

Effect of a Physical Examination Teaching Program on the Behavior of Medical Residents

Graham T. McMahon, MD MMSc,^{1,2} Ovidiu Marina, MEng,^{1,2} Patricia A. Kritek, MD,^{1,2} Joel T. Katz, MD^{1,2}

¹Department of Medicine, Brigham & Women's Hospital, Boston, Mass, USA; ²Harvard Medical School, Boston, Mass, USA.

CONTEXT: The reliance on physical examination as a diagnostic aid is in decline.

OBJECTIVE: To determine whether an educational program can increase the use of physical examination by medical residents.

DESIGN AND PARTICIPANTS: A series of educational workshops were provided to 47 second- and third-year medical residents at a large academic teaching hospital.

MEASUREMENTS: Interns and students reported the frequency and depth of clinical examination performance on morning rounds by their residents before and up to six months after the workshops. Behavior before and after the workshops was compared using a mixed model.

RESULTS: A total of 374 reports were returned (77% response). After adjusting for the type of service and observer, there was a statistically significant 23% increase ($P=.02$) in the performance of physical examination among residents who attended the course. Residents significantly increased the fraction of patients they examined on rounds (absolute increase 11%, $P=.002$) but did not increase the depth of their examination. The change was greatest on general medical teams, among whom the performance of physical examination had been least frequent. Teaching and feedback events on medicine teams by residents to their interns (2.8 and 1.1 events per 2 weeks, respectively) and medical students (5.9 and 2.8 events per 2 weeks, respectively) remained infrequent.

CONCLUSIONS: A skills improvement program can significantly increase the frequency of physical examination, but teaching and feedback events remain sporadic and infrequent.

KEY WORDS: medical education; physical examination; clinical skills.
DOI: 10.1111/j.1525-1497.2005.0159.x
J GEN INTERN MED 2005; 20:710-714.

The physical examination is a critical component of medical practice.¹⁻⁶ However, residents demonstrate considerable deficiencies in skills deemed important by program directors and during residency, and these skills do not typically improve beyond those of senior medical students.⁷⁻¹⁰

Despite the identification of these deficiencies, the teaching of physical examination skills is becoming less frequent. While in the early 1960s 75% of clinical teaching was at the bedside, by 1978, the figure had dropped to 16%.¹¹ As a result, physical examination skills have fallen into disuse among both trainees and attending physicians, leading to a spiral of decreased familiarity with signs, a decreasing sense of the utility of examination skills, and an increased reliance on laboratory and imaging tests.^{12,13}

Residents spend up to 3 hours a day teaching, and medical students and junior trainees receive large amounts of their teaching from residents.¹⁴⁻¹⁶ Residents have the greatest influence on how well medical students learn during their clerk-

ship, as measured by end-of-clerkship and standardized examinations.^{17,18} Residents therefore are a natural target for a focused educational effort to reverse this decline in the utilization of the clinical examination.

We hypothesized that a problem-based course provided to junior (postgraduate year (PGY)-2) and senior (PGY-3) medical residents would increase the frequency and depth of physical examination performance by medical residents on rounds, and would increase their teaching of these skills to more junior trainees.

DESIGN AND PARTICIPANTS

Recruitment

Second- and third-year medical residents in the Internal Medicine residency program at Brigham & Women's Hospital were recruited through housestaff announcements, and all participants provided written informed consent. The Institutional Review Boards of Brigham and Women's Hospital and Harvard Medical School approved the study.

Educational Program

Eight 90-minute small-group late afternoon sessions were held in November and December 2003. The content of the sessions was posted in advance and residents were free to attend up to 4 sessions of their choosing. During the sessions, residents rotated in groups of 3 to 6 around 4 stations, each supervised by a senior clinician, practicing the detection of signs on volunteer patients (appendix I, available online). The learning objectives were to refamiliarize the residents with physical findings for each system and to provide teaching strategies to facilitate learning among more junior trainees. Residents practiced these skills under observation by teaching their colleagues in the group. The content of the teaching sessions was determined by an *a priori* needs assessment sent to all residents in the program. A specific faculty development program prepared the senior clinicians.

