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.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Université d'Ottawa · University of Ottawa

Université d'Ottawa • University of Ottawa

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

PAPA, Linda

AUTEUR DE LA THÈSE - AUTHOR OF THESIS

M.Sc. (Epidemiology)

GRADE - DEGREE

Department of Epidemiology and Community Medicine

FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

A Feasability Study to Derive Two Clinical Decision Rules for Patients with Urinary Calculi: For the use of Urgent Intravenous Pyelography and For Predicting Complications of Delayed Urinary Calculus Passage

Ian G. Stiell and George Wells

DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

Yue Chen

Phil Wells

J.-M. De Koninck, Ph.D.

LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

SIGNATURE

DEAN OF THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

A Feasibility Study to Derive Two Clinical Decision Rules For Patients with Urinary Calculi:

For the Use of Urgent Intravenous Pyelography and For Predicting Complications of Delayed Urinary Calculus Passage

Linda Papa, MD
Department of Epidemiology & Community Medicine
University of Ottawa

Thesis submitted to the
School of Graduate Studies and Research
In partial fulfillment of the requirements
for the MSc in Epidemiology

Thesis Supervisors:
Ian G. Stiell, MD
George Wells, PhD

ABSTRACT

Background: About 5-15% of the North American and European populations pass a urinary stone in their lifetime.(1-3) Renal colic is a term used to describe the tremendous pain caused by the blockage of the urinary system by a calculus. In a small percentage of cases, urinary calculi can cause complete blockage of the urinary passage and lead to kidney dysfunction.(4-8) The most accurate way of detecting degree of urinary obstruction is with a radiograph called an intravenous pyelogram or IVP. Currently, emergency physicians show considerable variability when ordering pyelograms for patients with renal colic (9;10) and their clinical judgment is inaccurate in determining the degree of obstruction.(10;11) Therefore, some physicians are ordering too many unnecessary IVP's and others are missing severe urinary obstruction. A prospectively derived clinical guideline would help physicians eliminate or delay pyelography safely in many patients, improve and standardize patient care, save health care dollars and reduce time spent in the emergency department.(2;9;10;12;13) In addition, continued pain after discharge affects patients' ability to function and return to work. Currently, there are no clear guidelines to decide which patients need a procedure early to remove urinary calculi to prevent suffering, renal impairment and lost wages.(14)

Objectives: 1) To assess the feasibility of developing a clinical decision rule that would identify the clinical factors that are sensitive in predicting which patients with suspected ureteral calculi have severe obstruction and require urgent pyelography by:

i) defining severe obstruction; ii) determining the incidence of severe obstruction; iii) identifying potential predictor variables; and iv) calculating a sample size for the definitive decision rule study. 2) To prospectively assess clinical predictors of complications from urinary calculi after discharge from the ED by: i) determining the incidence of complications; ii) identifying preliminary predictor variables; iii) calculating a sample size for the definitive study.

Methods: This prospective cohort study included consecutive cases of suspected renal colic presenting to two tertiary care hospital emergency departments from January 1999 to September 1999. Patients with suspected renal colic having a 20-variable data form completed by an emergency physician and a pyelogram performed within 24 hours were included. All pyelograms were reviewed by a radiologist and a urologist in order to identify those cases with severe ureteral obstruction. Severe ureteral obstruction on pyelogram was defined by specialists in urology, radiology and emergency medicine in the Ottawa area who filled out surveys and agreed on its clinically significance via a panel discussion. Those with non-severe obstruction were discharged and followed via telephone at one and four weeks to assess development of complications. Complications were defined by: i) persistent pain by four weeks; or ii) development of fever (>38°C) as a result of an infection from their urinary tract; or iii) elevation of serum creatinine > 150 mmol/L by four weeks. Categorical data were analyzed using Fisher's Exact test, continuous variables assessed by independent sample 2-tailed t-test, and ordinal variables tested with Mann-Whitney U test.

Those variables found to be associated with the outcome measure of severe obstruction ($p<0.15$) or complications ($p<0.20$) were dichotomized and combined using logistic regression analysis. Preliminary models were compared to clinical judgment via receiver operating curves. Sample size was calculated from the results of the logistic regression analysis and on the precision of the sensitivity. Data collection methods were scrutinized.

Results: Results from the surveys and panel discussion, yielded a final definition of severe ureteral obstruction on pyelogram as urine extravasation and/or no visualization of contrast beyond the obstruction after two hours. The incidence of severe obstruction based on this definition was 15%. There were 437 suspected renal colic patients seen over 8 months in the two emergency departments and 119 had both a data form completed and a pyelogram performed. Of those patients meeting inclusion criteria, the mean age was 44 years (SD12), 69% were male, 47% had a previous history of renal colic, and 15% were admitted to the hospital. Logistic regression analysis was conducted for severe obstruction and used 72 of a possible 119 due to missing data. Four clinical variables were significantly correlated with having severe obstruction on IVP in this model: i) residual pain at discharge $\geq$ 1cm on VAS (OR=7.8, 95%CI 1.6-82.1); ii) rebound tenderness on abdominal palpation (OR=18.5, 95%CI 1.1-303.0); iii) vomiting (OR=7.8, 95%CI 1.0-62.0); and iv) persistent pain after 6 hours in the ED (OR=18.6, 95%CI 2.2-159.0). Hosmer-Lemeshow goodness-of-fit statistic was 7.4 (p=0.283). ROC curves comparing physician judgment to our preliminary clinical model yielded an area that was significantly larger (p=0.007) for our model.

Sample size calculations indicated that 50 cases of severe obstruction will be required for a definitive decision rule study for urgent imaging to achieve 91-100% confidence interval around a sensitivity of 100%.

Of those with non-severe urolithiasis on IVP, 52 were followed by telephone at one and four weeks, 29 (56%) patients developed complications. One patient (2%) developed fever from an infection of their urinary system, 4 patients (8%) developed renal failure, 48 (93%) had persistent pain at four weeks, and 3% had a combination of factors. Eighteen patients (35%) underwent a urologic procedure for their urinary calculi. Logistic Regression analysis used 38 of 52 potential patients due to missing data and yielded an unstable model with only one clinical variable: pain at discharge $\geq$ 2 cm on a VAS scale. Physicians were unable to predict complications after discharge. Due to the instability of the model we could not calculate a sample size.

Conclusions: This feasibility study is the first study to have prospectively assessed clinical variables that predict severe obstruction and complications in those patients with suspected renal colic. We defined and determined the incidence of severe obstruction and complications and built preliminary clinical models that were considerably better than current physician judgment for predicting severe obstruction and complications. We also estimated the sample size for the definitive study for urgent imaging. Although we lacked sufficient power to build a stable model for predicting complications, the data suggests that by using a larger sample size and adding calculus dimension and location to the list of potential variables a clinical guideline could be developed. Future studies will prospectively validate these models.

ACKNOWLEDGEMENTS

I would like to acknowledge the patience and assistance of my husband Giuseppe Ricci who shared the challenge of this thesis with me. His loving support and selflessness allowed me to fulfill a very important goal in my life, the pursuit of a research career.

To my parents Louisa and Joseph Papa, very special thanks for their constant guidance and encouragement which has always made all aspirations seem achievable and all trials seem trivial. Their example of perseverance, strength, and passion for knowledge will stay with me always.

To Lucia and Angelo Ricci, I offer thanks for their wonderful support in preparation of this thesis.

I am indebted to Dr. Ian Stiell, my mentor, for giving me the opportunity to pursue and prepare for a career in clinical research. His guidance was diligent and meticulous, ensuring the most comprehensive research mentoring possible. The dedication of his time and wisdom is valued and appreciated.

Many thanks to George Wells for guiding me through the maze of statistics and helping me to learn statistical principles with ease. Grazie!

Thanks to John Mahoney for his urologic expertise and the time he spent sifting through hundreds of pyelograms.

Thanks to Fay Draper for her diligence in always keeping me aware of the details that made this thesis possible.

Thanks to Irene Harris for helping me put the components of the thesis together into a neat package.

I would like to acknowledge the financial assistance of the Ontario Ministry of Health for funding me with a young investigator grant.

TABLE OF CONTENTS

TABLES:

FIGURES:

1. INTRODUCTION

1.1 The Use of Urgent Intravenous Pyelography

Renal colic (pain from urinary stones) is a very common condition seen in the emergency department. Renal colic is a term used to describe the tremendous pain and pressure caused by the blockage of a stone. Unfortunately, in a small percentage of cases, kidney stones can cause complete blockage of the urinary passage and lead to kidney failure. Complete blockage needs to be relieved promptly. The most accurate way of detecting blockage from kidney stones is with a radiograph called an intravenous pyelogram or IVP.

Since there are no clear guidelines for which patients with kidney stones should undergo an IVP urgently, physicians vary considerably when they order IVP and may order many more of these procedures than may be necessary. At the same time, however, blockages may be missed in those that do not have an IVP urgently.

This research project studied patients with acute renal colic in order to prepare for the development of a decision rule (guideline) for the use of urgent IVP. These guidelines will be designed to allow physicians to order urgent IVP more appropriately and therefore not miss serious blockage. Besides improving and standardizing the quality of patient care, widespread use of the guideline could lead to less time spent in the emergency department and to large savings for our health care system.

1.2 Predicting Complications Of Kidney Stone Passage

Fifty percent of kidney stones pass on their own in 2-3 days and about 80% pass within one month. For the rest, it may take several months before the stone passes and patients may experience prolonged pain and suffering. At present, we have no way of knowing who will have problems passing their stones and there are no clear guidelines to decide which patients need a procedure early to remove a kidney stone.

Urologists use several procedures to treat patients with delays or complications with stone passage. The least invasive method is extracorporeal shockwave lithotripsy that uses sound frequencies to break up the stones in the urinary system. Another method is ureteroscopy; performed using a scope (tube-like camera) passed into the bladder and into the ureters to extract the calculus. Percutaneous nephrolithotomy involves accessing the kidney directly through the skin in the patient's flank to drain the kidney. And, the most invasive surgical procedure involves opening the kidney surgically. The last method is very rarely performed at present.

This research project will follow patients with kidney stones after they leave the emergency department to see who develops problems. Then a second guideline (decision rule) will be developed using statistical modeling to help physicians detect patients who will have delayed stone passage.

These guidelines will help prevent prolonged patient suffering associated with difficulties passing their stones.

1.3 GOALS

The first goal of this thesis is to assess the feasibility of developing a clinical decision rule that will identify the clinical factors that are very sensitive in predicting which patients with suspected ureteral calculi have severe obstruction and require urgent pyelography (IVP). Severe obstruction impairs kidney function and requires urgent referral to a urologist. The second goal of this study is to determine the feasibility of developing a second decision rule for predicting complications such as prolonged pain, renal failure and infection from their urinary calculi developing after they leave the emergency department. The feasibility study will address several important issues for the definitive study: **a)** Sample size will be estimated for both decision rules by obtaining the prevalence of the primary and secondary outcomes through prospective emergency department data collection. **b)** The strength of association between the clinical predictor variables and primary and secondary outcomes will be tested on a sample of patients. In addition, modeling of the variables in a preliminary decision rule will be performed by multivariate analytic methods from prospective data collection. **c)** A clear set of criteria for identifying "significant radiological signs of severe obstruction" on the pyelogram will be formulated through a consensus of opinions from emergency physicians, urologists and radiologists because their are no absolute IVP criteria for grading obstruction in the literature at this time. **d)** The time frame for the definitive study will be planned and the number of centers involved for adequate accrual will be determined by prospective emergency department data collection. **e)** Data collection feasibility and follow-up will be examined.

If the above goals are met, the definitive decision rule study (Appendix 1) will be conducted. The definitive study will develop a decision rule for emergency physicians to decide who needs urgent pyelography so as to standardize and improve patient care, prevent patient exposure to unnecessary imaging, decrease time spent in the emergency department and lead to a more appropriate utilization of health care resources. The second decision rule will help primary care physicians decide upon disposition and help urologists intervene earlier to prevent complications and prolonged patient suffering from delayed calculus passage.

2. BACKGROUND

2.1 Literature on Imaging Practices in Renal Colic
(Decision Rule #1)

2.1.1 Epidemiology

Renal colic is a very common condition in North America and around the world. Renal colic is a term used to describe the tremendous pain and pressure caused by the over distention of the pelvicalyceal system, renal capsule and the ureters secondary to calculi.(15) About 5-15% of the North American and European populations passes a calculus in their lifetime.(1-3) It affects 2-5% of people in Asia and 20% of the population in Saudi Arabia.(16) Urolithiasis presents at an average age of 20 to 40 years with a 50% recurrence over 5 years.(3;9) The incidence and prevalence rates of urinary calculi may be affected by genetic, nutritional, and environmental factors.(16) It is one of the most common conditions seen in the emergency department.(1;3) There are no national data regarding the economic impact of urolithiasis in Canada. However, in the province of Ontario the cost for outpatient IVP's is in excess of 3.2 million dollars per year for technical and professional fees based on figures from the Ministry of Health of Ontario. In the United States the annual cost of outpatient evaluation of urolithiasis is about $278 million dollars.(17) The total estimated charge for evaluation, hospitalization and treatment of urinary calculi is $1.23 billion dollars per year.(17) These estimates were obtained from a summary of figures from the Agency for Health Care Policy and Research in the United States.

2.1.2 Current Emergency Department Imaging Practices

The initial investigation of suspected ureteral colic is currently not standardized (12;18) and there is considerable variability among emergency physicians in ordering pyelography. (1;9;10) Current indications for pyelography in the emergency department include uncertain diagnosis, identifying degree of obstruction from urinary calculi, and determining calculus size and location for prognostic purposes.(19;20) Empirical algorithms, based on "expert opinion", have been introduced by different groups of radiologists for deciding the need for imaging and each one differs from the other. Rhea recommended in his 1982 study on diagnostic testing in renal colic that imaging be considered only for those patients who did not experience classic relief of pain and passage of a calculus.(9) Subsequently, Smith et al. in 1999 suggested that imaging in renal colic is not necessary if the patient has a prior history of renal colic and presents with typical symptoms of renal colic. He also suggested the use of CT scan for all patients with acute flank pain without previous renal colic.(21) None of these algorithms have not been prospectively validated.

Emergency medicine research in renal colic has been dominated so much by the concern of missing an abdominal aortic aneurysm that the consequences of urinary obstruction have been inadequately studied. The popularity of the helical CT scan (computed tomography scan which is more rapid than the standard CT scan) in the diagnosis of renal colic can be explained by the fact that it can diagnose both abdominal aortic aneurysms and urinary calculi with a very high accuracy. Recent studies have suggested that the use of the CT scan is superior to IVP to diagnose ureteral calculi so many emergency departments

in North America are implementing its use for diagnostic purposes. However, helical CT is unable to assess kidney function as well as IVP.(22)

2.1.3 Radiological Methods of Investigation for Renal Colic

Radiological modalities available for investigating ureteral calculi include plain abdominal radiographs (also known as KUB), ultrasound, intravenous pyelography, and both enhanced and unenhanced CT scan.

Although 90% of renal calculi are radiopaque, many cannot be visualized on a plain radiograph due to bowel gas, overlying viscera, and other calcifications with similar appearance, which make interpretation difficult.(10) One investigator quoted only 39% of plain abdominal radiographs being positive(23) and another more recent study showed a sensitivity of 59%.(21) This modality cannot evaluate the degree of obstruction.

Some investigators have combined ultrasonography with plain radiographs to diagnose ureteral obstruction and have met with variable success.(13;18) Ultrasound studies show that calculus and hydroureter can be identified with a high degree of certainty with ultrasound, but it fails when filling is delayed and hydroureter is absent. As well, ultrasound does a very poor job at visualizing the ureter except at the most proximal and distal ends.(24;25) Degree of kidney function cannot be determined by ultrasound. With the advent of helical CT emergency physicians have started implementing it as an alternative method of diagnosing calculi in the emergency department for those with suspected renal colic. It eliminates the need for intravenous contrast, can be completed in less than 5 minutes and can demonstrate other causes of flank pain that are not urinary.(18) The diagnostic accuracy of CT scanning ureteral

calculi appears to be higher than intravenous pyelography.(22-24;26-29) Indeed, this is the case for diagnosis of calculi but not for determining the function of the urinary system when obstructed by calculi. The use of helical CT to evaluate the severity of obstruction from calculi has not been established. Authors who advocate CT still suggest that if the findings of helical CT are indeterminate or if it is necessary to find the degree of obstruction, then intravenous pyelography (IVP) is recommended.(18;30-32)

For many years intravenous pyelography was the radiological study of choice and was the gold standard for the diagnostic work-up of patients with renal colic.(8;25;33) Still today it remains the most common radiographic test for renal colic despite the introduction of helical CT.(15) It still gives the most detailed demonstration of the urinary system and of high-grade obstruction with inadequate filling of the ureter.(15) It delineates the location and size of the calculus (except small radiolucent calculi which may or may not show up as filling defects), the degree of obstruction, the functional status of the kidney, and the caliber and course of the ureter.(12)(10;32;33)

2.1.4 Defining Severe Obstruction

The first step in developing a clinical decision rule is to clearly define the outcome being predicted. (34-36) There are no absolute IVP criteria for grading obstruction in the literature or in radiological or urologic textbooks. (13;19) There have been some attempts to define severe obstruction in radiological literature. In the late 1980's it was felt that the length of the nephrogram correlated well with the degree of obstruction and patients' symptoms. Complete or near-complete obstruction was defined as a non-functioning kidney after

several hours. Chen et al. in 1997 described the radiological findings of acute urinary tract obstruction as a *nephrogram* (opacification of kidney with dye) that was delayed more than 15 minutes or the presence of *extravasation* (leakage of dye outside urinary system). Both were signs of severe urinary obstruction.(8) Stewart et al. classified severe obstruction by contrast material extravasation and time to complete ureteral filling longer than two hours.(13)

Since the definition of severe obstruction varies, so does the incidence. In one study out of New York, complete obstruction was detected in 14 of 40 intravenous pyelograms (35%).(12) In contrast, a study from Ottawa found only 3 out of 124 patients (2.4%) with complicated IVP's.(37) In another US study the incidence of extravasation was found to be about 11% and the incidence of dense nephrograms about 9%.(8) The incidence of high-grade obstruction in a study by Press was 6%.(1)

In the emergency literature severe obstruction is considered a criteria for admission.(38;39) Severe obstruction is a clinically significant outcome measure that is more objective and transportable than "need for admission" or "need for intervention" that have been used previously in the literature.

For the feasibility study we will have to both define clinically significant severe obstruction and determine its incidence.

2.1.5 Pathophysiology of Urinary Obstruction

The urologic literature clearly states that severe kidney obstruction demands immediate assessment when the kidney is non functional for several hours. (3;40). Impairment of urinary flow from prolonged ureteral obstruction can

cause renal function to deteriorate and can lead to renal damage.(15) The link between loss of renal function and obstruction was demonstrated in the 1970's through experiments on the hydrodynamics of the urinary system in dogs. They demonstrated that prolonged ureteral occlusion for days to weeks brought about a reduction in renal blood flow, decreased glomerular filtration and decreased renal output. In some cases these changes were irreversible.(4;5) In a study by Holm-Nielson et al., 11 of 31 patients with obstruction for more than four weeks developed irreversible renal damage.(6)

In cases of complete obstruction renal deterioration began within 18-24 hours of complete obstruction. Similarly, in patients with complete urinary obstruction irreversible renal damage can occur if obstruction is not relieved within four to seven days. (7;8) Although terminal renal failure occurs in few cases, milder renal impairment can lead to elevated blood pressure.(6)

Calculus size can often predict spontaneous passage but it cannot determine the degree of obstruction. Small and medium size calculi can cause severe obstruction and functional renal impairment as much as big calculi.(6) Evaluating kidney function with an IVP is justified if one suspects impaired function. There are no clinical criteria at this time to determine this. Hence the need for a clinical decision rule to order IVP urgently to identify these patients.

2.1.6 Urgent Imaging in Suspected Renal Colic

A review of the urologic and emergency medicine literature reveals that there is no set of validated criteria based on history, physical exam and laboratory tests to identify which patients need urgent pyelography.(9;12)

Diagnosing renal colic can be done clinically in 80% to 90% of cases.(15) Rhea et al. in a retrospective study found that the diagnostic correctness of flank pain, sudden onset of pain, hematuria, with and without KUB to be high enough to eliminate or delay the need for urgent IVP.(9) However, the authors recognized that these criteria could eliminate IVP as a diagnostic tool but not to identify the degree of obstruction. Similarly, Elton in 1993 retrospectively studied various predictors for diagnosing renal colic in the emergency department and found through univariate analysis that hematuria, flank pain, acute onset, and severe pain to be the highest ranking clinical predictors of renal colic with a 90% probability of a calculus.(2) Patients with the following four characteristics hematuria, flank pain, acute onset of pain and a positive KUB had a 98.5% probability of a calculus. It was concluded that emergency IVP may be deferred for diagnostic purposes and used more selectively.(2) Since emergency physicians can diagnose most cases of renal colic clinically with enough certainty, they do not require urgent imaging for diagnosing calculi. However, they would require urgent IVP in those select cases where there is a possibility of severe obstruction. Urgent pyelogram would then be considered necessary if severe obstruction were suspected.

2.1.7 Clinical Judgment in Determining Severity of Obstruction

There is considerable variability among emergency physicians in ordering pyelography based on clinical judgment. (1;9;10) However, clinical judgment has been shown to be inaccurate in estimating the degree of obstruction.(10;13) In a study done to determine whether emergency IVP adds significant information to clinical judgment in possible renal colic, it was shown that the degree of

obstruction was misjudged in at least 11% to 18% of cases.(10) In a retrospective study by Richards et al. published in 1999, it was shown that although many of the IVP's could have been eliminated, 8 out of 123 patients with high grade obstruction would have been missed based on current clinical practices.(11) A reliable and sensitive decision rule would permit physicians to be more accurate and discriminating in ordering pyelography.

2.1.8 Selective Use of Radiography

In a randomized two-arm trial in the emergency department by Tasso et al., one group of patients underwent an IVP and the second group had IVP done selectively based on the emergency physician's judgment (using various empirical clinical parameters not specified in the study). They showed that IVP could be delayed or eliminated without adversely affecting outcome.(12)

In North America, there is tremendous variability among emergency physicians in ordering IVP's.(10). If clinical judgment in suspected renal colic were sufficiently predictive using a carefully derived set of guidelines then pyelography could be delayed or eliminated in most patients, allowing for less expensive, less invasive testing and less time spent in the emergency department.(2;9;10;12;13)

Rhea suggested that a saving of approximately $5000 could be achieved if IVP's would have been used selectively and eliminated in 50 patients in his study.(9) Ontario Ministry of Health statistics show that 46,370 outpatient IVP's were performed in the province of Ontario in 1999. At the Ottawa Hospital, the cost of an IVP is approximately $146 dollars per patient, including the technical fee, the professional fee and the materials. This yields an annual cost of 6.77

million dollars. The estimated annual cost for IVP's ordered by the emergency departments of the Civic and General sites of the Ottawa Hospital alone is about $76,000. The potential savings associated with reducing this figure could be very significant on the health care budget.

The average time required to perform an IVP is one hour. However, times may range from 30 minutes to several hours depending on the findings of the IVP. Imaging patients in the emergency department, regardless of the type of radiological study, will impact length of stay. Total length of stay in patients with suspected renal colic having either an IVP or a helical CT performed in the emergency is 365 minutes and 336 minutes respectively.

2.2 Background On Treatment of Urinary Lithiasis
(Decision Rule #2)

2.2.1 Current Indications for Urologic Intervention

The urologist's decision to intervene also suffers from tremendous variability. Unfortunately, the current practice guidelines on methods of treating ureteral calculi set forth by the American Urologic Association do not provide definitive criteria for deciding who needs an intervention and at what stage of the disease.(14) Surgical intervention has changed with the introduction of extracorporeal shock wave lithotripsy which breaks-up the calculus through sound waves sent through skin and the development of minimally invasive endoscopic techniques such as ureteroscopy (extraction of the calculus in the ureter via a urinary scope) and percutaneous nephrolithotomy (drainage of the kidney through the flank). Most of these techniques are associated with very low morbidity.(6) Complications prompting urologic intervention include prolonged pain, calculi 6mm thick or larger, presence of complete obstruction, failure of the calculus to progress (impacted calculus), associated urosepsis, and impairment of renal function.(25) As well, various independent patient factors such as functional renal anatomy, associated ureteral pathology, patient occupation and the patients preferences are considered by the urologist in deciding who requires an intervention.

There is a tremendous burden to society for time lost from work from urinary calculi. Indirect costs for lost wages from urolithiasis are estimated to be $139 million dollars annually in the US.(17)

These figures were obtained from a study using figures from the Agency for Health Care Policy and Research in the United States.

Before the introduction of less invasive therapies, the watch and wait approach was the favored method of treatment for urinary calculi; even in patients with debilitating pain who required hospitalization and intravenous narcotics. Although many urologists still favor this approach, it is falling out of favor because of patient disability and the rising costs of hospitalization and health care. The current trend is for earlier intervention.(14;41) Unfortunately, no prospectively derived guidelines have been developed to address this issue.

2.2.2 Predictors of Complications

There has not been a prospective evaluation of predictors of outcome in acute renal colic. Retrospective studies have, however, shown degree of pain, persistence of pain, size of calculus, location along the ureter, and signs of infection are important factors in deciding to intervene.(40;42-45)

Recent studies have shown that up to 40% of patients have difficulty passing their calculi.(26;46)

In a 1998 study from Ottawa, it was observed that continued pain at 5-7 days was indicative of need for an intervention.(37) In addition, patients who had a procedure were found to have higher overall pain scores (based on analgesic requirements and pain quantity) than those who did not. The issue of pain as a predictor was also examined in another follow-up study. Ibrahim studied prognostic factors for outcome of urolithiasis and concluded that "duration of pain" was considered an important indicator of failure of conservative treatment if patients had continued pain 30 days after initial presentation.(44)

It has been demonstrated that calculus size is a predictor of spontaneous passage with calculi < 5mm passing 84-98% of the time and >5mm 25%of time.(14;47) Another study found that the rate of spontaneous passage was 84.4%, 69% and 0% for calculi measuring <4 mm, 4-6 mm, and >6 mm in lesser diameter respectively.(37) A retrospective review of 2700 patients showed that 65% of calculi passed spontaneously in four weeks.(40) When an intervention was performed within four weeks of symptom onset patients had fewer complications (7%) than those who had an intervention after four weeks of symptoms (20%).

2.3 Methodological Standards for Decision Rules

A clinical prediction rule is a decision-making tool for clinicians that include 3 or more variables from history, physical exam, or simple diagnostic tests. The standards for clinical decision rules can be summarized as follows(34-36;48):

a) The outcome or diagnosis to be predicted must be clearly defined and clinically important.

b) The assessment of the outcome should be made in a blinded fashion without knowledge of the status of the predictor variables.

c) The clinical predictors must be clearly defined, sensible, reproducible and collected prospectively without knowledge of the outcome.

d) In order for the results to be generalizable the patients in the study should be selected without bias and should represent a wide spectrum of clinical and demographic characteristics.

e) The mathematical techniques for deriving the rules must be adequately described and justified. Usually, the multivariate statistical methods used include logistic regression, recursive partitioning and discriminant function analysis.

f) As a general rule there should be an adequate sample size providing at least ten outcome events per independent variable in the multivariate statistical model.

g) Clinical decision rules should demonstrate clinical sensibility, i.e. be clinically sensible and relevant. A rule is more likely to be applied if it is sensible, easy for clinicians to use, and if it suggests a course of action.

h) The sensitivity and specificity of the decision rule should be specified. Generally, rules with high sensitivity will tend to rule out a disorder and rules with a high specificity will tend to rule one in.

i) It is essential to prospectively validate the rule on a different set of patients and preferably with different clinicians in order to test its accuracy, reliability and acceptability.

j) Properly developed and validated prediction rules can influence clinical practice. The ultimate test of a decision rule is when it is implemented to demonstrate the true effect on patient care. Transportability can be tested at this stage.

2.4 Rationale

This feasibility study will form the basis for deriving guidelines for the use of urgent intravenous pyelography in patients with suspected renal colic. The study will determine the proportion of severe obstruction, define clinically significant severe obstruction through surveys and expert opinion and identify the most predictive clinical factors for the decision rule study. A reliable and sensitive decision rule would permit physicians to be more accurate in their assessment of renal colic and more discriminating in ordering radiological tests. Presently, there are no guidelines to assist emergency physicians in deciding to order an intravenous pyelogram.(12) Clinical judgment has been shown to be inaccurate. In a study done to determine whether emergency IVP adds significant information to clinical judgment in possible renal colic, it was shown that the degree of obstruction was misjudged in at least 11% of cases.(10) In another study, 8% of high-grade obstructions could have been potentially missed had IVP not been performed.(11) Therefore, a clinical guideline to help physicians identify severe cases of urolithiasis is needed.

In the United States the cost of outpatient evaluation of urolithiasis is about $278 million dollars.(17) In another US study, it was suggested that a saving of approximately $5000 could be achieved if IVP would have been used selectively and eliminated in 50 patients in this study.(9)

Although there are no Canadian statistics available, records from the Ministry of Health in the province of Ontario show that in the 1999 fiscal year there were 44,465 IVP's performed on hospital outpatients, including emergency departments, and 1,916 IVP's performed in outpatient clinics. The Ontario Ministry reported that 3.3 million dollars were spent on technical and professional

fees alone, excluding the cost of materials and operation. We estimated, based on cost of an IVP provided by our radiology department (146$ per IVP), that the annual cost for all outpatient IVP's in the province of Ontario was in excess of 6.77 million dollars for all outpatient IVP's. The economic impact of reducing this figure nationwide could be very significant for the health care budget.

It has been suggested that less time could be spent in the emergency department by eliminating radiography. Although a few studies have suggested that imaging in renal colic patients could be delayed or eliminated, none have addressed the issue of how.(11;12) If clinical judgment in suspected renal colic were sufficiently predictive using a carefully derived set of guidelines then pyelography could be delayed or eliminated in most patients, allowing for less expensive, less invasive testing and less time spent in the emergency department.(10)

Since the diagnosis of renal colic can be made clinically in 80%-90% of cases(2;9;11;15;16) very few patients need urgent imaging for diagnostic purposes. However, there is a subset of patients who are suffering from severe urinary obstruction that should be identified early to avoid long-term renal dysfunction.(4-6;8;49) We are suggesting that imaging, particularly IVP, be performed selectively in the emergency department in those cases of suspected renal colic having potential for severe obstruction.

By identifying in the emergency department those patients who will develop complications after discharge, patients can be appropriately referred to a urologist for intervention early and avoid the unnecessary pain and suffering associated with delayed passage of their calculus and avoid functional renal impairment from obstructive calculi. Complications from an intervention for

ureteral obstruction are much higher when symptoms exceed four weeks than when interventions are performed within four weeks of onset. (40)

Clinically, in the emergency department, it is difficult to predict who will develop complications or who will have difficulty passing their urinary calculi once they leave to go home. This is important in order to prevent irreversible renal damage, which can occur in those patients with urinary obstruction if obstruction is impairing function and is not relieved within 2-4 weeks.(6;8;49) Also, we can help prevent prolonged patient suffering associated with difficulties passing their calculi and reduce time lost from work. Indirect costs for lost wages from urolithiasis are estimated to be $139 million dollars annually in the US.(17)

3. SPECIFIC OBJECTIVES

3.1 Urgent Use of IVP (Decision Rule #1)

The Primary Objective of this study is to determine the feasibility of a definitive study that will develop a clinical decision rule that is very sensitive in detecting severe obstruction in patients with suspected renal colic. Patients who had a pyelogram within one week of their emergency visit will be included.

