

Adv in Health Sci Educ
DOI 10.1007/s10459-012-9382-z

Factors influencing medical student attrition and their implications in a large multi-center randomized education trial

A. Kalet · R. H. Ellaway · H. S. Song · M. Nick · U. Sarpel ·
M. A. Hopkins · J. Hill · J. L. Plass · M. V. Pusic

Received: 20 December 2011 / Accepted: 30 May 2012
© Springer Science+Business Media B.V. 2012

Abstract Participant attrition may be a significant threat to the generalizability of the results of educational research studies if participants who do not persist in a study differ from those who do in ways that can affect the experimental outcomes. A multi-center trial of the efficacy of different computer-based instructional strategies gave us the opportunity

The members of the “The WISE Trial Study Collaborative” are given in the “[Appendix](#)” section.

Electronic supplementary material The online version of this article (doi:[10.1007/s10459-012-9382-z](https://doi.org/10.1007/s10459-012-9382-z)) contains supplementary material, which is available to authorized users.

A. Kalet (✉) · M. Nick · M. A. Hopkins · J. Hill · J. L. Plass · M. V. Pusic
New York University, 545 First Avenue Greenberg Hall, Room 6R., New York, NY 10016, USA
e-mail: adina.kalet@nyumc.org

M. Nick
e-mail: mike.nick@nyu.edu

M. A. Hopkins
e-mail: maryann.hopkins@nyumc.org

J. Hill
e-mail: jennifer.hill@nyu.edu

J. L. Plass
e-mail: jan.plass@nyu.edu

M. V. Pusic
e-mail: mvp19@columbia.edu

R. H. Ellaway
Northern Ontario School of Medicine, Thunder Bay, ON, Canada
e-mail: rachel.ellaway@nosm.ca

H. S. Song
Georgian Court University, 900 Lakewood Avenue, Lakewood, NJ 08701, USA
e-mail: hyuksong@gmail.com

U. Sarpel
Mount Sinai School of Medicine, 1 Gustave L. Levy Place, New York, NY 10029, USA
e-mail: umut.sarpel@mountsinai.org

to observe institutional and student factors linked to attrition from a study and the ways in which they altered the participation profile. The data is from a randomized controlled trial conducted at seven US medical schools investigating the educational impact of different instructional designs for computer-based learning modules for surgical clerks. All students undertaking their surgical clerkships at the participating schools were invited participate and those that consented were asked to complete five study measures during their surgery clerkship. Variations in study attrition rates were explored by institution and by participants' self-regulation, self-efficacy, perception of task value, and mastery goal orientation measured on entry to the study. Of the 1,363 invited participants 995 (73 %) consented to participate and provided baseline data. There was a significant drop in the rate of participation at each of the five study milestones with 902 (94 %) completing at least one of two module post-test, 799 (61 %) both module post-tests, 539 (36 %) the mid-rotation evaluation and 252 (25 %) the final evaluation. Attrition varied between institutions on survival analysis ($p < 0.001$). Small but statistically significant differences in self-regulation ($p = 0.01$), self-efficacy ($p = 0.02$) and task value ($p = 0.04$) were observed but not in mastery or performance goal orientation measures ($p = \text{NS}$). Study attrition was correlated with lower achievement on the National Board of Medical Examiners subject exam. The results of education trials should be interpreted with the understanding that students who persist may be somewhat more self-regulated, self-efficacious and higher achievers than their peers who drop out and as such do not represent the class as a whole.

Keywords Computer-aided instruction · Generalizability · Medical education · Randomized controlled trial · RCT · Reproducibility of results · Study design · Surgery education

Introduction

Participant drop out or study attrition is a significant threat to the validity and generalizability of all research involving human subjects (Jüni et al. 2001), not least because participants who persist in studies are likely to differ in their behaviours (and potentially other attributes) from those who drop out. This “volunteer bias”—where those who join and persist in studies are healthier and more likely to adhere to medical advice than those who do not—is well understood in epidemiological studies (Shrank et al. 2011) as a threat to the generalizability or external validity of the findings. A detailed knowledge of the potential confounders allows researchers to control for their impact in both the design of the study methods and the interpretation of its findings. Analogous biases have been described in studies of medical education interventions where it has been demonstrated that those who join and persist in such studies are academically stronger to begin with and have better academic performance (Callahan et al. 2007; Cook et al. 2008).

