasons given by physicians for withholding narcotics in 13 patients. In 10 cases, physicians stated that the patient had declined narcotics. One physician stated analgesics "May slow management process", one physician withheld narcotics due to " Borderline blood pressure", and in one case "Patient getting better" was reason for withholding narcotics.

5.4.8 Objective 2.9: Rate of Loss To Follow-Up

Overall loss to follow-up was low: records were located for all admitted patients, and only 9/677 discharged patients could not be contacted by telephone. Median time to phone contact was 16 days, with 75% of patients contacted by 18 days. Proxy respondents were used in only 89/677 attempted phone contacts (13%).

6. Discussion

6.1 Demographics: Overall Patient Population

The study eligibility criteria were designed to simulate those of a randomized clinical trial of the safety of narcotic analgesics in acute abdominal pain. The patient population captured by the current study was intended to be representative of the patient population that would be enrolled in a future trial. Is the patient population of the current study typical of the acute abdominal pain population in general ? To address this question, a comparison to previous studies of acute abdominal pain is presented below. Large discrepancies between the patient population of the current study and those found in the literature would have implications for the generalization of the findings of the current and proposed future studies to other settings.

As discussed in **section 3.3** of the literature review, only one abstract[26] and two clinical trials [24,25] have prospectively addressed safety of narcotics in acute abdominal pain. No demographic data are reported in the abstract[26] or the study by Zoltie[25] et al. to allow comparison. Attard[24] et al. studied only inpatients; none the less, gender distribution and mean age were similar when compared to the findings of the current study. Females accounted for 52% of the Attard study, compared to 61% in the current study, and their median age was 47 years compared to 40 years in this study. The most common diagnosis in the Attard trial was non-specific abdominal pain (22%), followed by appendicitis, perforated peptic ulcer, and biliary colic. Non-specific abdominal pain, or abdominal pain not yet diagnosed, was also the most common diagnosis in the current study accounting for the same

proportion of patients as the Attard study. However, gastro-enteritis was the second most common diagnosis in the current study, and biliary colic and appendicitis were less common diagnoses. This finding is not surprising, considering all patients in the Attard study had been referred and were all admitted.

In addition to the clinical trials, one observational study from an American university hospital emergency department documented the age distribution and diagnoses from 1000 consecutive cases of abdominal pain.[19] The authors report that 40% of patients were between 15 and 24 years of age, but unfortunately do not report the gender breakdown. Abdominal pain not yet diagnosed was the most common diagnosis there as well in 41%, followed by gastroenteritis in 6.9% and pelvic inflammatory disease in 6.7%. Appendicitis accounted for 4.3% of their patients and 27.4% of their patients were admitted to hospital.

No marked difference in demographics was found between the current study and the previous U.S. based studies. However, usage of emergency departments in Canada has been found to differ from usage patterns in the U.S.: in general, a higher proportion of the Canadian population uses the emergency department for primary care and have a lower acuity of disease compared to their U.S. counterparts, presumably due to universal insurance coverage. Therefore, there might be concerns regarding the applicability of findings of this study to American settings. If differences do exist between Canadian and US Emergency department populations, what would be the impact of such differences ? If fewer patients with acute disease present to Canadian Emergency Departments, the serious outcome rate determined in the present study might under-estimate similar rates in the U.S..

The high exclusion rate of patients with abdominal pain is a potential weakness of the current study: nearly three quarters of subjects with "abdominal pain" appear to have been excluded from the study. However, this is in part due to the definition of abdominal pain used and the manner in which data on ineligible patients were collected. On initial presentation to the emergency department, all patients were asked for their main complaint. The presenting complaints of all patients were then recorded on the emergency registration log, which was used to screen for potentially eligible patients which had not been enrolled by physicians. Any patient found on the registration log with a presenting complaint of abdominal pain or any other symptom referable to an abdominal or pelvic organs was screened by the research assistant for eligibility. All subjects thus screened were designated as "abdominal pain"; the definition of abdominal pain for screening purposes was deliberately over-inclusive to prevent missing eligible subjects. Thus, in 23% of patients screened for eligibility, chart review revealed that abdominal pain was not the primary complaint. In addition, it had been decided at the outset to focus on acute abdominal pain: symptoms greater than 72 hours was the most common reason for exclusion, accounting for 33% of abdominal pain. Although the time-frame used to define acute pain was arbitrary, it does reflect physicians perceptions of which patients are likely to experience acute complications.

Patients were entered into the present study 24 hours a day, seven days a week. Inter-physician agreement data was not collected from midnight to 8:00 a.m., as there was only one staff physician on duty. Thus, the results should not be influenced by diurnal, nor "day of the week" variation.

Thus, the patient population from the current study appears comparable to those reported in previous investigations into acute abdominal pain, and the exclusion criteria were not overly restrictive. Therefore, the findings of this study should be generalizable to the wider population of Canadian emergency department patients with abdominal pain for future research.

6.1.1 Demographics and Comparability of Patients With and Without Questionnaires

Questionnaires were returned for only 55% of patients. Do important differences exist between the two sub-groups of patients with and without questionnaires necessitating treating their data separately, or can the entire sample size of 860 patients be considered a single population ?

The groups were similar with respect to gender, admission rate, serious outcome rate, and operative rate. Although there was a statistically significant difference in mean age, the 3 year difference is not clinically important. However, the two groups did differ in emergency department diagnoses and rate of narcotic administration.

The differences in diagnosis between the two groups can be characterized by higher rates of appendicitis and biliary colic, among the questionnaire group, than in the group without returned questionnaires. Also, the frequency of urinary tract infections was lower in the questionnaire group compared to the group without questionnaires. These three diagnoses contributed three quarters of the total deviation from expected proportions in the overall χ^2 statistic. The higher rate of appendicitis is not surprising as physicians may have been more careful to complete questionnaires when they suspected the potentially serious

diagnosis of appendicitis. Patients with biliary colic typically have severe pain and dramatic presentations, and this may have reminded physicians to complete the form. Urinary tract infections, on the other hand, may present with other symptoms in addition to abdominal pain, and thus physicians may have believed such patients to be ineligible. Physicians may also have been reminded to complete questionnaires for patients when they considered ordering analgesics, accounting for the difference in narcotic administration rate between the two groups. Pain severity unfortunately can not be compared between the two groups, since this data was missing by definition from the group that did not receive a questionnaire.

Thus overall, few differences existed between the groups of patients with and without questionnaires. What differences did exist seem explainable by increased physicians compliance among patients of great interest to the study: patients with suspected appendicitis and biliary colic. Such patients are more likely than others to have an uncertain diagnosis while at the same time requiring analgesics due to pain severity. Thus, the controversy regarding the safety of narcotics amongst these patients is especially salient in this patient population. Conversely, it appears that physicians were selectively non-compliant in patients with suspected urinary tract infection who had abdominal pain as well. During data collection, several physicians commented that they did not think such patients should be eligible for the study, despite the fact that technically, they met all study eligibility criteria. Urinary tract infection rarely requires analgesics, and the diagnosis is often clear to the physician. There is a strong argument that patients in whom there was a strong clinical suspicion for the diagnosis of urinary infection should be excluded from future clinical trials.

