

Statistics: The stethoscope of a thinking urologist

Arun S. Sivanandam

Vattikuti Urology Institute, Henry Ford Hospital, Detroit, Michigan, United States

ABSTRACT

Understanding statistical terminology and the ability to appraise clinical research findings and statistical tests are critical to the practice of evidence-based medicine. Urologists require statistics in their toolbox of skills in order to successfully sift through increasingly complex studies and realize the drawbacks of statistical tests. Currently, the level of evidence in urology literature is low and the majority of research abstracts published for the American Urological Association (AUA) meetings lag behind for full-text publication because of a lack of statistical reporting. Underlying these issues is a distinct deficiency in solid comprehension of statistics in the literature and a discomfort with the application of statistics for clinical decision-making. This review examines the plight of statistics in urology and investigates the reason behind the white-coat aversion to biostatistics. Resources such as evidence-based medicine websites, primers in statistics, and guidelines for statistical reporting exist for quick reference by urologists. Ultimately, educators should take charge of monitoring statistical knowledge among trainees by bolstering competency requirements and creating sustained opportunities for statistics and methodology exposure.

Key words: Biostatistics, evidence-based medicine, level of evidence, statistics, urological literature

DOI: 10.4103/0970-1591.52935

INTRODUCTION

Clinical decision-making should be founded on the highest level of evidence available. According to current hierarchies, Randomized Control Trials (RCTs) govern the top echelon due to the lowest possible influence of bias.^[1] As such, well-executed RCTs are the gold standard for clinicians assessing therapeutic effectiveness and treatment options. Borawski *et al.*, performed the first formal evaluation of the levels of evidence in urological literature.^[2] Independent reviewers familiar with the level of evidence concept rated 600 studies using a standardized evaluation form adapted from the Center of Evidence Based Medicine. The studies were randomly selected from four major urology journals (The Journal of Urology, European Urology, BJU International, and Urology) in the periods 2000 and 2005. Overall, 60.3% of studies addressed questions of therapy/prevention, 11.5% addressed etiology/harm, 11.3% addressed prognosis, and 9.2% addressed diagnosis. Articles centered mainly on adult populations (86%) with oncology as the topic of choice (38.8%). Disturbingly, the levels of evidence provided by these studies were low: 5.3% Level I, 10.3% Level II, 9.8% Level III, and 74.5% Level IV.

From 2000 to 2005, the highest level of evidence did not significantly improve (16.0–15.3%, respectively).

The authors conclude by suggesting that the majority of studies in urological literature cannot adequately guide clinical decision-making as a result of such low level of evidence. Several barriers to providing the highest level of evidence among surgical subspecialties have been previously identified, such as lack of surgeon–patient equipoise about certain therapies, difficulty of standardizing quality of a given surgical procedure, and limited funding mechanisms.^[3,4] However, another looming possibility exists: Is there paucity in statistical sense among urologists?

In line with low levels of evidence, findings at scientific meetings do not see the light of full-text publication in many cases. Failure to publish is problematic for two main reasons: 1) Clinicians looking to apply research findings lack the necessary detail in abstracts to critically appraise a given study for validity and impact; 2) It is wasteful of resources, unethical, and can lead to unnecessary replication of studies. Smith *et al.*, reviewed clinical research abstracts accepted for publication at 2002 and 2003 AUA Meetings.^[5] Literature search follow-up of published articles was performed in 2005. Out of 1683 abstracts, not surprisingly, the most common topic was oncology (40.8%). The majority of abstracts from North America (62.5%), reported single institution efforts (68.2%) mainly in the domain of therapy/prevention (51.6%).

For correspondence: Dr. Arun S. Sivanandam, Vattikuti Urology Institute, Henry Ford Hospital, Detroit, Michigan, United States. E-mail: asivana2@hfhs.org

Forty-four percent of these abstracts were published with a median follow-up of 27.8 months and 54.2% indicated formal statistical hypothesis testing. Kaplan-Meier analysis showed less time to publication of abstracts that had statistical testing (912) compared to those that did not (771) (log-rank $P = 0.009$). Univariate analyses identified statistical hypothesis testing with time for publication along with other predictors as significant factors contributing to the difference in publication rates. This was confirmed in multivariable analysis, as reporting to statistical testing remained predictive (HR 1.2, 95% CI 1.1–1.4). The authors highlighted how 61% of studies are affected by nonpublication of research findings two years after presentation at the AUA meeting due to a lack of statistics.

