How Do We Promote Clinical Statistics Literacy of Emergency Medicine Residents: Is Clicker Technology the Answer?

Penny. S. Reynolds [1]

[1] Department of Emergency Medicine, Virginia Commonwealth University Medical Center and Virginia Commonwealth University Reanimation Science Center
1200 East Marshall St., Richmond VA 23298

Abstract
Physicians must be able to stay current with the vast and expanding medical research literature. Understanding of basic research concepts is also part of the competency requirements for medical resident education. However most residents lack even the most basic knowledge of clinical statistics required to meet these objectives. Wireless personal response system technology ("clickers") has been shown to improve retention rates of factual information in non-medical settings by encouraging active participation during lectures. The goal of this project was to increase Emergency Medicine resident reading comprehension of clinical statistics over 10 months by combining formal lectures with clicker technology. Emergency Medicine residents were given lectures covering 13 learning objectives in 3 research competency requirement areas. Lecture content de-emphasized computations in favor of concepts. Concept-based quizzes were given in each lecture, and in-class assessment tools ("clickers") provided immediate feedback. Supplemental readings were posted online before lectures. Assessment was based on pre- and post-test survey scores. *Results*: Reading comprehension of basic concepts remained poor. Median pre-test score was 30%; post-test was 35%. There was no change in knowledge of design, power, sensitivity and specificity, and bias. Understanding of risk increased from 25% to 70%; paradoxically, residents were unable to compute related metrics e.g. number needed to treat. On-line tracking showed that few residents availed themselves of supplementary resources.
Conclusions: Residents lack motivation to increase proficiency in basic statistical concepts when presented in traditional formats. Lack of time was cited as the most frequent barrier to utilizing information outside of in-house educational sessions. Clickers were useful in increasing immediate engagement and participation, but did not reinforce or encourage learning outside of the sessions.

Key Words: Audience response systems, medical education, physician; problem-based learning; statistics teaching,

1. Introduction

To understand and assess information about their field, physicians must be able to stay current with the vast and rapidly expanding body of medical research literature. Understanding of basic research principles is therefore part of the competency requirements for Emergency Medicine residency training, according to guidelines for resident education given by the Accreditation Council for Graduate Medical Education (ACGME)(Education, 2007), and the Council of Emergency Medicine Residency Directors (CORD) (Directors, 2007). A summary of basic criteria for the research competency, as identified by CORD, are given in Box 1. All areas of emphasis in the general Research Competency component (Practice-based Learning, Evidence-based Medicine, Interpretation of Medical Literature, and Performance of Research) require "familiarity" or "proficiency" [*sensu* CORD(Directors, 2007)] with major concepts and methods of clinical statistics. Additional goals for resident training in these areas include development of "self-directed, life-long learners", and the ability to perform "evidence-based practice". Skill sets to be developed include the ability to 'locate, critically appraise, and assimilate evidence' from clinical studies, and the application of 'knowledge of study designs and statistical methods to the appraisal of clinical studies and other information on diagnostic and therapeutic effectiveness'.

***Box 1*. Research competency areas identified by CORD for medical resident education**

- Identify major study designs (randomized controlled clinical trial, case-control, cohort, cross-sectional, case studies), & list the advantages & disadvantages of each.
- Identify the necessary conditions for study reliability & validity: randomization, blinding, allocation, Intention-to-treat analysis .
- Risk: Define & interpret odds ratio, relative risk/risk ratio. Determine strength of evidence for risk factors.
- Identify principles of statistical hypothesis testing: null & alternative hypotheses; alpha, beta, & statistical power; type I & type II errors as they relate to sample size & variance.
- Define major variable types: interval, ordinal, nominal, discrete, binary, continuous.
- Define & calculate summary statistics for continuous data: mean, median, mode, standard deviation, standard error, variance.
- Identify principles of summarizing non-continuous data; perform simple calculations.
- Statistical tests: t-test, paired t-test, analysis of variance, chi square, Fisher exact test, & non-parametric tests.
- Distinguish between statistical & clinical significance.
- Diagnostic tests: Define Incidence/prevalence, sensitivity, specificity, positive predictive value, & negative predictive value. Given a patient case scenario, be able to interpret probabilistic & frequentist statements in terms a patient can understand.
- Measures of association: Compare & contrast correlation & regression, & context for use. Distinguish between independent & dependent variables.
- Simple survival analyses (Kaplan-Meier, Cox proportional hazards): identify and interpret.

