Title Page

- Title of Project.

Identifying Unnecessary Irradiation of Patients with Suspected Renal Colic

- Principal Investigator and Team Members.

Christopher L. Moore MD, Cary Gross MD, Annette Molinaro PhD, Dinesh Singh MD, Brock Daniels MD, Christal Esposito, Seth Luty, Richelle Jessey.

- Organization.

Yale University School of Medicine, Department of Emergency Medicine

- Inclusive Dates of Project.

August 16, 2010 to July 31, 2014

- Federal Project Officer.

Marian James

- Acknowledgment of Agency Support.

We acknowledge that this project was supported by the Agency for Healthcare Research and Quality

- Grant Award Number.

R01HS018322

1. Structured Abstract

Purpose: To derive and validate an objective clinical prediction rule for the presence of uncomplicated ureteral stone in patients eligible for computed tomography imaging.

Scope: The derivation cohort used a random selection of subjects undergoing CT between April 2005 and November 2010; the validation cohort included consecutive prospectively enrolled patients from May 2011 to January 2013.

Methods: Multivariate logistic regression was used to determine the top five factors associated with ureteral stone and these were assigned integer points to create a scoring system that was stratified into low, moderate and high probability of ureteral stone. In the prospective phase this score was observationally derived blinded to CT results and compared to the prevalence of ureteral stone and important alternative causes of symptoms.

Results: The derivation sample determined the five factors found to be most predictive of ureteral stone: male gender, short duration of pain, non-black race, presence of nausea or vomiting, and microscopic hematuria, yielding a score of 0-13 (the "S.T.O.N.E. Score"). Prospective validation was performed on 491 subjects. In the derivation and validation cohorts ureteral stone was present in 8.3% and 9.2% of the "low" (score 0-5) group, 51.6% and 51.3% of the "moderate" group (score 6-9), and 89.6% and 88.6% of the "high" group (score 10-13), respectively. In the "high" group, acutely important alternative findings were present in 0.3% derivation cohort and 1.6% of the validation cohort.

Key Words: Kidney stone, ureteral stone, renal colic, computed tomography, CT, ultrasound, imaging, radiation, clinical decision rules, clinical prediction rules

2. **Purpose** (Objectives of Study).

Phase 1 and 2[1]:

Our primary goal was to derive and validate a clinical prediction score for symptomatic ureteral stones, identifying patients who have either a very high or a very low likelihood of having an uncomplicated ureteral stone. We hypothesized that patients who are highly likely to have a kidney stone are unlikely to harbor an important alternative diagnosis, and may be appropriate for imaging choices other than standard dose CT.

Additional purposes of phases 1 and 2 included:

- Determining the prevalence of acutely important alternate causes of symptoms in patients with suspected renal colic.[2]
- Demonstrating test characteristics of point-of-care renal ultrasound and its effect on the S.T.O.N.E. score.[3,4]
- Determining the prevalence of reduced dose CT usage in the United States for suspected renal colic.[5]
- Designing and testing a substantially reduced dose CT protocol for kidney stones.[6]

In phase 3 (manuscript in preparation[7]) we implemented a reduced dose CT protocol in conjunction with our clinical prediction rule to help guide imaging choices and determine the amount of radiation reduction and any potential adverse effects of this protocol.

3. **Scope** (Background, Context, Settings, Participants, Incidence, Prevalence).

Kidney stones are now estimated to occur at some point in nearly 1 in 11 persons in the United States, with flank or kidney pain resulting in over 2 million annual emergency department visits. Computed tomography (CT) scanning has been described as the "best imaging study to confirm the diagnosis of a urinary stone" and is now the first line imaging study for suspected kidney stone in the US. While accurate, CT scanning is costly, involves the use of ionizing radiation, and does not appear to have impacted patient-centered outcomes such as rates of diagnosis or hospital admission in patients with suspected kidney stone.

Many patients with flank pain will not benefit from a CT scan, as the majority of kidney stones will pass spontaneously. Moreover, it is unlikely that the a CT scan in the setting of flank pain will detect acutely important alternative findings in patients without signs of infection. Hence, an objective clinical prediction rule for renal colic that could reliably identify patients highly likely to have a ureteral stone (and thus unlikely to have an important alternative diagnosis) may allow patients to be safely managed without imaging, or imaged with other approaches such as ultrasound or reduced-dose CT scans.

