



Online article and related content
current as of June 3, 2009.

Internet-Based Learning in the Health Professions: A Meta-analysis

David A. Cook; Anthony J. Levinson; Sarah Garside; et al.

JAMA. 2008;300(10):1181-1196 (doi:10.1001/jama.300.10.1181)

<http://jama.ama-assn.org/cgi/content/full/300/10/1181>

Supplementary material	eTables http://jama.ama-assn.org/cgi/content/full/300/10/1181/DC1
Correction	Contact me if this article is corrected.
Citations	This article has been cited 10 times. Contact me when this article is cited.
Topic collections	Informatics/ Internet in Medicine; Internet; Medical Practice; Medical Education; Quality of Care; Evidence-Based Medicine; Review Contact me when new articles are published in these topic areas.
CME course	Online CME course available.
CME course	Online CME course available.
Related Letters	Internet-Based Education for Health Professionals . <i>JAMA</i> . ;301():598. et al. <i>JAMA</i> . ;301():599.

Subscribe
<http://jama.com/subscribe>

Permissions
permissions@ama-assn.org
<http://pubs.ama-assn.org/misc/permissions.dtl>

Email Alerts
<http://jamaarchives.com/alerts>

Reprints/E-prints
reprints@ama-assn.org

Internet-Based Learning in the Health Professions

A Meta-analysis

David A. Cook, MD, MHPE

Anthony J. Levinson, MD, MSc

Sarah Garside, MD, PhD

Denise M. Dupras, MD, PhD

Patricia J. Erwin, MLS

Victor M. Montori, MD, MSc

THE ADVENT OF THE WORLD Wide Web in 1991 greatly facilitated the use of the Internet¹ and its potential as an instructional tool was quickly recognized.^{2,3} Internet-based education permits learners to participate at a time and place convenient to them, facilitates instructional methods that might be difficult in other formats, and has the potential to tailor instruction to individual learners' needs.⁴⁻⁶ As a result, Internet-based learning has become an increasingly popular approach to medical education.^{7,8}

However, concerns about the effectiveness of Internet-based learning have stimulated a growing body of research. In the first decade of the Web's existence 35 evaluative articles on Web-based learning were published,⁹ whereas at least 32 were published in 2005 alone.¹⁰ Synthesis of this evidence could inform educators and learners about the extent to which these products are effective and what makes them more or less effective.⁶



CME available online at
www.jamaarchivescme.com
and questions on p 1245.

Context The increasing use of Internet-based learning in health professions education may be informed by a timely, comprehensive synthesis of evidence of effectiveness.

Objectives To summarize the effect of Internet-based instruction for health professions learners compared with no intervention and with non-Internet interventions.

Data Sources Systematic search of MEDLINE, Scopus, CINAHL, EMBASE, ERIC, TimeLit, Web of Science, Dissertation Abstracts, and the University of Toronto Research and Development Resource Base from 1990 through 2007.

Study Selection Studies in any language quantifying the association of Internet-based instruction and educational outcomes for practicing and student physicians, nurses, pharmacists, dentists, and other health care professionals compared with a no-intervention or non-Internet control group or a preintervention assessment.

Data Extraction Two reviewers independently evaluated study quality and abstracted information including characteristics of learners, learning setting, and intervention (including level of interactivity, practice exercises, online discussion, and duration).

Data Synthesis There were 201 eligible studies. Heterogeneity in results across studies was large ($I^2 \geq 79\%$) in all analyses. Effect sizes were pooled using a random effects model. The pooled effect size in comparison to no intervention favored Internet-based interventions and was 1.00 (95% confidence interval [CI], 0.90-1.10; $P < .001$; $n=126$ studies) for knowledge outcomes, 0.85 (95% CI, 0.49-1.20; $P < .001$; $n=16$) for skills, and 0.82 (95% CI, 0.63-1.02; $P < .001$; $n=32$) for learner behaviors and patient effects. Compared with non-Internet formats, the pooled effect sizes (positive numbers favoring Internet) were 0.10 (95% CI, -0.12 to 0.32; $P = .37$; $n=43$) for satisfaction, 0.12 (95% CI, 0.003 to 0.24; $P = .045$; $n=63$) for knowledge, 0.09 (95% CI, -0.26 to 0.44; $P = .61$; $n=12$) for skills, and 0.51 (95% CI, -0.24 to 1.25; $P = .18$; $n=6$) for behaviors or patient effects. No important treatment-subgroup interactions were identified.

Conclusions Internet-based learning is associated with large positive effects compared with no intervention. In contrast, effects compared with non-Internet instructional methods are heterogeneous and generally small, suggesting effectiveness similar to traditional methods. Future research should directly compare different Internet-based interventions.

JAMA. 2008;300(10):1181-1196

www.jama.com

Since 2001, several reviews (some of which also included non-Internet-based computer-assisted instruction) have offered such summaries.⁹⁻¹⁷ However, each had important methodological limitations, including incomplete accounting of existing studies, limited

Author Affiliations: College of Medicine (Drs Cook, Dupras, and Montori and Ms Erwin), Office of Education Research (Dr Cook), and Knowledge and Encounter Research Unit (Dr Montori), Mayo Clinic, Rochester, Minnesota, and McMaster University, Hamilton, Ontario (Drs Levinson and Garside).

Corresponding Author: David A. Cook, MD, MHPE, Division of General Internal Medicine, Mayo Clinic College of Medicine, Baldwin 4-A, 200 First St SW, Rochester, MN 55905 (cook.david33@mayo.edu).

assessment of study quality, and no quantitative pooling to derive best estimates of these interventions' effect on educational outcomes.

We sought to identify and quantitatively summarize all studies of Internet-based instruction involving health professions learners. We conducted 2 systematic reviews with meta-analyses addressing this topic, the first exploring Internet-based instruction compared with no intervention and the second summarizing studies comparing Internet-based and non-Internet instructional methods (media-comparative studies).

METHODS

These reviews were planned, conducted, and reported in adherence to standards of quality for reporting meta-analyses (Quality of Reporting of Meta-analyses and Meta-analysis of Observational Studies in Epidemiology standards).^{18,19}

Questions

We sought to answer (1) to what extent is Internet-based instruction associated with improved outcomes in health professions learners compared with no intervention, and (2) how does Internet-based instruction compare with non-Internet instructional methods? We also sought to determine factors that could explain differences in effect across participants, settings, interventions, outcomes, and study designs for each of these questions.

Based on existing theories and evidence,²⁰⁻²⁴ we hypothesized that cognitive interactivity, peer discussion, ongoing access to instructional materials, and practice exercises would improve learning outcomes. We also anticipated, based on evidence²⁵ and argument,²⁶ that Internet-based instruction in comparison to no intervention would have the greatest effect on knowledge, a smaller but significant effect on skills, and a yet smaller effect on behaviors in practice and patient-related outcomes. Finally, based on previous reviews and discussions,²⁷⁻²⁹ we expected no overall difference be-

tween Internet and non-Internet instructional modalities, provided instructional methods were similar between interventions.

Study Eligibility

We developed intentionally broad inclusion criteria in order to present a comprehensive overview of Internet-based learning in health professions education. We included studies in any language if they reported evaluation of the Internet to teach health professions learners at any stage in training or practice compared with no intervention (ie, a control group or preintervention assessment) or a non-Internet intervention, using any of the following outcomes³⁰: reaction or satisfaction (learner satisfaction with the course), learning (knowledge, attitudes, or skills in a test setting), behaviors (in practice), or effects on patients (BOX). We included single-group pretest-posttest, 2-group randomized and nonrandomized, parallel-group and crossover designs, and studies of "adjuvant" instruction, in which an Internet-based intervention is added to other instruction common to all learners.

Studies were excluded if they reported no outcomes of interest, did not compare Internet-based instruction with no intervention or a non-Internet intervention, used a single-group posttest-only design, or evaluated a computer intervention that resided only on the client computer or CD-ROM or in which the use of the Internet was limited to administrative or secretarial purposes. Meeting abstracts were also excluded.

Study Identification

A senior reference librarian with expertise in systematic reviews (P.J.E.) designed a strategy to search MEDLINE, Scopus, CINAHL, EMBASE, ERIC, TimeLit, Web of Science, Dissertation Abstracts, and the University of Toronto Research and Development Resource Base for relevant articles. Search terms included delivery concepts (such as *Internet*, *Web*, *computer-assisted instruction*, *e-learning*, *online*, *virtual*, and *distance*), study design concepts (such as

comparative study, *evaluative study*, *pretest*, or *program evaluation*), and participant characteristics (such as *education*, *professional*; *students*, *health occupations*; *internship and residency*; and *specialties, medical*). eTable 1 (<http://www.jama.com>) describes the complete search strategy. We restricted our search to articles published in or after 1990 because the World Wide Web was first described in 1991. The last date of search was January 17, 2008. Additional articles were identified by hand-searching reference lists of all included articles, previous reviews, and authors' files.

Study Selection

Working independently and in duplicate, reviewers (D.A.C., A.J.L., S.G., and D.M.D.) screened all titles and abstracts, retrieving in full text all potentially eligible abstracts, abstracts in which reviewers disagreed, or abstracts with insufficient information. Again independently and in duplicate, reviewers considered the eligibility of studies in full text, with adequate chance-adjusted interrater agreement (0.71 by intraclass correlation coefficient³¹ [ICC], estimated using SAS 9.1 [SAS Institute Inc, Cary, North Carolina]). Reviewers resolved conflicts by consensus.

Data Extraction

Reviewers abstracted data from each eligible study using a standardized data abstraction form that we developed, iteratively refined, and implemented electronically. Data for all variables where reviewer judgment was required (including quality criteria and all characteristics used in meta-analytic subgroup analyses) were abstracted independently and in duplicate, and interrater reliability was determined using ICC. Conflicts were resolved by consensus. When more than 1 comparison intervention was reported (eg, both lecture and paper interventions), we evaluated the comparison most closely resembling the Internet-based course (ICC, 0.77).

