COMMENTARY

The S.T.O.N.E. nephrolithometry scoring system: How valid is it?

Thomas Tailly, MD; Hassan Razvi, MD, FRCSC

Division of Urology, Western University, London, ON

With the observed trend toward the greater use of percutaneous nephrolithotomy (PCNL) in the management of urinary stone disease, the importance of preoperative risk assessment and estimation of treatment success has taken on an increasingly important role, especially as older and more ill patients are offered this procedure.[1] The potential benefits of preoperative scoring systems that can aid the urologist and his/her patient in the preoperative discussions about the risks and benefits of surgery are obviously welcome.

To date, 3 scoring systems have been proposed utilizing different metrics, including the Guy's Stone Score, S.T.O.N.E. nephrolithometry and the Clinical Research Office of the Endourological Society (CROES) nomogram.[2-4] The S.T.O.N.E system relies solely on information readily obtained from preoperative computed tomography (CT) imaging, whereas the other models also include patient variables. Four variables are shared with all 3 systems: stone location, size, number, and staghorn/multiple calyceal stone presence.

Equally important as the development of these prediction models is their external validation among different patient populations. Previous work comparing all 3 scoring systems demonstrated an equal ability for each to predict stone free status.[5] The authors of this manuscript attempted to validate the S.T.O.NE nephrolithometry score by studying 155 patients undergoing PCNL at their institution over a 4-year interval. While the study did validate the model, predicting a stone-free state with the S.T.O.N.E. score was no more accurate than using preoperative stone size or number of involved calyces. Moreover, the S.T.O.N.E. score did not predict postoperative complications.

These results should not diminish the potential usefulness of the model, but serve as a reminder that study design and patient heterogeneity may significantly affect study findings. The retrospective study design may have supplied an element of bias. The low number of events (patients with residual stone fragments) might have affected predictability as it has been suggested that at least 100 events are necessary for adequate discrimination ability of a model in an external validation study.[6] Lastly, a more diverse patient population reflecting not only patient factors but also different surgeons and institutional practices might also uncover other findings.

Prospective validation with a larger cohort involving multiple institutions should be the next step in assessing the various predictive models. To be most useful to clinicians, a single scoring system ideally will emerge, which will also promote standardized reporting of PCNL outcomes. We await those studies and congratulate the authors on their contribution to this growing field of research.

References

1. Ghani KR, Sammon JD, Bhojani N, et al. Trends in percutaneous nephrolithotomy use and outcomes in the United States. *J Urol* 2013;190:558-64. http://dx.doi.org/10.1016/j.juro.2013.02.036
2. Thomas K, Smith NC, Hegarty N, et al. The Guy's stone score: Grading the complexity of percutaneous nephrolithotomy procedures. *Urology* 2011;78:277-81. http://dx.doi.org/10.1016/j.urology.2010.12.026
3. Okhunov Z, Friedlander JI, George AK, et al. S.T.O.N.E. nephrolithometry: Novel surgical classification system for kidney calculi. *Urology* 2013;81:1154-9. http://dx.doi.org/10.1016/j.urology.2012.10.083
4. Smith A, Averch TD, Shahrour K, et al. A nephrolithometric nomogram to predict treatment success of percutaneous nephrolithotomy. *J Urol* 2013;190:149-56. http://dx.doi.org/10.1016/j.juro.2013.01.047
5. Labadie K, Okhunov Z, Akhavein A, et al. Evaluation and comparison of urolithiasis scoring systems used in percutaneous kidney stone surgery. *J Urol* 2015;193:154-9. http://dx.doi.org/10.1016/j.juro.2014.07.104
6. Vergouwe Y, Steyerberg EW, Eijkemans MJC, et al. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol* 2005;58:475-83. http://dx.doi.org/10.1016/j.jclinepi.2004.06.017

Correspondence: Dr. Hassan Razvi, St. Joseph's Health Centre, 268 Grosvenor St., London, ON N6A 4V2; fax: 519-646-6037; hrazvi@uwo.ca