Statistics in clinical research is critical to the branding of evidence-based medicine. Raw data are meaningless to the busy urologist without statistical transformation and presentation. Increasingly, statistical methodology has transitioned from the realm of statistical journals to medical research.^[6] With the advent and plethora of available statistical software, statistics provides a framework to test relevant clinical hypotheses and unproven assumptions. Not using statistics is one weakness, but making errors in statistical testing and reporting of results can compromise the health of research animals, human subjects, and ultimate recipients of therapies. In research literature, other specialties have shown errors in statistical usage.^[7,8] Scales and colleagues performed a systematic assessment of statistical usage in urology literature.^[9]

Using a single issue (August 2004) of four leading urology journals (Journal of Urology, British Journal of Urology, Urology, and European Urology), two independent raters with formal statistics training reviewed the articles using a standardized evaluation form developed with an experienced biostatistician. Out of 97 articles that met eligibility criteria, cohort design comprised the majority of studies (44%). Of the 12.4% of studies that were randomized trials, 42% detailed clinically significant differences, 50% detailed power calculations, and 30% described method of randomization. Overall, statistical tests were identified in 83% of studies. Descriptive statistics were widely reported (94%) and articles mainly included simple statistical comparisons of two groups (77%). Distressingly, 71% of studies with statistical comparisons had at least one statistical error, including incorrect test (28%), faulty use of a parametric test (22%), and failure to adjust for multiple comparisons (65%). In addition, overfitting a regression model was a common problem (39%) in the 29% of studies that applied multivariable analysis. Such flawed application of statistics can potentially increase the likelihood of type I error and should be identified as a potential threat to validity of conclusions. The authors clearly show that statistical methods are used inappropriately in urology literature.

Statistics is paramount to success for the urologist as a researcher and as a clinician in urology. The remainder of this review will focus on probing the underlying problem of statistical use among clinicians and offer solutions that can be applied to rectify this situation.

THE PROBLEM DEFINED

To exercise evidence-based medicine (EBM), physicians need access to full-fledged research reports to critically evaluate study analysis and interpretation. However, surveys dating back to the 1980s identified physicians who had a poor grasp of statistical tests and interpretation of statistical results due to a lack of formal training in biostatistics.^[10–12] This problem is even more explosive today in light of increased complexity of statistical methods used in the literature.^[13] In response, graduate medical educators have increased training in biostatistics throughout the expanse of medical education. Medical schools have incorporated statistics courses and Accreditation Council for Graduate Medical Education (ACGME) guidelines since residency competency stipulates that residents must have a solid basic foundation in statistical methodology as it pertains to scientific research.^[14] While residency programs address this issue through EBM curricula and journal clubs,^[15–17] a few, if any, programs focus on selection and interpretation of statistical results.^[18]

To broadly assess residents' knowledge and skills in EBM, Windish *et al.*, conducted a seminal multiprogram assessment of 11 internal medicine residency programs in Connecticut.^[19] By first reviewing research articles in six leading general medical journals between January and March 2005 on the basis of statistical methods used, the researchers developed a survey instrument of questions focused on identifying and interpreting results in the most frequently occurring statistical tests. Questions were multiple-choice, centered on a clinical vignette, and required no calculations. Attitudes and confidence questions were adapted from surveys on the Assessment Resource Tools for Improving Statistical Thinking website, rated on a 5-point Likert scale. This instrument was validated and reformulated by pilot testing the questions on 5 internal medicine faculty with advanced training in biostatistics and 12 primary care internal medicine residents.

In terms of respondent characteristics, out of 277 residents, 48% were female, 60.8% aged 26–30 years with no advanced degrees (85.1%), and a modest distribution of years since medical school (35.0% <1 year, 26.8% 1–3 years, 30.1% 4–10 years). Of the foreign medical graduates in the population, 38.6% completed their medical school training outside the U.S., 68.8% had previous coursework in biostatistics [69.5% of which were during medical school (15.9% college, 3.2% residency)]. Over 50% had previous training in epidemiology and EBM, and regularly read medical journals. Interestingly, the number of residents

who could correctly identify and interpret statistical results was low. Approximately 25.6% could correctly identify chi-squared analysis, 13.0% could correctly identify Cox proportional hazard regression, 11.9% could interpret a 95% CI and statistical significance, and only 10.5% could interpret Kaplan-Meier analysis results. Using a forward stepwise regression model, advanced degrees, successive years since medical school, and prior biostatistics training were all factors found to be independently associated with knowledge scores. In terms of attitudes and confidence, 95% of residents agreed that knowledge of statistics is essential to being an intelligent reader of literature and 77% indicated they would like to learn more statistics. While over 58% of residents reported using statistics in forming opinions or making clinical decisions, 75% indicated they did not fully understand the statistics reported in literature. Only 38% of residents felt confident assessing the appropriateness of statistical testing used and respondents with a higher confidence level in statistical knowledge fared better on the knowledge questions.