Multi-institutional studies can be a highly valuable, although resource-intensive, way of evaluating the efficacy of educational interventions (Torgerson 2002; Todres et al. 2007; Bleakley et al. 2011). However, there may be considerable variation in student participation both within and between study sites. This participation variation can be a two-edged sword; on one hand it may confound the treatment effect leading to inaccurate results but on the other hand, discernible treatment effects that rise above the inter-institutional variation are more likely to generalize well (in particular if the variation is known and accounted for). Either way, accounting for the dynamics of student participation in single- or multi-institutional studies is an important part of ensuring their quality and utility.

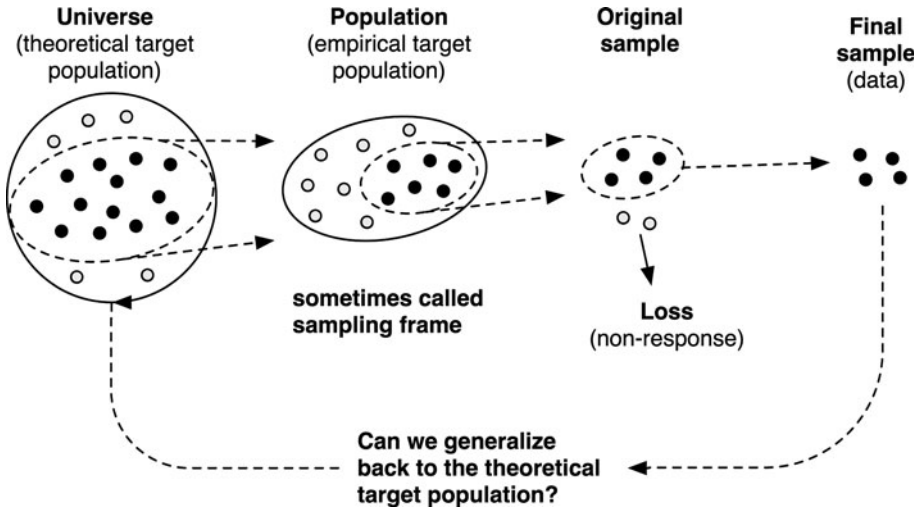


Fig. 1 Conceptual framework. Baseline measures of student traits were collected for all participants on entry to the study. Baseline measures for participants in the final sample were compared to those of participants who became lost/non-responsive at various points of the study

This paper considers volunteer bias as it impacted the WISE Trial, a large prospective trial that involved medical students at seven different US undergraduate medical schools using multimedia instructional materials during their surgery clerkship rotations. The original trial compared the impact of four theory-based instructional design features on medical student learning in surgical clerkships. In this paper we use the WISE Trial dataset to describe institutional and student factors associated with student participation and attrition rates. The goal of this paper is to inform the design, reporting and application of future controlled trials in medical education by illustrating the need to carefully profile student participants in research studies who persist in or leave the studies over time and in doing so their representativeness of the populations from which they are drawn. Figure 1 illustrates this concept.

Methods

In this paper we refer to WISE-MD (a collection of online modules) and WISE Trial (a large multicenter field experiment using WISE-MD modules). The Web Initiative in Surgical Education for Medical Doctors (WISE-MD) has been creating case-based computer-assisted learning materials for surgical clerks since 2003 (Kalet et al. 2007) and had, as of February 2012, produced computer-assisted instructional (CAI) modules on 18 core clerkship topics. WISE Trial, a multi-centre randomized trial was started in 2009 to investigate the educational impact of four different evidence supported CAI designs. The Institutional Review Boards of all seven US medical schools participating in WISE Trial reviewed and approved the study.

All participating institutions ran a rotating clerkship model, which meant that different participants undertook their surgery clerkships at different times during the academic year. Six of the schools had an 8-week long clerkship while one school had a 12-week clerkship. All third-year medical students at the seven participating schools (N = 1,363) rotating

through their required surgery clerkships between July 2009 and August 2010 were eligible and invited to participate in the study.

Participants at each study site received an email on the first day of their surgery clerkship from the WISE Trial team and the local clerkship director inviting them to participate in the study and complete an online consent form. The clerkship coordinators or directors at each school also gave a brief verbal introduction to the trial during orientation. Local administrators and faculty were not aware of student consent status. The two study modules were mandatory at all sites though they made up only a small percentage of the final grade (<5 %). No incentives were offered for participation in the study.