Any differences that did exist between subjects with and without questionnaires will have had the greatest impact on the data contained within the questionnaire itself: namely, physicians presumptive diagnoses, assessment of peritoneal irritation, intention to order narcotics, reasons for withholding narcotics, and pain severity assessments. The potential effect of selection bias will be discussed for each of these questions separately below.

6.2 Principal Outcomes

6.2.1 Serious Outcome Rate

Among eligible patients, the rate of serious outcomes was eight times greater than the 1% of all eligible patents predicted in **section 4.10**, and thus should be feasible as the primary endpoint for a randomized clinical trial into the safety of narcotics in acute abdominal pain. No previous study was found that addressed the rate of serious outcomes in an undifferentiated abdominal pain population to allow direct comparison. One study did examine serious outcomes at three weeks among emergency department patients discharged with a diagnosis of non-specific abdominal pain[2]. Of 403 such patients, 3.5% required subsequent admission, only 0.5% experienced a serious outcome. This lower rate of serious outcomes is not surprising given the pre-selection for patients with non-specific abdominal pain.

One potential weakness of the current study was the definition of serious outcomes employed. Although carefully thought out, this definition had not been previously validated nor is it an accepted standard definition: no such outcome criteria exists. The rationale

behind the definition chosen for this study, and alternative outcomes that were rejected, are discussed below.

Mortality rate is often used as a primary outcome in medical research; one of the fundamental pursuits of medicine is to prevent untimely death. In addition, it has the advantages of being objective, and not prone to misclassification bias. Potential for increased mortality is of great concern to physicians deciding whether or not to administer narcotic analgesics in acute abdominal pain. Two problems prevent the use of mortality alone as the sole endpoint. First, mortality was anticipated to be rare in acute abdominal pain, thus it was questionable as to whether the current study would have sufficient precision to be informative about the true mortality rate. More importantly, while mortality is a necessary element in proof of the safety of narcotic analgesics in this setting, it is not sufficient. Clearly, physicians, allied health care workers providers, and the patients themselves would also concerned about any additional morbidity associated with the use of opioids in undiagnosed abdominal pain. The definition of serious outcomes in this study attempted to encompass the range of complications that might occur as a result of delayed or missed diagnosis due to narcotic analgesics. However, as such, the definition is arbitrary and open to criticism.

For example, the inclusion of partial bowel obstruction as a serious outcome might be questioned, as such patients often respond to conservative therapy alone. However, the definition of serious outcomes was designed to be convincing to clinicians who are currently skeptical about the safety of narcotics in undiagnosed abdominal pain. Thus, partial obstruction was included in the definition of serious outcome because it was felt that many

physicians would be unwilling to administer narcotics if this resulted in higher rates of partial obstruction.

Another potential criticism of the definition employed is that there was no attempt to directly link serious outcomes as being caused by the administration of narcotics. However, attempts to create an operational definition of "serious outcome due to use of analgesia" that was not susceptible to bias by the rater failed. Given a randomized controlled trial, where the only factor consistently varied between the experimental and control groups is use of narcotic analgesia, any differences between the two groups could then be attributed to the use of analgesia.

6.2.2 Need for Opioids Analgesics: Definition Controversy

The second fact needed to determine the feasibility of the proposed randomized clinical trial was the proportion of patients with abdominal pain who required analgesics. Overall, approximately one third of patients with acute abdominal pain required narcotic analgesia, although the estimate varied depending on how "need for analgesia" was defined. Using the physician's assessment alone produced an estimate 20% lower than the most liberal definition.

As mentioned in section 4.9.2, "Primary Outcomes: Narcotic Requirement", there is no universally accepted standard to determine who requires narcotic analgesics. For the purpose of determining eligibility in any future randomized clinical trial, it was decided that the physicians opinion on analgesic requirement should be used, because physicians would ultimately determine patient eligibility in such a future trial. However, it was recognized that at present, some physicians withhold narcotics because of fears of masking the patients

diagnosis. To avoid missing this group of patients that would be eligible for the proposed clinical trial, physicians were asked to indicate their reason for withholding narcotics. Patients in whom narcotics were withheld because of concerns of masking the diagnosis, or that a consultant might object, were included in the estimate of potentially eligible subjects. As well, those cases in whom the physician stated that although the patients pain was severe enough to warrant a narcotic analgesic, they would try another treatment first were also considered as eligible for the proposed clinical trial. Finally, in the remaining patents where "Other" was given as the reason for withholding narcotics, physicians felt a narcotic was indicated, but patients stated they did not wish narcotic analgesics. Such patients would also be eligible for the proposed randomized trial, although this 2 % of patients may represent subjects who would decline to participate. In total, 20% of subjects had narcotics withheld by physicians despite having sufficient pain severity to merit analgesia. Using the physicians assessment in the subgroup of patients with returned questionnaires, and taking into account physician's reasons for withholding narcotics, 51% of subjects required narcotic analgesia.

However, questionnaires were returned on only 52% of patients in this study. Furthermore, patients with completed questionnaires were more likely to have actually received narcotics than subjects without questionnaires (34% vs. 24% respectively, $p<0.001$). Given this selection bias, the proportion of subjects with completed questionnaires requiring narcotics may not be representative of the prevalence requiring narcotics in the abdominal pain population overall. What is the impact of such a selection bias ?

If we assume an extreme scenario in which all subjects requiring narcotic analgesics from the entire abdominal pain population had questionnaires completed, and none of the

subjects without questionnaires required narcotics, then the rate of narcotic requirement would be calculated by taking these 224 patients and dividing by the entire study population of 860 for a narcotic requirement rate of 26%. However, we are not completely without data on the narcotic requirements of the 410 subjects in whom the physician did not answer the question "Will you order a narcotic analgesic?". As shown in **Figure 8**, 97/410 (24%) of subjects without a questionnaire actually received a narcotic. Thus, combining the 224 with a questionnaire plus the 97 subjects without questionnaire that actually received narcotics, a minimum of 321/860 (37%) required analgesia, assuming that no subjects in the group without questionnaires that required narcotics were denied analgesics. Such an assumption is conservative, given that 20% of subjects who physicians felt required narcotics had narcotics withheld in the questionnaire group. It is interesting to note that this 37% composite estimate of narcotic requirement is similar to the proportion of patients that felt they themselves needed a strong pain killer.

Thus, the 51% rate of narcotic requirement found in the questionnaire group is likely an over-estimate, due to selection bias. However, the 26% of patients who received narcotics is an under-estimate of the true rate of narcotic requirement, and the composite measure of 37% probably best estimates the true rate of abdominal pain severe enough to require narcotics.

6.2.3 Serious Outcome among Patients Requiring Opioids Analgesics: Implications for Future Randomized Clinical Trials

The proportion of all emergency department patients who were eligible for the study, (2.71%), was slightly lower than the 2.85% to 4.75% anticipated in section **4.10**. Considering the exclusion criteria used, the results of the current investigation are consistent with many previous studies which have found abdominal pain represents between 4-10% of the total emergency department census in both the U.S. and Canada.[1,2,3,4,5,6] No previous study has used eligibility criteria similar to the current investigation.

