A Systematic Review: The Effect of Clinical Supervision on Patient and Residency Education Outcomes

Jeanne M. Farnan, MD, MHPE, Lindsey A. Petty, MD, Emily Georgitis, MD, Shannon Martin, MD, Emily Chiu, Meryl Prochaska, and Vineet M. Arora, MD, MAPP

Abstract

Purpose
To summarize the literature regarding the effect of clinical supervision on patient and educational outcomes, especially in light of the recent (2010) Accreditation Council for Graduate Medical Education report that recommends augmented supervision to improve resident education and patient safety.

Method
The authors searched the English-language literature from 1966 to 2010 using electronic databases and a hand search. They included studies that described a controlled design, and they have relayed the effects of supervision on patient- and education-related outcomes. Two authors abstracted prescribed data from the reviewed studies. The authors rated the quality of each study using the Medical Education Research Study Quality Instrument.

Results
Twenty-four articles across a variety of specialties (i.e., psychiatry, emergency medicine, surgery, anesthesia, and internal medicine) met inclusion criteria.

Conclusions
Enhanced clinical supervision of trainees has been associated with improved patient- and education-related outcomes in published studies. Future work should focus on developing validated measures of the effects of clinical supervision.

The Residency Review Committee of the Accreditation Council for Graduate Medical Education (ACGME) has put forth general program requirements, which cover certification, training, and the availability of clinical supervisors. These requirements inform institutional policies regarding resident supervision. Across institutions, policy statements share the common themes of patient safety, clinical care, and trainee education while also highlighting the importance of affording housestaff graduated levels of responsibility throughout the course of their training. Trainees themselves have identified inadequate supervision as one of the most common causes of the medical errors that occur during a patient’s hospitalization.

Specific subspecialties had, prior to the 2010 changes to duty hours regulations, defined their own levels of supervision; some specialties have required the physical presence of an attending during the delivery of key aspects of care since as early as 1990. In 2008, after revisiting the 2003 duty hours regulations, the Institute of Medicine (IOM) recommended in its report Resident Duty Hours: Enhancing Sleep, Supervision and Safety that graduate medical training programs should augment resident supervision, including on nights and weekends, through immediate access to an on-site, residency-approved supervisor. Further, the most recent iteration (2010) of the ACGME recommendations for duty hours include (in addition to further limitations in the consecutive duty hours of postgraduate year 1 [PGY-1] trainees) explicit recommendations for trainee supervision. These new recommendations include defining explicit requirements for trainee supervision in subspecialties not previously defined, and assessing a trainee's ability to supervise more junior colleagues before he or she actually takes on a supervisory role.

Whereas narrative reviews and observational studies have concluded that enhanced attending supervision results in improved education and patient outcomes, to date no systematic review of the literature on the effects of attending supervision has been published. Systematic review is critical to identify whether, how, and what type of supervision improves resident education and patient outcomes. Examining the evidence is especially important given the current debate over how to implement the above-mentioned 2010 ACGME recommendations regarding the supervision of trainees. To this end, we performed a systematic review of the literature to describe the effect of clinical supervision on residency patient- and education-related outcomes.

Method
The 2009 PRISMA Statement on the content and reporting of systematic review manuscripts informed our research methods. We searched the English-language literature for studies regarding clinical supervision of housestaff (PGY-1 or greater) in U.S.- and Canadian-based training programs from 1966 (the beginning of electronic indexing of references) to 2010 using the electronic databases PubMed and JSTOR. Given the fundamental differences in the current approach to and regulation of supervision of trainees in programs throughout the world, we decided to...
focus exclusively on medical training programs in the United States and Canada. After consulting a medical librarian, we performed a comprehensive search in August 2009 and again in November 2010 using the following strategy: [“Education, Medical, Graduate” [Mesh] OR “Internship and Residency” [Mesh]] AND (supervis* OR oversigh*). We conducted a hand search for journals not electronically indexed between 1966 and 2010 (e.g., Medical Education). We also included pertinent references from the studies that we included in the abstraction process. Finally, we reviewed proceedings from medical-education-focused meetings (e.g., the Association of American Medical Colleges annual meeting and the ACGME annual meeting) from 2007 through 2010, using electronic and hand searching methods.

We included studies (1) that described controlled study designs (e.g., randomized controlled trials [RCTs], concurrent control, or pre–post studies), (2) that examined the implementation of a supervision-related intervention among graduate medical education trainees (PGY-1 to highest level of postgraduate year of specialty in question) in any field of practice in any clinical setting (e.g., academic or community based; inpatient or ambulatory), and (3) that relayed patient- or education-related outcomes.

All of us reviewed the titles and abstracts of all the articles our search uncovered (see Figure 1). Four of us (L.P., M.P., S.M., and E.C.) performed the initial data abstraction of, respectively, about a quarter of the selected studies, and two of us (V.A., J.F.) reviewed those data to confirm results and resolve any discrepancies. We used a standardized abstraction tool (see Supplemental Digital Appendix 1, http://links.lww.com/ACADMED/A76) that requested the following:

- the role of the supervisor and supervisee,
- the study design and site (e.g., academic medical center [AMC], community-based program),
- the type of supervision (see below),
- the setting (e.g., inpatient unit, ambulatory clinic),
- the year(s) of the study,
- the intervention,
- the number of participants,
- duration of follow-up, and
- supervision-related outcomes.

The data abstraction tool defined four types of supervision: (1) direct, with the supervisor physically in the presence of the patient and the trainee, (2) indirect, with the supervisor immediately available at the site of care or (3) indirect, with the supervisor available via phone or pager and able to be physically present within a short period of time, and (4) oversight, in which the supervising physician reviews the care after the care has been delivered.6,8,15

To assess the quality of study designs, we used the previously validated Medical Education Research Study Quality Instrument (MERSQI).16 The MERSQI evaluates six domains of study quality: design, sampling, type of data, validity, data analysis, and outcomes. Items are scored on an ordinal scale with a maximum 18 allowable points.16 At least one of us rated the quality of each included study, and we resolved all disagreements by consensus.