Once consented into the main study, participants were randomized by computer algorithm to one of 4 study arms for the original trial. Each participant was asked to complete two online modules and three assessments based on the content of these modules. Each module, without assessments, represented approximately an hour of study time. Students were free to decide when and how quickly they completed the modules. Log data on participants' use of the WISE Trial modules and their completion of the study assessments was recorded by the online WISE-MD system.

Measures

Subject baseline characteristics

Table 1 lists the metacognitive and motivational measures that were administered to each participant immediately after having consented to join the trial. These measures were selected as representing important modifiers of academic achievement in CAI and associated with “persistence” in learning (Pintrich and de Groot 1990; Vandewalle 1997; Bong 2001; Song et al. 2010). Each was derived from an evidence-based measure, tailored for use with clinical clerkship students, and piloted and revised extensively before being fielded.

Scores on the US National Board of Medical Examiners (NBME) surgery—subject examination (mandated for all students at the seven participating institutions) were used as a measure of overall achievement for the surgery clerkship. The time of year that participants took their surgery clerkships was also tracked as educational outcomes tend to improve in general as the clerkship year proceeds (Kies et al. 2010).

Attrition study measures

The main outcome measure for this attrition study was the degree to which the medical students continued to participate in the WISE-Trial. Full participation in the study required them to complete the following five study procedures:

1. Consent and baseline measures (during the first week of the rotation)
2. First module including a 12-item domain specific declarative and conceptual knowledge post-test (timing at students' discretion)
3. Second module including a 12-item domain specific declarative and conceptual knowledge post-test (timing at students' discretion)
4. Mid-clerkship review made up of 27-item knowledge test clinical reasoning test on both topics (available at the mid-point of the rotation)
5. End of clerkship review made up of a 15-item, knowledge and clinical reasoning test on the content of both modules (last week of their clerkship rotation).

Table 1 Description of instruments used to measure participant characteristics

	Description	Example element	Maximum score	Minimum score	Minimal significant increment	Reference
Self-regulation score—usage	Self-generated thoughts, feelings, and actions for attaining academic goals—usage measure counts how many different types of strategy are used	“I re-read online materials until I understand” (reviewing records)	9	0	1	Song et al. (2010)
Self-regulation score—frequency	Self-generated thoughts, feelings, and actions for attaining academic goals—frequency measure counts how often self-regulation strategies are used	List your study methods (as many as you can)	∞	0	1	Song et al. (2010)
Mastery goal orientation	The individual preference for an absolute or intrapersonal standard in learning or achievement	“I enjoy challenging and difficult tasks”	35	5	1	Vandewalle (1997)
Performance approach—goal orientation	The preference for setting individual goals based on a positive normative standard in learning or achievement	“I enjoy it when others are aware of how well I am doing”	28	4	1	Vandewalle (1997)
Performance avoidance—goal orientation	The preference for setting individual goals based on a negative normative standard in learning or achievement	“I prefer to avoid situations where I might perform poorly”	28	4	1	Vandewalle (1997)
Task values	A person’s subjective values of a learning task or activity	“I think what I learn about surgery is important”	21	3	1	Bong (2001)
Self-efficacy	The judgment of personal capability in a specific domain	“I’m confident I can learn the basic concepts related to surgery”	35	5	1	Pintrich and de Groot (1990)
NBME subject exam score	Objectively scored examination of core content in surgery including organ systems and clinical reasoning	Multiple choice questions	Percentile			NBME website

Completing both modules and their associated measures took students on average a total of 175 min over the course of the 8–12 week clerkship.

Data analysis

Survival analysis techniques were used to model participant persistence in the study as it varied with categorical predictor variables (institution, study group assignment, rotation order) that were unchanged over the entire clerkship (Statacorp 2009). The “survival” event was defined as the study milestone after which a participant dropped out of the study. Participation was not always sequential (see Web Appendix); where it was not, the participation was classified based on the number of events that they completed. The log-rank test was used to assess for statistical significance. Since schools differed widely on how much of their final clerkship grades were based on NBME Subject examination scores (0–60 %), we dichotomized the data (cut-point: Shelf Exam \geq 40 % of grade) to examine the association between the influence of the Subject exam proportion of final grade and study persistence using survival analysis (log rank test). The other institution-level fixed variable we assessed was whether the medical school was situated in an urban or suburban/rural area.