Approximately 1% (321/31,772) of patients seen during the study period satisfied eligibility criteria and met the definition of narcotic requirement given in section **4.9.2.4**. The serious outcome rate was 11.2% among patients meeting this definition of narcotic requirement (95% Confidence Interval, 8.1-15.3%). Given the proportion of emergency department patients who would be eligible, and the rate of serious outcomes among such patients, what sample size would be required to assess the safety of narcotics and placebo with adequate precision in the proposed clinical trial?

To perform the sample size calculations, the probability of type I and type II errors that will be tolerated need to be chosen. As discussed in section **4.11**, the null hypothesis for equivalence trials differ from standard trials, and this influences the probability of type I and II error used. The null for the proposed equivalence trial stipulates that patients in the narcotics group would experience a serious outcome rate greater than placebo by the minimum amount considered clinically important. Rejection of this null allows acceptance of the alternate hypothesis that the serious outcome rates are equivalent. The commonly accepted 5%

probability of type I error was used in sample size calculations. A type II error in the context of the proposed equivalence trial would occur if it was concluded that there was a higher rate of serious outcomes with narcotic analgesic use when none truly existed.

The following table presents sample size calculations under various assumptions regarding the baseline rate of the proposed primary endpoint, serious outcomes. Required sample sizes are calculated assuming a serious outcome rate corresponding to the lower 95% confidence boundary of 8.1%, the point estimate of 11.2% and the upper 95% confidence limit of 15.3%. As well, sample sizes were calculated within each strata of baseline rate of serious outcomes using relative differences of 25% and 33% as the "clinically significant" difference in rate of serious outcomes detectable. Sample size calculations were based on **Equation 1**, given in section **4.11**.

Table 24: Projected sample size requirements for proposed randomized clinical trial

Baseline Serious Outcomes	Z Alpha	% Type I Error (1 tailed)	Relative Difference Detectable	Absolute Difference Detectable	Z Beta	% Type II Error	Sample Size
8.1%	1.645	5.0%	25%	2.03%	0.842	20.0%	4492
8.1%	1.645	5.0%	33%	2.70%	0.842	20.0%	2526
11.2%	1.645	5.0%	25%	2.80%	0.842	20.0%	3130
11.2%	1.645	5.0%	33%	3.74%	0.842	20.0%	1760
15.3%	1.645	5.0%	25%	3.83%	0.842	20.0%	2192
15.3%	1.645	5.0%	33%	5.10%	0.842	20.0%	1232

The detectable difference in serious outcome rate chosen has more influence on determining sample size than the baseline rate of serious outcomes. Decreasing the detectable difference by half quadruples the sample size, whereas decreasing the baseline rate of serious outcomes by half only doubles sample size requirements. Attention should be focused on

what clinicians would accept as a "clinically important" difference in serious outcome rate detectable.

The scenarios presented above demonstrate that a clinical trial into the safety of narcotic analgesics that randomized 3200 patients could detect a relative difference of approximately 33% between treatment groups under a variety of assumptions. Abdominal pain requiring narcotics represented approximately 1% of all emergency department visits, thus 330,000 emergency departments visits would be required to meet study sample size requirements. Such a randomized clinical trial could be conducted at three emergency departments with annual censuses of 55,000 in approximately two years.

6.3 Secondary Outcomes

6.3.1 Physician and Patients Assessment of Narcotic Requirement and Pain Severity

It is often stated in the literature that physicians under-estimate patients' analgesic requirements, however few studies have attempted to quantify this disagreement, and no study has examined patient-physician agreement on need for narcotic analgesics in the acute abdominal pain population.[15,16,38,39,40,41,42,44,45,46,47] In one study of post-operative pain control, 41% of patients still had moderately severe pain after assessment and treatment with analgesia.[15] In a study of 37 consecutive patients admitted to a medical floor for painful conditions, 31% had severe distress despite being treated for 48 hours with intramuscular narcotics. In a review of 198 Emergency Department patients with a variety of painful conditions,[40] only 44% received any analgesics. Looking specifically at patients with a diagnosis of an intra-abdominal condition, 41% were given analgesics. Thus, the current

literature suggests many patient with painful conditions are not currently treated with analgesics, but the documentation of patient-physician agreement on need for analgesics is limited.

Patient-physician agreement on need for analgesics in the current study was only moderate, correcting for chance agreement, with a kappa of 0.53. Physicians stated they would withhold narcotic in 39% of the 147 patients who stated they needed a strong pain killer. Agreement did not improve when physicians reasons for withholding narcotics were considered.

The results of this study confirm the sub-optimal agreement between patients and physicians regarding need for narcotics found in previous studies.[15,16,38,39,40,41,42,44,45,46,47] But what are the reasons behind this poor agreement ? Todd[89] describes a theoretical framework of pain assessment and treatment that may be helpful in understanding how patient-physician disagreement on pain severity and treatment arise. The patient must first perceive the pain, and externally express their experience of pain to the physician. This expression of pain is then assessed by the physician, who integrates the information and makes decisions regarding pain management. Finally, the patient reacts to the physicians treatment. Events during each of these stages may contribute to the overall patient-physician disagreement. First, the amount of pain an individual patient experiences from a given injury may vary greatly from "average", and this may contribute to physician under or over-estimation of pain severity. Second, the patient may not express their pain to the physician, due to a stoic personality or communication barriers. Conversely, the physician may not be adequately receptive to the pain being expressed by the patient, due to time pressures or inattention. Finally, given that

the patient and physician do agree on pain severity up to this point, they may still disagree regarding how to manage the pain: physician or patient concern regarding side-effects may outweigh perceived need for analgesics, for example.

A further difficulty with the interpretation of agreement in pain severity is the lack of a validated tool to compare pain severity assessments by different raters. Although the visual analog scale has been used extensively in experimental pain studies and in analgesics studies, no studies have validated the visual analog scale as a measurement tool to compare pain severity perception between patients and physicians. Emergency physicians have, in general, witnessed a wider range of pain severity than patients, therefore it is likely that their respective definitions of "unbearable pain" will differ. Given such differences in scale, even if a patient and physician agree on pain severity, this may not translate into similar measures on the visual analog scale. A proposed method for developing and validating a comparative pain severity tool are discussed at the end of this section.

Similarly, the comparison of the patients' perception of narcotic requirement by the question "Do you need a strong pain killer" to physicians' perception assessed by the question "Will you order a parental narcotic in this patient?" may have introduced measurement error. As discussed in section **5.3.2**, the question "Do you need a strong pain killer" was used in accordance with standard survey research methodology, which suggest a grade 4-6 reading level for survey questions intended for the general population; otherwise non-response rates are excessive. An alternative approach which had been considered was to have physicians ask "Do you need a narcotic pain killer ?" and then have physicians explain what a narcotic was. However, it is uncertain that the patients' and the physicians' understanding of "narcotic

analgesics" would be the same even after explanation. Thus, it was felt that the incremental time that such an explanation would require was not justified.