1.1 Barriers

Barriers to implementing these admittedly ambitious programme goals can be categorised in terms of objectives, motivation, and preparation, and institutional support. First, the three levels of competency defined by CORD — "Mastery", "Proficiency", Familiarity" – are designed for assessing clinical competencies in the management of acute illness and injury, plus associated cognitive and motor skills. However it could be argued that the primary goal of the clinical research component should be directed towards establishing statistical literacy (Gal, 2003), rather than statistical competence per se; therefore the aims, tools, and skill sets to be developed will differ.

The second barrier – lack of motivation – arises from the severe time constraints on residents and competing education goals for the resident education curriculum. The research component tested in board exams is small and therefore penalties for not knowing the prescribed information are minimal. There are simply far more important items to be mastered in the residency curriculum relating more directly to medical practice. Many residents have negative associations with statistics; attitudes vary from dislike to terror (*personal communications*).

The third barrier – inadequate preparation – results from residents lacking even basic knowledge of clinical statistics. Formal training in statistics is minimal, usually early in the curriculum, and therefore information is forgotten by entry into residency. A recent anonymous cross-sectional survey of 277 residents in 11 internal medicine residency programs across Connecticut showed that at least 95% felt that it was “important to understand statistical concepts”, nevertheless the median knowledge scores tested at 39% across all subject areas described as necessary for research competency. Almost 40% of those surveyed had no prior biostatistics training; knowledge scores tended to decrease with successive years since medical school graduation (Windish et al., 2007). There is likely to be minimal understanding or retention of complex statistical information if background preparation is lacking.

Finally institutional support for biostatistics training for residents is somewhat unstructured. During residency, the research competency is usually addressed through journal clubs. Journal club format consists of monthly gatherings to review papers relevant to clinical practice; ‘statistical education’ is provided by input from attending faculty. Because this model is unstructured, infrequent, & informal, there was little or no communication, understanding, or retention of complex statistical concepts. In past years, the clinical research requirement was also addressed through a two-hour lecture given once yearly to incoming interns at the beginning of the two-week intern orientation period. Because there are typically over 20 participants in any one session, and conference sessions are 8 or more hours per day, time limitations and fatigue often prevented full participation by every individual or complete discussion of complex topics. Consequently, even though time was allowed for questions, residents complained that they still could not understand specific topics. Retention of difficult concepts, such as sensitivity & specificity & interpretation of risk was minimal.

1.2 Closing the gap: Past methods and a proposed remedy

As previous instructional strategies were inadequate to cover the CORD requirements, a more intensive lecture series was proposed, to take place within the weekly in-house educational sessions. Because audience participation during lectures has been demonstrated to increase with the use of compact electronic wireless classroom performance, or personal response, systems (“clickers”), it was proposed to incorporate this system into the teaching of medical statistics. Increased active audience involvement has been shown with personal response systems; clicker use has been demonstrated to result in improved information retention in non-medical settings (Slain et al., 2004; Wit, 2003; Wood, 2004), and is finding increased use in medical schools (Schackow et al., 2004).

2. Methods

2.1 Project design

I conducted a prospective before and after study to measure change in clinical statistics knowledge resulting from resident exposure to an intensive clinical statistics lecture series. This project was conducted at a university-affiliated emergency medicine residency training program in a Level I trauma medical center located in Richmond VA.

The study involved 25-30 emergency medicine residents at our institution. Each month, a 1-hour lecture topic from the clinical research competency requirements list was presented in the regular weekly in-house education sessions. All lectures were presented using the Classroom Performance System™ (CPS™; eInstruction®, Denton, TX) integrated with Powerpoint presentation software (Microsoft); each participant was provided with a CPS Pulse™ handheld response device. Supplementary material (lecture material, Powerpoint slides, reference articles, lecture notes) was posted on Blackboard.

2.2 Assessment format

Three assessment metrics were measured: pre-test-post-test scores; topic-specific assessments, and resource use assessment. All residents attending the in-house conference were given a pre-intervention assessment quiz adapted from Windish et al. (2007). Questions were modified to reflect content relevant to emergency medicine practice. A post-assessment quiz was administered after the completion of the series at the end of the residency year.