MEASUREMENTS

Survey Administration and Timing

Participants on in-patient clinical rotations at Brigham and Women's and Faulkner Hospitals as well as their interns and students ("observers") were surveyed before and periodically in the 6 months after the educational program. Residents provided self-reports, while medical students and interns provided observed reports. Surveys (appendix II, available online) were accessed via a web-link posted into individual email accounts and managed online (surveymonkey.com,

Address correspondence and requests for reprints to: Dr. McMahon: Department of Medicine, Brigham & Women's Hospital, 221 Longwood Ave., RFB2, Boston, MA 02115 (e-mail: gmcMahon@partners.org).

Received for publication March 21, 2005
and in revised form March 21, 2005
Accepted for publication March 21, 2005

Madison, Wis). Interns (PGY-1) and students were surveyed when either their own or their residents' rotation ended, as long as they had spent at least 6 days with that resident. Residents were surveyed only at the end of each rotation. Services were divided into 3 categories: intensive care units (ICUs), general medical, and medical subspecialty (cardiology, oncology, bone marrow transplant) services.

Physical examination practice was evaluated by asking respondents what fraction of patients, and in what depth, residents examined patients on rounds with their team. Depth of examination was evaluated by asking respondents to select the percentage of patients who were examined in a given depth, ranging from 0 to 4+ systems. We defined the depth–frequency index (DFI) as the product of the percentage of patients on the team being examined and the average number of systems examined on each patient. The DFI has a maximum of 4 (100% of patients having 4 or more systems examined on rounds), and a minimum of zero.

The surveys also invited self-reports of skill among residents, interns, and medical students in examination techniques, teaching behavior, source of teaching received, and the value placed on the utility of physical examination skills. They also inquired about self-perceived skill at physical examination and the value placed on its role in the practice of medicine. The non-numeric survey questions used a 5-point Likert categorical scale with response options ranging from 1 = “strongly disagree” to 5 = “strongly agree,” resulting in an “agreement score” (AS).

Statistical Analyses

Analysis of the change in the behavior of residents in their performance of physical examination was performed separately for observers' reports and for residents' self-reports. For observers' reports, the mean change in behavior was estimated under a mixed effects model with fixed effects for service (intensive care, specialty, general medicine) and time (before or after the program) and random effects for resident and observer. These random effects were used to account for correlations between observations made on the same resident and between observations made by the same observer. Because each resident or observer could appear in different service types during the course of the study, we fitted different covariance structures, allowing the variance of the random effects to vary by service type; the final covariance model was chosen based on Akaike and Bayesian information criteria. This analysis was performed for each of the 3 behavior outcomes (DFI, fraction of patients examined, and depth examined). We also evaluated the behavior change within each service type, using a mixed-effects model with fixed effect for time and random effects for resident and observer. The residents' self-reports were analyzed via similar mixed-effects models but without the observer random effect.

The secondary outcomes include agreement scores and teaching and feedback event frequency. Because the agreement scores in 5-point Likert scale were ordinal in nature and the distribution of the event frequencies was highly skewed, the between-group differences in these outcomes were assessed using the nonparametric Wilcoxon rank-sum and Kruskal–Wallis tests. In addition, the signed rank test was used to evaluate whether the number of teaching and feedback events changed after the teaching program. Tukey's correction

was made for multiple pairwise comparisons between the groups. All tests were performed using 2-sided significance levels. Significance was accepted for $P < .05$. The analyses were performed using SAS 9.0 (SAS Institute Inc., Cary, NC).

RESULTS

Teaching sessions were provided to 47 residents. Four hundred eighty-four (484) comprehensive surveys were distributed, and 374 (77%) were returned. The response rates for residents, interns, and medical students were 89%, 69%, and 82%, respectively. The median number of patients on each ICU, specialty, and general medicine team remained constant at 11 (interquartile range/IQR 8 to 11), 8 (IQR 8 to 14), and 13 (IQR 11 to 14) throughout the study.