Did the disagreements between patients and physicians regarding pain severity and need for narcotics in this study represent actual differences in perception of pain severity, or different interpretations of the questions used to assess narcotic need ? Clearly, both factors are involved in the overall disagreement, but scrutiny of the data suggest that differences in perception of pain severity do contribute to disagreement regarding need for analgesics. Physicians pain severity scores were less than patients' self assessments in three quarters of cases, and this discrepancy was as large as 30 mm in a third of comparisons. Differences in pain severity scores were higher when patients felt they needed a strong painkiller, but physicians withheld narcotics. In the sub-group where physicians reason for withholding narcotics was that they felt the patients pain was not severe enough, but the patient felt they needed a strong pain killer, mean differences in pain severity assessments were the highest of all at 32.5 mm. Although agreement on pain severity as assessed by the intraclass correlation coefficient was greater than expected by chance, 0.60, it was less than the recommended minimum value considered meaningful of 0.75.[90] Note that agreement expressed as the intraclass coefficient was again much lower in cases where the patient felt a narcotic was needed but the physician disagreed (0.14). The responsiveness of the visual analog scale under these different conditions suggests a true difference exists in pain severity perception rather than purely being an artifact of different use of the VAS by patients and physicians.

Thus, patient-physician disagreement on need for narcotics was at least in part due to different perceptions of pain severity. Was differing opinion on how to manage pain an

important contributor to patient-physician disagreement as well ? In other words, did physicians withhold narcotics from patients who had pain severe enough to require such narcotics in the physicians own assessment ? Fear of masking the diagnosis or that a consultant might object was the reason physicians withheld narcotics in only 7% of cases where this question was answered (see **Table 22**). The physicians mean pain severity scores was 58.9 mm in this group compared to a mean pain score of 33.6 mm when physicians did not think narcotics were indicated. Thus, in a minority of cases where physicians and patients agreed that pain was sufficiently severe that narcotics were required, physicians none the less withheld analgesics.

Further research is required to better define what underlies patient-physician disagreement regarding pain severity and need for analgesia. Specifically, a validated measurement tool that allows comparison of patients' and physicians' pain severity perception is needed. This may involve validation of the comparison of patient and physician visual analog scales, using the same methodology used in the validation of non-comparative use of the visual analog scale. However, a new multidimensional psychometric measurement tool may need to be developed and validated, if the visual analog scale proves to be inadequate as a comparative measure.

The results of the current study could then be verified to determine if true differences in the perception of pain severity exist. Finally, the reasons for such disagreement should be explored, thus allowing improvement in the patient-physician interaction.

6.3.2 Objective 2.2: Inter -Physician Agreement on the Abdominal Exam

Inter-physician agreement on the abdominal exam was only moderate. Location of maximum tenderness was the most reproducible physical sign The agreement on location of maximum tenderness improved to 0.70, moderate to good, when partial credit was given for "near misses" via the weighted Kappa. Implicit in the weighting system used was the assumption that geographic proximity should be taken into consideration when evaluating maximum tenderness. However, the weights were arbitrarily assigned, and other weights (e.g. clustering according to the probability of tenderness location given suspected diagnoses) could be considered in future research.

Agreement on location of maximum tenderness varied for different locations; right lower quadrant pain, right upper quadrant pain and epigastric pain were the most reliable, whereas agreement on the remaining locations was not better than expected by chance. Right lower quadrant pain is a cardinal sign of appendicitis, thus physicians training in the examination of the abdomen focuses on this finding. Similarly, right upper quadrant and epigastric tenderness are also indicative of specific organ dysfunction (e.g. cholecystitis and gastric disorders). However, agreement may have been improved due to characteristic symptom patterns for disease affecting these locations. Repeated examination by physicians blinded to the patients history is one experimental design that could be used to explore this issue.

Murphy's sign followed by presence of peritoneal irritation were the next most reproducible signs, and had similar point estimate for inter-observer agreement in the moderate range. However, given the small number of positive cases for presence of

peritoneal irritation, confidence intervals for this sign were wide. Four of the physical signs tested had agreement no better than expected by chance: presence of distention, guarding, rebound, and the percussion sign.

The results of the current study confirm the poor to moderate inter-physician reliability of the abdominal exam found in two previous studies. [74,75] Excellent agreement had been found by Greene et al [75] for "deep tenderness to palpation", however, as no definition for this sign was reported in the abstract, it could not be reproduced in the current study.

The lack of a standard definition of physical signs is a potential reasons for poor inter-observer agreement. However, physicians in the current investigation were instructed on the physical findings during initial training sessions, and written definitions were included on study forms (see **Appendix 1:** Interobserver Reliability of Physical Exam

). Physicians were not given individual, bedside instructions, nor were the definitions re-enforced during the study, and this may have contributed to the poor agreement observed.

Physiologic and anatomical factors specific to the abdomen may also have contributed to the poor agreement in this study. The abdominal organs lie relatively deep within the peritoneal cavity, and are innervated by visceral sensory nerves. Both factors contribute to the poor ability of patients to describe and localize pain in the abdomen as compared to the exquisite ability to localize pain in the hand, for example.

The findings of this study are weakened by the small number of patients on whom inter-rater reliability was performed: only 44 patients were assessed by two physicians. Were these patients representative of the acute abdominal pain population in general ? There was

no difference in age, gender, final emergency department diagnosis, disposition, rate of narcotic administration, surgical rates, nor rates of serious outcomes between subjects who did and did not have inter-observer forms completed. Thus, although the sample size was small, it appears to be representative.

As previously noted by Greene et al. [75], the low baseline rate of inter-physician reliability for the abdominal exam should be taken into account in the planning of trials which intend to use abdominal findings as outcomes. Given that the best inter-observer reliability for a physical sign in the current study was only 0.52, (see **Table 15**) future randomized studies using these signs as outcome measures may be unable to detect clinically important differences between groups. The presence of peritoneal irritation is a logical choice as an outcome measure. however, future studies are needed to define inter-rater reliability of this sign more precisely than was possible in the current study.

6.3.3 Objective 2.3: Relationship between Narcotic Analgesic and Serious Outcome

Serious outcomes occurred among 12.4% of patients who received narcotic analgesics, compared to 5.7% of subjects who had not. In a previous randomized clinical trial by Zoltie et al., serious outcomes were not reported as an outcome measure. In the trial by Attard et al.[24], serious outcomes were only reported among the sub-group classified as incorrectly diagnosed. No serious outcomes occurred in the analgesic group, and one perforation and one obstruction occurred in the placebo group ($p < 0.0001$, Fisher's Exact test). Thus, no previous study has compared serious outcomes between analgesic and non-analgesic groups.