Settings and Participants:

The Yale New Haven Hospital emergency department is an urban, tertiary care teaching hospital and trauma center that sees over 80,000 adult ED visits annually. The Shoreline Medical Center (SMC) emergency department is a free-standing 8-bed suburban ED without residents that sees approximately 20,000 adult and pediatric patients annually.

Derivation phase

Dictated reports of all patients receiving a CT “flank pain protocol” (the name given at both sites to a non-contrast enhanced CT protocol for suspected kidney stone) at either of the two emergency department sites between April 2005 to November 2010 were electronically retrieved. Subjects were eligible if the CT was performed in the emergency department and they were 18 years of age or older at the time of CT. From an original set of over 5,000 CTs, approximately one third of records (estimated to yield about 1000 records that met inclusion criteria) were selected for full record review using a random number spreadsheet function (Microsoft Excel, Redmond WA). Exclusion criteria were lack of any flank or back pain, history of trauma, evidence of infection (subjective or objective fever or presence of leukocytes on urine dipstick analysis), known active malignancy, known renal disease (including creatinine >1.5mg/dl or 133μmol/L), or prior urologic procedure (including lithotripsy or ureteral stent).[2]

Validation phase

From May of 2011 to February of 2013, consecutive subjects presenting during defined periods of time to the emergency department sites in whom the clinician intended to obtain a CT scan for kidney stone were approached for enrollment. Both clinicians and enrolling personnel were not aware of the specific elements of the rule derived in the retrospective phase. Defined enrollment shifts included overnights, weekends, and holidays, and an automatic paging system was set up to notify the research associate of all CTs ordered for renal colic. Review of the hospital imaging system was conducted daily to monitor any subjects missed during enrollment and/or when enrollment was not taking place.

4. **Methods** (Study Design, Data Sources/Collection, Interventions, Measures, Limitations).

Phase 1 and 2: retrospective derivation and prospective validation of a clinical prediction rule for uncomplicated ureteral stone.

Study design and setting

We performed a retrospective derivation and prospective validation of a clinical scoring system for symptomatic ureteral stones in two separate emergency departments that utilize the same medical record systems. At the time of this study, both sites utilized a templated, hand-written, scanned emergency department patient care record (Lynx Medical Systems, Bellevue WA) with laboratory and dictated radiology reports on Sunrise Clinical Manage (Eclipsys, Atlanta GA). The derivation (retrospective) phase was approved by the Human Investigation Committee of the Yale Institutional Review Board with a waiver of informed consent, and the validation (prospective) phase involved written informed consent from all subjects.

Data abstraction

Based on clinical experience and review of the literature five physician co-investigators from three specialties (emergency medicine, internal medicine and urology) identified an *a priori* list of factors thought to potentially be predictive of ureteral stone (online appendix 1). Literature review was conducted using key word searches in PubMed and relevant citations through Web of Science (Thomson Reuters). These factors were then abstracted from medical records blinded to CT reports. The Lynx medical record used by emergency clinicians during the study period is a templated, hand written chart that specifically prompts clinicians for the presence or absence of factors related to the chief complaint selected (typically flank or back pain), and was well suited to determining the presence or absence of factors. These factors were abstracted into a standardized form on an electronic database (Filemaker Pro 12, FileMaker Inc., Santa Clara CA).

Results of the dictated CT reports were blindly abstracted and categorized as previously described.[2] CT reports were reviewed primarily to determine whether a kidney stone was causing symptoms, or whether the CT showed another cause of symptoms. A kidney stone was considered symptomatic if it was located from the renal pelvis to the ureterovesical junction (parenchymal stones were not considered symptomatic) or if signs of passed ureteral stone were specifically mentioned in the CT report. Acutely important alternative causes of symptoms (such as appendicitis, diverticulitis, and others) were noted.[2] Other factors associated with kidney stone including stone size, location, presence and degree of hydronephrosis or hydroureter, presence of perinephric or ureteral stranding, and asymptomatic stones as well as incidental findings (defined as unrelated to patient symptoms) were also documented. CT results were abstracted into a standard form on a separate FileMaker database.

