

# Assessing Electronic Note Quality Using the Physician Documentation Quality Instrument (PDQI-9)

P.D. Stetson<sup>1,2</sup>; S. Bakken<sup>1,3</sup>; J.O. Wrenn<sup>4</sup>; E.L. Siegler<sup>5</sup>

<sup>1</sup>Department of Biomedical Informatics, Columbia University; <sup>2</sup>Department of Medicine, Columbia University; <sup>3</sup>School of Nursing, Columbia University; <sup>4</sup>Department of Medical Informatics, Vanderbilt University; <sup>5</sup>Division of Geriatrics and Gerontology, Weill Cornell Medical College

## Keywords

Electronic health record, documentation, note, quality, instrument

## Summary

**Objective:** To refine the Physician Documentation Quality Instrument (PDQI) and test the validity and reliability of the 9-item version (PDQI-9).

**Methods:** Three sets each of admission notes, progress notes and discharge summaries were evaluated by two groups of physicians using the PDQI-9 and an overall general assessment: one gold standard group consisting of program or assistant program directors ( $n = 7$ ), and the other of attending physicians or chief residents ( $n = 24$ ). The main measures were criterion-related validity (correlation coefficients between Total PDQI-9 scores and 1-item General Impression scores for each note), discriminant validity (comparison of PDQI-9 scores on notes rated as best and worst using 1-item General Impression score), internal consistency reliability (Cronbach's alpha), and inter-rater reliability (intraclass correlation coefficient (ICC)).

**Results:** The results were criterion-related validity ( $r = -0.678$  to  $0.856$ ), discriminant validity (best versus worst note,  $t = 9.3$ ,  $p = 0.003$ ), internal consistency reliability (Cronbach's alphas =  $0.87-0.94$ ), and inter-rater reliability (ICC =  $0.83$ , CI =  $0.72-0.91$ ).

**Conclusion:** The results support the criterion-related and discriminant validity, internal consistency reliability, and inter-rater reliability of the PDQI-9 for rating the quality of electronic physician notes. Tools for assessing note redundancy are required to complement use of PDQI-9. Trials of the PDQI-9 at other institutions, of different size, using different EHRs, and incorporating additional physician specialties and notes of other healthcare providers are needed to confirm its generalizability.

## Correspondence to:

Peter D. Stetson, MD, MA  
Columbia University Medical Center  
650 West 168th Street  
BB2-239  
New York, NY 10032  
Email: pds14@columbia.edu  
Phone: 212.342.0029

## Appl Clin Inf 2012; 3: 164–174

doi:10.4338/ACI-2011-11-RA-0070

received: November 21, 2011

accepted: March 29, 2012

published: April 11, 2012

**Citation:** Stetson PD, Bakken S, Wren JO, Siegler EL. Assessing electronic note quality using the physician documentation quality instrument (PDQI-9). Appl Clin Inf 2012; 3: 164–174  
<http://dx.doi.org/10.4338/ACI-2011-11-RA-0070>

## 1. Background

Electronic notes (e.g. initial visit, follow-up, consult notes) are a core feature of “fully functional” ambulatory electronic health records (EHRs) [1], and a “core” component of proposed Stage 2 Meaningful Use criteria [2]. Electronic note-writing in inpatient EHRs has been associated with a reduction in mortality during hospitalization [3]. Use of electronic visit notes in the ambulatory setting has been associated with improved Healthcare Effectiveness Data and Information Set (HEDIS) quality measures [4]. As EHR adoption continues to grow, several key questions remain unanswered regarding electronic clinical notes. Evaluating the effects of entry modality (free text vs. structured templates[5]) on the “professionalism” and “readability” of EHR note outputs, the relationship of note quality to team communication and care coordination, and the relationship of note quality to care quality, requires the development of valid measurement tools.

### 1.1 Brief History of Note Quality

It is likely that the quality of hospital notes has always been mixed; even charts on Osler’s service were poor and failed to reflect the care or bedside clinical discussion that was taking place during the last decade of the 19<sup>th</sup> century [6].