While their report was confined to internal medicine residents, high internal consistency, good discriminative validity, and similarity in results among different residency programs lend credibility to the illustrated problem.^[19] The authors direct the poor knowledge and understanding of biostatistics to insufficient training. A comprehensive review of biostatistics teaching indicates that 90% of medical schools taught biostatistics in preclinical years only with varying breadth and depth of education.^[20] While basic statistics were frequently addressed, advanced methods were seldom included. Another pressing issue is that senior residents performed worse than junior residents, indicating a time correlation. Most likely, loss of knowledge over time, coupled with lack of adequate reinforcement could lead to loss of statistical competency. This lack of AGCME competency comes at a great cost. If clinicians cannot evaluate appropriate statistical tests and accurately interpret results, risks could be carried over to incorrect clinical decision-making.

West and colleagues performed a similar study in 2005 on 301 medical students, internal medicine residents and faculty, about their attitudes toward biostatistics in medicine.^[21] According to their findings, 48.3% of those surveyed felt biostatistics is a difficult subject, 87.3% felt that understanding biostatistics would help their careers, and 17.6% felt their training in biostatistics was adequate for their needs. Furthermore, 23.3% of respondents could evaluate appropriateness of statistical methods used in a study, 88% felt knowledge of statistics is necessary for evaluating medical literature, and 48.5% felt that biostatistics is a necessary skill for clinicians not involved in research. In essence, the survey strongly indicated that clinicians were uncomfortable with biostatistics and even more dissatisfied with this cognizance. It is unclear why physicians are queasy

regarding statistics even though they use statistics in their daily routine.

Perhaps the finding that 20% of respondents felt their biostatistics coursework was taught effectively calls into question as to how clinicians are being educated about statistics in healthcare fields. Can understanding of statistics be improved to avoid erroneous interpretation and application? Traditional teaching methods in schools employ a stepwise approach entailing formulae, data, and spoon-fed instructions. This does not relate well to patients or analysing scientific papers. Medical statistics are often taught as abstract concepts removed from clinical relevance. Bordering on a moral quandary is the question of whether expectations for the average urologist are too high. Would the urologist who is not a researcher be better suited to appraise practice guidelines, derived by experts with the necessary statistical knowledge, rather than interpret statistics? Urology is a highly competitive field that is constantly evolving and as such, expectations will continue to be shattered and stacked higher. The current consensus will most likely rest on the urologist having a strong statistical repertoire because research is an increasingly integral component of residency and fellowship programs, because guidelines can change given new information, and because treatment accountability ultimately rests with the physician's ability to evaluate evidence and make decisions.

Most of the studies examining the use of statistics and knowledge of clinicians have thus far been centered in the U.S. In urology, only major journals have been examined leaving other international journals indexed in MEDLINE, such as Brazilian Journal of Urology and Indian Journal of Urology out of the loop. It is vital to assess how these journals and how urology practitioners in these regions fare in comparison to the current data through future investigations of this nature.

RATIONAL SOLUTIONS

So what is a busy urologist to do? Although errors in statistics and a lack of comprehensive understanding in methodology are common in the literature,^[22,23] modifications to current mindsets can still be made in the best interests of the patient. Curran-Everett and Benos have proposed guidelines for reporting statistics in journals published by the American Physiological Society.^[24] A set of 10 guidelines, ranging from advice to consult a biostatistician to interpretation based on confidence intervals and *P*-values, address reporting of statistics in the Materials and Methods, Results, and Discussion sections of a manuscript. A cursory look at additional references cited in the manuscript provides additional resources for urologists interested in looking at the framework of statistics and presentation issues. In addition, a commentary aimed at the publication of these guidelines by Murray Clayton provides an excellent critique of when to use the guidelines.^[25] Clayton argues that the

algorithmic approach of guidelines may not always serve the practitioner or peer-reviewer well as situational cues dictate statistical testing and interpretation. As such the word is still out as to whether these guidelines truly represent the best practices in statistics.