3.1.1 Estimating Proportion of Clinically Significant Severe Obstruction

Estimate the proportion of clinically significant severe obstruction on pyelogram (definition to be determined in objective 3.1.3) in patients with suspected urinary lithiasis who had an IVP. This will help estimate accrual rates and sample size for the definitive study.

3.1.2 Testing Clinical Predictors

Preliminarily determining the clinical predictors that will be most highly associated with the primary outcomes of severe obstruction and practice modeling of the rule.

3.1.3 Defining Clinically Significant Obstruction

Define "clinically significant severe ureteral obstruction" on IVP using surveys of urologists, radiologists and emergency physicians and an expert panel review.

3.1.4 Evaluating Data Collection Methods

Evaluating the feasibility of data collection methods that will make data collection in the emergency department more complete for the formal study.

3.1.5 Testing Predicted Probability of Severe Obstruction

Test the physicians' accuracy in assessing the pre-test probability of renal colic and severe obstruction.

3.2 Predicting Complications (Decision Rule #2)

The Secondary Objective for this feasibility study is to collect data on the complications (def 4.4.2.1) associated with the delayed passing of a ureteral calculus after discharge from the emergency department in order to identify clinical features that will predict who will have a complicated course. This will be accomplished by prospectively following patients with confirmed ureteral calculi on intravenous pyelography who are not severely obstructed and:

3.2.1 Estimating the Proportion of Complications

Estimating the proportion of patients with ureteral calculi who develop complications (fever, renal failure and prolonged pain - see definition in 4.4.2.1).

3.2.2 Testing Clinical Predictors

Testing the clinical predictors that will be most highly associated with the primary outcomes of complications and practice modeling of the rule.

3.2.3 Estimating Rate of Surgical Intervention

Estimating the rate of surgical intervention (ureteroscopy, percutaneous nephrostomy, extracorporeal shockwave lithotripsy, or open surgery - see definition in 4.4.2.2) in patients with proven renal colic.

3.2.4 Estimating Losses to Follow-up.

Estimating the number of patients lost to follow-up.

3.2.5 Testing Predicted Probability of Complications

Test the physicians' accuracy in assessing the pre-test probability of developing complications after discharge.

4. METHODS

4.1 Study Design

4.1.1 Study Design for Urgent Use of IVP *(Decision Rule #1)*

The first goal of the feasibility study was achieved using a prospective observational cohort.

4.1.2 Study Design for Predicting Complications *(Decision Rule #2)*

The second goal was accomplished via a prospective observational cohort.

4.2 Study Population

The study population included all patients of the Ottawa area presenting to the emergency departments of two university teaching hospitals (Civic and General sites of the Ottawa Hospital) with suspected renal colic. The study sampled patients presenting to the emergency department over a 7-month data collection period. The Ottawa Hospital is situated in Ontario Canada and has access to IVP and urologists 24 hours a day. These sites were selected after a chart review (Appendix 9) identified a high volume of patients with renal colic here, and confirmed that IVP's are ordered routinely to investigate renal colic at both sites.

4.2.1 Inclusion Criteria for Urgent Use of IVP *(Decision Rule #1)*

1) Adults presenting to the emergency department with suspected renal colic were eligible. Patient eligibility was determined by the attending staff emergency physician based on the patient having the following characteristics on arrival(2):

a) flank pain

b) acute onset of pain within 24 hours of emergency visit

2) Patients in whom an intravenous pyelogram was performed within one week of their emergency department visit were included.

4.2.2 Inclusion Criteria for Predicting Complications *(Decision Rule #2)*

1) Those patients who had a positive pyelogram showing evidence of a ureteral calculi without severe obstruction. Radiographic findings included: **i)** Prolonged dense nephrogram <15 minutes, **ii)** visualization of a calculus, **iii)** columnization, **iv)** delayed filling of the ureter <2 hours, **v)** hydronephrosis, **vi)** interureteric ridge edema, **vii)** sharp cutoff at some point in the of the ureter.
2) Available for telephone follow-up.

4.2.3 Exclusion Criteria for Urgent Use of IVP *(Decision Rule #1)*

1) Under 18 years old

Reason: Considered consenting age. Patients attending our institutions are usually older than 18 years old.

2) Patient already enrolled.

Reason: Excess influence of repeated measures in the same patient.

3) Allergy to contrast media.

Reason: Potential adverse reactions, morbidity and mortality associated.

4) Unstable vital signs.

Reason: Indicative of another medical condition. Urinary lithiasis should not cause instability in vital signs.

5) Pregnancy

Reason: Not safe to use IVP in pregnancy.

4.2.4 Exclusion Criteria for Predicting Complications *(Decision Rule #2)*

1) Same exclusions as 1,2, 5 above

2) Inability to contact patient for follow-up (no telephone, language barrier)

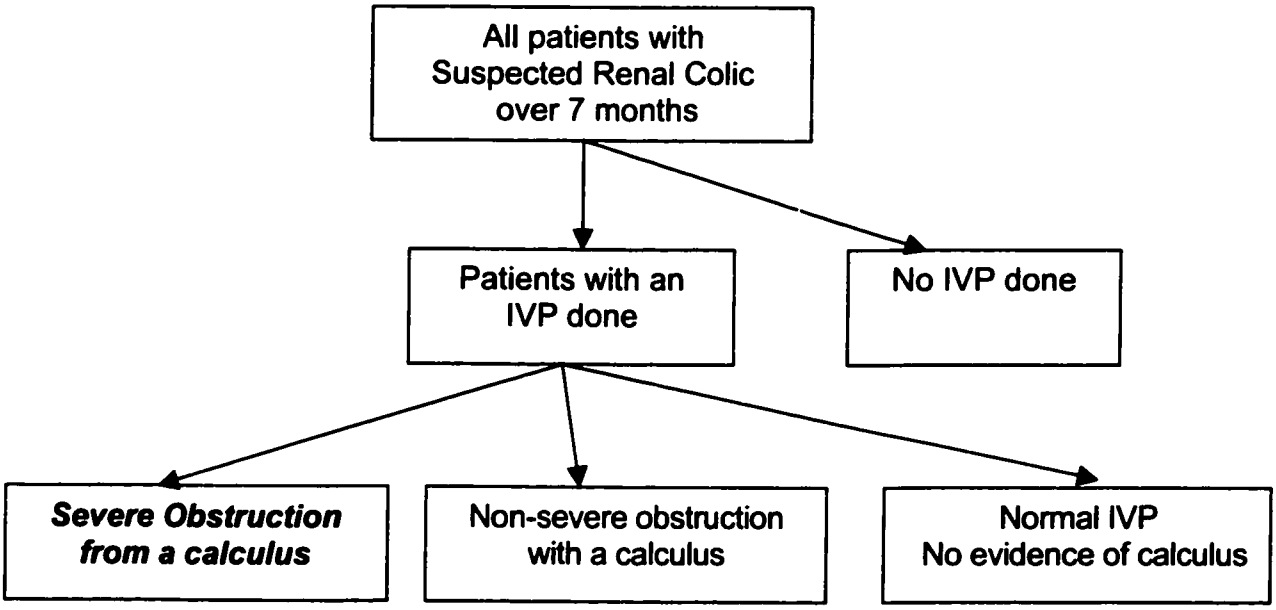
Reason: Patients require follow-up to determine outcome.

4.3 Standardized Patient Assessment

4.3.1 Sample for Urgent Use of IVP *(Decision Rule #1)*

a) The sample for objective 3.1.1 (proportion of severe obstruction) was obtained prospectively from the radiology reports of the intravenous pyelograms ordered by the emergency department over the 7-month data collection period. The cohort of patients used to determine the number of severe obstruction cases for sample size calculations is summarized in Figure 1. Qualified radiologists staffing the Ottawa Hospital read the pyelograms. Those IVP's showing the slightest evidence of severe obstruction according to the radiologist were subsequently reviewed by a qualified urologist using the new consensus definition of "clinically significant obstruction" defined during the course of this study (objective 3.1.3) (Appendix 2)

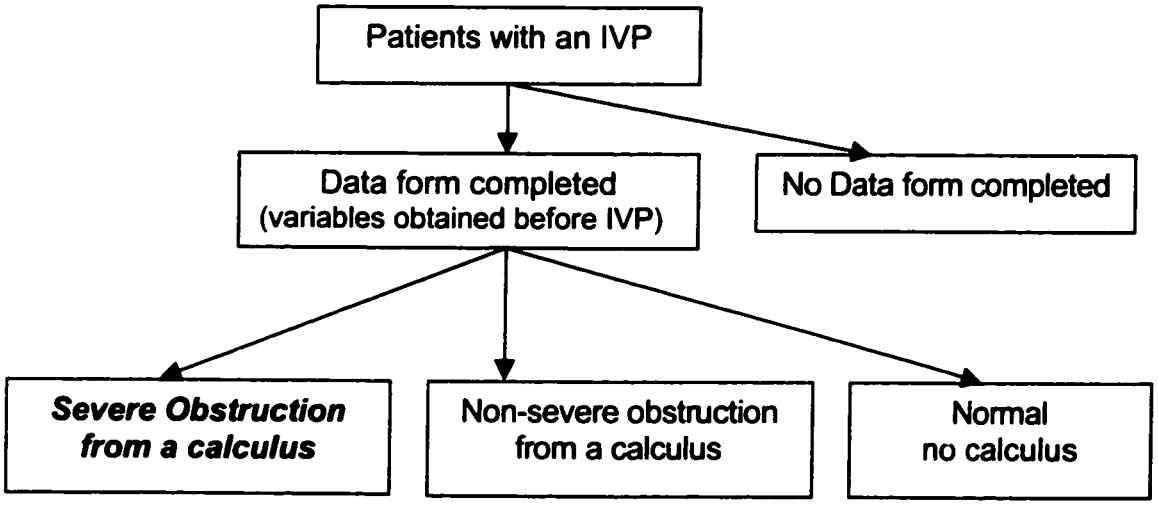
Figure 1
Proportion of Severe Obstruction



b) The sample for objective 3.1.2 (variables associated with severe obstruction) was obtained over the 7-month data collection period from the 20-variable data form completed by the emergency physician prior to the patient going for IVP. Figure 2 outlines the flow of patients for this part of the study. Based on other cohort studies in our department we expected an a priori data form completion rate of 50%.

A teaching session was provided for all emergency staff, including clerks, nurses and doctors, on how to enroll patients. Emergency department clerks and triage nurses flagged potential patients 24 hours a day as patients first arrived in the department. Emergency physicians then assessed the flagged patients and determined their eligibility. The treating emergency physician treated the patient as per his/her usual practice except for filling out the data form. The data form consisted of 20 variables including the patient's pain, medical history, physical exam and pre-test probabilities of disease. (Appendix 3)

Figure 2
Variables from the Data Form Associated with Severe Obstruction



c) For objective 3.1.3 (defining clinically significant severe obstruction), patient assessment was not required. Surveys were sent to all radiologists, urologists and emergency physicians practicing in the two university teaching hospitals of the Ottawa region. (Appendices 4 & 5 & 6) Surveys asked physicians to define clinically significant urinary obstruction based on their current clinical practice. Those physicians not responding to the initial questionnaire were sent a second copy. If the physicians did not respond to the second questionnaire they were contacted by telephone and reminded to complete the surveys. Results were then analyzed and prepared for the review by an expert panel composed of a radiologist, an emergency physician and 5 urologists. Having a panel composed primarily of urologists was necessary to define "clinically significant obstruction" since urologists are ultimately responsible for determining the urgency of this condition. The panel discussed the results of the surveys and agreed upon a final definition for clinically severe urinary obstruction. This definition was then used to analyze the data.

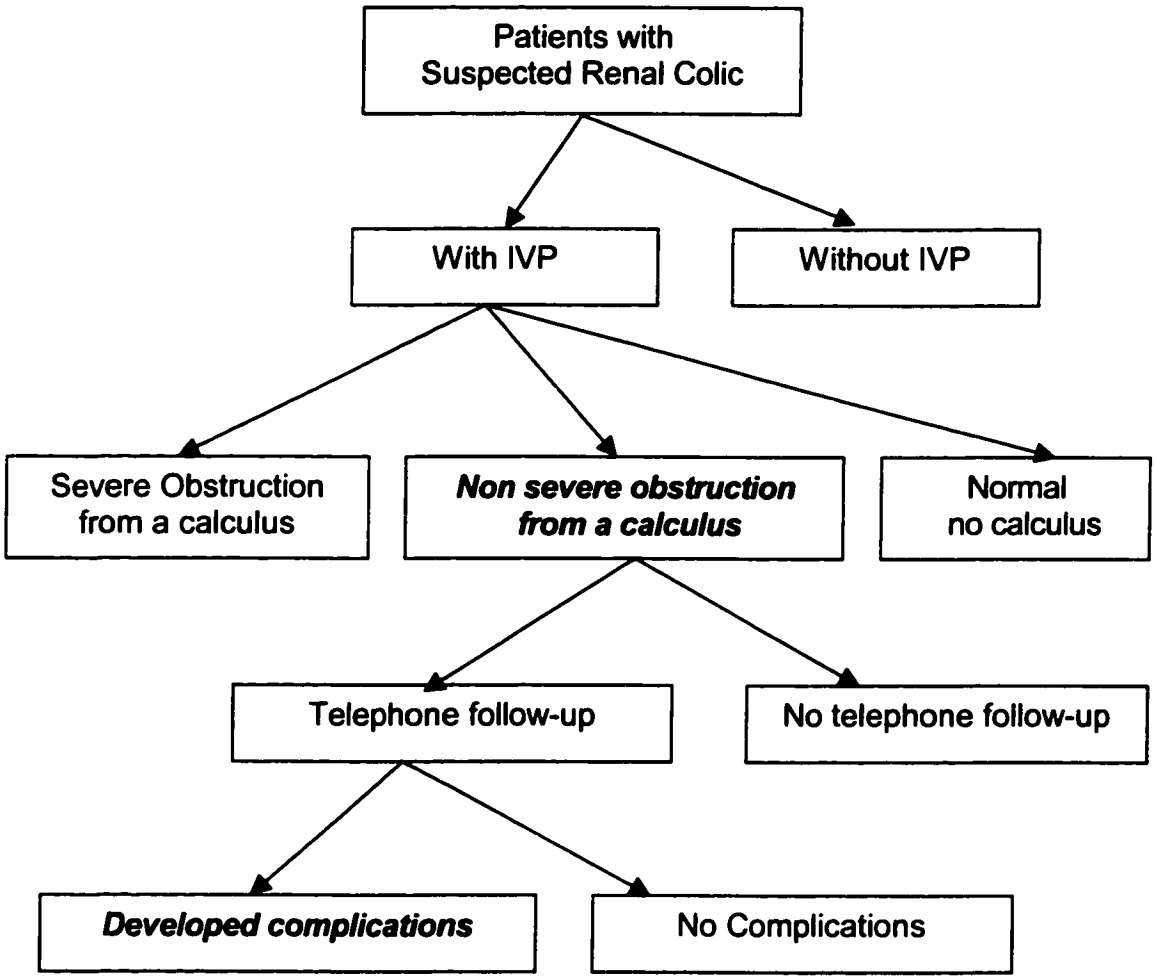
d) Assessment of the data collection methods (objective 3.1.4) did not require patient assessment. Cooperation from staff in patient enrollment, data form completion, and compliance with answering the questions on the data form were all assessed. This was meant to identify potential flaws in the data collection process and correct them for the formal study.

4.3.2 Sample for Predicting Complications *(Decision Rule #2)*

Once patients were discharged and had a pyelogram done, they were eligible for decision rule #2.

a) The sample for objective 3.2.1 (proportion of complications) was composed of patients with ureteral calculi confirmed by IVP without severe obstruction who were available for telephone follow-up. Those who developed complications after discharge were identified (fever from a urinary infection, renal failure or have continued pain after one month). Figure 3 is the cohort of patients used to determine the proportion of complications in patients with non-severe urinary obstruction.

Figure 3
Proportion of Complications

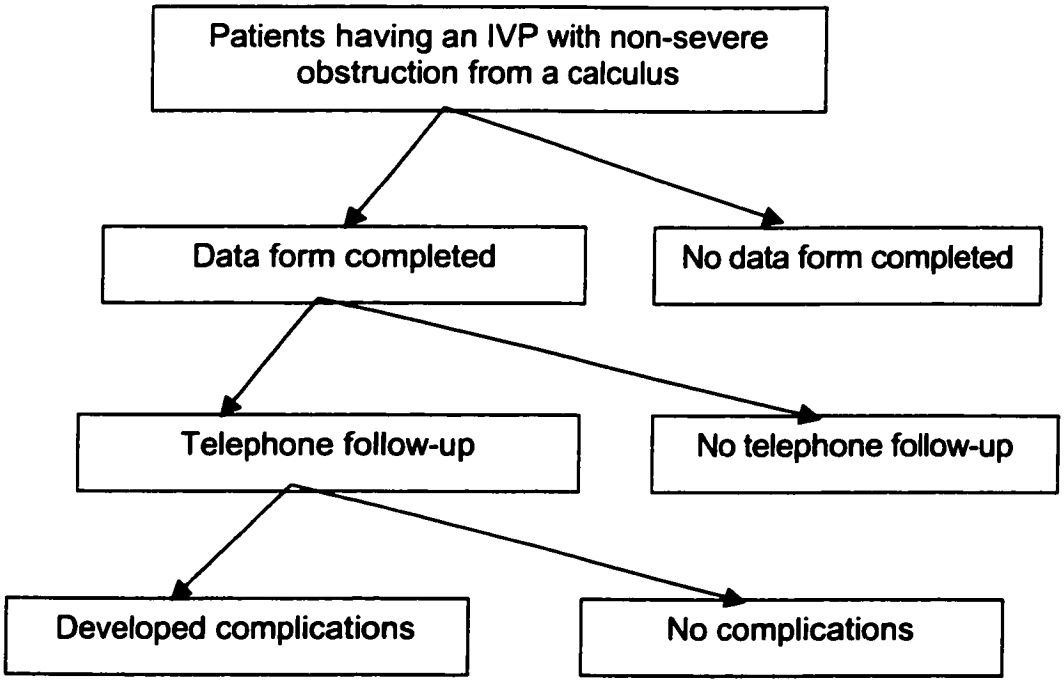


The research staff conducted a structured telephone interview at one week and at four weeks. (Appendix 7 & 8). One week and four week follow-up were chosen because over 50% of urinary calculi pass within 48-72 hrs (37) and over 80% of calculi pass within one month.(37)

During the telephone follow-up patients were asked about persistence of pain, urinary symptoms, vomiting, fever from a urine infection, revisits to a physician, and about any interventions performed on them. All procedures were described to patients in simple terms. (Appendix 8)

b) The sample for objective 3.2.2 (variables associated with complications) was obtained from the sample of patients with non-severe obstruction on IVP who had telephone follow-up and physician-filled data forms collected over the 7-month data collection period. The cohort for used for these calculations of association is shown in Figure 4.

Figure 4
Variables from the Data Forms Associated with Complications



c) The sample for objective 3.2.3 (rate of surgical intervention) was obtained from those patients who had an IVP confirming urinary calculi and who were followed for a one-month period. (Figure 4)

d) The sample for objective 3.2.4 (lost to follow-up) was obtained from those eligible for follow-up that could not be reached.

4.3.3 Ethical Consideration

In no way did the study interfere with the usual management of the patient. Consent for the telephone follow-ups were obtained from the patients by the research nurses. The ethics board of the Ottawa Hospital approved this study.

4.3.4 Selection of Potential Variables for both Decision Rules

Variables (clinical factors) were selected based on data from the literature and upon the clinical experience of emergency physicians and urologists from the two institutions. To date, the association of clinical factors with the degree of obstruction has not been previously assessed. Therefore, the clinical variables chosen for this study correlate with diagnosis, admission, and intervention of urinary calculi. Those diagnostic predictors of urinary calculi having a relative risk of greater than 1.0 in previous studies in the emergency department include: a) presence of a urinary calculus on plain radiography; b) hematuria; c) severity of pain; d) flank pain; e) acuity of onset <24 hours; f) radiation of pain; g) visceral pain; h) nausea/vomiting; I) history of ureteral calculi; and j) male gender.(2) In a prospective study on selective IVP use in the

emergency department, clinical signs such as refractory pain, vomiting, and temperature over 38°C were all criteria for admission.(12)

Hematuria is the presence of blood in urine and it has been shown to be closely associated with the presence of urinary calculi. Detection of hematuria has been shown to be more accurate using a combination of urinanalysis and urine dipstick than either one alone.(1) We can measure the amount of red blood cells and white blood cells in urine using a microscope, a procedure called urinanalysis. On urinanalysis, the number of red blood cells and white blood cells visible in the high power field of the microscope indicates the amount of these cells in urine Amounts of greater than 100 red blood cells per high power field (RBC/HPF) on urinalysis have not been shown to significantly correlate with bad outcome. But there is a trend toward increased surgical intervention with amounts greater than 100 red blood cells per high power field on urinanalysis.(44) Pyuria is the presence of greater than 10 white blood cells per high power field of urine. This has been shown to correlate with the need for intervention. If the white blood cells per high power field were less than 10, then, surgical management of renal calculi was required only 35% of the time versus 65% if it was greater than 10 on urinanalysis.(44)

In a study describing the impact of pyelography on patient management, it was found that the diagnostic yield of intravenous pyelography was greater in patients over fifty years of age.(50)

In two separate urologic studies, duration of pain over 30 days has consistently been shown to predict the need for surgical intervention.(37;44)

Studies have shown that the need for intervention can be based on the following parameters: severity and duration of pain, signs and symptoms of infection, history of calculi and patient occupation. (26)

4.3.5 Variables from History

a) Age (Years),(50) **b) Gender** (Male/Female),(47;51;52) **c) Previous history of renal colic, d) Radiation of pain to groin**(2;12) **e) Duration of pain prior to arriving in the emergency department** (hours),(37;44) **f) vomiting,** (13) **g) Time to becoming pain free in the emergency department** (hours) it is postulated that the longer it takes to becoming pain free, the more severe the obstruction, (37;44) **h) Degree of pain at presentation** (measured with a validated pain score called a visual analogue scale (VAS). It is a 10 cm line with scores from 0 to 10 on a continuous scale),(53) **i) Persistent pain for days** (days) those with persistent pain may indicate a more severe obstruction, (44;47) **j) Dysuria** is pain on urination (Yes/No), and **k) Urinary frequency** is urinating more frequently than usual (Yes/No)

4.3.6 Variables from Physical Exam

a) Temperature greater than 38°C,(12;54;55) **b) presence of abdominal tenderness** (pain to palpation anywhere on the abdomen), **c) rebound tenderness** (pain to palpation with signs of peritoneal irritation anywhere on the abdomen), and **d) costovertebral angle tenderness** (pain when tapping over the kidney area on the back) .(12)

4.3.7 Variables from Diagnostic Tests

a) Degree of hematuria on urine dipstick (negative is less than 5 red blood cells per microliter (RBC/μl), small is 5-50 red blood cells per microliter, moderate is between 50 and 250 red blood cells per microliter, and large is greater than 250 red blood cells per microliter), **b) Degree of hematuria on urinalysis** (small is 0-10 red blood cells per high power field, moderate is 10-50 per high power field, large is 50 per high power field,(1;13) **c) pyuria on urine dipstick** (negative is less than 10 white blood cells per microliter, small is 10-25 white blood cells per microliter, moderate is 75 white blood cells per microliter, large >500 white blood cells per microliter. (44) **d) Serum Creatinine >150mmol/L**(8;12)

4.4 Outcome Measures

4.4.1 Primary Outcome for Urgent Use of IVP

4.4.1.1 Severe Urinary Obstruction

The primary outcome, severe acute urinary obstruction, is defined as a severe urinary tract blockage secondary to a urinary calculus (stone) on pyelography (8;13;39) and is characterized presently in the radiological literature as:

i) urine extravasation (leakage of urine into the pericalyceal tissue due to increased intrapelvic and tubular pressure in the kidney from obstruction),

ii) ureteral filling longer than 2 hours (dye appears in the ureter after 2 hours, the longer the filling the worse the obstruction)

iii) prolonged dense nephrogram with delayed filling of the ureter greater than 15 minutes

This definition was used to screen IVP's for cases of severe obstruction. Subsequently, they were reviewed by a urologist using a consensus definition (the final outcome definition) of severe obstruction.

4.4.1.2 Consensus Definition

This definition will serve as the basis for clearly defining our primary outcome of "clinically significant" urinary obstruction. A survey of specialists in the field of emergency medicine, urology and radiology was conducted via a questionnaire. Surveys were sent by hospital mail to all radiologists, urologists and emergency physicians practicing at the two sites of the Ottawa Hospital. A copy of the surveys is found in Appendices 4, 5 and 6. Physicians were asked to select from a list of radiological criteria those findings they currently use in

practice to identify clinically significant urinary obstruction. Those physicians not responding to the initial questionnaire were sent a second copy by hospital mail. If the physicians did not respond to the second questionnaire, they were contacted at their offices by telephone and reminded to complete the surveys. Results were then presented to a review panel of experts composed of a radiologist, an emergency physician and 5 urologists. Since urologists ultimately make this decision in clinical practice it was decided to have a majority of urologists on the panel. Patients with "clinically significant obstruction" were defined as a subset of patients with renal colic who have a greater risk of acute renal damage due to urinary calculi than other patients with the disease.

The panel reviewed the results of the surveys and after a discussion, agreed upon a final definition for clinically severe urinary obstruction. This definition was then used to analyze the data.

4.4.1.3 Feasibility of Data Collection Methods

Data collection feasibility was evaluated by: i-the number of completed data forms versus those not completed; ii- the physician ordering rates of IVP in patients with suspected renal colic; and iii- the pitfalls of obtaining adequate sample sizes.

4.4.2 Primary Outcome for Predicting Complications

4.4.2.1 Complications

Complications from prolonged obstruction were assessed after discharge from the emergency department in those patients whose intravenous pyelogram showed findings of ureteral calculus without severe obstruction. Evidence of a non obstructing calculus on IVP is defined as having *at least one* of the following radiological signs: (i)Prolonged dense nephrogram < 15 minutes (parenchymal contrast medium opacification of the kidney); ii)visualization of a calculus, iii)columnization ("column" of dye along the ureter); iv)delayed filling of the ureter < 2 hours (delay in dye excretion); v)hydronephrosis or ureteropyelocaliectasis (dilatation or effacement of the ureter, renal pelvis or calyx); vi)interureteric ridge edema (ureteral edema from recent passage of a calculus); and vii)cutoff of some portion of the ureter.(2;8;13;19)

Complications were defined by, at least, one of the following: i) development of a temperature >38°C from a urinary source as measured in the emergency department by tympanic membrane thermometer reading or after discharge by a physician seen in a follow-up visit; (12;54;55) (56) ii) persistent or recurrent pain at four weeks after emergency department discharge; (37) and iii) elevation in Creatinine > 150mmol/L in the emergency department or after discharge by a physician seen in a follow-up visit.(8;12)

4.4.2.2 Surgical Intervention

Surgical intervention was defined as requiring a urologic procedure to relieve the obstruction including ureteroscopy, percutaneous nephrostomy, extracorporeal shockwave lithotripsy, or open surgery. Patients who had

telephone follow-up at four weeks after discharge from the emergency department were asked about any procedures performed on them within the last month. The charts on these patients were obtained from medical records and procedures were recorded from the operative reports in the chart. In addition, all admitted patients had their medical charts reviewed for operative notes.

4.4.2.3 Loss to Follow-up

Losses to follow-up were recorded. Reasons for losses were also noted.

4.5 Data Analysis

4.5.1 Proportion of Severe Obstruction and Complications

For objectives 3.1.1, 3.2.1 and 3.2.3 the proportion of severe obstruction, complications and surgical intervention were expressed as proportions using 95% confidence intervals.

4.5.2 Strength of Association

For objectives 3.1.2 and 3.2.2 the strength of association between the clinical predictor variables and outcomes of severe obstruction and complications were tested using univariate analysis. Emergency physicians used standard data forms to assess clinical variables prospectively over the study period. Severe obstruction and complications were analyzed as dichotomous outcomes (yes or no). This process aided in the selection of the best variables for the logistic regression analysis.

The appropriate univariate technique was chosen according to the type of data: for nominal data, the chi-square test with continuity correction; for ordinal variables, the Mann-Whitney U test; and, for continuous variables, the unpaired 2-tailed t-test, using pooled or separate variance estimates as appropriate. All analyses were performed using the statistical software package SPSS version 10.0.5.

4.5.3 Multivariate Analysis

Preliminary modeling of a decision rule was conducted in this feasibility study. Those variables found to be strongly associated with the outcome measure ($P<0.05$) in the univariate analysis were combined using logistic

regression for a preliminary model for both severe obstruction and development of complications. The continuous variables in the univariate analysis that were significant were dichotomized. Cut points were determined by doing a chi-square analysis using Fisher's Exact 2-tailed test. Therefore all variables entered into the logistic model were dichotomous. The software package SPSS version 10.0.5 was used to perform the Logistic Regression. We obtained odds ratios and 95% confidence intervals on each variable in the preliminary logistic model for both severe obstruction and development of complications. Probabilities of the outcome given the combination of the variables in the logistic model were calculated. Sensitivity and specificity of the logistic models at various levels of probability were calculated and a receiver operating characteristics curves (ROC curve) were constructed. .(57) The areas under the ROC curve with 95% confidence intervals were calculated.

Using the results from the logistic regression model we determined the feasibility of pursuing a definitive study for determining severe obstruction and for development of complications. Sample size for the definitive study was explored using information from the logistic regression data. Two methods of sample size determination were used. The first method is discussed in a paper by Flack et Eudey (58) and uses a priori information on covariates. The second method estimates the precision around the sensitivity of the derived rule.

For the definitive study, sample size calculations will be based on the logistic regression data obtained from this feasibility study. Using logistic regression to obtain sample size estimates where pilot data is available is discussed in a paper by Flack et Eudey.(58) This method uses a priori information on covariates and is suitable for regressions with a small number of

explanatory variables, as in our feasibility study. In order to estimate sample sizes using this logistic method the feasibility data should: firstly, demonstrate that the logistic model with the chosen variable χ is plausible for fitting the proportions and, secondly, allow estimation of the logistic parameter vector β and its covariance matrix Σ.

For a given set of patient characteristicsχ, standard asymptotic theory gives the limiting distribution of the logit:

$$(\sqrt{\eta})\,(\log(\rho\hat{}/(1-\rho\hat{})) - \Lambda = (\sqrt{\eta})(\chi\beta\hat{} - \chi\beta) \rightarrow N(0, \chi S\chi')$$

For a given sample size η, we estimate the $(1 - \alpha)$ confidence interval for the logit as:

$$\chi\beta\hat{} \pm Z_{(1-\alpha/2)}\,(\sqrt{\chi S\chi'/\eta})$$

Where, $\pm Z_{(1-\alpha/2)}$ is the usual Gaussian $_{1-\alpha/2}$ percentile.

We can then apply the inverse transformation: $\rho = \exp(\chi\beta\hat{})/(1 + \exp(\chi\beta\hat{}))$ to the upper and lower bounds of the confidence interval for the logit to obtain the corresponding confidence interval (ρ_L, ρ_U). We can then explore the tightness of the confidence interval and determine an adequate sample size.

The second method of sample size calculation is the precision around the sensitivity method used in highly referenced decision rule studies.(34;59) It is calculated by determining 95% confidence intervals around various levels of sensitivity of the rule. A sample size of the outcome is then determined by selecting a desired confidence interval around the sensitivity.

4.5.4 Data Collection Feasibility

For objectives 3.1.4 and 3.2.4 the number of completed data forms, the IVP ordering rates and the losses the follow-up were expressed as proportions

using 95% confidence intervals. The time frame to obtain adequate sample sizes was based on the sample size calculations.

