



# Using the 4MAT Framework to Design a Problem-based Learning Biostatistics Course

[Amy S. Nowacki](#)  
Cleveland Clinic

*Journal of Statistics Education* Volume 19, Number 3(2011),  
[www.amstat.org/publications/jse/v19n3/nowacki.pdf](http://www.amstat.org/publications/jse/v19n3/nowacki.pdf)

Copyright © 2011 by Amy S. Nowacki all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the author and advance notification of the editor.

---

**Key Words:** Active learning; Student attitudes; Curriculum assessment; Course evaluation.

## Abstract

The study presents and applies the 4MAT theoretical framework to educational planning to transform a biostatistics course into a problem-based learning experience. Using a four-question approach, described are specific activities/materials utilized at both the class and course levels. Two web-based instruments collected data regarding student satisfaction with the course and perception of the field of biostatistics (Attitudes Toward Statistics Survey). Student satisfaction and perception increased significantly following implementation of the new curriculum as compared to previous ratings. The results indicated that students felt more strongly that the seminars were well-organized, encouraged participation/discussion and integrated concepts across the curriculum. Additionally, recommendations for implementation are provided regarding problem-based learning techniques and the adaptation of our approach to more general settings are addressed.

## 1. Introduction

Current research in cognitive science demonstrates that students learn best when they are actively engaged in a meaningful social context while constructing understanding and solving problems ([National Research Council 2001](#)). This construction of new knowledge from one's personal experiences is the foundation of constructivist theory ([Bransford 2000](#)). Eight characteristics differentiate constructivist learning environments: multiple representations of reality, complexity of the real world, emphasis on knowledge construction (instead of knowledge reproduction), authentic tasks in a meaningful context, avoid predetermined sequences of instruction, encourage thoughtful reflection, enable context- and content-dependent knowledge

construction, and support collaborative construction of knowledge through social negotiation (not competition among learners for recognition) ([Jonassen 1994](#)). In the classroom, the constructivist view of learning usually means encouraging students to use active techniques (experiments, real-world problem solving) to create more knowledge and then to reflect on and talk about what they are doing and how their understanding is changing. The instructor strives to create "a learning environment in which the student can learn to restructure the new information and their prior knowledge into new knowledge about the content and to practice using it" ([McManus 2001, p. 425](#)). The base of such an educational philosophy assumes learners are self-motivated and self-directed. Learners have diverse backgrounds and are interested in contextual learning, in other words, they want to see relevance ([Knowles 1973](#)).

We utilized a framework developed by McCarthy, based on Kolb's work ([Kolb 1984](#)), called the 4MAT system ([Guild & Garger 1998](#)) to employ a constructivist learning environment. McCarthy describes 4MAT as a process for delivering instruction in a way that engages, informs, and allows for practice and creative use of material learned within each lesson. Students journey through the learning process by asking four simple questions. Why? Learners seek a reason or motivation for learning. What? Learners identify and seek knowledge. How? Learners actively try out and apply knowledge allowing them to understand how they individually are going to use what they are learning. If? Learners develop extensions of their learning to create new experiences. Our aim was to provide students with a problem-based learning set of experiences designed to engage them with authentic problems that would motivate them and require them to engage with the material to apply statistics in a context they are likely to see as practicing professionals.

For several years, the biostatistics faculty at the Cleveland Clinic Lerner College of Medicine (CCLCM) of Case Western Reserve University struggled to deliver an engaging course. In a program designed to motivate and train the physician investigators of tomorrow ([Fishleder, Henson, & Hull 2007](#)), the biostatistics course is an integral component to the mission of the school. Concerned that the course may dispirit students from research, faculty decided to redesign the course format. The goal was to align the biostatistics course with the educational philosophy of CCLCM by developing a curriculum emphasizing active learning techniques and to foster a student's inherent interest in and perception of the role of statistics in medical research.

This paper describes the development of a new biostatistics course format and the learning theory that provided a framework for that design. Data is presented regarding students' responses to the course itself, their shifting perceptions' of the discipline and their certification examination performance. In addition, recommendations for implementation are provided and the adaptation of our approach to more general settings is addressed.

## **2. METHODS**

### **2.1 Course Design**

CCLCM was created in 2002 to increase the number of physician-investigators in the U.S. With an annual class size of 32, this unique five-year program emphasizes early clinical experiences,

discovery learning methods, research experiences, and competency-based assessment ([Fishleder et al. 2007](#)). The program contains a formal research curriculum extending from year one to graduation. Providing a foundation for this research curriculum are the first two summers, the basic and translational research block and the clinical research block, respectively.

