The Role of the 24-Hour Urine Collection in the Prevention of Kidney Stone Recurrence

Ryan S. Hsi,* Thomas Sanford, David S. Goldfarb and Marshall L. Stoller

From the Department of Urologic Surgery, Vanderbilt University School of Medicine (RSH), Nashville, Tennessee, Department of Urology, University of California-San Francisco (RSH, TS, MLSI), San Francisco, California, and Nephrology Section, New York Harbor Veterans Affairs Healthcare System and New York University Langone Medical Center, New York University School of Medicine (DSG), New York, New York

**Purpose:** Kidney stone prevention relies on the 24-hour urine collection to diagnose metabolic abnormalities and direct dietary and pharmacological therapy. While its use is guideline supported for high risk and interested patients, evidence that the test can accurately predict recurrence or treatment response is limited. We sought to critically reassess the role of the 24-hour urine collection in stone prevention.

**Materials and Methods:** In addition to a MEDLINE® search to identify controlled studies of dietary and pharmacological interventions, evidence supporting the AUA (American Urological Association) and EAU (European Association of Urology) guidelines for metabolic stone prevention were evaluated. Additionally, the placebo arms of these studies were examined to assess the stone clinic effect, that is the impact of regular office visits without specific treatment on stone recurrence.

**Results:** The 24-hour urine test has several limitations, including the complexity of interpretation, the need for repeat collections, the inability to predict stone recurrence with individual parameters and supersaturation values, the unclear rationale of laboratory cutoff values and the difficulty of determining collection adequacy. Only 1 prospective trial has compared selective dietary recommendations based on 24-hour urine collection results vs general dietary instructions. While the trial supported the intervention arm, significant limitations to the study were found. Placebo arms of intervention trials have noted a 0% to 61% decrease in stone recurrence rate and a remission rate during the study of 20% to 86%.

**Conclusions:** Whether all recurrent stone formers benefit from 24-hour urine collection has not been established. Additional comparative effectiveness trials are needed to determine which stone former benefits from selective therapy, as guided by the 24-hour urine collection.

**Key Words:** kidney, urolithiasis, secondary prevention, urine specimen collection, recurrence

The goal of metabolic management for kidney stones is to prevent recurrent stone episodes, which occur in up to 50% of individuals after 10 years.1 National guideline panels have recommended that a 24-hour urine collection should represent the specific metabolic evaluation to be performed in high risk stone formers.2,3 The rationale for testing is that...
identifying specific metabolic abnormalities in urine would enable the clinician to make individualized diet and pharmacological interventions to correct these abnormalities and reduce the recurrence risk. There is weighty reliance on this test to guide the metabolic stone management. According to AHRQ (Agency for Healthcare Research and Quality), additional studies are needed “to estimate the effectiveness, cost-effectiveness and harms of different kidney stone evaluation, treatment and followup strategies vs a control strategy to prevent stone recurrence.” Which stone formers should undergo this test, when and how often is unclear. We examine the evidence supporting the role of the 24-hour urine collection for the prevention of future stone events.

ROLE IN CONTEMPORARY PRACTICE
Both AUA and EAU guidelines recommend 24-hour urine collections in the high risk individual. AUA also recommends testing in the motivated first-time stone former. In the EAU guidelines, 2 consecutive collections are recommended initially, again after initiating dietary or pharmacological prevention for 8 to 12 weeks and every 12 months thereafter. AUA recommends 1 or 2 collections performed initially, an additional collection within 6 months of intervention to assess adherence and response, and yearly collection thereafter.

The utilization of the collection is increasingly viewed as a quality metric in nephrolithiasis care. In current practice, however, this test is uncommonly performed. Among privately insured Americans identified as high risk for stone recurrence, the prevalence of testing is only 7%. Of those tested, 16% with an abnormality in the initial collection undergo repeat collections within 6 months. Why overall use is so low is unclear but before broad quality metrics tied to reimbursement are implemented, more study is needed to determine who benefits most from this test.

24-HOUR URINE COLLECTION
Rationale
A systemic search was done in PubMed® for original publications involving adult kidney stone disease formers up to 2016. The search terms included were kidney stones, urolithiasis, nephrolithiasis, medical management, prevention, 24-hour urine, Litholink and randomized clinical trial. We reviewed relevant systemic reviews and clinical guidelines for studies relating to the usefulness of the 24-hour urine collection.