Serious attempts to understand documentation and its value began in the late 1960s and early 1970s in anticipation of computerization of the record, including physician’s notes. Weed’s Problem Oriented Medical Record [7] and Fries’ Time Oriented Record [8] were both designed to organize information in a way that would render it accessible to readers, beneficial to writers, and useable by programmers. Reiser [9] described erroneous, missing and poorly synthesized data, imprecise language, and limited narrative. Burnum [10] wrote more bluntly about many of the same issues and concluded, “...all medical record information should be regarded as suspect; much of it is fiction” (p 484).

### 1.2 Unintended Consequences of Electronic Documentation

Certain factors that diminish handwritten note quality, such as illegibility or absence of data, have disappeared with the advent of the EHR [11]. However, EHRs introduced features like the copy and paste function, giving rise to new concerns about documentation quality [12–18]. Our own experience implementing electronic documentation modules within commercial EHRs has uncovered provider concerns about the perceived lack of high-quality EHR note outputs (reading like “boilerplate”). Our providers pride themselves on well-crafted notes, believing that the quality of their written communication helps their patients, and helps to maintain robust referring/consulting relationships. Other investigators have confirmed that the “attractiveness” of notes is at risk with electronic note-writing systems [19].

Note quality has remained an elusive concept. Studies examining the quality of the medical record have generally focused on data completeness and accuracy [20–22], but quality notes encompass more than just data organized into relatively standard components. Rosenbloom et al [23], who interviewed providers to determine the factors that influenced satisfaction with clinical documentation systems, found that document system time efficiency, availability, expressivity, structure, and quality were most important. Note quality was defined by legibility, accuracy, thoroughness, and compliance with administrative documentation standards, but no assessment tool emerged from this work.

No valid and reliable instrument exists to evaluate the quality of physician notes, despite the need for such a tool. We have previously published preliminary evidence for the construct validity and internal consistency reliability of a 22-item assessment tool, the Physician Documentation Quality Instrument (PDQI) [24]. The terms used in the original tool were derived from a combination of literature review and expert opinion and then subjected to formal psychometric analysis.

## 2. Objectives

Our objectives in this study were to refine attributes of a high quality note, and validate a usable tool to assess electronic note quality. Our approach involved extending our prior work on the PDQI to create a simpler tool for real-world application in clinical care, and confirm its validity and reliability for rating the quality of electronic admission, progress, and discharge notes. We also assessed provider use of the PDQI to detect redundancy (copy/paste) by comparison to an automated method for redundancy detection. We sought to create a tool that could be used for teaching purposes with medical students and residents, as a guide to EHR developers, and as a research tool for assessing whether there is a correlation of documentation quality with care quality.

## 3. Methods

Study methods involved item reduction from the 22-item PDQI to develop the 9-item PDQI, creation of the note corpus for rating using the PDQI-9 and 1-item General Impression score, PDQI-9 data collection to support establishment of a gold standard, and descriptive and psychometric analyses (criterion-related validity, discriminant validity, internal consistency reliability, and inter-rater reliability). This study was approved by our Institutional Review Board.

### 3.1 Development of the 9-item PDQI

We used several strategies to reduce the number of items on the original PDQI instrument. First, we conducted additional exploratory factor analyses and internal consistency reliability assessments on our original data set with the goal of eliminating items that were not relevant to all three note types of interest (admission, progress, discharge). This was done in order to allow assessment of the generalizability of the tool for all three note types. Second, we examined the remaining items for possible item reduction, and eliminated items that did not negatively affect the internal consistency reliability of the factor scores on the 22-item PDQI and were judged by an expert in note quality to be redundant. For example, brief, concise, and succinct all loaded on the same original PDQI factor, and succinct was selected as the single item to convey the concept. The item reduction strategies resulted in a 9-item PDQI consisting of *Up-to-Date*, *Accurate*, *Thorough*, *Useful*, *Organized*, *Comprehensible*, *Succinct*, *Synthesized*, and *Consistent*. Each item included a five-point Likert scale, with a descriptor of ideal characteristics anchoring the highest value of the scale. General terms “Not at all” and “Extremely” appeared over the numbers 1 and 5, respectively. A copy of the PDQI-9 appears in the Appendix (see supplementary file).