We abstracted information on the number and training level of learners,

Box. Definitions of Study Variables**Participants****Health professions learners**

Students, postgraduate trainees, or practitioners in a profession directly related to human or animal health; for example physicians, nurses, pharmacists, dentists, veterinarians, and physical and occupational therapists.

Interventions**Internet-based instruction**

Computer-assisted instruction—instruction in which “computers play a central role as the means of information delivery and direct interaction with the learner (in contrast to the use of computer applications such as PowerPoint), and to some extent replace the human instructor.”⁶—using the Internet or a local intranet as the means of delivery. This included Web-based tutorials, virtual patients, discussion boards, e-mail, and Internet-mediated videoconferencing. Applications linked to a specific computer (including CD-ROM) were excluded unless they also used the Internet.

Learning environment (classroom vs practice setting)

Classroom-type settings were those in which most learners would have attended had the course not used the Internet (ie, the Internet-based course replaced a classroom course or supplemented a classroom course, or other concurrent courses were in a classroom). Practice-type settings were those in which learners were seeing patients or had a primary patient care responsibility (ie, students in clinical years, postgraduate trainees, or on-the-job training).

Practice exercises

Practice exercises included cases, self-assessment questions, and other activities requiring learners to apply information they had learned.

Cognitive interactivity

Cognitive interactivity rated the level of cognitive engagement required for course participation. Multiple practice exercises typically justified moderate or high interactivity, although exercises for which questions and answers were provided together (ie, on the same page) were rated low. Essays and group collaborative projects also supported higher levels of cognitive interactivity.

Discussion

Face-to-face discussion required dedicated time for instructor-student or peer-peer interaction, above and beyond the questions that might arise in a typical lecture. Online discussion required provision for such interactions using synchronous or asynchronous online communication such as discussion board, e-mail, chat, or Internet conferencing.

Tutorial

Tutorials were the online equivalent of a lecture and typically involved learners studying and completing assignments alone. These often comprised stand-alone Internet-based applications with varying degrees of interactivity and multimedia.

Synchronous or asynchronous communication

Synchronous communication involved simultaneous interaction between 2 or more course participants over the Internet, using methods such as online chat, instant messaging, or 2-way videoconferencing.

Internet conferencing

Internet conferencing involved the simultaneous transmission of both audio and video information. Video information could comprise an image of the instructor, other video media, or shared projection of the computer screen (ie, whiteboard).

Repetition (single-instance vs ongoing access)

Repetition evaluated the availability of interventions over time; coded as single instance (learning materials available only once during the course) and ongoing access (learning materials accessible throughout the duration of the course).

Duration

The time over which learners participated in the intervention.

Outcomes**Satisfaction (reaction)**

Learners' reported satisfaction with the course.

Knowledge

Subjective (eg, learner self-report) or objective (eg, multiple-choice question knowledge test) assessments of factual or conceptual understanding.

Skills

Subjective (eg, learner self-report) or objective (eg, faculty ratings, or objective tests of clinical skills such as interpretation of electrocardiograms or radiographs) assessments of learners' ability to demonstrate a procedure or technique.

Behaviors and patient effects

Subjective (eg, learner self-report) or objective (eg, chart audit) assessments of behaviors in practice (such as test ordering) or effects on patients (such as medical errors).

learning setting (classroom vs practice setting; ICC, 0.81), study design (pretest-posttest vs posttest-only, number of groups, and method of group assignment; ICC range, 0.88-0.95), topic, instructional modalities used, length of course (ICC, 0.85), online tutorial (ICC, 0.68) or video-conference (ICC, 0.96) format, level of cognitive interactivity (ICC, 0.70), quantity of practice exercises (ICC, 0.70), repetition (ICC, 0.65), presence of online discussion (ICC, 0.85) and face-to-face discussion (ICC, 0.58), synchronous learning (ICC, 0.95), and each outcome (subjective or objective [ICC range, 0.63-1.0] and descriptive statistics). When outcomes data were missing, we requested this information from authors by e-mail and paper letter.

Recognizing that many nonrandomized and observational studies would be included, we abstracted information on methodological quality using an adaptation of the Newcastle-Ottawa scale for grading the quality of cohort studies.³² We rated each study in terms of representativeness of the intervention group (ICC, 0.63), selection of the control group (ICC, 0.75), comparability of cohorts (statistical adjustment for baseline characteristics in non-randomized studies [ICC, 0.49], or randomization [ICC, 0.93] and allocation concealment [ICC, 0.48] for randomized studies), blinding of outcome assessment (ICC \geq 0.74), and completeness of follow-up (ICC, 0.37 to 0.79 depending on outcome).

Data Synthesis

We analyzed studies separately for outcomes of satisfaction, knowledge, skills, and behaviors or patient effects. For each outcome class we converted means and standard deviations to standardized mean differences (Hedges *g* effect sizes).³³⁻³⁵ When insufficient data were available, we used reported tests of significance (eg, *P* values) to estimate the effect size. For crossover studies we used means or exact statistical test results adjusted for repeated measures or, if these were not reported, we used

means pooled across each intervention.^{36,37} For 2-group pretest-posttest studies we used posttest means or exact statistical test results adjusted for pretest or, if these were not reported, we used differences in change scores standardized using pretest variance. If neither *P* values nor any measure of variance was reported, we used the average standard deviation from all other included studies.

To quantify inconsistency (heterogeneity) across studies we used the I^2 statistic,³⁸ which estimates the percentage of variability across studies not due to chance. I^2 values greater than 50% indicate large inconsistency. Because we found large inconsistency ($I^2 \geq 79\%$ in all analyses), we used random-effects models to pool weighted effect sizes across studies using StatsDirect 2.6.6 (StatsDirect Ltd, Altrincham, England, <http://www.statsdirect.com>).

We performed subgroup analyses to explore heterogeneity and to investigate the questions noted above regarding differences in participants, interventions, design, and quality. We used a 2-sided α level of .05. We grouped studies with active comparison interventions according to relative between-intervention differences in instructional methods; namely, did the comparison intervention have more, less, or the same amount of interactivity, practice exercises, discussion (face-to-face and Internet-based discussion combined), and repetition.

We conducted sensitivity analyses to explore the robustness of findings to synthesis assumptions, with analyses excluding low-quality studies, studies with effect size estimated from inexact tests of significance or imputed standard deviations, 1 study³⁹ that contributed up to 14 distinct Internet-based interventions, studies of blended (Internet and non-Internet) interventions, and studies with major design flaws (described below).

RESULTS

Trial Flow

The search strategy identified 2045 citations, and an additional 148 poten-

tially relevant articles were identified from author files and review of reference lists. From these we identified 288 potentially eligible articles (FIGURE 1). Following a single qualitative study reported in 1994, the number of comparative or qualitative studies of Internet-based learning increased from 2 articles published in 1996, to 16 publications in 2001, to 56 publications in 2006. We contacted authors of 113 articles for additional outcomes information and received information from 45. Thirteen otherwise eligible articles contained insufficient data to calculate an effect size (ie, sample size or both means and statistical tests absent) and were excluded from the meta-analyses. Ultimately we analyzed 201 articles, 5 of which contributed to both analyses, representing 214 interventions. TABLE 1 summarizes key study features and eTable 2 (<http://www.jama.com>) provides detailed information.

Study Characteristics

Internet-based instruction addressed a wide range of medical topics. In addition to numerous diagnostic and therapeutic content areas, courses addressed topics such as ethics, histology, anatomy, evidence-based medicine, conduct of research, biostatistics, communication skills, interpretation of electrocardiograms and pulmonary function tests, and systems-based practice. Most interventions involved tutorials for self-study or virtual patients, while over a quarter required online discussion with peers, instructors, or both. These modalities were often mixed in the same course. Twenty-nine studies (14.4%) blended Internet-based and face-to-face instruction. Non-Internet comparison interventions most often involved face-to-face courses or paper modules but also included satellite-mediated videoconferences, standardized patients, and slide-tape self-study modules.

The vast majority of knowledge outcomes consisted of multiple-choice tests, a much smaller number comprised other objectively scored methods, and 18 of 177 studies assessing

knowledge (10.2%) used self-report measures of knowledge, confidence, or attitudes. Skills outcomes included communication with patients, critical appraisal, medication dosing, cardio-pulmonary resuscitation, and lumbar puncture. These were most often assessed using objective instructor or standardized patient observations. Skills outcomes were self-reported or the method could not be determined for 7 of 26 studies (26.9%). Behavior and patient effects included osteoporosis screening rates, cognitive behavioral therapy implementation, workplace violence events, incidence of postpartum depression, and various perceived changes in practice. Ten of 23 articles (43.5%; representing nearly two-thirds of the interventions) used self-reported behavior or patient effects outcomes. Most objective assessments used chart review, although 1 study used incognito standardized patients.

Study Quality

TABLE 2 summarizes the methodological quality of included studies, and eTable 3 (<http://www.jama.com>) contains details on the quality scale and individual study quality. Nine of 61 (14.8%) no-intervention 2-group comparison studies determined groups by completion or noncompletion of elective or “required” Internet-based instruction. Although such groupings are susceptible to bias, sensitivity analyses showed similar results when these studies were excluded. Eight of 43 studies (18.6%) assessing satisfaction, 42 of 177 (23.7%) assessing knowledge, 4 of 26 (15.4%) assessing skills, and 5 of 23 (21.7%) assessing behaviors and patient effects lost more than 25% of participants from time of enrollment or failed to report follow-up. The mean (SD) quality score (6 points indicating highest quality) was 2.5 (1.3) for no-intervention controlled studies, and 3.5 (1.4) for non-Internet comparison studies.

