Comparative Effectiveness of Shock Wave Lithotripsy and Ureteroscopy for Treating Patients With Kidney Stones

Charles D. Scales Jr, MD, MSHS; Julie C. Lai, MPH; Andrew W. Dick, PhD; Jan M. Hanley, MS; Jeroen van Meijgaard, PhD; Claude M. Setodji, PhD; Christopher S. Saigal, MD, MPH; for the Urologic Diseases in America Project

IMPORTANCE Shock wave lithotripsy (SWL) and ureteroscopy (URS) account for more than 90% of procedural interventions for kidney stones, which affect 1 in 11 persons in the United States. Efficacy data for SWL are more than 20 years old. Advances in URS, along with emerging evidence of reduced efficacy of modern lithotripters, have created uncertainty regarding the comparative effectiveness of these 2 treatment options.

OBJECTIVE To compare the effectiveness of SWL and URS to fragment or remove urinary stones in a large private payer cohort.

DESIGN, SETTING, AND PARTICIPANTS We performed a retrospective cohort study of privately insured beneficiaries who had an emergency department visit for a kidney stone and subsequently underwent SWL or URS. Using an instrumental variable approach to control for observed and unobserved differences between the 2 groups, we created a bivariate probit model to estimate the probability of repeat intervention following an initial procedure.

MAIN OUTCOMES AND MEASURES A second procedure (SWL or URS) within 120 days of an initial intervention to fragment or remove a kidney stone.

RESULTS Following an acute care visit for a kidney stone, 21,937 patients (45.8%) underwent SWL and 25,914 patients (54.2%) underwent URS to fragment or remove the stone. After the initial URS, 4,852 patients (18.7%) underwent an additional fragmentation or removal procedure compared with 5,186 patients (23.6%) after the initial SWL (P < .001). After adjusting for observed and unobserved variables, the estimated probabilities of repeat intervention were 11.0% (95% CI, 10.9-11.1) following SWL and 0.327% (95% CI, 0.325-0.329) following URS.

CONCLUSIONS AND RELEVANCE Among privately insured beneficiaries requiring procedural intervention to remove a symptomatic stone, repeat intervention is more likely following SWL. For the marginal patient (as opposed to the average patient), the probability of repeat intervention is substantially higher.

Published online May 16, 2014.
Kidney stones impose a significant burden of disease in the United States. The prevalence of kidney stones has increased dramatically since 1976, a change that is likely driven by the obesity epidemic.\textsuperscript{1,2} Stone disease affects approximately 1 of every 11 persons in the United States, a prevalence similar to that of diabetes mellitus.\textsuperscript{3} Up to 50\% of individuals forming stones will have a recurrence within 5 years.\textsuperscript{4} Dietary and lifestyle factors contribute substantially to the risk of stone disease, and emerging physiological data suggest that stone disease should be considered a metabolic disorder, punctuated by attacks of periodically formed, symptomatic kidney stones.\textsuperscript{5-8}

The costs to society of kidney stones are significant. In terms of aggregate annual medical expenditures, kidney stones represent one of the most costly urologic conditions, with greater than $10 billion in expenditures in 2006 for treating patients with kidney stones.\textsuperscript{9,10} Indirect costs such as work loss and temporary disability are also an important contribution to the disease burden, particularly because kidney stones affect a largely working-age population.\textsuperscript{11} Among the important cost drivers of kidney stone disease are procedural interventions to treat patients with symptomatic stones.