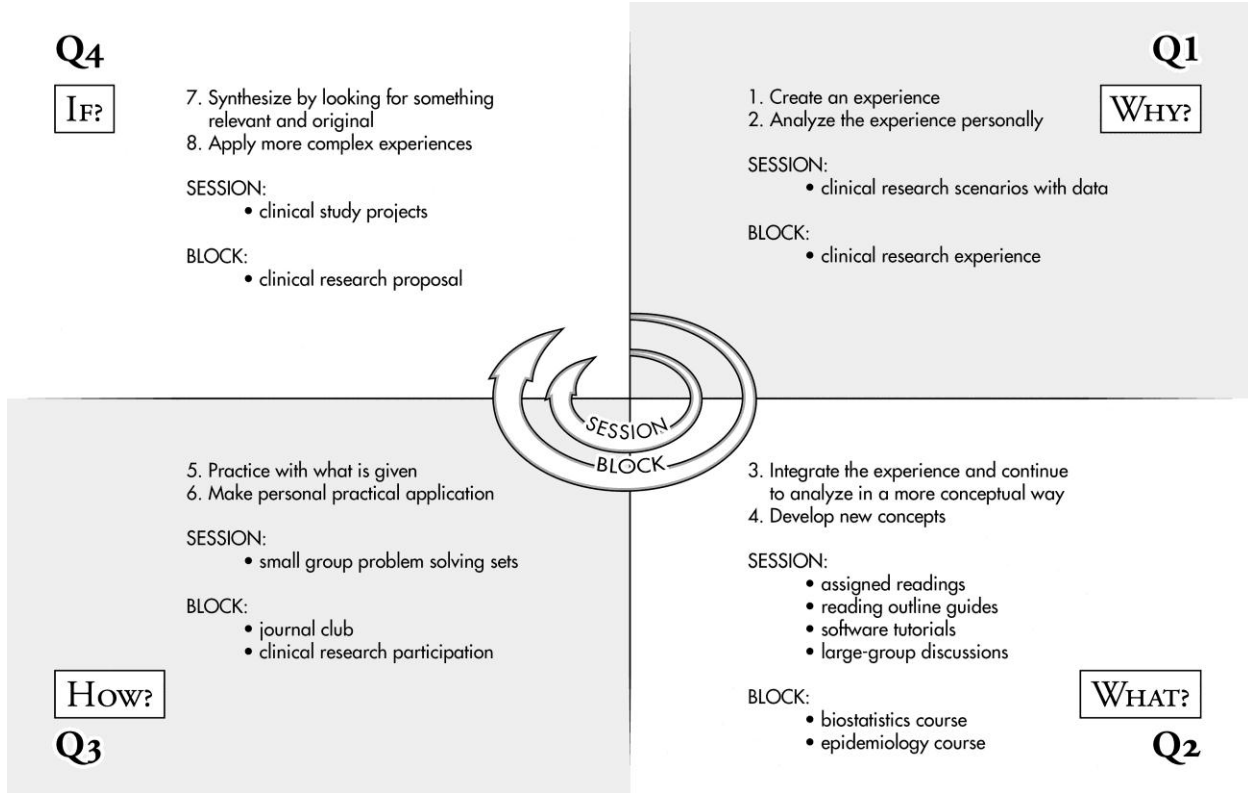
The year two summer is a nine-week clinical research block consisting of four components: biostatistics, epidemiology, journal club, and a clinical research experience. The biostatistics course meets for 2 hours twice a week covering in detail typical introductory biostatistics topics such as central tendency, variability, probability, z-scores, sampling distributions, estimation, and hypothesis testing. Also introduced are nonparametric tests, linear regression, logistic regression and survival analysis. The statistical software package JMP, a point and click interactive product of SAS, is utilized. Each student is assigned to an in-class working group (4 students) and an out-of-class project group (2 students). Additionally, every other week the class divides into two sections to present student projects. The presentations are the primary venue for student assessment. While there are no grades at CCLCM, student progress is determined through competency assessments. Students create individual learning portfolios that provide ongoing documentation of their developing skills and expertise ([Dannefer & Henson 2007](#)).

## **2.2 Instructional Activities/Materials**

The faculty organized instructional activities using a framework that develops the 4MAT system into a guideline for educational planning ([Armstrong & Parsa-Parsi 2005](#)). This design incorporates a range of carefully planned teaching strategies. In essence, the new format shifts the goals of instruction from emphasizing facts and concepts during class with application/integration occurring outside of the classroom (homework), to preparation outside of class with emphasis on problem solving and application within the classroom.

This section describes the course activities designed to engage students as they progress through the four questions of the 4MAT system ([Figure 1](#)). Progression occurs and is described at two levels: throughout each session and throughout the entire summer block. While the session level is general and applicable to most courses, it is recognized that the summer block level is rather specific to the CCLCM program.

**Figure 1. Learning goals and activities that align with educational planning framework at both the session and block level.**



***Quadrant 1: Personal meaning and motivation should be set for what is to follow ... the Why?***

All biostatistics sessions begin with the presentation of an actual clinical research scenario with accompanying data. Students more readily see the significance of what they are about to learn and tend to be more motivated.

The clinical research experience, where each student joins an active clinical research group at the Cleveland Clinic, is one of the four components of the summer which motivates the block. Students are joining projects at their various stages, thus during this experience students are involved in an assortment of activities such as reading clinical research papers, participating in study design, obtaining internal review board (IRB) approval, collecting, coding or analyzing data and discussing results at project meetings (see [Appendix A](#) for sample clinical research experience descriptions). Being personally involved in a research project provides a context within which to frame the conceptual material of the didactic courses; thus, students are not learning material in isolation.

***Quadrant 2: Acquisition of new knowledge and concepts ... the What?***

Prior to each class, students receive a reading assignment covering the topics of that session. Students are provided with reading outline guides to accompany their reading. These guides relate the topic to the previous lessons and build on prior knowledge as well as aid in

identification of core concepts (see [Appendix B](#) for sample reading outline guide). This component is in response to student feedback expressing difficulty with statistical reading comprehension and key point recognition resulting in hesitation of application. Additionally, students receive statistical software tutorials prior to class with instructions for obtaining and navigating software output. When class time is already limited, how best to integrate new software? By providing tutorials, the student is in charge of his/her learning and does not consume valuable class time, thus allowing the focus to remain on concepts and application instead of programming and troubleshooting.

After presenting the clinical research scenario with data, most biostatistics sessions include a large group discussion. Here, students identify how the scenario differs from previous scenarios (integrating the experience) and why previous analytical approaches are now inappropriate. Then, faculty briefly review the new concepts from the reading in the context of the presented clinical research scenario.

The biostatistics and epidemiology courses represent the acquisition of knowledge components for the clinical research block at large. The clinical research experience where students are immersed in and contribute towards an actual research project would not be possible without well-organized and substantive courses supporting it.

### ***Quadrant 3: Practical application ... the How?***

After each large group discussion, students break into small working groups and are provided with problem-solving sets. This is an opportunity for students to practice on real data pulled from the literature (citations always provided). All problems are in the context of a real clinical research study with guiding questions to identify and direct the analyses (see [Appendix C](#) for sample small-group problem-solving set). The vignettes are designed to provide both interesting medical facts and a clinical review of various disease processes, symptoms or treatments, so as to achieve contextual learning and integration. Faculty and facilitators are available to answer clarifying questions or redirect a group when necessary.

The journal club and activities of the clinical research experience affords an opportunity for students to apply their new knowledge and skills. Although each individual student may only experience a portion of the clinical research process (study design, writing a protocol, enrolling subjects, collecting data, analyzing data, manuscript preparation, etc.) due to the 9 week time limitation, the collective class achieves full exposure with crosstalk among students proving to be valuable exposure.

### ***Quadrant 4: Synthesis and extension ... the If?***

Students are also assigned to project groups that convene outside of class. Faculty delegate to each project group a real clinical study with larger data sets and more variables. The projects consist of a number of questions more closely resembling those that arise in clinical practice (not so guided) giving the students an opportunity to reflect on previous sessions and to apply the concepts to more complex experiences (see [Appendix D](#) for sample project group assignment). It is here with extension and reflection that the new knowledge becomes prior knowledge as the

learner restarts the cycle ([Figure 1](#)). Thus, after mastering one topic, the learner builds on that knowledge and progresses.