Examination of Patients

The primary outcome was the observed change in the residents' performance of physical examination on rounds after the program. Table 1 illustrates the overall increase in DFI (accounted for almost entirely by the increase in frequency of physical examination) after the teaching program by service, an increase that persisted through the 6-month follow-up period.

Residents' self-reports corroborate the changes observed by their interns and medical students; the correlation of DFI scores for a given resident between medical students and interns was 0.54 ($P = .002$), 0.66 ($P < .001$) between medical students and residents, and ($P < .001$) between interns and residents 0.81. Figure 1 demonstrates the relationship between the observed and the residents' self-reported behavior by service. There were no significant differences between junior and senior medical residents in the baseline DFI, percent of patients examined, or depth of examination on rounds. Both junior and senior residents increased their DFI by similar amounts over the course of the trial.

Skill in Examination

Residents perceived themselves to have a lower average physical examination skill than either medical students or interns, with only 23% of residents considering themselves “skilled.” By contrast, 53% of medical students and 44% of interns rate themselves as skilled ($P = .02$ for differences between groups).

After the teaching program, residents reported that their physical examination skills had improved (AS 3.70 before vs 4.03 after the intervention, $P = .05$); medical students and interns (who had not received the teaching intervention) reported no change in their own physical examination skills over the same time period. Additionally, over the 6-month follow-up, participating residents reported a significant increase in the frequency with which they used physical examination for diagnosis (AS 3.02 vs 3.48, $P = .01$). The majority (57%) of participants felt that the program was directly responsible for any changes in their physical examination performance.

Teaching Physical Examination Skills. The residents reported a significant improvement in their confidence in teaching physical examination skills (AS 3.2 vs 3.9, $P < .001$). In contrast, interns, who were not subject to the teaching program, showed no improvements in their confidence. Greater confidence in teaching correlated strongly with having a plan to

Table 1. Observer Reports* on Changes in the Depth-Frequency Index, Fraction Examined, and Depth of Examination by their Residents

Service Type	Before Program		After Program		Estimated Change [‡] (SE)	P Value
	N [†]	Mean (SD)	N [†]	Mean (SD)		
<i>(a) Depth-frequency index</i>						
All services	32	1.37 (1.05)	166	1.56 (1.16)	0.32 (0.14)	0.02
Intensive care	4	3.01 (1.39)	24	3.66 (0.48)	0.43 (0.25)	0.10
Specialty	17	1.38 (0.82)	58	1.46 (0.95)	0.08 (0.26)	0.76
General medicine	11	0.77 (0.51)	84	1.04 (0.67)	0.24 (0.23)	0.29
<i>(b) Fraction examined</i>						
All services	32	0.47 (0.27)	166	0.55 (0.30)	0.11 (0.04)	0.002
Intensive care	4	0.86 (0.28)	24	1.00 (0.00)	— [§]	— [§]
Specialty	17	0.47 (0.25)	58	0.56 (0.29)	0.08 (0.08)	0.33
General medicine	11	0.32 (0.15)	84	0.42 (0.22)	0.09 (0.07)	0.22
<i>(c) Depth examined</i>						
All services	32	2.74 (0.74)	166	2.61 (0.86)	0.05 (0.13)	0.68
Intensive care	4	3.33 (0.83)	24	3.66 (0.48)	0.06 (0.20)	0.76
Specialty	17	2.85 (0.66)	58	2.52 (0.84)	-0.23 (0.23)	0.33
General medicine	11	2.34 (0.67)	84	2.37 (0.75)	-0.00 (0.26)	0.99

*There were a total of 200 observer reports made by 98 observers on 42 residents; 32 of the observer reports were before the program (made by 29 observers on 20 residents) and 168 were after the program, made by 82 observers on 39 residents.