Inter-rater reliability

A subset of 50 randomly selected records were blindly re-reviewed to determine inter-rater reliability of elements abstracted from the medical record. *Apriori*, any element with a kappa of below 0.6 was not eligible for inclusion in the prediction rule. Inter-rater reliability of categorization of CT scan results was performed from a random selection of records.[2]

Constructing the scoring system

All variables included in Table 2 were considered via univariate logistic regression analysis with estimation of prevalence and odds ratios with corresponding 95% confidence intervals. Multivariate logistic regression was performed employing forward selection and 10-fold cross-validation for model selection including estimation of two measures of prediction accuracy: the misclassification rate and the area under the receiver operating characteristic curve (AUC). Misclassification is a measure of prediction error, ranges from 0 to 1, with lower scores indicating less error in prediction. AUC ranges from 0.5 to 1, with higher scores indicating better prediction. The best model was the one that had a low cross-validated misclassification rate and had a high AUC. Subsequently, all observations were included to provide the most accurate estimates of the coefficients for the selected model and to derive a corresponding integer scoring system following the methodology used in the Framingham study. The simplicity of this scoring system allows calculation of a patient's risk without the need for a calculator. Initially the variables in the final multivariate model are organized into meaningful categories, each with a specific reference value. A referent risk factor profile is selected as the base category for each risk factor and assigned 0 points in the scoring system, such that a higher point total conveys more risk. Next, the difference in terms of regression units between each category and the corresponding base category is calculated. The constant, B, for the points system is set as the number of regression units that corresponds to one point. The points for each risk factor's risk categories are computed as the difference in regression units between each category and its base category divided by B. Subsequently the risk associated with each point total is calculated via the multiple logistic regression equation. A weighted kappa test is used to verify the agreement between risk estimates based on the point system and those based on the multivariate logistic regression model. In addition to estimating AUC for summarizing the model's discrimination, the Hosmer and Lemeshow test is used to test for goodness of fit and calibration.

While the odds ratios (coefficients) from the multivariate regression analysis can be used to estimate the probability of an event (in this case ureteral stone), we sought to construct a more straightforward scoring system for clinical use without the use of complicated calculations. Integer points were assigned to the presence of risk factors for ureteral stone using the coefficients from a multivariate analysis based on all observations, as described in methods used to estimate risk of cardiovascular disease in the Framingham study. Points for each factor were computed as the difference in regression units between each category and its base category, which was given a value of zero.

In order to assess the difference in accuracy between the integer point system and the logistic regression model we calculated the misclassification rate, AUC, and weighted kappa based on differences in classification for each model. In addition to estimating AUC for summarizing the model's discrimination, the Hosmer and Lemeshow test was used to test for goodness of fit and calibration.

After the point system was constructed from the derivation phase but prior to analysis of prospective data the research team selected three categories for risk (low, moderate, and high) based on estimated clinical utility for the probability of ureteral stone by point total in each category.

Prospective Validation

Prior to analysis of the validation data the scoring system was developed from the derivation set as described above, yielding a 0-13 point scale. Also prior to analysis of the prospective data this scale was stratified based on estimated clinical utility into “low” (~10%), “moderate” (~50%), and “high” (~90%) likelihood of ureteral stone. Estimated clinical utility of cut points on the scale were arrived at via consensus of all investigators, including physicians from emergency medicine, internal medicine, and urology. Stratification into three groups provided the ability to compare the derivation and validation sets in terms of clinical utility for discrimination of risk as well as allowing estimates of the prevalence of more rare important alternative findings in each group.

Enrolled patients had all relevant factors (listed in online Appendix 1) from the derivation phase recorded by the research associate prior to results of CT being known. Research associates were not aware of the elements of the S.T.O.N.E. score when prospective data was collected. Point values of 0-13 and category of risk were assigned to each subject in the validation cohort blinded to the CT result, and the CT result was categorized blinded to the clinical factors (except laterality of pain) and point total. Hosmer-Lemeshow test and discrimination (AUC) with AUC point estimates and 95% confidence intervals were estimated by bootstrapping.