However, narcotics should not be considered as the cause of serious outcome based on the results of this observational study. "Confounding by indication" a form of "reverse causality bias", occurs when factors that are the result of a disease process are erroneously concluded to exert causative influence on an outcome. Is analgesic use in this study simply a marker of more severe pain and consequently a greater baseline risk for a serious outcome ? Or did narcotics actually result in more serious outcomes ?

Logistic regression analysis might have been used to attempt to control for confounding variables in the examination of the relationship between narcotic analgesics and serious outcome. However, variables that have previously been shown to be important predictors of serious outcomes, white blood cell count and temperature,[1] were not prospectively collected in the current study. Thus, logistic regression analysis was not performed. Although this might be a cost effective method of examining any association between narcotic analgesics and serious outcomes, the risk of residual bias would still have to be considered.

Thus, the effect of narcotic analgesics on serious outcomes in acute abdominal pain is still unknown. A randomized clinical trial would provide the most convincing evidence regarding the safety of narcotic analgesics in acute abdominal pain.

6.3.4 Objectives 2.4: Baseline Surgical Rates

The positive and negative rate of surgical procedures are potential outcome measures for comparing narcotic and placebo. The overall surgical rate was 10.1%, with surgical pathology found in 94.2% of these procedures. In a retrospective review of 1000 emergency

department patients presenting with abdominal pain to a U.S. teaching hospital in 1972, Brewer et al found 15% required surgery, and 86.6% of these operations found pathology "requiring immediate surgical intervention"[19]. Differences between the previous and current studies in the surgical rate and rate of positive findings at surgery may be attributable to secular trends towards non-operative management; improved imaging technology has become increasingly available which may have decreased the rate of negative laparotomies. As well, local practice patterns, or U.S.-Canadian differences in practice patterns, may have produced the discrepancy in negative operative rate, and future research should attempt to determine the negative operative rate in multiple centers.

6.3.5 Objectives 2.5: Return Visits and Subsequent Hospitalization

Forty-three of the 677 subjects initially discharged from hospital subsequently returned to the emergency department. No subjects reported receiving emergency care at another center. Of the 43 patients who returned, 9 were admitted, and six of these initially discharged subjects subsequently required an operation (0.7%). These results are similar to Brewer et al.[19], who reported 11 out of 1000 of their subjects (1.1%) had an acute surgical condition that was not recognized prior to discharge and subsequently required surgery.

The 34 patients who were discharged from the ED a second time were not followed up for subsequent events, and this is a potential weakness with the study. As well, the reasons patients returned to the ED were not explored.

The rate of discharged patients that subsequently returned to the ED, required admission, or required surgery, are potential principal or secondary outcome measures for

future clinical trials into the safety of narcotic. The rationale for using return visits as a principal outcome is that it is an objective measurable outcome, and may serve as a surrogate measure of misdiagnosis. However, quality of medical care is merely one of many factors that influence patients' satisfaction with emergency care and their decision to return to the Emergency Department. As well, worse medical care may paradoxically lower return visit rates, as subjects seek care elsewhere. Thus, while of interest as a secondary endpoint, return visits may not be suitable as the primary endpoint for a future randomized clinical trial. Subsequent hospitalization may be a better candidate for primary outcome in future clinical trials; patients requiring admission within two weeks of discharge from the emergency department are likely to have been misdiagnosed on their initial presentation. This is supported by the high rate, (66%) of return admission that required surgery in the current study. The rate of missed surgical pathology may be the best potential principal outcome for future trials on the safety of narcotics in abdominal pain, because it is objective, clinically significant , and the baseline rate is well described: Brewers' study and the current project have both found similar rates of 1.1% and 0.7% respectively. However, missed surgical pathology has the disadvantage of being relatively rare: studies using this as a primary endpoint would require sample sizes in excess of 10,000 to demonstrate a 33% relative difference. In addition, significant complications following the mismanagement of acute abdominal pain may occur which are not surgically treated (sepsis and partial bowel obstructions, for example).

6.3.6 Objective 2.6: Narcotics Dose And Route

Physicians at our center used meperidine the majority of the times a narcotic was given, and half of patients were given a 25 mg dose initially. A dose of 1.75 mg/kg is the recommended loading dose for the treatment of acute pain.[91] Thus, the 25 mg dose of meperidine most frequently used in this study would correspond to 0.4 mg/kg assuming a weight of 60 kg, and may have been insufficient to provide adequate analgesia. In the study of Zoltie et al.[25], 200 and 400 μg of buprenorphine were used sublingually (0.003 and 0.006 mg/kg); the recommended parenteral loading dose is 0.004 mg/kg. In the Attard study,[24] 20 mg of papaveritum (0.33 mg/kg) was administered, with a recommended loading dose of 0.4 mg/kg. Burke et al. titrated IV morphine to a maximum dose of 20 mg (0.3 mg/kg, recommended dose 0.25 mg/kg). A potential weakness of the study was that only the initial and not total dosage of narcotics used was recorded. Patients may have received further doses of analgesia, titrated until comfortable.

Future research should attempt to define and standardize what is meant by adequate analgesia for studies into acute pain. Currently, both standard dosage based on body weight and titrated doses of narcotics are reported in the literature. If it is accepted that dose titration should be employed, what should the end point of such titration be ? Should patients determine if more analgesic is necessary, should physicians make this decision, or should the definition of pain relief be determined by a pre-specified change on pain severity on the VAS ? Resolution of these issues would be important prior to undertaking the proposed randomized clinical trial.

6.3.7 Objective 2.7: Physicians Reasons for Withholding Narcotics

Physicians felt that pain was not severe enough to require narcotics in two thirds of cases where they withheld opioids. Physicians gave "Other treatment may relieve pain" as the second most common reason given by for withholding in narcotics. "Consultant might object" or "May mask diagnosis" was given as the reason for withholding narcotics only rarely. This is the first study to the authors knowledge which examined physicians reasons for withholding narcotics in the acute abdominal pain setting. Physicians stated the reason for withholding narcotics was that the patient did not wish analgesics in 3.2%, suggesting that in some cases, physicians felt a narcotic was required when the patient did not.

Assessing inner cognitive states, such as physicians reasons for withholding narcotics, is a difficult area to study, and state of the art methodology involves the development and validation of multi-dimensional psychometric tools.[92] Since this was a secondary objective, such resource intensive methods were not employed: rather, a single question stem with several response categories including an open ended "other" category was employed. Physicians may have been unwilling to admit their true motivation for withholding narcotics: as alluded to in **section 1**, withholding narcotics because of fears of masking a patients diagnosis is increasingly perceived as "old-fashioned" in the emergency medicine community. During study education sessions, the author emphasized that no evidence exists as to the safety or danger of using narcotics in acute abdominal pain, and reinforced that concerns over the safety of using narcotics in acute abdominal pain were valid. However, it is difficult to assess the success of these education sessions with the simple measurement tool employed.

A specific psychometric tool should therefore be development prior to further assessment of physicians reasons for withholding narcotic. The issue of "social undesirable" responses will need special attention when developing such a tool.

6.3.8 Objective 2.8: Loss to Follow-up: Potential Impact on Study Results

Losses to follow-up were low at 1% of all subjects. Even if all these subjects experienced a serious outcome, the impact on the primary endpoint would be minimal.