4.5.5 Pre-test Probabilities

For objective 3.1.5 the physicians estimated the patients' pre-test probability of having severe obstruction prior to patients having an IVP by circling a % from 0 to 100% on the data forms. This was compared to the actual presence or absence of severe obstruction. Sensitivity and specificity of physician judgment at various levels of probability were calculated and a receiver operating characteristics curve (ROC curve) was constructed.(57) The area under the ROC curve was calculated using SPSS software and compared with the area of the ROC curve for the preliminary model using ROC Analyzer.
For objective 3.2.5 physicians estimated the patients' pre-test probability of developing complications on the data form. The same procedure was followed as for 3.1.5 except that probabilities were measured on a Likert scale.

4.6 Sample Size

4.6.1 Patients Required

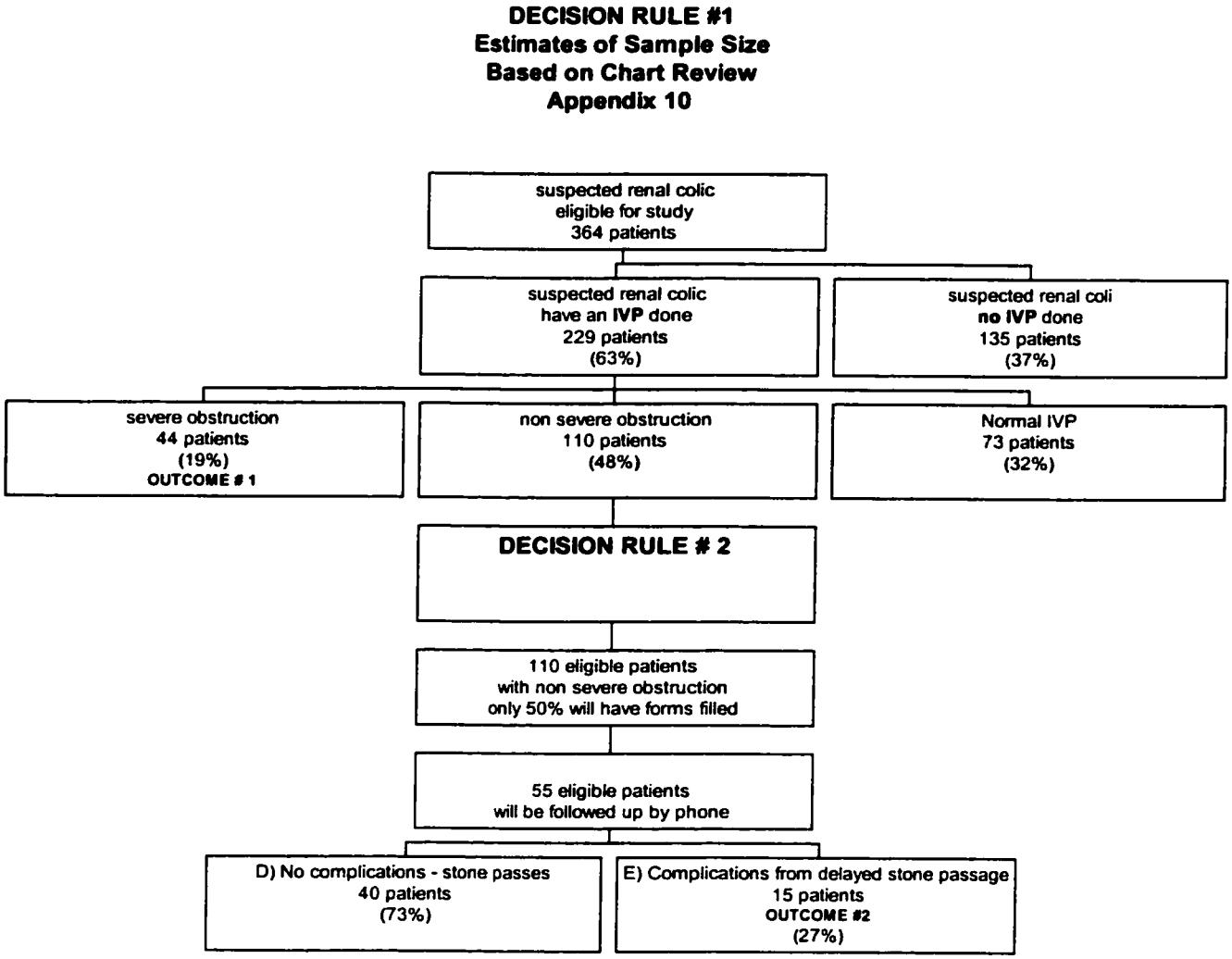
Based on a seven-week chart review done at the Ottawa Hospital Civic site from January 1998 to February 1998 we found an average of 7 cases of renal colic per week. Since the volumes of both emergency departments are equal (50,000 visits/year) we estimated about 14 cases of potential renal colic cases per week at both sites. We expected that 364 patients with suspected renal colic would present to the emergency departments of the Ottawa Hospital over 6 months.

The chart review showed a 63% IVP ordering rate at the Ottawa Hospital (Civic site) with a prevalence of severe obstruction of 19% of all IVP's. Therefore, we estimated that 229 patients would have IVP's done and that 44 of these IVP's would show severe obstruction.

In addition, we estimated that 110 patients would be discharged with a pyelogram that showed non-severe obstruction from a renal calculus.

Based on a 1998 study from our institution looking at complication rates for renal colic cases, (37) we expected a 27% complication rate. We followed by telephone a convenience sample of patients with non-severe obstruction in whom physicians had completed data forms. We anticipated a 50% completion rate for data forms by physicians. Since no hypothesis was being tested and this was a preliminary feasibility study, sample size was based on a reasonable time frame and an anticipated yield of approximately 15 patients with complications. Our sample size estimate is shown in Figure 5.

Figure 5
Sample Size Estimates



5. RESULTS: URGENT USE OF IVP

5. 1 Descriptive Statistics

5.1.1 Patient Enrollment

Data were collected at the Ottawa Hospital from January 1999 to August 1999. During this 7 month data collection period 437 patients were identified as having suspected renal colic: 312 (71.4%) at the Civic site and 125 (28.6%) at the General site. Enrollment of suspected renal colic cases was consistent throughout the study period but slightly higher in the spring and summer months (Figure 6).

Figure 6
Distribution of Suspected Renal Colic Patients Over the Data Collection Period

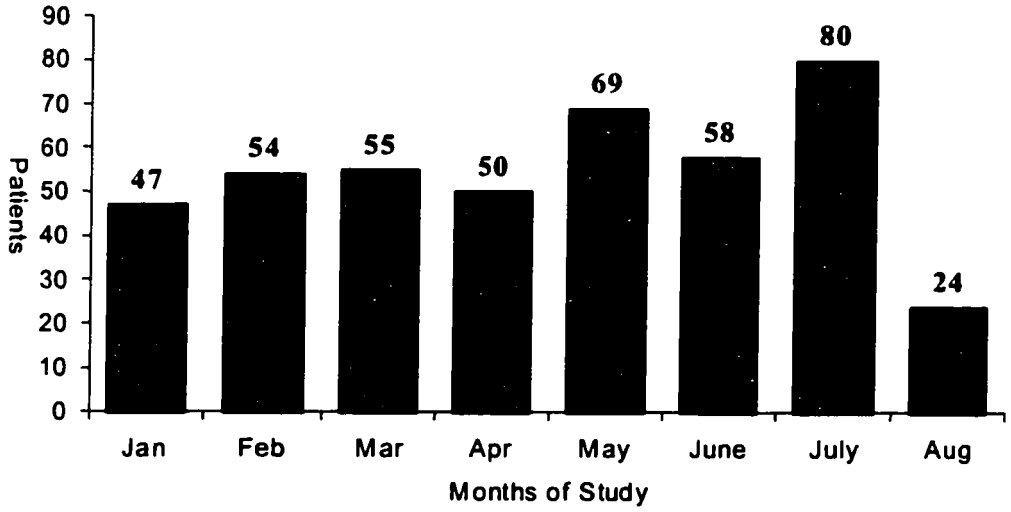
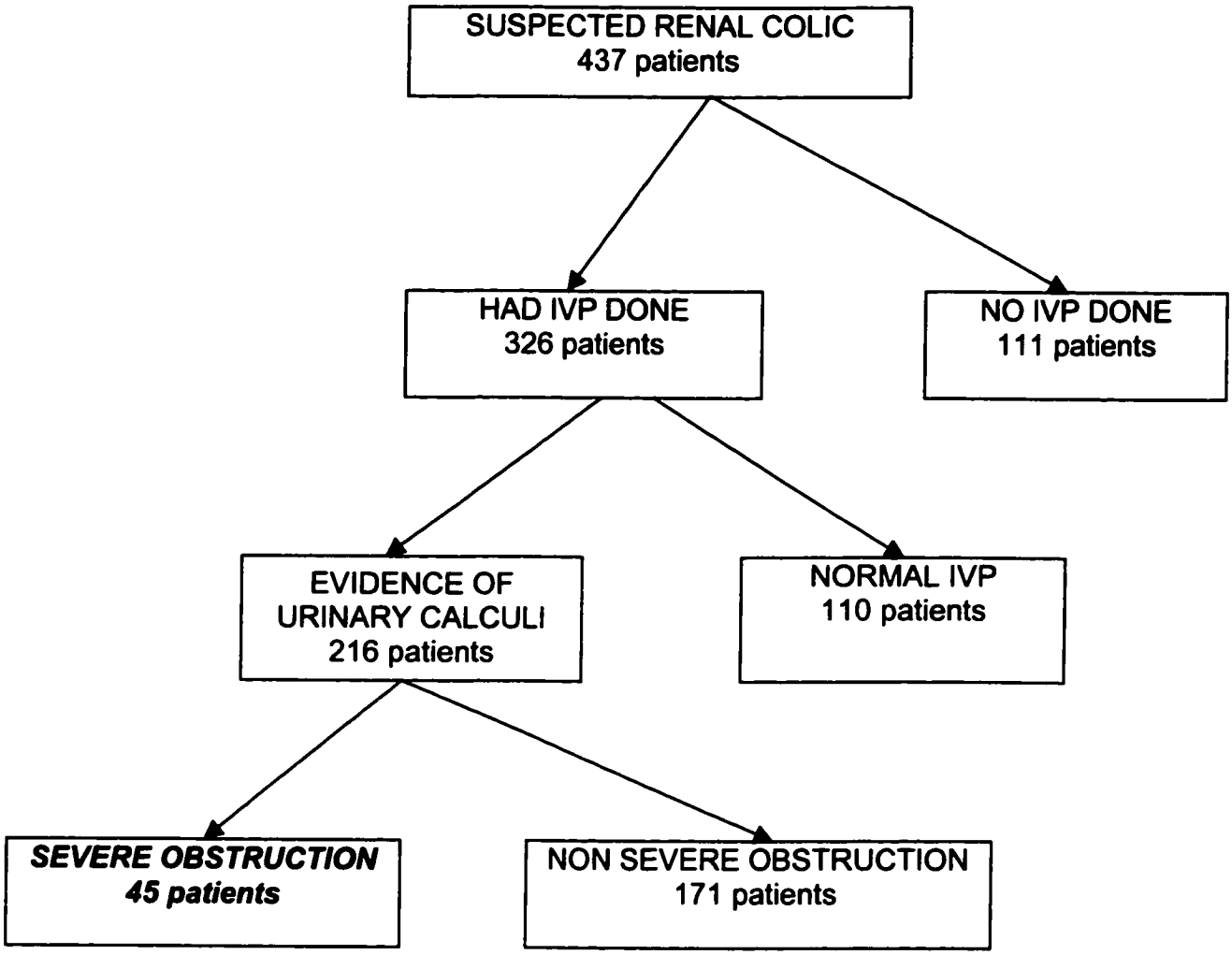


Figure 7 describes the flow of all patients in the study who had an IVP performed. The proportion of patients with suspected renal colic who had severe obstruction on IVP was 45 out of 326 patients or 14.0%. The definition of severe

obstruction used to calculate this proportion was a consensus definition that will be

Figure 7
Cohort for Severe Obstruction All Patients Presenting with Suspected Renal Colic Having had an IVP Performed Regardless of Data Form Completion

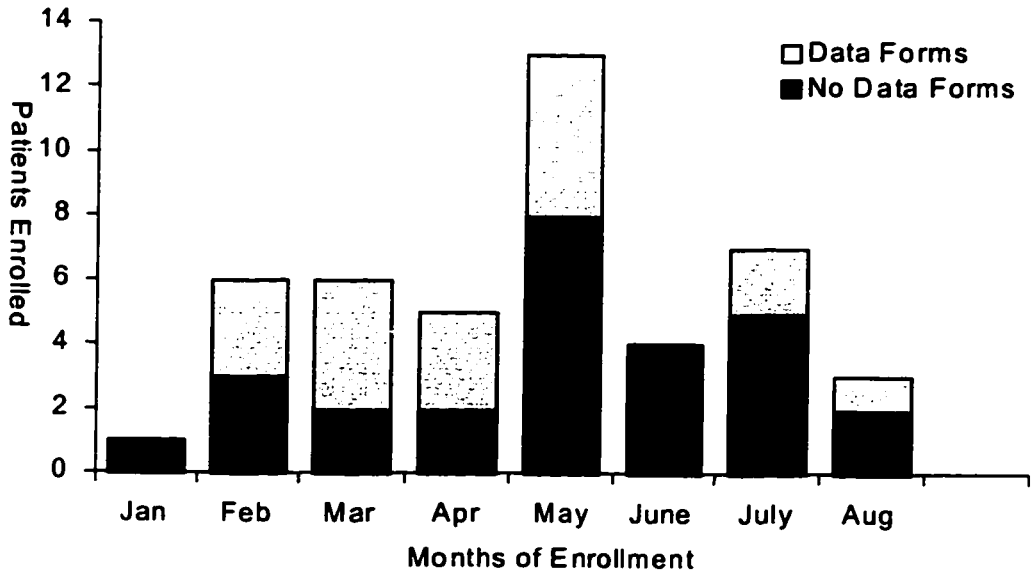


described in the next section (see 5.2.2). Of the 45 IVP's demonstrating severe obstruction, 7 IVP's showed extravasation alone, 34 IVP's showed delayed filling over two hours alone and 4 IVP's had both findings.

Summarized in Figure 8 is a comparison of data form completion and non-completion in the cases with severe obstruction over the eight month study

period. The number of patients with severe obstruction peaked in the month of May. The proportion of patients having severe obstruction in those with (15%) and without (13%) data forms was not statistically different (p=0.62).

Figure 8
Distribution of Cases of Severe Obstruction with and without Data Forms Over the Data Collection Period

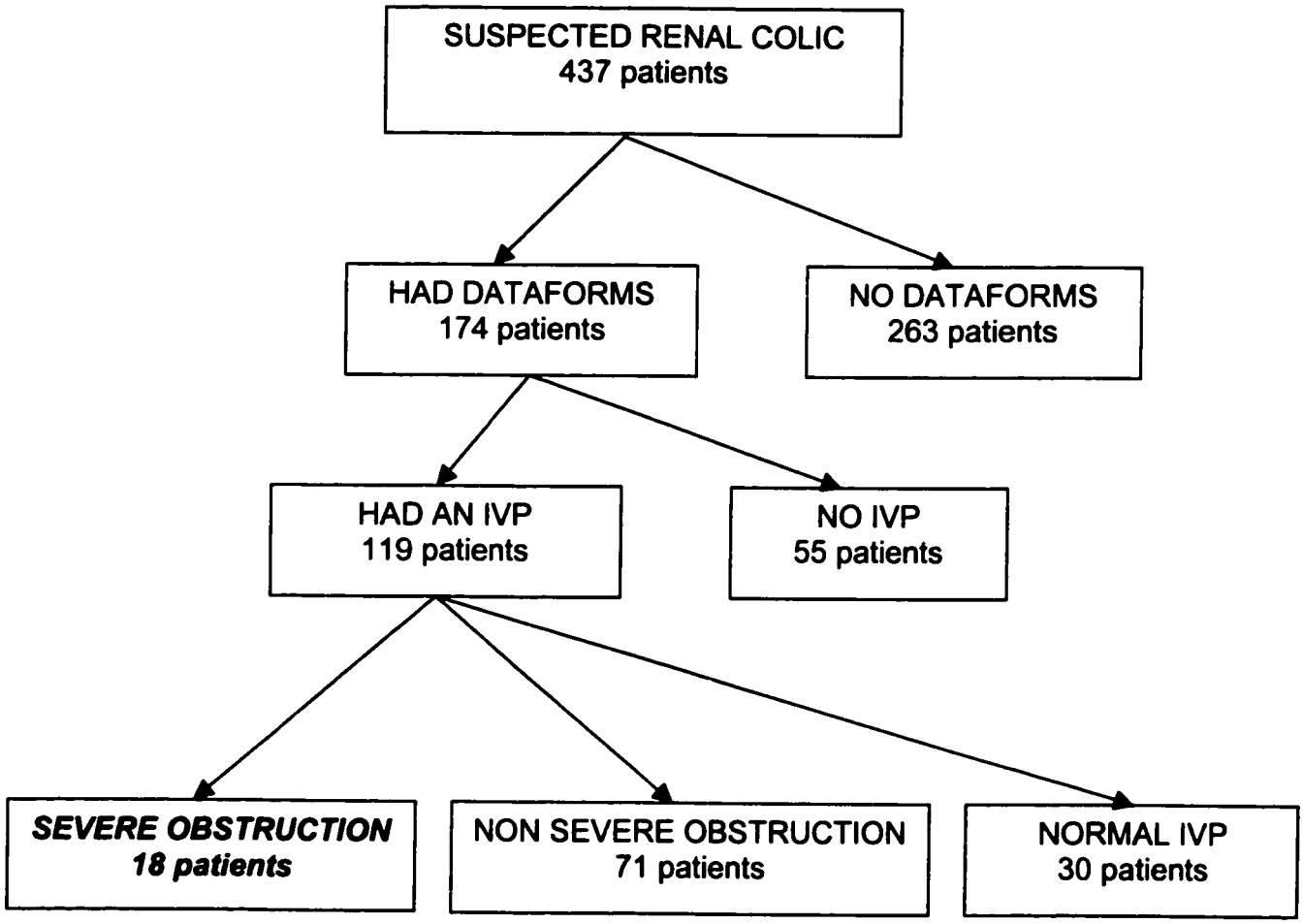


In comparing patients with and without completed data forms, it was found that there were no statistical differences in age, admission rate, history of renal colic, time spent in the emergency department or tests ordered. However, there were two statistically significant differences between the groups. In those with data forms completed 68% were male and in those without data forms completed 59% were males (p=0.04). Also, IVP's were ordered more frequently (79% versus 68%) in those with data forms (p=0.02). We attribute this to the higher prevalence of the disease in males rather than gender bias. In addition, IVP's were ordered more frequently in patients with data forms probably as a

result of the physician's higher clinical suspicion of urolithiasis in the patients for which they completed a data form.

Overall, physicians completed the data forms in 37% of those who had an IVP done. Therefore, of the 326 patients who had an IVP, 119 of them had a data form completed by a physician. Figure 9 describes how, from the potential 45 cases of severe obstruction observed over the study period, only 18 were available for further analysis because of incomplete data forms.

Figure 9
Cohort of Severe Obstruction Patients with both a Data Form Completed and an IVP Performed



At our institutions, IVP was the imaging study of choice as an initial imaging procedure for patients with suspected renal colic in the emergency department. Of the 437 patients with suspected renal colic, physicians ordered IVP in 326 (75%) of them, with 229 ordered from the Civic site and 97 from the General site. Other imaging modalities used to initially assess patients included plain abdominal radiographs (18%), ultrasounds (11%) and CT scans (3.5%).

Of the 326 IVP's performed, urinary calculi was detected on IVP in 216 patients. Table 1 compares the characteristics of four groups of patients: all patients with suspected renal colic, patients with data forms completed, patients who had an IVP done and patients with evidence of urinary calculi on IVP. The patient profiles are very similar in all four groups except for a higher percentage of patients admitted and referred to urology in those whose IVP's confirmed urinary calculi.

In addition, patient profiles are very similar to those reviewed in the background literature. Patients are on average 44 years of age with a male to female ratio of about 2:1. About 18% of patients with positive IVP's were admitted. About 60% of cases had a previous history of renal colic. In those patients with confirmed urinary lithiasis on IVP, hematuria was present in 90% on a urine dipstick test in the emergency department and present in 95% when a microscopic urinanalysis was performed in the laboratory. Physicians ordered a serum creatinine level on 85% of patients.

Overall, patients with suspected renal colic had an average length of stay in the emergency department of 5 hours and 45 minutes. Length of stay for those patients who had an IVP was 6 hours and 20 minutes and those not

having an IVP stayed 4 hours. Thus, patients having an IVP stayed an average of 2 hours and 20 minutes longer than those without.

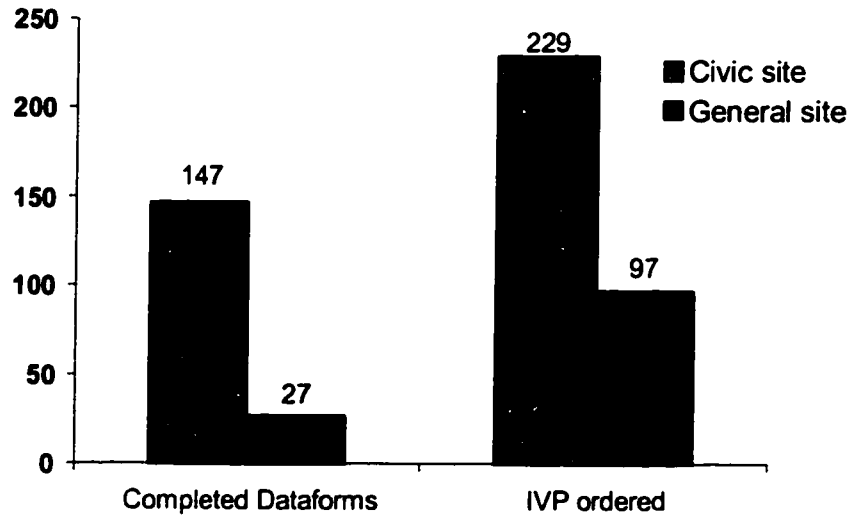
Table 1
Comparison of the Characteristics Between Four Groups of Patients

Characteristics	*Patients with suspected renal colic N=437*	*Patients with data forms N=174*	*Patients having an IVP done N=326*	*Patients whose IVP showed calculi N=216*
Demographics				
Mean Age in years (±SD)	44 (15)	43 (14)	44 (14)	45 (14)
Male Gender (%)	63	68	60	67
Admitted (%)	12	11	15	18
Referred to Urology (%)	64	67	69	80
Data Forms completed (%)	40	100	37	41
Time of registration to time seen by MD (min±SD)	60(126)	61(162)	54(105)	53 (123)
Time seen by MD to time of discharge (min±SD)	294 (187)	291 (183)	327 (188)	329 (170)
Total length of stay in ED from registration to discharge (min±SD)	344 (190)	335 (192)	373 (195)	373 (185)
MD thought patient was exaggerating pain on arrival (%)	5	5	4	5
Tests Ordered in Emergency				
IVP (%)	75	68	100	100
KUB (%)	18	17	14	14
Ultrasound (%)	11	9	9	6
CT Scan (%)	3	1	3	2
Serum Creatinine ordered (%)	84	84	87	86
Urine dipstick performed (%)	89	93	87	85
Urinanalysis performed (%)	26	22	27	20
Results of Tests				
IVP evidence of calculus (%)	49	51	75	100
KUB evidence of calculus (%)	53	54	50	63
Ultrasound evidence calculus (%)	26	14	30	60
CT Scan evidence of calculus (%)	46	0	50	100
Hematuria on urine dipstick (%)	85	90	83	90
Pyuria on urine dip (%)	20	18	21	17
Hematuria on urinanalysis (%)	78	87	77	95
Mean serum Creatinine (μmol/L±SD)	88 (57)	96 (86)	85 (23)	89 (23)

5.1.2 Evaluation of Data Collection Process

Part of the feasibility study was to identify potential pitfalls in the data collection process. There was a discrepancy in patient enrollment and data form completion between the two sites of the Ottawa Hospital. Of the 437 patients identified as having suspected renal colic 71% were enrolled from the Civic site alone. However, if we compare the ordering rate at the two sites it is very similar. Civic physicians ordered IVP's in 73% of suspected renal colic patients and physicians at the General site ordered IVP's in 77% of suspected cases. Data form completion rate at the General site was substantially lower 27 (22%) than that at the Civic site 147 (48%). Figure 10 compares the two sites of the Ottawa Hospital for the number of IVP's ordered as well as the data forms completed.

Figure 10
Comparison of the 2 Sites of the Ottawa Hospital in Data Form Completion and IVP Ordering Rate



Another source of missing data can be observed on the individual data forms. Variable completion on the data forms was quite high for most of the 20

variables. Those variables most often left blank were those requiring patient reassessment after initial evaluation of the patient. Discharge VAS pain scores were missing on 30% of forms, time to pain relief was missing on 25% of forms and time of discharge was not recorded in 10% of cases. All other variables completed at the time of the initial patient evaluation were recorded consistently and were seldom left incomplete.

5.2 Defining Clinically Significant Obstruction

5.2.1 Survey of Urologists, Radiologists and Emergency Physicians

In January of 1999 surveys were sent to urologists, radiologists and emergency physicians of the Ottawa Hospital asking them to define clinically severe obstruction (Appendices 4,5, and 6). A total of 73 surveys were sent out. Those physicians not responding to the initial questionnaire after one month were sent a second copy of the same questionnaire. If the physicians did not respond they were contacted by telephone and reminded to complete the surveys. The overall response rate was 47 (64%). The physicians sampled had an average of 14 years practice experience.

Most physicians agreed that the terms "severe" and "complete" obstruction are clinically part of the same spectrum of obstructive disease with "complete" being at the extreme. It was felt that severe obstruction would encompass more patients and would be more easily defined than complete obstruction. The results of the surveys are provided in Table 2.

There were three findings on IVP consistently selected by all three groups of specialists to define significant obstruction. These 3 radiological findings on IVP were: i) prolonged dense nephrogram greater than 15 minutes duration; ii) extravasation; and iii) delayed filling of the ureter below the obstruction greater than 2 hours. These results are compatible with the current definition found in the literature and in the urologic textbooks.(8;12;13;19;37)

Table 2
Defining Significant Obstruction on IVP
Results of Surveys to Urologists, Radiologists and Emergency Physicians

	Overall N=47	*Urologists* N=10	*Radiologists* N=15	*ED Physicians* N=22
Years in practice (mean±SD)	14 (7.8)	16 (1.7)	17 (6.8)	10 (7.3)
No difference between Complete vs. Severe obstruction (%)	43	30	53	45
Radiological indicators of Severe obstruction				
Dense nephrogram >15 min(%)	66	60	87	55
Extravasation (%)	81	60	87	86
Delayed Filling >2 hrs (%)	60	30	87	55
Delayed Filling >4 hrs (%)	19	40	0	23
Delayed Filling >6 hrs (%)	15	30	7	14
Site of Calculus(%)	13	10	0	23
Size of Calculus (%)	17	0	7	32
Degree Hydronephrosis (%)	19	20	13	23

5.2.2 Consensus Definition

An expert panel, composed of a radiologist, an emergency physician and 5 urologists were assembled to discuss the results of the surveys and to agree on a final definition for clinically severe urinary obstruction. The results of the panel discussion are summarized in Table 3. It was agreed that the most significant findings that should be used to determine clinically important obstruction were: i) extravasation and ii) delayed visualization of contrast beyond the obstruction greater than 2 hours. It was believed that prolonged dense nephrogram could not be used as an independent criterion to suggest severe

obstruction. There was some discussion among the urologists on the panel whether 2 hours or 4 hours should be

Table 3
Panel Definition of Clinically Significant Urinary Obstruction From Calculi on IVP

Finding on IVP suggestive of severe obstruction	*Said yes (%)*	*Accepted after discussion*
Prolonged dense nephrogram >15 min	29	No
Extravasation	100	Yes
Delayed filling below obstruction >2hr	71	Yes
Delayed filling below obstruction >4 hrs	29	No
Delayed filling below obstruction >6 hrs	0	No
Site of calculus	0	No
Size of calculus	0	No
Degree hydronephrosis	0	No

used as the cutoff for delayed visualization of contrast below the obstruction. It was felt by the majority of the panel that being conservative was the best option and it was agreed to use the 2-hour delay instead of the 4-hour delay in contrast visualization. Figure 11 demonstrates actual radiographic pictures of IVP's with these findings.

The final definition of *clinically significant severe urinary obstruction on intravenous pyelogram was*:

1) Presence of extravasation of dye along the urinary system

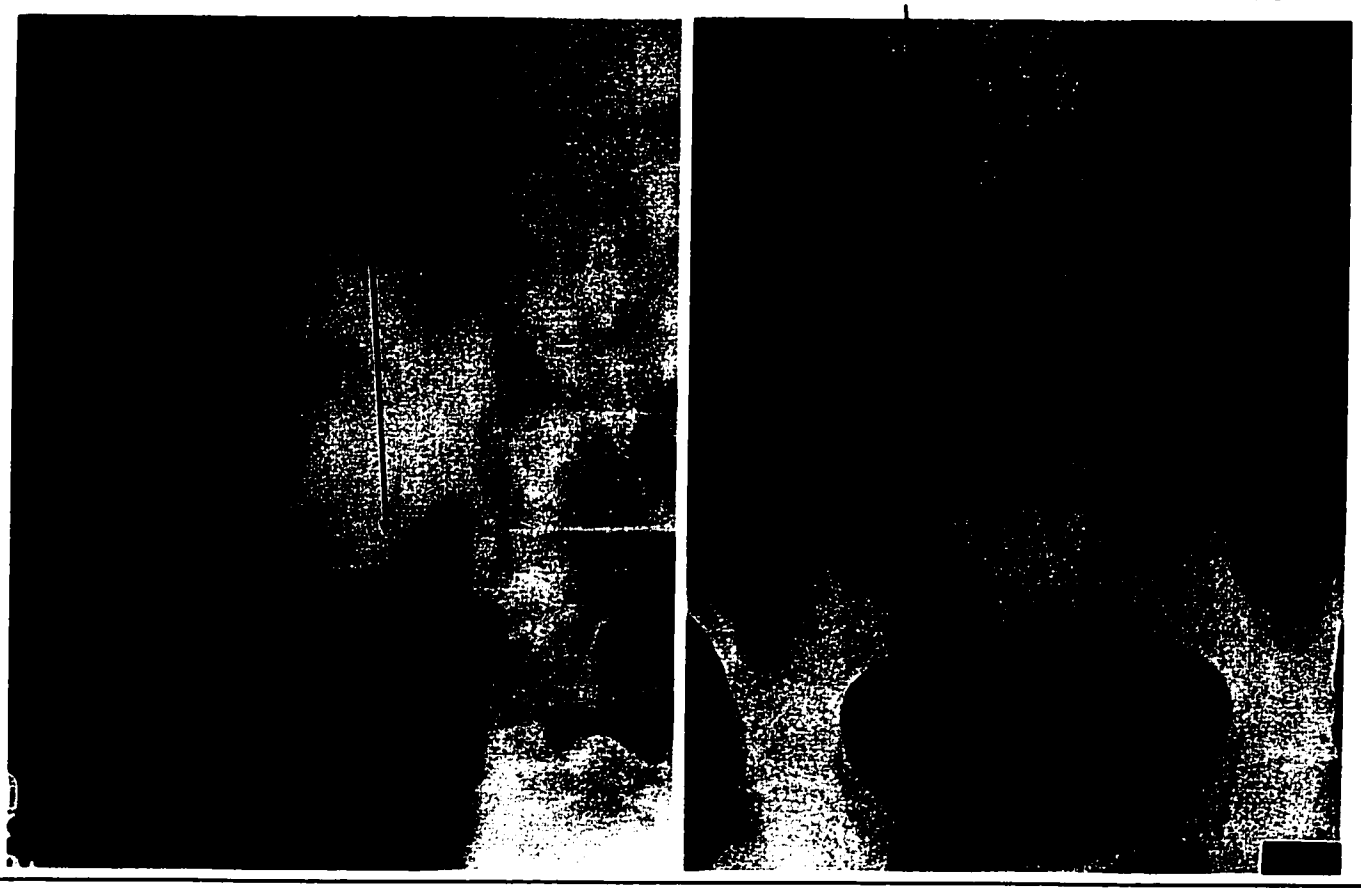
and/or

2) Delayed visualization of contrast beyond the site of obstruction along the urinary system that is greater than 2 hours

Figure 11 RadiologicFindings of Severe Obstruction on

Extravasation of contrast

Dye unable to pass obstructing calculus 2 hrs after iv contrast



5.3 Testing Clinical Predictors

5.3.1 Univariate Analysis

The cohort of patients used in the univariate analysis is outlined in Figure 9. Only those clinical variables obtained prospectively from the physician data forms were used in the univariate analysis. The results of the univariate analysis are presented in Table 4. Those variables having a statistically significant ($p<0.05$) association with severe obstruction included: i) vomiting at any time from symptom onset with (67%) in the severely obstructed group versus (39%) in the non obstructed group; ii) patient's VAS pain score on arrival with a mean of 9.0 cm (SD1.3) in the severe group versus 7.0 cm (SD3.0) in the non severe group; and iii) physician's estimate of patient's VAS pain score on arrival with 8.2 cm (SD1.9) in the severe group versus 6.3 cm (SD2.9) in the non severe group.