Quantitative Data Synthesis:

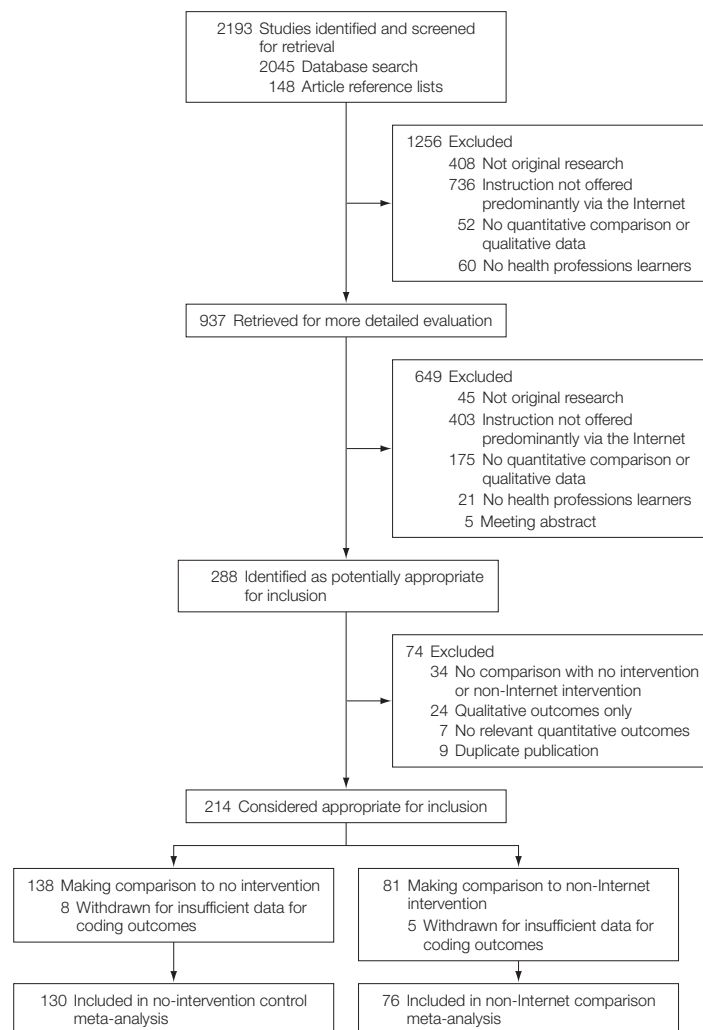
Comparisons With No Intervention

FIGURES 2, 3, 4, and eTable 4 (<http://www.jama.com>) summarize the re-

sults of the meta-analyses comparing Internet-based instruction with no intervention. Satisfaction outcomes are difficult to define in comparison to no intervention, and no studies reported meaningful outcomes of this type. We used inexact *P* values to estimate 17 of 174 effect sizes (9.8%), and we imputed standard deviations to estimate 10 effect sizes (5.7%). eTable 4 (<http://www.jama.com>) contains detailed results of the main analysis and sensitivity analyses for each outcome. Sensitivity analyses did not affect conclusions.

Knowledge. One hundred seventeen studies reported on 126 interventions using knowledge as the outcome. The pooled effect size for these interventions was 1.00 (95% confidence interval [CI], 0.90-1.10; $P < .001$). Because effect sizes larger than 0.8 are considered large,⁴⁰ this suggests that Internet-based instruction typically has a substantial benefit on learners' knowledge compared with no intervention. However, we also found large inconsistency across studies ($I^2=93.6\%$), and individual effect sizes ranged from -0.30 to 6.69. One of the

Figure 1. Trial Flow



Five studies compared the Internet-based intervention with both no intervention and a non-Internet comparison intervention.

2 interventions yielding a negative effect size⁴¹ was an adjunct to an existing intensive and well-planned course on lung cancer. The other⁴² compared Internet-based educational order sets for medical students on a surgery clerk-

ship to students at a different hospital without access to these order sets, which could arguably be construed as an active comparison intervention.

In subgroup analyses exploring this inconsistency, we failed to confirm our

hypotheses that high interactivity, ongoing access to course materials, online discussion, or the presence of practice exercises would yield larger effect sizes (P for interaction $\geq .15$) (Figure 2). However, we found a significant interaction with study quality, with studies scoring low on the modified Newcastle-Ottawa scale showing a greater effect than high-quality studies (mean score, 1.07; 95% CI, 0.96-1.18 vs mean score, 0.71; 95% CI, 0.51-0.92; P for interaction = .003).

Skills. Sixteen interventions used skills as an outcome. The pooled effect size of 0.85 (95% CI, 0.49-1.20; $P < .001$) reflects a large effect. There was large inconsistency across trials ($I^2=92.7\%$), and effect sizes ranged from 0.02 to 2.50.

The pooled effect size for interventions with practice exercises was significantly higher than those without (pooled effect size, 1.01; 95% CI, 0.60-1.43 vs pooled effect size, 0.21; 95% CI, 0.04-0.38; P for interaction $< .001$), but once again interactivity, repetition, and discussion did not affect outcomes (P for interaction $\geq .30$) (Figure 3).

Behaviors and Effects on Patient Care. Nineteen studies reported 32 interventions evaluating learner behaviors and effects on patient care. These studies demonstrated a large pooled effect size of 0.82 (95% CI, 0.63-1.02; $P < .001$) and large inconsistency ($I^2=79.1\%$). Effect sizes ranged from 0.06 to 7.26.

In contrast to skills outcomes, practice exercises were negatively associated with behavior outcomes (0.44; 95% CI, 0.33-0.55 if present; 2.09; 95% CI, 1.38-2.79 if absent; P for interaction $< .001$) (Figure 4). We also found statistically significant differences favoring tutorials, longer-duration courses, and online peer discussion.

Quantitative Data Synthesis: Comparisons With Non-Internet Interventions

FIGURES 5, 6, 7, and 8 and eTable 4 (<http://www.jama.com>) summarize the results of the meta-analyses comparing Internet-based instruction with

Table 1. Description of Included Studies^a

Study Characteristic	No Intervention Comparison		Non-Internet Comparison	
	No. (%) of Studies	No. of Participants ^b	No. (%) of Studies	No. of Participants ^b
All studies	130	19 234	76	7218
Study design				
Posttest-only 2-group	33 (25.4)	5565	47 (61.8)	4516
Pretest-posttest 2-group	28 (21.5)	4107	29 (38.2)	2702
Pretest-posttest 1-group	69 (53.1)	9562	0 (0)	0
Setting				
Classroom	38 (29.2)	5702	46 (60.5)	4166
Practice	90 (69.2)	13 414	29 (38.2)	3014
Undefined	2 (1.6)	118	1 (1.3)	38
Participants ^c				
Medical students	40 (30.8)	5851	20 (26.3)	2491
Physicians in postgraduate training	31 (23.9)	4376	5 (6.6)	413
Physicians in practice	27 (20.8)	4824	5 (6.6)	443
Nursing students	8 (6.2)	673	15 (19.7)	1312
Nurses in practice	20 (15.4)	967	8 (10.5)	612
Dental students	2 (1.6)	148	3 (4.0)	126
Dentists in practice	1 (0.8)	17	0 (0)	0
Pharmacy students	11 (8.5)	646	5 (6.6)	324
Pharmacists in practice	6 (4.6)	142	1 (1.3)	4
Other	21 (16.2)	2820	21 (27.6)	1736
Interventions ^d				
Interactivity high	79 (60.8)	12 541	50 (65.8)	4584
Practice exercises present	78 (60.0)	11 576	46 (60.5)	4313
Repetition ongoing access	72 (55.4)	8208	35 (46.0)	3364
Duration ≤ 1 wk	56 (43.1)	10 057	37 (52.9)	2909
Tutorial	112 (86.2)	15 957	63 (82.9)	6035
Discussion	28 (21.5)	3906	33 (43.4)	3314
Synchronous	5 (3.9)	84	14 (18.4)	1169
Comparison vs face-to-face	NA	NA	57 (75)	5723
Comparison vs paper	NA	NA	14 (18.4)	1303
Outcomes ^e				
Satisfaction	0	0	43 (56.6)	4370
Knowledge	117 (90.0)	18 053	63 (82.9)	5781
Skills	16 (12.3)	1708	12 (15.8)	1029
Behaviors and patient effects	19 (14.6)	2159	6 (7.9)	822
Quality ^e				
Newcastle-Ottawa ≥ 4 points	22 (16.9)	3343	38 (50.0)	3362

Abbreviation: NA, not applicable.

^aA total of 201 studies representing 214 interventions were included in the meta-analysis. This table presents data with studies as the unit of analysis. Five studies compared the Internet-based intervention with both no intervention and a non-Internet comparison intervention and are counted separately. See eTable 2 online (<http://www.jama.com>) for details on individual studies.

^bNumbers reflect the number of students enrolled. The number of participants for subgroups may total more than the number for all studies when characteristics are not mutually exclusive.

^cPercentages total more than 100% because several studies included more than 1 learner group or reported multiple outcomes.

^dInterventions refer to Internet-based intervention except when noted otherwise. Numbers total more than 100% because these characteristics are not mutually exclusive.

^eSee text and eTable 3 (<http://www.jama.com>) for details on the modified Newcastle-Ottawa scale.