Ordinal logistic regression was used to assess the degree to which continuous predictors were associated with attrition. Participation was classified according to a 9-point ordinal scale (1 = consented; then 2 points each for participating in any of the remaining four study milestones; the details of this are set out in Table 2) to assess potential associations of participation with participant characteristics that are measured as continuous variables. Using this ordinal scale univariate and multivariate ordinal logistic regression analyses were carried out where the participation scale was the dependent variable and the participant factors (self-regulation, self-efficacy, perceived task value, mastery goal orientation, and NBME subject examination score) were the predictor variables. Ordinal logistic regression makes no assumptions as to the distance between cut points on an ordinal scale, such as was used in this case. In addition, it does not require a linear relationship between dependent and independent variables and the dependent variable need not be normally distributed (Anderson and Philips 1981). The proportional odds assumption was verified using the likelihood ratio test.

For those predictors that were associated with study persistence on univariate regression at the $p < 0.10$ level, multivariate ordinal logistic regression was performed to determine which participant factors remained predictive of study persistence, after adjustment for the other variables.

Results

Of the 1,363 eligible students, 995 entered the study representing an overall response rate of 73 % (ranging from 44 to 81 % across the seven schools). There was a significant drop-off at each study milestone with 902 (66 %) completing at least one module, 799 (59 %) completing both modules, 539 (40 %) completing the mid-rotation evaluation and 252 (18 %) completing the final evaluation. Overall participation rates by school are represented in Fig. 2.

Attrition from the study did not differ among the four study arms of the RCT ($p = 0.59$) nor by rotation order ($p = 0.12$).

Table 2 Factors Associated with Higher Participation

	Consented to participate but NO modules completed (1 point)			Completed ONE module but no other measures (3 points)			Completed TWO modules but not the final two measures (5 points)			Completed TWO modules and mid-rotation measure but not final measure (7 points)			Completed ALL measures (9 points)			Statistical test*
	N	M	SD	N	M	SD	N	M	SD	N	M	SD	N	M	SD	
Self-regulation score—usage	48	3.5	1.5	67	3.8	1.6	183	3.5	1.6	209	3.7	1.5	137	4.1	1.6	$p < 0.01^{**}$
Self-regulation score—frequency	48	4.7	2.1	67	5	2.3	183	4.7	2.4	209	5	2.4	137	5.7	2.5	$p < 0.01$
Mastery goal orientation	67	29.3	3.9	86	28.9	4.1	272	28.9	4.4	301	29.1	4.4	203	29.7	4.2	$p = NS$
Performance approach	67	19.6	4.1	89	18.7	4.9	274	18.9	5	309	19.1	4.7	204	19.3	4.9	$p = NS$
Performance avoidance	66	14.2	5.7	87	13.8	5.1	279	14.3	5.2	316	13.7	5.2	205	13.8	5.3	$p = NS$
Task values	68	18.5	2.9	89	17.8	3.1	281	18.4	2.9	314	18.5	2.9	207	18.9	2.6	$p = 0.02^{**}$
Subject examination scores	20	72.1	6.8	36	73.2	8.7	170	75.2	9.4	230	76.8	8.9	152	79.6	10.3	$p < 0.001^{**}$

* p value for ordered logistic regression coefficient on univariate analysis (see “Methods”)