Topic-specific assessments were performed during each lecture. Multiple choice questions were interspersed throughout each lecture in pre-test/post-test format; one or two pre-test questions preceded brief instructional content, which in turn was followed by several post-test questions. Quiz questions were based primarily on *BMJ Statistics Endgames* material by Philip Sedgwick, modified for emergency medicine relevance. Instructional content consisted of 15-30 minutes explanation of new concepts and/or simple computations. Audience members keyed in their best answer using their CPS keypad and tallied results were displayed anonymously. Resident use of supplementary resource material was tracked by the number of Blackboard web hits.

3. Results

3.1. Resident characteristics. Nearly all respondents in this survey reported receiving some biostatistics training during college or medical school. However of these, 69% had less than 40 hours of formal instruction; generally this consisted of a 2 week course in epidemiology, given in the first year of medical school. Only 3 individuals had at least a semester course in college.

Nearly all respondents agreed or strongly agreed that understanding biostatistics was essential to understanding the clinical literature (95%). However, at least half (50%) stated that lack of time was the primary barrier to reading and understanding the literature, while a further 40% reported uncertainty as to whether or not to believe research results. The second–ranked barrier to understanding was stated as 'the statistical gobledegook' (48%). In this survey, residents reported that reading journal abstracts (rather than the entire journal article) was their primary method of keeping current (55%), followed by opinion pieces in leading journals (30%), and talks at meetings (15%). Interestingly, none reported using alternative methods of obtaining clinical information, such as various media outlets or pharmaceutical representatives.

3.2 Knowledge of statistical methods. Overall test results were disappointing. Pre-test median scores were 31%; post-test medial scores were 35%; scores ranged from 10% to 85%. In common with the Windish et al. study, residents scored highest in recognition of blinding (80%), systematic bias (75%) and interpretation of relative risk (75%). Understanding of power, sensitivity and specificity, and bias were generally poor, and did not change between pre-test and post-test scores. Residents were least able to correctly interpret odds ratio, the meaning of a P-value, or differences between case-control and cohort studies (< 10%). Understanding of risk increased from 25% to 70%; paradoxically, residents were unable to compute related metrics e.g. number needed to treat (5%). On-line tracking showed that very few residents availed themselves of supplementary resources outside of instructional sessions.

3.3 Simple statistical literacy tests. In-class mini-quiz results assessed by clickers were extremely useful for tailoring and modifying instructional input to match the needs of the participants. For example, the concepts involved in diagnostic tests are notoriously difficult for many clinicians to grasp (Gigerenzer et al., 2007; Loong, 2003). The lecture on this topic was constructed in three parts: a brief case study with a multiple-choice question involving calculation of positive predictive value; a brief 5-minute presentation of working definitions of the key terms, and a participant mini-quiz on three simple tests of probability. Each section was assessed by a multiple-choice question involving either comprehension of terms or simple calculations. At the end of the section the initial case study question was repeated. Initially nearly all residents (24/25) were unable to calculate positive predictive value from the presented information. There was good comprehension of the meanings of key definitions (sensitivity, specificity, prevalence, etc.). However, resident understanding of the concepts of basic probability was surprisingly poor. Residents were presented with three questions involving simple calculations: expected number of heads in a fair coin toss, converting a percent to a proportion, and converting a proportion to a percent [Box 2; adapted from (Schwartz et al., 1997)].

Box 2. Testing resident comprehension of basic probability concepts. Questions were adapted from (Schwartz et al., 1997) .

1. **Basic probability**
   Imagine that you flip a coin 1000 times. *How many times* would you expect the coin to come up heads? _____ times out of 1000

2. Convert a percent to a proportion
A person taking a given drug has a 1% chance of having an allergic reaction. If 1,000 people take this drug, *how many* would you expect to have an allergic reaction? ____ person(s) out of 1000

3. **Convert a proportion to a percent**
A person taking a given drug has a 1 in 1000 chance of an allergic reaction. What *percent* of people taking this drug will have an allergic reaction? ___ %

The percentage of participants giving correct answers to these questions was expected to be 100%, 90% and 75% for the three respective questions, based on data presented in (Gigerenzer et al., 2007) for physicians. However, clicker-based assessments were 95%, 67%, and 25%, much lower than expected, although in agreement with results obtained from a nationally-representative sample of American adults with some college education [Schwartz and Woloshin 2000 unpublished, reported in (Gigerenzer et al., 2007)]. Remedial material could therefore be directed towards developing skill sets in basic calculations before tackling the more complex calculations required for solving the case study problem. At the end of the session, the majority of participants (80%) were able to correct compute the positive predictive value for the case-based scenario presented initially.