[†]N denotes the number of observer reports included in the analysis. Because of missing values, N may be smaller than the total number of observer reports.

[‡]Adjusted for correlations between observations from the same observer and correlations between observations made on the same resident. For the analysis on all services combined, service type was also included in the model.

[§]Analysis was not performed because there was no variability in the fraction of patients examined after the program; the fraction was 100% on all observer reports.

increase teaching ($P < .001$). The frequency of teaching of physical examination skills by residents to their interns increased (3.6 vs 4.0 episodes per 2 weeks, $P = .11$) during the study. Medical students reported receiving substantially more teaching sessions from residents than from interns or attendings after the teaching program (5.9 vs 2.7 vs 3.2 episodes per 2 weeks, $P < .001$). For interns, the ICU rotation represented a significant opportunity for learning about physical diagnosis (mean 3.6 vs 2.8 teaching events per 2 weeks in the ICU compared with the non-ICU rotations).

Feedback about Physical Examination Skills. Feedback events about physical examination skills were consistently less frequent than teaching events, at a ratio of 1:2. Medical students reported receiving an average of 2.8 feedback events per 2-week block (IQR 0 to 5) from their residents, higher than the rate at which residents report giving such feedback (mean

2.5, IQR 0 to 3). Interns provided significantly less feedback to students than did residents (1.4 vs 2.8 events per 2 weeks, respectively, $P < .001$), with 52% of students reporting that they had received no feedback from their interns in the previous 2 weeks. Interns noted that they received feedback on their physical examination skills from their residents only rarely, with a mean of 1.2 events per 2 weeks; 58% reported no feedback at all during the same time period. Medical students received most of their feedback from residents, and the balance was split between their intern and attending. Feedback between medical students regarding physical examination skills was rare.

Learning about Physical Examination Skills. Residents reported that they received most of their teaching from their team attending (38%) and conferences (34%), while interns reported that their resident (42%) and team attending (35%) provided the majority of teaching. Medical students received most of their teaching from their resident (49%), with the balance split between their intern (22%) and team attending (27%). A majority of medical students (92%) report that they enjoy learning about these skills and seek more teaching in this area.

All groups agreed that more emphasis should be placed on the teaching of physical examination. Residents were more likely than students ($P = .01$) or interns ($P < .001$) to make this assertion (AS 4.43 vs 4.13 vs 4.10, respectively). More than 90% of all groups agreed that there is sufficient evidence to justify the acquisition of physical examination skills, and that such skills could increase the quality of patient care.

CONCLUSIONS

Even in this era of burgeoning diagnostic technology, the physical examination remains among the physician's most accurate tools.¹ Although residents in this and other studies place great importance on physical examination, they demonstrate

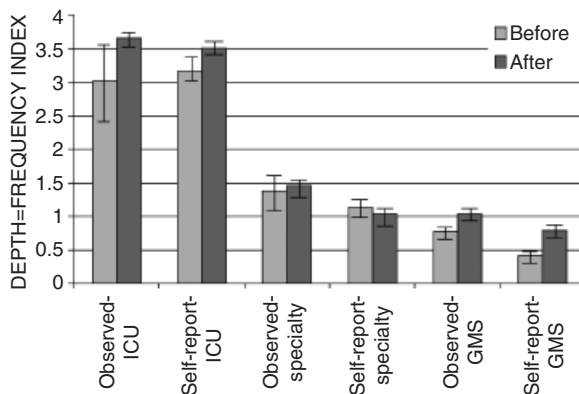


FIGURE 1. Observed and self-reported depth-frequency index at baseline and following the teaching program (mean ± SE). GMS, general medicine service

highly variable skill levels and performance frequency.^{7,9} Residents cannot learn these skills on their own,¹⁰ and facilitating skill development has been a persistent challenge to educators. Although other groups have demonstrated increased confidence in trainees after exposure to a training program,^{19–22} a year of monthly lectures on physical examination did not increase knowledge attainment among residents as assessed by written examination.²³ We have not found a prior study that measured plasticity in clinical examination behavior.

