



Are Students Less Likely to Report Pertinent Negatives in Post-encounter Notes?

Anne Walling, MB, ChB; Scott E. Moser, MD; Gretchen Dickson, MD; Rosalee E. Zackula, MA

BACKGROUND AND OBJECTIVES: In completing post-encounter notes (PENs), students are believed to under-report about 30% of the important information obtained in the medical history. The resulting incomplete clinical notes can contribute to adverse patient care and medicolegal outcomes. We hypothesized that pertinent negative items would be more likely to be under-reported than positive items. We compared reporting rates for pertinent positive and negative items on two cases in a clinical skills assessment (CSA) taken by all 55 third-year students. Based on standardized patient (SP) checklists, students obtained 87% of both positive and negative items. Scoring of PENs found significant differences in the reporting rates for positive (75%) and negative (52%) items. These results were consistent for each case. Students appear to be able to elicit pertinent negative information from patients but, although these items may be crucial in the medical history, they are significantly more likely than positive items to be omitted from the clinical note.

(Fam Med 2012;44(1):22-5.)

The value of including post-encounter notes (PENs) in clinical assessments of students is well established.^{1,2} PENs provide unique information that is complementary to other data such as checklist scores completed by standardized patients (SPs) or assessments by observing faculty.³ Psychometric studies show high correlation between scores in data gathering (obtaining key items by history and physical examination) and written communication (PENs).^{4,5} Despite the clinical and educational significance, student notes consistently fail to report all the pertinent information obtained during clinical encounters, even in high-stakes examinations.⁶⁻⁸ One study found that students failed to report 32% of 13 key elements of the

medical history in one case during a clinical skills examination (CSA).⁷ On a study of four CSA cases, we found that students failed to report 27% of key history items.⁸

Few studies have commented on which items are most likely to be omitted from clinical notes by students or clinicians.^{7,9,10} In our previous study, we noted that pertinent negative items appeared less likely to be reported than other key items, but the study included too few pertinent negative items for a valid comparison.⁸ We therefore used data from a subsequent formative mid-third year CSA to compare the reporting of pertinent negative and pertinent positive items in two SP cases. One case concerned abdominal pain and the other headache.

We limited the study to the medical history as these items can be more reliably validated than physical examination items. Physical examination may be technically difficult to view on camera or video-recordings and/or require judgment by SPs or observers as to how well a student performed against an established standard technique.^{11,12} These problems may contribute to the very high rates of inaccurate documentation reported for physical examination items in some studies.^{13,14} Our hypothesis was that students are more likely to report pertinent positive than pertinent negative items from the medical history, ie, that under-reporting would be significantly higher for pertinent negative items.

Methods

During the end of the first semester CSA, 55 mid-third-year students were assessed on the two cases portrayed by SPs. Faculty measured student performance on 15 key items from the medical history. After each SP encounter, students recorded information in the PEN, and SPs used checklists to score obtained/not obtained for each key item. Independently trained observers monitored the accuracy of SP performance. Upon completion of the PEN, the 15 key items selected by faculty were

From the Department of Family and Community Medicine, University of Kansas School of Medicine-Wichita.

scored and recorded as reported/not reported. A single faculty member, who was blinded from the SP report, scored and reported all PENS. Overall performance validity was evaluated by subsequently reviewing recordings for approximately 20% of the encounters and making comparisons to SP checklists.

Two faculty members, who were not involved in evaluating the formative CSA and were blinded to student scores, classified items as pertinent negatives or pertinent positives. Pertinent negatives were defined as key items whose absence was significant to the case. Examples included significant risk factors (such as smoking) or symptoms (such as blood in the stool) that the SP denied if asked. Of the 30 items (15 per case), eight pertinent negative and 18 pertinent positive items were identified by consensus. The remaining four questions were qualitative or descriptive and were excluded from the study.

Descriptive statistics were used to summarize results and compare proportions of information obtained per the SP checklist and reported per the PEN by case and in total. Binomial

two-sample proportion z-tests were conducted to assess significant differences. The retrospective analysis of these data was determined to be an exempt study by the Institutional Review Board's Human Subjects Committee.