Phase 3 methods:

Prospective interventional study with follow-up. Consecutive consenting adult patients undergoing CT for suspected renal colic were enrolled. Likelihood of ureteral stone was stratified as “high” (>90%), “moderate” (~50%), or “low” (<10%) using our previously validated S.T.O.N.E. score. All subjects with a “high” score underwent ULDCT, and subjects with a “moderate” score could undergo ULDCT at clinician discretion. A priori the identification of any ureteral stone on the ULDCT report was considered adequate to identify possible need for urologic intervention, and specific etiology (e.g. appendicitis) for non-urologic intervention; this was abstracted blinded to follow-up data. Subjects underwent phone follow-up (blinded to CT results) to determine if any surgical intervention (urologic or non-urologic) was performed within 90 days.

5. **Results** (Principal Findings, Outcomes, Discussion, Conclusions, Significance, Implications).

Derivation sample

Of 5,383 CT FPPs performed in the emergency departments on patients 18 years of age or older during the retrospective period, 1,853 (34.4%) were randomly selected for full record review. Of these, there were 1,040 complete records that had no exclusion criteria (enrollment diagram, online figure 1). Demographics are shown in table 1. Approximately half (49.5%; 515 of 1040) of the patients had a symptomatic ureteral stone on their CT, while 2.9% (30 out of 1040) had an acutely important alternative causes of symptoms. Inter-rater reliability for categorization of CT result yielded a kappa of 0.75-.80, indicating "excellent" agreement. Factors that were significant for the presence or absence of ureteral stone on univariate analysis are shown in table 2.

The S.T.O.N.E. score[1]

Multivariate analysis yielded five factors that were most significantly associated with the presence of a ureteral stone: male gender, acute onset of pain, non-black race, presence of nausea or vomiting, and microscopic hematuria (Table 3). Prior ED visits were also significantly associated with lower likelihood of ureteral stone, but were not included in the model in order to maximize generalizability between centers. These five factors were incorporated into the "S.T.O.N.E. Score" with associated integral point values as shown in table 3, yielding a total score ranging from 0-13. The multivariate logistic regression model had a misclassification rate of 0.229 (95% CI 0.224 to 0.234) and an AUC of 0.857 (95% CI 0.786 to 0.929) while the "S.T.O.N.E. Score" had a misclassification rate of 0.229 (95% CI 0.224 to 0.234) and an AUC of 0.820 (95% CI 0.741 to 0.900). Agreement between the risk estimates based on the "S.T.O.N.E. Score" and those based on the multivariate logistic regression model demonstrated a weighted kappa of 0.87 (95% CI 0.86 to 0.87), indicating minimal loss of accuracy by assigning integral points to the factors.

Prospective validation

From May 25th 2012 to January 24th 2013 491 patients without exclusion criteria were enrolled (enrollment diagram, online figure 2). Demographics of subjects not approached were not significantly different from those that were not approached (Table 1). For the validation cohort, the S.T.O.N.E. Score grouped into 3 levels of risk had an AUC of 0.792 (95% CI 0.756 to 0.828) and the Hosmer-Lemeshow X^2 = 1.95 was not significant (p = 0.38), indicating good discrimination and calibration.

Comparison of derivation and validation sets

In the derivation and validation sets there were (respectively) 19.8% and 15.5% of patients that fell in the "low" group, 49.6% and 46.8% classified as "moderate", and 30.6% and 37.7% stratified as "high" likelihood of kidney stone. The prevalence of ureteral stone by group in the derivation and validation sets was (respectively) 8.3% and 9.2% in the "low" group, 51.6% and 51.3% in the "moderate" group, and 89.6% and 88.6% in the "high" group (Figure 1). Overall, acutely important alternative causes of symptoms were found on CT scan in 2.9% and 3.7% of the derivation and validation cohorts, with acutely important alternative causes in 0.3% and 1.6% of the "high" group, respectively. Causes and frequency of acutely important alternative findings in the overall derivation and validation sets are shown in Table 4.