Table 2. Quality of Included Studies^a

Study Characteristics	No. of Studies	Representativeness	Selection	Comparability ^b		Blinded Outcome ^c	Follow-up ^c
				1 Point	2 Point		
No-Intervention Controlled Studies							
All studies	130	45	45	20	11	98	97
Study design							
Posttest-only 2-group	33	19	21	9	3	24	29
Pretest-posttest 2-group	28	11	24	11	8	22	18
Pretest-posttest 1-group	69	15	0	0	0	52	50
Setting							
Classroom	38	19	12	4	2	25	30
Practice	90	25	31	15	9	71	65
Undefined	2	1	2	1	0	2	2
Participants ^d							
Medical students	40	22	14	6	2	32	33
Physicians	52	12	17	11	7	40	39
Nurses	27	5	8	2	2	19	15
Other	36	9	11	5	4	27	28
Interventions							
Interactivity high	79	32	35	18	5	56	58
Practice exercises present	78	32	34	14	8	59	61
Repetition ongoing access	72	29	32	15	8	49	50
Duration ≤1 wk	56	18	16	10	3	47	46
Tutorial	112	38	37	19	9	85	85
Discussion	28	6	12	4	2	17	22
Synchronous	5	1	1	1	0	2	5
Outcomes							
Knowledge	117	39	37	17	9	89	86
Skills	16	6	10	6	2	7	13
Behaviors and patient effects	19	6	10	6	2	7	15
Non-Internet Comparison Studies							
All studies	76	45	57	27	10	51	61
Study design							
Posttest-only 2-group	47	26	32	12	1	24	39
Pretest-posttest 2-group	29	19	25	15	9	27	22
Setting							
Classroom	46	29	35	19	1	27	37
Practice	29	16	22	8	9	24	24
Undefined	1	0	0	0	0	0	0
Participants ^d							
Medical students	20	12	15	8	2	16	17
Physicians	10	8	8	3	5	10	8
Nurses	22	11	14	6	0	13	18
Other	30	16	23	11	3	16	24
Interventions							
Interactivity high	50	25	37	21	5	37	41
Practice exercises present	46	23	33	20	6	35	38
Repetition ongoing access	35	21	25	10	4	22	27
Duration ≤1 wk	37	23	32	18	7	30	30
Tutorial	63	37	47	25	10	45	53
Discussion	33	19	22	7	5	18	26
Synchronous	14	6	9	2	3	9	11
Comparison vs face-to-face	57	36	40	16	6	36	45
Outcomes							
Satisfaction	43	26	33	12	8	0	35
Knowledge	63	39	46	24	9	48	51
Skills	12	6	9	6	1	9	11
Behaviors and patient effects	6	2	4	2	2	3	5

^aData presented as number of studies. Quality was assessed using a modification of the Newcastle-Ottawa scale³² that rated each study in terms of representativeness of the intervention group (1 point), selection of the control group (1 point), comparability of cohorts (2 points), blinding of assessment (1 point), and completeness of follow-up (1 point). See text and eTable 3 (<http://www.jama.com>) for details regarding this scale.

^bThe columns for comparability of cohorts are additive, eg, 31 no-intervention controlled studies had at least 1 point.

^cExcept for the Outcomes categories, blinding and completeness of follow-up were counted as present if this was done for any reported outcome. Responses for Outcomes categories are specific to that outcome.

^dNumber of studies do not appear to match those in Table 1 because several studies included learners at more than 1 level (eg, both medical students and physicians, or nurses in training and in practice).

non-Internet instruction. We used inexact *P* values to estimate 1 of 124 effect sizes (0.8%), and we imputed standard deviations to estimate 5 effect sizes (4.0%). Sensitivity analyses did not alter conclusions except as noted.

Satisfaction. Forty-three studies reported satisfaction outcomes comparing Internet-based instruction to non-Internet formats. The pooled effect size (positive numbers favoring Internet)

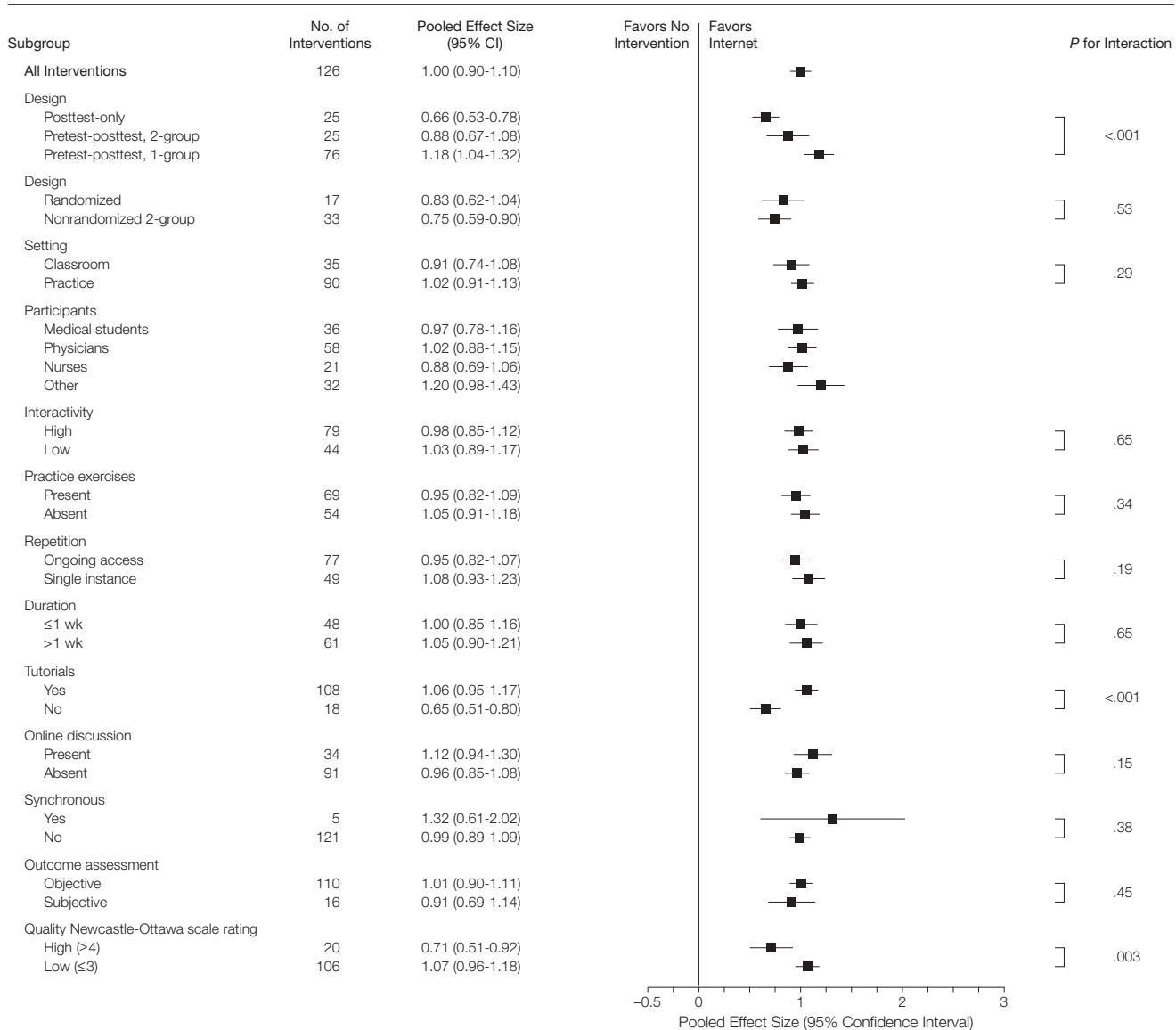
was 0.10 (95% CI, -0.12 to 0.32), with $I^2=92.2\%$. This effect is considered small⁴⁰ and was not significantly different from 0 ($P=.37$). Individual effect sizes ranged -1.90 to 1.77.

We had no a priori hypotheses regarding subgroup comparisons for satisfaction outcomes, but we found statistically significant treatment-subgroup interactions favoring short courses, high-quality studies, and

single-instance rather than ongoing-access Internet-based interventions (Figure 5).

Knowledge. Sixty-three non-Internet-controlled studies reported knowledge outcomes. Effect sizes ranged from -0.98 to 1.74. The pooled effect size of 0.12 (95% CI, 0.003 to 0.24) was statistically significantly different from 0 ($P=.045$) but small and inconsistent ($I^2=88.1\%$). A sensitivity analysis ex-

Figure 2. Random-Effects Meta-analysis of Internet-Based Learning vs No Intervention: Knowledge Outcomes



Boxes represent the pooled effect size (Hedges *g*). *P* values reflect paired or 3-way comparisons among bracketed subgroups. Participant groups are not mutually exclusive; thus, no statistical comparison is made. There are 126 interventions because the report by Curran et al³⁹ contributed 10 separate interventions to this analysis. I^2 for pooling all interventions is 93.6%.

cluding blended interventions yielded a pooled effect size of 0.065 (95% CI, -0.062 to 0.19; $P = .31$).

In accord with our hypothesis, effect sizes were significantly higher for Internet-based courses using discussion vs no discussion (P for interaction = .002) (Figure 6). A statistically significant interaction favoring longer courses was also found (P for interaction = .03). However, our hypotheses regarding treatment-subgroup interactions across levels of interactivity, prac-

tice exercises, and repetition did not find support.