Due to the small numbers of our primary outcome in this feasibility study, we decided to include those variables from the univariate analysis with a $p<0.15$ in the multivariate analysis. This allowed us to recruit three additional variables: i) length of time for patient to achieve a pain free state in the emergency department; ii) rebound tenderness on physical exam; and iii) mean VAS score at discharge. It took longer to control a patient's pain in the emergency department if they had severe obstruction (5.5 hrs) than if they did not (3.66 hrs). Rebound tenderness occurred 5 times more frequently in those with severe obstruction. Mean VAS pain scores at discharge were higher (2.3 cm) in those with severe obstruction than those without (0.9 cm).

In preparation for the logistic regression analysis the four variables that were continuous were dichotomized at various cut points and their association tested for significance by chi-square analysis (Table 5). They included patient's pain score on arrival, physician's estimate of patient's pain score on arrival, patient's pain score at discharge and time to a pain free state.

Table 4

Univariate Analysis Comparing Severe vs. Non Severe Obstruction in 119 Patients with IVP and Completed Data Forms

Variables	*Severe (n=18)*	*Not severe/ Normal (n=101)*	*P-Value*
Demographics			
Age (mean in years±SD)	44 (12.0)	43 (13.7)	0.72
Gender (male) %	78	68	0.58
Admitted %	33	12	0.03
Revisits (n=61) %	50	47	1.00
Referred to Urology (n=117) %	83	73	0.55
Prior History of Renal Colic (n=118) %	39	49	0.46
Signs & Symptoms			
Temperature (n=101) (°C ±SD)	36.2 (0.8)	36.1 (0.9)	0.89
Vomiting %	67	39	0.04
Dysuria %	11	18	0.73
Frequency %	17	28	0.40
Abdominal Tenderness %	50	41	0.61
Rebound Tenderness %	11	2	0.11
CVA Tenderness %	61	60	1.00
Pain Measures			
Hours of continuous pain before presenting to ED (n=118) (hrs±SD)	4.8 (6.3)	4.8 (6.0)	1.00
Days of pain prior to visit (n=118) (days±SD)	2.2 (4.5)	3.1 (5.6)	0.51
Patient's VAS at arrival (n=117) (cm±SD)	9.0 (1.3)	7.0 (3.0)	<0.001
Patient's VAS at discharge (n=83) (cm±SD)	2.3 (2.4)	0.9 (1.6)	0.09
Physician's VAS at patient's arrival (n=116) (cm±SD)	8.2 (1.9)	6.3 (2.9)	0.001
Time to a pain free state in the ED (n=95) (min±SD)	317 (248)	203 (140)	0.13
Laboratory Results			
Hematuria present on urine dip (n=107) %	94	87	0.69
Degree of hematuria on urine dip (n=107)			0.90
Degree of pyuria on urine dip (n=107)			0.99
Degree of hematuria on urinalysis (n=25)			0.37
Pyuria present on urine dip (n=107) %	18	18	1.00
Hematuria present on urinalysis (n=25) %	100	83	1.00
Serum BUN (n=99) (μmol/L±SD)	5.9 (2.1)	5.8 (1.9)	0.93
Serum Creatinine (n=104) (μmol/L±SD)	92.1 (20.5)	86.1 (21.8)	0.30

Table 5

Cut points of Continuous Variables Dichotomized for the Logistic Regression for Detecting Severe Obstruction

Variable Cut points	*Pearson Chi-Square*	*Fisher's Exact Test (2-sided)*
Patient's VAS score on arrival (in cm)		
VAS ≥ 7.0	4.1	0.053
VAS ≥ 8.0	3.7	0.073
VAS ≥ 8.5	6.2	0.019
VAS ≥ 9.0	4.6	0.038
Physician's estimate of patient's VAS score on arrival (in cm)		
VAS ≥ 6.0	6.8	0.011
VAS ≥ 7.0	12.8	<0.0001
VAS ≥ 8.0	3.4	0.074
VAS ≥ 9.0	3.0	0.092
Patient's VAS score at discharge (in cm)		
VAS ≥ 1.0	7.4	0.011
VAS ≥ 2.0	2.4	0.210
VAS ≥ 3.0	4.9	0.049
VAS ≥ 4.0	4.5	0.068
VAS ≥ 5.0	2.3	0.178
Time for patient to achieve a pain free state in the ED (in minutes)		
T ≥ 240 (4 hrs)	0.6	0.540
T ≥ 300 (5 hrs)	2.7	0.176
T ≥ 360 (6 hrs)	10.4	0.005
T ≥ 420 (7 hrs)	6.6	0.029

This resulted in 6 variables with dichotomous outcomes available for the logistic regression analysis: i) vomiting, ii) patient's VAS score greater than or equal to 8.5 cm on arrival, iii) physician's estimate of patient's VAS score of greater than or equal to 7 cm on arrival, iv) rebound tenderness, v) continued pain after 6 hours in the emergency department, and vi) patient's VAS score greater than or equal to 1 cm at discharge.

5.3.2 Multivariate Analysis using Logistic Regression – Model #1

Using the significant clinical variables from the univariate analysis we performed logistic regression analysis and modeled two preliminary rules. These models will be used to help calculate the sample size for the definitive study.

Logistic regression analysis used 72 of a possible 119 cases due to missing data. Data was incomplete for 2 cases of patients' VAS score at arrival, 3 cases of physicians' estimated VAS, 24 cases of time to pain relief and 37 cases of discharge VAS pain scores. Stepwise backward logistic regression analysis yielded a model of four variables (Table 6).

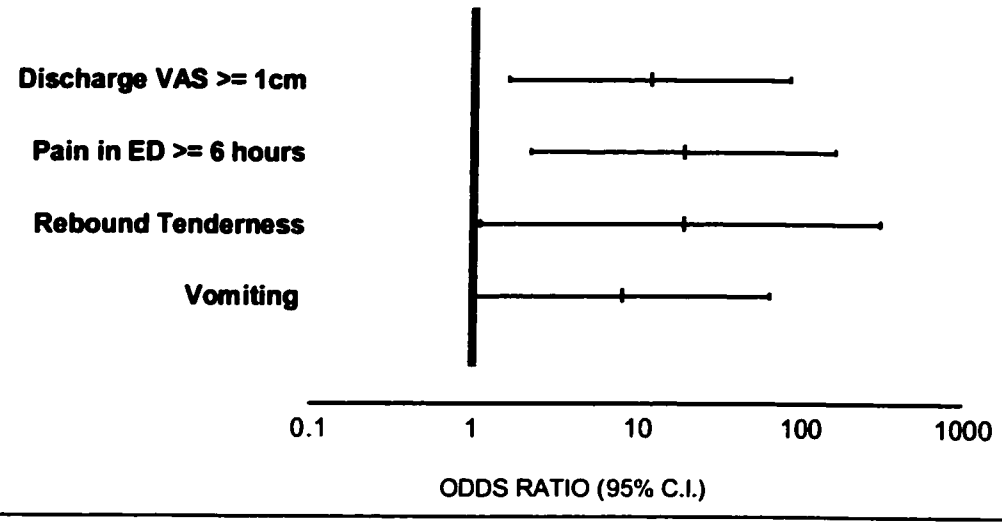
Table 6
Variables in Logistic Regression Model #1 For Severe Obstruction

Variable	*β*	*SE*	*Wald statistic*	*p-value*	*Odds Ratio*	*95% C.I.*
Discharge VAS ≥ 1cm	2.5	1.0	6.0	0.014	11.6	1.6 – 82.1
Pain after 6 hours in ED	2.9	1.1	7.1	0.008	18.6	2.2 – 158.8
Rebound tenderness	2.9	1.4	4.2	0.041	18.5	1.1 – 302.9
Vomiting	2.1	1.1	3.8	0.052	7.8	1.0 – 61.9

The odds ratio for each variable, adjusted for the other variables, is displayed with 95% confidence intervals in Figure 12. The Hosmer-Lemeshow

goodness-of-fit test yielded a chi-square of 7.4 (6 degrees of freedom (df)) with a significance of 0.283 for this Model #1.

Figure 12
Odds Ratio with 95% C.I. for Predictor Variables in Model #1



The probability of severe obstruction given a specific set of predictor variables can be expressed with the following formula:

$$P(X) = \frac{1}{1 + e^{-(\alpha + \sum \beta i x i)}} = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$$

$$= -5.171 + 2.452\,(x_1) + 2.922\,(x_2) + 2.918\,(x_3) + 2.054\,(x_4)$$

where, x_1 = discharge VAS $\geq$ 1 cm

x_2 = continued pain after 6 hrs in the ED

x_3 = rebound tenderness

x_4 = vomiting

If a patient were negative for all variables in the model (i.e. $x_1=0$, $x_2=0$, $x_3=0$, $x_4=0$) that patient would have a probability of severe obstruction of 0.6%.

$$P(X) = \frac{1}{1 + e^{-(-5.171)}} = 0.006 = 0.6\%$$

Alternatively, if a patient were positive for all variables in the model (i.e. $x_1 = 1$, $x_2 = 1$, $x_3 = 1$, $x_4 = 1$) that patient would have a probability of severe obstruction of 99.4%

$$P(X) = \frac{1}{1 + e^{-(5.175)}} = 0.994 = 99.4\%$$

If we test all possible combinations of these four variables we obtain a series of probabilities for each combination as shown in Table 7.

Table 7
Probability of Severe Obstruction based on the Presence or Absence of the 4 Predictor Variables from the Logistic Regression Model #1

# of vari-ables	Discharge VAS ≥ 1 cm (x_1)	No pain relief after 6 hrs (x_2)	Rebound tender-ness (x_3)	Vomiting (x_4)	Prob of severe obstruct ion P(x)	%	Sensitivity %
One	1	0	0	0	0.062	6	90
	0	1	0	0	0.095	10	90
	0	0	1	0	0.095	10	90
	0	0	0	1	0.042	4	90
Two	1	1	0	0	0.055	56	20
	1	0	1	0	0.550	55	30
	1	0	0	1	0.340	34	60
	0	0	1	1	0.450	45	60
	0	1	0	1	0.451	45	60
	0	1	1	0	0.661	66	20
Three	1	1	0	1	0.905	91	0
	1	1	1	0	0.958	96	0
	1	0	1	1	0.905	91	0
	0	1	1	1	0.938	94	0
Four	1	1	1	1	0.994	99	0

With any one variable there is a 4-10% probability of severe obstruction, with any two variables there is a 34-66% probability of severe obstruction and with any three variables there is a 91-96% probability of severe obstruction. As we can see from the logistic regression equation (- 5.171 + 2.452 (x_1)+ 2.922 (x_2) + 2.918 (x_3) + 2.054 (x_4)), the coefficient of each variable ranges from 2.054 to 2.922. Although, each variable has a different weight in predicting severe obstruction, the coefficients are all within close range of each other. From Table 7 we observe that having any one variable present will give a 90% sensitivity.

5.3.3 Classification Performance – Model #1

If we use this logistic regression model as a preliminary decision rule based on the predicted probability threshold of severe obstruction of 0.5 (50%) we would correctly classify 89% of patients but we would miss 70% of patients with severe obstruction. If, however, we reduce the probability threshold to a 0.1, (10%) predicted probability of severe obstruction, we would correctly classify 86% but our sensitivity increases to 90% and we would only miss 10% of cases. This occurs at the expense of the specificity being lowered to 85% (Table 8).

Classification performance of the regression model using predicted probabilities of severe obstruction of 0.1, 0.5 and 0.75 are shown in Table 8. As the probability threshold decreases the sensitivity increases. An ROC curve was generated for the logistic regression model by calculating the sensitivity and specificity for the various probabilities of severe obstruction from 0.01 to 0.99.

Table 8
Classification Performance of the Logistic Regression Model #1 Using Predicted Probabilities of Severe Obstruction of 0.1, 0.5 and 0.75

Predicted Probability of 0.1

	Actual Severe Obstruction	
	YES	**NO**
Predicted Severe Obstruction		
YES	9	37
NO	1	25

Sensitivity	90%
Specificity	85%
Accuracy	86%
Positive Predictive Value	50%
Negative Predictive Value	98%

Predicted Probability of 0.5

	Actual Severe Obstruction	
	YES	**NO**
Predicted Severe Obstruction		
YES	3	1
NO	7	61

Sensitivity	30%
Specificity	98%
Accuracy	89%
Positive Predictive Value	75%
Negative Predictive Value	90%

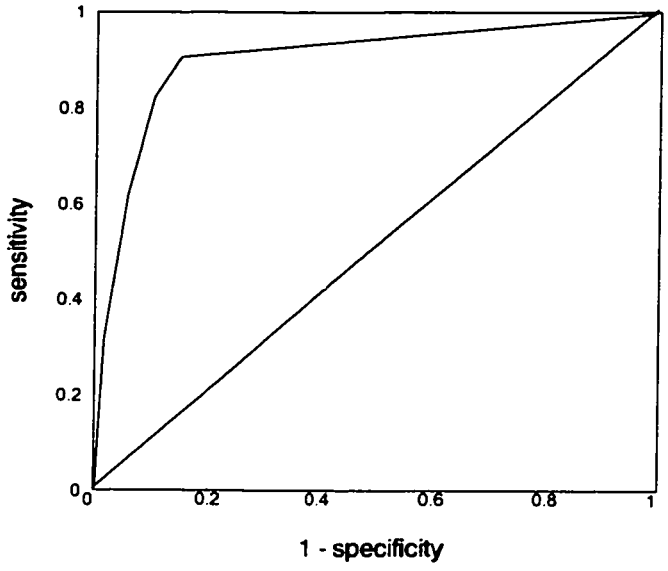
Predicted Probability of 0.75

	Actual Severe Obstruction	
	YES	**NO**
Predicted Severe Obstruction		
YES	2	0
NO	8	62

Sensitivity	20%
Specificity	100%
Accuracy	89%
Positive Predictive Value	100%
Negative Predictive Value	89%

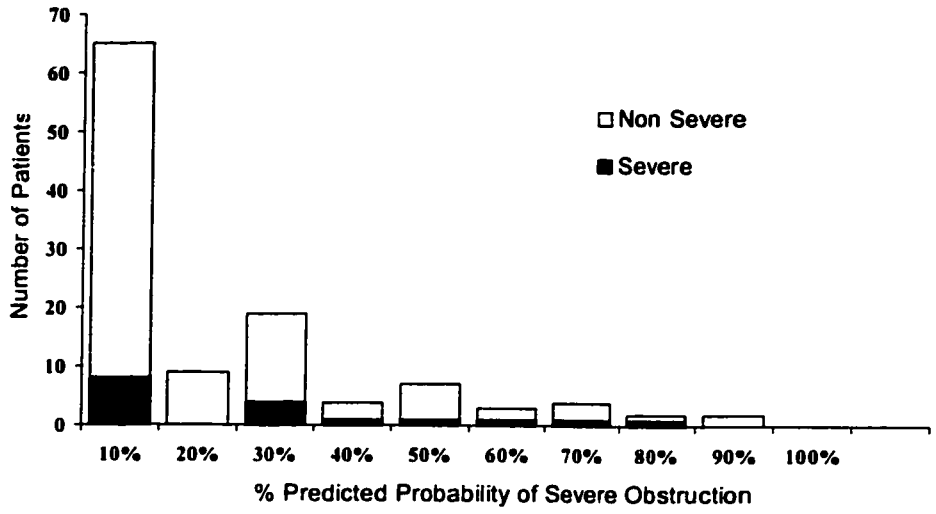
The area under the ROC curve for this model is 0.89 (95%CI 0.74-1.03) ($p<0.001$). (Figure 13). The shape of the curve shows very good discrimination between those with and without severe obstruction. A sensitivity of 90% is reached quickly with specificity of 85%, but a sensitivity of 100% is not achieved until the specificity is 0%.

Figure 13
ROC Curve of the Classification Performance of the Logistic Regression Model #1 for Severe Obstruction



We tested emergency physician's ability to predict severe obstruction based on their expert clinical evaluations. Severe obstruction occurred frequently when the physicians' clinical suspicions of its presence were low and occurred infrequently when physicians had a high suspicion (Figure 14).

Figure 14
Physicians' Predicted Probability of Severe Obstruction versus Presence of Severe Obstruction on IVP



Classification performance of the physicians' predicted probability of severe obstruction with thresholds of 0.01, 0.5 and 0.75 for ordering an IVP are shown in Table 9. Physicians would have missed 83% of severe obstruction cases if the threshold for ordering an IVP to detect severe obstruction had been their own predicted probability of at least 50%. An ROC curve for physicians' predicted probability as a model for detecting severe obstruction is shown in Figure 15.

Figure 15
ROC Curve of the Classification Performance of Physicians' Predicted Probability of Severe Obstruction

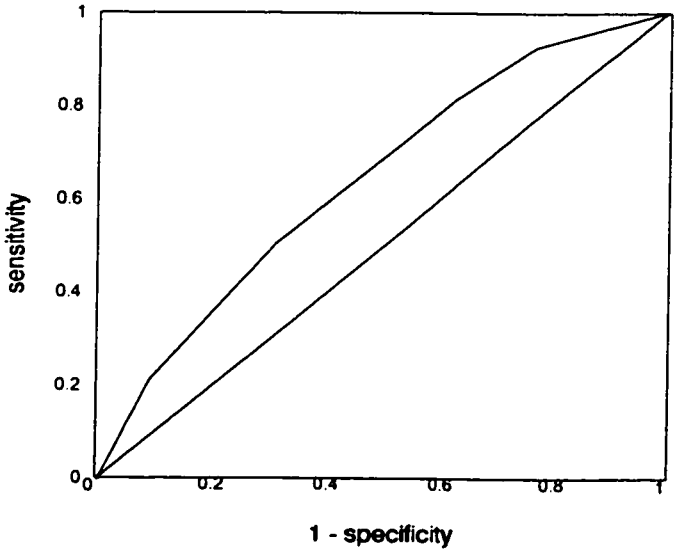


Table 9
Classification Performance of the Physician's Predicted Probability of Severe Obstruction with Probability Thresholds of 0.1, 0.5 and 0.75

Predicted Probability of 0.1

	Actual Severe Obstruction	
	YES	**NO**
Predicted Severe Obstruction		
YES	9	41
NO	9	59

Sensitivity	50%
Specificity	59%
Accuracy	58%
Positive Predictive Value	18%
Negative Predictive Value	87%

Predicted Probability of 0.5

	Actual Severe Obstruction	
	YES	**NO**
Predicted Severe Obstruction		
YES	3	8
NO	15	92

Sensitivity	17%
Specificity	92%
Accuracy	81%
Positive Predictive Value	27%
Negative Predictive Value	86%

Predicted Probability of 0.75

	Actual Severe Obstruction	
	YES	**NO**
Predicted Severe Obstruction		
YES	1	3
NO	17	97

Sensitivity	6%
Specificity	97%
Accuracy	83%
Positive Predictive Value	33%
Negative Predictive Value	85%

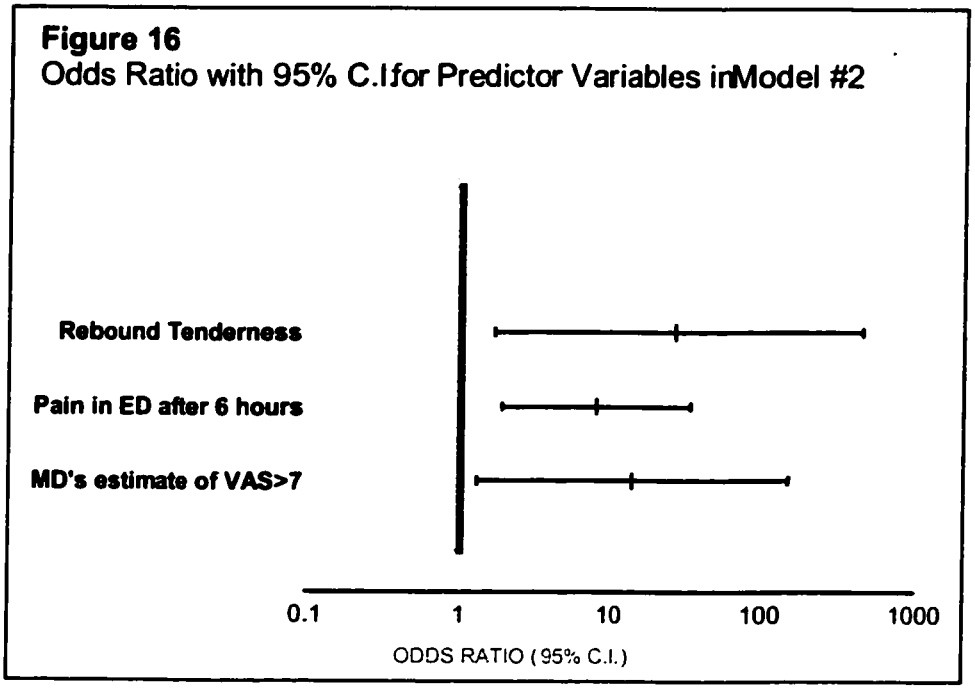
The area under the curve is 0.64 (95%CI 0.51-0.77) (p=0.063). If we compare the area of the our clinical Model #1 versus that of physicians' judgment our clinical model is significantly more accurate (p=0.007) in predicting severe obstruction.

Despite the small sample size of this feasibility study, these initial results are very promising and show that this preliminary rule should be developed further.

5.3.4. Multivariate Analysis using Logistic Regression - Model #2

Discharge VAS was the variable that was missing most frequently of all variables in the model. In order to increase the sample size available for the logistic regression analysis, we removed the variable discharge VAS ≥ 1 cm and performed another logistic regression. This allowed us to include an additional 18 cases for the analysis. This logistic regression analysis used 90 of a possible 119 cases. Stepwise backward logistic regression analysis yielded a model consisting of three variables (Table 10). The odds ratio for each variable, adjusted for the other variables, is displayed with 95% confidence intervals (Figure 16).

Figure 16
Odds Ratio with 95% C.I for Predictor Variables in Model #2



The Hosmer-Lemeshow goodness-of-fit test yielded a chi-square of 2.4 (5 df) with a significance of 0.79 for this Model #2.

Table 10
Variables in Logistic Regression Model #2 For Severe Obstruction

Variable	*β*	*SE*	*Wald statistic*	*p-value*	*Odds Ratio*	*95% C.I.*
Rebound tenderness	3.3	1.4	5.5	0.019	28.0	1.7 – 450.0
Pain after 6 hours in ED	2.1	0.747	7.8	0.005	8.1	1.9 – 35.0
MD's VAS over 7cm	2.6	1.213	4.7	0.030	14.0	1.3 – 150.8

The probability of severe obstruction given a specific set of predictor variables can be expressed with the following formula:

$$P(X) = \frac{1}{1 + e^{-(\alpha + \sum \beta i x i)}} = .$$

$$\alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 = -4.653 + 3.334\,(x_1) + 2.092\,(x_2) + 2.639\,(x_3)$$

where,

x_1 = rebound tenderness

x_2 = continued pain after 6 hrs in the ED

x_3 = physician's VAS estimate of ≥ 7 cm

If a patient were negative for all variables in the model (i.e. $x_1=0$, $x_2=0$, $x_3=0$) that patient would have a probability of severe obstruction of 0.9%.

$$P(X) = \frac{1}{1 + e^{-(-4.653)}} = 0.009 = 0.9\%$$

Alternatively, if a patient were positive for all variables in the model (i.e. $x_1=1$, $x_2=1$, $x_3=1$,) that patient would have a probability of severe obstruction of 96.8%.

$$P(X) = \frac{1}{1 + e^{-(3.412)}} = 0.968 = 96.8\% .$$

If we test all possible combinations of these 3 variables we obtain a series of probabilities for each combination as shown in Table 11.

Table 11
Probability of Severe Obstruction based on Presence or Absence of the 3 Predictor Variables from the Logistic Regression Model #2

# of variables	Rebound tenderness (x_1)	No pain relief after 6 hrs (x_2)	VAS MD >=7 cm (x_3)	P(x)	%	sensitivity
one	0	1	0	0.072	7	92
	0	0	1	0.118	12	62
	1	0	0	0.211	21	62
two	1	1	0	0.684	68	15
	0	1	1	0.519	52	15
	1	0	1	0.789	79	0
three	1	1	1	0.968	97	0

With any one variable there is a 7-21% probability of severe obstruction, and with any two variables there is a 52-79% probability of severe obstruction. As we can see from the logistic regression equation $-4.653 + 3.334 (x_1) + 2.092 (x_2) + 2.639 (x_3)$, the β of each variable ranges from 2.092 to 3.334.

5.3.5 Classification Performance – Model #2

If we use this logistic regression model as a preliminary decision rule we would correctly classify 89% of patients but we would miss 38% of patients with severe obstruction, based on the predicted probability threshold of severe obstruction of 0.5 (50%). If, however, we lower the probability threshold to 0.1 (10%), this will increase our sensitivity to 92% but will lower our specificity to 44%. As a result, the overall accuracy of the rule falls from 89% to 51%.

Classification performance of the regression model using predicted probabilities of severe obstruction of 0.1, 0.5 and 0.75 are shown in Table 12. As the probability threshold decreases the sensitivity increases. An ROC curve was generated for the logistic regression model by calculating the sensitivity and specificity for the various probabilities of severe obstruction from 0.01 to 0.99. (Figure 17). The shape of the curve shows a good discrimination between those with and without severe obstruction and the area under the curve is 0.82 (95%CI 0.69-0.96) ($p<0.001$). In Table 9 the classification performance of the physicians' predicted probabilities of severe obstruction with thresholds of 0.01, 0.5 and 0.75 are provided.

Figure 17
ROC Curve of the Classification Performance of the Logistic Regression Model #2 for Severe Obstuction

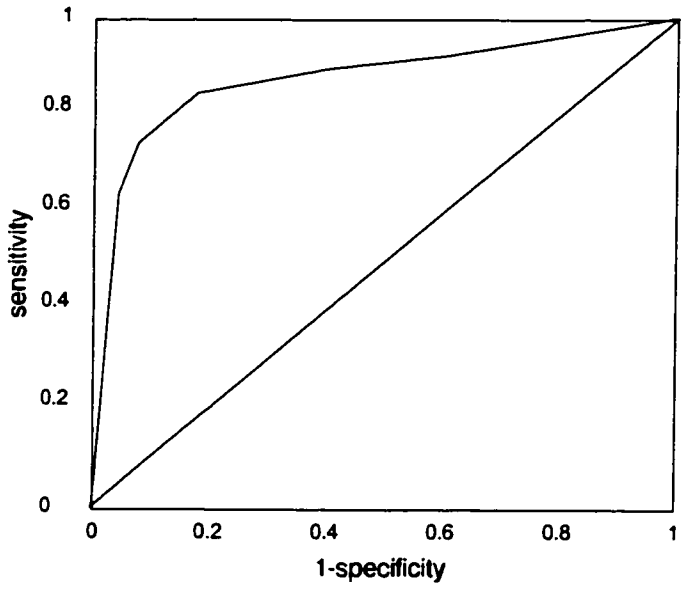


Table 12
Classification Performance of the Logistic Regression Model #2 Using Predicted Probabilities of Severe Obstruction of 0.1, 0.5 and 0.75

Predicted Probability of 0.1

	Actual Severe Obstruction YES	Actual Severe Obstruction NO
Predicted Severe Obstruction YES	12	43
Predicted Severe Obstruction NO	1	34

Sensitivity	92%
Specificity	44%
Accuracy	51%
Positive Predictive Value	22%
Negative Predictive Value	97%

Predicted Probability of 0.5

	Actual Severe Obstruction YES	Actual Severe Obstruction NO
Predicted Severe Obstruction YES	8	5
Predicted Severe Obstruction NO	5	72

Sensitivity	62%
Specificity	94%
Accuracy	89%
Positive Predictive Value	62%
Negative Predictive Value	94%

Predicted Probability of 0.75

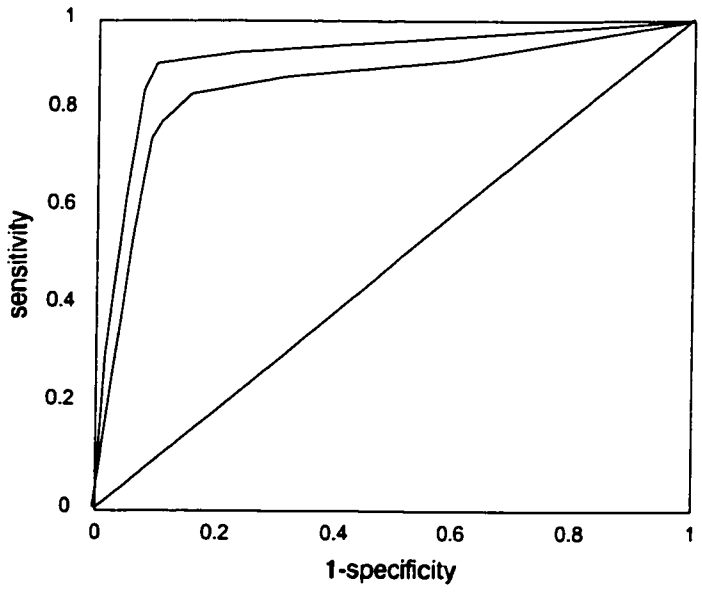
	Actual Severe Obstruction YES	Actual Severe Obstruction NO
Predicted Severe Obstruction YES	2	0
Predicted Severe Obstruction NO	11	77

Sensitivity	15%
Specificity	100%
Accuracy	88%
Positive Predictive Value	100%
Negative Predictive Value	88%

In Figure 15 we can see that the ROC curve for physician judgment yields an area of 0.64 (95%CI 0.51-0.77) (p=0.063). The area is much larger for the logistic regression Model #2 in comparison to physician judgment (p=0.014).

If we compare the performance of Model #1 versus Model #2 using the area under the ROC curve, the difference is not statistically significant (p=0.273). However, if we look at the curves in Figure 18 it becomes evident that in model #1 a high sensitivity (90%) is reached with a high specificity (85%); whereas in Model #2 a high sensitivity (92%) is reached at a much lower specificity (44%). Both models are more accurate than current physician judgment.