Phase 3 results:[7]

To date 173 subjects have been enrolled, 56% are male. S.T.O.N.E. score was "high" in 49 (28.3%), "moderate" in 86 (49.7%), and "low" in 38 (22.0%). ULDCT was performed in 114 (65.9%) of subjects. Ureteral stone was identified in 44 (89.8%), 49 (57.0%) and 4 (10.5%) of high, moderate, and low STONE score groups respectively. Mean dose length product (DLP) was 100.8 +/- 40.9 mGy-cm in those receiving ULDCT and 884.7 +/- 397.3 in those receiving regular dose CT, representing a mean dose reduction of 88.6%. To date follow-up is complete on 73 patients; 14 had urologic intervention for ureteral stone. ULDCT alone was performed in 9 of 14 and detected a ureteral stone in all 9, yielding a sensitivity of 100% (95% CI 70-100%). There were 6 acutely important alternate findings on CT, with ULDCT done in 3 of 6 and findings seen on ULDCT in all 3. There were 2 non-urologic surgical interventions, both in subjects undergoing regular CT.

DISCUSSION

Principal findings

To our knowledge, this is the first clinical scoring system to be derived and validated for prediction of uncomplicated ureteral stone in ED patients in whom CT imaging is deemed indicated. Our data show that the quantitative effects of the five factors incorporated into the "S.T.O.N.E. Score" can accurately predict ureteral stone and allow stratification of ED patients with suspected kidney stone into one of three groups: low (~10% or less chance of stone), moderate (~50% chance of stone), and high (~90% chance of stone).

Additionally, we found that the likelihood of an acutely important alternative finding is inversely proportional to the likelihood of a ureteral stone as predicted by the S.T.O.N.E. score. While the overall presence of acutely important alternative findings was 2.9% in the derivation set and 3.8% in the validation set, the prevalence of clinically important alternative diagnoses in the "high" stone probability group was less than half of this: 0.3% and 1.6% in the derivation and validation cohorts.

Clinical and policy implications

In deriving and validating this clinical prediction rule (rather than a decision rule), we are not necessarily stating that patients with a high stone score should not undergo any CT imaging at all - though this may not be an unreasonable approach in certain situations. In any clinical situation the risk of a test (in this case from radiation) and the resources required to do the test will need to be balanced against the tolerance for uncertainty and risk of misdiagnosis on the part of both the clinician and the patient. In some patients - perhaps particularly younger ones who are more susceptible to radiation and less likely to have certain diagnoses such as diverticulitis, aortic pathology, or malignancy - this score may be used to provide objective data to help balance the cost and risk of performing a CT at all. The other possibility is that this clinical prediction rule could be used to determine which patients may be most appropriate for a substantially reduced dose CT, which has been shown to reliably identify ureteral stones, particularly large ones that may require intervention.

CT use in the United States and public health implications

Since the landmark paper by Smith et al. in 1996, CT scanning has become the first-line test for kidney stone in the United States. However, despite a 10-fold increase in the utilization of CT scanning for diagnosis of kidney stone from 1996-2007, the proportion of patients diagnosed with kidney stone, findings of significant alternative diagnoses, or hospital admission has not changed. This suggests that the increase in CT use for diagnosis of this condition may not be substantially improving patient-centered outcomes. Outside of the United States, CT scanning is not necessarily the first line test for suspected kidney stone. In 2011 the European Urology Association released comprehensive guidelines on urolithiasis in which they state that “ultrasonography should be used as the primary procedure”. In 2007, the yearly rate of CT scanning in the United States was nearly 228 per 1,000 persons – more than double the rate in Canada and nearly four times the rate in the United Kingdom. These data are not specific to imaging in kidney stones and do not include patient outcomes, but the presence of wide regional variation (particularly in a condition that is not life-threatening) suggests an opportunity for more appropriate utilization.

While the health risk attributable to a single CT is small, in a country of 310 million people (approximate United States population) it is important to note a lifetime incidence of nephrolithiasis of approximately 10%. If half of these people undergo a CT to detect it (likely a conservative estimate as kidney stones are frequently recurrent and many patients undergo multiple CTs), we could expect 15 million CTs to be performed on current United States residents. In addition to the cost of this imaging, it could be estimated that exposure to ionizing radiation from CT could cause between 10,000 and 30,000 additional malignancies (using risk estimates of between 1 in 500 and 1 in 1500 for renal colic CT scans).