To date, there has been only 1 prospective trial evaluating the role of the 24-hour urine collection in the dietary prevention of stone recurrence. In a study by Kocvara et al, 242 idiopathic calcium based, first-time stone formers were evenly assigned to a group with a specific dietary regimen tailored by a 24-hour urine collection or to a second group in which general, nonselective dietary measures were begun. After 3 years, stone recurrence and/or growth occurred in 7% of the intervention group and in 23% of the nonselective group (p < 0.01).

While these results would appear to favor the 24-hour urine collection, several issues should be noted. The intervention group had regular follow-ups, including repeat collections and dietary adjustment, while the nonselective group had no followup. Urine collections were performed only in the intervention group and, except for an improvement in uricosuria, there was no appreciable improvement in any other urinary parameters. In fact, there were significantly higher mean levels of urinary calcium and oxalate at the end of the study compared to baseline in the selective therapy group. The difference in stone recurrence rates in the 2 groups may be explained more by the stone clinic effect, which has been documented to have a significant effect on recurrence.

Drawbacks
No laboratory test is perfect but the interpretation of the 24-hour urine collection has several issues (see Appendix). Often more than 1 abnormality is present, placing the clinician in a quandary about which abnormality to address. Borderline values may have clinical importance. Stone composition provides value in the role of medical management, especially if the composition is uric acid or cystine, for example. However, it is less helpful for the predominant calcium oxalate stone former in whom a 24-hour urine collection is performed.

The 24-hour urine collection negates diurnal and nocturnal variations in the urinary constituents related to diet and metabolism, so that often a single test is difficult to interpret. While performing 1 or 2 consecutive collections as part of the initial workup has been debated, variation is intrinsic to patient diets and activities, so that it remains uncertain whether doing more collections and making more diagnoses increases clinical value. Patients may prefer to do collections at home, on the weekends, while having different dietary habits during the week, or while at work or school, leading to further confusing variation.

The 24-hour urine collection is limited in its ability to predict recurrence and prognosticate risk. The definition of stone recurrence based on symptoms and/or radiographic findings has varied among different trials, which this contributes to the difficulty of relating 24-hour urine results to specific
hypercalciuria as an inclusion criterion. The baseline urinary citrate. The same is true for recurrent stone formers with normal range baseline urinary citrate. The same is true for several thiazide trials, which did not require hypercalciuria as an inclusion criterion. The only positive dietary intervention trial, that by Borghi et al, did not require more than hypercalciuria among eligible participants. Furthermore, the rates of metabolic abnormalities were similar among single stone formers compared to recurrent stone formers. Therefore, additional study is needed to further refine the analysis of urinary risk factors to predict stone recurrence.

Stone formers can have normal 24-hour urine collections and nonstone formers can have abnormal collections, raising the question of whether current laboratory cutoff parameters are appropriate. In a study of 5,942 geriatric (age greater than 65 years) stone formers, the rate of finding no metabolic abnormalities was more than 35%. In a separate analysis of 1,392 stone formers, after excluding low urine volume in 23% and infection in 2.5% as causes for stone formation, 1.1% had no metabolic abnormality identified. Curhan et al found that specifically among nonstone formers, the rates of hypercalciuria, hyperoxaluria, hyperuricosuria, hypocitruria and low urine volume were 14% to 27%, 7% to 43%, 8% to 40%, 3% to 9% and 7% to 20%, respectively. As they noted, the cut points for abnormal values are arbitrary with no rational basis for men and women to have different thresholds for the definition of hypercalciuria. Whether different cutoffs for defining urinary abnormalities should be used among different ethnic groups has also been debated.

The adequacy of the 24-hour urine collection, judged by the urinary creatinine-to-body weight ratio, is an imprecise measure. Generally, reference ranges of 15 to 20 mg/kg for women and 18 to 24 mg/kg for men are used to determine collection completeness. In a study of 381 initial collections, 51% were outside the reference range, 37% of patients had an “undercollection” and 14% had an “overcollection.” In a separate study of 1,502 patients with 24 or 48-hour collections, 51% of patients similarly submitted an inadequate sample based on the same definition. Often the clinician is faced with the uncomfortable situation of confronting the patient regarding the adequacy issue, which usually leads to repeat testing. Patients often acknowledge drinking more fluids and adhering more tightly to prescribed dietary regimens when performing the collections, in a gamesmanship that leads to further limits to the interpretation of results.