This study demonstrates that the frequency of physical examination, although not the depth of examination, can be increased through a skills-based educational program delivered to medical trainees. Important aspects of our program appeared to be the use of expert clinician-educators, the recruitment of appropriate and willing patients, using small groups for the learning exercises, and the creation of a safe environment where participating residents could practice their clinical and teaching skills and receive immediate feedback.

Residents are expected to be skilled in physical examination²⁴ but are rarely taught these skills.¹¹ Our finding that residents had poorer self-rating than medical students and interns is evidence that they perceive this deficiency with greater clarity as they progress through their residency. Despite the general enthusiasm we perceive for increased emphasis on the teaching of physical exam skills and the favorable response from patients noted in other studies,^{12,25–27} teaching and feedback about physical examination skills have become rare, and this is most apparent in non-ICU rotations. Even the daily performance of physical examination on clinical rounds has become the exception rather than the rule.

There are several putative reasons for the discrepancy between interest in and the performance of clinical examination. The most recent generation of physicians has been trained with the ready availability of minimally invasive diagnostic tools that do not depend on their own skills, but rather on the expertise of others. Increased throughput in hospitals has diminished the time attendings have for teaching trainees. As patients remain hospitalized for shorter periods, the opportunity for repeated and group examinations has also been reduced. Increasingly, the rotations of attendings, residents, and interns begin and end on different schedules, which can be disruptive to teaching. That interns report more teaching and feedback about physical diagnosis in the ICU may reflect the ongoing presence of daily bedside rounds with the attendings in these units.²⁸ Such rounds are rapidly becoming a thing of the past in general medicine services and are increasingly threatened by the implementation of shift work among residency programs.

As it was not possible to blind the participants in what was necessarily an open trial, we used the objectivity of multiple random observers to determine outcome. The exclusion of observers from informational mailings and the extended study duration makes substantial observer bias less likely, and the robust interobserver correlations suggest a relatively accurate measure of performance. Exposure to alternative learning opportunities does not appear to be a likely contributor to behavior change, as nonparticipants denied any change in their performance of physical examination in the follow-up survey. Our finding that junior and senior residents had similar baseline depth and frequency measures that increased by a similar amount over the course of the study suggests that the teaching

intervention played a larger role than seniority in the behavior change.

Further research in the area is warranted.^{1,29} In particular, research to characterize the link between competence in physical examination and patient outcomes would be valuable. Despite changes in the clinical environment, patient-centered teaching remains essential to the training of future clinicians. This study illustrates that the declining performance of physical examinations can be partially addressed through the introduction of a dedicated skills improvement program.

This program was funded by a Support for Excellence in Educational Development (SEED) grant from the department of Medicine at Brigham & Women's Hospital. GTM is supported by the National Institutes of Health (NHLBI) K30 Grant Number HL04095.

We appreciate the participation of the medical house-staff and teaching faculty of the Brigham & Women's Hospital, and are indebted to Mei-Chiung Shih, PhD, for helpful discussions and comments on the statistical analyses.