Figure 18
ROC Curve of the Classification Performance of the Logistic Regression Model #1 vs #2 for Severe Obstuction



5.4 Sample Size for the Definitive Study

5.4.1 Estimating Sample Size Using Flack and Eudey Method

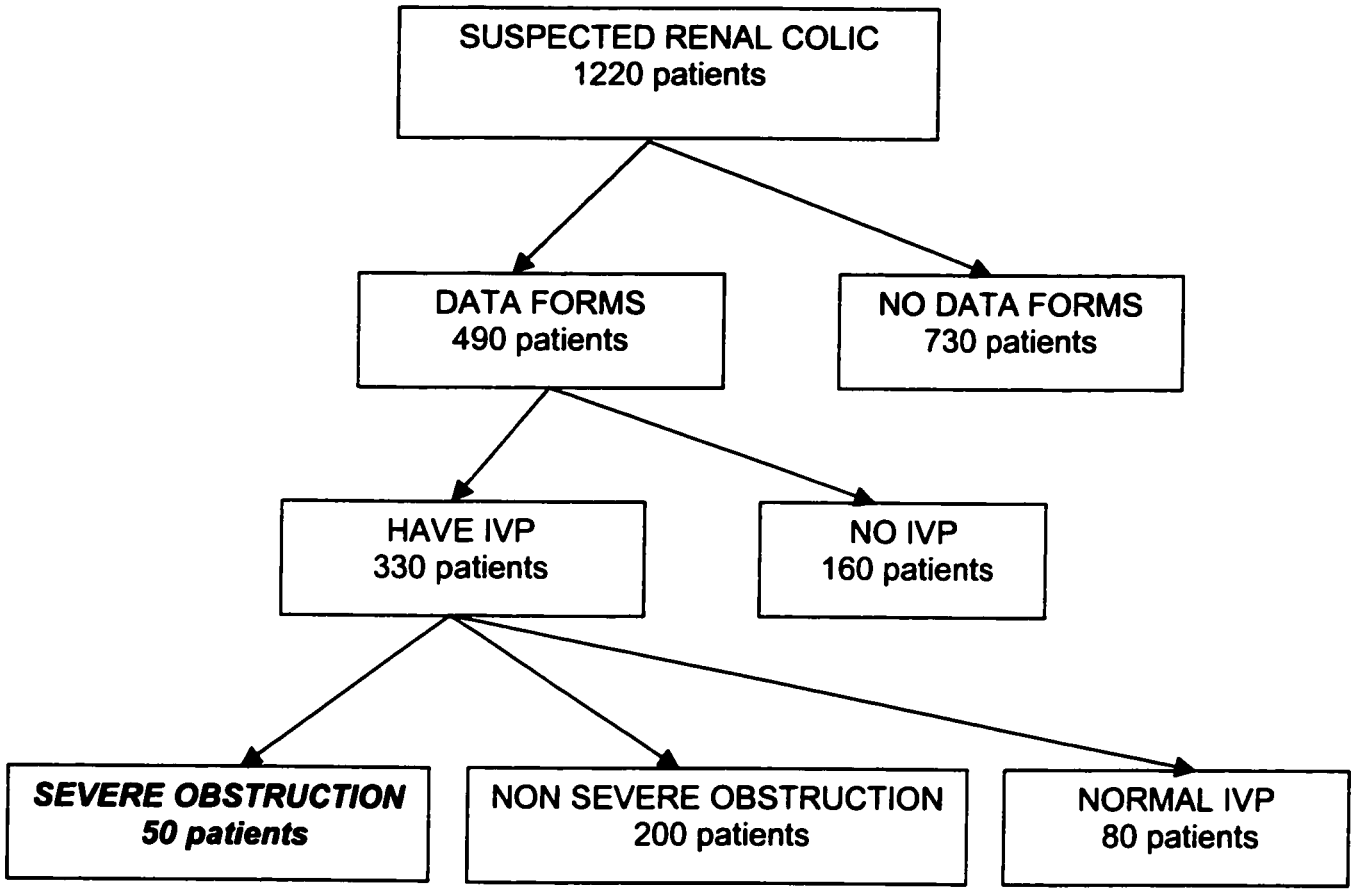
For the definitive study, sample size calculations will be based on the logistic regression data obtained from this feasibility study. Using the method by Flack et Eudey (58) we explored the bound of error around varying sample sizes for both logistic regression models, constructed graphs and then determined sample sizes for desired confidence intervals. The details of the sample size calculations using this method are given in Appendix 10.

In calculating the bounds of error for each model, we found that the total sample size (i.e. the number of patients having an IVP & completed data form) required for our chosen confidence intervals was the same for both models. A confidence interval width of 0.1 (95% CI) would be achieved with a sample size of 200 patients having data forms and IVP, yielding 30 cases of severe obstruction. For a confidence interval width of 0.05 (97.5%CI) we would require 700 patients with data forms and IVP, providing 106 cases of severe obstruction. For a confidence interval width of 0.01 (99%CI) we would require over ten thousand patients with a data form and IVP, yielding 1513 cases of severe obstruction.

Using our preliminary logistic regression models to estimate sample size, we have estimated that in order to achieve 95% confidence interval we would require 200 patients with IVP's and data forms. Assuming the same accrual rate and patient profile as the feasibility study for the definitive study we would need to enroll a total of 734 cases of suspected renal colic over one year to obtain these results. Figure 19 illustrates the expected flow of patients for the definitive study given these figures. In order to achieve a 97.5% confidence interval we

would require 700 patients with IVP's and data forms for a total of 2570 cases of suspected renal colic over 41 months (3.5 years). Based on a 99% confidence interval, a sample size of 10 000 patients with data forms and an IVP would require 36 700 cases of suspected renal colic over 588 months (49 years).

Figure 19
Expected Cohort of Suspected Renal Colic Patients For Definitive Decision Rule Study of Urgent use of IVP for 95% CI



5.4.2 Estimating Sample Size Using Precision of Sensitivity of Rule

In the definitive study we would like to develop a decision rule that will be highly sensitive with the highest possible specificity. By calculating the confidence interval around the sensitivity we can estimate the number of patients with severe obstruction required for a desired sensitivity of the decision rule. Using the formula in Fleiss et al. (60) for inferences about a single proportion we can calculate the 95% confidence interval with the following formula:

$$\text{CI width} = P_{upper} - P_{lower}$$

where,

$$P_{lower} = \frac{(2np + C_{\alpha/2}^2 - 1) - C_{\alpha/2}\sqrt{C_{\alpha/2}^2 - (2 + 1/n) + 4p(nq + 1)}}{2\,(n + C_{\alpha/2}^2)}$$

$$P_{upper} = \frac{(2np + C_{\alpha/2}^2 + 1) + C_{\alpha/2}\sqrt{C_{\alpha/2}^2 + (2 - 1/n) + 4p(nq - 1)}}{2\,(n + C_{\alpha/2}^2)}$$

n = sample size of severe obstruction cases

p = sensitivity of decision rule

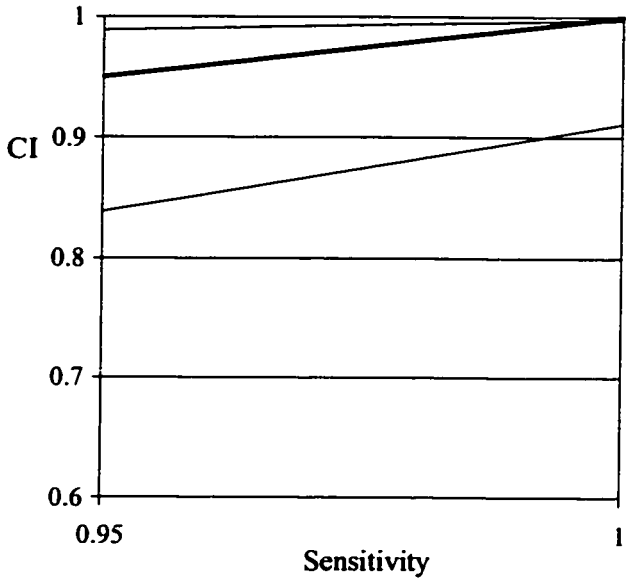
q = 1 – p

$C_{\alpha/2} = 1.96$

As a rule of thumb, we should have at least ten outcome events per independent variable in the prediction rule.(35;36) Currently we have 3-4 variables and should therefore start with at least 40 outcomes of severe obstruction. Sample sizes from 50 to 200 were used to calculate confidence intervals around given sensitivities.

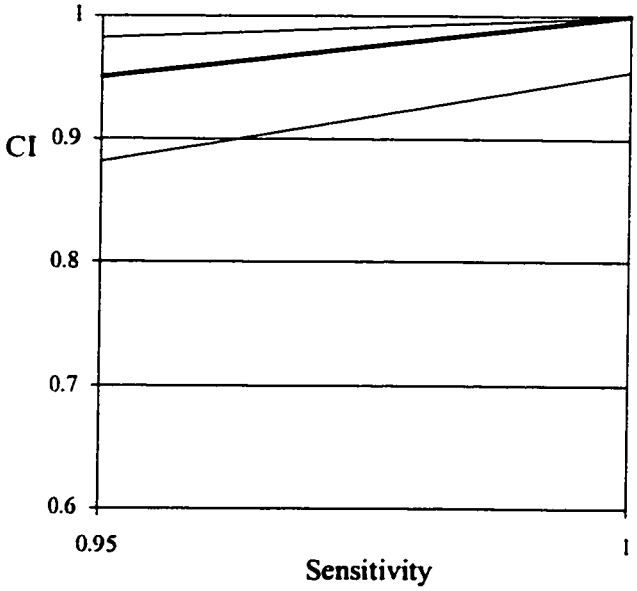
For a sensitivity of 100% and a confidence interval of 91% to 100%, 50 cases of severe obstruction would be required (Figure 20). In this case 330 patients with data forms and IVP's would be required for the definitive study.

Figure 20
Confidence Intervals for Sensitivity of 95% for a Sample Size of 50 cases of Severe Obstruction



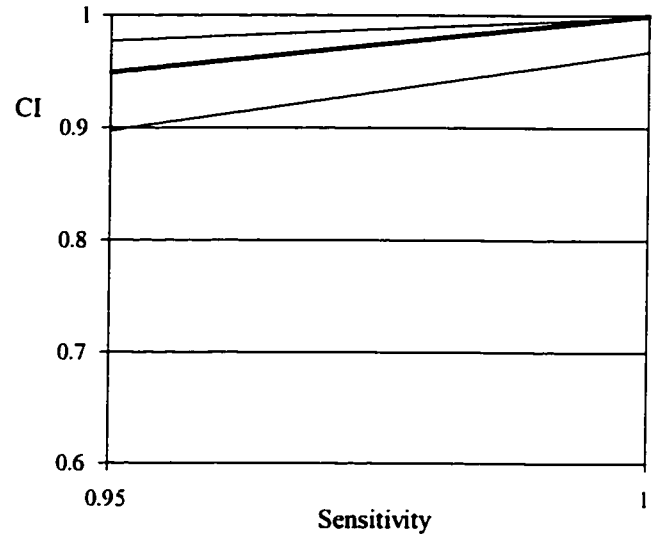
For a sensitivity of 100% and a confidence interval of 95% to 100%, 100 cases of severe obstruction would be needed (Figure 21). Therefore, 660 with data forms and IVP's would be necessary for this confidence interval.

Figure 21
Confidence Intervals for Sensitivity of 95% for a Sample Size of 100 cases of Severe Obstruction



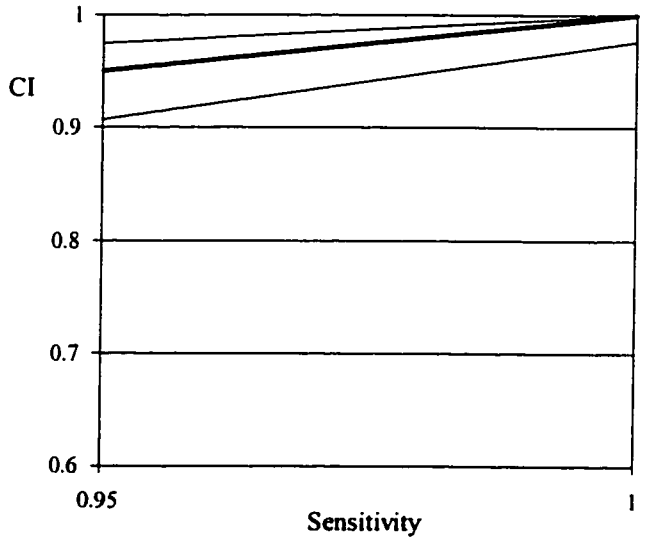
For a sensitivity of 100% and a confidence interval of 97% to 100%, 150 cases of severe obstruction would be required (Figure 22). Hence, 990 patients with data forms and IVP's would have to be enrolled for this confidence interval.

Figure 22
Confidence Intervals for a Sensitivity of 95% for a Sample Size of 150 cases of Severe Obstruction



And for a sensitivity of 100% and a confidence interval of 98% to 100% 200 cases of severe obstruction would be required (Figure 23). About 1300 patients with data forms and IVP's would have to be enrolled.

Figure 23
Confidence Intervals for a Sensitivity of 95% for a Sample Size of 200 cases of Severe Obstruction



5.4.3 Comparing Both Sample Size Methods

For a confidence interval width of 0.1 the Flack and Eudey method yielded a sample size of 30 cases of severe obstruction and the sensitivity method yielded 50 cases of severe obstruction. For a confidence interval width of 0.05 the Flack and Eudey method yielded 106 cases of severe obstruction versus 100 for the sensitivity method. Most physicians are comfortable with the statistical standard of 95%. Since there is no standard for detecting severe obstruction at this time it would be acceptable to emergency physicians to derive a decision rule in the definitive study with 95% confidence that the estimated parameter is within ±0.05 of the true value; i.e. a confidence interval of 0.1.

Therefore, we will use a 95% CI for determining the sample size for our definitive decision rule study. The second method of sample size determination yielded the more conservative sample size estimate of 50 cases of severe obstruction, whereas the Flack Eudey method yielded 30. Based on this estimate we will require 330 cases of renal colic having and IVP and data form completed. Assuming the accrual of patients will mimic that in this feasibility study, we estimate a total of 1220 suspected renal colic cases would be enrolled over a 20-month period.

6. RESULTS: PREDICTING COMPLICATIONS

6. 1 Descriptive Statistics

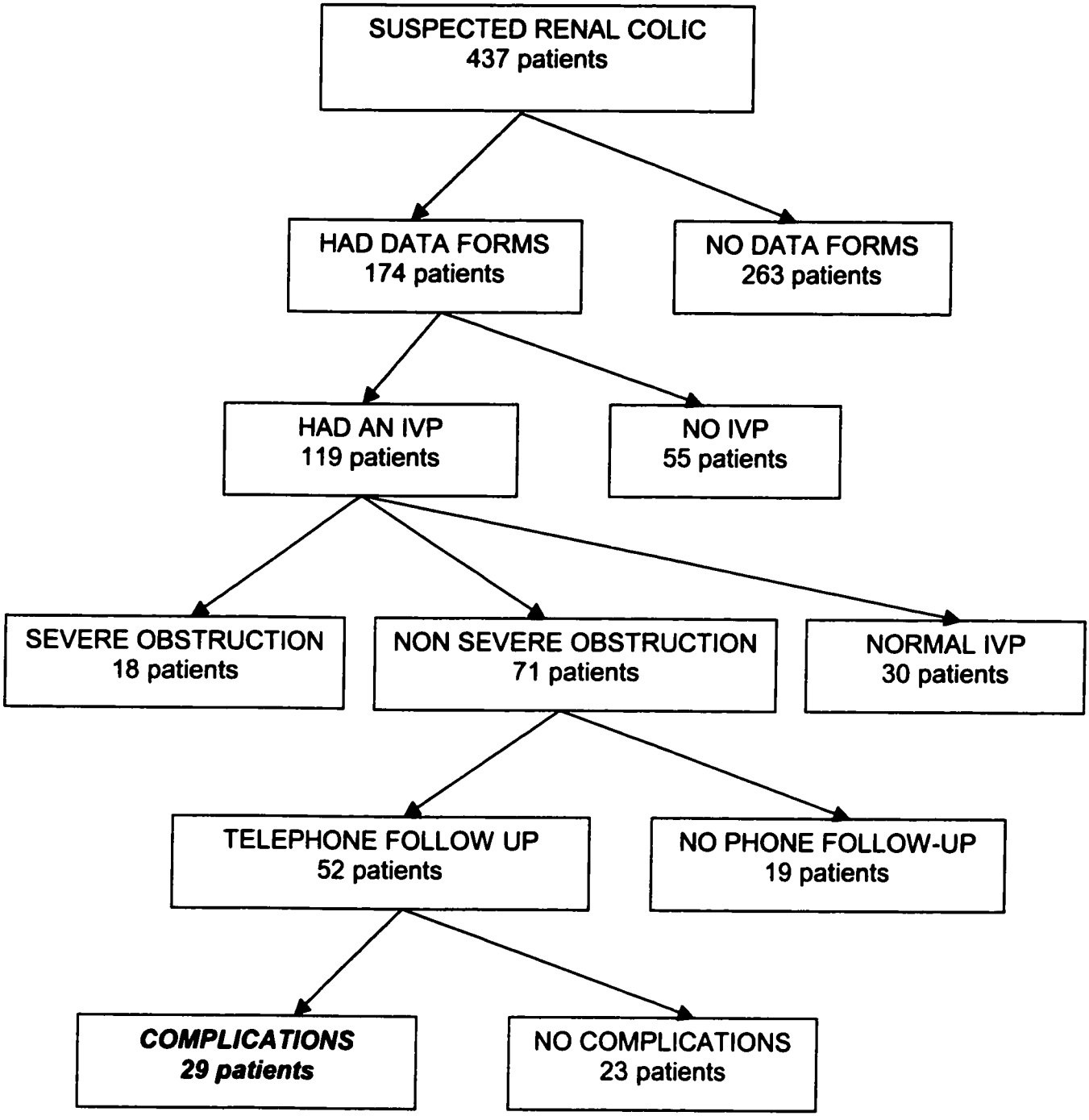
6.1.1 Patient Evaluation

Data was collected for this part of the study at the Ottawa Hospital from January 1999 to September 1999. We obtained our sample of renal colic patients for this part of the study from those we had evaluated in the emergency department in the first part of this feasibility study. Figure 24 is a flow diagram summarizing the cohort. All those patients who were discharged from the emergency department after having had a data form completed by the emergency physician and an IVP showing non severe urinary obstruction from calculi were followed prospectively at one and four weeks. Of the 71 patients who were eligible to participate in this part of the study only 52 were contacted for follow-up. Three patients refused to participate, 4 patients could not be reached at the number provided and 12 were not called within the four week follow-up period. We compared those with telephone follow-up with those lost to follow-up and found no important clinical difference no statistically important differences between the two groups in age (p=0.54), gender (p=0.38), admission rate (p=0.08), time spent in the emergency department (p=0.71), return visits (p=1.0) or previous history of renal colic (p=1.0). In addition, there were no statistical differences in any of the clinical parameters measured.

Of those patients who received a telephone follow-up, 4 patients (8%) were from the General Campus and 48 (92%) were followed from the Civic Campus. The mean age of patients followed at one and four weeks was 45 years old, 73% were male. About 50% had had a previous history of renal colic

before the current episode. Complications developed in 29 out of 52 patients (56%).

Figure 24
Cohort for Predicting Complications in the Feasibility Study



Persistent pain was the major complication (93%). Renal failure accounted for only 2% of complications and fever 8% of the total. Three percent (3%) of the 29 patients had a combination of persistent pain at four weeks with either renal failure or fever from a urinary source. Comparing the patient characteristics of those with and without complications in the demographics portion of Table 13, we noted no statistical differences in age, gender, admission rate and referral rate between the two groups. The return rate to the emergency department was higher (63%) in those patients having complications than those without (37%) complications.

6.1.2 Results of Follow-up

At one week follow-up 68% of patients were still experiencing pain from their calculi. Furthermore, 53% of those contacted at four weeks admitted to persistent pain. Table 14 describes the symptoms patients were experiencing from their renal colic at one and four weeks. This demonstrates the impact renal colic has on the activities of daily living of those suffering from it. Half of patients followed at four weeks after their initial visit to the emergency department were unable to return to a normal work routine and were having trouble eating and sleeping as a result of the pain from their calculi.

Table 13
Univariate Analysis Comparing 52 Patients With and Without Complications

Variables	*Complications (n=29)*	*No Complications (n=23)*	*P-Value**
Demographics			
Age (mean in years±SD)	47 (16.0)	44 (13.3)	0.52
Gender (male) %	76	70	0.76
Admitted %	21	4	0.12
Revisits (n=35) %	58	44	0.51
Referred to Urology %	90	83	0.69
Prior History of Renal Colic (n=51) %	43	57	0.41
Length of stay in ED (minutes)	410 (248)	300 (151)	0.07
Registration time to discharge time (n=49) (min±SD)			
Signs & Symptoms			
Temperature (n=50) (°C ±SD)	36.1 (0.6)	36.0 (0.8)	0.55
Vomiting %	38	43	0.78
Dysuria %	17	13	1.00
Frequency %	31	39	0.57
Abdominal Tenderness %	45	35	0.57
Rebound Tenderness %	7	0	0.50
CVA Tenderness %	52	43	0.59
Pain Measures			
Hours of continuous pain before presenting to ED (hrs±SD)	5.7 (9.3)	3.4 (1.6)	0.20
Days of pain prior to visit (days±SD)	1.8 (2.5)	2.2 (2.9)	0.61
Patient's VAS at arrival (n=51) (cm±SD)	6.8 (3.0)	7.2 (2.9)	0.56
Patient's VAS at discharge (cm±SD) (n=39)	1.3 (1.9)	0.1 (0.3)	0.008
Physician's VAS at patient's arrival (n=49) (cm±SD)	5.9 (2.7)	6.6 (2.8)	0.37
Time to patient becoming pain free (n=47) (min±SD)	195 (133)	178 (122)	0.66
Laboratory Results			
Hematuria present on urine dip (n=44) %	96	100	1.00
Degree of hematuria on urine dip (n=44)			0.62
Degree of pyuria on urine dip (n=44)			0.50
Degree of hematuria on urinalysis (n=11)			0.37
Hematuria present on urinanalysis (n=11)%	100	100	
Pyuria present on urine dip (n=44) %	20	11	0.68
Serum BUN (n=42) (μmol/L±SD)	6.4 (1.9)	5.7 (1.9)	0.26
Serum Creatinine (n=44) (μmol/L±SD)	92.7 (22.8)	89.1 (22.5)	0.60

Table 14
Symptoms of 52 Patients with Non Severe Urolithiasis at 1 and 4 Weeks after Discharge from the Emergency Department

Characteristics	*1 week*	*95%CI*	*4 weeks*	*95%CI*
Symptoms				
Still experiencing pain from their stone (%)	68	(54-82)	53	(39-67)
Mean pain score (rated 1-10 in cm)	4.5	(3.5-5.5)	4.5	(3.5-5.5)
Pain kept patient from working (%)	47	(26-65)	46	(27-66)
Pain kept patient from sleeping (%)	59	(41-77)	50	(30-70)
Pain kept patient from eating (%)	44	(26-62)	36	(17-55)
Patient experienced fever/chills (%)	31	(14-48)	29	(11-46)
Patient vomiting (%)	28	(12-45)	14	(0-28)
Patient having pain with urination (%)	50	(32-68)	61	(41-80)
Patient noticed blood in their urine (%)	34	(17-52)	39	(20-59)
Patient had side effects from the medications (%)	22	(0-33)	26	(0-46)

6.1.3 Description of Urologic Intervention

Data on interventions performed were obtained from the 52 patients followed by telephone. Eighteen 18 (35%) underwent a urologic procedure for their urinary calculi after leaving the emergency department. Extracorporeal shockwave lithotripsy was performed on 6 of the 18 patients (33%), ureteroscopy on 9 (50%) patients, percutaneous nephrostomy on 1 (6%) patient and open surgery on 2 (11%) patients.

6.2 Testing Clinical Predictors

6.2.1 Univariate Analysis

The variables from the data forms were tested for their strength of association with the development of complications over four weeks. Complications were analyzed as a dichotomous outcome (yes or no). Only those clinical variables obtained prospectively from the physician data forms were used in the univariate analysis. The results of the univariate analysis are given in Table 13. The only variable having a statistically significant ($p<0.05$) association with complications was the patient's VAS pain score at discharge from the emergency department. The pain score obtained from the patient at discharge was significantly higher in those patients who developed complications 1.3 (SD1.9) than those who did not 0.1 (SD0.3). Hence, those patients who left the emergency department having some residual pain were at greater risk of complications than those who left without any pain.

In order to assess additional parameters for our multivariate analysis, we decided to included those variables from the univariate analysis with a $p<0.20$. This allowed us to explore two additional variables: i) length of stay in the emergency department; and ii) number of hours patient experienced pain before coming to the emergency department. Patients with complications had a longer stay in the emergency department 410 minutes (SD250) than those without complications 300 minutes (SD150). In addition, complications were higher in those patients experiencing pain for a longer time before coming to the emergency 5.7 hours (SD9.3) than who had pain for a shorter time 3.4 hours (SD1.5).

In preparation for the logistic regression analysis we dichotomized the 3 continuous variables being analyzed: patient's discharge pain score; length of stay in the emergency department; and hours of pain before presenting to the emergency department. They were dichotomized at various cut points and their association tested for significance by chi-square analysis in Table 15.

Table 15
Cut points of Continuous Variables Dichotomized for Logistic Regression For Predicting Complications

Variable Cut points	*Pearson Chi-Square*	*Fisher's Exact Test (2-sided)*
Patient's VAS score at discharge (in cm)		
VAS = 1.0	2.7	0.139
VAS = 2.0	6.1	0.022
VAS = 3.0	3.8	0.110
VAS = 4.0	2.8	0.235
Number of hours of pain before coming to the ED (in hours)		
5	0.0	1.0
6	2.9	0.117
7	4.4	0.059
8	4.4	0.059
9	2.5	0.245
10	2.5	0.245
Length of stay in the ED (in minutes & hours)		
240 (4 hrs)	1.5	0.352
300 (5 hrs)	1.6	0.258
360 (6 hrs)	2.1	0.235
390 (6.5 hrs)	2.3	0.143
420 (7 hrs)	1.2	0.333
450 (7.5 hrs)	0.6	0.506

Therefore, we had 3 variables with dichotomous outcomes available for the logistic regression analysis: i) discharge VAS pain score greater than or equal to 2 cm; ii) stay in the emergency department greater than or equal to 6.5 hours,

and; iii) Continuous pain for greater than or equal to 7 hours before coming to the emergency department.

6.2.2 Multivariate Analysis using Logistic Regression

Logistic regression analysis used 38 of a possible 52 cases due to missing data. Discharge VAS pain scores were missing on 14 patients. Stepwise backward logistic regression analysis yielded a model consisting of one variable (Table 16).

Table 16
Variable in Logistic Regression Model For Predicting Complications

Variable	*β*	*SE*	*Wald statistic*	*p-value*	*Odds Ratio*	*95% C.I.*
Discharge VAS ≥ 2 cm	9.4	44.6	0.04	0.83	11909	0 - ∞
Discharge VAS ≥ 2cm 1 added to each cell	1.96	1.1	3.0	0.083	7.1	0.77 – 64.4

Estimate of the odds ratio was not stable with an extremely wide confidence interval. Our logistic regression model yielded an odds ratio of 110909 (95%CI 0 - ∞). (Table 17) The probability of developing complications given this predictor variable can be expressed with the following formula:

$$P(X) = \frac{1}{1 + e^{-(\alpha + \sum \beta i x i)}} = \alpha + \beta_1 x_1 = -0.182 + 9.385\,(x_1)$$

where,

x_1 = Discharge VAS ≥ 2 cm

Table 17
Classification Performance of the Logistic Regression Model Using Predicted Probabilities of Complications of 0.01, 0.5 and 0.9

Predicted Probability of 0.01

		Actual Complications	
		YES	NO
Predicted Complications			
	YES	20	18
	NO	0	0

Sensitivity	100%
Specificity	0%
Accuracy	53%
Positive Predictive Value	53%
Negative Predictive Value	0%

Predicted Probability of 0.5

		Actual Complications	
		YES	NO
Predicted Complications			
	YES	5	0
	NO	15	18

Sensitivity	25%
Specificity	100%
Accuracy	61%
Positive Predictive Value	100%
Negative Predictive Value	55%

Predicted Probability of 0.9

		Actual Complications	
		YES	NO
Predicted Complications			
	YES	5	0
	NO	15	18

Sensitivity	25%
Specificity	100%
Accuracy	61%
Positive Predictive Value	100%
Negative Predictive Value	55%

If a patient is negative for this variable and has a VAS pain score of less than 2 cm at discharge (i.e. x_1=0) that patient would have a probability of developing complications of 45%.

$$P(X) = \frac{1}{1 + e^{-(-0.182)}} = 0.454 = 45\%.$$

Alternatively, if a patient is positive for this variable and had a discharge VAS of greater than or equal to 2 cm at discharge (i.e. χ_1 =1) that patient would have a probability of developing complications of 99.9%.

$$P(X) = \frac{1}{1 + e^{-(9.203)}} = 0.999 = 99.9\% .$$

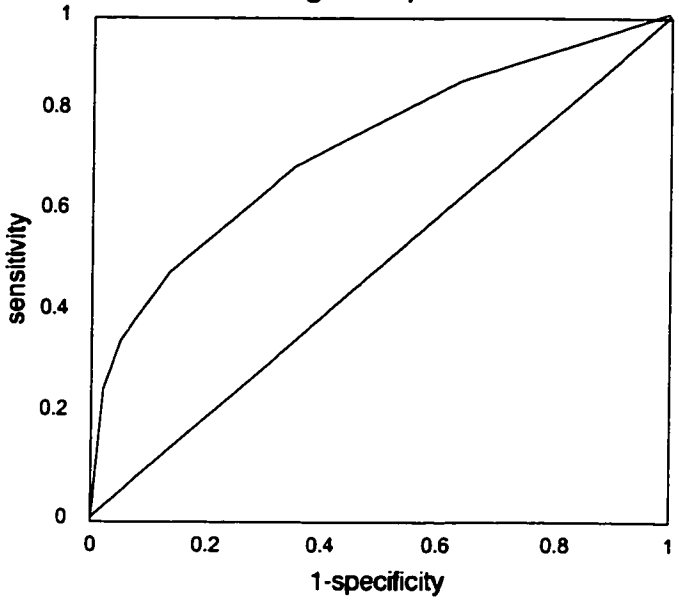
If we use this logistic regression model as a preliminary decision rule we would correctly classify 61% of patients and we would miss 75% of patients who would develop complications.

6.2.3 Classification Performance

Classification performance of the regression model using a low, moderate and high predicted probability of developing complications is shown in Table 17. The ROC curve for the various probabilities of developing complications from 0.01 to 0.99 for this model yielded an area of 0.64 (95%CI 0.47 – 0.82) with a p-value of 0.13. (Figure 25).