The relationship between individual 24-hour urinary abnormalities is poorly understood and undermines our understanding of how to interpret the test. The best example of this is the relationship of urinary citrate and pH in response to citrate supplementation. Urinary citrate inhibits calcium stone formation but also leads to bicarbonaturia and urinary alkalization. That latter effect is used successfully in the prevention of uric acid and cystine stones. However, some patients taking citrate supplements have more citruria and others have more alkalization. For example, in a study of 572 patients, the correlation between urinary citrate and pH was poor (r = −0.04, p = 0.36) and the finding persisted when controlled for age, gender, body weight, urinary volume and thiazide use. These findings suggest that additional factors beyond citrate supplementation and excretion influence urinary pH and interpretation of the results must take into account the complexity of the multiple, distinctly measured values.

It is unclear how to utilize calculated urinary supersaturation indexes in the 24-hour urine report. It has been shown that these supersaturation indexes correlate closely with stone type and this analysis has been advocated as a method to inform future stone risk and monitor treatment. However, in these studies individual supersaturation indexes were compared in isolation with stone type and not with the combination of multiple urinary values on the report, which often can have multiple abnormal values. Moreira et al studied 503 stone formers using stone composition data. While univariate comparisons of supersaturation and stone types showed strong associations, a multivariate model correctly predicted stone type in only 64% of cases. Additional study is needed on how to interpret urinary supersaturation indexes in the context of the other measured and calculated urinary parameters.

In addition, to our knowledge the degree to which reducing supersaturation for calcium stones is necessary to prevent new stones has not been tested in carefully controlled trials. A rule of thumb to decrease calcium oxalate supersaturation by 50% is not rigorously evidence-based. It is worth noting that the effect of citrate supplementation to prevent calcium stones is in part attributed to its ability to inhibit crystal aggregation and agglomeration, properties that are not reflected by reductions in supersaturation. Further complexity is demonstrated by the controversy about whether citrate prevents calcium phosphate stones. Anecdotal evidence suggests that citrate inhibits calcium phosphate stone recurrence, while the associated
bicarbonate and increased urine pH tend to decrease calcium phosphate supersaturation. Calculating supersaturation by alternative equations such as JESS (Joint Expert Speciation System) may lead to different interpretations of these effects.²⁷

An additional issue is that urine collections can be expensive, especially when accounting for repeat testing. Out-of-pocket costs for Litholink (Chicago, Illinois) are more than $400 per test at the full rate and less than $200 if discounted. Not all metabolic stone conditions require frequent monitoring. For example, the management of idiopathic uric acid nephrolithiasis does not necessitate 24-hour urine collections for pH monitoring. Spot urine pH testing is sufficient, inexpensive and easy.

We also noted that testing is not available in many less developed countries. Even in the United States, some municipal health care settings do not cover testing and with the large number of Americans lacking health care coverage, many individuals do not have access to testing. Therefore, it is incumbent on the kidney stone community to consider how to prevent stone recurrence when testing is not possible.

**Is Collection Needed for All High Risk Stone Formers?**

Not all stone formers need a complete metabolic evaluation, as the risk of recurrence is different for each individual. Among idiopathic calcium based stone formers, more than 50% will have only 1 recurrence in a lifetime and 10% will have more than 3 recurrences.²⁸ Placebo arms of the different intervention trials evaluating thiazides, allopurinol, citrate, phosphate and magnesium supplementation for recurrent calcium based stone formers have shown a 0% to 61% decrease in the stone recurrence rate (the rate of stone events after vs before intervention) and a remission rate (the absence of stone events and/or stone growth during the trial) of 20% to 86% (see Table). The stone clinic effect refers to regular clinic visits at which fluid and dietary reminders alone are associated with significantly decreased recurrence rates. Among 108 patients with single and mostly recurrent calcium based stones during 5 years, without pharmacological intervention 58% had no evidence of stone growth or new stone formation.⁸ Of all patients, 71% of those with hypercalciuria and 47% with hyperuricosuria did not form new stones.

**DISCUSSION AND FUTURE DIRECTIONS**

For a substantial population of stone formers, the 24-hour urine collection offers hope in revealing the specific etiology of the disease. The alternative of remaining ignorant of urinary chemistry is admittedly unsatisfactory to patients and their physicians. The test provides screening for certain metabolic conditions, such as primary hyperoxaluria and cystinuria. While it serves as the mainstay of the complete metabolic evaluation, there are has limitations to test interpretability. To our knowledge data from the test have not reliably been used to predict recurrence or prognosticate risk in the context of different dietary and pharmacological management strategies.