### Results

Our initial search uncovered 1,451 articles. Of those, we selected 252 for detailed review. We excluded a total of 228 studies: 112 were not research reports, 95 were observational studies, 20 were not relevant to clinical supervision, and 1 occurred outside of North America (see Figure 1). We report here on 24 English-language studies evaluating clinical supervision interventions on patient-care- and education-related outcomes in a controlled fashion. Appendix 1, organized by specialty and quality score, summarizes some of the key results of each of the individual studies that we reviewed.

### Specialties represented and types of supervision evaluated

The 24 studies that met inclusion criteria represent the following specialties: psychiatry,17-21 emergency medicine,10,22-24 surgery and its subspecialties,25-30 obstetrics–gynecology,31 internal medicine,32-35 anesthesiology,36 pediatrics,37 and radiology.38,39 Of the included studies, six examined the impact of a direct supervision intervention,10,17,24,26,28,31,34 seven evaluated oversight,20,22,23,32,33,35,38,39 and four evaluated indirect models of supervision.26,29,30,36 Seven studies compared the types of supervision with one another; four evaluated direct versus oversight,18,19,21,37 and three compared direct and either type of indirect supervision.23,25,27 Only three studies were RCTs,18,20,33 The remaining studies used either a pre–post design,17,22,23,32,34,36 case–control design,29 retrospective cohort design,19,25,27,28,30,31,35,36 or prospective cohort design.10,21,24,26,37,39 Most studies were single site with small sample sizes.

### Setting of supervision

Of the 24 studies we reviewed, 5 (20.8%) examined supervision in the ambulatory setting,18-20,33,34,35 and the remaining 19...
Outcomes evaluated and quality of studies

Twenty-one of the studies evaluated the effect of supervision on patients,10,18–28,30–36,38,39 and six studies20,29,34,35,37 evaluated educational outcomes (three of the studies examined the effect of supervision on both patients and education). Using the MERSQI scoring criteria to assess study quality,16 scores ranged from 8.5 to 16, with a mean score of 11.8. The average MERSQI for the 21 studies evaluating patient-related outcomes was 11.9, and that for the 6 studies evaluating education-related outcomes was 11.0.

Patient-related outcomes: Overview

The 21 studies that evaluated the impact on patient-related outcomes examined complications related to diagnostic or therapeutic procedures,25–28,30,31,36 the treating physician’s perception of the severity of the patient’s illness,18,21,34 supervisor-initiated changes to missed or inappropriate diagnoses,22,24,30,32 greater success at continuing therapy,18,20 and resident practice patterns influenced by attending, such as resource utilization19,23,33,35 or compliance with protocols.10,32

Complications related to diagnostic or therapeutic procedures. Eight studies evaluated the impact of a supervisory intervention on the performance of an invasive procedure.25–28,30,31,36 A retrospective cohort study of nearly 40,000 cases by Itani and colleagues25 revealed no significant difference in mortality between patient groups with and patient groups without an attending physically present, indicating that the degree of trainee supervision did not affect the overall clinical outcomes for patients, when adjusted for patient factors. A study that evaluated the impact of direct attending supervision of trainees on postoperative pain26 scores showed no significant difference between procedures performed by the attending surgeon alone versus those performed by a resident with direct attending-level supervision.

A medical record audit showed that greater attending presence was associated with significantly lower mortality and patient complication rates when procedures were stratified according to elective (e.g., plastic) surgical specialties and nonelective (e.g., trauma, cardiac) surgical specialties at a teaching hospital or AMC.27 However, Rogers and colleagues28 demonstrated no significant differences in mortality in patient groups at two tertiary care centers: one with a resident in-house (RIH) alone, and one with an in-house, board-certified attending surgeon. These researchers did note a greater chance of both technical and provider error among the RIH group.28

Lastly, from the surgery literature, a retrospective chart review of bilateral mammoplasty at an AMC showed no significant differences in complication rates between the side operated on by the attending alone versus the side operated on by the supervised resident.30

In related obstetrics–gynecology and anesthesia literature, a pre–post study of anesthesia resident intubations outside of the operating room revealed a significant decrease in intubation-related complications when attending presence was instituted.36 Finally, a medical record audit in obstetrics–gynecology revealed that the operative time for the teaching cases (i.e., those with residents and attendings) was longer (not significantly) than for the nonteaching cases and that a significantly higher incidence of seromas postoperatively occurred in the nonteaching cases.31

These studies of resident supervision in the operating room demonstrated either a positive or neutral effect on patient-specific outcomes of all types (e.g., complications, mortality, operating room time, pain) when trainees were directly or indirectly supervised by attending providers.

Physician perception of patient illness severity. Three studies focused on the perception of an illness’s severity, specifically on how attending providers rated patient acuity.18,21,34 Gennis and Gennis34 found that attendings who personally evaluated patients judged the patients to be more severely ill than did the resident. Similarly, Stein and colleagues21 noted a statistically significant difference in acuity ratings of patient illness; attendings who heard cases (but were not present for the patient interview) judged the illness to be of significantly higher acuity on validated scales than did the residents who actually (without supervision) interviewed the patients. These authors did not note differences between the acuity ratings of the resident and the attending if both were present for the resident’s interview of the patient; however, the overall acuity ratings of attendings and residents (across cases) were generally higher if both the resident and attending were present for the interview (versus lower overall acuity ratings for the cases in which the resident conducted the interview without supervision).