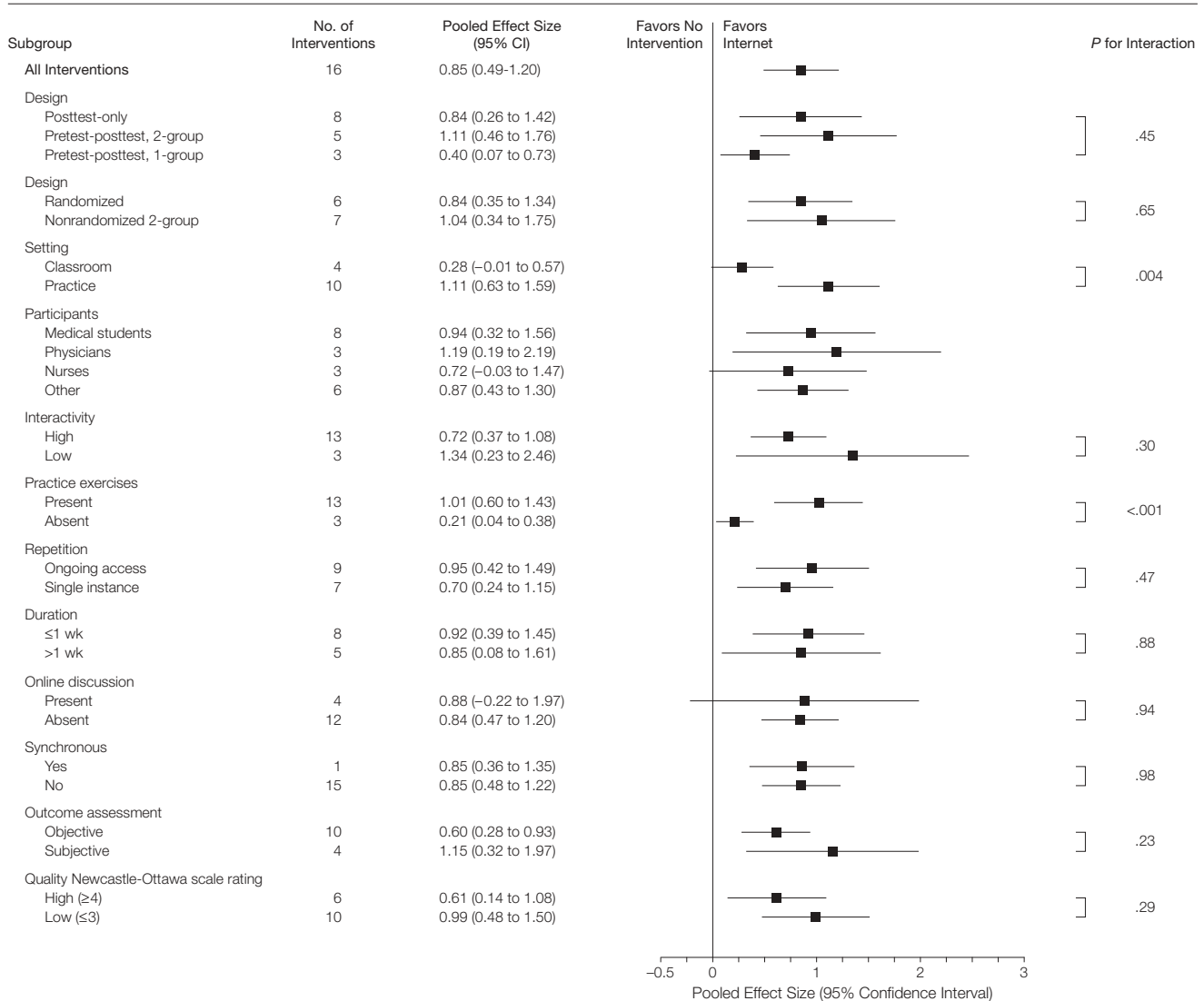
Skills. The twelve studies reporting skills outcomes demonstrated a small pooled effect size of 0.09 (95% CI, -0.26 to 0.44; $P = .61$). As with other outcomes, heterogeneity was large ($I^2 = 89.3\%$). Effect sizes ranged from -1.47 to 0.93.

We found statistically significant treatment-subgroup interactions (P for interaction $\leq .04$) favoring higher levels of interactivity, practice exercises,

and peer discussion (Figure 7). However, these analyses were limited by very small samples (in some cases only 1 study in a group). Contrary to our expectation, single-instance interventions yielded higher effect sizes than those with ongoing access (P for interaction = .02).

Behaviors and Effects on Patient Care. Six studies reported outcomes of behaviors and effects on patient care. The pooled effect size of 0.51 (95% CI, -0.24 to 1.25) was moderate in size, but

Figure 3. Random-Effects Meta-analysis of Internet-Based Learning vs No Intervention: Skills Outcomes



For a definition of figure elements, see the legend to Figure 2. All interventions were tutorials; hence, no contrast is reported for this characteristic. I^2 for pooling all interventions is 92.7%.

not statistically significant ($P = .18$). Inconsistency was large ($I^2 = 94.6\%$) and individual effect sizes ranged from -0.84 to 1.66 .

We again found a statistically significant treatment-subgroup interaction favoring discussion (P for interaction = .02); (Figure 8) but as with skills outcomes, the results are tempered by very small samples. Once again, single-instance interventions yielded higher effect sizes than those with ongoing access (P for interaction = .006).

COMMENT

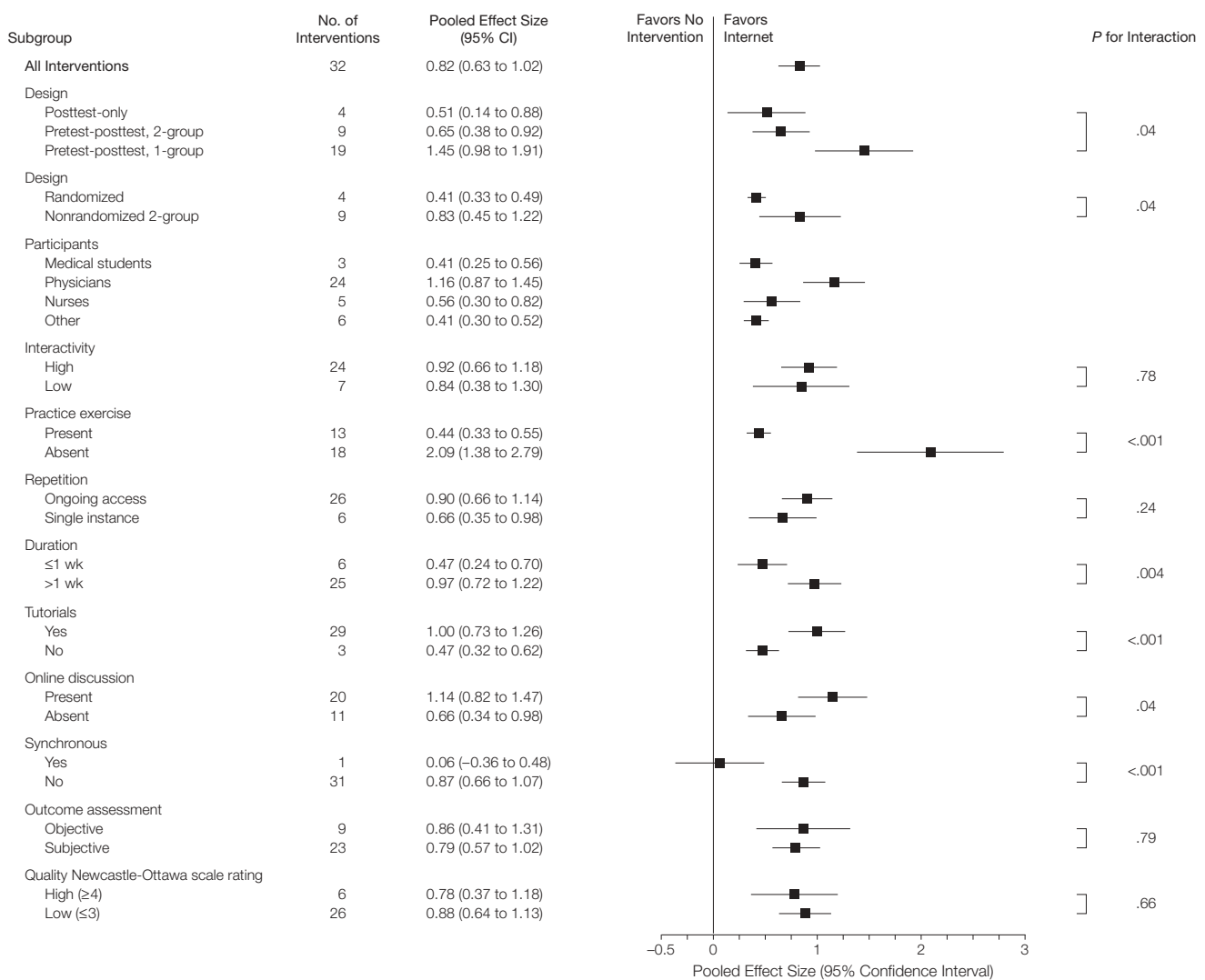
We found that Internet-based learning compared with no intervention has a consistent positive effect. The pooled estimate of effect size was large across all educational outcomes.⁴⁰ Furthermore, we found a moderate or large effect for nearly all subgroup analyses exploring variations in learning setting, instructional design, study design, and study quality. However, studies yielded inconsistent (heterogeneous) results, and subgroup com-

parisons only partially explained these differences.

The effect of Internet-based instruction in comparison to non-Internet formats was likewise inconsistent across studies. In contrast, the pooled effect sizes were generally small (≤ 0.12 for all but behavior or patient effects) and nonsignificant (CIs encompassing 0 [no effect] for all outcomes except knowledge).

Heterogeneity may arise from variation in learners, instructional methods, outcome measures, and other as-

Figure 4. Random-Effects Meta-analysis of Internet-Based Learning vs No Intervention: Behaviors in Practice and Effects on Patients