** Variable was not significant on multivariate analysis

Participation Rates

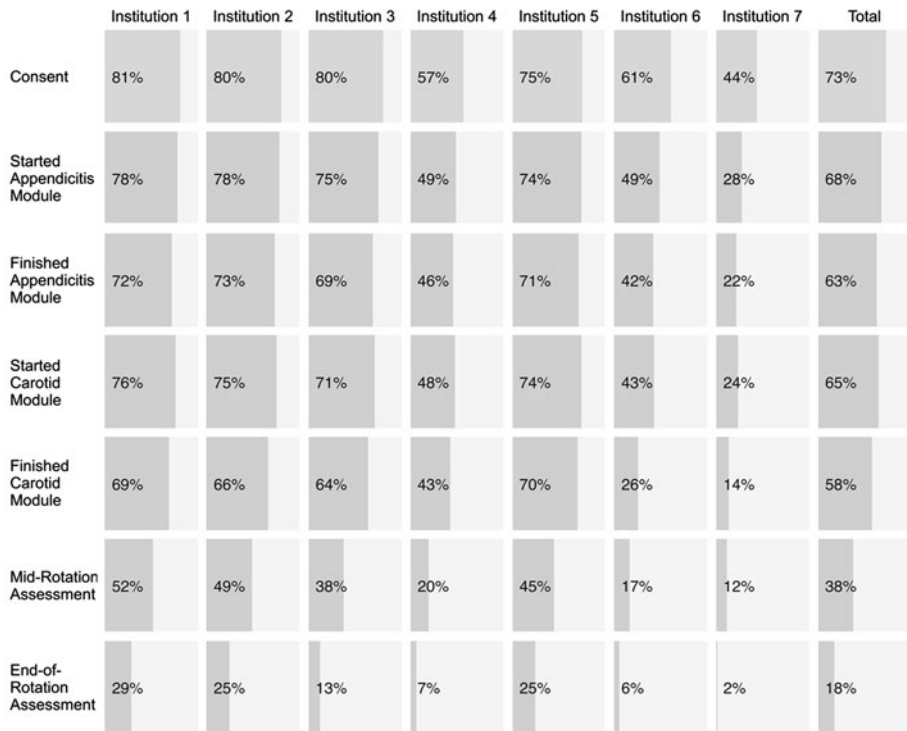


Fig. 2 Voluntary participation rates at the seven institutions. The width of the bar is proportional to the percentage of eligible participants at the given institution. Absolute numbers (N) are not included to maintain institution confidentiality. Consent was carried out within 2 days of the beginning of the surgery clerkship rotation. The Mid-Review (3 weeks) and End-Review (6 weeks) were done at specific time points. It was up to the participants as to when the two tutorials (Appendicitis and Carotid Stenosis) were completed. Institutions differed significantly in their participation rates (Log-Rank survival analysis $p < 0.001$)

Study persistence differed between schools on survival analysis ($p < 0.001$), though most of this difference was due to two schools with considerably higher attrition rates than the other five (See Fig. 2). The NBME Subject exam was worth 20–30 % of the final clerkship grade in four of the programs while it was worth 45–60 % of the grade in the two with higher attrition rates, suggesting that the greater the portion of the overall clerkship grade attributed to by the NBME subject examination, the less likely participants were to persevere ($p < 0.001$). No difference was found between urban and non-urban schools.

While lower participant baseline self-regulation and self-efficacy and final NBME Subject Exam scores were correlated with attrition from the study, goal orientation measures (Mastery, Performance Approach and Performance Avoidance) did not correlate with attrition (see Table 2).

Participants' endorsement of the task value of learning surgery had a small but statistically significant association with study attrition ($p = 0.02$) (Table 2), which was no longer evident on multivariate analysis. We used the variables shown to be statistically significantly ($p < 0.10$) associated with study attrition on univariate analysis (self-regulation frequency, Task Value, Self-Efficacy, NBME Subject Exam score and Subject exam

proportion) as predictor variables in an ordered logistic regression model to predict study attrition. Self-regulation frequency, Task Value, Self-Efficacy, and NBME Subject Exam score remained significant, independently explaining some of the variation in attrition.

In the study procedure, there were both scheduled events (date of consent and baseline measures, date of mid-clerkship evaluation and end-clerkship evaluation) and variable events (when the participants completed the modules). The stage in the clerkship when participants elected to start the modules was examined with respect to the scheduled milestone events. (See Web Appendix: Table 3) Participants who started their modules early on in the clerkship had higher self-regulation and higher NBME Subject Exam scores but there was no discernible correlation between early and late module completion and self-efficacy, task values, mastery goal orientation or post module test scores.

Discussion

While a large number of students were eligible to participate in this multi-center randomized controlled trial and there was a good initial consent rate, there was steady attrition from the study, with only 18 % of the participants completing all study components during their 8–12 week surgery clerkships. While this presents a major challenge for the main study, the variability and the size of the study as a whole presented an excellent opportunity for us to examine factors predicting study attrition. Factors associated with lower participant attrition included both learner's metacognitive characteristics and institutional factors (i.e. the emphasis placed on standardized tests for grading). In addition, learners who demonstrated greater persistence with the study performed better on the NBME subject examination, a measure of general academic achievement on the clerkship.

Field medical education trials, such as this one, are intrinsically bound to the contexts in which they take place (Hawe et al. 2004). Ideally these studies afford better measures of retention of learning over time than studies of educational interventions that measure outcomes immediately post instruction, designs which (Cook et al. 2008) trade off high compliance with the outcome assessment against the risk of over-estimating the longer-term impact of the intervention. There is similar tension in other research fields. In health behavior research, for instance, where interventions are complex and, in order to be ecologically valid, must be “hosted” in real world environments, maintaining study participation is challenging, especially for those individuals who might benefit most from the intervention (Amico 2009; Driscoll et al. 2009).

