Variation Among Internet Based Calculators in Predicting Spontaneous Resolution of Vesicoureteral Reflux

Jonathan C. Routh,* Edward M. Gong, Glenn M. Cannon, Jr., Richard N. Yu, Patricio C. Gargollo and Caleb P. Nelson†

From the Department of Urology (JCR, EMG, GMC, RNY, CPN) and Harvard Pediatric Health Services Research Fellowship Program (JCR), Children's Hospital Boston, Boston, Massachusetts, and Department of Urology, University of Texas–Southwestern Medical Center, Dallas, Texas (PCG)

Abbreviations and Acronyms

CHB = Children's Hospital Boston

UTI = urinary tract infection

VUR = vesicoureteral reflux

Submitted for publication July 24, 2005.

* Supported by Agency for Healthcare Research and Quality Grant T32-HS000063.

† Correspondence: Department of Urology, Children's Hospital Boston, 300 Longwood Ave., HU-359, Boston, Massachusetts 02115 (telephone: 617-355-3776; FAX: 617-730-0474; e-mail: caleb.nelson@childrens.harvard.edu).

Purpose: An increasing number of parents and practitioners use the Internet for health related purposes, and an increasing number of models are available on the Internet for predicting spontaneous resolution rates for children with vesicoureteral reflux. We sought to determine whether currently available Internet based calculators for vesicoureteral reflux resolution produce systematically different results.

Materials and Methods: Following a systematic Internet search we identified 3 Internet based calculators of spontaneous resolution rates for children with vesicoureteral reflux, of which 2 were academic affiliated and 1 was industry affiliated. We generated a random cohort of 100 hypothetical patients with a wide range of clinical characteristics and entered the data on each patient into each calculator. We then compared the results from the calculators in terms of mean predicted resolution probability and number of cases deemed likely to resolve at various cutoff probabilities.

Results: Mean predicted resolution probabilities were 41% and 36% (range 31% to 41%) for the 2 academic affiliated calculators and 33% for the industry affiliated calculator ($p = 0.02$). For some patients the calculators produced markedly different probabilities of spontaneous resolution, in some instances ranging from 24% to 89% for the same patient. At thresholds greater than 5%, 10% and 25% probability of spontaneous resolution the calculators differed significantly regarding whether cases would resolve (all $p < 0.0001$).

Conclusions: Predicted probabilities of spontaneous resolution of vesicoureteral reflux differ significantly among Internet based calculators. For certain patients, particularly those with a lower probability of spontaneous resolution, these differences can significantly influence clinical decision making.

Key Words: forecasting; Internet; models, statistical; prognosis; vesico-ureteral reflux

In the quest to forecast more accurately relevant outcomes of diseases and their treatment clinical researchers have become increasingly sophisticated in the use of prediction models such as logistic regression based nomograms and risk scores, artificial neural networks and recursive decision trees. In pediatric urology these techniques have been applied to vesicoureteral reflux, specifically to predict the probability of its spontaneous resolution.[1–4]

Several of these models have now been published as Internet based calculators, allowing providers and parents

0022-5347/10/1834-1568/0
THE JOURNAL OF UROLOGY®

Vol. 183, 1568-1573, April 2010
Printed in U.S.A.
DOI:10.1016/j.juro.2009.12.042

to enter pertinent clinical features of a patient and to determine the probability of spontaneous VUR resolution. The accuracy and consistency of such Internet calculators are of particular relevance to pediatricians and pediatric specialists, since in recent years there has been a dramatic increase in the use of the Internet to disseminate health information. Among adults using the Internet 75% to 85% report having gone online specifically to obtain health information.[5] A recent poll revealed that 86% of adults who search for information on the Internet believe that the information they find there is reliable.[5,6] As might be expected, teenagers and young Americans are even more dependent on the Internet as an information source than their elders, with approximately 90% of American teenagers accessing the Internet regularly, often for health related topics.[6,7] Given the potential clinical ramifications of differing predictions, we determined whether currently available Internet based calculators significantly differ from one another in terms of predicted probabilities for spontaneous resolution of pediatric VUR.