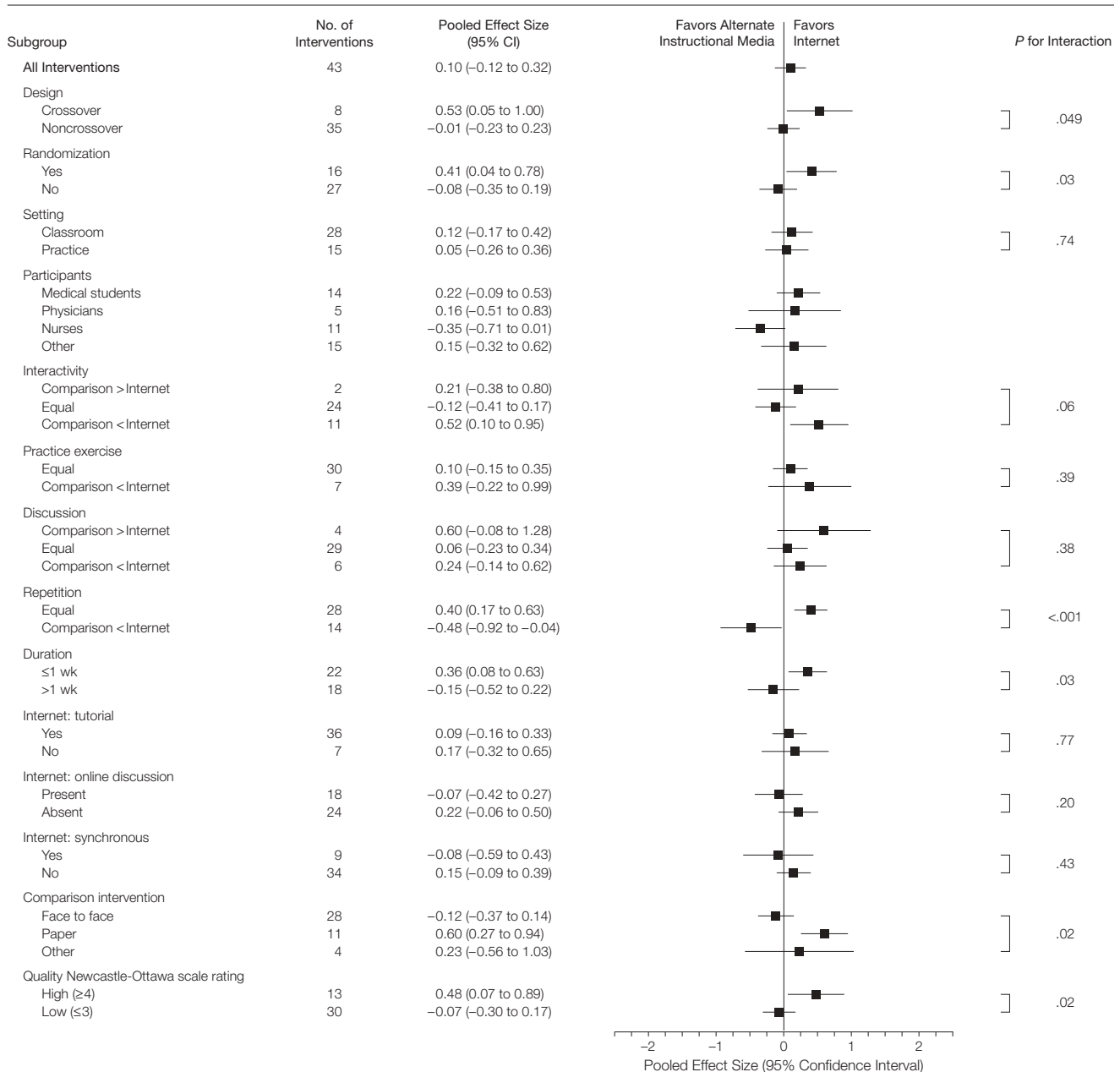
For a definition of figure elements, see the legend to Figure 2. All interventions occurred in a practice setting; hence, no contrast is reported for this characteristic. There are 32 interventions because the report by Curran et al³⁹ contributed 14 separate interventions to this analysis. I^2 for pooling all interventions is 79.1%.

pects of the educational context. For example, only 2 no-intervention controlled studies^{41,42} had negative effect sizes, and in both instances the lack of benefit could be ascribed to an educationally rich baseline or comparison.

Our hypotheses regarding changes in the magnitude of benefit for variations in instructional design were generally not supported by subgroup analyses, and in some cases significant differences were found in the direc-

tion opposite to our hypotheses. These findings were not consistent across outcomes or study types. Unexplained inconsistencies would allow us to draw only weak inferences if not for the preponderance of positive effects on all out-

Figure 5. Random-Effects Meta-analysis of Internet-Based Learning vs Alternate Instructional Media: Satisfaction Outcomes



Studies are classified according to relative between-intervention differences in key instructional methods; namely, did the comparison intervention have more (comparison >Internet), less (comparison <Internet), or the same (equal) amount of interactivity, practice exercises, discussion (face-to-face and Internet-based discussion combined), and repetition. Boxes represent the pooled effect size (Hedges *g*). *P* values reflect paired or 3-way comparisons among bracketed subgroups. Participant groups are not mutually exclusive; thus, no statistical comparison is made. All outcomes were subjectively determined; hence, no contrast is reported for this characteristic. Crossover studies assessed participant preference after exposure to Internet-based and non-Internet-based interventions. *I*² for pooling all interventions is 92.2%.

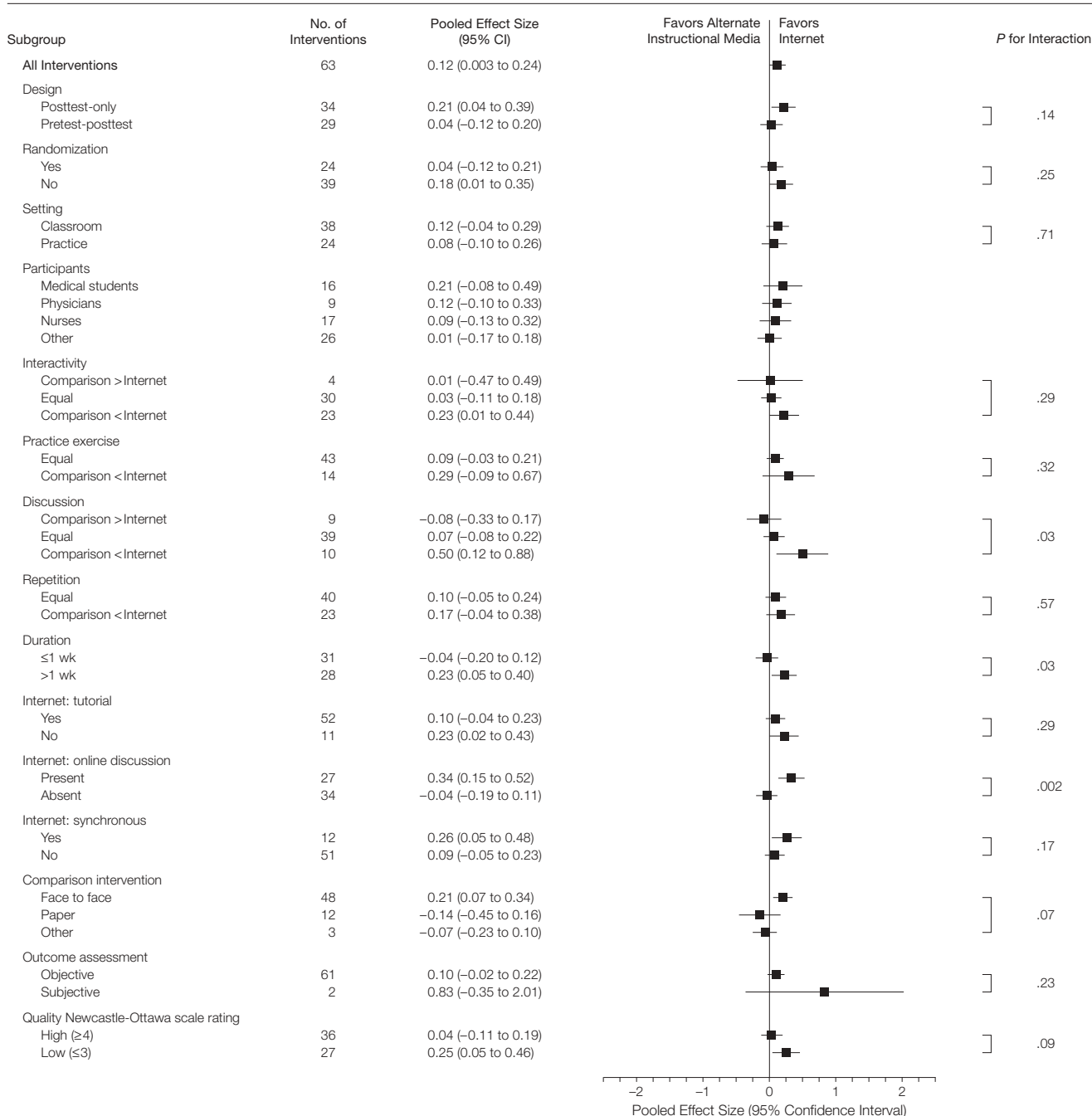
comes in the no-intervention comparison studies. For comparisons with non-Internet formats these inconsistencies make inferences tenuous. Additional research is needed to explore the inconsistencies identified in this review.

Limitations and Strengths

Our study has several limitations. First, many reports failed to describe key elements of the context, instructional design, or outcomes. Although the review process was conducted in duplicate,

coding was subjective and based on published descriptions rather than direct evaluation of instructional events. Poor reporting might have contributed to modest interrater agreement for some variables. Although we obtained addi-

Figure 6. Random-Effects Meta-analysis of Internet-Based Learning vs Alternate Instructional Media: Knowledge Outcomes



For a definition of figure elements and study parameters, see the legend to Figure 5. I² for pooling all interventions is 88.1%.

tional outcome data from several authors, we still imputed effect sizes for many studies with concomitant potential for error. Sparse reporting of validity and reliability evidence for assessment scores precluded inclusion of such evidence. Furthermore, methodological quality was generally low. However, subgroup and sensitivity analyses did not reveal consistently larger or smaller effects for different study de-