Within the clinical research experience, students work closely with their physician advisors to develop a focused clinical research proposal for written and oral presentation at the end of the summer block. The proposals are *not* to be about the students' summer research; instead, it is a proposed future study (often an extension of their summer work) focusing on core issues of hypotheses generation, study design and data analysis. The purpose is to integrate further the somewhat abstract concepts developed in the formal coursework with the scientific needs and medical environment of a clinical research area. These proposals often require complex designs and analyses not addressed in the introductory courses; hence, providing an opportunity to assess the students' understanding of the material and prompting students to initiate discussion on more advanced topics.

### 2.3 Evaluation

With IRB approval, students completed two web-based questionnaires. Formal end-of-course evaluations were utilized to assess student satisfaction using a four-point scale ranging from 1 (poor) to 4 (excellent). Evaluations assessed course components, teaching effectiveness, resources, and student likes and dislikes. A survey instrument, the Attitudes Toward Statistics (ATS) ([Wise 1985](#)), was administered pre- and post- course to assess student perception of the field. ATS is a self-administered questionnaire consisting of 29 statements. Students rate their perceptions about each statement using a scale from 1 (strongly disagree) to 5 (strongly agree). The instrument has two subscales: *Field* – reflecting attitude toward the usefulness of statistics in general or in the student's field; and *Course* – reflecting attitude toward the statistics course the student is currently enrolled.

The United States medical licensing exam step 1 (USMLE1) is the first of three examinations that medical students must successfully pass to qualify for licensure. This exam assesses one's ability to apply knowledge, concepts, and principles that are important in health and disease and that constitute the basis of safe and effective patient care. Results from the Biostatistics and Epidemiology component of the USMLE1 assess student knowledge.

The analyses utilized course evaluation data on five cohorts longitudinally, class of 2010 through class of 2014. Means and standard errors summarize all web-based questionnaires. ATS data were collected pre- and post- course for the class of 2013 and statistically compared using the Wilcoxon signed-rank test for ordinal data. Reliability of the ATS subscales was estimated by Cronbach's alpha. All data were de-identified for analyses. Statistical analyses were performed using SAS version 9.2 (SAS Institute, Inc., Cary, NC).

## 3. RESULTS

Student response to the redesigned course was remarkably optimistic. [Table 1](#) presents means and standard errors of pertinent course evaluation questions for the five offerings of the biostatistics course. Students were asked to "Rate the extent to which the biostatistics seminars helped you learn clinical research". After the redesign, there was a dramatic improvement. The



class of 2013 and 2014 report a weighted mean score for this statement of 3.6 compared to the three previous years (classes of 2010 – 2012) for which the weighted mean score was only 2.2. Most notably, the classes receiving the newly designed course thought more strongly that the seminars were well-organized, encouraged participation/discussion and integrated concepts across the curriculum. These classes largely favored the problem-solving approach to learning (weighted mean 3.6) and 84% disagreed or strongly disagreed that a didactic or lecture style would have been a preferred approach (weighted mean 2.0). Results indicate a significant improvement in all aspects after the course redesign.

**Table 1. Sampling of course evaluation questions and student responses over the last five years.**

	Class of 2010	Class of 2011	Class of 2012	Redesigned Course	
				Class of 2013	Class of 2014
	mean (SE)	mean (SE)	mean (SE)	mean (SE)	mean (SE)
Class size of 32 students - number completing survey (%)	29 (91%)	27 (84%)	26 (81%)	32 (100%)	30 (94%)
Rate the extent to which the biostatistics seminars helped you learn clinical research concepts	<b>2.2</b> (0.2)	<b>2.3</b> (0.2)	<b>2.2</b> (0.2)	<b>3.7</b> (0.1)	<b>3.4</b> (0.2)
The required readings (textbooks, journal articles, handouts, etc.) contributed to my learning	<b>2.2</b> (0.2)	<b>2.0</b> (0.2)	<b>2.3</b> (0.2)	<b>3.0</b> (0.2)	<b>3.2</b> (0.1)
The seminars were well organized	<b>2.2</b> (0.2)	<b>2.4</b> (0.1)	<b>2.4</b> (0.2)	<b>3.8</b> (0.1)	<b>3.7</b> (0.1)
The seminars encouraged student participation/discussion	<b>2.8</b> (0.1)	<b>2.3</b> (0.1)	<b>2.0</b> (0.2)	<b>3.5</b> (0.1)	<b>3.6</b> (0.1)
The seminars helped me integrate concepts across the curriculum	<b>2.6</b> (0.2)	<b>2.4</b> (0.1)	<b>2.5</b> (0.2)	<b>3.5</b> (0.1)	<b>3.4</b> (0.1)
The problem-solving session approach contributed to my learning				<b>3.7</b> (0.1)	<b>3.5</b> (0.1)
I would have preferred a more didactic/lecture style approach				<b>2.0</b> (0.2)	<b>1.9</b> (0.1)

1 = Strongly Disagree 2 = Disagree 3 = Agree 4 = Strongly Agree; SE = Standard Error

For the pre-course and the post-course questionnaire, [Table 2](#) shows mean scores and the mean-difference (post-pre) along with the Wilcoxon signed-rank level of significance (2-tailed). Students obtained statistically significant increases of practical importance in both the ATS Field and ATS Course subscales; 2.9 and 3.6 units respectively. Demonstrating high reliability of approximately 0.90, Cronbach's alpha for the two subscales is consistent with other studies ([Wise 1985](#)) ([Vanhoof, Sotos, Onghena, Verschaffel, Van Dooren, & Van den Noortgate 2006](#)). Six out of twenty-nine statements observed a statistically significant difference. These data indicate that after having taken the biostatistics course, the CCLCM class of 2013 more strongly agrees that with respect to their medical careers, statistical training will be worthwhile. Conversely, after the course, they more strongly disagree that the thought of another statistics course makes them nervous; statistics seems mysterious, they have trouble seeing how statistics relates to medicine; another course would be an unpleasant experience, and that they wish they could have avoided taking this course. Marginal significance was achieved regarding agreement that statistical training will be useful, relevant to their performance, and will better help them understand the medical research; along with disagreement that statistics is too complicated. Given the importance of preparing medical practitioners to understand and be able to use

statistics, improvement in student attitudes and motivation alone are sufficiently useful outcomes to justify the reform effort.