Results

Table 1 compares and summarizes results from SP checklists and student PENS. Overall congruence between checklists and PENS was similar to that reported in previous studies.⁷ For all 30 items, students obtained 84.8% and reported 66.7% of items performed. They under-reported (obtained but did not report) 23.8% of the items (Table 1). Students obtained high rates of both pertinent negative (87.7%) and pertinent positive (87%) items. Students were significantly more likely to report positive items than negative items, 75.1% compared to 52.3% respectively, $z=8.53$, $P<.001$. Conversely, in under-reporting, the proportion of positive items was significantly lower than negative items, 17.6% compared to 40.7%, $z=9.35$, $P<.001$.

Overall, results were similar between cases (Table 2). Students were

significantly more likely to report pertinent positive items than negative items, 74.1% (abdominal) and 76.0% (headache) compared to 53.9% (abdominal) and 50.9% (headache), respectively. Detailed examination of the question content showed the highest reporting of a negative item was for smoking, whereas the lowest was for radiation of pain. For positive items, the highest reporting was for specific location of abdominal pain and the lowest was for use of prescription medication.

Discussion

Despite obtaining almost all of the pertinent positive and negative items, students tended to omit from the PEN (under-report) pertinent negative items significantly more than positive items.

The PEN has been described as the "benchmark" for assessing physician competence as well as providing the medicolegal record of the patient encounter.¹⁴ Failure to report all the essential information obtained in a patient history can have many negative consequences, including inaccurate diagnosis, inappropriate follow-up, under-billing,

Table 1: CSA Results Comparing SP Checklist and PENS (55 Students, Two Cases)

	SP Checklist				Total (%)	
	Obtained (%)		Not Obtained (%)			
PENS: 30 items						
Reported	1,007	(61.0)	94	(5.7)	1,101	(66.7)
Not reported	393	(23.8)	156	(9.5)	549	(33.3)
Total	1,400	(84.8)	250	(15.2)	1,650	(100.0)
Pertinent negative: 8 items*						
Reported	207	(47.0)	23	(5.3)	230	(52.3)
Not reported	179	(40.7)	31	(7.0)	210	(47.7)
Total	386	(87.7)	54	(12.3)	440	(100.0)
Pertinent positive: 18 items*						
Reported	687	(69.4)	56	(5.7)	743	(75.1)
Not reported	174	(17.6)	73	(7.4)	247	(24.9)
Total	861	(87.0)	129	(13.0)	990	(100.0)

CSA—clinical skills assessment

SP—standardized patient

PENS—post-encounter notes

* Note: results are rounded such that only marginal percentages will total appropriately

Table 2: CSA Pertinent Items Obtained and Reported, Summarized by Type and by Case (55 Students)

Case Description	Obtained (%)		Reported (%)		Under-reported (%)		Obtained and Reported (%)	
Negative items, n=220								
Abdominal pain	194	(88.2)	118	(53.6)	93	(42.3)	101	(45.9)
Headache	192	(87.3)	112	(50.9)	86	(39.1)	106	(48.2)
Positive items, n=495								
Abdominal pain	419	(84.6)	367	(74.1)	88	(17.8)	331	(66.9)
Headache	442	(89.3)	376	(76.0)	86	(17.4)	356	(71.9)
Comparison (z, P)								
Abdominal pain	1.25	0.212	5.42	<.001	7.09	<.001	5.29	<.001
Headache	0.79	0.432	6.64	<.001	6.27	<.001	6.13	<.001

CSA—Clinical Skills Assessment

and underestimation of the quality of clinical data gathering.^{13,15} Nevertheless, several studies document that students, residents, and physicians substantially under-report important items obtained during the medical interview.^{6-10,13,14,16} Understanding which items get forgotten or ignored in documenting the clinical encounter is an important element in designing strategies to improve clinical care. This study validates the common clinical impression that pertinent negatives are significantly more likely than positive data to be omitted from PENs.