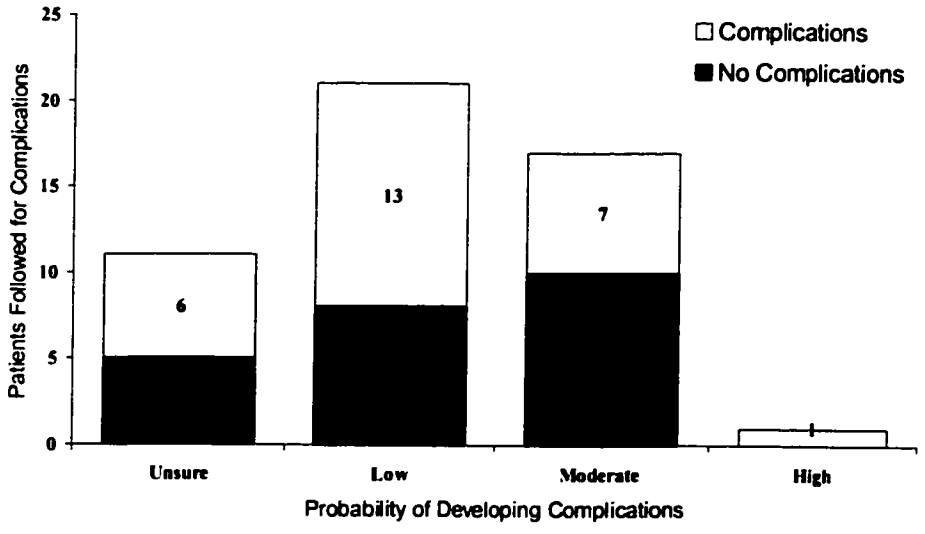
We tested emergency physician's ability to predict complications based on their clinical evaluations while the patients were in the emergency department. Physicians were unable to accurately predict which patients would develop complications after being discharged based on the patients' clinical presentation.

Figure 25
ROC Curve of the Classification Performance of the Logistic Regression Model for Predicting Complications



Complications developed frequently (62%) when physicians had a low suspicion and occurred infrequently (38%) when they had a moderate to high suspicion of the patient developing complications (Figure 26).

Figure 26
Physicians' Predicted Probability of Complications vs Development of Complications



Classification performance of the physicians' predicted probabilities of developing complications of 0.1, 0.5 and 0.9 are shown in Table 18.

Table 18
Classification Performance of the Physicians' Predicted Probability of Complications with Thresholds of 0.1, 0.5 and 0.9

Predicted Probability of 0.1

	Actual Complications	
	YES	**NO**
Predicted Complications		
YES	14	15
NO	13	8

Sensitivity	52%
Specificity	35%
Accuracy	44%
Positive Predictive Value	74%
Negative Predictive Value	38%

Predicted Probability of 0.5

	Actual Complications	
	YES	**NO**
Predicted Complications		
YES	8	10
NO	19	13

Sensitivity	30%
Specificity	57%
Accuracy	42%
Positive Predictive Value	44%
Negative Predictive Value	41%

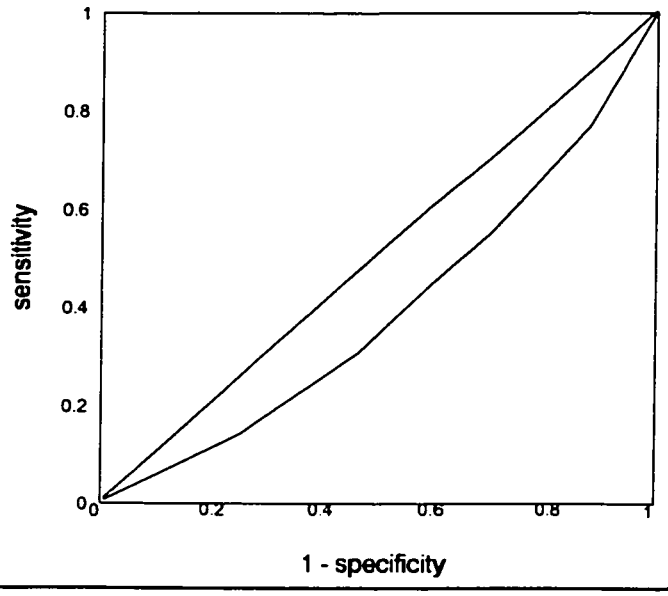
Predicted Probability of 0.9

	Actual Complications	
	YES	**NO**
Predicted Complications		
YES	1	0
NO	26	23

Sensitivity	4%
Specificity	100%
Accuracy	48%
Positive Predictive Value	100%
Negative Predictive Value	47%

Physicians would have missed 70% of complications if the threshold for suspicion had been their own predicted probability of at least 50%. With the threshold reduced to 10% probability of developing complications they would still have missed 13 patients with complications (48%). The area of the ROC curve for physicians' predicted probability, as a model for predicting complications, is 0.43 (95%CI 0.24 – 0.61) shown in Figure 27.

Figure 27
ROC Curve of theClassification Performance of Physicians' Predicted Probability of Complications



Given that the odds ratio was difficult to interpret in the original model and due to the presence of a 0 in one of the cells, we added 1 to each of the 4 cells and ran the logistic regression again. The analysis used 42 patients with this adjustment and our odds ratio became 7.12 (95% CI 0.77-65.48). The odds ratio still includes 1 and is not statistically significant but this yielded a more stable model. Due to the instability of the model we could not calculate a sample size for the definitive study.

6.2.4 Subgroup Analysis

Our definition of "complications" encompassed three distinct entities, persistent pain at four weeks, fever from a urinary source and renal failure into one definition. Of the 29 patients with complications 27 (93%) had persistent pain at four weeks. We did a subgroup analysis on the group of patients with persistent pain at four weeks to see which variables would be predictive of this outcome. We repeated the univariate analysis using persistent pain at four weeks as the outcome measure (Table 19) and found that only discharge VAS was significant.

Table 19
Univariate Analysis Comparing 51 Patients With and Without Persistent Pain at 4 Weeks Follow-up

Variables	*Pain at 4 weeks (n=27)*	*No Pain at 4 weeks (n=24)*	*P-Value**
Demographics			
Age (mean in years±SD)	48 (14.0)	42 (15.8)	0.16
Gender (male) %	74	71	1.00
Admitted %	15	8	0.67
Revisits (n=34) %	56	44	0.73
Referred to Urology %	89	83	0.69
Prior History of Renal Colic (n=51) %	42	54	0.57
Length of stay in ED	400 (248)	315 (165)	0.17
Registration time to discharge time (n=49) (min±SD)			
Signs & Symptoms			
Temperature (n=49) (°C ±SD)	36.0 (0.6)	36.0 (0.8)	0.95
Vomiting %	37	46	0.58
Dysuria %	19	13	0.71
Frequency %	33	38	0.78
Abdominal Tenderness %	48	33	0.39
Rebound Tenderness %	7	0	0.49
CVA Tenderness %	56	42	0.40
Pain Measures			
Hours of continuous pain before presenting to ED (hrs±SD)	5.9 (9.6)	3.4 (1.5)	0.19
Days of pain prior to visit (days±SD)	1.8 (2.5)	2.2 (2.9)	0.72
Patient's VAS at arrival (n=50) (cm±SD)	6.7 (3.1)	7.3 (2.9)	0.49
Patient's VAS at discharge (cm±SD) (n=39)	1.3 (1.9)	0.1 (0.3)	0.008
Physician's VAS at patient's arrival (n=48) (cm±SD)	5.7 (2.7)	6.7 (2.8)	0.24
Time to patient becoming pain free (n=46) (min±SD)	177 (110)	181 (120)	0.91
Laboratory Results			
Hematuria present on urine dip (n=43) %	96	100	1.00
Degree of hematuria on urine dip (n=43)			0.67
Degree of pyuria on urine dip (n=43)			0.39
Degree of hematuria on urinalysis (n=11)			0.24
Hematuria present on urinanalysis (n=11)%	100	100	
Pyuria present on urine dip (n=43) %	22	10	0.68
Serum BUN (n=41) (μmol/L±SD)	6.4 (1.9)	5.7 (1.9)	0.26
Serum Creatinine (n=43) (μmol/L±SD)	92.7 (22.8)	89.1 (22.5)	0.60

By including variables with a $p<0.20$ we were able to recruit 3 other variables: i) length of stay in the emergency department, ii) number of hours patient experienced pain before coming to the emergency department, and iii) age. The only other variable different from the previous model was age. We performed a logistic regression using appropriate cut points for continuous variables (Table 20).

Table 20
Cut points of Continuous Variables Dichotomized for Logistic Regression For Persistent Pain at 4 Weeks

Variable Cut points	*Pearson Chi-Square*	*Fisher's Exact Test (2-sided)*
Patient's VAS score at discharge (in cm)		
VAS = 1.0	2.7	0.139
VAS = 2.0	6.1	0.022
VAS = 3.0	3.8	0.110
VAS = 4.0	2.8	0.235
Number of hours of pain before coming to the ED (in hours)		
5	0.2	0.749
6	3.5	0.103
7	4.9	0.052
8	4.9	0.052
9	2.0	0.364
Length of stay in the ED (in minutes & hours)		
240 (4 hrs)	1.1	0.364
300 (5 hrs)	1.0	0.396
360 (6 hrs)	1.2	0.377
390 (6.5 hrs)	1.3	0.364
420 (7 hrs)	0.3	0.742
Age (years)		
35	0.8	0.485
40	0.4	0.575
45	2.7	0.154
50	0.4	0.767

There was no sensible cut point for length of stay, so it was not considered for the model. Using i) discharge VAS ≥ 2, ii) age ≥ 45 years, and iii) pain ≥ 7 hours prior to presentation in the logistic regression, we obtained a final model (Table 21). Neither discharge VAS ≥ 2 nor age ≥ 45years had a significant odds ratio.

Table 21
Variable in Logistic Regression Model For Persistent Pain at 4 Weeks

Variable	*β*	*SE*	*Wald statistic*	*p-value*	*Odds Ratio*	*95% C.I.*
Discharge VAS ≥ 2 cm	9.3	42.8	0.05	0.83	10996	0 - ∞
Age ≥ 45 years	1.4	0.7	3.3	0.068	3.9	0.91 – 16.8

7. DISCUSSION

7.1 Urgent Imaging

7.1.1 Rationale

There is a need for a clinical guideline to help physicians decide whether immediate imaging in a patient with suspected renal colic is necessary. Patient care, accuracy, time, and cost are the factors driving this issue. Through this feasibility study we have identified that there are a significant number of patients that develop severe urinary obstruction from urolithiasis. Physician accuracy in identifying them is poor and therefore the risk of long-term renal dysfunction is significant. As emergency departments become overcrowded, expediting patients' stays with validated guidelines, becomes very important. Length of stay in our study was much shorter in those not having a radiological test performed. By reducing the current indiscriminant use of radiography with a clinical decision rule could significantly impact the health care budget and standardize care.

Use of IVP for suspected renal colic was quite high in the two participating Ottawa Hospital emergency departments. Seventy five percent (75%) of suspected renal colic patients had an IVP and only 14% had evidence of severe obstruction. The estimated annual cost for IVP's ordered by the emergency departments of the Civic and General sites of the Ottawa Hospital alone is about $76,000. Figures obtained from the Ontario Ministry of Health show that the technical and professional costs for outpatient IVP's is in excess of 3.2 million dollars per year. The economic impact on the Canadian health system is unknown. The annual cost of outpatient evaluation of urolithiasis in the United

States is about $278 million dollars based on figures from the Agency for Health Care Policy and Research.(17).

For decades IVP has been considered the gold standard for diagnosing and following patients with urolithiasis. However, there is a North American trend toward adopting spiral CT as the primary imaging modality for patients with renal colic because of its rapidity (5 minutes) and negative risk of allergic reactions to contrast. As a diagnostic tool, CT has shown a slightly higher sensitivity and specificity than IVP for detecting urolithiasis. Despite this, IVP still has the advantage of delineating the urinary tract and determining renal function. Spiral CT is unable to do this unless enhanced with contrast. Since it was our primary objective to determine severity of obstruction, IVP was the preferred imaging modality for this outcome.

Despite renal colic being one of the most common conditions seen in the emergency department, physician judgment to image these patients is inconsistent and inaccurate. (18;25;61) We found that in the majority of cases physicians estimated that the probability of severe obstruction was below 10%. Physicians estimated that there was a greater than 60% probability of severe obstruction in only 6% of cases. The area under the physicians' ROC curve for detecting severe obstruction was only 0.64. This inconsistency is likely due to a lack of dependable, prospectively derived clinical guidelines, hence the rationale for this study. Our study is the first in North America to date to prospectively define and describe criteria to do this. It has been suggested that less time could be spent in the emergency department by eliminating radiography. Although a few studies have suggested that imaging in renal colic patients could be delayed or eliminated, none have addressed the issue of how.(11;12)

In our study, length of stay in the emergency department for those patients undergoing an IVP was significantly higher than those without. This is a result of the time it takes to undergo an IVP. Factors such as waiting time to be called to the radiology department for the procedure, transport time, actual procedure time and time for interpretation are the culprits for this delay. Similarly, other studies have suggested that imaging selectively would reduce length of stay. (2;10;12;13) If we could be selective in our use of imaging in suspected renal colic with a decision rule, we could reduce length of stay by 2.5 hours in those who don't require imaging.

7.1.2 Patient Enrollment

Unlike many of the retrospective renal colic studies published to date, our data was gathered prospectively in a manner that reflects actual physician practice.

The patient characteristics of the renal colic cases in this study are comparable to those described in the emergency medicine and urologic literature.(10;12;19;25) They are young, predominantly male, often have a prior history of renal colic and present more frequently in the spring and summer months. We observed that only 14% of confirmed renal colic cases had severe obstruction. Those patients with severe obstruction are more likely to have very high pain scores and to have pain that is more difficult to control with intravenous narcotics.

Overall, data form completion rates were poor with 48% at the Civic site and half that (22%) at the General site. Despite both sites having similar volumes of patients, the Civic site enrolled the majority of the patients. This may

have been a result of Civic physicians being more accustomed to such studies and, therefore, more familiar with completing data forms. Form completion was dependent on emergency physicians and busy shifts make form completion difficult. In addition, there are select physicians who prefer not to participate in such studies. Regardless, the physicians' IVP ordering rate was similar at both the Civic (73%) and the General (77%) sites. Patients were recruited more during the early summer months due to the seasonal rise in urolithiasis during the spring and summer.(52;62)

In comparing Figure 7 to Figure 9, the number of severe obstruction cases fell from a potential 45 cases to only 18 cases for analysis due to incomplete data forms. In an effort to examine selection bias in those with data forms we compared the patient characteristics in those with and without data forms and found no statistically significant differences between the two. It was noted that physicians were more likely to complete data forms in male patients and we assumed this to be a result of the higher prevalence of the disease in males rather than gender bias. In addition, IVP's were ordered more frequently in patients with data forms probably as a result of the physician's higher clinical suspicion of urolithiasis in the patients for which they completed a data form. Since this was a prospective cohort with no intervention it is interesting that in Table 1 all four groups of patients consistently show the same demographic characteristics and this reflects a good sample.

We found that although most cases of confirmed renal colic had microscopic hematuria, 5-10% did not. This finding is consistent with previous studies, which quote negative microscopic hematuria to be from 5 to 17%.(1;2;13)

7.1.3 Evaluation of Data Collection

Identifying data collection pitfalls is an integral part of a feasibility study. We have identified a few areas that will need improvement for the definitive study that will minimize missing data. Incomplete data forms was the culprit in 60% of lost cases and missing data points were responsible for another 40% lost for analysis. By reducing incomplete data forms and ensuring that individual variables are completed could increase our sample dramatically.

Data collection involves the clerks, nurses and emergency physicians. In order to improve data form completion all three groups need to be targeted. Strategies such as training sessions for staff, simplifying the data collection forms, introducing incentives to reward those who fill in data forms, and providing frequent friendly reminders to those who leave forms incomplete. Ideally, a research nurse should be available 24 hours a day to identify potential patients and ensure form completion. Gathering forms for data entry and ensuring that missing variables are tracked when possible should be the responsibility of a diligent research assistant. Trends in missing data should be identified early and corrected promptly.

Another problem with data completion was the lack of recording of patients' progress while in the department, including pain control and discharge pain scales. This can be explained by the fact that physicians are very busy in the department and often lack the time to record these items during a hectic shift. The presence of a research nurse in the department that could follow up on these variables could alleviate the burden from the physicians in this matter.

7.1.4 Consensus Definition of Severe Obstruction

One of the standards for clinical decision rules is that the outcome measure must be clearly defined and clinically relevant. Severe obstruction is a very important diagnosis to make because it carries a risk of complications if not treated. Moreover, we have defined it through surveys and via the expert opinions from three different medical and surgical specialties.

There are three issues surrounding our definition of severe obstruction that should be addressed: the choice of severe obstruction as an outcome measure, the accuracy of the definition chosen and the generalizability of the outcome.

Firstly, why use severe obstruction as an outcome measure? To date, there are no studies clearly defining severe obstruction and using it as a sole outcome. The focus of emergency research in patients with flank pain has been to distinguish the ominous abdominal aortic aneurysm from urolithiasis, detracting from research on the consequences of obstruction. Abdominal aortic aneurysms are rare but quickly fatal if undetected. The morbidity of severe obstruction from urinary calculi is less dramatic but includes an important spectrum of renal disease from hypertension to renal failure. Animal studies have clearly shown the morbidity associated with ureteral obstruction.(4;5;49) and human studies as well. (6;63) In our study 85% of suspected renal colic patients had a serum creatinine measured because of concern over deterioration in renal function from calculi. With the current inconsistency in ordering imaging for these patients and the inaccuracy in detecting severe obstruction demonstrated by this study and others (10;11;13) we felt it was important to

develop a guideline for physicians. In summary, severe obstruction is an objective measure on pyelogram that is transportable and clinically important.

Secondly, how accurate is a definition based on surveys and a panel discussion? Our definition is actually a refinement of the current definition used in the radiological literature.(1;8;12;13) Our survey encompassed 3 specialty groups of physicians closely involved in the care of renal colic patients to verify their practice patterns. This is the most comprehensive method for clarifying this definition described in the renal colic literature to date.(8;12;13) Agreement was high among the specialists and the IVP findings selected reflect current North American practice. We wanted to develop a "clinically meaningful" definition, not just a radiological one that would identify those patients who needed urgent consultation with a urologist. Consequently, we held an expert panel discussion to determine which findings on IVP would warrant urgent urologic follow-up.

Thirdly, how generalizable is our expert panel definition to places outside a tertiary care setting? Although our sample is taken from a tertiary care milieu, we feel it is reflective of current North American practice for the following reasons. The definition is compatible with that found in the current literature. Many of the physicians surveyed also practice in community hospitals in Ottawa and have a perspective into community-based medicine. Many of the physicians surveyed did their training outside of Ottawa and are familiar with practice outside the Ottawa region. The panel discussion allowed us to refine an already existing definition to a more clinically relevant one without much compromise.

7.1.5 Testing Clinical Predictors

This study has prospectively tested clinical variables to predict obstruction in urolithiasis. The variables selected reflect frequently cited variables in the renal colic literature. What was novel about this study was the extensive use of validated pain scales as variables to predict outcome. The theory that degree of pain correlates well with severity of a disease is substantiated in this study.

All clinical predictors were gathered prospectively. Pain scores appeared to correlate well with severity of obstruction in our study. A clinically significant change in the VAS score was measured in a study by Todd (64) and was defined as a difference in VAS of 13 mm. The mean difference in VAS scores between the severe and not severely obstructed groups was 20 mm, hence VAS scores were statistically and clinically significant. These VAS measures were the scores obtained from both patients and physicians at the time of initial assessment. Therefore, patients with more severe disease had higher pain scores. Vomiting also correlated well with severity of disease occurring almost twice as often in the severe obstruction group.

Since our study lacked sufficient sample size to detect all possible associated variables at a $p<0.05$ we chose to include in the multivariate analysis all variables with a $p<0.15$ on univariate analysis. This allowed us to recruit more variables. We found that it took longer to control the pain in those patients with severe obstruction than those without severe obstruction. The reasoning behind this is speculative but it was observed that pain was stronger and recurred more frequently in those with significant obstruction.

Rebound tenderness was another finding associated with severe obstruction. This parameter was not tested for interobserver reliability in our

study and therefore we do not know how consistent this finding would be among different physicians. In addition, there were so few cases where rebound tenderness was actually found on exam that we lack the power to make a firm conclusion about this predictor.

Another significant pain score was the discharge VAS. Those patients with severe obstruction left the department in more pain than those without severe obstruction with a mean difference of 14 mm. Patients without severe obstruction were more likely to leave completely pain free at discharge.

It is impractical for emergency physicians to calculate probabilities for continuous variables. We dichotomized our variables in order to simplify its clinical use choosing cut points for each continuous variable that were most significantly associated with severe obstruction. The cut points also made clinical sense. This method has been used in high caliber decision rule studies.(34)

7.1.6 Logistic Regression Modeling

The purpose of the feasibility study was to assess clinical predictors and practice modeling. A commonly used "rule of thumb" for sample size adequacy is that there should be at least ten outcome events per independent variable in the prediction rule.(35;48) Since we only had 18 outcomes of interest and 6 potential variables in our feasibility study, our sample did not have enough power for a definitive rule. We were, however, able to use logistic regression analysis to help calculate the sample size for the definitive study.

Of the potential 119 cases included in our model, only 72 were used for modeling because of missing data. Recording of time to pain relief and

discharge pain scores were the two variables found missing most frequently. We suspect this occurred as a result of shift changes and the hectic nature of the emergency department, which made it difficult to record follow-up notes on these patients.

The first model performed quite well. The presence of any one variable could detect severe obstruction with a 90% sensitivity. Having all variables absent would exclude severe obstruction with a specificity of almost 100%. When interpreting the Hosmer-Lemeshow goodness-of-fit statistic large values of chi-square indicate a poor fit and small values indicate a good fit. The P-value is interpreted as a good fit if it is near 1 and a poor fit if it is near 0. Our first model yielded a chi-square of 7.4 and p-value of 0.283, suggesting a good fit of the model. Since the sample size was inadequate for a definitive rule the Hosmser-Lemeshow is only speculative at this time. Nonetheless, the clinical model was significantly superior to physician judgement (p=0.007).

The second model had 18 more cases available for analysis at the expense of dropping a significant variable from the equation. It also performed better than physician judgment. The Hosmer-Lemeshow goodness-of-fit yielded a chi-square of 2.4 with a p-value 0.79. This actually yielded a better fit than the first model.

7.1.7 Sample Size Determination

Determining an adequate sample size is essential to developing an accurate decision rule. In order to calculate a robust sample size for our decision rule study we felt that testing two distinct methods of sample size calculation would be allow us to do this with confidence. The Flack & Eudey

Method used the information on covariates from the logistic regression models. Generally, this method results in smaller sample sizes and smaller prediction standard errors than other approaches. The "precision around the sensitivity" method has been used in highly referenced decision rule studies.(34;59) We found the two methods complimented each other as they both yielded very similar results given a specified confidence interval. We selected a 95% confidence interval for our sample size because it was desirable and feasible. This yielded a sample of 50 cases of severe obstruction. Since there is no standard for detecting severe obstruction at this time it would be acceptable to emergency physicians to derive a decision rule in the definitive study with 95% confidence that the estimated parameter is within ±0.05 of the true value; i.e. a confidence interval of 0.1. In addition, most physicians are comfortable with the statistical standard of 95%. To ensure physician comfort with this they will be surveyed prior to the definitive study.

If we assume the accrual to be exactly the same as the feasibility study, it would take 20 months to enroll 50 cases of severe obstruction. However, if we eliminate incomplete data forms, we could complete the study in 8 months! This shows us how important it is to have good and complete enrollment.

7.1.8 Methodological Issues

Although this is a preliminary decision rule study it is important to adhere to the methodological standards for decision rules outlined in section 2.3 that are applicable. We will review these principles and indicate the strengths and weaknesses of our methodology.

The outcome, clinically significant severe obstruction, was clearly defined based on objective findings on an IVP. The clinical importance of severe obstruction lies in the spectrum of renal disease that could develop if it is not detected. An investigator blinded to the predictor variables assessed outcome. Physicians prospectively identified clinical predictor variables prior to knowledge of the outcome in a standardized manner using data forms. Physicians were blinded to the outcome during their assessments.

Interobserver agreement for the physical exam variables was not measured during this study, as this was only a preliminary assessment of variables for a future decision rule. The only predictor variable in the model requiring assessment of interobserver reliability was rebound tenderness. VAS pain scores and vomiting are objective findings that are reported by patients and not by the physician assessor. The reproducibility of VAS pain scores has been shown in other studies.(53)

The generalizability of the study findings is dependent on the sample taken. Our sample of patients had similar characteristics to renal colic patients in other studies and was typical of suspected renal colic patients commonly seen in emergency departments in North America. Data on these patients was collected 24 hours per day with the cooperation of all the emergency physicians of the Ottawa Hospitals. However, many eligible patients were overlooked because of incomplete data forms. This study is the first prospective sample of suspected renal colic patients to have clinical variables tested against the outcome of severe obstruction.

The statistical analyses were clearly identified and described for both univariate and multivariate tests. Using the well-described mathematical

technique of logistic regression, we built a preliminary model that was considerably better than current physician judgment. Recursive partitioning was not considered for the feasibility study because it was only a preliminary model but will be used in the building of the definitive model.

The study attempted to build preliminary models that were simple for clinicians to apply. Predictor variables were dichotomized and categorized in a sensible manner so they could be used at bedside without the need for any tabulation.

We estimated the number of potential outcomes prior to the feasibility study via a chart review. Since no statistics were available to calculate a sample size we performed this feasibility study to estimate sample sizes for a definitive study. Our study accomplished its goal but was inadequate for a definitive rule.

The preliminary models yielded sensitivities of 90% to 92% for predicting severe obstruction. Both models did considerably better than current physician judgment. Prior to undertaking the definitive study, we will survey physicians to determine an acceptable error rate for missing severe obstruction. The higher the sensitivity, the better the rule will be at detecting severe obstruction; however, it will be at the expense of specificity.

The next phase of this study will be to develop a definitive decision rule using adequate sample sizes and to further validate it.

7.1.9 Limitations of the Study

The purpose of a feasibility study is to recognize limitations so as to correct them for the definitive study. Selection of patients was based on the judgment of the physicians using acute onset of severe flank pain within 24

hours of presentation as guiding criteria. If we would have specified hematuria we may have missed up to 15% of cases. The reliability of these criteria was based on a retrospective study conducted to diagnose urolithiasis using clinical markers. Flank pain, acute onset of pain, severe pain and hematuria were found to be the highest ranking clinical predictors of urolithiasis with a 90% probability of calculi. Most emergency and urology textbooks use flank pain and hematuria as criteria to identify potential renal colic patients.(19;20) Based on these clinical symptoms eligible patients with atypical presentations may have been missed and ineligible patients may have been included. However, we believe that this selection bias will not threaten the validity of the study as it represents actual practice.

There is potential for assessment bias in this feasibility study. Although predictor variables were assessed independently of outcome in an a priori fashion not all patients with suspected renal colic had an IVP performed. There is a possibility that cases of severe obstruction were missed in the 111 patients who did not undergo an IVP. This was justifiable by the fact that to expose 100% of patients to IVP would not have been acceptable to the emergency physicians or the ethics committee. These patients were not followed after discharge to assess their outcome, but this is a consideration for the next study.

Ensuring data form completion was difficult because we were dependent on the emergency physicians for their evaluation of patients. Patients lost to incomplete data forms and missing variables reduced our sample size and reduced the statistical power of our analysis. Although data forms were left incomplete on eligible patients, the characteristics of patients with and without forms were comparable and we believe did not invalidate our results. Notably,

the cooperation of emergency physicians was voluntary and essential to the fulfillment of the study.

There was initial concern that using the diagnostic criteria, hematuria, as a predictor variable would create a selection bias. Our purpose, however, was to examine the association of the "degree" of hematuria with the outcome of severe obstruction and not the presence of hematuria alone. The precedent for this was set in another study testing the association of degree of hematuria and obstruction in renal colic.(13)

Another criticism of the study is that severe obstruction patients were never followed-up after their disposition from the emergency department. Although we clearly recognize the risk of renal damage from severe obstruction, we never addressed how these patients were treated or followed. In the definitive study, a follow-up telephone call would help to address this issue. In order to assess the development of long-term complications such as hypertension, one would need a long-term cohort study to evaluate these patients over several years. This is not a feasible option for us at this time.

We chose to compare our severe obstruction group to a non-severe group including those with a normal IVP. We chose to do this to follow the actual practice situation that would apply when using the rule. When suspected renal patients are evaluated in the emergency department, physicians will have to decide who has severe obstruction on a group of patients whose pyelogram is unknown. Regardless of whether the pyelogram shows non-severe obstruction or is normal, the decision rule will identify those who need pyelography and patients with normal IVP's do not need imaging. So we felt that the better comparison group should include all normal IVP's, as well as those with non-

severe obstruction. In addition, some urinary calculi are very tiny and are not visualized, while others pass before imaging is performed.

7.2 Predicting Complications

7.2.1 Rationale

Clinically, in the emergency department, it is difficult to predict who will develop complications or who will have difficulty passing their urinary calculi once they leave the hospital. By identifying in the emergency department those patients who will develop complications after discharge, patients can be appropriately referred to a urologist for intervention early and avoid the unnecessary pain and suffering associated with delayed passage of their calculus as well as time lost from work.

There has not been a prospective study evaluating the clinical predictors of complications in renal colic. Retrospective studies have, however, shown degree of pain, persistence of pain, size of calculus, location along the ureter, and signs of infection are important factors in deciding to intervene.(40;42-45) Our goal was to identify clinical symptoms and signs that would allow emergency physicians and urologists to identify these patients early and consider intervention to expedite passage of the calculus.

Using "need for intervention" as an outcome measure is too subjective and not generalizable because current treatment practices are extremely variable among urologists.(14) Fever from infected urinary calculi and renal failure, are very important consequences of urolithiasis and require prompt treatment with an intervention. In addition, continued pain at four weeks was selected as an outcome measure because it is an objective measure that is transportable and applicable to any study population. It is significant because studies show that most patients who fail conservative treatment and require intervention have continued pain 30 days after initial presentation. (37;44)

Furthermore, complications from an intervention for ureteral obstruction are much higher when symptoms exceed four weeks than when interventions are performed within four weeks of onset. (40)

A considerable number (56%) of patients developed complications by our definition. Persistent pain was the most prevalent complication (90%), renal failure and fever together accounted for only 10% of the total complications. Most of these patients still had difficulty returning to normal daily living at four weeks. The economic impact of days off work for these patients is considerable. Indirect costs for lost wages from urolithiasis are estimated to be $139 million dollars annually in the US.(17) These figures are based on statistics from the Agency for Health Care Policy and Research.

7.2.2 Patient Evaluation

The patient demographics for this part of the study were almost identical to the first part. The follow-up rate of 73% for this study was adequate. Most of the patients were recruited from the Civic site of the Ottawa hospital because of a higher rate of data form completion. There was a 16% follow-up loss simply as a result of clerical error in forgetting to make the telephone calls at four weeks. These patients were given an initial call at one week and the four-week call was not made. This much clerical error was unacceptable and could easily be improved with better study coordination. Patient cooperation was very good with only 4% refusing to participate and 7% being inaccessible by telephone. The impact of the patients lost to follow-up on the results of the study were likely minimal since the demographics of the two groups were comparable.

7.2.3 Follow-up and Intervention

Prospective studies following patients after emergency department discharge are nonexistent. This study was unique in that it prospectively described these outcomes in renal colic patients. Quality of life issues are often not taken into consideration when studying renal colic and this study demonstrates how significant it actually is. Future studies need to address this issue.

Very few patients developed fever from infected calculi and renal failure in our sample of patients with non-severe obstruction after discharge. However, a significant number of patients continued to have pain after four weeks of onset of symptoms. The pain impaired activities of daily living such as work, sleep and appetite in at least half its victims. Treating these patients in a timely fashion by early identification becomes important for improving quality of life and reducing lost wages.