**Table 2. Attitudes Toward Statistics before and after the redesign of the biostatistics course, Cleveland Clinic Lerner College of Medicine (n = 32).**

	Pre-course Mean*	Post-course Mean*	Difference (Post - Pre)	P value**
<b>Total Score</b>				
Attitudes Toward Statistics score (29 statements)†	112.4	118.9	6.5	<b>0.0008</b>
<b>Subscales</b>				
Attitudes Toward Field (20 statements)†	80.8	83.7	2.9	<b>0.04</b>
Cronbach's Alpha	0.91	0.93		
Attitudes Toward Course (9 statements)†	31.7	35.3	3.6	<b>&lt;.0001</b>
Cronbach's Alpha	0.88	0.89		
<b>Individual Statements</b>				
I feel that statistics will be useful to me in my profession	4.44	4.66	0.22	0.09
The thought of being enrolled in a statistics course makes me nervous (R)	3.03	3.72	0.69	<b>&lt;.0001</b>
A good researcher must have training in statistics	4.41	4.59	0.19	0.24
Statistics seems very mysterious to me (R)	3.28	4.00	0.72	<b>0.0004</b>
Most people would benefit from taking a statistics course	4.09	4.19	0.09	0.70
I have difficulty seeing how statistics relates to my field of study (R)	4.28	4.59	0.31	<b>0.008</b>
I see being enrolled in a statistics course as a very unpleasant experience (R)	3.31	4.09	0.78	<b>0.0002</b>
I would like to continue my statistical training in an advanced course	2.88	3.09	0.22	0.29
Statistics will be useful to me in comparing the relative merits of different objects, methods, programs, etc.	4.22	4.28	0.06	0.81
Statistics is not really very useful because it tells us what we already know anyway (R)	4.16	4.38	0.22	0.17
Statistical training is relevant to my performance in my field of study	4.09	4.41	0.31	0.06
I wish that I could have avoided taking my statistics course (R)	3.75	4.31	0.56	<b>0.007</b>
Statistics is a worthwhile part of my professional training	4.25	4.50	0.25	<b>0.04</b>
Statistics is too math oriented to be of much use to me in the future (R)	4.28	4.28	0.00	1.00
I get upset at the thought of enrolling in another statistics course (R)	3.91	3.88	-0.03	0.87
Statistical analysis is best left to the "experts" and should not be part of a lay professional's job (R)	3.84	3.91	0.06	0.79
Statistics is an inseparable aspect of scientific research	4.38	4.56	0.19	0.24
I feel intimidated when I have to deal with mathematical formulas (R)	3.69	3.88	0.19	0.20
I am excited at the prospect of actually using statistics in my job	3.47	3.59	0.13	0.47
Studying statistics is a waste of time (R)	4.31	4.50	0.19	0.18
My statistical training will help me better understand the research being done in my field of study	4.19	4.47	0.28	0.08
One becomes a more effective "consumer" of research findings if one has some training in statistics	4.28	4.44	0.16	0.31
Training in statistics makes for a more well-rounded professional experience	4.34	4.38	0.03	0.96
Statistical thinking can play a useful role in everyday life	3.91	3.94	0.03	1.00
Dealing with numbers makes me uneasy (R)	4.06	4.22	0.16	0.27
I feel that statistics should be required early in one's professional training	3.84	3.88	0.03	0.83
Statistics is too complicated for me to use effectively (R)	3.78	4.06	0.28	0.07
Statistical training is not really useful for most professionals (R)	3.78	3.94	0.16	0.39
Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write	2.19	2.22	0.03	0.73

\* 1 = Strongly Disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Strongly Agree; reverse-keyed items indicated by an "(R)"

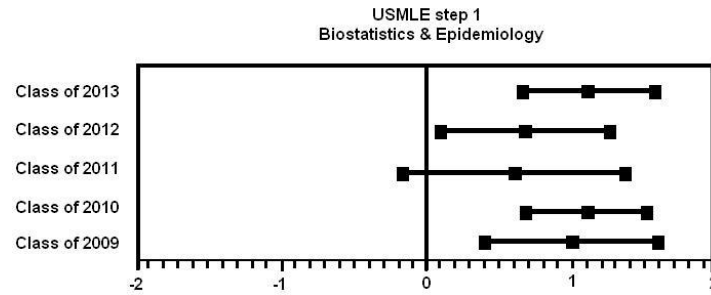
\*\* Score data statistically compared using the Wilcoxon signed-rank test for ordinal data (nonparametric test)

† Score data adjusted for reverse keyed items

Performance of our students on the USMLE1 has remained strong (Figure 2). Because CCLCM students do represent highly motivated and intelligent students who would excel in most instructional settings, the USMLE1 scores are presented more to ensure no harm. The Class of 2013, the first cohort receiving the new format for which these scores are available, has demonstrated outstanding marks. Since performance was already high with little room to improve, it was important that the redesign did not have a substantial negative impact. Scores for the Class of 2014 are not yet available.



**Figure 2. Score distribution of first takers from CCLCM relative to the distribution for all U.S./Canadian first takers.**



All scores are scaled in standard score units. The mean performance is represented by the solid vertical line at 0. Roughly 68% scored within one standard deviation of the mean, between -1 and 1. The center box on each line depicts the mean performance of first takers from CCLCM and the distance from the center box to one end of the line indicates one standard deviation.

## 4. DISCUSSION

This curriculum redesign demonstrates the feasibility of applying a theoretical framework to educational planning to transform a biostatistics course into a problem-based learning experience. Overall, there was high praise for the new format. Students commended the student-centered approach, the substantial student involvement/engagement and the heavy emphasis on application. As course director, I was astounded by the students' initiative in their own learning and the breadth and depth of class discussion. The teaching experience was both exciting and challenging, as the discussions were student directed and less in my control. While finding the motivation to redesign a course is difficult, particularly when reward structures do not encourage taking the time needed for extensive change ([Dancy & Henderson 2008](#)), encouragement and lessons learned follow for those eager and interested in implementation.

You can "cover" enough material. Philosophically, one must decide; is it about transferring facts/details or learning how to think? An instructor-centered passive learning technique is a more efficient form of information transmission, but is that really your goal? Recommendations are to spend class time only on the most critically important and conceptually difficult parts of the lesson, leaving the students to cover the rest for themselves ([Felder & Brent 1996](#)). The class time saved with such an approach should be more than sufficient for problem-based learning exercises. Through trial and error, we were able to streamline the sessions by identifying and then deemphasizing topics students grasped from the reading/preparation and targeting the challenging concepts for the classroom. In this instance, the classes were more lively and effective while covering the entire syllabus.

The material is not too technical for such methods. Initially, there was reluctance that a student-centered approach would work for biostatistics inundated with terminology, formulas and facts. However, the facts were the easiest for students to grasp in their session preparation. Some take comfort in discovering that lectures can have a place as part of the learning process (quadrant 2 in [Figure 1](#)) so long as one does not begin and end the experience there. Following up an explanation of technical concepts with an application activity better engages the learner and facilitates assessment of the student's knowledge acquisition.