Educationally, the ability to prepare an accurate, complete, yet concise PEN is recognized as a core competency for both medical students and residents,^{17,18} yet little has been published on how to teach this skill.¹⁹ For educators, analysis of which items are selectively reported (and conversely those that are elicited but not reported) also provides insights on student utilization of clinical knowledge and decision making.^{20,21} Our students elicited the pertinent negative information, but we do not know why they failed to report so much of it in the PEN. The items that were omitted came from all sections of the history, ie, presenting complaint, review of systems, and social history. The magnitude of the effect and its pervasive nature suggest that simple

forgetting does not adequately explain the under-reporting. Students may under-value the significance of negative information in clinical decision making. Pertinent negative information may also be inadequately emphasized in clinical teaching across several clerkships. Such an effect may be too small to be apparent in individual courses and only apparent in a general assessment such as the CSA. Educators have been criticized for underutilizing such curricular-wide feedback from a CSA in improving educational programs.²²

This preliminary study has several limitations. Data were only gathered on 55 mid-third-year students and two SP cases. As previously reported, reporting accuracy may vary substantially by case^{6,8,13} and by student.^{7,23} Studies based on SP checklist reports and faculty scoring of PENs also face several methodological challenges in the design of items on checklists⁶ and the use of scoring instruments.²⁴⁻²⁶ We endeavored to minimize the principal problems by training the SPs and the physician scorer and also by validation of SP performances and checklists by observers. Reliability of PEN scores has not been shown to improve by duplicate scoring by a second trained physician.²⁷ The study was also limited as a retrospective analysis of examination data. Some questions could not be classified as either

“positives” or “negatives,” and others may have suffered from ambiguity. In editing cases for a future study, greater attention is being paid to design of checklist and PEN scoring rubrics as well as specific SP training in the areas of interest.²⁸ The editing process also allows validation of the clinical significance of each item²⁹ and a more even distribution of positive and negative items in the cases.

Finally, the study was limited by its role as an end of first semester assessment. Since students had only completed half of the required clerkships, this CSA focused on general attributes of patient assessment and clinical content that could be expected of all students, regardless of clinical experiences. Both students and residents are reported to improve performance on SP assessments with increasing clinical experience.³⁰ A larger CSA held at the end of the third year has the potential to include up to 12 cases and to address a much wider scope of clinical information. Future studies could also investigate any relationship between attention to pertinent negatives and other indicators of student performance.

CORRESPONDING AUTHOR: Address correspondence to Dr Walling, University of Kansas School of Medicine-Wichita, Department of Family and Community Medicine, 1010 N. Kansas, Wichita, KS 67214. 316-293-2607. Fax: 316-293-2696. awalling@kumc.edu.