Fragmentation and removal of symptomatic stones relieves pain and prevents harm to renal function from chronic obstruction. Two dominant modalities exist for procedural intervention, namely, shock wave lithotripsy (SWL) and ureteroscopy (URS). Shock wave lithotripsy is noninvasive and uses high-energy acoustic waves to fragment stones. Ureteroscopy is a minimally invasive endoscopic technique that can access all parts of the ureter and renal collecting system, typically using a laser to fragment stones. Together, these 2 modalities represent more than 90\% of the procedures performed to remove renal and ureteral stones in the United States.\textsuperscript{12} Both are considered first-line options for the management of symptomatic stones, although SWL may have been slightly less efficient than URS in published series.\textsuperscript{13,14} More recently developed lithotripters seem to be less effective than the original device (modified HM3 lithotripter, Dornier MedTech), which was used to generate most of the clinical trial data comparing URS and SWL.\textsuperscript{15,16} In addition, ureteroscopic technology has experienced important advances during the past decade.\textsuperscript{17-19}

Reducing waste and avoiding retreatment are important factors to drive value in the provision of health care.\textsuperscript{20,21} Prior analyses of clinical trial data suggest that a trade-off for the completely noninvasive nature of SWL is a greater need for retreatment.\textsuperscript{13,14,22} In addition, claims data suggest that unadjusted retreatment rates for SWL can be up to triple those of URS.\textsuperscript{23} However, that analysis did not account for important clinical factors such as stone size and location, nor did it incorporate other potential confounders such as patient preferences.

Given that both technologies are well established and only incrementally evolving, it is unlikely that new randomized clinical trials comparing the 2 will occur. In addition, creating a large registry of URS and SWL procedures, along with data regarding patient preferences, to understand the comparative effectiveness of these technologies would be time-consuming and expensive. Therefore, we performed an administrative claims-based analysis using econometric techniques to compare repeat interventions among patients initially treated with SWL or URS after adjusting for important unmeasured confounders such as stone size and location and patient preferences. The use of appropriate econometric techniques such as instrumental variable analysis seeks to balance observed and unobserved confounders among treatment groups using naturally occurring variation in observational data.\textsuperscript{23-25} Through the use of a bivariate probit model in a large cross-section of commercially insured beneficiaries, we sought to compare the effectiveness of SWL vs URS in the treatment of patients with urinary stones.

Methods

Data Source

The institutional review board of the RAND Corporation determined that the study design was exempt from the requirement for review given the nature of the data. To identify patients who were likely to have a symptomatic stone, we used data from the MarketScan Commercial Claims and Encounters data set (http://www.bridgetodata.org/node/987) to select beneficiaries with an emergency department or urgent care center clinical encounter (the index encounter) for a kidney or ureteral stone between January 1, 2002, and December 31, 2010. The MarketScan data include encounter-level health claims for clinical services and pharmacy claims. Although they are not nationally representative, these data represent a convenience sample of commercially insured beneficiaries in the United States.

Study Cohort

We further restricted our cohort of patients receiving emergent or urgent care to those who had no claim for any clinic visit or procedure for urinary stones within 12 months before the index encounter and who had at least 8 months of continuous enrollment following the index encounter. Finally, we limited our cohort to those beneficiaries who underwent URS or SWL within 4 months of the index encounter. These exclusion criteria resulted in a study sample of 47 851 patients who underwent treatment to fragment or remove a presumably symptomatic urinary stone.

Treatment Identification and Patient Covariates

We identified procedures for treating patients with kidney or ureteral stones using Current Procedural Terminology codes (eTable 1 in the Supplement).\textsuperscript{26} We created a binary treatment variable for SWL vs URS.

Demographic characteristics for each patient were identified from MarketScan data. These included age and sex; patient race/ethnicity is not included in this data set. The year of treatment was included to address secular trends in practice during the 9-year study period. Preexisting comorbid conditions were summarized by calculating the Charlson Comorbidity Index using claims submitted within 365 days before the index encounter. Geographic variation in the epidemiology of
stone disease exists, as well as variation in practice patterns. Therefore, we included census regions, as reported in the MarketScan data.

Outcome Measures
The primary outcome was a second procedure to fragment or remove stones within 120 days of the initial intervention. We elected a window of 120 days to capture potential second procedures performed outside of the global period for these procedures. A second URS or SWL was considered an outcome event, regardless of which initial procedure the patient underwent.