Conversely, Jaynes and colleagues18 found that residents who evaluated patients under direct attending supervision judged more patients to be “moderately ill” and fewer patients to be “severely ill” than did residents whose attendings simply provided oversight. In sum, the perception of patient acuity varied with the direct presence of the attending, with two studies demonstrating direct evaluation to result in more severe acuity ratings.21,34

Supervisor-initiated changes or missed diagnoses. Four studies involved supervisor-initiated changes made to missed or inappropriate diagnoses: two from the radiology literature38,39 and two from emergency medicine.22,24 Albano and colleagues38 demonstrated nearly identical faculty and resident interpretations of abdominal CT scans in the setting of suspected appendicitis. Similarly, Velhamos and colleagues59 prospectively evaluated discrepancies between resident and attending interpretations of trauma-related cases, and only 11% of faculty interpretations were found to be different than the preliminary resident interpretation. Notably, the resident postgraduate year correlated with the likelihood of missed findings: the more experienced the...
resident, the lower the likelihood of a discrepancy between resident and faculty.39
The emergency medicine studies focused on change in patient management. Holliman and colleagues22 conducted a pre–post study, in which attendings reviewed cases after residents performed traditional patient presentations. Their results showed that attendings not only noted significant missed findings, including some that were life threatening, but also changed treatment plans in their postpresentation review of cases.22 There was no stratification for postgraduate year, although PGY-1 trainees were excluded. Finally, Sacchetti34 performed a prospective review of resident treatment plans, specifically looking for the degree of change attendings noted after they reviewed the plan; a third of cases required minor modifications, and several required major ones.

Studies examining oversight provided by attending physicians in the delivery of direct patient care did demonstrate that the attending physician frequently modified the treatment or diagnostic plan of the resident physician, to either minor or major degrees; however, when the oversight occurred in a more objective manner (e.g., review of a radiologic image), investigators noted fewer discrepancies between faculty and trainees.

**Continuing therapy.** Only two studies looked at the effect of supervision on patients’ continuing therapy.18,20 The investigators of one RCT18 judged the patients of directly supervised residents in an ambulatory clinic in psychiatry to have overall improved status and greater success at continuing treatment. An additional RCT in psychiatry20 noted similar findings: Directly supervised residents reported higher ratings of patient outcomes on a nominal scale.

**Resident practice patterns.** Six studies evaluated resident practice patterns—specifically, resource utilization19,23,33,35 and compliance with protocols10,32—as influenced by the attending supervisor. Two internal medicine studies evaluated resource utilization from the ambulatory33 and inpatient perspective.35 An RCT by Everett33 revealed that residents’ outpatient laboratory testing utilization was not influenced by the practice patterns of their supervising attending physician. Feinglass and colleagues35 examined the relationship between resource use and perceived level of autonomy in a retrospective review; their results revealed that higher perceived autonomy predicted significantly lower resource use. Jin et al19 evaluated resource utilization with respect to constant observation use in the field of psychiatry. They found that residents who received training and had more intensive attending-level supervision were less likely than residents with no training and standard oversight to order constant observation for patients. Kroboth and colleagues33 performed an assessment among the admitting practices of emergency medicine residents who experienced varying degrees of supervision, and found the following among residents with greater supervision: a significant increase in hospital admission rates, a significant decrease in the use of the observation unit, and a significant decrease in patients leaving without an examination. The impact of attending presence on resource utilization is equivocal, with increased admission rates demonstrated but decreased utilization of intensive patient observation orders.

Two studies examine protocol adherence. Blumberg and colleagues32 evaluated resident compliance with inpatient protocols for high-risk patients and noted that, after the community teaching hospital initiated faculty review and discussion of the medical records that housestaff completed, resident compliance became more consistent with prescribed protocol. Sox et al10 demonstrated that directly supervised residents had a higher mean rating of compliance with the process-of-care guidelines in the emergency department setting than did other residents.

Overall, attending supervision demonstrated a positive effect in resident adherence to quality-of-care guidelines but did not appear to significantly affect resource utilization, either negatively or positively.

**Education-related outcomes: Overview**

Six studies evaluated the impact of supervision on education-related outcomes—specifically, attendings’ perception of residents’ diagnostic or procedural skills,29,34 residents’ performance on standardized testing,17 residents’ perception of autonomy,35 and residents’ overall satisfaction.17

**Diagnostic or procedural skills.** Two studies report the effect of supervision on diagnostic skills related to history taking, physical examination, and procedures. Gennis and Gennis34 conducted a pre–post study in which attendings rated the diagnostic skills (i.e., history taking and physical examination) of internal medicine housestaff first after hearing the resident’s case presentation exclusively and, then, after examining the patient with the resident. The attendings’ ratings of housestaff skills decreased significantly after attendings performed the actual patient examination.34 With respect to procedural skills, Eroso and colleagues29 performed a case–control study of the use of teleproctoring on surgical skills in general surgery residents. They compared the residents’ scores on proctored and unproctored surgeries and found that all scores were significantly superior for the teleproctored surgeries.

**Performance on standardized testing, perception of autonomy, and overall satisfaction.** Osborn and colleagues37 conducted a prospective study of residents who self-selected into one of three types of continuity clinics (university based, community based, or private practice) and experienced varying types of attending-level supervision relative to their location. They found that the differences in supervision did not change the residents’ performance on the pediatrics in-training exam when analyzed by site.37 As mentioned previously, Feinglass and colleagues35 found, in their study of resource use and resident autonomy, that higher mean autonomy ratings were significant predictors of lower resource use and shorter length of patient stay. And, finally, in terms of the perception of the quality of the educational experience, one RCT20 compared cases reviewed by an attending with peer-reviewed cases. The investigators found no statistically significant difference in the perceived quality of peer versus attending case review.