Initial development can be time intensive depending upon the technique. To implement case discussion, problem-based learning (PBL) or problem-solving sets, one must dedicate significant faculty time to researching and writing the activities; although once completed, cases are available for repeated use. For the development of our course material, one primary faculty member received one day a week protected time for ten months. Positives and negatives exist to such an approach. Although this places a great deal of responsibility on one individual, it does help create a coherent curriculum. Others prefer the use of diverse faculty teams to ease the burden. If unready or unable to make such time commitments, other active learning techniques such as question-based outlines (Huerta 2007) or minute-papers (Stead 2005) can easily be incorporated into any session with minimal development. Publishing more examples in the medical literature will allow faculty to borrow ideas, designs and materials making activities more accessible and further reducing the necessary time for development. Similarly, online venues are emerging where educators can publish peer reviewed instructional materials (e.g. MedEdPortal, <https://www.mededportal.org/>, Health Education Assets Library (HEAL), <http://www.healcentral.org/>, Multimedia Educational Resource for Learning and Online Teaching (MERLOT), <http://www.merlot.org>).

Many of the techniques have students working in small groups, which does not have to be too faculty intensive. Based on our experience, we recommend at least one faculty for every 4 working groups, a nontrivial request. Our solution is the utilization of group facilitators or teaching assistants, an inexpensive and efficient alternative allowing the course to be conducted with only *one* faculty member who is more of a team leader. The balance between saving faculty time and facilitator or teaching assistant training and development, however, must be considered.

Active learning techniques exist for classes of all sizes. Based on our experience, the small working groups approach is best suited for groups of 4 – 6 students. With fewer than four, the workload becomes overwhelming and group discussion suffers, whereas with more than six, not every student has a role within the group and achieving group consensus becomes difficult. Some components of our hybrid design (outline reading guides and large-group discussions) are amenable to a class of any size. Small-group projects are still feasible but with written solution reports submitted instead of oral presentations. Alternatively, Marantz et al. (2003) describes the successful implementation of case-discussion in large cohorts.

While the results presented do not prove that the new format caused the increase in student satisfaction and student perception, we feel that such substantial changes are largely due to the instructional approach. It is possible that the observed differences reflect the new faculty course director or are cohort specific. However, different faculty taught the course with each cohort prior to the revamping and as seen in Table 1, faculty and student cohort changes over the previous three years had negligible effect.

Evaluation of the course redesign has been challenging considering some of the unique characteristics of the CCLCM program. Common practice is to report final course grades before and after an intervention to assess the knowledge component of a course. As mentioned previously, there are no grades at CCLCM, thus such an internal measure is unavailable. Alternatively, we present an external measure of performance (USMLE1 scores).

A reasonable assumption is that the pre-course ATS data is representative of any CCLCM class before taking a biostatistics course. However, with no data on previous years, the observed pre-to post-course changes, while informative, could be the result of simply taking a biostatistics course and not necessarily the result of the instructional approach. Therefore, the ATS results indicate the success of the new course format and are not comparable to previous formats.

## **5. CONCLUSION**

Problem-based learning requires more than just listening; students must prepare in advance and discuss/solve problems. The 4MAT framework helps design classroom activities to support students throughout the learning process. This approach supports motivated learners to acquire, apply, generalize and reflect on new knowledge. The model, however, applies to any discipline and its flexibility allows the incorporation of various teaching techniques. We hope to inspire others to build on learning theories and to engage students. Knowledge disseminated does not equate to knowledge learned.

## **APPENDIX A**

### **Sample Clinical Research Experience Descriptions**

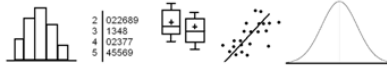
#### Example 1

My clinical research experience involved a study where the aim was to determine whether adjunctive passive cooling during transport affects temperature variability and neurodevelopmental outcomes in infants with hypoxic ischemic encephalopathy treated with hypothermia therapy. My role in this study was to conduct a literature review of hypothermia therapy for infants with hypoxic ischemic encephalopathy, and then to design a database for collection on a variety of relevant variables (including birth parameters, temperatures at initiation and throughout cooling, hospital course, and 12 month neurodevelopmental outcomes). Subsequently, I conducted a retrospective chart review for a small cohort of fourteen infants born in outborn hospitals and transported to the Cleveland Clinic neonatal intensive care unit. I was able to perform the data analysis resulting in an abstract submitted to the American Society of Pediatrics Research.

#### Example 2

My clinical research experience involved a study where the aim was to investigate whether surgery or a sling is the better treatment for clavicle fractures as determined by frequency of patients developing complications. Additionally, we assessed whether a variety of factors such as fracture type, mechanism of injury, and others could be predictive for the development of complications. Working with an orthopedic traumatologist at a level 1 trauma center in Cleveland, Ohio, my role in this study was to obtain a firm grasp of the relevant literature, requiring me to read, review, and synthesize numerous papers. Next, I created a data capture form and built an accompanying database. Then I reviewed patient charts, collected the relevant data, and performed data entry. I analyzed the data and presented the findings at one of our biweekly departmental orthopedic research seminars.

## APPENDIX B Sample Reading Outline Guide



CMED 402: Statistical Science in Medical Research

Seminar 11: Simple Linear Regression

Preparation:	Required Reading:  Rosner B. Fundamentals of Biostatistics. 6th ed. California: Thomson Brooks/Cole; 2006.  11 Regression and Correlation Methods, pp. 464 – 487, 492 - 496
--------------	---

There is often an approximately linear relationship between variables from a population. Simple linear regression allows us to quantify such relationships.

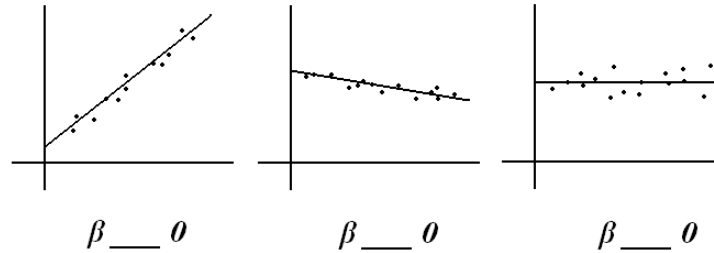
The full linear regression model takes the form:  $y = \alpha + \beta x + e$

We fit the linear model to our data to obtain:  $\hat{y} = E[y|x] = a + bx$

Match the simple linear regression component to its definition.