Participants reported that they enjoyed the anonymity of clicker-based response systems and encouraged their participation more than being singled out class could.

4. Conclusions

Little Jack Horner
Sat in the corner,
Eating a Christmas pie;
He put in his thumb,
And pulled out a plum,
And said 'What a good boy am I!'

Trad. English nursery rhyme

Physicians need tools to weigh clinical evidence and those tools involve an understanding of often highly-complex statistical concepts and skill sets (Switzer and Horton, 2007). Equally, it is apparent that medical residents have neither the understanding nor the skills to interpret much of the statistical reporting found in the majority of clinical journals. This study reported here was a before-and-after test of the proposal that intensive lecture courses combined with interaction technology should improve clinical statistical literacy in emergency medicine residents at a single institution. Although a small sample and limited by incomplete participation by all residents at this institution, this survey confirms what has been determined by larger, more rigorous studies (Windish et al., 2007); namely, that medical residents are hampered in their ability to critically evaluate published clinical data, and that current professional expectations and pedagogy geared towards fulfilling those expectations may not be adequate to bridge that gap.

Residents generally agree that understanding clinical statistics is critical for intelligent interpretation of the clinical literature. However, because of time constraints and inadequate statistics background, clinicians rely on numerous short cuts (such as reading abstracts, and opinion pieces) to remain current. This 'Jack Horner strategy' – relying on others to pull out the information "plums" – while presumably saving time and effort, nevertheless involves a drastic filtering of information, with concomitant losses in both reliability and content. A random sample of 44 abstracts in five clinical journals found that over 20% reported findings or conclusions inconsistent with the full article. More recently it was found that 'spin' – presenting research results more favourably than indicated by the available evidence – was found in at least 25% of articles surveyed and was more prevalent in the abstracts than in the full articles (Boutron et al., 2010). A survey of 148 randomised controlled trials presented at American College of Cardiology meetings between 1999-2002 found that 41% differed in efficacy estimate of primary outcome from later reports published in the primary literature (Altman, 2002). Finally, Marcia Angell, former editor of the *New England Journal of Medicine* has stated "It is simply no longer possible to believe much of the clinical research that is published……" (Angell, 2009). Physicians are hampered in their ability to weigh and assess evidence without an adequate statistical background (Altman and Bland, 1991).

(Guyatt et al., 2004) have identified the biggest challenge to evidence-based medicine as one of knowledge translation; that is, ensuring that "clinicians base their day to day decision-making on the right principles and on current best evidence". However failures in translation occur because of failures in application, either through lack of knowledge of available evidence or inability to translate it to clinical practice. It is now apparent that such translational failures may result from the statistical literacy gap between expectations defined by CORD and ACGME, and the reality of resident educational background. It is easy to underestimate the prevalence of limited statistical literacy among even well-educated and motivated individuals; as a consequence it is also easy to over-estimate the ability of physicians to understand the information they are given. The fundamental issue appears to be one of statistical literacy, in which case there will need to be a profound shift in instructional emphasis. Instruction of core statistical topics appropriate for students majoring in the quantitative sciences may be neither appropriate or logistically feasible if the goal is the development of increased statistical literacy (Gal, 2002, 2003), especially in a highly-selected population of trainee physicians. Consequently, it is critical to understand exactly what the primary needs of residents are