Statistical Analysis
We used the χ² test and t test, as appropriate, to compare patient-level covariates between procedures. The χ² test was also used to compare the proportions of patients undergoing a second procedure between SWL and URS.

An important consideration in any observational analysis is the potential for estimation bias due to unmeasured variables that can confound the results and limit causal inference. In particular, if unobserved factors that affect the treatment decision also influence the outcomes, standard estimates of treatment effects might be biased. For example, the size and location of urinary stones are important determinants of treatment (SWL vs URS) and outcomes. Therefore, treatment effects may be biased by the lack of equivalence of the unobserved factors. Previous claims-based comparisons of SWL and URS have failed to address this potential confounding.

We adopted an alternative approach in which we specify and estimate a simultaneous equations model that addresses this concern. Our approach, which is analogous to instrumental variable analysis, used an econometric technique designed to balance the effect of measured and unmeasured confounders among treatment groups and, if specified correctly, allows for causal inference of treatment effects.

The validity of our approach is contingent on the availability of instrumental variables that satisfy 2 conditions. First, they are strongly correlated with treatment decisions. Second, conditional on other observable factors, they are only related to the outcome (in our case, a second intervention to remove stones) through their influence on treatments.

We identified 3 potential instruments for the analysis. We hypothesized that the distance to a hospital with SWL vs URS capability would influence the probability of patients receiving one of the treatments and that this is the only mechanism by which the distance to a hospital with SWL vs URS capability would influence the receipt of a second procedure within 120 days of the initial intervention. Similar distances have been used effectively in other instrumental variable analyses.

We hypothesized that the per capita density of urologists and surgeons would reflect the availability of surgical services at the local area level, as well as the intensity of potential competition among providers, and that these factors would influence the probability of patients receiving one of the treatments (eMethods and eTable2 in the Supplement).

Given that the regression coefficients of probit models are not intuitive to interpret, we then calculated model-derived probabilities of repeat treatment for ease of interpretation. As a sensitivity analysis, we constructed separate models for patients treated from 2002 to 2005 and from 2006 to 2010. Statistical testing was 2-sided, with the type I error rate set at 5%. Analyses were performed using commercial software (SAS version 9.2; SAS Institute Inc and STATA version 11.0; StataCorp LP).

Results
The cohort included 47 851 individuals who underwent an initial procedure for fragmentation or removal of a urinary stone. Among these, 25 914 (54.2%) underwent URS as the first procedure and 21 937 (45.8%) underwent SWL as the first procedure. The individuals undergoing SWL were slightly older than those undergoing URS (Table 1). Women were more likely to undergo URS as the first procedure (41.3% in the URS group vs 37.6% in the SWL group, P < .001). Minor regional variation in the use of the 2 procedures existed. Most patients had a Charlson Comorbidity Index of 0.

Within 120 days of the initial procedure, approximately 1 in 5 individuals (10 038 of 47 851 [21.0%]) underwent an additional procedure to fragment or remove urinary stones. After the initial URS, 4852 patients (18.7%) underwent an additional fragmentation or removal procedure compared with 5186 patients (23.6%) after the initial SWL (P < .001). Those undergoing initial SWL were more likely to require a second procedure than those undergoing initial URS (Table 2). The relative risk of a second procedure following SWL was approximately 25% higher than after URS (relative risk, 1.26; P < .001). On multivariate analysis (naïve model), differences in the proportions of individuals undergoing additional fragmentation or removal procedures persisted after adjusting for observable confounders such as age, sex, and census region. Adjustment for these potential confounders did not substantively change the predicted probability of additional treatments.

We then used a 2-equation bivariate probit model to adjust for observed and unobserved potential confounders. The likelihood ratio test of the 2-equation model demonstrated that the results were statistically significant (P < .001). The 2-equation model substantially reduced the predicted probability of additional treatments for SWL and URS (Table 2). After adjusting for observed and unobserved potential confounders, the predicted probability of an additional procedure for the marginal patient (as opposed to the average patient) undergoing SWL was 11.0%. In contrast, after adjusting for observed and unobserved potential confounders, the predicted probability of an additional procedure for the marginal patient undergoing URS was 0.3%.