Multiple sampling clearly increases opportunities for participant attrition, however as long as persistence is demonstrably independent of group assignment, then causal inferences about the effectiveness and efficacy of the intervention (internal validity) are not necessarily compromised. However, with high attrition, generalizing the results to all students (external validity) is problematic (Cook and Campbell 1979). Characterizing what types of students do and do not persist in studies allows us to better appraise the potential impact of attrition of students from an educational trial.

Once a student agrees to participate in educational research, we can show that compliance with the study measures can correlate with higher levels of learner metacognitive skills and general achievement just as compliance with drug study procedures has been shown to be an independent marker for better health outcomes, independent of a drug's effectiveness (Petitti 1994; DiMatteo et al. 2002). It is interesting to speculate why we found an association between attrition from study participation and greater emphasis on the NBME subject examination as a component of students' clerkship grades. Educational

institutions where test scores are emphasized may lead students to avoid spending time on activities that divert study time from exam preparation.

Rates of attrition may also correlate with how much control over recruitment and retention there is within a study. If a trial requires N subjects to achieve sufficient power, the investigator often faces a dilemma: whether to try to achieve the required N at their home institution, where they presumably have maximum control, or to collect the data across multiple institutions to normalize the impact of local biases and confounders. Given the lower levels of investigator control at the remote sites, attrition rates were expected to be higher, something our study confirmed. By comparison, the institution where WISE-MD originated (and where the study team was based) had one of the highest recruitment and retention rates. The amount of control and authority associated with an intervention (or the lack thereof) can alter the behaviours of those who engage with the intervention. For example, in situations where the education intervention and its assessment is high-stakes or mandatory, recruitment and retention rates can approach 100 %. Of course this further limits the applicability to the low-stakes contexts (that are a more common locus for using CAI).

There were several limitations to this study. Because this dataset was from a study designed to assess educational effectiveness of CAI, and not attrition from educational research, we measured only those student characteristics known to impact learning therefore missing the opportunity to fully understand attrition by measuring relevant characteristics. (e.g. social responsibility, altruism, attitudes toward being subjects of research). When asked directly about this in focus groups, student participants at each study site consistently identified the main cause of attrition as competing time demands. Clearly our findings raise questions that will need to be explored in future research.

The anticipated benefits of increased ecological validity of multi-center studies needs to be tempered by the realities of increased study attrition and the attendant biases that have been demonstrated in this study. In the end, tolerating a certain amount of attrition may be an acceptable price to pay for the other benefits of multi-center enrollment, such as ecological validity and the development of learning interventions that can scale across institutions. The findings of medical education studies that are subject to volunteer bias should be interpreted in light of our findings that the participants tend to be somewhat more self-efficacious, self-regulating and higher achieving but that institutional factors can be even more important.

Conclusions

The significant student attrition rates in this study provided an opportunity to assess participant and to a lesser extent institutional characteristics associated with persistence. Learners who elected to persist with their participation in the study displayed different behavioural and cognitive characteristics from those that, having joined the study, did not persist with it. Our findings provide guidance for medical education researchers designing large-scale distributed studies and generalizing from their findings.

Acknowledgments This study was funded by an investigator-initiated grant from the National Library of Medicine (Kalet, 1LM009538-R01). The authors wish to acknowledge the following in their contribution to this study: Thomas S. Riles, M.D., Barry Mann, M.D., Ajit Sachdeva, M.D., Patrice G. Blair, M.P.H., Debra DaRosa, Ph.D., Susan Kaiser, M.D., John M. Stewart, Marc Triola, M.D., Thomas Lynch, M.D., Adnan Alseidi, M.D., James Hebert, M.D., D. Scott Lind, M.D., Kathleen Liscum, M.D., Fred Luchette, M.D., Kimberly Schenarts, Ph.D., Paula Termuhlen, M.D., Roy Phitayakorn, M.D., Kevin Foster, M.D., Lisa

Goodrich, Heather Keane, Brenda Barr, Bert Ongkeo, Grace Piaia, Gainosuke Sugiyama, M.D., Karen Taylor, Beverly Vaughn, Sherry Weitz, Kevin Wolfe, Ph.D., and Mari Suzuki.