Sometimes no urinary abnormalities are found. The AUA guidelines committee has recognized the dilemma of the “normal” 24-hour urine collection in the recurrent calcium stone former with no metabolic abnormalities.² In guideline 17, which is listed as a Standard, it states that in these individuals, thiazides and/or potassium citrate should be considered for those with “normal” test results.² Perhaps one may question what value the test adds in such cases.

The stone clinic effect may reduce the real benefit of the 24-hour urine collection. For a large subset of recurrent stone formers, perhaps it is warranted to trial a period of nonselective dietary changes rather than repeat tests. Indeed, it is well established that high fluid intake²⁹ and a low salt diet³⁰ significantly decrease stone recurrence without the addition of pharmacological treatment.⁴ Fluid intake is low cost, accessible and safe with minimal side effects.

**Stone clinic effect shown by stone recurrence and remission rates in placebo arms in recurrent stone formers**

<table>
<thead>
<tr>
<th>Study</th>
<th>Control Regimen</th>
<th>Pretreatment On Study</th>
<th>% Formation Change</th>
<th>% Remission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coe FL: Ann Intern Med 1977; 87: 404</td>
<td>Fluids + diet</td>
<td>0.31</td>
<td>0.27</td>
<td>-13</td>
</tr>
<tr>
<td>Ettinger B et al: Am J Med 1979; 67: 245</td>
<td>Placebo + diet, diet</td>
<td>0.78, 0.57</td>
<td>0.33, 0.32</td>
<td>-58, -44</td>
</tr>
<tr>
<td>Johansson G et al. J Urol 1980; 124: 770</td>
<td>None</td>
<td>0.50</td>
<td>0.22</td>
<td>-56</td>
</tr>
<tr>
<td>Brooks P et al: Lancet 1981, 2: 124</td>
<td>Placebo</td>
<td>0.70</td>
<td>0.11</td>
<td>-84</td>
</tr>
<tr>
<td>Scholz D et al: J Urol 1982; 128: 303</td>
<td>Placebo</td>
<td>0.70</td>
<td>0.11</td>
<td>-84</td>
</tr>
<tr>
<td>Lærum and Larsen ²²</td>
<td>Placebo, diet + fluids</td>
<td>0.56</td>
<td>0.33</td>
<td>-41</td>
</tr>
<tr>
<td>Ettinger B et al: J Urol 1980; 124: 770</td>
<td>Placebo, diet + fluids</td>
<td>0.56</td>
<td>0.33</td>
<td>-41</td>
</tr>
<tr>
<td>Ettinger et al³⁰</td>
<td>Placebo</td>
<td>0.71</td>
<td>0.26</td>
<td>-63</td>
</tr>
<tr>
<td>Okhawa M et al: Br J Urol 1992; 69: 571</td>
<td>Placebo, diet + fluids</td>
<td>0.57</td>
<td>0.22</td>
<td>-61</td>
</tr>
<tr>
<td>Barcelo F et al, J Urol 1993; 150: 1781</td>
<td>Diet + fluids</td>
<td>Not applicable</td>
<td>0.31</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Hofbauer et al²⁸</td>
<td>Placebo</td>
<td>1.10</td>
<td>1.10</td>
<td>0</td>
</tr>
<tr>
<td>Ettinger et al³⁰</td>
<td>Placebo, diet + fluids</td>
<td>1.90</td>
<td>0.70</td>
<td>-61</td>
</tr>
<tr>
<td>Ettinger et al³⁰</td>
<td>Placebo + diet</td>
<td>0.57</td>
<td>0.27</td>
<td>-52</td>
</tr>
</tbody>
</table>

References

2. Lærum and Larsen
5. Barcelo F et al, J Urol 1993; 150: 1781
6. Hofbauer et al²⁸
7. Ettinger et al³⁰

Dochead: Adult Urology FLA 5.4.0 DTD ■ JURO14096_proof ■ 4 January 2017 ■ 4:16 pm ■ EO: JU-16-1644
It is appreciated that stone formation and growth are more than physical chemistry phenomena in the urinary milieu. The importance of trace elements, including heavy metals, must be further defined and potentially incorporated into the metabolic evaluation. The role of urinary proteins and their involvement in promoting or inhibiting matrix formation needs additional study. Tissue level factors also must be accounted for, as the role of Randall plaques in idiopathic calcium based stones is well accepted. Finally, the interpretation of the multitude of urinary values may fall under the umbrella of complex data necessitating computer modeling, extrapolation and simulation.