Focusing on urology, Scales and colleagues have produced two publications that can serve as a starting point of quick statistical reference. First, they provided a series of non-technical explanations of basic statistical concepts encountered in urological literature.^[26] In terms of results, they discuss various outcome measures, how to summarize continuous data, how to summarize non-normal continuous and ordinal data, how to summarize unordered categorical data, how to interpret CIs, how to interpret RRs, the difference between Odds, OR, and RR, how to interpret a KM curve, and how to interpret multivariable analyses. In addition, the authors provide examples of common statistical flaws involving Type I and II errors, sample size calculations, multiple comparisons, and confounding variables to increase awareness of study limitations in light of statistical restrictions. By providing a statistical roadmap, the authors provide advice on choosing appropriate statistical tests as a brief introductory roundup for the practicing urologist.

Scales and colleagues also provided a complementary companion primer on evidence-based clinical practice (EBCP) for urologists using examples from the literature.^[27] Principles of EBCP are discussed followed by a step-by-step approach to implementing EBCP. Sources of evidence are discussed along with methods to evaluate a study for therapeutic effectiveness. With appendices that summarize levels of evidence, electronic databases of primary evidence, and web addresses of online EBCP centers, this primer can provide urologists with the tools and questions that can aid in accumulating evidence and clinical decision-making.

Faculty who are implementing biostatistics curricula can access these teaching resources. Without a doubt, teaching of statistics to medical students, residents, and fellows can be improved. Rather than sparse statistical exchanges during journal clubs, medical education should be expanded to make biostatistics less daunting and more meaningful to urologists in practice. More time should be allotted to biostatistics education in medical school in a clinical problem-based learning format.

Rather than a one-shot infusion of statistics through an isolated course or a seminar, reinforced and integrated learning simulating research experiences should be fostered. Ideally, medical students will have exposure to statistics throughout their training. In residency, this can be complemented by recurring seminars from available biostatisticians or visiting faculty from nearby universities. These can be in the form of a retreat with a distribution of problem-sets at the end. Small-group work

can be encouraged for a gathering and review of solutions a week later. Yet another option is online-educational courses offered by a variety of universities. For instance, Harvard University Extension School offers a semester-long course on introductory graduate biostatistics.^[28] Students can view streamed video lectures, post questions on an online discussion board, ask questions from professors and teaching assistants and receive feedback on homework and examinations as if they were partaking in a live course. While mailing outside the U.S. for graded assignments poses a time-lag problem, courses such as these provide an alternative if the means of quality education and expertise are lacking in the area. Such courses provide the welcome opportunity of immersing oneself in statistical software and learning the realities behind a particular formula.

Ultimately, broader facilitation should be imparted at the departmental level to enable urologists to better answer research questions. Considering a hectic schedule of surgeries in the OR and clinic presence, accessibility of literature for review, adequate data management infrastructure, availability of statistics know-how, and project supervision by faculty are the key factors that can dissuade even the most curious physician. Urology training programs need to be more trainee-centered to imbibe a statistical way of thinking to work around the areas of uncertainty. Statistical software that can transform raw data from a database into meaningful results using a core set of statistical tests should be freely available for use. Softwares such as STATA, SPSS, SAS, Sigmaplot, R, JMP, and Comprehensive Meta Analysis, to name a few, understandably require institutional licenses. Although these licenses are expensive, the investment is worthwhile because residents and fellows will get hands-on exposure to working with numbers. If such expenses are prohibitive, regional collaborations are encouraged to allow such software packages to be transitive in distribution. Departmental oversight of this nature can help ensure competency in fields of data management, statistical formula application, critical analysis, and study interpretation.

Competencies should be expanded in medical school and residency to mandate a certain level of proficiency in order to progress from one training year to the next. In conjunction with better education of urologists, attitudes toward, and use of statistics will continue to improve.

CONCLUSION

Medicine is evolving at a rapid pace with publications increasing to the rate that journals have a backlog of articles that see print six months after acceptance. At this pace, urologists need to be less intimidated by biostatistics. As important as the stethoscope, statistical sense is crucial to evaluate research findings and examining patient research. If not just for clinical decision-making, at least the physicians

have a mechanism of expressing to patients why they are making a particular decision. The current problem of a low level of statistical evidence in urology literature coupled with a significant lag between abstract presentation and a full-text publication represent a lack of understanding of, and comfort with, statistics. This is reflected in errors in statistical usage that can be corrected by increased awareness of the problem and readiness to act by improving medical education of statistics.

ACKNOWLEDGMENTS

I would like to thank biostatistician Mireya Insua-Diaz for her critical review and insightful inputs on the manuscript.