**Overnight supervision**

The one study17 that examined the impact of attending-level supervision during a previously unsupervised rotation, the overnight shift, evaluated the self-reported effect of nighttime attending supervision on psychiatry.
residents in the emergency department of an AMC. Attendings were to be physically present in the emergency department in order to provide a direct, case-based educational experience and to provide assistance in the event of high patient load. Housestaff were not placed onto the night rotation until they had completed six months of training, and night duty was divided equally amongst residents. Trainees completed surveys both before and after these changes; the postimplementation surveys revealed a nonsignificant increase in comfort, a nonsignificant decrease in anxiety and isolation, and a nonsignificant increase in the sharing of responsibility. Costs of increased supervision

None of the studies we reviewed commented either on the cost of maintaining attending-level supervision or on the additional cost of maintaining an attending in-house in lieu of trainees independently. The study that demonstrated longer operative times when attending-level providers directly supervised residents did not comment on the impact of longer case time on the overall system, with respect to cost and throughput.

Discussion

The results of this review indicate that enhanced attending supervision of trainees, especially in activities already supervised to at least the oversight degree, result in positive changes in patient- and education-related outcomes. Few studies showed that attending presence had no effect on specified outcomes, but two showed increased incidence of surgical complications in nonteaching environments. Faculty personally evaluating a patient, with or without the trainee present, often resulted in a change in the overall assessment of the diagnostic or therapeutic plan and, more often than not, resulted in patients being assessed as less severely ill. Missed diagnoses when attendings were absent, as well as major changes to care plans after attendings reviewed residents’ work, did occur, albeit infrequently, and most studies showed that these errors and changes were related to the level of training of the resident. Lack of data and objective, standardized metrics to measure supervisory quality present challenges to making strong conclusions and represent areas for future research to better delineate the impact of supervision on trainees.

This review has implications for the current debate on the necessity of continuous, on-site, attending-level supervision initially recommended by the 2008 IOM report and subsequently explicitly endorsed by ACGME regulation. Importantly, this review does suggest that enhanced attending-level supervision, in a variety of ways, improves both patient- and education-related outcomes in the majority of subspecialties and a majority of clinical settings. The benefits for trainee education are present in both ambulatory and inpatient settings. Further, increased supervision during procedure-based scenarios generally results in improved patient-related and educational outcomes. As mentioned, our review revealed that enhancing attending-level presence in already-supervised activities has positive benefit; however, only one study examined the effect of introducing formalized supervision in previously unsupervised activities, such as the overnight or on-call period (the effects were not statistically significant).

The fact that, to date, only one study has examined the impact of enhanced nighttime supervision on patient- and education-related outcomes limits the ability to make valid conclusions regarding the specific recommendation of overnight presence of faculty. Based on the results of this review, there is not sufficient evidence to definitively support continuous, on-site, attending-level supervision for trainees; however, in the absence of such evidence, patient safety concerns should drive current decisions regarding overnight supervision practices. Additional rigorous studies will need to examine the impact of such a policy change before mandating a costly intervention with questionable efficacy.

The lack of evidence for mandating enhanced supervision must be considered in the context of the current ACGME recommendations, which specify the need for access to a “qualified supervisor.” A “qualified” supervisor is not necessarily someone who has fully completed his or her postgraduate training; it may be a more senior resident colleague—someone who has only marginally more training than the supervisee. A small number of studies did demonstrate that the number of changes or corrections made to a patient assessment by a faculty physician depended on the postgraduate year of the trainee; more advanced trainees’ clinical assessments required fewer modifications. These findings suggest that supervision is not a one-size-fits-all approach and that the degree of faculty oversight depends on the trainee’s perceived and actual levels of skill. Objective criteria either to measure the readiness of a trainee to perform in the role of supervisor or to indicate the level of supervision required are lacking, as is education for supervisors of all levels. Again, this lack speaks to the limited work in assessing the effects of supervision and indicates directions for possible future endeavors, such as defining milestones to identify more senior trainees who may be prepared to supervise their more junior colleagues.

The absence of objective outcomes related to the quality of supervision limits the rigor with which the results of supervision can be explored—specifically, quantifying and qualifying the effects of clinical supervision on patient- and education-related outcomes. Although many investigators have performed observational studies of current strategies in an attempt to formulate a framework for analysis, no validated instrument is available either to quantify the quality of supervision (given or received) or to relate supervision to clinical outcomes. Many of the studies we reviewed relied on nonvalidated instruments or nonobjective measures to evaluate the impact of the intervention.

There are several limitations to our review including the small overall number of studies included. Of those studies included, many occurred at a single institution and, as mentioned, used a small sample size. In addition, few represented RCTs. The majority either used exclusively unvalidated survey data or lacked objective assessments of the clinical supervision provided. Publication bias could also influence the results of our review. We did not conduct a search of abstracts or papers indexed within EMBASE, which may have decreased the number of studies included in the final review. Finally, we chose to
focus specifically on graduate medical education and did not examine the impact or use of supervisors in other health professions such as nursing.

Despite these limitations, this work provides a review of the existing literature and helps to shape the current understanding of the effect of supervision on patient- and education-related outcomes. It suggests that enhanced attending supervision of care provided by residents in a variety of clinical settings may improve both patient- and education-related outcomes. However, there is insufficient evidence at present to mandate continuous, on-site attending supervision as previously recommended by the IOM. Importantly, this lack of objective evidence should not cloud the ultimate goal, which is the delivery of safe, quality patient care. Future research should be conducted to establish the effects of enhanced on-site supervision, at any level, particularly during nights and weekends. Lastly, objective measures of the quality and extent of supervision provided to trainees are needed to improve the rigor and further the work in this area of research.