The majority of patients who had an intervention underwent ureteroscopy and extracorporeal shock wave lithotripsy. Lithotripsy is usually the first line treatment of most urinary calculi as per the 1997 recommendations of a urology task force on urolithiasis. (14) In our center ureteroscopy was performed more frequently because our center lacked a lithotripsy machine and patients who needed the procedure had to drive two hours for treatment. The more invasive procedures such as percutaneous nephrostomy and open surgery were performed least frequently and on the patients in whom non-invasive therapy had failed.

7.2.4 Univariate Analysis

There are a few studies that have addressed pain as correlating with complications of urolithiasis in our review of the literature but none have used a validated pain scale to do so. This is one of the first studies to have prospectively tested clinical variables to predict complications in urolithiasis.

Patients with complications had higher pain scores at discharge than those without. Due to the inadequacy of our sample size we recruited additional variables to test in our model by raising the inclusion p-value and found that patients having pain for a longer time prior to visiting the emergency department were more likely to have pain at four weeks post discharge. This has been shown to be true in two previous studies, one retrospective study (44) as well as one prospective study from Ottawa (37). One possible explanation can be found in the research on pain. Current pain literature suggests that patients with prolonged exposure to pain become more sensitive to painful stimuli. This "hyperalgesia" occurs when nerve fibers not normally associated with pain are recruited due to prolonged irritation. (65) In other words, the longer time you have pain for the more difficult it is to control it. In addition, patients with complications had longer stays in the emergency department suggesting they required more time for observation, investigation and possibly pain control.

7.2.5 Multivariate Analysis

A sample of 38 patients is inadequate to draw any conclusions from the multivariate analysis. Since our goal was to practice modeling we proceded with the logistic regression to test the parameters from the univariate analysis.

Given that the odds ratio was difficult to interpret in the original model and due to the presence of a 0 in one of the cells, we added 1 to each of the 4 cells and ran the logistic regression again. The analysis used 42 patients with this adjustment and our odds ratio became 7.12 (95% CI 0.77-65.48). The odds ratio still includes 1 and is not statistically significant. Although the logistic regression model for predicting complications appears to be more accurate than physician judgment it is still mediocre for predicting complications.

Possible explanations for these results include: i) that the sample size was inadequate and that our study lacked the power to determine a meaningful outcome, ii) that the variables used to describe the outcome of complications were simply not predictive of this outcome, or iii) that the outcome itself was not correctly defined and not a good end point.

Of the three explanations above we suspect that the small sample size is the main reason for only one variable entering the model and that a larger sample size would allow more variables to enter the model.

The second explanation is also important because the variables traditionally used to predict need for intervention have included calculus size and calculus location. Since we did not include these variables in our study because our aim was to use only clinical criteria to predict complications we may have excluded some valuable parameters from the model.

Our definition of complications encompassed three distinct entities: persistent pain at four weeks, fever from infected calculi and renal failure. Other studies have used "need for intervention" as an outcome measure but we felt that this outcome was too subjective and would suffer from practice variation bias. Complications prompting urologic intervention include prolonged pain,

calculi 6mm thick or larger, presence of complete obstruction, failure of the calculus to progress (impacted calculus), associated urosepsis, impairment of renal function, associated renal pathology, patient occupation, and patient preference. (25) The three most objective and clinically relevant outcomes were selected for our definition of complications. Persistent pain comprised the majority of our outcome and may have caused us to underestimate the importance of other variables that would have predicted renal failure of development of fever if analyzed separately. Since the number of patients who developed fever and renal failure were so small, we could not perform a subgroup analysis on them. When we explored persistent pain separately in a subgroup analysis we found no significant correlation with any of the variables. This issue will need to be addressed further in the definitive study by perhaps only predicting one of the three entities instead of combining them or doing a subgroup analysis on a larger sample of patients.

7.2.6 Methodological Issues

Although this is a preliminary decision rule study it is important to adhere to the methodological standards for decision rules outlined in section 2.3 that are applicable. We will review these principles and indicate the strengths and weaknesses of our methodology.

The outcome, complications, was clearly defined but was a compilation of three discrete complications. Each individual entity is clinically important and affects the clinical course of urinary calculi but combining them may have underestimated the variables that would have predicted fever from an infected calculus or renal failure. This issue will need to be addressed further in the

definitive study by perhaps only predicting one of the three entities instead of combining them. An investigator blinded to the predictor variables assessed outcome. Physicians assessed clinical predictor variables prior to knowledge of the outcome. Predictor variables were assessed in a standardized manner using data forms.

Interobserver agreement for the physical exam variables was not measured during this study, as this was only a preliminary assessment of potential variables for the definitive rule study. Since the generalizability of study findings is dependent on the sample, we feel our sample of patients reflected suspected renal colic patients commonly seen in emergency departments in North America because they had similar characteristics to renal colic patients in other published studies. Data on these patients was collected 24 hours per day with the cooperation of all the emergency physicians of the Ottawa Hospitals. However, many eligible patients were overlooked because of incomplete data forms and losses to follow-up. This study is the first prospective sample of suspected renal colic patients to have clinical variables tested against the outcome of complications.

The statistical analyses were clearly described for both univariate and multivariate tests. The predictor variables from the univariate analysis were dichotomized and categorized in a sensible manner so they could be used easily at the bedside without the need for any tabulation. Using the well-described mathematical technique of logistic regression, we built a preliminary model that was better than current physician judgment but that was statistically unstable. Sample size for the feasibility study was based on a 1997 study of complications of renal colic performed at the Ottawa Hospital two years earlier.(37) It was

difficult to predict a sample size for the feasibility study because our outcome measure had never been studied in this manner before. The next phase of this study will be to increase our sample size and develop a decision rule.

7.2.7 Limitations of the Study

The purpose of a feasibility study is to recognize limitations so as to correct them for the definitive study.

There is potential for assessment bias in this feasibility study. Although predictor variables were assessed independently of outcome in an a priori fashion not all patients with suspected renal colic had an IVP performed. There is a possibility that cases of non-severe obstruction were missed in the patients who did not undergo an IVP. This was justifiable by the fact that to expose 100% of patients to IVP would not have been acceptable to the emergency physicians or the ethics committee.

Ensuring data form completion was difficult because we were dependent on the emergency physicians for their evaluation of patients. Patients lost to incomplete data forms reduced our sample size and limited our analysis.

Clerical error due to the mismanagement of the four-week follow-up telephone call also reduced our sample size.

A small sample size limited our exploration of the data. Since the number of patients who developed fever and renal failure were so small, we could not perform a subgroup analysis on them and we may have underestimated the importance of predictor variables for renal failure or fever. This issue will need to be addressed further in the definitive study by perhaps only predicting persistent pain at four weeks and doing a subgroup analysis on a larger sample of patients.

Although this feasibility study lacked sufficient power to detect a significant association between the variables and the outcome, we feel that this study should be pursued. Using a larger sample size and adding stone dimension and location to the list of potential variables to be explored we are confident that predictive criteria could be derived. Data from new CT imaging techniques could make this very interesting.

7.3 Future Studies

7.3.1 Definitive Study for Urgent Imaging

The objective of this next phase will be to prospectively derive a clinical decision rule on a new set of patients. We will gather information for the rule on 50 cases of severe obstruction. Assuming the accrual of patients will mimic the feasibility study, we estimate that 1220 suspected emergency staff physicians would assess renal colic cases in the same manner as the first study.

The outcome of severe obstruction will be determined by a qualified radiologist and by an urologist. This evaluation will proceed without knowledge of clinical variables.

Ensuring data form completion will be very important in the definitive study. If we will continue to depend on the emergency physician evaluation of patients we will improve data form completion by simplifying the data collection forms, introducing incentives to reward those who fill in data forms, and providing frequent friendly reminders to those who leave forms incomplete. If we have a research assistant available to work in the emergency department data form completion will be facilitated by identifying eligible patients and presenting the data form to physicians directly for completion. Gathering forms for data entry and ensuring that missing variables are tracked will be the responsibility of a diligent research assistant. Trends in missing data will be identified early and corrected promptly such as discharge pain scales and time to pain relief in order to minimize losses for analysis.

The clinical decision rule will be derived using logistic regression analysis and will be assessed by calculating the sensitivity, specificity, positive predictive value and negative predictive value for detecting which patients have severe

obstruction. All variables assessed in the feasibility study will be retained for the definitive study. Interobserver agreement will be performed on the physical exam variables such as abdominal tenderness, rebound and CVA tenderness on a subset of patients and their kappa coefficients will be calculated. We will ask physicians their comfort level with eliminating or delaying radiography in patients with suspected renal colic. In addition, we will ask physicians about their comfort in having a clinical decision rule that has a 95% confidence for detecting severe obstruction.

For the definitive study we would like to follow our severe obstruction patients after discharge to observe their outcome.

The increasing use of helical CT scan in the evaluation of patients with suspected renal colic suggests that future studies should address this issue. If we could perform both CT scan and IVP on patients and compare the predictive validity for both imaging modalities the results would be more generalizable and current.

7.3.2 Definitive Study for Complications

The objective of this next phase will be to prospectively derive a clinical decision rule on a new set of patients. We will increase our sample size and add more variables from to the list of potential predictors. Methods of enrollment will mimic that of the feasibility study.

In the definitive study we will add variables from the pyelogram such as calculus size and location to the list of predictor variables. Although we attempted to predict complications without the need of a radiograph we may

have excluded these two valuable parameters from the model shown to be predictive in other studies.(40;42)

Persistent pain comprised the majority of our outcome and may have caused us to underestimate the importance of other variables that would have predicted renal failure or development of fever if analyzed separately. Since the number of patients who developed fever and renal failure were so small, we could not perform a subgroup analysis on them. We will consider predicting persistent pain at four weeks as a main outcome measure and doing a subgroup analysis on those who develop renal failure or infected calculi.

A rigorous tracking system for following patients by telephone will have to be implemented in the definitive study to avoid clerical error in follow-up telephone call also reduced our sample size. Ensuring data form completion will be performed using the methods described in the previous section 7.3.1.

Although this feasibility study lacked sufficient power to detect a significant association between the variables and complications, we feel that this study should be pursued further. By using a larger sample size, using persistent pain at four weeks as an outcome and adding calculus dimension and location to the list of potential variables to be explored, we are confident that predictive criteria could be derived.

8. CONCLUSION

8.1 Urgent Imaging

Presently, there are no guidelines to assist emergency physicians in deciding to order immediate imaging. A reliable and sensitive decision rule would permit physicians to be more accurate in their assessment of renal colic and more discriminating in ordering radiological tests. We are suggesting that IVP be done selectively in those cases having potential for severe obstruction.

This feasibility study has accomplished a number of goals. This study transformed the nebulous radiological definition of severe urinary obstruction into a clinically significant one. The proportion of severe obstruction was determined. We have shown the variability among emergency physicians in imaging renal colic patients in the emergency department and their inability to predict degree of obstruction. Also, the study demonstrated that selective imaging could reduce length of stay by 2.5 hours in those who don't require imaging.

As per the methods for developing clinical decision rules we have clearly defined the outcome of severe obstruction. Clinical predictor variables were prospectively identified and collected prior to knowledge of the outcome. Using the well-described multivariate mathematical technique of logistic regression, we practiced modeling and built a preliminary model that was considerably better than current physician judgment. The clinical predictors were sensible and easy to apply to clinical practice. The sample of patients had similar characteristics to renal colic patients in other studies.

The sample size estimated for the definitive study using our preliminary models was a feasible number and could be obtained in a reasonable timeframe.

This study determined that it was indeed feasible to proceed to a definitive study to predict the need for urgent imaging in renal colic patients in the emergency department.

With current costs of 6.77 million dollars for outpatient IVP's in Ontario and $278 million dollars in the USA, the potential for cost savings is tremendous.

Although there are no prospective outcome studies to date on humans with severe obstruction we hope to inspire others to do so. A reliable and sensitive decision rule would permit physicians to be more accurate in their assessment of renal colic, save time spent in the emergency department and save health care dollars.

8.2 Predicting Complications

This part of the feasibility study allowed us to look at the proportion of complications and to explore its relationship to clinical variables. Through these results we were able to determine the need for more variables and gain insight into the characteristics of patients with complications, the proportion of urologic procedures and the natural course of urolithiasis after discharge. We have shown that physicians cannot predict who will develop complications or who will have difficulty passing their urinary calculi once they leave the emergency department. We do know that a significant number of patients have their daily lives impaired by the pain of this disease. Currently, urologists show tremendous variability in intervening on patients with urinary calculi. Indirect costs for lost wages from urolithiasis are estimated to be $139 million dollars annually in the US. By identifying, in the emergency department, those patients who will develop complications after discharge, patients can be appropriately referred to a urologist for intervention early and avoid the unnecessary pain associated with delayed passage of their calculus, avoid functional renal impairment from obstructive calculi and reduce lost wages.

9. REFERENCES

(1) Press MP. Incidence of negative hematuria in patients with acute urinary lithiasis presenting to the emergency room with flank pain. Urology 45[5], 753-757. 2000.

(2) Elton TJ. A clinical prediction rule for the diagnosis of ureteral calculi in the ED. J Gen Intern Med 8, 57-62. 1993.

(3) Stewart C. Nephrolithiasis. Emerg Med Clin North Am 6[3], 617-630. 1988.

(4) Rose JG. Pathophysiology of ureteral obstruction. Am J Physiol 225[4], 830-837. 1973.

(5) Rose JG. Effects of obstruction on ureteral function. Urology 12[2], 139-145. 1978.

(6) Holm-Nielsen A. The prognostic value of probe renography in ureteric stone obstruction. Br J Urol 53, 504-507. 1981.

(7) Stecker JF. Experimental partial ureteral obstruction. I. Alteration in renal function. Invest Urol 8[4], 377-385. 1971.

(8) Chen MYM. Radiologic findings in acute urinary tract obstruction. J Emerg Med 15[3], 339-343. 1997.

(9) Rhea JT. Evaluation of a sequence of diagnostic tests using the workup of ureteral stone as a model. Medical Care 20[8], 843-848. 1982.

(10) Wrenn K. Emergency IVP in the setting of possible renal colic, is it indicated? Ann Emerg Med 26[3], 304-307. 1995.

(11) Richards JR. Intravenous urography in the emergency departement: When do we need it. Eur J Emerg Med 6, 129-133. 1999.

(12) Tasso SR. Effectiveness of selective use of IVP in patients presenting to the emergency departememt with ureteral colic. Acad Emerg Med 4[8], 780-784. 1997.

(13) Stewart DP. Microscopic hematuria and calculus related ureteral obstruction. J Emerg Med 8, 693-695. 1990.

(14) Segura JW. Ureteral stones clinical guidelines panel summary report on the management of ureteral calculi. J Urol 158, 1951-1921. 1997.

(15) Honey JD. Current management of ureteral stones. Can J CME 12[5], 89-102. 2000.

(16) Pak CYC. Kidney stones. Lancet 351, 1797-1801. 1998.

(17) Clark JY. Economic impact of urolithiasis in the United States. J Urol 154, 2020-2024. 1995.

(18) Chia SJ. Ureteric Colic: Value of initial investigations and the outcome. Ann Acad Med Singapore 24, 366-369. 1995.

(19) Walsh PC. Campbell's Urology , 7th ed. 1998. WB Saunders Company.

(20) Hardwood-Nuss A. Emergency Medicine Concepts and Clinical Practice. Stone Disease. Chapter 96[3rd ed.]. 1994. Mosby Year Book.

(21) Smith RC. A modern approach to diagnosis and management. Seminars in Ultrasound 20[2], 108-135. 1999.

(22) Miller OF. Prospective comparison of unenhanced spiral computed tomography and intravenous urogram in the evaluation of acute flank pain. Urology 52, 982-987. 1998.
Ref Type: Generic

(23) Smith RC. Noncontrast helical CT in the evaluation of acute flank pain. Abdominal Imaging 23, 10-16. 1998.

(24) Smith RC. Diagnosis of acute flank pain: Value of unenhanced helical CT. AJR 166, 97-101. 1996.

(25) Haddad MC. Renal Colic: diagnosis and outcome. Radiology 84, 83-88. 1992.

(26) Dalrymple NC. The value of unenhanced helical CT in the management of acute flank pain. J Urol 159, 735-740. 1998.

(27) Katz DS. Unenhanced helical CT of ureteral stones: incidence of associated urinary tract findings. AJR 166, 1319-1322. 1996.

(28) Sheley RC. Helical CT in the evaluation of renal colic. Am J Emerg Med 17[3], 279-282. 1999.

(29) Sourtzis S. Radiologic investigation of renal colic: unenhanced helical CT compared with excretory urography. AJR 172, 1491-1494. 1999.

(30) Older RA. Stone Disease. Urol Clin North Am 27[2], 215-229. 2000.

(31) Spencer BA. Helical CT and ureteral colic. Urol Clin North Am 27[2], 231-241. 2000.

(32) Preminger GM. Urolithiasis detection and management with unenhanced spiral CT---- A Urologic Perspective. Radiology 207, 308-309. 2000.

(33) Van Arsdalen K. Radiographic imaging and urologic decision making in the managemant of renal and ureteral calculi. Urol Clin North Am 17[1], 17-90. 1990.

(34) Stiell IG. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. Ann Emerg Med 21, 384-390. 1992.

(35) Wasson JH. Clinical prediction rules: applications and methodologic standards. NEJM 313, 793-799. 1985.

(36) Laupacis A. Clinical Prediction rules: a review and suggested modifications of methodologic standards. JAMA 277[6], 488-494. 1997.

(37) Mechaalni N. The use of pain scores and other characteristics as predictors for the success of conservative treatment of urinary calculi. Canadian Association of Urology Meeting 1998 Abstract. 1998.

(38) Mutgi A. Utility of Pain Abdominal Roentgenogram. Arch Intern Med 151, 1589-1592. 1991.

(39) Loberant N. Emergency imaging of the urinary tract. Emerg Med Clin North Am 10[1], 59-91. 1992.

(40) Hubner WA. Natural history and current concepts for the treatment of small ureteric calculi. Eur Urol 24, 172-176. 1993.

(41) LeRoy AJ. Diagnosis and treatment of nephrolithiasis: Current perspectives. AJR 163, 1309-1313. 1994.

(42) Jewett M. Comparative costs of the various strategies of urinary stone disease management. Urology 46[3A], 15-22. 1993.

(43) Ohkawa M. Spontaneous passage of upper urinary tract calculi in relation to composition. Urol Int 50, 153-158. 1993.

(44) Ibrahim AIA. Prognostic factors in the conservative treatment of ureteric stones. Br J Urol 67, 358-361. 1991.

(45) Ueno A. Relations of spontaneous passage of ureteral calculi to size. Urology 10[6], 544-546. 1997.

(46) Morse RM. Ureteral calculi: natural history and treatment in an era of advanced technology. J Urol 145, 263-265. 1991.

(47) Morris SB. Should all patients with ureteric colic be admitted. Ann R Coll Surg Engl 77, 450-452. 1995.

(48) Stiell IG. Methodologic standards for the development of clinical decision rules in emergency medicine. Ann Emerg Med 33[4], 437-447. 1999.

(49) Katul MJ. Evaluatiion of renal function during experimental hydronephrosis dy means of radioscope renogram. Surg Gynecol Obst 126, 563-571. 1968.

(50) Collie DA. The diagnostic yield of IVU: a demographic study. Br J Urol 73, 603-606. 1994.

(51) McDonald MW. Urinary stone disease:a practical guide to metabolic evaluation. Geriatrics 52, 38-56. 1997.

(52) Prince CL. A statistical analysis of ureteral calculi. J Urol 83[5], 561-565. 1960.

(53) McCormack HM. Clinical applications of visual analogue scales: a critical review. Psych Med 18, 1007-1019. 1988.

(54) Brennan DF. Reliability of infrared tympanic thermometry in the detection of rectal fever in children. Ann Emerg Med 25[1]. 1995.

(55) Jaffe DM. What's hot and what's not: the gold standard for thermometry in emergency medicine. Ann Emerg Med 25[1], 97-99. 1995.

(56) Pearle MS. Optimal method of urgent decompression of the collecting system for obstruction and infection due to ureteral calculi. J Urol 160, 1260-1264. 1998.

(57) Hanley JA. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143[1], 29-36. 1982.

(58) Flack VF, Eudey TL. Sample size determinations using logistic regression with pilot data. Statistics in Medicine 12, 1079-1084. 1993.

(59) Stiell IG et al. The Canadian CT Head Rule for patients with minor head injury. Lancet 357, 1391-1396. 2001.

(60) Fleiss JL. Statistical methods for rates and proportions. chapter 1[2nd edition]. 1982. New York, Wiley & Sons.

(61) Simon J. Conservative management of ureteral stones. Acta Urologica Belgica 65[2], 7-9. 1997.

(62) Soucie JM. Relation between geographic variability in kidney stones prevalence and risk factors for stones. Am J Epidemiol 143, 487-495. 1996.

(63) Stecker JF. Alterations in renal metabolism occurring in ureteral obstruction in vivo. Surg Gynecol Obst 133[5], 846-848. 1971.

(64) Todd KH. The minimum clinically important difference in physician-assigned visual analog pain scores. Acad Emerg Med 3, 142-146. 1996.

(65) Ducharme J. Acute pain and pain control: state of the art. Ann Emerg Med 35[6], 592-603. 2000.

Appendix 1

ABSTRACT FOR DEFINITIVE STUDY

BACKGROUND: About 5-15% of the North American and European populations pass a urinary calculus in their lifetime.(1-3) Currently, emergency physicians are ordering pyelograms based on inaccurate clinical judgment.(9;10) A carefully derived set of guidelines could eliminate or delay pyelography in most patients, improve and standardize patient care, save health care dollars and reduce time spent in the emergency department.(2;9;10;12;13) In addition, a number of patients develop complications from their urinary calculi after discharge. Continued pain after discharge affects patients' ability to function and return to work. Currently, there are no clear guidelines to decide which patients need a procedure early to remove urinary calculi to prevent suffering, renal impairment and lost wages.(14;46)
OBJECTIVES: 1) To predict which patients with suspected renal colic have severe obstruction and require urgent urinary tract imaging. 2) To prospectively assess clinical predictors of complications from urinary calculi after discharge from the ED.
METHODS: We will conduct a prospective cohort study of suspected renal colic cases presenting to the hospital emergency department (ED). Patients will have a 20-variable data form completed by ED physicians and an IVP performed within 24 hours. All IVP's will be reviewed by a radiologist and a urologist to identify those cases with severe ureteral obstruction (urine extravasation and/or no visualization of contrast beyond the obstruction after 2 hours). Those with non-severe obstruction will be discharged and followed via telephone at 1 and 4 weeks to assess development of complications. Complications will be defined by persistent pain by 4 weeks. Subgroup analysis will be performed on those who i) develop of fever (>38°C) from a urinary infection; and ii) develop an elevation of serum creatinine > 150 mmol/L by 4 weeks. Categorical data will be analyzed using Fisher's Exact test, continuous variables assessed by independent sample 2-tailed t-test, and ordinal variables tested with Mann-Whitney U test. Those variables found to be associated with the outcome measure of severe obstruction or complications will be combined using logistic regression analysis.
SAMPLE SIZE: Based on the logistic regression model of our feasibility study we will require 50 cases of severe obstruction for a 91%-100% confidence interval around a desired sensitivity of 100% for the decision rule.
CONCLUSIONS: This study will derive a clinical decision guideline based to help physicians decide the suspected renal colic patients who need urgent imaging in the emergency department. It will improve the detection of severe obstruction from urinary calculi by standardizing care of suspected renal colic patients; it will potentially save health care dollars and reduce time spent in the emergency department. This study will also derive criteria to identify which renal colic patients will have continued pain at 4 weeks after discharge so that early referral and intervention can be instituted. This will prevent prolonged patient disability and reduce time lost from work. Once derived, the decision rules will need to be validated in subsequent studies.

Appendix 2

SUBJECT NUMBER:_____
UNIQUE/HOSPITAL #:_____

RADIOLOGIST'S REPORT OF THE INTRAVENOUS PYELOGRAM

Date of ED visit: _____
Date of IVP: _____

1	Is it a normal IVP?	YES O NO O
2	Is there evidence of a stone?	YES O NO O
3	Is the stone visible?	YES O NO O nm O
4	Size of stone	_____mm
5	What side is it on?	Rt O Lt O both O
6	Where is it located?	KidO UPJO midO UVJO
7	Is there hydronephrosis?	YES O NO O nm O
8	What is the degree of hydronephrosis?	mild □ mod □ severe □ nm□
9	What is the degree of obstruction?	mild □ mod □ severe □ nm□
10	Is there extravasation?	YES O NO O nm O
11	Is there delayed visualization of contrast in the ureter > 2hrs?	YES O NO O nm O
12	Is there a prolonged dense nephrogram?	YES O NO O nm O
13	How long is the nephrogram present?	_____min
14	Is there any other renal pathology?	YES O NO O nm O
15	Is there any other abdominal pathology?	YES O NO O nm O

nm = not mentioned

Was the IVP interpreted by the emergency physician on the chart?
YES O NO O

Copy their comments:_____
_____

After presenting results of the other radiologist was there any change in opinion?
YES O NO O

_____
_____

Date:_____

Appendix 3

THE PARC STUDY

(PREDICTING OUTCOME IN ACUTE RENAL COLIC)

➡Patient's Name:_____Unique/hospital number:_____

➡Patient's Telephone Number: daytime:_____Evenings:_____

(Please inform patient that they will be telephoned in 1 week and at 4 weeks)

➡Physician's Name: _____ Time of initial assessment:_____

Goals of the PARC Study:

This study will assess the feasibility of developing a clinical decision rule that will identify clinical factors that are both sensitive and specific in predicting which patients with renal colic will have severe obstruction. The decision rule should prevent patient exposure to unnecessary imaging and lead to less time spent in the emergency department.

In addition, we'd like to identify which patients with renal colic will have a complicated course from delayed passage of their stone after discharge from the emergency department. This will allow for prioritization of referrals to the urologist and help urologists decide upon the need for further intervention.

Inclusion Criteria:

- Patients > 18 years old
- Suspected Renal Colic a) Flank pain b) Hematuria c) Acute pain within 24 hours

NB. (Pain may have been present for days but it has intensified within the last 24 hours.)

Exclusion Criteria:

- Patient already enrolled
- Allergy to contrast media
- Unstable vital signs
- Inability to contact patient for follow-up post discharge
- Pregnancy

➡Is this patient eligible for the PARC study? □ □ Yes □ No

If yes, please fill in areas marked with an arrow (➡)

Thank you for your time

Name of clerk or Triage Nurse who enrolled the patient_____

MEASURE PATIENT'S PAIN

The Visual Analogue Scale (VAS) is used to measure the degree of pain the patient is experiencing. Please have the patient place a mark along the 10 cm line to indicate the pain they are experiencing both on arrival and on discharge from the department.

➡A)VAS on arrival: 0 (No Pain) 10 (Worst pain ever)

➡B)VAS at discharge: 0 (No Pain) 10 (Worst pain ever)

C) **History of present illness:**

➡Previous history of renal colic: O YES O NO

➡Duration of today's episode of pain (episode leading to visit)?:_____(HOURS)

➡How many days of this pain before coming to the ED:_____(DAYS)

➡Vomiting: O YES O NO

➡Dysuria: O YES O NO

➡Frequency: O YES O NO

D)**Physical Exam**

➡Abdominal tenderness to palpation: O YES O NO

➡Rebound tenderness: O YES O NO

➡CVA tenderness: O YES O NO

➡E) Place a mark along the line where YOU think the patient's pain is?

0 (No Pain) 10 (Worst pain ever)

➡F) Do you think this patient is exaggerating their pain? O YES O NO

Explain (eg. drug seeking, cultural, low threshold)_____

➡G)What is the probability of this patient having renal colic? *(Please circle)*

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

➡H) What is the probability that an IVP will change your management?

O high probability O moderate probability O low probability O none O unsure

➡I) What is the probability of this patient having severe obstruction on IVP (extravasation, no ureteral filling > 2hours)?

0%, 1, 2, 3, 4, 5, 6, 7. 8, 9, 10....15.....20.....25.....30% 40% 50% 60% 70% 80% 90% 100%

➡J) What is the probability this patient will have problems passing their stone after discharge?

O high probability O moderate probability O low probability O none O unsure

Appendix 4

SURVEY OF RADIOLOGISTS IN OTTAWA

Name:_____

Number of years of practice: _____

1- In your opinion, is there a difference between complete and severe urinary obstruction from urinary stones? O YES O NO

What is the difference?_____

2- In your opinion, what radiological signs on IVP indicate severe obstruction in a patient with renal colic? (check any number of boxes)

- O Prolonged dense nephrogram (contrast outlining the kidney > 15min)
- O Pericalyceal extravasation of dye
- O Delayed visualization of contrast: O no filling of ureter >2 hrs
  O no filling of ureter >4 hrs
  O no filling of ureter >6 hrs
- O Site of obstruction along ureter (proximal vs. distal ureter):
- O Size of stone *(please indicate size)*: > _____mm
- O Degree of hydronephrosis
- O other *(Please specify)*

_____

3- In your opinion, what radiological signs on IVP indicate clinically significant severe obstruction (requiring urgent urologic intervention) in a patient with renal colic? (check any number of boxes)

- O Same as question 2
- O Prolonged dense nephrogram (contrast outlining the kidney > 15min)
- O Pericalyceal extravasation of dye
- O Delayed visualization of contrast: O no filling of ureter >2 hrs
  O no filling of ureter >4 hrs
  O no filling of ureter >6 hrs
- O Site of obstruction along ureter (proximal vs. distal ureter):
- O Size of stone *(please indicate size)*: > _____mm
- O Degree of hydronephrosis
- O other *(Please specify)*

_____

4- Do you think that every patient with suspected renal colic should get an IVP? O YES O NO

5- What is your imaging study of choice for diagnosing urinary stones in patients with renal colic?

O KUB O IVP O ULTRASOUND OSPIRAL CT

6- What is your imaging study of choice for determining degree of obstruction from urinary stones?

O KUB O IVP O ULTRASOUND OSPIRAL CT

Appendix 5

SURVEY OF UROLOGISTS IN OTTAWA

Name:_____
University or community hospital: _____
Number of years of practice: _____

1- In your opinion, is there a difference between <u>complete</u> and <u>severe</u> urinary obstruction from stones? O YES O NO
Why?_____
_____

2- How many cases of <u>severe</u> urinary obstruction secondary to stones do you see in 1 month? _____

3- Currently, there is no consensus on what is considered clinically significant "severe ureteral obstruction" on intravenous pyelography (IVP). In your opinion, what radiological signs on IVP indicate <u>severe obstruction</u> in a patient with renal colic? (check any number of boxes)

- O Prolonged dense nephrogram (contrast outlining the kidney > 15min)
- O Pericalyceal extravasation of dye
- O Delayed visualization of contrast: O no filling of ureter >2 hrs
  O no filling of ureter >4 hrs
  O no filling of ureter >6 hrs
- O Site of stone (location of stone along urinary tract):
- O Size of stone *(please indicate size)*: > _____mm
- O Degree of hydronephrosis
- O other *(Please specify)*

_____
_____

4- How long do you wait before intervening in a patient with <u>severe</u> obstruction?

5- How long do you wait before intervening in a patient with complete obstruction?

6- What are your criteria for intervening (i.e. performing one of the following procedures: percutaneous nephrostomy, lithotripsy, ureteroscopy,

or open surgery) in patients with renal colic? (check any number of boxes)

- O Fever (T > 38°C)
- O Uncontrolled pain (failure of intravenous narcotics for 12 hours)
- O Pain > 2 weeks
- O Pain > 4 weeks
- O Pain > 6 weeks
- O Solitary kidney
- O Protracted vomiting (> 3 episodes of vomiting in 1 day)
- O Elevated white blood cell count (WBC)
- O Elevated BUN and Creatinine (BUN > 8.9 and Cr > 150)
- O Signs of severe obstruction on IVP
- O Signs of complete obstruction on IVP
- O Patient's occupation
- O Patient's preference
- O Other_____

_____

7- Do you think that every patient with suspected renal colic should get an IVP? O YES O NO

8- What is your imaging study of choice for following patients with stones?