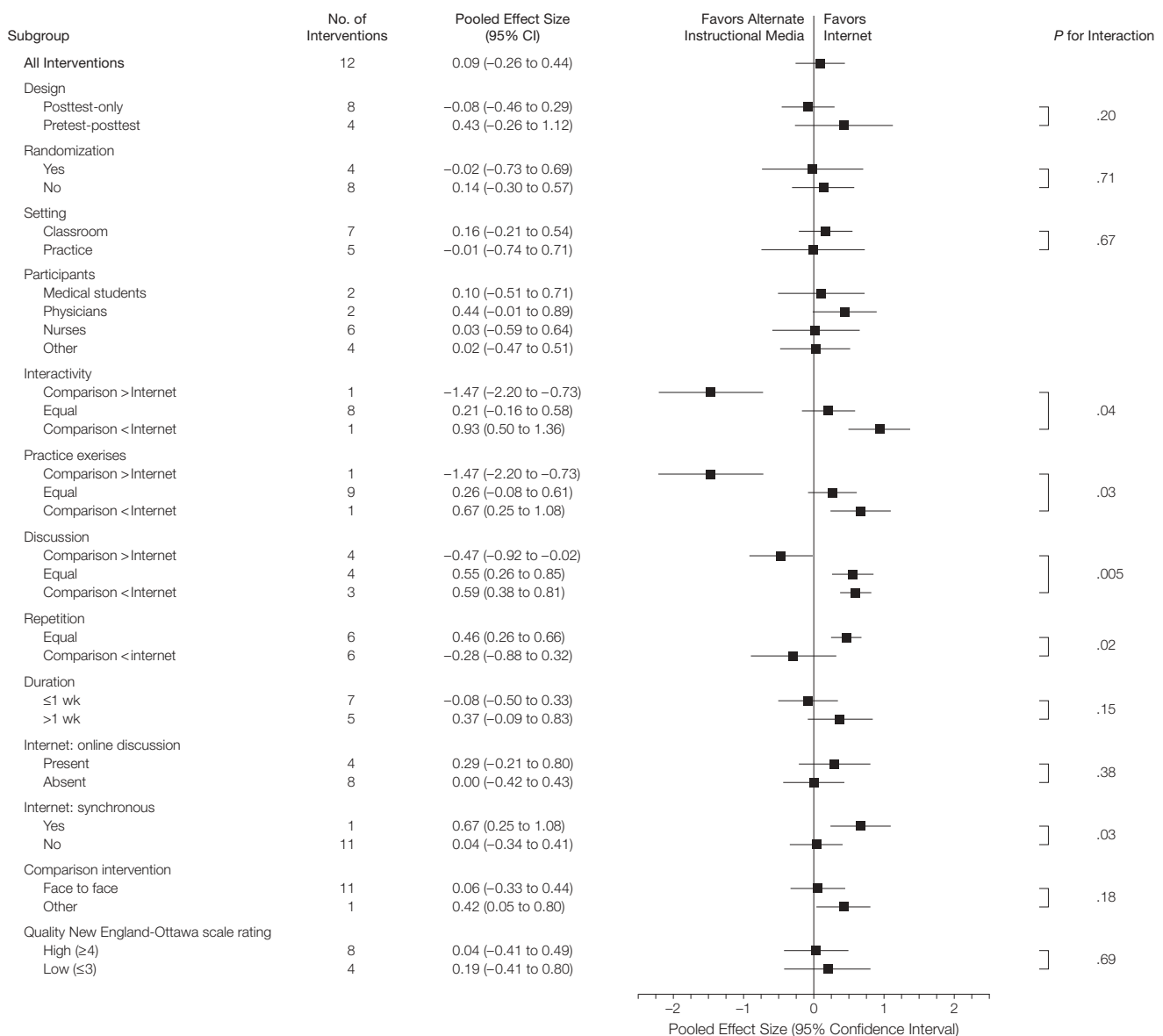
signs or quality or after excluding imputed effect sizes.

Second, interventions varied widely from study to study. Because nearly all no-intervention comparison studies found a benefit, this heterogeneity suggests that a wide variety of Internet-based interventions can be used effectively in medical education. Alternatively, this finding may indicate publication bias with negative studies remaining unpub-

lished. We did not use funnel plots to assess for publication bias because these are misleading in the presence of marked heterogeneity.⁴³

Third, we report our results using subgroups as an efficient means of synthesizing the large number of studies identified and simultaneously to explore heterogeneity. However, subgroup results should be interpreted with caution due to the number of compari-

Figure 7. Random-Effects Meta-analysis of Internet-Based Learning vs Alternate Instructional Media: Skills Outcomes



For a definition of figure elements and study parameters, see the legend to Figure 5. All interventions were tutorials, and all outcomes were objectively determined except for 1 study in which the method of assessment could not be determined; hence, no contrasts are reported for these characteristics. I^2 for pooling all interventions is 89.3%.

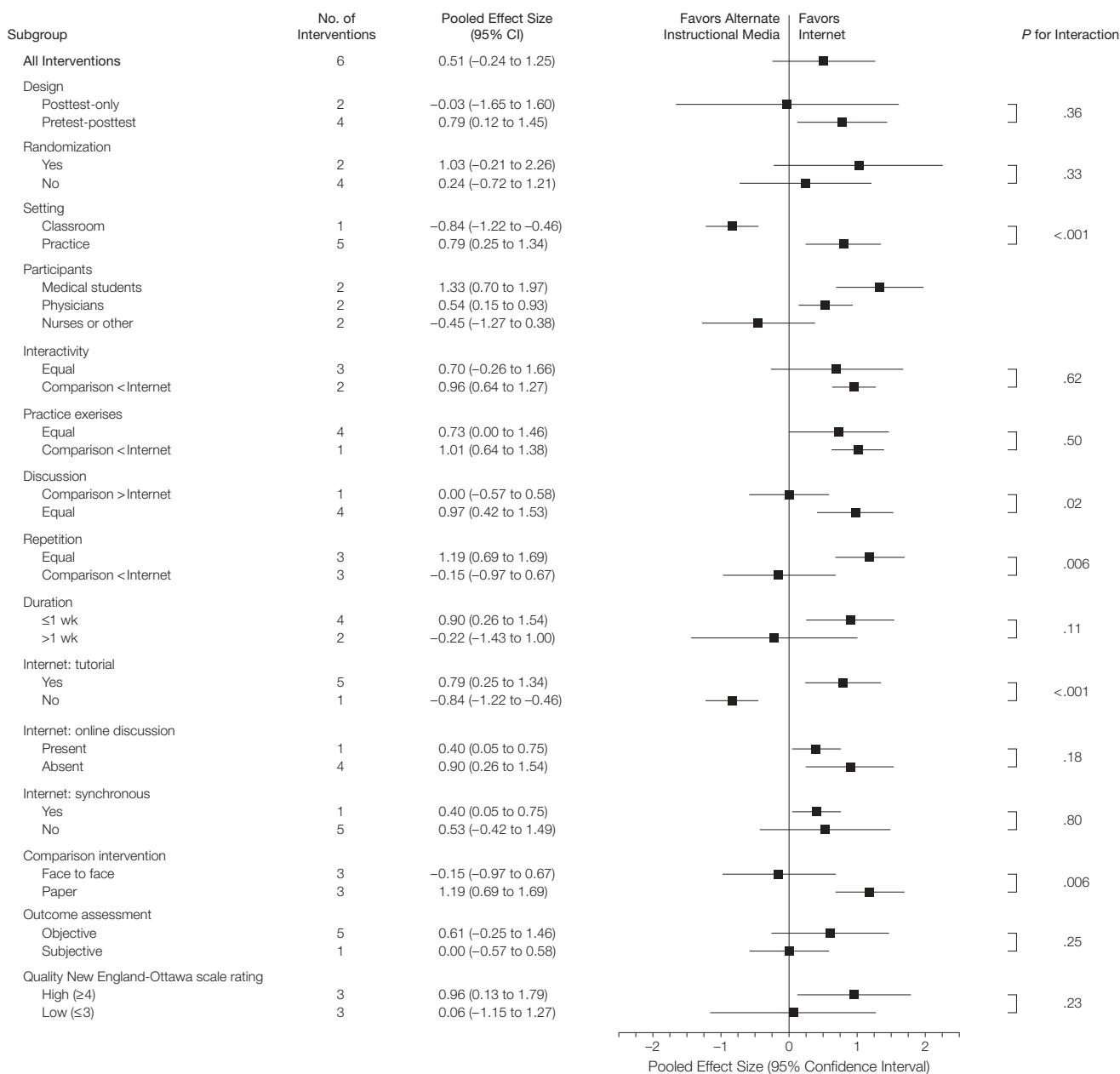
sons made, the absence of a priori hypotheses for many analyses, the limitations associated with between-study (rather than within-study) comparisons, and inconsistent findings across outcomes and study types.⁴⁴ For example, we found contrary to expectation that interventions with greater repetition (Internet-based course per-

mitting ongoing access vs non-Internet intervention available only once) had lower pooled effect sizes than interventions with equal repetition. These results could be due to chance, confounding, bias, or true effect. Another example is the finding that practice exercises were associated with higher effect sizes for skills outcomes

and lower effect sizes for behavior or patient effects; this could be explained by true differential effect of the interventions on these outcomes, variation in responsiveness across outcomes, unrecognized confounders, or chance.

Finally, by focusing our review on Internet-based learning, we of necessity

Figure 8. Random-Effects Meta-analysis of Internet-Based Learning vs Alternate Instructional Media: Behaviors in Practice and Effects on Patients



For a definition of figure elements and study parameters, see the legend to Figure 5. I² for pooling all interventions is 94.6%.

ignored a great body of literature on non-Internet-based computer-assisted instruction.

Our review also has several strengths. The 2 study questions are timely and of major importance to medical educators. We intentionally kept our scope broad in terms of subjects, interventions, and outcomes. Our search for relevant studies encompassed multiple literature databases supplemented by hand searches. We had few exclusion criteria, and included several studies published in languages other than English. All aspects of the review process were conducted in duplicate with acceptable reproducibility. Despite the large volume of data, we kept our analyses focused, conducting relatively few planned subgroup analyses to explain inconsistency and sensitivity analyses to evaluate the robustness of our findings to the assumptions of our meta-analyses.

Comparison With Previous Reviews

The last meta-analyses of computer-assisted instruction in health professions education^{45,46} were published in or before 1994, and computer-assisted instruction has changed dramatically in the interim. To the 16 no-intervention controlled and 9 non-Internet comparative studies reported in the last comprehensive review of Web-based learning,⁹ we add 176 additional articles as well as a meta-analytic summary of results. This and other reviews^{11-14,16,17,47} concur with the present study in concluding that Internet-based learning is educationally beneficial and can achieve results similar to those of traditional instructional methods.

Implications

This review has implications for both education and research. Although conclusions must be tempered by inconsistency among studies and the possibility of publication bias, the synthesized evidence demonstrates that Internet-based instruction is associated with favorable outcomes across a wide

variety of learners, learning contexts, clinical topics, and learning outcomes. Internet-based instruction appears to have a large effect compared with no intervention and appears to have an effectiveness similar to traditional methods.

The studies making comparison with no intervention essentially asked whether a Web-based course in a particular topic could be effective. The answer was almost invariably yes. Given this consistency of effect and assuming no major publication bias, there appears to be limited value in further research comparing Internet-based interventions against no-intervention comparison groups. Although no-intervention controlled studies may be useful in proof-of-concept evaluations of new applications of Internet-based methods (such as a study looking at rater training on the Web⁴⁸), truly novel innovations requiring such study are likely to be increasingly rare and will infrequently merit publication.