Acknowledgments: The authors wish to thank Dr. Kathryn Fletcher from the Medical College of Wisconsin for her valuable insights regarding systematic review, Dr. Darcy Reed from Mayo Clinic College of Medicine for the use of the Medical Education Research Study Quality Instrument (MERSQI), and Ms. Christa M. Moshiedel, biomedical bibliographer at the John Crerar Library at the University of Chicago, for technical assistance with our search strategy.

Funding/Support: This study was funded by the University of Chicago Department of Medicine.

Other disclosures: None.

Ethical approval: Not applicable.

Disclaimer: None.

Previous presentations: Portions of this review were presented at the 2010 University of Chicago Medical Education Day; 2010 Society of General Internal Medicine Meeting, Minneapolis, Minnesota; April 29, 2010.

Dr. Farman is assistant professor, Department of Medicine, Pritzker School of Medicine, University of Chicago, Chicago, Illinois.

Dr. Petty is a third-year internal medicine and pediatrics resident, Department of Medicine, Pritzker School of Medicine, University of Chicago, Chicago, Illinois.

Dr. Georgitis was, at the time of this research, a resident, Department of Medicine, Pritzker School of Medicine, University of Chicago, Chicago, Illinois, and is now a fellow, Pulmonary and Critical Care, Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania.

Dr. Martin is a fellow in hospital medicine, Department of Medicine, Pritzker School of Medicine, University of Chicago, Chicago, Illinois.

Ms. Chiu is an undergraduate student studying economics, University of Michigan, Ann Arbor, Michigan.

Ms. Prochaska is a project manager, Department of Medicine, Pritzker School of Medicine, University of Chicago, Chicago, Illinois.

Dr. Arora is associate professor, Department of Medicine, Pritzker School of Medicine, University of Chicago, Chicago, Illinois.

References


30. Patel SP, Gauger PG, Brown DL, Englesbe MJ, Cederna PS. Resident participation does...
## Appendix 1

### Summary Including Some Findings of Reviewed Studies That Examined the Effect of Clinical Supervision on Trainee and Patient Outcomes, Organized by Specialty and Quality Score

| Study, year | Design and participants | Type of supervision | Supervision intervention | Education-related outcomes | Patient-care-related outcomes | Study quality (MERSQI score)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Psychiatry</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jaynes et al, 1979</td>
<td>Randomized controlled trial (RCT) of 78 psychiatry residents at an academic medical center (AMC)</td>
<td>Direct versus oversight</td>
<td>Attending psychiatrist directly present in resident-led outpatient clinic visit versus review of outpatient cases after evaluation by trainee</td>
<td>Not applicable (N/A)</td>
<td>• Directly supervised group had fewer patients judged (by attendings) to be severely ill compared with case-review group (4% versus 28%, (P &lt; .05))&lt;sup&gt;16&lt;/sup&gt; • Directly supervised group demonstrated the following trends compared with case-review group (chi-square test level of significance ranged from (P &lt; .06) to (P &lt; .13)) — Improved patient status (44% versus 28%) — Greater success of patients continuing treatment (64% versus 28%)</td>
<td>16</td>
</tr>
<tr>
<td>Kass et al, 1984</td>
<td>RCT of 83 patient encounters staffed by psychiatry residents at an AMC</td>
<td>Oversight</td>
<td>Outpatient cases reviewed by attending psychiatrist versus peer-reviewed outpatient cases versus no supervision</td>
<td>Trainees reported both senior and peer supervision to be positive educational experiences</td>
<td>• Supervised residents reported higher ratings of patient outcome (including continuing therapy) on a 7-point scale (1 = patient much better, 7 = patient much worse) than did both peer-reviewed group and control group: attending supervision (2.96, standard deviation [SD] 0.98) versus peer (3.32, SD 0.81), control (3.59, SD 0.88); 1-way ANOVA* (F = 3.35, P &lt; .05)</td>
<td>14</td>
</tr>
<tr>
<td>Jin et al, 2000</td>
<td>Retrospective cohort study of 104 psychiatry residents at an AMC</td>
<td>Direct versus oversight</td>
<td>Comparison of residents with training and with direct supervision in consultation–liaison psychiatry ordering constant observation (CO) for patients versus control group (no training, and standard oversight)</td>
<td>N/A</td>
<td>• Controls requested significantly higher percentage of CO orders (44.1%) compared with supervised residents (15.4%) (\chi^2 = 12.1, df = 1, P &lt; .001) • Supervised residents ordered CO less frequently during regular hours (2.8%) when attending physicians were available compared with after hours when no attending presence was available (22.1%) (\chi^2 = 6.72, df = 1, P &lt; .01)</td>
<td>12.5</td>
</tr>
</tbody>
</table>