- |           |       |   |
|-----------|-------|---|
| $y$       | _____ | A. The value of the independent variable                  |
| $a$       | _____ | B. The estimate of the intercept (regression coefficient) |
| $\beta$   | _____ | C. The random error                                       |
| $x$       | _____ | D. The value of the dependent variable                    |
| $e$       | _____ | E. The slope  |
| $\hat{y}$ | _____ | F. The estimate of the slope (regression coefficient)     |
| $E[y x]$  | _____ | G. The average value of $y$ for a given value of $x$      |
| $a$       | _____ | H. The average value of $y$ for a given value of $x$      |
| $b$       | _____ | I. The intercept  |

The interpretation of the regression line for different value of  $\beta$  ( $=, <, >$ )



For any sample point  $(x_i, y_i)$ , the residual component of that point about the regression line is defined by  $y_i - \hat{y}_i$ . This is the difference between the actual value and the predicted value. The regression component of that point about the regression line is defined by  $\hat{y}_i - \bar{y}$ . The best fitting regression line has large regression components and small residual components. The worst fitting regression line has small regression components and large residual components. One approach to quantifying how good a regression line fits the data is to square the deviations about the mean,  $y_i - \bar{y}$ , sum them up over all points, and decompose this sum of squares into regression and residual components.

Decomposition of the Total Sum or Squares into Regression and Residual Components:

$$\sum_{i=1}^n (y_i - \bar{y})^2 \text{ is the } \underline{\hspace{2cm}} \text{ sum of squares (Total SS)}$$

$$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \text{ is the } \underline{\hspace{2cm}} \text{ sum of squares (Reg SS)}$$

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 \text{ is the } \underline{\hspace{2cm}} \text{ sum of squares (Res SS)}$$

$R^2$  can be thought of as the proportion of the variance of  $y$  that can be explained by  $x$ .  
 If  $R^2 = \underline{\hspace{1cm}}$ , then all variation in  $y$  can be explained by variation in  $x$ , and all data points fall on the regression line. If  $R^2 = \underline{\hspace{1cm}}$ , then  $x$  gives no information about  $y$ , and the variance of  $y$  is the same with or without knowing  $x$ .



The simple linear regression model is a great tool, but its answers will only be useful if it is the right model for the data. We need to check the assumptions before using the model. The four assumptions are:

- (1)
- (2)
- (3)
- (4)

The main question of interest in simple linear regression is whether or not a linear relationship exists between  $x$  and  $y$ . This can be tested by the slope parameter.

$H_0: \beta = 0$  (No linear association between  $x$  and  $y$ )

$H_1:$

Compute the test statistic:  $t = \frac{b}{SE(b)}$  where  $SE(b) = \sqrt{\frac{\text{Res SS} / n - 2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$

Compare to critical value:  $t_{n-2, 1-\alpha/2}$

If  $H_0: \beta = 0$  is not rejected, what is the best “guess” for the value of the response variable for any value of the predictor variable?

Confidence Intervals for the Slope of the Regression Line:

Lower Bound:

Upper Bound:

One important use for regression lines is in making predictions. For a given value of  $x$ , an estimate from the regression line is denoted  $\hat{y} = ax + b$ . Frequently, the accuracy of these predictions must be assessed. How accurate a prediction is depends on whether we are making predictions for **one specific subject** or for the **mean value of all subjects of a given  $x$** . Intervals for an individual from the population are called \_\_\_\_\_ intervals. Intervals for the mean of the population are called confidence intervals.

Both are of the form:  $\hat{y} \pm t_{n-2, 1-\alpha/2} \cdot SE(\hat{y})$

but when predictions are made from regression lines for individual observations,

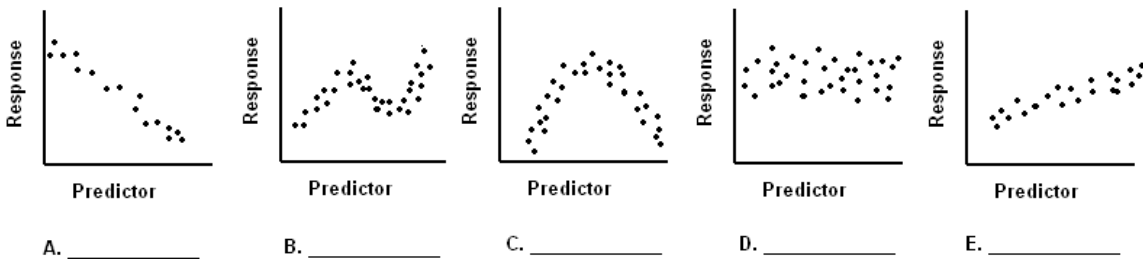
$$SE(\hat{y}) = \sqrt{\frac{\text{Res SS}}{n-2} \left[ 1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right]}$$

and when predictions are made from regression lines for the mean value of  $y$  for a given  $x$ ,

$$SE(\hat{y}) = \sqrt{\frac{\text{Res SS}}{n-2} \left[ \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right]}$$

Simple Linear Regression focuses on predicting one dependent variable ( $y$ ) from an independent variable ( $x$ ). Often we are interested not in predicting one variable from another but rather in investigating whether or not there is a relationship between two variables. The \_\_\_\_\_ is a useful tool for quantifying the **linear** relationship between two variables and is better suited for this purpose than the regression coefficient.

Label each scatter diagram as a linear, nonlinear or no relation between predictor and response.



Properties of the Linear Correlation Coefficient

- Always between \_\_\_ and \_\_\_, inclusive
- If  $r = \underline{\hspace{1cm}}$ , there is a perfect positive linear relation between the two variables
- If  $r = \underline{\hspace{1cm}}$ , there is a perfect negative linear relation between the two variables
- If  $r > 0$ , then the variables are said to be positively correlated – as  $x$  increases,  $y$  tends to \_\_\_\_\_, whereas as  $x$  decreases,  $y$  tends to \_\_\_\_\_
- If  $r < 0$ , then the variables are said to be negatively correlated – as  $x$  increases,  $y$  tends to \_\_\_\_\_, whereas as  $x$  decreases,  $y$  tends to \_\_\_\_\_
- If  $r$  is close to \_\_\_, this implies no *linear* relation between the two variables.

True    False      For the least-squares regression model, we require that the independent variable,  $x$ , be normally distributed.

True    False       $R^2$  is defined as Res SS / Total SS

True    False      The point  $(\bar{x}, \bar{y})$  falls on the regression line.