Of note, in this setting CT was performed nearly as often in women as in men in both phases of the study (48.1% of CTs in women in the derivation phase; 44.4% in the validation phase). However, the diagnostic yield (percentage of patients with ureteral stones on CT) for men was much higher: 68.8% in the derivation phase and 66.7% in the validation phase compared to women (28.7% in the derivation phase and 41.7% in the validation phase). The lower diagnostic yield in women coupled with a higher risk from radiation of the pelvis with CT suggests that women (especially younger women) may be a group that could benefit from more judicious use of CT radiation.

Use of the score to select appropriate patients for reduced dose CT or ultrasound

In terms of potential clinical utility, if a CT scan is being considered for suspected kidney stone and a patient has a “high” S.T.O.N.E. Score (which occurred in about a third of patients: 30.6% in the derivation cohort and 37.7% in the validation cohort), then the patient is very likely to have a kidney stone and very unlikely to have an important non-kidney stone cause of symptoms. Thus, if the S.T.O.N.E. Score is high a CT might be avoided entirely or a reduced dose CT could be performed (to ensure there is not a large stone that may require intervention). It is important to note that it is still possible to miss an important alternative diagnosis in the "high" group if CT is not performed (of the ~10% of patients in the "high" group about 10% of these or 1-2% of the overall group had an important alternative finding), However, the S.T.O.N.E. score thus offers objective data to both the clinician and patient that could help guide shared decision-making about CT scanning, which is not without risk in terms of radiation and incidental findings that may lead to further testing or intervention. Our hope is that this score can be incorporated into imaging decisions for suspected renal colic to decrease radiation exposure and reduce imaging over-utilization (i.e. imaging without improvement in patient care). Further investigation, potentially including a randomized trial may help to elucidate this.

The majority of kidney stones (smaller stones, about 80% in this study as is generally the case) will pass spontaneously with symptomatic treatment. Patients with a very high likelihood of ureteral stone thus may not require any imaging at all and could be managed with pain control and medication to enhance stone expulsion, with definitive diagnosis using a urine strainer. However, clinicians may still want to perform a CT to exclude potentially serious alternative causes of symptoms and to determine the size and location of a stone, if present (with implications for prognosis and intervention). In this case, patients with a high S.T.O.N.E. score may be ideally suited for substantially reduced dose CT scanning. While data on low dose protocols has been published outside of the United States. and the American College of Radiology states reduced-dose techniques are "preferred", data from the Dose Imaging Registry (part of the American College of Radiology National Radiology of Data Registry: www. nrdr.acr.org) indicates that the mean institutional dose for CT for renal colic is still greater than 10mSv, and reduced dose techniques are rarely used in U.S. hospitals (manuscript in press).

Reduced dose CT has been shown to be accurate for kidney stones, particularly larger ones that may require intervention, but has not been widely used in the US, likely because of concerns about accuracy in an unselected population. Reluctance to implement reduced dose CT protocols for renal colic may be due to fear of missing other pathology. An investigator looking at reduced dose CT for renal colic noted that to put these reduced dose protocols into practice they "would want to target it at patients who have a high pretest probability of calculi and obstructive uropathy, since the ability to detect other pathology is hindered." In addition to predicting kidney stone, our data show that the group that is most likely to have kidney stones is also very unlikely (<2%) to have an important alternative cause of symptoms. A likelihood of disease under 2% has been identified as a testing threshold (point at which the negatives of a test outweigh the positives) for CT use in detecting other significant diseases, such as pulmonary embolism. Identifying patients in this group could safely direct some patients with suspected kidney stone to low- or ultra-low dose CT.

Ultrasound is another option that may be used for imaging in suspected renal colic, and this is often a first line test outside of the United States. Ultrasound has the advantage of avoiding radiation entirely, and is sometimes definitively diagnostic: identifying the presence, size and location of a symptomatic kidney stone. Often, however, ultrasound may show indirect evidence of obstruction (hydronephrosis) without visualizing the actual ureteral stone, which may be obscured by bowel. We did find the presence of hydronephrosis on CT to be highly predictive of ureteral stone, and future work will incorporate the presence of hydronephrosis on ultrasound into the S.T.O.N.E. score.