Discussion
In this novel analysis, we use an econometric approach to compare the effectiveness of SWL and URS. Our findings demonstrate that, on an unadjusted basis and after adjusting for observable potential confounders, the relative risk of an additional
stone fragmentation or removal procedure is approximately 25% higher following SWL compared with URS. After using a bivariate probit model to address unobserved potential confounders (eg, stone size), the probability of additional procedures decreases for SWL and URS, but the marginal patient undergoing SWL is substantially more likely to experience an additional procedure to remove a previously treated stone compared with the same patient undergoing URS.

The efficacy of SWL and URS for fragmenting and removing urinary stones is well established. For ureteral stones, SWL and URS are considered first-line therapy. However, for ureteral stones exceeding 10 mm, ureteroscopic fragmentation generally results in higher stone-free rates with fewer procedures. For renal stones 10 mm or less, SWL and URS have similar efficacy in clinical trials.

In contrast to abundant efficacy data, the comparative effectiveness of SWL and URS outside of controlled clinical trials is poorly documented. One claims-based cost analysis suggested that the mean number of procedures per patient is slightly higher for SWL (1.22 vs 1.12 for URS). Prior work from the Urologic Diseases in America Project examined practice patterns for urinary stone fragmentation and removal procedures among unselected Medicare beneficiaries. In a longitudinal cohort, 38% of Medicare beneficiaries undergoing SWL had an additional stone fragmentation or removal procedure within 120 days compared with 12% of those undergoing URS. Most important, neither of these analyses adjusted for important covariates such as stone size and location, nor did they focus on patients with symptomatic stones (as opposed to patients potentially treated for incidentally detected stones).

Our findings suggest that, for the marginal patient with a symptomatic stone amenable to SWL or URS, first-line treatment with SWL would result in approximately 11% of patients undergoing an additional stone removal procedure. By comparison, the model results suggest that the same patient undergoing first-line treatment with URS faces less than a 1% chance of an additional stone removal procedure. Most important, our analytic approach should control for observable factors (eg, patient age, sex, comorbid conditions, and census region) and factors that we could not observe in our analysis (eg, patient obesity, race/ethnicity, size and location of the targeted stone, ureteral stent placement, bilateral stones, and patient preference). This approach strengthens the causal inference of our findings.

The contrast between efficacy data and comparative effectiveness data for these technologies is noteworthy. Some of this difference could be due to evolving technical capabilities of the 2 procedures. Much of the efficacy data for SWL was generated using the original lithotripter (modified HM3) from more than 20 years ago. Newer lithotripter models have smaller focal zones and deliver shock waves at faster rates.

Table 2. Estimated Probability of a Second Procedure, by Adjustment Method

<table>
<thead>
<tr>
<th>Effect</th>
<th>Retreatment Probability, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Shock Wave Lithotripsy</td>
</tr>
<tr>
<td></td>
<td>Unadjusted</td>
</tr>
<tr>
<td></td>
<td>Naive model, adjusting for observed confounders*</td>
</tr>
<tr>
<td></td>
<td>Bivariate probit model, adjusting for observed and unobserved confounders*</td>
</tr>
</tbody>
</table>

* Adjusted for age, sex, Charlson Comorbidity Index, census region, and year of treatment.