Studies making comparison to alternate instructional media asked whether Internet-based learning is superior to (or inferior to) traditional methods. In contrast to no-intervention controlled studies, the answers to this question varied widely. Some studies favored the Internet, some favored traditional methods, and on average there was little difference between the 2 formats. Although the pooled estimates favored Internet-based instruction, for all but behavior or patient effects the magnitude of benefit was small and could be explained by sources of variation noted above or by novelty effects.²⁷ These findings support arguments that computer-assisted instruction is neither inherently superior to nor inferior to traditional methods.^{10,27-29} Few non-Internet comparison studies reported skills and behavior or patient effects outcomes, and the CIs for these pooled estimates do not exclude educationally significant effects. Additional research, using outcome measures responsive to the intervention and sensitive to change, would be required to improve the precision of these esti-

mates. However, inconsistencies in the current evidence together with conceptual concerns^{27,28} suggest limited value in further research seeking to demonstrate a global effect of Internet-based formats across learners, content domains, and outcomes.

The inconsistency in effect across both study types suggests that some methods of implementing an Internet-based course may be more effective than others. Thus, we propose that greater attention be given to the question, "How can Internet-based learning be effectively implemented?" Elucidating how to effectively implement Internet-based instruction will be answered most efficiently through research directly comparing different Internet-based interventions.^{7,10,27-29,49} Inconsistency may also be due to different learning contexts and objectives, and thus the question, "When should Internet-based learning be used?" should be considered as well.¹⁰

Finally, although our findings regarding the quality of this body of research are not unique to research in Internet-based instruction,⁵⁰⁻⁵² the relatively low scores for methodological quality and the observed reporting deficiencies suggest room for improvement.

Author Contributions: Dr Cook had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Cook, Levinson, Dupras, Garside, Erwin, Montori.

Acquisition of data: Cook, Levinson, Dupras, Garside, Erwin.

Analysis and interpretation of data: Cook, Montori.

Drafting of the manuscript: Cook.

Critical revision of the manuscript for important intellectual content: Cook, Levinson, Dupras, Garside, Erwin, Montori.

Statistical analysis: Cook, Montori.

Obtained funding: Cook.

Administrative, technical or material support: Cook, Montori.

Study supervision: Cook.

Financial Disclosures: None reported.

Funding/Support: This work was supported by intramural funds and a Mayo Foundation Education Innovation award.

Role of Sponsor: The funding source for this study played no role in the design and conduct of the study; in the collection, management, analysis, and interpretation of the data; or in the preparation of the manuscript. The funding source did not review the manuscript.

Additional Information: Details on included studies and their quality, and on the meta-analyses are available at <http://www.jama.com>.

Additional Contributions: We thank Melanie Lane, BA, Mohamed Elamin, MBBS, and M. Hassan Murad, MD, from the Knowledge and Encounter Research Unit, Mayo Clinic, for assistance with data extraction and meta-analysis planning and execution, and Kathryn Trana, from the Division of General Internal Medicine, Mayo Clinic, for assistance in article acquisition and processing. These individuals received compensation as part of their regular employment.

REFERENCES

- Berners-Lee T, Cailliau R, Luotonen A, Nielsen HF, Secret A. The World-Wide Web. *Commun ACM*. 1994; 37(8):76-82.
- Friedman RB. Top ten reasons the World Wide Web may fail to change medical education. *Acad Med*. 1996; 71(9):979-981.
- MacKenzie JD, Greenes RA. The World Wide Web: redefining medical education. *JAMA*. 1997;278(21):1785-1786.
- Ruiz JG, Mintzer MJ, Leipzig RM. The Impact of e-learning in medical education. *Acad Med*. 2006; 81(3):207-212.
- Cook DA. Web-based learning: pro's, con's, and controversies. *Clin Med*. 2007;7(1):37-42.
- Effective Use of Educational Technology in Medical Education: Summary Report of the 2006 AAMC Colloquium on Educational Technology*. Washington, DC: Association of American Medical Colleges; 2007.
- Tegtmeyer K, Ibsen L, Goldstein B. Computer-assisted learning in critical care: from ENIAC to HAL. *Crit Care Med*. 2001;29(8)(suppl):N177-N182.
- Davis MH, Harden RM. E is for everything—e-learning? *Med Teach*. 2001;23(5):441-444.
- Chumley-Jones HS, Dobbie A, Alford CL. Web-based learning: sound educational method or hype? a review of the evaluation literature. *Acad Med*. 2002; 77(10)(suppl):S86-S93.
- Cook DA. Where are we with Web-based learning in medical education? *Med Teach*. 2006;28(7):594-598.
- Greenhalgh T. Computer assisted learning in undergraduate medical education. *BMJ*. 2001;322(7277):40-44.
- Lewis MJ, Davies R, Jenkins D, Tait MI. A review of evaluative studies of computer-based learning in nursing education. *Nurse Educ Today*. 2001;21(1):26-37.
- Wutoh R, Boren SA, Balas EA. eLearning: a review of Internet-based continuing medical education. *J Contin Educ Health Prof*. 2004;24(1):20-30.
- Chaffin AJ, Maddux CD. Internet teaching methods for use in baccalaureate nursing education. *Comput Inform Nurs*. 2004;22(3):132-142.
- Curran VR, Fleet L. A review of evaluation outcomes of Web-based continuing medical education. *Med Educ*. 2005;39(6):561-567.
- Hammoud M, Gruppen L, Erickson SS, et al. To the point: reviews in medical education online computer assisted instruction materials. *Am J Obstet Gynecol*. 2006;194(4):1064-1069.
- Potomkova J, Mihal V, Cihalik C. Web-based instruction and its impact on the learning activity of medical students: a review. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 2006;150(2):357-361.
- Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet*. 1999;354(9193):1896-1900.
- Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA*. 2000;283(15):2008-2012.
- Bransford JD, Brown AL, Cocking RR, et al. *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press; 2000.
- Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor-Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *JAMA*. 1999;282(9):867-874.
- Mayer RE. Cognitive theory of multimedia learning. In: Mayer RE, ed. *The Cambridge Handbook of Multimedia Learning*. New York, NY: Cambridge University Press; 2005:31-48.
- Cook DA, Thompson WG, Thomas KG, Thomas MR, Pankratz VS. Impact of self-assessment questions and learning styles in Web-based learning: a randomized, controlled, crossover trial. *Acad Med*. 2006; 81(3):231-238.
- Cook DA, McDonald FS. E-learning: is there anything special about the "E"? *Perspect Biol Med*. 2008; 51(1):5-21.
- Marinopoulos SS, Dorman T, Ratanawongsa N, et al. Effectiveness of continuing medical education. *Evid Rep Technol Assess (Full Rep)*. 2007;149:1-69.
- Shea JA. Mind the gap: some reasons why medical education research is different from health services research. *Med Educ*. 2001;35(4):319-320.
- Clark RE. Reconsidering research on learning from media. *Rev Educ Res*. 1983;53:445-459.
- Cook DA. The research we still are not doing: an agenda for the study of computer-based learning. *Acad Med*. 2005;80(6):541-548.
- Friedman CP. The research we should be doing. *Acad Med*. 1994;69(6):455-457.
- Kirkpatrick D. Revisiting Kirkpatrick's four-level model. *Train Dev*. 1996;50(1):54-59.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979; 86:420-428.
- Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm. Accessed June 16, 2008.
- Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol Methods*. 2002; 7(1):105-125.
- Dunlap WP, Cortina JM, Vaslow JB, Burke MJ. Meta-analysis of experiments with matched groups or repeated measures designs. *Psychol Methods*. 1996; 1:170-177.
- Hunter JE, Schmidt FL. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Thousand Oaks, CA: Sage; 2004.
- Curtin F, Altman DG, Elbourne D. Meta-analysis combining parallel and cross-over clinical trials, I: continuous outcomes. *Stat Med*. 2002;21(15):2131-2144.
- Higgins JP, Green S. *Cochrane Handbook for Systematic Reviews of Interventions (Version 5.0.0)*. Available at: <http://www.cochrane.org/resources/handbook/index.htm>. Updated February 2008. Accessed 29 May 2008.
- Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003; 327(7414):557-560.
- Curran V, Lockyer J, Sargeant J, Fleet L. Evaluation of learning outcomes in Web-based continuing medical education. *Acad Med*. 2006;81(10)(suppl): S30-S34.
- Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.
- Mehta MP, Sinha P, Kanwar K, Inman A, Albanese M, Fahl W. Evaluation of Internet-based oncologic teaching for medical students. *J Cancer Educ*. 1998; 13(4):197-202.
- Patterson R, Harasym P. Educational instruction on a hospital information system for medical students during their surgical rotations. *J Am Med Assoc*. 2001;285(2):111-116.
- Lau J, Ioannidis JPA, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *BMJ*. 2006; 333(7568):597-600.
- Oxman A, Guyatt G. When to believe a subgroup analysis. In: Hayward R, ed. *Users' Guides Interactive*. Chicago, IL: JAMA Publishing Group; 2002. <http://www.usersguides.org>. Accessed August 14, 2008.
- Cohen PA, Dacanay LD. Computer-based instruction and health professions education: a meta-analysis of outcomes. *Eval Health Prof*. 1992;15: 259-281.
- Cohen PA, Dacanay LD. A meta-analysis of computer-based instruction in nursing education. *Comput Nurs*. 1994;12(2):89-97.
- Lewis MJ. Computer-assisted learning for teaching anatomy and physiology in subjects allied to medicine. *Med Teach*. 2003;25(2):204-206.
- Kobak KA, Engelhardt N, Lipsitz JD. Enriched rater training using Internet based technologies: a comparison to traditional rater training in a multi-site depression trial. *J Psychiatr Res*. 2006;40(3):192-199.
- Keane DR, Norman G, Vickers J. The inadequacy of recent research on computer-assisted instruction. *Acad Med*. 1991;66(8):444-448.
- Cook DA, Beckman TJ, Bordage G. Quality of reporting of experimental studies in medical education: a systematic review. *Med Educ*. 2007;41(8): 737-745.
- Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. *JAMA*. 2007;298(9):1002-1009.
- Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach*. 2005; 27(1):10-28.