REFERENCES

- Petrisor BA, Keating J, Schemitsch E. Grading the evidence: Levels of evidence and grades of recommendation. *Injury* 2006;37:321-7.
- Borakowski KM, Norris RD, Fesperman SF, Vieweg J, Preminger GM, Dahm P. Levels of evidence in the urological literature. *J Urol* 2007;178:1429-33.
- Solomon MJ, MacLeod RS. Surgery and the randomized clinical trial: Past, present and future. *Med J Aust* 1998;169:380-3.
- Evans CP. Evidence-based medicine for the urologist. *BJU Int* 2004;94:1-2.
- Smith WA, Cancel QV, Tseng TY, Sultan S, Vieweg J, Dahm P. Factors associated with the full publication of studies presented in abstract form at the annual meeting of the American Urological Association. *J Urol* 2007;177:1084-9.
- Altman DG, Goodman SN. Transfer of technology from statistical journals to the biomedical literature: Past trends and future predictions. *JAMA* 1994;272:129-32.
- Hall JC, Mills B, Nguyen H, Hall JL. Methodologic standards in surgical trials. *Surgery* 2001;119:466-72.
- Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other misuses of baseline data in clinical trials. *Lancet* 2000;355:1064-9.
- Scales CD Jr, Norris RD, Peterson BL, Preminger GM, Dahm P. Clinical research and statistical methods in the urology literature. *J Urol* 2005;174:1374-9.
- Berwick DM, Fineberg HV, Weinstein MC. When doctors meet numbers. *Am J Med* 1981;71:991-8.
- Weiss ST, Samet JM. An assessment of physician knowledge of epidemiology and biostatistics. *J Med Educ* 1980;55:692-7.
- Wulff HR, Andersen B, Brandenhoff P, Guttler F. What do doctors know about statistics? *Stat Med* 1987;6:3-10.
- Horton NJ, Switzer SS. Statistical methods in the journal. *N Engl J Med* 2005;353:1977-9.
- Accreditation Council for Graduate Medical Education (ACGME). Outcome project: Enhancing residency education through outcomes assessment. Available from: <http://acgme.org/Outcome/>. [last accessed on 2008 Sep 28].
- Green ML. Evidence-based medicine training in internal medicine residency programs: A national survey. *J Gen Intern Med* 2000;15:129-33.
- Dellavalle RP, Stegner DL, Deas AM, Hester EJ, McCeney MH, Crane LA, et al. Assessing evidence-based dermatology and evidence-based internal medicine curricula in US residency training programs: A national survey. *Arch Dermatol* 2003;139:369-72.
- Alguire PC. A review of journal clubs in postgraduate medical education. *J Gen Intern Med* 1998;13:347-53.
- Green ML. Graduate medical education training in clinical epidemiology, critical appraisal, and evidence-based medicine: A critical review of curricula. *Acad Med* 1999;74:686-94.
- Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. *JAMA* 2007;298:1010-22.
- Looney SW, Grady CD, Steiner RP. An update on biostatistics requirements in US medical schools. *Acad Med* 1998;73:92-4.
- West CP, Ficalora RD. Clinician attitudes toward biostatistics. *Mayo Clin Proc* 2007;82:939-43.
- Altman DG. Statistics in medical journals: Some recent trends. *Stat Med* 2000;19:3275-89.
- Altman DG, Bland JM. Improving doctors' understanding of statistics. *J R Stat Soc Ser A* 1991;154:223-67.
- Curran-Everett D, Benos DJ. Guidelines for reporting statistics in journals published by the American Physiological Society. *Adv Physiol Educ* 2004;28:85-7.
- Clayton MK. How should we achieve high-quality reporting of statistics in scientific journals? A commentary on "Guidelines for reporting statistics in journals published by the American Physiological Society". *Adv Physiol Educ* 2007;31:302-4.
- Scales CD, Peterson B, Dahm P. Interpreting statistics in the urological literature. *J Urol* 2006;176:1938-45.
- Scales CD Jr, Preminger GM, Keitz SA, Dahm P. Evidence based clinical practice: A primer for urologists. *J Urol* 2007;178:775-82.
- Evans SR, Wang R, Yeh TM, Anderson J, Hajja R, Mcbratney-Owen PM et al. Evaluation of distance learning in an "Introductory Biostatistics" class: A case study. *Stat Educ Res J* 2007;6:59-77.

How to cite this article: Sivanandam AS. Statistics: The stethoscope of a thinking urologist. *Indian J Urol* 2009;25:241-5.

Source of Support: Nil, **Conflict of Interest:** None declared.