Neither the original PDQI nor the 9-item instrument is designed to assess the presence or absence of specific note components (e.g. “reason for admission” in an admission note), but rather they focus on descriptive characteristics. This was a design choice intended to

1. reflect the terms were published in the literature review conducted for the original PDQI instrument [24], and
2. to ensure that the tool had the broadest applicability possible across different clinical note template types which might have different required sections (e.g. “reason for referral” in a consult note).

The PDQI-9 is included in the Appendix (see supplementary file).

### 3.2 Creation of Note Corpus for Testing the PDQI-9

We selected three charts from admissions to the teaching internal medicine service of the Columbia or Allen campuses of the New York Presbyterian Hospital using inclusion criteria developed for a study examining redundancy in the EHR [25]. All hospitalizations lasted at least 72 hours and contained physician-authored admission notes, progress notes and discharge summaries. The notes in the three charts were written in an inpatient EHR by physicians who typed in the notes on a variety of unstructured and semi-structured templates. The notes chosen for analysis represented a wide

range of quality to support assessment of the discriminant validity of the PDQI-9. The notes were chosen from two time periods (one in 2006, the other in 2009) in order to represent as broad a possible range of note entry modalities (it was possible the entry templates used in our EHRs changed over time).

We prepared a de-identified, hard-copy of the three charts with three sections divided by tabs. The “Notes” section contained, in chronological order, all of the progress notes, the admission note and the discharge summary, in addition to a discharge summary from the prior admission (if applicable). We also created separate sections for the laboratory data and for other relevant data such as electrocardiogram interpretations and radiology reports. They were arranged in reverse chronological order as they are within our EHR.

### 3.3 Data Collection Procedures

The study sample comprised two types of participants. For the initial analysis, seven physician leaders who had experience with house staff training were chosen to establish a gold standard. The rationale for these physician leaders to comprise the gold standard were

1. these faculty members were not part of the research team, and
2. were routinely involved in the structured evaluations of house staff performance, including communication skills and chart reviews in both the inpatient and outpatient setting (including past and current Residency and Fellowship Program Directors, and Key Faculty Advisors).

The remainder of the sample consisted of 24 volunteers from among present medical chief residents and attending physicians on the two campuses, including internal medicine attendings and subspecialists. All participants volunteered their time.

Using a random number generator ([www.random.org](http://www.random.org)) to establish order, we presented the three charts (named “C,” “D,” and “E”) in random order, the notes within each chart in random order, and the two evaluation tools for each note in random order. Participants were required to spend time reviewing these charts at the beginning of the review session to familiarize themselves with the patients (as though they were “picking up the patient”). The seven physician leaders assessed the quality of the admission note, a designated progress note, and the discharge summary from each of three charts using the PDQI-9 and a single item General Impression of note quality on a five point Likert Scale from “Terrible” to “Excellent.” The use of the two measures allowed us to capture specific ratings with the PDQI-9 and the overall “Gestalt” judgment regarding note quality. The 1-item General Impression score was used to compare a single “gestalt” score to a total score for the 9 items in the PDQI. We subsequently repeated these same procedures with the additional 24 participants.

Following their assessment of the nine notes, we also asked the seven physician leaders to provide general comments regarding the PDQI-9 and suggestions for its improvement. Some other participants informally provided comments after their note rating sessions. Observations of the data collection sessions were recorded.

### 3.4 Descriptive and Psychometric Analyses

#### 3.4.1 General data management and analysis

We used PASW Version 18 for data management and analysis. A Redundancy score was assessed for each note (the percentage of the note that was copied from other notes or reports) using the algorithm from our prior study (determined by a novel Levenshtien distance calculation) [25]. We included this step to assess whether the note reviewers would miss copy/paste in their assessments of the notes, even though we included all notes from the hospital stay in our prepared charts. One set of PDQI-9 and General Impression ratings was discarded to remove potential bias because the participant recognized the notes. We used mean item substitution for PDQI-9 score in an instance in which one item was not rated. When more than one rating was missing, we excluded all ratings for that reviewer from the analysis for the specific note. Narrative comments and observations were summarized by the senior author (ES).