PATIENTS AND METHODS

Identification of Internet Based Calculators

We identified relevant Internet based calculators by performing a systematic search modeled on previously described methods.[8] Using the search engine Google, we queried the term "vesicoureteral reflux," sequentially combining this term with each of the terms "spontaneous," "resolution," "spontaneous resolution," "prediction," "model," "prediction model," "calculator," "Internet based" and "neural network." We identified 4 Internet based calculators, including 1 produced by Children's Hospital Boston (available at www.childrenshospital.org/vurcalculator), 2 produced by the University of Iowa (available at www.urocomp.net) and 1 produced by Q-Med Scandinavia (available at www.deflux.com). The 2 available calculators from Iowa differ primarily in that 1 incorporates nuclear renal scan data and the other does not.

Since many patients with newly diagnosed VUR will not undergo nuclear renal scan (at least not at initial diagnosis), and since the other 2 calculators do not incorporate these data, we chose to use only the Iowa calculator that does not require renal scan data. Because the Iowa calculator provides the probability of reflux resolution only at 2 years, we recorded only the 2-year probability for all 3 Internet based calculators. Similarly while the CHB and Q-Med calculators provide a resolution probability, the Iowa calculator provides the odds of resolution, from which we calculated the resolution probability. For patients with bilateral VUR grade and phase of onset were assumed to be equal on each side. Parameters considered by each calculator are listed in the Appendix.

Generation of Random Patient Sample and Statistical Analysis

We generated a random sample of 100 hypothetical patients with VUR using SAS®, version 9.1. For each patient we randomly selected several clinical variables, including age (integer values ranging from 0 to 11 years), VUR grade (I to V), gender, presentation (afebrile UTI, febrile UTI, prenatal hydronephrosis or sibling screen), laterality (unilateral or bilateral), cystogram demonstrated phase of VUR onset (bladder filling or voiding), ureteral duplication (present or absent), dysfunctional voiding (present or absent) and percent bladder volume (50% to 150% in 10% increments). We used a previously published definition of percent bladder volume, ie maximum bladder volume for a given patient normalized for age expected bladder capacity.[3] Overall, 9 clinical variables with up to 12 ordinal levels were randomly assigned for a total of 84,480 possible permutations.

To represent better a typical patient distribution, we over sampled patient age to simulate a Poisson distribution (median age 2 years), VUR grade to simulate a Gaussian distribution (median grade III with approximately a 1:2:4:2:1 ratio) and percent predicted bladder volume also to simulate a Gaussian distribution (median volume 100% of predicted). Additionally we programmed our randomization scheme to reflect a female-to-male gender ratio of 3:1 and a single-to-duplicated ureter ratio of 3:1. All parameter ranges were selected based on typical values for the patient population at our institution. Following randomization, the sample was screened to ensure that no variable combinations were duplicated and that each randomly generated patient was unique.

We then entered data for each patient into each of the 3 Internet based calculators and recorded the probability of spontaneous VUR resolution. For patients with unilateral VUR the left ureter was always assigned as the refluxing ureter. When data exceeded the maximum possible setting for a calculator the maximum possible setting was selected, eg if a patient was 11 years old but the calculator maximum age was 10, an age of 10 was entered.

We determined mean probabilities of spontaneous resolution across the cohort and 95% confidence intervals for each Internet based calculator. We compared the mean predicted probabilities of spontaneous resolution derived from each calculator using 1-way ANOVA. To determine further the clinical relevance of differences in predicted resolution probabilities, we compared the calculators in terms of ability to discriminate between cases that were likely to resolve spontaneously and those that were not. Specifically we compared the number of hypothetical patients whose predicted resolution probabilities were above or below a given threshold or cutoff value. Because no consensus exists regarding the most appropriate threshold value, we varied this threshold across a wide range (5%, 10%, 25%, 33%, 50% and 75% probability of resolution at 2 years). The numbers of cases above or below the threshold value were then compared between calculators using Fisher's exact test. All tests were 2-sided and p values of 0.05 or less were considered significant.

RESULTS

For all 100 hypothetical patients the mean predicted resolution probability was 40.7% (95% CI 35.7 to 45.7) for the CHB calculator, 36% (31.1 to 41) for the Iowa calculator and 32.5% (27.6 to 37.5) for the