(Appendix continues)
### Appendix, continued

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Design and participants</th>
<th>Type of supervision</th>
<th>Supervision intervention</th>
<th>Education-related outcomes</th>
<th>Patient-care-related outcomes</th>
<th>Study quality (MERSQI score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finlayson et al, 1979&lt;sup&gt;17&lt;/sup&gt;</td>
<td>Pre–post study of 24 psychiatry residents at an AMC who had at least 6 months of inpatient training</td>
<td>Direct</td>
<td>Implementation of attending-level supervision overnight with shifts divided between all residents</td>
<td>N/A</td>
<td>N/A</td>
<td>9.5</td>
</tr>
<tr>
<td>Stein et al, 1975&lt;sup&gt;21&lt;/sup&gt;</td>
<td>Prospective cohort study of 16 psychiatry residents at an AMC</td>
<td>Direct versus oversight</td>
<td>Resident interviewed patient alone and presented case to attending at later time versus attending present at interview but did not participate; attending and resident then rated psychopathology, motivation, insight, and prognosis of patient independently on accepted rating scales</td>
<td>N/A</td>
<td>Attendings who were present for interview rated patient acuity significantly higher than did attendings who heard resident presentation only</td>
<td>9.5</td>
</tr>
<tr>
<td>Itani et al, 2005&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Retrospective cohort of 39,577 surgical cases at an AMC-affiliated Veterans’ Administration hospital</td>
<td>Direct versus indirect, with supervision immediately available</td>
<td>Varying levels of supervision: Level 3 = attending not present in operating room (OR) suite but immediately available; Level 2 = attending in the OR but not scrubbed; Level 1 = attending in the OR assisting; Level 0 = attending performed surgery</td>
<td>N/A</td>
<td>Overall complication rate lower for patients operated on under Level 3 supervision than other levels (Level 3 0.13 versus other levels 0.16, ( P &lt; .001 ))</td>
<td>14.5</td>
</tr>
</tbody>
</table>

### Emergency medicine, surgery and anesthesia

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Design and participants</th>
<th>Type of supervision</th>
<th>Supervision intervention</th>
<th>Education-related outcomes</th>
<th>Patient-care-related outcomes</th>
<th>Study quality (MERSQI score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finlayson et al, 1979&lt;sup&gt;17&lt;/sup&gt;</td>
<td>Pre–post study of 24 psychiatry residents at an AMC who had at least 6 months of inpatient training</td>
<td>Direct</td>
<td>Implementation of attending-level supervision overnight with shifts divided between all residents</td>
<td>N/A</td>
<td>N/A</td>
<td>9.5</td>
</tr>
<tr>
<td>Stein et al, 1975&lt;sup&gt;21&lt;/sup&gt;</td>
<td>Prospective cohort study of 16 psychiatry residents at an AMC</td>
<td>Direct versus oversight</td>
<td>Resident interviewed patient alone and presented case to attending at later time versus attending present at interview but did not participate; attending and resident then rated psychopathology, motivation, insight, and prognosis of patient independently on accepted rating scales</td>
<td>N/A</td>
<td>Attendings who were present for interview rated patient acuity significantly higher than did attendings who heard resident presentation only</td>
<td>9.5</td>
</tr>
<tr>
<td>Itani et al, 2005&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Retrospective cohort of 39,577 surgical cases at an AMC-affiliated Veterans’ Administration hospital</td>
<td>Direct versus indirect, with supervision immediately available</td>
<td>Varying levels of supervision: Level 3 = attending not present in operating room (OR) suite but immediately available; Level 2 = attending in the OR but not scrubbed; Level 1 = attending in the OR assisting; Level 0 = attending performed surgery</td>
<td>N/A</td>
<td>Overall complication rate lower for patients operated on under Level 3 supervision than other levels (Level 3 0.13 versus other levels 0.16, ( P &lt; .001 ))</td>
<td>14.5</td>
</tr>
</tbody>
</table>

However, more patients in the Level 3 supervision group required emergency surgery (12.84%) compared with the other levels of supervision (6.79%, \( P < .001 \)).

No significant difference in mortality between patient groups with and without attending presence, when adjusting for patient factors.

(Appendix continues)
### Study, year  
### Design and participants  
### Type of supervision  
### Supervision intervention  
### Education-related outcomes  
### Patient-care-related outcomes  
### Study quality (MERSQI score)  

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Design and participants</th>
<th>Type of supervision</th>
<th>Supervision intervention</th>
<th>Education-related outcomes</th>
<th>Patient-care-related outcomes</th>
<th>Study quality</th>
</tr>
</thead>
</table>
| Fallon et al, 1993 | Retrospective medical record audit of 4,417 cases at an AMC | Direct versus indirect, with supervision immediately available | Faculty supervision was rated on a nominal scale (grade 1 attending scrubbed to grade 4, attending not in house but available) for all procedures in the OR and for all operations or resuscitations on the trauma service, then level of supervision data were matched to outcome variables (i.e., death or complication) | N/A | • Attendings were either scrubbed or present in the OR for 91.8% of cases  
• Greater attending presence associated with significantly lower mortality and complication rates  
  — Complications 3.1% (attending present) versus 9.25% (attending absent), *P* < .0005  
  — Death 0.72% (attending present) versus 9.5% (attending absent), *P* < .0005 | 14 |
| Rogers et al, 1993 | Retrospective cohort study comparing tertiary care at an AMC with in-house, board-certified surgeon (IHBCS) versus one with a resident in house (RH) | Direct | Comparison of outcomes and provider-related complications for patients treated at 2 trauma centers, 1 with IHBCS and 1 with RH | N/A | • No significant difference was noted between patient groups at either center  
• Mortality rates were not significantly different between the 2 centers, although the mortality at the RH trauma center was higher than predicted by a preceding study  
• Provider-related complications were significantly greater at the RH center than at the IHBCS center  
  — Chance of error in judgment (21% versus 2.8%; OR 9.5, *P* < .05)  
  — Chance of error in technique (16% versus 7.8%; OR 2.2, *P* < .05) | 13 |
| Patel et al, 2010 | Retrospective chart review of bilateral reduction mammoplasty at an AMC | Indirect, with direct supervision immediately available | 259 patients underwent bilateral reduction mammoplasty; the right breast was operated on independently by the attending, and the left breast was operated on by the resident alone with attending supervision; charts were reviewed for major complications | N/A | • 23 of 295 patients (7.8%) experienced complications  
• 19 patients had complications in only 1 breast, and 4 had complications in both breasts  
• No significant difference was noted in major complication rates between resident side and attending side (4.7% vs 4.4%, *P* = .99) | 12.5 |
## Appendix, continued