### 3.4.2 Criterion-related validity

The physician leaders' General Impression ratings were considered to be the gold standard or criterion. To assess criterion-related validity, we calculated Pearson Correlation coefficients between the General Impression score and the PDQI-9 Total score.

### 3.4.3 Discriminant validity

We assessed the PDQI-9's discriminant ability by comparing PDQI-9 scores on the note judged to be of lowest quality versus the note judged to be of highest quality by the majority of all 31 participants using the single-item General Impression score.

### 3.4.4 Internal consistency reliability

Internal consistency reliability for the PDQI-9 for each type of note was measured using Cronbach's alpha.

### 3.4.5 Inter-rater reliability

We assessed inter-rater reliability using the intraclass correlation for consistency of average measures on the PDQI-9 total scores for each of the nine notes. In a two-way mixed model, each note was considered as a fixed effect and each rater as a random effect.

## 4. Results

Most participants spent 60–90 minutes to review charts and score the nine notes. The bulk of the time was spent reading the chart to gain an understanding of the patient's hospital course.

### 4.1 Descriptive statistics

► Table 1 includes the mean Total PDQI-9, General Impression, and Redundancy scores for each note. Although there was a broad range of scores and wide standard deviation for certain items in certain notes, there was no pattern, i.e., no item had a standard deviation of greater than 1 for more than three of the nine notes, and no one note type (e.g., progress) was responsible for the items with a higher standard deviation. For example, the item "thorough" scored a high standard deviation for the admission note of chart C, the progress note of chart C, and the discharge summary of chart E. It performed well for all other notes.

There was no significant correlation ( $r = -0.423$ ,  $p$  (2 tailed) = 0.256) between total PDQI-9 scores and redundancy scores.

### 4.2 Criterion-related validity

► Table 2 lists correlation coefficients between Total PDQI-9 scores and General Impression scores for each note. Correlations were highly significant for all notes with both groups of raters, with the exception of the admission note of chart C for the physician leader group, likely due to the small size ( $n = 7$ ).

### 4.3 Discriminant validity

Discharge note "D" was rated as Terrible or Bad by 58.1% of the participants and its mean PDQI-9 score was 26.2; 87.1% of the participants rated admission note "D" as Good or Excellent and its mean PDQI-9 score was 36.6. In a paired samples t-test of PDQI-9 scores,  $t = 9.3$  and  $p = 0.003$ .

### 4.4 Internal consistency reliability

Cronbach's alphas were admission note: 0.87, progress note: 0.93, and discharge summary: 0.94.

## 4.5 Inter-rater reliability

The intraclass correlation was .83 (Confidence Interval = 0.72–0.91).

## 4.6 Narrative comments and observations

No reviewer had difficulty using or understanding the PDQI-9, and some commented that the exercise encouraged them to think about their own notes or their judging styles. There was clear variability in quality of assessments – some reviewers gave the same rating to all attributes (e.g., “straight 3s”), causing a “gravitational pull” that lessened the discriminatory value of the individual items. For example, one reviewer commented, “too succinct” but scored succinct as a 3.

A common, although not universal, concern was uncertainty about evaluating the discharge summary; reviewers felt that there was a lack of unanimity about the amount and type of data it should capture, how it could be constructed when it was based on notes from multiple authors.

Some reviewers questioned how specific measures (e.g., the presence of inappropriate abbreviations or the lack of information about immunizations) or the use of copy/paste ought to affect the assessment of overall note quality; occasionally there was a very strong negative response to a note, where a “fatal flaw” led to a low overall score, even though the PDQI-9 score was not as negative because the note scored fair or better on specific individual attributes.

## 5. Discussion

This study offers preliminary evidence for the criterion-related and discriminant validity, internal consistency reliability, and inter-rater reliability of the PDQI-9 for rating of admission, progress, and discharge notes. Seeking simplicity and broad applicability while taking advantage of the psychometric strengths of the original 22-item instrument [24], we created a shortened instrument that captured the essential attributes while eliminating redundant terms. This work on note quality extends that of others [19, 26–32], through psychometric validation, and the creation of a brief tool that is applicable across documentation systems.