- True False A regression coefficient is another name for a correlation coefficient.
- True False A correlation coefficient of 0 means no relation
- True False The sample correlation coefficient is only meaningful if the two variables are normally distributed

## APPENDIX C

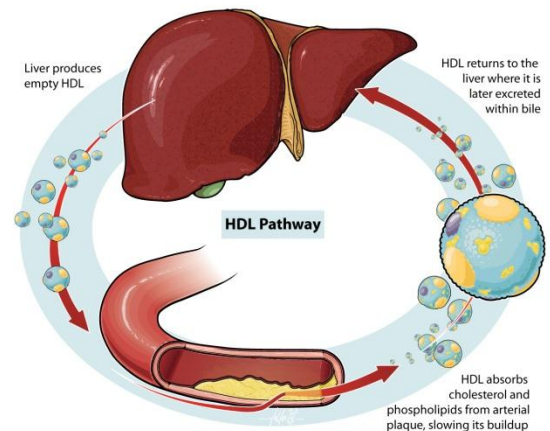
### Sample Small-group Problem-solving Set

CMED 402: Statistical Science in Medical Research  
Seminar 11: Simple Linear Regression

Objectives:	<ul style="list-style-type: none"> <li>• Test the claim that a linear relation exists between two quantitative variables</li> <li>• Compute a confidence interval about the slope</li> <li>• Construct confidence intervals about a predicted value</li> <li>• Construct prediction intervals about a predicted value</li> <li>• Compute and interpret the linear correlation coefficient</li> </ul>
-------------	--

1. **Age versus HDL Cholesterol** Cholesterol can not dissolve in the blood. It has to be transported to and from the cells by carriers called lipoproteins. About one-fourth to one-third of blood cholesterol is carried by high-density lipoprotein (HDL). HDL cholesterol is known as “good” cholesterol, because high levels of HDL seem to protect against heart attack. Low levels of HDL (less than 40 mg/dL) also increase the risk of heart disease. Medical experts think that HDL tends to carry cholesterol away from the arteries and back to the liver, where it's passed from the body. Some experts believe that HDL removes excess cholesterol from arterial plaque, slowing its buildup. An investigator wants to determine whether there is a relation between a male’s age and his HDL cholesterol. He randomly selects 17 of his patients and determines their HDL cholesterol. He obtained the following data:

Age	HDL Cholesterol	Age	HDL Cholesterol
38	57	38	44
42	54	66	62
46	34	30	53
32	56	51	36
55	35	27	45
52	40	52	38
61	42	49	55
61	38	39	28
26	47		



**Source:** Data based upon information obtained from the National Center for Health Statistics

Look at a scatter diagram of the data, treating age as the predictor variable. What type of relation, if any, appears to exist between age and HDL cholesterol?

What type of experiment does this represent – independent or dependent samples?

Take the clinical question and put it into statistical hypotheses that can be formally tested. The hypotheses are:

$$H_0:$$

$$H_1:$$

Test the claim that a linear relation exists between age ( $x$ ) and HDL cholesterol ( $y$ ) at the  $\alpha = 0.05$  level of significance. Using technology, find:

- The test statistic:  $t =$
- Degrees of freedom:  $df =$
- Critical value  $t_{n-2, 1-\alpha/2} =$
- The  $p$ -value:  $p$  - value =
- Conclusion:

What are the unbiased estimates of  $\alpha$  (intercept) and  $\beta$  (slope)?

Obtain a 95% confidence interval about the slope of the true least-squares regression line. Does your 95% CI include the null value of 0? How do these results compare with that of hypothesis testing?

Suppose a 42-year-old male patient visits the doctor's office. Would you recommend using the least-squares regression line to predict the HDL cholesterol of this patient? Why? What would be a good guess as to the HDL cholesterol of this patient?

As an exercise, predict the mean HDL cholesterol of males who are 51 years old using the regression equation?

Obtain a 95% confidence interval about the mean HDL cholesterol of males who are 51 years old.

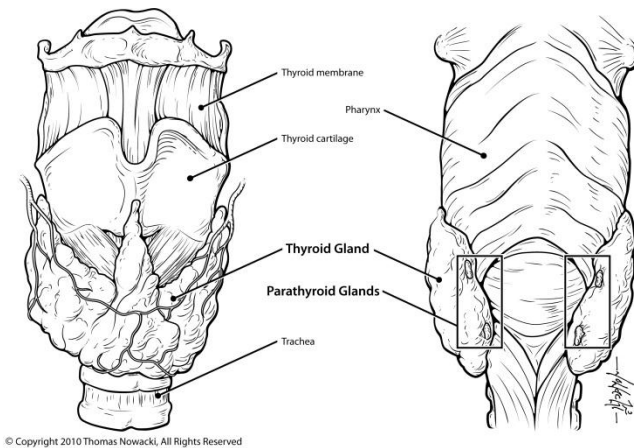
Predict the HDL cholesterol of a randomly selected male patient who is 51 years old using the regression equation?

Obtain a 95% prediction interval about the HDL cholesterol of a randomly selected male patient who is 51 years old.

Explain the difference between the two previous predictions. Which interval is wider, the confidence interval or prediction interval? Explain why you might think this is so.

Imagine that the researcher is interested in measuring the strength of the linear relation between two variables. What would be an appropriate statistic to report? Obtain an estimate of this measure and interpret the findings.

**2. Pediatrics, Endocrinology** Thyroid hormones are required for normal development of the human brain. Preterm infants are particularly vulnerable to an adverse neurodevelopmental outcome, but as yet, appropriate serum levels of thyroid hormones to achieve optimal brain maturation have not been quantified. Transient hypothyroxinemia, a common finding in premature infants, is not thought to have long-term consequences, or to require treatment. A study was performed to investigate whether hypothyroxinemia in premature infants is a cause of subsequent motor and cognitive abnormalities. Blood thyroxine values were obtained on routine screening in the first week of life from 526 infants who weighed 2000 g or less at birth and were born at 33 weeks gestation or earlier. The data presented concerns the relationship between mean thyroxine level and gestational age.



Gestational age	Mean thyroxine level ( $\mu\text{g/dL}$ )
$\leq 24^*$	6.5
25	7.1
26	7.0
27	7.1
28	7.2
29	7.1
30	8.1
31	8.7
32	9.5
33	10.1

\*Treat as 24 in subsequent analyses

Source: Reuss, M.L. *NEJM* 1996, 334(13): 821-27.

Using regression methods, derive a formula predicting mean thyroxine as a function of gestational age.

Interpret the intercept and slope.