References

- Boulet JR, Rebecchi TA, Denton EC, McKinley DW, Whelan GP. Assessing the written communication skills of medical school graduates. *Ad Health Sci Educ* 2004;9:47-60.
- Scoles PV, Hawkins RE, LaDuca A. Assessment of clinical skills in medical practice. *J Contin Educ Health Prof* 2003;23:182-90.
- Colliver JA, Travis TA, Robbs RS, et al. Assessment of uniqueness of information provided by post-encounter written scores on standardized patient examinations. *Eval Health Prof* 1992;15:465-74.
- Margolis MJ, Clauser BE, Swanson DB, Boulet JR. Analysis of the relationship between score components on a standardized patient clinical skills examination. *Acad Med* 2003;78:S68-S71.
- Harik P, Clauser BE, Grabovsky I, et al. Relationships among subcomponents of the USMLE Step 2 Clinical Skills Examination, the Step 1, and the Step 2 clinical knowledge examinations. *Acad Med* 2006;81:S21-S24.
- MacMillan MK, Fletcher EA, De Champlain AE, Klass DJ. Assessing post-encounter note documentation by examinees in a field test of a nationally administered standardized patient test. *Acad Med* 2000;75:S112-S114.
- Worzala K, Rattner SL, Boulet JR, et al. Evaluation of the congruence between students' post-encounter notes and standardized patients' checklists in a clinical skills examination. *Teach Learn Med* 2008;20:31-6.
- Walling A, Moser SE. Accuracy in post-encounter notes: are key items underreported, facetiously reported, or not obtained? *Teach Learn Med* 2010;22:63-4.
- Dresselhaus TR, Peabody JW, Lee M, et al. Measuring compliance with preventive care guidelines: standardized patients, clinical vignettes, and the medical record. *J Gen Intern Med* 2000;15:782-8.
- Luck J, Peabody JW, Dresselhaus TR, et al. How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. *Am J Med* 2000;108:642-9.
- Lammers RL, Davenport M, Korley F, et al. Teaching and assessing procedural skill using simulation: metrics and methodology. *Acad Emerg Med* 2008;15:1079-87.
- Peitzman SJ, McKinley D, Curtis M, et al. International medical graduates' performance of techniques of physical examination, with comparison of US citizens and non-citizens. *Acad Med* 2000;75:S115-S117.
- Szauter KM, Ainsworth MA, Holden MD, Mercado AC. Do students do what they write and write what they do? The match between the patient encounter and patient note. *Acad Med* 2006;81:S11-S12.
- Dresselhaus TR, Luck J, Peabody JW. The ethical problems of false positives: a prospective evaluation of physician reporting in the medical record. *J Med Ethics* 2002;28:291-4.
- Zierler-Brown S, Brown TR, Chen D, Blackburn RW. Clinical documentation for patient care: models, concepts, and liability considerations for pharmacists. *Am J Health Syst Pharm* 2007;64:1851-8.
- Langewitz WA, Loeb Y, Nubling M, Huniker S. From patient talk to physician notes—comparing the content of medical interviews with medical records in a sample of outpatients in internal medicine. *Pat Educ Couns* 2009;76:336-40.
- The Medical School Objectives Writing Group. Learning objectives for medical student education—guidelines for medical schools: Report 1 of the Medical School Objectives Project. *Acad Med* 1999;74:13-8.
- ACGME Outcome Project. www.acgme.org.
- McCarty T, Parkes MV, Anderson TT, et al. Improved patient notes from medical students during Web-based teaching using faculty-calibrated peer review and self-assessment. *Acad Med* 2005;80:S67-S70.
- Donnelly WJ. The language of medical case histories. *Ann Intern Med* 1997;127:1045-8.
- Nendaz MR, Gut AM, Perrier A, et al. Beyond clinical experience: features of data collection and interpretation that contribute to diagnostic accuracy. *J Gen Intern Med* 2006;21:1302-5.
- Petrusa ER. Taking standardized patient-based examinations to the next level. *Teach Learn Med* 2004;16:98-110.
- Van der Vleuten CPM, Swanson DB. Assessment of clinical skills with standardized patients: state of the art. *Teach Learn Med* 1990;2:58-76.
- Norcini J, Boulet J. Methodological issues in the use of standardized patients for assessment. *Teach Learn Med* 2003;15:293-7.
- Howley L, Szauter K, Perkowski L, et al. Quality of standardized patient research reports in the medical education literature: review and recommendations. *Med Educ* 2008;42:350-8.
- Whelan GP, Boulet J, McKinley, et al. Scoring standardized patient examinations: lessons learned from the development and administration of the ECFMG clinical skills assessment (CSA). *Med Teach* 2005;27:200-6.
- Boulet JR, Ben-David M, Hambleton RK, et al. An investigation of the sources of measurement error in the post-encounter written scores form standardized patient examinations. *Adv Health Sci Educ* 1998;3:89-100.
- Huber P, Baroffio A, Chamot E, et al. Effects of item and rater characteristics on checklist recording: what should we look for? *Med Educ* 2005;39:852-8.
- Boulet JR, van Zanten M, de Champlain A, et al. Checklist content on a standardized patient assessment: an ex post facto review. *Adv Health Sci Educ Theory Pract* 2006;13:59-63.
- Haydon R, Donnelly M, Schwartz R, et al. Use of standardized patients to identify deficits in student performance and curriculum effectiveness. *Am J Surg* 1994;168:57-65.