Statistical literacy for adults as (as opposed to students actively learning statistics) has been proposed to refer to ability to first, 'interpret and critically evaluate statistical information, data-related arguments' encountered in different contexts, and second 'to discuss or communicate their reactions to such statistical information (Gal, 2002, 2003); that is data awareness, comprehension and interpretation. Physicians are unlikely to value clinical statistics training or see it as clinically relevant unless it can be shown how it can be immediately applied in clinical practice. Motivation to learn clinical statistics can only come from the appreciation that clinical data are a part of current everyday practice, clinical data are often misused, resulting in misinformation and possible harm, and clinical decisions made based on data can have a strong impact on practice and patient outcome. Appropriate instructional formats should probably be re-framed, not in the context of mathematical subject matter but in such a way that participants can use statistical thinking on problems containing data, and appreciate the role of statistics in decision making (Schield, 2010). Therefore, specific content in their daily clinical education must refer to publications of relevant trials and other studies to show how research methods integrate with clinical practice. To make this framework 'operational' in the instructional setting, the instructors should devise suitable tasks, such as mini-quizzes for comprehension assessment, or assigning a patient-based case scenario with the appropriate literature references to support or refute diagnostic or therapeutic claims. Finally the instructing statistician should be attuned to possible limitations in statistical literacy among health care professionals in the audience, then modify communication methods appropriately to match the needs of those individuals.

Acknowledgements

This project was supported through the generosity of Virginia Commonwealth Uuniversity Center for Teaching Excellence (CTE) Small Grants Program, and the cooperation of Drs. T Evans and D Franzen, Emergency Medicine, VCU Medical Center. Special acknowledgements to Dr. Philip Sedgwick, St George's, University of London, for the brilliant BMJ *Endgames,* examples from which provided essential teaching material for this project. Parts of this work were presented at the 2011 Joint Statistical Meetings (Teaching Statistics in Health Sciences), Miami Beach FL.

References

Altman, D. G. 2002. Poor-quality medical research: What can journal do? JAMA 287: 2765-2767.

Altman, D. G., and J. M. Bland. 1991. Improving doctors' understanding of statistics. Journal of the Royal Statistical Society. Series A (Statistics in Society) 154: 223-267.

American Council for graduate Medical Education. 2007. Program requirements for graduate medical education in emergency medicine 2007.

Angell, M. 2009. Drug companies & doctors: A story of corruption New York Review of Books No. 56. p 3.

Boutron, I., S. Dutton, P. Ravaud, and D. G. Altman. 2010. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes JAMA 303: 2058-2064.

Council of Emergency Medicine Residency Directors (CORD). 2007. Model of the clinical practice of emergency medicine 2007.

Gal, I. 2002. Adults' statistical literacy: Meanings, components, responsibilities. International Statistical Review 70: 1-51.

Gal, I. 2003. Teaching for statistical literacy and services of statistical agencies. American Statistician 57: 80-84.

Gigerenzer, G., W. Gaissmaier, E. Kurz-Milcke, L. M. Schwartz, and S. Woloshin. 2007. Helping doctors and patients make sense of health statistics. Psychological Science in the Public Interest 8: 53-96.

Guyatt, G., D. Cook, and B. Haynes. 2004. Evidence based medicine has come a long way. BMJ 329: 990-991.

Loong, T.-W. 2003. Understanding sensitivity and specificity with the right side of the brain. BMJ 327: 716-719.

Schackow, T. E., M. Chavez, L. Loya, and M. Friedman. 2004. Audience response system: Effect on learning in family medicine residents. Family Medicine 36: 496-504.

Schield, M. 2010. Assessing statistical literacy: Take care. In: P. Bidgood, N. Hunt and F. Jolliffe (eds.) Assessment methods in statistical education. John Wiley & Sons, Chichester, UK

Schwartz, L. M., S. Woloshin, W. C. Black, and H. G. Welch. 1997. The role of numeracy in understanding the benefit of screening mammography. Annals of Internal Medicine 127: 966-972.

Slain, D., M. Abate, B. M. Hodges, M. K. Stamatakis, and S. Wolak. 2004. An interactive response system to promote active learning in the doctor of pharmacy curriculum. American Journal of Pharmaceutical Education 68: 1-9.

Switzer, S. S., and N. J. Horton. 2007. What your doctor should know about statistics (but perhaps doesn't...). Chance 20: 17-19.

Windish, D. M., S. J. Huot, and M. L. Green. 2007. Medicine residents' understanding of the biostatistics and results in the medical literature. JAMA 298: 1010-1022.

Wit, E. 2003. Who wants to be… the use of a personal response system in statistics teaching. MSOR Connections 3: 13-20.

Wood, W. B. 2004. Clickers: A teaching gimmick that works. Developmental Cell 7: 796-798.