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Design and participants</th>
<th>Type of supervision</th>
<th>Supervision intervention</th>
<th>Education-related outcomes</th>
<th>Patient-care-related outcomes</th>
<th>Study quality (MERSQI score†16)</th>
</tr>
</thead>
</table>
| Schmidt et al, 200836 | Pre–post study of anesthesia resident intubation of 322 patients requiring emergent intubation outside of the OR at an AMC | Indirect, with direct supervision immediately available | Attending physician assigned to attend emergency intubations at predetermined time periods; 115 patients were intubated by residents in the presence of attendings, and 207 patients were intubated by anesthesia residents alone | N/A                         | • Attending supervision associated with significant decrease in complications (e.g., esophageal intubation, traumatic intubation, aspiration, dental injury, endobronchial intubation) (6.1% versus 21.7%, $P < .0001$)  
• Intubations with attending physicians present had a significantly higher use of muscle relaxants (46% versus 17%; $P < .001$)  
• No difference in ventilator-free days, 28-day mortality rates, patient disposition, intubation attempts  
• No differences were noted by level of postgraduate training | 12.5 |
| Sox et al, 199810 | Prospective cohort study of 1,120 emergency room patient encounters, during which residents were supervised by attendings at an AMC | Direct | Assessment of the effect of direct attending physician supervision on resident physicians on the quality of care in emergency departments | N/A                         | Directly supervised residents had significantly higher adjusted mean percentage compliance with process-of-care guidelines, as measured by complaint-specific (e.g., chest pain) data forms (64%), than other residents (55%, $P < .001$) | 12.5 |
| Ereso et al, 201029 | Case–control evaluation of the effect of remote surgical telemonitoring§ | Indirect, with direct supervision available | Comparison of mean performance outcomes of general surgery residents performing 3 procedures, first unproctored, then with teleproctoring by surgical subspecialist; resident performance evaluated by appropriate subspecialty proctor on 5-point rating scale, with higher score representing better performance | N/A                         | • All performance scores were superior in all scenarios where residents were teleproctored ($P < .001$)  
• Mean performance for individual metrics such as instrument handling, procedure time, and general knowledge of anatomy were significantly better when teleproctored ($P < .001$) | 12 |
## Appendix, continued

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Design and participants</th>
<th>Type of supervision</th>
<th>Supervision intervention</th>
<th>Education-related outcomes</th>
<th>Patient-care-related outcomes</th>
<th>Study quality (MERSQI score)</th>
</tr>
</thead>
</table>
| Holliman et al, 1995<sup>22</sup> | Pre–post study of 3,133 visits to an AMC emergency room | Oversight | Cases reviewed by attending; residents (excluding all first-year residents) rated on a data score sheet that reflected minor and major changes to the resident’s plan of care | N/A | • Attending physicians made 153 major changes (e.g., missed X-ray finding) and 353 minor changes (e.g., charting deficiency)  
• 9% of patients had more than 1 change recorded per encounter  
• No change in care for 590 patients (59%)  
• 17 major changes judged to be life- or limb threatening | 12 |
| Akingba et al, 2008<sup>31</sup> | Retrospective cohort study of 424 hysterectomies performed at an AMC | Direct | 265 surgeries involving a resident (assisting) versus 159 surgeries performed by an attending physician alone | N/A | • Mean operative time was noted to be 13 minutes longer in the supervised resident group, but this was not clinically or statistically significant  
• Statistically significantly higher incidence of seromas (2.5% versus 0.0%, P = 0.2) and “other miscellaneous adverse outcomes” (5.0% versus 0.8%, P = .007) in the teaching cases  
• Teaching cases were likely to be performed by attendings with fewer years of experience | 11.5 |
| Kroboth et al, 1982<sup>23</sup> | Pre–post study of emergency medicine residents at AMC | Direct and indirect, with direct supervision immediately available | Institution of supervision (of junior residents) by general medicine faculty and emergency medicine faculty versus previous supervision by chief residents | N/A | • After institution of faculty supervision  
• 20% increase in hospital admissions (17.5% versus 20.5%, P < .01)  
• Decreased utilization of observation unit (2.8% versus 2.4%, P < .01)  
• Fewer patients leaving without being examined (4.3% versus 2.2%, P < .02) and  
• No change in number of patients leaving against medical advice or utilization of routine lab tests | 10.5 |
| Nguyen et al, 2008<sup>26</sup> | Prospective cohort study of patients undergoing office-based vasectomy | Indirect, with direct supervision immediately available | Comparison of patient pain scores of indirectly supervised residents performing office-based vasectomy versus staff surgeon alone | N/A | No significant differences noted between patients’ pain scores for those operated on by supervised trainee, after adjusting for age or side of surgery, versus those operated on by attending-level providers (1.01 versus 1.76, P = 0.10) | 9.5 |
### Appendix, continued