Table 1. Characteristics of the Study Cohort

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Ureteroscopy (n = 25 914)</th>
<th>Shock Wave Lithotripsy (n = 21 937)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>42.5 (11.8)</td>
<td>43.5 (11.1)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Female sex, No. (%)</td>
<td>10 715 (41.3)</td>
<td>8252 (37.6)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Charlson Comorbidity Index, No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>23 030 (88.9)</td>
<td>19 520 (89.0)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2430 (9.4)</td>
<td>2063 (9.4)</td>
<td>.50</td>
</tr>
<tr>
<td>≥2</td>
<td>454 (1.8)</td>
<td>354 (1.6)</td>
<td></td>
</tr>
<tr>
<td>Census region, No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>1427 (5.5)</td>
<td>1530 (7.0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>North Central</td>
<td>8221 (31.7)</td>
<td>6048 (27.6)</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>12 993 (50.1)</td>
<td>11 326 (51.6)</td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>3273 (12.6)</td>
<td>3032 (13.8)</td>
<td></td>
</tr>
<tr>
<td>Year of treatment, No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>1264 (4.9)</td>
<td>1127 (5.1)</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>1866 (7.2)</td>
<td>1674 (7.6)</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>2160 (8.3)</td>
<td>1931 (8.8)</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>2218 (8.6)</td>
<td>1937 (8.8)</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>3614 (13.9)</td>
<td>3179 (14.5)</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>4254 (16.4)</td>
<td>3665 (16.7)</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>5324 (20.5)</td>
<td>4335 (19.8)</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>5214 (20.1)</td>
<td>4089 (18.6)</td>
<td></td>
</tr>
</tbody>
</table>
which is believed to decrease the efficiency of stone fragmentation.\textsuperscript{15} At the same time, substantial technical progress has been made with endoscopic technology, which has improved the capability of URS.\textsuperscript{17-19} Technical limitations of modern lithotripters have stimulated intense interest in new, more efficient designs.\textsuperscript{34-35} Our findings provide additional support for the importance of engineering research to improve the technical capabilities of SWL.

Our results have important implications for counseling patients regarding treatment options. Kidney stones affect approximately 9% of the United States population,\textsuperscript{1} and up to half of the patients who develop a kidney stone will experience recurrence within 5 years.\textsuperscript{4} Many of these patients will undergo procedures to fragment or remove stones. A substantial proportion of patients undergoing SWL will require a second procedure to fragment or remove stones. However, SWL may be preferred by some patients because it is a noninvasive procedure.\textsuperscript{29} Given this trade-off, physicians should formally assess patient-centered outcomes and preferences and ideally develop a tool to facilitate shared decision making for those patients who require procedural intervention.

These findings also have important policy implications regarding payment structures for health care delivery. Under the current fee-for-service structure, the intervention most likely to require additional treatment (SWL) is reimbursed at a higher rate than the technology most likely to remove a stone in a single session (URS).\textsuperscript{36} Even when surgeon fees are limited by a global payment period, substantial ancillary costs (ie, anesthesia and facility payments) still pertain. This misaligned economic structure would seem to exacerbate the costs of treating patients with kidney stones. Not only are these costs growing rapidly, but kidney stones are one of the most costly urologic conditions from the perspective of aggregate health care expenditures.\textsuperscript{9-10} Shifting toward a payment structure based on an episode of care would more closely align payer and provider incentives and potentially reduce the overall costs of treating patients with kidney stones.

Our results should be considered in the context of several limitations. The study population was young, generally healthy, and privately insured. Therefore, our results may not generalize to older populations (eg, those covered by Medicare) or others without private insurance. The claims-based nature of the analysis prevented direct examination of the influence of important clinical factors such as obesity and stone size and location. Similarly, we were unable to incorporate patient preferences for SWL or URS into our investigation. However, assuming that the instrumental variables function appropriately, our analytic approach should balance these and other unobserved confounders. Finally, the study findings pertain to the marginal patient, rather than the outcomes of treatment for the average patient. Clinical experience likely represents an average picture of patient outcomes. Therefore, even when expressed as the probability of repeat treatment, the findings may not be intuitive from a clinical perspective. Conceptually, patients undergoing URS or SWL for stones comprise 3 groups, namely, those who always undergo URS (eg, for a 6-mm distal ureteral stone), those who always undergo SWL (eg, for a 10-mm intermittently symptomatic renal stone treated by an SWL specialist), and those who might undergo either procedure. The average outcome (ie, clinical experience) likely represents the comparison of all patients undergoing URS vs all patients undergoing SWL. In contrast, the marginal patient represents the comparison of only those individuals who might undergo either procedure.\textsuperscript{37}