We also acknowledge that spot urine collections are not superior in providing clinically meaningful metabolic diagnoses or judging responses to therapy, except for monitoring urine pH in the management of uric acid stones. It should also be clear that the consequent diagnosis of primary hyperoxaluria is unlikely to be made without a 24-hour urine collection. Finally, the extent of dietary sodium ingestion, which is often occult and related to eating processed foods, is best assessed and communicated to patients by the 24-hour collection.

REFERENCES

16. Sawyer MD, Dietrich MS, Pickens RB et al: May require repeat testing.
28. Sawyer MD, Dietrich MS, Pickens RB et al: May require repeat testing.
34. Guideline supported Complexity in interpretation
35. Objective, quantifiable data Limited ability to predict recurrence
36. Can check compliance with fluids and medications Limited ability to predict response
37. Some view as convenient May require repeat testing
38. Limits diet or medication prescription to specific issues
39. Gives hope to lifelong disease Cost

APPENDIX

Pros and Cons of 24-Hour Urine Collection

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guideline supported</td>
<td>Complexity in interpretation</td>
</tr>
<tr>
<td>Objective, quantifiable data</td>
<td>Limited ability to predict recurrence</td>
</tr>
<tr>
<td>Can check compliance with fluids and medications</td>
<td>Limited ability to predict response</td>
</tr>
<tr>
<td>Some view as convenient</td>
<td>May require repeat testing</td>
</tr>
<tr>
<td>Limits diet or medication prescription to specific issues</td>
<td></td>
</tr>
<tr>
<td>Gives hope to lifelong disease</td>
<td>Cost</td>
</tr>
</tbody>
</table>

CONCLUSIONS

The metabolic evaluation of stone formers with 24-hour urine collections is recommended by current guidelines and is increasingly viewed as a quality metric in kidney stone care. However, the evidence demonstrating that treatment based on the test is superior to empirical or nonselected therapy is limited. The 24-hour urine collection is imperfect in predicting stone recurrence, stone composition or responses to treatment. Interpretation of the study is complex and often subjective. Determining which stone forming populations benefit most from the test and developing additional tools to determine recurrence risk are needed.


EDITORIAL COMMENT

The value of 24-hour urine collection is put under suspicion for urolithiasis diagnosis and recurrence. Parks et al noted that only a single 24-hour urine collection is not enough.1 This analysis has limitations, including more than 1 biochemical abnormality present, borderline values and weekend collection which could vary the diet. However, 2 consecutive collections decrease variability and make results trustworthy. Weekend diet differences are the same for stone formers compared to non-formers with a genetic component in the latter (forming stones is not a wish but a capacity).

Defining recurrence is difficult but countless papers show that correcting biochemical abnormalities decreases recurrence (reference 19 in article).2 Empirical treatment benefits with citrate potassium and/or thiazides are possible but followup studies to eventually make the diagnosis have not been previously reported. Diet and increasing fluids are of great value but not always sufficient, especially in hypercalciuric patients who require thiazides and increasing doses to achieve metabolic control according to followups.

Rodolfo Spivacow
Medical Department
Instituto de Investigaciones Metabólicas
Buenos Aires, Argentina

REFERENCES


REPLY BY AUTHORS

The perspective of the commentator is what precisely motivated our effort to challenge how we prescribe and interpret the 24-hour urine collection. Our intention was not to focus on whether performing 1 or 2 collections is better or which day of the week is optimal to perform this test. Rather, as a starting point, our intention was to generate a discussion on how we can better define and refine who should receive testing.

The stone clinic effect is known to be associated with a decrease in stone recurrence in the absence of 24-hour urine testing. Empirical drug therapy with potassium citrate and/or thiazides without testing may be as good as specific therapy guided by a 24-hour urine collection for the majority of idiopathic calcium stone formers.

The prognostic capacity of 24-hour urine collection values individually and in combination, along with supersaturation indexes, has not been rigorously studied. Until we perform the necessary prospective studies, we will not know the answers to these questions. Indeed, the 24-hour urine collection provides value for the care of many of our patients with kidney stone. We should not let our reliance on this test lead to stagnation of the development of additional tools to help prevent kidney stone recurrence.