Appendix: The WISE Trial Study Collaborative

M. A. Hopkins, New York University, maryann.hopkins@nyumc.org; D. Risucci, New York Medical College, donald_risucci@nymc.edu; R. Nesbit, Medical College of Georgia, rnesbit@georgiahealth.edu; W. Grant, University of Nebraska Medical Center, wgrant@unmc.edu; L. Dresner, State University of New York at Downstate, lisa.dresner@downstate.edu; T. James, University of Vermont, ted.james@vtmednet.org; M. Hochberg, New York University, mark.hochberg@nyumc.org; G. Isenberg, Jefferson Medical College, gerald.isenberg@jefferson.edu

References

- Amico, K. R. (2009). Percent total attrition: A poor metric for study rigor in hosted intervention designs. *American Journal of Public Health, 99*, 1567–1575.
- Anderson, J. A., & Philips, P. R. (1981). Regression, discrimination and measurement models for ordered categorical variables. *Journal of the Royal Statistical Society, 30*(1), 22–31.
- Bleakley, A., Bligh, J., & Browne, J. (2011). *Lack, trajectories and ruptures in medical education research, medical education for the future* (pp. 215–225). The Netherlands: Springer.
- Bong, M. (2001). Between- and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task value, and achievement goals. *Journal of Educational Psychology, 93*(1), 23–34.
- Callahan, C. A., Hojat, M., & Gonnella, J. S. (2007). Volunteer bias in medical education research: An empirical study of over three decades of longitudinal data. *Medical Education, 41*, 746–753. doi: [10.1111/j.1365-2923.2007.02803.x](https://doi.org/10.1111/j.1365-2923.2007.02803.x).
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Cook, D. A., Levinson, A. J., Garside, S., Dupras, D. M., Erwin, P. J., & Montori, V. M. (2008). Internet-based learning in the health professions: A meta-analysis. *The Journal Of The American Medical Association, 300*(10), 1181–1196.
- DiMatteo, R. M., Giordani, P. J., Lepper, H. S., & Croghan, T. W. (2002). Patient adherence and medical treatment outcomes: A meta-analysis. *Medical Care, 40*(9), 794.
- Driscoll, K. A., Killian, M., Johnson, S. B., Silverstein, J. H., & Deeb, L. (2009). Predictors of study completion and withdrawal in a randomized clinical trial of a pediatric diabetes adherence intervention. *Contemporary clinical trials, 30*(3), 212–220.
- Hawe, P., Shiell, A., & Riley, T. (2004). Complex interventions: How “out of control” can a randomised controlled trial be? *BMJ, 328*(7455), 1561–1563.
- Jüni, P., Altman, D. G., & Egger, M. (2001). Assessing the quality of controlled clinical trials. *BMJ, 323*(7303), 42–46.
- Kalet, A. L., Coady, S. H., Hopkins, M. A., Hochberg, M. S., & Riles, T. S. (2007). Preliminary evaluation of the web initiative for surgical education (WISE-MD). *American Journal of Surgery, 194*(1), 89–93.
- Kies, S. M., Roth, V., & Rowland, M. (2010). Association of third-year medical students’ first clerkship with overall clerkship performance and examination scores. *JAMA, 304*(11), 1220–1226.
- Petitti, D. B. (1994). Coronary heart disease and estrogen replacement therapy can compliance bias explain the results of observational studies? *Annals of Epidemiology, 4*(2), 115–118.
- Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33–40.
- Shrank, W. H., Patrick, A. R., & Brookhart, M. A. (2011). Healthy user and related biases in observational studies of preventive interventions: A primer for physicians. *Journal of General Internal Medicine, 26*(5), 546–550.
- Song, H. S., Kalet, A. L., & Plass, J. L. (2010). Assessing medical students’ self-regulation as aptitude in computer-based learning. *Advances in Health Sciences Education, 16*(1), 97–107.
- StataCorp. (2009). *Stata survival analysis and epidemiological tables reference manual* (11th ed.). College Station, Texas: Stata Press.

- Todres, M., Stephenson, A., & Jones, R. (2007). Medical education research remains the poor relation. *BMJ*, *335*, 333–335.
- Torgerson, C. T. (2002). Educational research and randomised trials. *Medical Education*, *36*, 1002–1003.
- Vandewalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and Psychological Measurement*, *57*(6), 995–1015.