**Conclusions**

These limitations notwithstanding, our findings retain significant validity and have important policy implications. Observed differences in the comparative effectiveness suggest that further exploration of additional outcomes (eg, unplanned postprocedural care, complication rates, and costs) will inform the comparison of outcomes for these procedures. Formal assessment of patient preferences, incorporating potential trade-offs in the nature and outcomes of these procedures, could reduce decisional conflict and improve patient satisfaction. Finally, efforts to improve the efficiency of modern lithotripters should continue to be supported.

**Funding/Support:** This study was supported by the Robert Wood Johnson Foundation Clinical Scholars Program and the US Department of Veterans Affairs (Dr Scales). The Urologic Diseases in America Project is supported by grant HHSN276201200016C from the National Institute of Diabetes and Digestive and Kidney Diseases and the National Library of Medicine.

**Role of the Sponsor:** The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation or review of the manuscript; or decision to submit for publication. The Robert Wood Johnson Foundation and US Department of Veterans Affairs had no role in the approval of the manuscript. The Executive Committee of the Urologic Diseases in America Project includes representatives from the National Institute of Diabetes and Digestive and Kidney Diseases; the Executive Committee approves

**Conflict of Interest Disclosures:** None reported.

**Author Contributions:** Robert Wood Johnson Foundation Clinical Scholars Program, University of California, Los Angeles, Los Angeles (Scales); Veterans Affairs Greater Los Angeles Healthcare System, US Department of Veterans Affairs, Los Angeles, California (Scales); Department of Urology, David Geffen School of Medicine, University of California, Los Angeles (Scales); Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles (Scales); currently with the Duke Clinical Research Institute, Division of Urologic Surgery, Duke University Medical Center, Durham, North Carolina (Scales); RAND Corporation, Santa Monica, California (Lai, Hanley, Saigal); RAND Corporation, Boston, Massachusetts (Dick); Department of Health Policy and Management, Fielding School of Public Health, University of California, Los Angeles (van Meijgaard), RAND Corporation, Pittsburgh, Pennsylvania (Setodji).

**Author Affiliations:** Robert Wood Johnson Foundation Clinical Scholars Program, University of California, Los Angeles, Los Angeles (Scales); Veterans Affairs Greater Los Angeles Healthcare System, US Department of Veterans Affairs, Los Angeles, California (Scales); Department of Urology, David Geffen School of Medicine, University of California, Los Angeles (Scales); Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles (Scales); currently with the Duke Clinical Research Institute, Division of Urologic Surgery, Duke University Medical Center, Durham, North Carolina (Scales); RAND Corporation, Santa Monica, California (Lai, Hanley, Saigal); RAND Corporation, Boston, Massachusetts (Dick); Department of Health Policy and Management, Fielding School of Public Health, University of California, Los Angeles (van Meijgaard), RAND Corporation, Pittsburgh, Pennsylvania (Setodji).

**Funding/Support:** This study was supported by the Robert Wood Johnson Foundation Clinical Scholars Program and the US Department of Veterans Affairs (Dr Scales). The Urologic Diseases in America Project is supported by grant HHSN276201200016C from the National Institute of Diabetes and Digestive and Kidney Diseases and the National Library of Medicine.

**Role of the Sponsor:** The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation or review of the manuscript; or decision to submit for publication. The Robert Wood Johnson Foundation and US Department of Veterans Affairs had no role in the approval of the manuscript. The Executive Committee of the Urologic Diseases in America Project includes representatives from the National Institute of Diabetes and Digestive and Kidney Diseases; the Executive Committee approves

**Conflict of Interest Disclosures:** None reported.
REFERENCES