At our institution, the S.T.O.N.E. score has been incorporated into the computerized physician order entry system (Epic, Verona WI). When a clinician orders a CT for kidney stone the questions are asked and a S.T.O.N.E. score with risk category accompanies the radiology order. This has been welcomed by the radiologists who were often unsure of the perceived likelihood of kidney stone on the part of the ordering physician. We have found that the S.T.O.N.E. score is easily entered and calculated using our electronic health record. We are also currently using the S.T.O.N.E. score in a prospective study to select patients who are appropriate for either expectant management (no CT) or an "ultra-low dose CT", with a radiation dose that is about 90% lower than conventional CT (effective dose of ~1mSv, about that of a plain abdominal radiograph). On a population basis, assuming the no threshold linear model suggested by Bier VII, an equivalent reduction in cancer risk could be expected. Current average effective dose of CT in the U.S. is 11.2mSv, with only 2% of CTs done using "low dose".[5]

Strengths and limitations of this study

An important limitation of this study is that gestalt clinician pre-test probability for kidney stone has not been thoroughly investigated and it is possible that it would perform similarly to an objective clinical prediction rule. A study by Abramson et al. showed that the pre-test probability of emergency department physicians obtaining CT for suspected kidney stone clustered in the 41-60% and 71-90% ranges. However, the use of a relatively objective scoring system has the advantage that it is not dependent on clinician experience. In pulmonary embolism, for example, while gestalt pre-test probability has been shown to be reasonably accurate, authors comparing gestalt pre-test probability to objective scoring systems conclude that they “advocate the use of a clinical prediction rule because it has been shown to be accurate and can be used by less-experienced clinicians”. This study is also limited by being derived and validated in the same clinical setting; it is not known how well it would perform in other settings.

Conclusion

In conclusion, we have derived and validated a clinical prediction score for the presence of symptomatic ureteral stone. Multi-center validation and evaluation of incorporating the S.T.O.N.E. score into imaging algorithms is warranted.

In phase 3 we achieved a dose reduction of nearly 90% while maintaining sensitivity of imaging to detect conditions potentially requiring intervention.

Our future directions include a submission to the AHRQ for a dissemination grant based on these findings.

6. List of Publications and Products

1. Moore CL, Bomann S, Daniels B, et al. Derivation and validation of a clinical prediction rule for uncomplicated ureteral stone--the STONE score: retrospective and prospective observational cohort studies. BMJ 2014;348; March 26, 2014 doi: 10.1136/bmj.g2191.

2. Moore CL, Daniels B, Singh D, Luty S, Molinaro A. Prevalence and clinical importance of alternative causes of symptoms using a renal colic computed tomography protocol in patients with flank or back pain and absence of pyuria. Acad Emerg Med 2013;20(5):470–8.

3. Herbst MK, Rosenberg G, Daniels B, et al. Effect of Provider Experience on Clinician-Performed Ultrasonography for Hydronephrosis in Patients With Suspected Renal Colic. Ann Emerg Med 2014;64:269–76.

4. Daniels B, Gross CP, Singh D, Luty S, Moore CL. Addition of Renal Point-of-Care Limited Ultrasound Improves a Clinical Prediction Score for Uncomplicated Ureteral Stones in Emergency Department Patients with Suspected Renal Colic: The S.T.O.N.E. PLUS Score. Acad Emerg Med 2014;21(5):S189.

5. Lukasiewicz A, Bhargavan-Chatfield M, Coombs L, et al. Radiation Dose Index of Renal Colic Protocol CT Studies in the United States: A Report from the American College of Radiology National Radiology Data Registry. Radiology 2014;271(2):445–51.

6. Moore CL, Daniels B, Ghita M, et al. Accuracy of Reduced-Dose Computed Tomography for Ureteral Stones in Emergency Department Patients. Ann Emerg Med 2014; ePub ahead of print doi: 10.1016/j.annemergmed.2014.09.

7. Moore C, Gunabushanam G, Daniels B, et al. Implementation of an Ultra-Low Dose CT Protocol for ED Patients with Suspected Kidney Stone. Acad Emerg Med 2014;21(5):S302–S303.