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Design and participants</th>
<th>Type of supervision</th>
<th>Supervision intervention</th>
<th>Education-related outcomes</th>
<th>Patient-care-related outcomes</th>
<th>Study quality (MERSQI score)</th>
</tr>
</thead>
</table>
| Sacchetti et al, 1992 | Prospective study of 408 patient encounters in an AMC emergency department                | Direct              | Initial patient plan of care made by resident, reviewed with attending physician, who personally interviewed and examined the same patient and then modified the resident’s plan and recorded changes | N/A                       | • 16 patients (4%) required major changes to management plan (e.g., missed fracture, disposition changes, missed diagnoses)  
• 134 patients (33%) required minor modifications (not explicitly defined in manuscript)  
• No changes necessary for the remaining 358 patients (63%) | 8.5                          |
| Blumberg et al, 1995  | Pre–post study of 20 residents and attendings at a community teaching hospital            | Oversight           | Institution of faculty physician review of care protocols for high-risk patients (e.g., patients with chest pain, congestive heart failure, intoxication); feedback provided to resident physicians by staff physician on residents’ compliance with protocols | N/A                       | • Resident compliance with protocols was more consistent after initiation of faculty review of medical records (mean compliance of 95% [SD 3%] versus 92% [SD 8%]) | 13                           |
| Everett 1985         | RCT of 2,492 patient visits, seen by residents at AMC                                    | Oversight           | For 9-month period, all laboratory ordering initiated by residents required approval by attending physician; residents were randomized according to receiving training on test utilization | N/A                       | Regardless of residents’ exposure to educational intervention on test ordering, test utilization styles of attending physicians did not have a significant effect on test utilization styles of residents | 12.5                         |
| Feinglass et al, 1991| Retrospective paired chart review and survey data of internal medicine residents an AMC   | Oversight           | Examination of relationship between use of hospital resources and resident-perceived level of attending supervision; autonomy ratings were established by surveying residents about their perception of their decision-making independence under each attending on a scale of 1 to 5; residents’ ratings of attendings were compared | Mean autonomy rating for the 14 raters (65) whose responses demonstrated the highest level of agreement was 2.83 (1–5 scale) | • Mean autonomy ratings were highly significant predictors of resource use and length of stay ($P < .0001$), with greater perceived autonomy predicting significantly lower resource use and shorter length of stay  
• A 1-unit increase in attending autonomy rating was associated with $311 saving in resource use and 0.38-day decrease in stay length | 11.5                         |

---

*(Appendix continues)*
### Appendix, continued

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Design and participants</th>
<th>Type of supervision</th>
<th>Supervision intervention</th>
<th>Education-related outcomes</th>
<th>Patient-care-related outcomes</th>
<th>Study quality (MERSQI score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Osborn et al, 1993&lt;sup&gt;37&lt;/sup&gt;</td>
<td>Prospective cohort with pre–post assessment of 54 pediatrics residents</td>
<td>Direct versus oversight</td>
<td>Residents self-selected into 1 of 3 tracks for continuity clinic experience (university based; private practice; or publicly funded, community-based clinics); assessment of performance on in-training examination, rating of supervision provided, and overall clinic experience</td>
<td>• Number of visits per half-day varied by site: 4.2 (AMC) versus 6 (private) versus 3.5 (community-based); ( P &lt; .0001 )</td>
<td></td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Supervision of residents at AMC versus community-based practice were similar — Attendings consulted on 69% and 60% of visits, respectively</td>
<td>• Attendings infrequently observed the residents directly as the latter provided patient care</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• At private offices, residents consulted less frequently with their attending (38% of visits) but were directly observed providing care at least once per week</td>
<td>• No change in performance on the pre–post in-training exam when comparison performed by site</td>
<td></td>
</tr>
<tr>
<td>Gennis and Gennis, 1993&lt;sup&gt;34&lt;/sup&gt;</td>
<td>Pre–post study of 19 residents in 74 patient interactions at an AMC</td>
<td>Direct</td>
<td>Attending physicians participated in the outpatient clinic visits with residents</td>
<td>• After personally seeing the patient, attendings rated the quality of housestaff history and physical exam lower on a 5-point scale than before seeing the patient — History rating pre 3.57 versus post 3.29; ( P &lt; .05 ) — Physical rating pre 3.64 versus post 3.40; ( P &lt; .05 )</td>
<td>• At least 1 change in management in 58.1% of cases</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Changes in plan were independent of level of training of residents</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>(Appendix continues)</sup>
<table>
<thead>
<tr>
<th>Study, year</th>
<th>Design and participants</th>
<th>Type of supervision*8,9</th>
<th>Supervision intervention</th>
<th>Education-related outcomes</th>
<th>Patient-care-related outcomes</th>
<th>Study quality (MERSQI score†16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Velhamos et al, 200139 | Prospective comparison of faculty versus preliminary resident read | Oversight | Examination of 6 months of trauma-related radiology interpretations, comparing preliminary reads by resident versus reads by covering attending radiologist around the clock (CARAC); a discrepancy was defined as a difference resulting in a significant change in management after the CARAC read | N/A | • In 11% (42/383) of patients, a discrepancy was noted between reads  
• Most discrepancies were found on abdominal CT scans, which were associated with longer hospital stay and higher mortality  
• The lower the postgraduate year of the trainee performing preliminary read, the higher the likelihood of discrepancy | 12.5 |
| Albano et al, 200138 | Retrospective comparison of AMC resident preliminary read to attending final read of suspected appendicitis | Oversight | Resident preliminary review and attending final review of CT and surgical specimens; faculty members were aware of the preliminary read; if a discrepancy was noted, scans were reexamined by a third independent faculty member | N/A | • Final faculty interpretation and the preliminary resident read were identical in 96/103 patients (93%; 95% CI: 87.8–97.2)  
• Only 1 negative preliminary read (by a resident) was changed to a positive by the attending radiologist | 10 |

* Type of supervision defined as direct, which requires the physical presence of the supervisor with the patient and trainee; indirect, with direct supervision immediately available (immediate ability of supervisor to be at the bedside/site of care) or indirect, with direct supervision available (supervisor available via phone or pager and able to be physically present); or oversight, in which a supervising physician reviews the care after care has been delivered.

† MERSQI (Medical Education Research Study Quality Instrument) rates quality of research studies on a 16- to 18-point scale for study design, type of data, sampling, instrument validation, data analysis, etc.18

‡ ANOVA indicates analysis of variance.

§ Teleproctoring involved live audio and video with continuous real-time updating of subspecialist while the resident performed 1 of 3 procedures on an animal model.