The loss to follow-up rate is in keeping with numerous publications from the same center, using a similar data-collection technique.[76,77,78] In contrast, Lukins et al.[2] found that 28% of patients discharged from an American teaching hospital were lost to follow-up. It is possible that patients did seek medical attention elsewhere and were reluctant to report this at phone follow-up, however it is unlikely that no patient would have reported this; if patients were dissatisfied with care, they presumably would have been more likely to complain to the telephone interviewer. However, patients admitted to other centers due to a serious outcome, would not have been available to provide telephone follow-up and were at greater risk of having been lost to follow-up. No proxy respondents indicated that the patient had been admitted to another center.

Finally, although searching the records of other local hospitals for patients that were lost to follow-up had been considered, the additional resources required to do so was considered unwarranted.

7. Conclusions

Abdominal pain severe enough to require narcotic analgesics is a common problem, and a significant proportion of patients with acute abdominal pain experience serious outcomes. This study also underlines that patient-physician agreement on need for analgesia was sub-optimal. Frequently, when the patient felt a strong pain killer was required, the physician disagreed. Further research is required to better describe the cause of this disagreement.

Inter-physician agreement on the abdominal exam was poor in general. Therefore, physical signs may not be useful as primary endpoints for future clinical trials, although subsequent research should estimate the reliability of the presence of peritoneal irritation with greater precision than was achieved in this study.

Physicians seldom admitted to withholding narcotics for fear of masking the diagnosis or because of concerns that a consulting service might object to the use of analgesics. However, the possibility of a Hawthorne effect can not be excluded, and it is likely that physicians reasons for withholding analgesics vary from center to center.

Although a higher rate of serious outcomes was observed among patients who received narcotic analgesics, the results of this observational study should not be interpreted as implying a causal influence of narcotics on serious outcome. Rather, it is likely that sicker patients, experiencing more pain, were given analgesics more often than other patients. A randomized clinical trial would provide the strongest evidence as to the effect of narcotics on patient outcomes.

At present, it is unknown whether patients who are currently denied analgesics are suffering needlessly, or whether patients currently given narcotic analgesics are being exposed to unwarranted risks. Thus it is important to definitely determine whether opioid analgesics can be given safely to patients suffering with acute abdominal pain. Is a randomized clinical trial comparing analgesics to placebo, and using serious outcome as the principle outcome, feasible ?

The results of this study suggest that such a trial ***is feasible***, although to properly address this question would require a large sample size and multicenter collaboration. Whether or not analgesics are found to be safe, the cost of such a study would be justified. If analgesics are indeed safe, the results of the study could be used to prevent unnecessary suffering among abdominal pain patients. However, if analgesics are unsafe, then the proposed trial could avert further needless morbidity.

INTER-OBSERVER QUESTIONNAIRE: PHYSICIAN #1

STAFF PHYSICIANS ONLY

Complete only when two staff physicians available

Date: (yy/mm/dd) ___/___/___

Patient's Last Name: _____

Unique #: _____

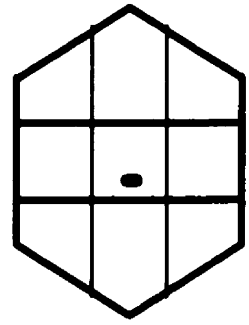
Physician #1 Initials: _____

INTER-OBSERVER ABDOMINAL EXAM: PHYSICIAN #1

Time of initial exam : _____

Distention	❑No	❑Yes
Bowel Sounds	❑Yes	❑No
Percussion Sign	❑No	❑Yes
Movement Pain	❑No	❑Yes
Guarding	❑No	❑Yes
Rebound	❑No	❑Yes
Rigid Abdomen	❑No	❑Yes
Murphy's Sign	❑No	❑Yes

Indicate Location of Maximum Tenderness (X):



Mark one box only please

(**Circle diagram** if non-tender)
(**Cross off** if uniformly tender)

Does the patient have peritoneal irritation ? : ❑No ❑Yes

How certain are you of this?

❑Very Certain ❑Certain ❑Neutral ❑Uncertain ❑Very Uncertain

Definitions:

Distention	Abnormal or new abdominal protrusion on inspection
Bowel Sounds	Any sounds present (recommend 30 seconds auscultation)
Movement Pain	Pain elicited by rocking patient or stretcher side to side
Percussion Sign	Pain elicited by percussion of abdomen
Guarding	Abdominal muscular contraction on palpation that persists during respiration
Rebound	Severe pain elicited or worsened by abrupt withdrawal of palpation pressure
Rigid Abdomen	Severe, diffuse, sustained abdominal muscular contraction on palpation
Murphy's Sign	Pain worsened by deep inspiration during RUQ palpation

PLEASE TEAR OFF FOLLOWING PAGE AND GIVE TO PHYSICIAN #2 TO COMPLETE

INTER-OBSERVER QUESTIONNAIRE: PHYSICIAN #2

STAFF PHYSICIANS ONLY

Complete only when two staff physicians available

Date: (yy/mm/dd) ___/___/___

Patient's Last Name: _____

Unique #: _____

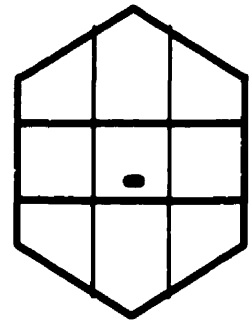
Physician #2 Initials: _____

INTER-OBSERVER ABDOMINAL EXAM: PHYSICIAN #2

Time of initial exam : _____

Indicate Location of Maximum Tenderness (X):

Distention	❑No	❑Yes
Bowel Sounds	❑Yes	❑No
Percussion Sign	❑No	❑Yes
Movement Pain	❑No	❑Yes
Guarding	❑No	❑Yes
Rebound	❑No	❑Yes
Rigid Abdomen	❑No	❑Yes
Murphy's Sign	❑No	❑Yes



Mark one box only please

(**Circle diagram** if non-tender)
(**Cross off** if uniformly tender)

Does the patient have peritoneal irritation ? : ❑No ❑Yes

How certain are you of this?

❑Very Certain ❑Certain ❑Neutral ❑Uncertain ❑Very Uncertain

Definitions:

Distention	Abnormal or new abdominal protrusion on inspection
Bowel Sounds	Any sounds present (recommend 30 seconds auscultation)
Movement Pain	Pain elicited by rocking patient or stretcher side to side
Percussion Sign	Pain elicited by percussion of abdomen
Guarding	Abdominal muscular contraction on palpation that persists during respiration
Rebound	Severe pain elicited or worsened by abrupt withdrawal of palpation pressure
Rigid Abdomen	Severe, diffuse, sustained abdominal muscular contraction on palpation
Murphy's Sign	Pain worsened by deep inspiration during RUQ palpation

9. Appendix 2: Telephone Interview

"Hello, this is _____ . I'm a research assistant calling from the Ottawa Civic Hospital to follow-up about your abdominal pain. Do you have two minutes to answer a few questions ?"