What is  $R^2$  for this regression line? What does  $R^2$  mean?

Test for the statistical significance of the regression line using the  $t$  test. Present the test statistic,  $p$ -value and conclusion.

After looking at the scatterplot and the  $R^2$  value, did you expect your  $t$  test conclusion? Explain.



In order to use this regression model, the assumptions need to be checked. This simple example has only 10 data points; however, typically one has many more. To demonstrate that you can perform assumption checks, do the following:

1. Discuss with your group the assumption of independence.
2. Obtain a residual by predicted plot for assessing linearity and homoscedasticity.
3. Obtain a normal quantile plot of the residuals to assess normality.

## APPENDIX D

### Sample Project Group Assignment

The dataset represents data from The WHO Young Infants Study Group (“Clinical prediction of serious bacterial infections in young infants in developing countries”. *The Pediatric Infectious Disease Journal* 1999; 18(10): S23-S31).

Project 2: Acute Respiratory Infections (ARI)

#### QUESTIONS:

Using formal statistical inference, answer the following clinical questions. Be sure to present hypotheses, test statistic, p-value and conclusions when appropriate.

1. Assume that body mass index is approximately normally distributed. Use your sample to estimate the mean and variance, oximeter  $\sim N(\text{____}, \text{_____})$ . Use this information to answer the following questions. Show all of your work.
    - What is the probability that a randomly selected individual has a heart-rate oximeter between 150 and 170?
    - What is the probability a random sample of 30 individuals will have a mean heart-rate oximeter of less than 145?
  2. Obtain an estimate and 95% confidence interval on the proportion with definite lower chest wall indrawing. Obtain a 90% confidence interval on the proportion with definite lower chest wall indrawing. Which one is wider? Explain.
  3. In the US, infant convulsions occur in 1 per 100 live births for term babies; the incidence is even higher in preterm infants. Do infants from Ethiopia have a similar incidence? Do infants from Papua New Guinea have a similar incidence?
  4. Most doctors — and the [American Academy of Pediatrics \(AAP\)](#) — agree that a normal body temperature for a healthy baby is between 97 and 100.4 degrees Fahrenheit (36 to 38 degrees Celsius). If your baby's rectal temperature is above this range, he has a fever. Do the infants in this study with severe pneumonia have feverish temperatures?
  5. It is thought that an infant's ability to sustain a suck is associated with their weight. Is the weight of the infants who sustain a suck different from that of the infants unable to sustain a suck?
  6. Is there an association between infection status (none vs. cold/cough) and adjusted respiratory rate? Is there an association between infection status (none vs. pneumonia) and adjusted respiratory rate? Provide one possible explanation for your results.
-

## References

- Armstrong, E. and Parsa-Parsi, R. (2005), "How Can Physicians' Learning Styles Drive Educational Planning?," *Academic Medicine*, 80, 680-684.
- Bransford, J. (2000), *How people learn brain, mind, experience, and school* (Expanded ed.). Washington, D.C.: National Academy Press.
- Dannefer, E.F. and Henson, L.C. (2007), "The Portfolio Approach to Competency-Based Assessment at the Cleveland Clinic Lerner College of Medicine," *Academic Medicine*, 82, 493-502.
- Dancy, M.H. and Henderson, C. (2008), "Barriers and promises in STEM reform," Paper presented at the National Academies of Science Promising Practices Workshop. [http://www7.nationalacademies.org/bose/Dancy\\_Henderson\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Dancy_Henderson_CommissionedPaper.pdf).
- Felder, R.M. and Brent, R. (1996), "Navigating the Bumpy Road to Student-Centered Instruction," *College Teaching*, 44, 43-47.
- Fishleder, J., Henson, L.C., and Hull, A.L. (2007), "Cleveland Clinic Lerner College of Medicine: An Innovative Approach to Medical Education and the Training of Physician Investigators," *Academic Medicine*, 82, 390-396.
- Guild, P.B. and Garger, S. (1998), *Curriculum: McCarthy's 4mat system*. In: *Marching to Different Drummers*. 2<sup>nd</sup> Edition, Alexandria, VA: Association for Supervision & Curriculum Development.
- Huerta, J.C. (2007), "Getting Active in Large Lecture," *Journal of Political Science Education*, 3, 237-249.
- Knowles, M. (1973), *The Adult Learner: A Neglected Species*, Houston, TX: Gulf Publishing Company.
- Jonassen, D. (1994), Thinking technology. *Educational Technology*, 34(4), 34-37.
- Kolb, D.A. (1984), *Experiential Learning Experience as the Source of Learning and Development*, Englewood Cliffs, NJ: Prentice Hall.
- Marantz, P.R., Burton, W., and Steiner-Grossman, P. (2003), "Using the Case-discussion Method to Teach Epidemiology and Biostatistics," *Academic Medicine*, 78, 365-371.
- McManus, D.A. (2001), "The Two Paradigms of Education and the Peer Review of Teaching," *Journal of Geoscience Education*, 49, 423-434.
- National Research Council (2001), *Knowing what students know: The science and design of educational assessment*. Committee on the Foundations of Assessment. Pelligrino J, Chudowsky

N, and Glaser R., editors, Board of Testing and Assessment, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Stead, D.R. (2005), "A review of the one-minute paper," *Active Learning in Higher Education*, 6, 118-131.

Vanhoof, S., Sotos, A., Onghena, P., Verschaffel, L., Van Dooren, W., and Van den Noortgate, W. (2006), "Attitudes Toward Statistics and Their Relationship with Short- and Long-Term Exam Results," *Journal of Statistical Education* [Online], 14(3).

<http://www.amstat.org/publications/jse/v14n3/vanhoof.html>

Wise, S.L. (1985), "The development and validation of a scale measuring attitudes toward statistics," *Educational and Psychological Measurement*, 45, 401-405.

---

Amy S. Nowacki  
Department of Quantitative Health Sciences  
Cleveland Clinic and Cleveland Clinic Lerner College of Medicine  
9500 Euclid / JJN3 – 01  
Cleveland, OH 44195  
Email: [nowacka@ccf.org](mailto:nowacka@ccf.org)  
Phone: 216-444-3773  
Fax: 216-444-8021

---

[Volume 19 \(2011\)](#) | [Archive](#) | [Index](#) | [Data Archive](#) | [Resources](#) | [Editorial Board](#) | [Guidelines for Authors](#) | [Guidelines for Data Contributors](#) | [Guidelines for Readers/Data Users](#) | [Home Page](#) | [Contact JSE](#) | [ASA Publications](#)