O KUB O IVP O ULTRASOUND OSPIRAL CT

9- What is your imaging study of choice for suspected renal colic in the emergency department? O KUB O IVP O ULTRASOUND OSPIRAL CT

Appendix 6

SURVEY OF EMERGENCY PHYSICIANS IN OTTAWA

Name:_____

1- In your opinion, is there a difference between complete and severe urinary obstruction from stones? O YES O NO

Why?_____

_____

2- How many cases of severe urinary obstruction secondary to stones do you see in 1 month? _____

3- Currently, there is no consensus on what is considered clinically significant "severe obstruction" from urinary stones on intravenous pyelography (IVP). In your opinion, what radiological signs on IVP indicate severe obstruction in a patient with renal colic? (check any number of boxes)

O Prolonged dense nephrogram (contrast outlining the kidney > 15min)
O Pericalyceal extravasation of dye (dye outside the kidney)
O Delayed visualization of contrast: O no filling of ureter >2 hrs
O no filling of ureter >4 hrs
O no filling of ureter >6 hrs
O Site of stone along ureter (where stone is along urinary tract):
O size of stone *(please indicate size)*: > _____mm
O Degree of hydronephrosis
O other *(Please specify)*

4- For what % of suspected renal colic patients that you see in the ED do you order an IMMEDIATE IVP? (Please circle response)

5- Why should IVP's be performed on patients with suspected renal colic? (choose as many as you consider appropriate)

O To detect obstruction
O To decide on admission
O To confirm diagnosis
O To decide upon referral
O To detect complications
O To see size of stone

Appendix 7

Telephone follow-ups sample script

"Hello M._____ I am Dr. _____ from the Ottawa Hospital. I am calling you to follow-up on your recent visit to the emergency department, where you were treated for kidney stones. Would you mind answering a few brief questions for a project we are doing on kidney stones here at the Ottawa Hospital? It should take no longer than 15 minutes. Your cooperation in answering these questions will not affect how you will be treated in any way and your participation is optional. May I continue with some questions.

QUESTIONS

Are you still experiencing pain from your stone?

If yes, what is your pain rated on a scale from 0 to 10?

Has the pain kept you from working or sleeping or eating?

Have you experienced any fever, chills, vomiting, pain when you urinate, blood in your urine or have you felt the stone pass?

Have you experienced any side effects from the medications?

Have you seen a physician since being seen in the emergency department?

Have any new tests or operations been performed since your visit?

M._____ I would like to thank you for your cooperation. With your permission I will follow up with you on one more occasion by telephone."

Appendix 8

SUBJECT NUMBER:_____

FOLLOW-UP PHONE CALL -

PATIENT'S PHONE #:_____ / _____
DATE & TIME OF FOLLOW-UP ATTEMPTS:_____
:_____
:_____
DATE REACHED: (Y/M/D): _____

40.) Are you still experiencing pain from your stone? YES O NO O

If yes,

a) What is your pain rated on a scale from 0 to 10?_____
(0 means no pain and 10 is the worst imaginable pain)

b) Has the pain kept you from working YES O NO O

c) sleeping YES O NO O

d) eating YES O NO O

41.) Have you experienced any fever or chills YES O NO O

b) vomiting YES O NO O

c) pain when you urinate YES O NO O

d) blood in your urine YES O NO O

e) have you felt the stone pass? YES O NO O

42.) Have you experienced any side effects from the medications? YES O NO O
e.g. a) O skin rash b) Odizziness, c) Oheadache, d) Onausea, e) Ovomiting, f) Odiarrhea, g) Osore stomach

_____

43.) Have you seen a physician since being seen in the emergency department?
YES O NO O

If yes, O Family Doctor O Emergency department O Urologist
Name of facility_____ Name of doctor_____

44) Have any tests or operations been performed since your visit? YES O NO O
a) O Blood tests b) O X-rays c) O Ultrasound
d) O Percutaneous **nephrostomy,** e) O Extracorporeal shockwave **lithotripsy**
f) O Ureteroscopy (**scope** with or without basket extraction), g) O Open surgery
results_____

Appendix 9

OTTAWA HOSPITAL

CHART REVIEW OF RENAL COLIC CASES

February 1998 - March 1998

At the Civic Campus of the Ottawa Hospital there is a database called EPIS that codes all visits to the emergency department. Diagnostic codes are based on the final diagnosis written on the chart.

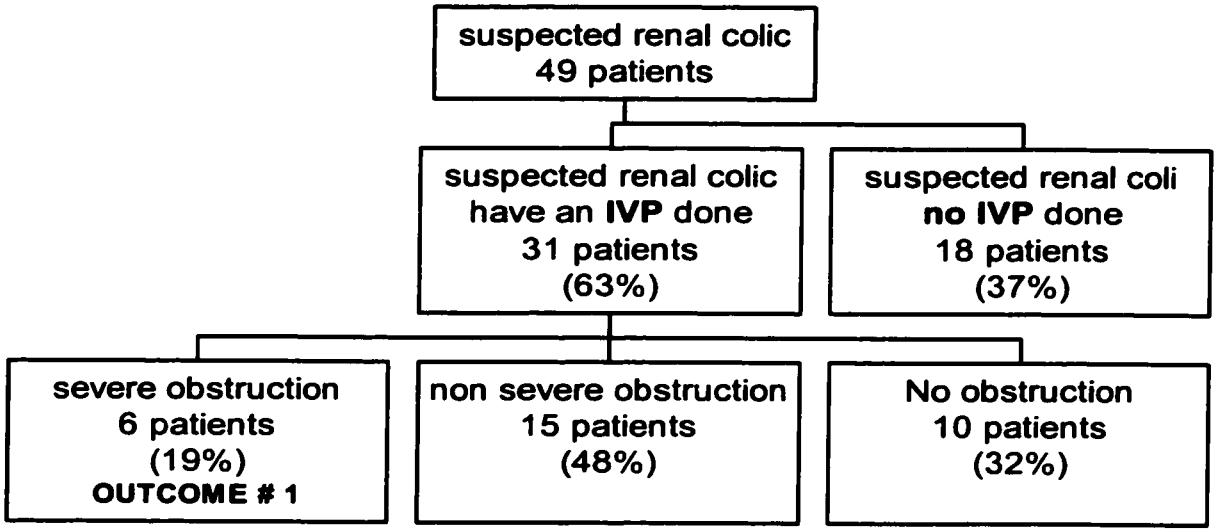
Over a 7-week period from February 1998 to March 1998 charts were reviewed for potential renal colic cases. The EPIS database was searched for all charts having a diagnosis of renal colic, ureteral stones/calculi, and kidney stones/calculi. These charts were then pulled from the medical records department and reviewed to identify all potential renal colic cases. Potential renal colic cases were selected by final diagnosis & clinical findings (acute onset of severe flank pain with or without hematuria) & by confirmatory imaging.

Forty-nine cases of potential renal colic cases were found. In these patients 63% had an IVP performed. Of the 31 IVP's performed, 6 had evidence of severe obstruction (based on the current radiological definition) and 10 were found to be normal with no evidence of a stone.

Limitations to this chart review include the limitations associated with using a database, namely coding. The chart review was not meant to draw conclusions; only to estimate number of renal colic cases.

Number Of Patients Obtained from a Chart Review at the
Civic Site of Ottawa Hospital
(Numbers Estimate Severe Obstruction Cases Expected in the Feasibility Study)

DECISION RULE #1
CHART REVIEW RESULTS



Appendix 10

Calculations of Confidence Intervals for Sample Sizes

Using Model #1:

From our logistic regression data using SPSS software, we obtained a correlation matrix. We converted our correlation matrix into a covariance matrix by the following formula:

$$C_{ij} = R_{ij} S_i S_j$$

C_{ij} = value in the covariance matrix (calculated)

R_{ij} = value in the correlation matrix (from logistic regression)

S_i = standard error of variable$_i$ our logistic regression model

S_j = standard error of variable$_j$ our logistic regression model

The standard error and $\beta^{\wedge\prime}$ value of each variable in the Model #1 are displayed in Table 10.1 and the calculated covariance matrix for Model #1 in Table 10.2. The clinical variables from our logistic regression model are represented as follows:

x_c = constant

x_1 = discharge VAS $\geq$ 1 cm

x_2 = pain in ED after 6 hours

x_3 = rebound tenderness

x_4 = vomiting

Since all the variables in the logistic model #1 are categorical, the variable x is either 0 or 1. Let us take two examples of CI calculations in which all variables are either positive (x=1) or negative (x=0).

Example 1:

If we assume a situation where all variables are positive, (i.e. x=1) and the constant is 1 we can make the following calculations using the values in Tables 10.1 & 10.2:

$$x_c = \text{constant},\ x_1 = 1,\ x_2 = 1,\ x_3 = 1,\ x_4 = 1$$

$$C.I. = x\hat{\beta}' \pm Z_{(1-\alpha/2)}(\sqrt{xSx'/\eta})$$

$Z_{(1-\alpha/2)} = 1.96$ (assuming a 95% confidence interval)

S = Covariance Matrix

$\chi\hat{\beta}' = (x_c, x_1, x_2, x_3, x_4)$ $\quad (\hat{\beta}'_c)\ (\hat{\beta}'_1)\ (\hat{\beta}'_2)\ (\hat{\beta}'_3)\ (\hat{\beta}'_4)$

$x\hat{\beta} = (1,1,1,1,1)$ $\quad \begin{matrix}(-5.171)\\(2.452)\\(2.922)\\(2.918)\\(2.054)\end{matrix}$

$x\hat{\beta}' = 5.175$

$xSx' = (1,1,1,1,1)$

1.904	-0.912	-0.938	-0.895	-1.109	1
-0.912	0.996	0.553	0.292	0.246	1
-0.938	0.553	1.199	0.532	0.256	1
-0.895	0.292	0.532	2.033	0.427	1
-1.109	0.246	0.256	0.427	1.117	1

$xSx' = 4.153$

$C.I. = 5.175 \pm 1.96\ (\sqrt{4.153/n})$

By varying the n we can observe different widths of the confidence intervals and determine an appropriate accuracy for estimation of the study. We took a series of sample sizes from 20 to 1000 and calculated confidence intervals on these sample sizes using this formula. After obtaining the confidence intervals for the logit we applied the inverse transformation

$$P = \exp(CI)/(1 + \exp(CI))$$

to the upper and lower bounds of the confidence interval for the logit to obtain the corresponding confidence interval (p_L,p_U) for estimating the conditional probability of patient response, given x.

The results of these calculations are shown in Table 10.3. As n increased the confidence interval width decreased. The limit of error for each confidence interval width was also calculated. Given this combination of variables, a confidence interval width of 0.1 (95% CI) could be achieved with a sample size of 20 patients. For a confidence interval width of 0.05 (97.5%CI) we would require 100 patients. For a confidence interval width of 0.01 (99%CI) we would require 1000 patients. Figure 10.1 illustrates how the width of the confidence interval tapered quickly down to a sample size of 300 and then became almost constant once n=500.

Example 2:

If we assume a situation where all variables have a value of 0, (i.e. x=0) and the constant is 1 we can make the following calculations using the values in Tables 10.1 & 10.2:

x_c = constant, $x_1 = 0$, $x_2 = 0$, $x_3 = 0$, $x_4 = 0$

$Z_{(1-\alpha/2)} = 1.96$ (assuming a 95% confidence interval)

S = Covariance Matrix

$x\beta\hat{}' = (x_c, x_1, x_2, x_{x3}, x_4)$ $(\beta\hat{}'_c)$ $(\beta\hat{}'_1)$ $(\beta\hat{}'_2)$ $(\beta\hat{}'_3)$ $(\beta\hat{}'_4)$

$x\beta\hat{}' = (1,0,0,0,0)$ (-5.171) (2.452) (2.922) (2.918) (2.054)

$x\beta^{\wedge\prime} = -5.171$

$xSx' = (1,0,0\ 0,0)$

1.904	-0.912	-0.938	-0.895	-1.109	1
-0.912	0.996	0.553	0.292	0.246	0
-0.938	0.553	1.199	0.532	0.256	0
-0.895	0.292	0.532	2.033	0.427	0
-1.109	0.246	0.256	0.427	1.117	0

$xSx' = 1.904$

$C.I. = -5.171 \pm 1.96\ (\sqrt{1.904/n})$

Again, we took a series of sample sizes from 20 to 1000 and calculated confidence intervals on them using this formula and then applied the inverse transformation: P = exp (CI)/(1 + exp (CI)) to the upper and lower bounds of the confidence interval for the logit to obtain the corresponding confidence interval (p_L,p_U). Table 10.4 shows how the confidence interval and the bounds of error changed with increasing sample sizes. This is depicted graphically in Figure 10.2. Given this combination of variables, a confidence interval width of 0.05 (97.5%CI) could be achieved with a sample size of 50 patients. For a confidence interval width of 0.01 (99%CI) we would require 500 patients.

In order to explore the tightness of confidence interval widths for the model we extended these two examples to include all possible combinations of the 4-variable Model #1. Using the above formulae for all 16 combinations we graphed both the confidence interval width and the bound of error in Figure 10.3 and Figure 10.4 respectively. These graphs demonstrate how the slope of the graphs decreased markedly once sample sizes approached 300 to 400 and changed minimally after 600 to 700. Given the all parameters of Model #1, a confidence interval width of 0.1 (95% CI) would be achieved with a sample size

of 200 patients. For a confidence interval width of 0.05 (97.5%CI) we would require 700 patients. For a confidence interval width of 0.01 (99%CI) we would require over ten thousand patients.

Calculations of Confidence Intervals for Sample Sizes Using Model #2:

$$C.I. = x\hat{\beta}' \pm Z_{(1-\alpha/2)}(\sqrt{xSx'/\eta})$$

From our logistic regression data for Model #2 we calculated our covariance matrix from the correlation matrix by the following formula:

$$C_{ij} = R_{ij} S_i S_j$$

C_{ij} = value in the covariance matrix (calculated)

R_{ij} = value in the correlation matrix (from logistic regression)

S_i = standard error of variable$_i$ our logistic regression model

S_j = standard error of variable$_j$ our logistic regression model

The standard error and $\hat{\beta}'$ value of each variable in the model #2 are displayed in Table 10.5 and the calculated covariance matrix for model #2 in Table 10.6. The clinical variables from Model #2 are represented by the following mathematical symbols:

x_c = constant

x_1 = rebound tenderness

x_2 = pain after 6 hours in ED

x_3 = MD's VAS $\geq$ 7 cm

Since all our variables in the logistic Model #2 are categorical, the variable x is either 0 or 1. Let us take two examples of CI calculations; the first in which all

variables are positive in the model (x=1) and the second in which all variables are negative in the model (x=0).

Example 1:

If we assume a situation where all variables are 1, (i.e. x=1) and the constant is 1 we can make the following calculations using the values in Tables 10.5 & 10.6:

x_c = constant, $x_1 = 1$, $x_2 = 1$, $x_3 = 1$

$Z_{(1-\alpha/2)} = 1.96$ (assuming a 95% confidence interval)

S = Covariance Matrix

$x\beta\hat{}' = (x_c, x_1, x_2, x_3)$ $\begin{matrix}(\beta\hat{}'_c)\\(\beta\hat{}'_1)\\(\beta\hat{}'_2)\\(\beta\hat{}'_3)\end{matrix}$

$x\beta\hat{}' = (1,1,1,1)$ $\begin{matrix}(-4.653)\\(3.334)\\(2.092)\\(2.639)\end{matrix}$

$x\beta\hat{}' = 3.412$

$xSx' = (1,1,1,1)$

1.510	-0.823	-0.257	-1.374	1
-0.823	2.132	0.240	0.639	1
-0.257	0.240	0.558	0.034	1
-1.374	0.639	0.034	1.471	1

$xSx' = 2.589$

C.I. = $3.412 \pm 1.96 (\sqrt{2.589/n})$

By varying the η from 20 to 1000 we constructed a graph with varying confidence intervals to determine the best value meeting the desired study estimation accuracy. After obtaining the confidence intervals for the logit we applied the inverse transformation to the upper and lower bounds to obtain the corresponding confidence interval (p_L,p_U). In Table 10.7 we observe how

confidence interval width and its bound of error decreased as η increased. Given this combination of variables, a confidence interval width of 0.05 (97.5%CI) could be achieved with a sample size of 20 patients. For a confidence interval width of 0.01 (99%CI) we would require 200 patients. Figure 10.5 illustrates how the width of the confidence interval tapered quickly down to a sample size of 200 and then changed minimally thereafter.

Example 2:

If we assume a situation where all variables have a value of 0, (i.e. x=0) and the constant is 1 we can make the following calculations using the values in Tables 10.8 & 10.9:

x_c = constant, $x_1 = 0$, $x_2 = 0$, $x_3 = 0$

$Z_{(1-\alpha/2)} = 1.96$ (assuming a 95% confidence interval)

S = Covariance Matrix

$x\beta\hat{}' = (x_c, x_1, x_2, x_3)$ $\quad \begin{matrix}(\beta\hat{}'_c)\\(\beta\hat{}'_1)\\(\beta\hat{}'_2)\\(\beta\hat{}'_3)\end{matrix}$

$x\beta\hat{}' = (1,0,0,0)$ $\quad \begin{matrix}(-4.653)\\(3.334)\\(2.092)\\(2.639)\end{matrix}$

$x\beta\hat{}' = -4.653$

$xSx' = (1,0,0\ 0)$

1.510	-0.823	-0.257	-1.374	1
-0.823	2.132	0.240	0.639	0
-0.257	0.240	0.558	0.034	0
-1.374	0.639	0.034	1.471	0

$xSx' = 1.510$

$C.I. = -4.653 \pm 1.96\ (\sqrt{1.510/n})$

Using this formula we varied the sample sizes from 20 to 1000 and calculated confidence intervals and transformed them from the logit to obtain the results in Table 10.10. Given this combination of variables, a confidence interval width of 0.01 (99%CI) could be achieved with a sample size of 50 patients.

In order to explore the tightness of confidence interval widths for the entire model we extended these two examples to include all possible combinations of the 3-variable Model #2. Using the above formulae for all 8 combinations we graphed the confidence interval width in Figure 10.7. This graph demonstrates, like Figure 10.3 in model #1, how the slope of the graph decreased markedly once sample sizes approached 300 to 400 and changed minimally after 600 to 700. This is seen consistently over all 8 combinations of the variables. Given the all parameters of Model #2, a confidence interval width of 0.1 (95% CI) would be achieved with a sample size of 200 patients. For a confidence interval width of 0.05 (97.5%CI) we would require 700 patients. For a confidence interval width of 0.01 (99%CI) we would require over ten thousand patients.

Evidently, both models yielded the same results.

Table 10.1
Results of Logistic Regression for Predicting Severe Obstruction
Variables in Model #1

Variable	β^	*SE* (β^)
Constant	-0.5171	1.380
Discharge VAS ≥ 1 cm	2.452	0.998
Pain in ED after 6 hrs	2.922	1.095
Rebound tenderness	2.918	1.426
Vomiting	2.054	1.057

Table 10.2
Covariance Matrix for Predicting Severe Obstruction in Model #1

Parameter	*Constant*	*Discharge VAS ≥ 1cm*	*Time to pain relief in ED> 6 hrs*	*Rebound tenderness*	*Vomiting*
Constant	**1.904**	**-0.912**	**-0.938**	**-0.895**	**-1.109**
Discharge VAS ≥ 1 cm	**-0.912**	**0.996**	**0.553**	**0.292**	**0.246**
Pain in ED after 6 hrs	**-0.938**	**0.553**	**1.199**	**0.532**	**0.256**
Rebound tenderness	**-0.895**	**0.292**	**0.532**	**2.033**	**0.427**
Vomiting	**-1.109**	**0.246**	**0.256**	**0.427**	**1.117**

Table 10.3
Confidence Intervals For Various Sample Sizes
Using the Inverse Transformation to convert from the Logit Model #1
Given ($x_1 = 1, x_2 = 1, x_3 = 1, x_4 = 1$)

n	*P lower*	*P upper*	*Interval Width*	*Bound of Error*
20	0.9864	0.9977	0.0113	0.0113
50	0.9901	0.9968	0.0067	0.0067
100	0.9916	0.9962	0.0046	0.0046
200	0.9926	0.9958	0.0032	0.0032
300	0.9929	0.9955	0.0026	0.0026
400	0.9931	0.9954	0.0023	0.0023
500	0.9933	0.9953	0.0020	0.0020
600	0.9934	0.9952	0.0018	0.0018
700	0.9935	0.9952	0.0017	0.0017
800	0.9935	0.9951	0.0016	0.0016
900	0.9936	0.9951	0.0015	0.0015
1000	0.9936	0.9950	0.0014	0.0014

Table 10.4
Confidence Intervals For Various Sample Sizes
Using the Inverse Transformation to convert from the Logit Model #1
Given ($x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0$)

n	*P lower*	*P upper*	*Interval Width*	*Bound of Error*
20	0.0031	0.0102	0.0071	0.6997
50	0.0039	0.0083	0.0044	0.5321
100	0.0043	0.0074	0.0031	0.4189
200	0.0047	0.0068	0.0021	0.3088
300	0.0048	0.0066	0.0018	0.2727
400	0.0049	0.0065	0.0016	0.2461
500	0.0050	0.0064	0.0014	0.2188
600	0.0050	0.0063	0.0013	0.2063
700	0.0051	0.0063	0.0012	0.1905
800	0.0051	0.0062	0.0011	0.1774
900	0.0052	0.0062	0.0010	0.1613
1000	0.0052	0.0062	0.0010	0.1613

Table 10.5
Results of Logistic Regression for Predicting Severe Obstruction
Variables in the Model #2

Variable	*β^*	*SE (β^)*
Constant	- 4.653	1.229
Rebound tenderness	3.334	1.416
Pain in ED after 6 hrs	2.092	0.747
VAS estimated by MD ≥ 7 cm	2.639	1.213

Table 10.6
Covariance Matrix for Predicting Severe Obstruction in Model #2

Parameter	*Constant*	*Rebound tenderness*	*Time to pain relief in ED> 6 hrs*	*VAS estimated by MD to be > 7 cm*
Constant	**1.510**	**- 0.823**	**- 0.257**	**- 1.374**
Rebound tenderness	**- 0.823**	**2.132**	**0.240**	**0.639**
Pain in ED after 6 hrs	**- 0.257**	**0.240**	**0.558**	**0.034**
VAS estimated by MD to be > 7 cm	**- 1.374**	**0.639**	**0.034**	**1.471**

Table 10.7
Confidence Intervals For Various Sample Sizes
Using the Inverse Transformation to convert from the Logit Model #2
Given ($x_1 = 1$, $x_2 = 1$, $x_3 = 1$)

N	*P lower*	*P upper*	*Interval width*	*Bound of Error*
20	0.9374	0.9840	0.0466	0.0474
50	0.9510	0.9793	0.0283	0.0289
100	0.9568	0.9765	0.0197	0.0202
200	0.9604	0.9743	0.0139	0.0143
300	0.9619	0.9732	0.0113	0.0116
400	0.9628	0.9726	0.0098	0.0101
500	0.9634	0.9722	0.0088	0.0091
600	0.9638	0.9718	0.0080	0.0082
700	0.9642	0.9716	0.0074	0.0076
800	0.9644	0.9714	0.0070	0.0072
900	0.9647	0.9712	0.0065	0.0067
1000	0.9649	0.9710	0.0061	0.0063

Table 10.8
Variables in Logistic Regression Model #2 For Severe Obstruction

Variable	*β*	*SE*	*Wald*	*Sig.*	*Odds Ratio*	*95% C.I.*
Rebound tenderness	3.3	1.4	5.5	0.019	28.0	1.7 – 450.0
Pain after 6 hours in ED	2.1	0.747	7.8	0.005	8.1	1.9 – 35.0
MD's VAS over 7 cm	2.6	1.213	4.7	0.030	14.0	1.3 – 150.8

Table 10.9
Probability of Severe Obstruction based on Presence or Absence of the 3 Predictor Variables from the Logistic Regression Model #2

# of variables	Rebound tenderness (x_1)	No pain relief after 6 hrs (x_2)	VAS MD >=7 cm (x_3)	P(x)	%	sensitivity
one	0	1	0	0.072	7	92
	0	0	1	0.118	12	62
	1	0	0	0.211	21	62
two	1	1	0	0.684	68	15
	0	1	1	0.519	52	15
	1	0	1	0.789	79	0
three	1	1	1	0.968	97	0

Table 10.10
Confidence Intervals For Various Sample Sizes
Using the Inverse Transformation to convert from the Logit Model #2
Given ($x_1 = 0$, $x_2 = 0$, $x_3 = 0$)

N	*P lower*	*P upper*	*Interval width*	*Bound of Error*
20	0.0055	0.0156	0.0101	0.6474
50	0.0067	0.0132	0.0065	0.4924
100	0.0074	0.0120	0.0046	0.3833
200	0.0080	0.0112	0.0032	0.2857
300	0.0082	0.0108	0.0026	0.2407
400	0.0084	0.0106	0.0022	0.2075
500	0.0085	0.0105	0.0020	0.1905
600	0.0086	0.0104	0.0018	0.1731
700	0.0086	0.0103	0.0017	0.1650
800	0.0087	0.0103	0.0016	0.1553
900	0.0087	0.0102	0.0015	0.1471
1000	0.0088	0.0102	0.0014	0.1373

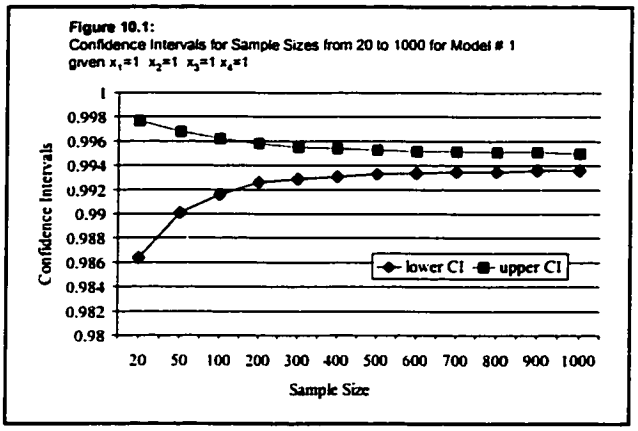


Figure 10.1:
Confidence Intervals for Sample Sizes from 20 to 1000 for Model # 1
given x1=1 x2=1 x3=1 x4=1
Confidence Intervals
1
0.998
0.996
0.994
0.992
0.99
0.988
0.986
0.984
0.982
0.98
lower CI
upper CI
20 50 100 200 300 400 500 600 700 800 900 1000
Sample Size

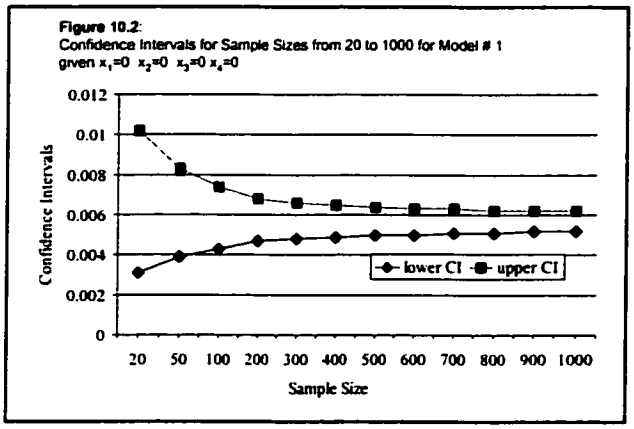


Figure 10.2:
Confidence Intervals for Sample Sizes from 20 to 1000 for Model # 1
given x1=0 x2=0 x3=0 x4=0
Confidence Intervals
0.012
0.01
0.008
0.006
0.004
0.002
0
lower CI
upper CI
20 50 100 200 300 400 500 600 700 800 900 1000
Sample Size

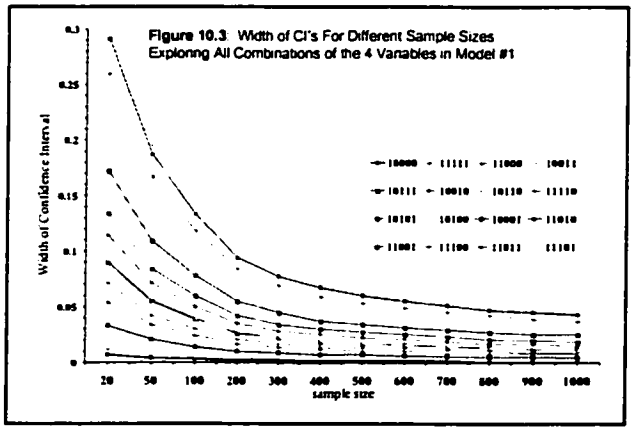


Figure 10.3 Width of CI's For Different Sample Sizes
Exploring All Combinations of the 4 Variables in Model #1
Width of Confidence Interval
sample size

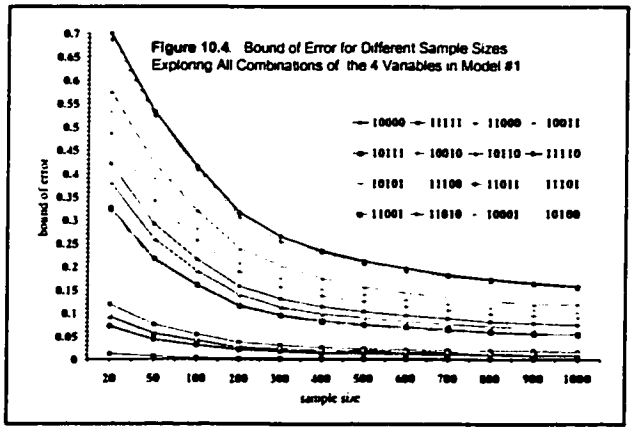


Figure 10.4 Bound of Error for Different Sample Sizes
Exploring All Combinations of the 4 Variables in Model #1
bound of error
sample size

Figure 10.5:
Confidence Intervals for Sample Sizes from 20 to 1000 for Model # 2
given $x_1=1$ $x_2=1$ $x_3=1$

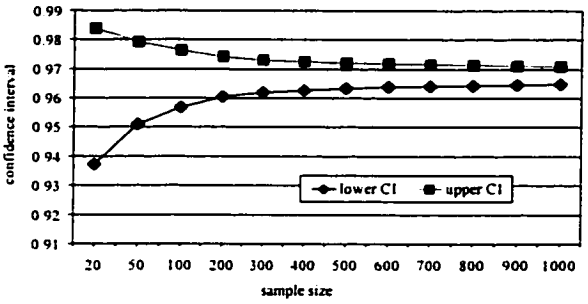


Figure 10.6:
Confidence Intervals for Sample Sizes from 20 to 1000 for Model # 2
given $x_1=0$ $x_2=0$ $x_3=0$

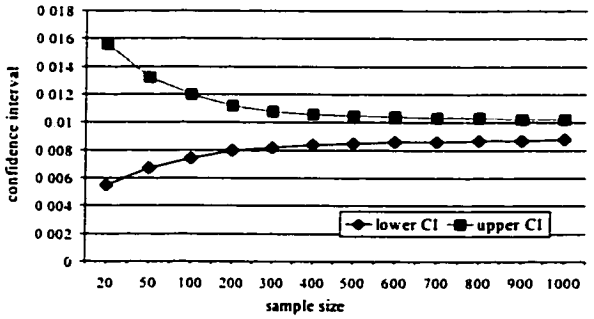


Figure 10.7:
Width of CI's for Different Sample Sizes
Exploring All Combinations of the 3 variables in Model #2