"Have you gone back to any Emergency Department for your abdominal pain ?"

If yes, "Which Emergency Department did you go to ?"

If yes "What day did you go back to the Emergency Department ?"

"Where you admitted to the hospital for your abdominal pain ?"

If yes, "Which hospital where you admitted to ?"

If yes "What day were you admitted to the hospital ?"

"Did you have an operation for your abdominal pain ?"

If yes, "Which hospital were you operated at ?"

If yes, "What day were you operated on ?"

"Thank you very much for your assistance "

10. Bibliography

[1] Graff L, Radford MA, Werne C. Probability of appendicitis before and after observation. Ann Emerg Med 1991;20:503-507.

[2] Lukens TW, Emerman C, Effron D. The natural history and clinical findings in undifferentiated abdominal pain. Ann Emerg Med 1993;22(4):690-696.

[3] McCaig LF. National hospital ambulatory medical care survey: 1992 Emergency Department Summary. Advance data CDC1994;245(710):1-12:

[4] Bertinieri G, Bonino P, Caser L, Ceffa GC, Martina C Calogero G. Il medico di pronto soccorso di fronte alla patolgia addominale acuta. Gastroenterologia 1987;78:1597-1600.

[5] Brown EM, Goel V. Factors related to emergency department use: results from the Ontario Health Survey, 1990. Ann Emerg Med 1994;24(6):1083-91.

[6] Hadas N, Eschar J, Kochavi D, Scapa E. Gastrointestinal patients in the emergency room. Israel J Med Sci 1995;31(11):700-1.

[7] Cope Z. Early diagnosis of the acute abdomen. 15th Ed., revised Silen W., New York: Oxford University Press, 1979

[8] Quayle J. Analgesia and the acute abdomen. BMJ 1979;ii:1290.

[9] Editorial (Anonymous). Analgesia and the acute abdomen. BMJ 1979;ii: 1093.

[10] Angell JC. A touched up text. BMJ 1979;i:485-486;

[11] Hughes TJ. Opiates in acute abdominal pain. (letter) BMJ 1979;ii:1145.

[12] Shepard J. Analgesia and the acute abdomen. (letter) BMJ 1979;ii:1289.

[13] Quayle J. Analgesia in the acute abdomen. BMJ 1979;ii:1290.

[14] Hamilton WAP. Analgesia and the acute abdomen. BMJ 1979;ii:1363.

[15] Sriwatankul K, Weiss OF, Alloze JL et al. Analysis of narcotic analgesic usage in the treatment of postoperative pain. JAMA 1983;250:926-929.

[16] Perry S, Heidrich G. Management of pain during debridement: A survey of United States burn units. Pain 1982;13: 267-280.

[17] OVID Version 6.1. OVID Technologies, 1996: New York.

[18] Staniland JR, Ditchburn J, de Dombal FT. Clinical presentation of acute abdomen: Study of 600 patients. BMJ 1972;3;393-398.

[19] Brewer RB, Golden GT, Hitch DC, Rudolf LE, Wangensteen SL. Abdominal pain. An analysis of 1,000 consecutive cases in a university hospital emergency room. Am J Surg 1976; 131;219-223.

[20] Bugliosi TF, Meloy TD, Vokov LF. Acute abdominal pain in the elderly. Ann Emerg Med 1986;19(12):1383-1386.

[21] Weil TP. Clinton's health reform and emergency department volumes. A return visit. Ann Emerg Med 1993;22;852-54.

[22] Lavallee M. Analgesia and the acute abdomen. BMJ 1979;ii:1289.

[23] Walker AP, Condon RE. Peritonitis and Intra-abdominal Abscesses. In Shwartz SI, Shires GT, Spencer FC. *Principles of Surgery* 5th ed, New York, McGraw-Hill 1989; 1472.

[24] Attard AR, Corlett MJ, Kidner NJ, Leslie, AP, Fraser IA. Safety of early pain relief for acute abdominal pain. BMJ 1992;305:554-556.

[25]Zoltie N, Cust MP. Analgesia in the acute abdomen. Ann Royal Coll Surg Eng 1986;691:209-210.

[26] Burke T, Mack S, Coward B, Pace S. The use of intravenous morphine for early pain relief in patients with acute abdominal pain. (abstract) Acad Emerg Med 1994;1:A51.

[27] Rusnack RA, Borer JM, Fastow JS. Misdiagnosis of acute appendicitis: Common features discovered in cases after litigation. Am J Emerg Med 1994;12;397-402.

[28] Jones PF. Early analgesia for acute abdominal pain. (letter) BMJ 1992;6835(305):554-556.

[29] Fleiss JL. Determining sample size. In *Statistical methods for rates and proportions*. 2nd ed, New York: John Wiley & sons 1981; 260.

[30] Boey JH. The acute abdomen. In: Way LH, ed. *Current surgical diagnosis and treatment.* 9th ed., Connecticut: Prentice Hall, 440.1991;

[31] Hobson T, Rosenman D. Acute appendicitis-when is it right to be wrong? Am J Surg 1964;108; 306-312.

[32] Howie JGR. The place of appendectomy in the treatment of young adult patients with possible appendicitis. Lancet 1968; ii:1365-67.

[33] Kazarian K, Roeder WJ, Mersheimer WL. Decreasing mortality and increasing morbidity from acute appendicitis. Am J Surg 1970;119:681-685.

[34] Lewis AR, Holcroft JW, Boey J, Dunphy E. Appendicitis A critical review of diagnosis and treatment in 1,000 cases. Ann Surg 1975;110:677-683.

[35] Shwartz SI. Appendix. In Shwartz SI, Shires GT, Spencer FC. *Principles of Surgery* 5th ed, New York, McGraw-Hill, 1324.1989;

[36] Howie JG. Death from appendicitis and appendectomy. Lancet 1966;ii:1334-1337.

[37] Howie JGR. The place of appendectomy in the treatment of young adult patients with possible appendicitis. Lancet 1968;ii:1365-1366.

[38] Marks RM, Sachar EJ: Undertreatment of medical inpatients with narcotic analgesics. Ann Int Med 1973;78:173-181.

[39] Selbst SM, Clark M. Analgesic use in the emergency department. Ann Emerg Med 1990;19:1010-1013.

[40] Wilson JE, Pendelton JM. Oligoanalgesia in the emergency department. Am J Emerg Med1989; 7(6):620-623.

[41] Paris PM. No pain, no pain. Am J Emerg Med 1989;7(6):660-662.

[42] Read JV. Perceptions of nurses and physicians regarding pain management of pediatric emergency room patients. Ped Nurs 1994;29(3):314-318.

[43] Paris PM. Pain management in the child. Emerg Med Clin 1987;5:699-706.

[44] Lewis LM, Lasater LC, Brooks CB. Are emergency physicians too stingy with analgesics? South Med J 1994;87(1):7-9.

[45] Quinn JV. Debunking the myths about analgesia. (letter) 1994;CMAJ 7;151-152.