REFERENCES

1. **Sackett DL, Rennie D.** The science of the art of the clinical examination. *JAMA.* 1992;267:2650–2.
2. **Mangione S, Peitzman SJ.** Physical diagnosis in the 1990s. Art or artifact? *J Gen Intern Med.* 1996;11:490–3.
3. **Reilly BM.** Physical examination in the care of medical inpatients: an observational study. *Lancet.* 2003;362:1100–5.
4. **Sandler G.** The importance of the history in the medical clinic and the cost of unnecessary tests. *Am Heart J.* 1980;100:928–31.
5. **Peterson MC, Holbrook JH, Von Hales D, Smith NL, Staker LV.** Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses. *West J Med.* 1992;156:163–5.
6. **Crombie DL.** Diagnostic process. *J Coll Gen Pract.* 1963;54:579–89.
7. **Mangione S, Nieman LZ, Gracely E, Kaye D.** The teaching and practice of cardiac auscultation during internal medicine and cardiology training. A nationwide survey. *Ann Intern Med.* 1993;119:47–54.
8. **Schwinn CJ, Boehler ML, Foise R, Dunnington G, Markwell SJ.** Development of physical examination skills in a third-year surgical clerkship. *Am J Surg.* 2001;181:338–40.
9. **Oddone EZ, Waugh RA, Samsa G, Corey R, Feussner JR.** Teaching cardiovascular examination skills: results from a randomized controlled trial. *Am J Med.* 1993;95:389–96.
10. **Mangione S, Burdick WP, Peitzman SJ.** Physical diagnosis skills of physicians in training: a focused assessment. *Acad Emerg Med.* 1995;2:622–9.
11. **LaCombe MA.** On bedside teaching. *Ann Intern Med.* 1997;126:217–20.
12. **Thibault GE.** Bedside rounds revisited. *N Engl J Med.* 1997;336:1174–5.
13. **Obel J.** *Losing the Touch.* Washington, DC: The Washington Post; 2003:1.
14. **Brown RS.** House staff attitudes toward teaching. *J Med Educ.* 1970;45:156–9.
15. **Bing-You RG, Sproul MS.** Medical students' perceptions of themselves and residents as teachers. *Med Teach.* 1992;14:133–8.
16. **Morrison EH, Hafler JP.** Yesterday a learner, today a teacher too: residents as teachers in 2000. *Pediatrics.* 2000;105:238–41.
17. **Irby DM, Papadakis M.** Does good clinical teaching really make a difference? *Am J Med.* 2001;110:231–2.
18. **Roop SA, Pangaro L.** Effect of clinical teaching on student performance during a medicine clerkship. *Am J Med.* 2001;110:205–9.
19. **Wipf JE, Orlander JD, Anderson JJ.** The effect of a teaching skills course on interns' and students' evaluations of their resident—teachers. *Acad Med.* 1999;74:938–42.
20. **Spickard A III, Corbett EC Jr, Schorling JB.** Improving residents' teaching skills and attitudes toward teaching. *J Gen Intern Med.* 1996;11:475–80.

21. **Edwards JC, Kissling GE, Plauche WC, Marier RL.** Evaluation of a teaching skills improvement programme for residents. *Med Educ.* 1988; 22:514-7.
22. **Frattarelli LC, Kasuya R.** Implementation and evaluation of a training program to improve resident teaching skills. *Am J Obstet Gynecol.* 2003;189:670-3.
23. **Mangione S, Peitzman SJ, Gracely E, Nieman LZ.** Creation and assessment of a structured review course in physical diagnosis for medical residents. *J Gen Intern Med.* 1994;9:213-8.
24. **Accreditation Council for Graduate Medical Education.** Program Requirements for Residency Education in Internal Medicine 2004. http://www.acgme.org/acWebsite/RRC_140/140_prIndex.asp. Accessed on February 24, 2005.
25. **Lehmann LS, Brancati FL, Chen MC, Roter D, Dobs AS.** The effect of bedside case presentations on patients' perceptions of their medical care. *N Engl J Med.* 1997;336:1150-5.
26. **Simons RJ, Baily RG, Zelis R, Zwillich CW.** The physiologic and psychological effects of the bedside presentation. *N Engl J Med.* 1989;321:1273-5.
27. **Romano J.** Patients' attitudes and behavior in ward round teaching. *JAMA.* 1941;664-7.
28. **Nair BR, Coughlan JL, Hensley MJ.** Student and patient perspectives on bedside teaching. *Med Educ.* 1997;31:341-6.
29. **Landon BE, Normand SL, Blumenthal D, Daley J.** Physician clinical performance assessment: prospects and barriers. *JAMA.* 2003; 290:1183-9.

Supplementary Material

The following supplementary material is available for this article online:

- Appendix I. Workshop Organization and Structure.**
- Appendix II. Sample Survey Instrument.**