[46] Thompson E. Lack of pain control a tragedy. (letter) CMAJ 1994;3;152.

[47] Melzack R. IASP President's address- The tragedy of needless pain: a call for social action. In Dubner R, Gerhart GF, Bond MR, (eds). *Proceedings of the 5th world congress on pain.* New York: Elsevier Science Publishing: 1988.

[48] Shepherd J. Analgesia and the acute abdomen (letter). BMJ 1979;ii:1289.

[49] Jones PF. Early analgesia for acute abdominal pain. BMJ 1992; 6835(305):554-555.

[50] Cusheri A. The acute abdomen and disorders of the peritoneal cavity. In: Cusheri A, Giles GR, Moossa AR, eds. *Essential surgical practice.* 2nd ed. London: Wright PGS; 1986:1238.

[51] Hickey MS, Kiernan GJ, Weaver KE. Evaluation of abdominal pain. Emerg Med Clin Nor Am 1989;79(30):437-452.

[52] *Webster's encyclopedic unabridged dictionary of the English language.* New York: Portland House, 1989;1036.

[53] Melzack R, Wall PD. Pain mechanisms: a new theory. Science 1965;150; 971.

[54] Melzack R. The puzzle of pain. In Melzack r (ed.). *The puzzle of pain.* New York: Basic Books, 1973: 15-20.

[55]Melzack R. Concepts of pain measurement. In Melzack R (ed.). *Pain measurements and assessment.* New York: Raven, 1983; 1-50.

[56] Reading AE. Testing pain mechanisms in persons in pain. In Wall PD, Melzack R (eds.). *Textbook of pain.* New York: Churchill Livingston, 1989: 269-280.

[57] Calliet R. Neuroanatomy of pain mechanisms. In Calliet R. *Pain mechanisms and management.* Philedelphia: F A Davis, 1993; 1-28.

[58] Chapman CR, Syrjala KL. Measurement of pain. In Bonica JJ ed. *The management of pain.* Lea & Febiger:, Philidelphia,1990;580-595.

[59] Melzack R, Torgerson WS. On the language of pain. Anesthesiology 1971;34-50.

[60] Corson JA, Schneider MJ. The Dartmouth pain questionnaire: An adjunct to the McGill Pain Questionnaire. Pain 1984;19:59.

[61] Reading AE. An analysis of the language of pain in chronic and acute patient groups. Pain 1982;13:185-192.

[62] Reading RE. The McGill pain questionnaire: an appraisal. In Melzack R (ed.). *Pain measurements and assessment.* Philedelphia: F A Davis, 1993; 55-61.

[63] Budzynski TH, Stoyva JM, Adler CS, Mullaney DJ. EMG biofeedback and tension headaches: a controlled outcome study. Psychosom Med 1973;25:484-496.

[64] Kremer E, Atkinson JH, Ignelzi RJ. Measurement of pain: patient preference does not confound pain measurement. Pain 1981;10:241-248.

[65] AhlesTA, Ruckdeschel JC, Blanchard EB. Cancer-related pain. II Assessment with visual analog scales. J of Psychosom Res 1984;28:121-124.

[66] Reading A, A comparison of pain rating scales. J Psychosom Res 1980;119-124.

[67] Scott J, Huskisson EC. Graphic Representation of pain. Pain 1975;2:175-184.

[68] Huskisson EC. Visual Analogue Scales. In Melzack R (ed.). *Pain measurements and assessment* New York: Raven, 1983; 33-40.

[69] Huskisson EC, Shenfield GM, TaylorRT, Hart FD. A new look at ibuprofen. Rheum Phys Med Supp 1970; 88-92.

[70] Aitken RCB, Zealley AK. Measurement of moods. Brit J Hosp Med 1970;4:215-225.

[71] McCormack HM, De La Horne DJ, Sheather S. Clinical application of visual analogue scales: a critical review. Psych Med 1988;18: 1007-1019:

[72] Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. Res Nurs Health 1990;13:227-236.

[73] Sriwantanakul K, Kelvie W, Lasagna L , Calimlim JF, Weiss OF, Mehta G. Studies with different types of visual analog scales for measuring pain. Clin Pharmacol Ther 1983;34(2):234-239.

[74] Bjerregaard B, Brynitz S, Holst-Christensen J, Jess P, Kalaja E, Lund-Kristensen J, Thomsen C. The reliability of medical history and physical examination in patients with acute abdominal pain. Meth Inform Med 1983;22:15-18.

[75] Greene TJ, Meador GR. Interrater agreement of palpation of signs in focal abdominal pain. (abstract) Acad Emerg Med 1995;2(5):378.

[76] Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Worthington JR. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. Ann Emerg Med 1992;21:384.

[77] Stiell IG, Greenberg GH, McKnight RD, et al. Decision rules for the use of radiography in acute ankle injuries: refinement and prospective validation. JAMA 1993;269:1127-1132.

[78] Stiell IG, McKnight RD, Greenberg GH, et al.. Implementing of the Ottawa ankle rules. JAMA 1994;271(11):827-32.

[79] Personal communication, Dr J. R. Worthington, Director, Division of Emergency Medicine, Ottawa Civic Hospital, 1995.

[80] Talan DA. Recent developments in our understanding of sepsis: Evaluation of anti-endotoxin antibodies and biologic response modifiers. Ann Emerg Med 1993;22(12):1871-1890.

[81] de Dombal FT. The diagnosis of acute abdominal pain with computer assistance: worldwide perspective. Ann de Chirg 1991;45(4):273-277.

[82] McAdams WA, Brock BM, Armitage T, Davenport P, Chan M, de Dombal. Twelve years experience of computer-aided diagnosis in a district general hospital. Ann Roy Coll Surg Eng 1990;72(2):140-6.

[83] de Dombal FT. Acute abdominal pain: An O.G.M.E. survey. Scand. J. Gastroent. 1978;14(sup):30-43.

[84] Donner A. Approaches to sample size estimation in the design of clinical trials-a review. Stats in Med 1984;3:199-214.

[85] Makuch R, Johnson M. Issues in planning and interpreting active equivalence studies. J Clin Epidemiol 1989;42 (6):503-511.

[86] Vollset SE. Confidence intervals for a binomial proportion. Stats in Med 1993;12:809-824.

[87] Kramer MS, Feinstein AR. Clinical biostatistics LIV. The biostatistics of concordance. Clin Pharm Ther 1981;29(1):111-123.

[88] Fleiss J. Relaibility of measurement. In Fleiss J. *Deisgn and analysis of clinical experiments.* New York: Wiley, 1985: 1-32.

[89] Todd, K. Standardized pain control research. JAMA 1994;271:925-928.

[90] Burdock EI, Fleiss J, HardestyAS. A new view of interobserver agreement. Pers Psychol 1963;16:373-384.

[91] Forrest JB. *Management of acute pain.* London, Canada: EMIS-Canada 1993: 71.

[92] Himmelfarb S. The measurement of attitudes. In Eagley AH, Chaiken S. eds. *The psychology of attitudes.* San Diego: Harcourt Brace Janovich,1992; 23-87.