The PDQI-9 was able to discriminate between good and bad notes and delineate the note's qualities. It functioned best in the evaluation of admission and progress notes. It also had acceptable reliability and validity scores for discharge summaries, but some caution is appropriate; scorers felt uncomfortable rating the summaries, saying that there was no unanimity about the qualities of the ideal discharge summary.

The PDQI-9 did not correlate with the extent of copy/paste as measured by the redundancy score. This is not an unexpected finding. The PDQI-9 can flag lack of succinctness and internal inconsistencies that result from poor use of copy/paste, but it does not assess originality. Moreover, a copied note may appear to be of acceptable quality by the PDQI-9 measures, especially if the note from which it is derived is good and the patient's course has not changed dramatically, or the note's author reviewed and edited copied material carefully. None of the reviewers in our study detected when notes were mostly copied/pasted, even though there were several highly redundant notes (► Table 2). While one study suggests redundancy may not be as tightly linked with note quality [32], and not all redundancy is inappropriate [25, 38], we recommend that redundancy scores or overall assessment of originality should complement the PDQI-9 and serve as a “second look” [25].

For pedagogical purposes, we recommend that the PDQI-9 be used as part of a three-phase evaluation that assesses the Total PDQI-9 score of a note, followed by a review of the item-specific PDQI-9 scores, and a separate assessment of the amount of copy/paste used to generate the note. The learner easily comprehends the total score; the PDQI-9 item-specific scores provide a breakdown of the underlying attributes (the “why” of the total score) and targets areas for improvement; and the direct comparison with other notes enables the reviewer to assess the note's originality and its role in the larger medical record. Although the nine notes in this study required up to 90 minutes to review, we have confirmed through subsequent use in practice by the authors (ES and PDS) that the PDQI-9 can be completed in less than five minutes when the attending knows the patient.

We agree with Bates and O'Malley [33–35] in their contention that EHRs will certainly require new tools that go beyond notes to support coordinated Team Care in the Patient-centered Medical Home model. Furthermore, electronic note-writing (whether authored as currently supported by EHRs; with speech recognition;  $\pm$  natural language processing [5]; using “Wikis” [36] or “Walls”) will continue to serve as a foundation for articulating the medical decision-making and ongoing care of the patient, and the quality of care provided. This is supported by the association of electronic documentation with improved clinical measures and outcomes [3–4].

With the PDQI-9 in hand, future studies could be directed at assessing the effects of entry modality on note quality, and the effects of electronic notes of varying quality on care processes, outcomes, and medical liability. EHR vendors might benefit from the application of a tool like the PDQI-9 as part of their development and quality control processes for note output quality and attractiveness [19]. Information from studies such as these might inform whether electronic notes should be re-introduced as a Meaningful Use measure [37].

There are a number of limitations to this study. First, the chart was a simplified, de-identified paper version composed of printouts from an electronic record; the test situation was thus somewhat artificial. However, we do not think that reading an abbreviated paper version of the chart failed to convey the sense of a chart or distracted the participants from the task of documentation assessment. Second, the raters did not know the patients and were required to determine diagnosis and hospital course by reading the chart prepared for them. Although we chose hospitalizations of intermediate length to try to limit the work of review, comprehending the hospital course required time and concentration and may have increased the likelihood that the rater would miss note inaccuracies or redundancies. Third, there was no training period for the reviewers, and their ease and accuracy likely increased after they had reviewed a number of notes. Although it was not possible to eliminate training effect altogether, we randomized chart and note order in order to limit its impact on any one note type. Fourth, the reviewers were internal medicine attendings and chief residents. Physicians of other specialties may have scored these notes differently with the PDQI-9. Fifth, the PDQI-9 score for a given note may be affected by the quality of the “surrounding notes” for a patient’s admission. We did not assess the notes that were scored in isolation from the full chart to assess this effect. Sixth, the notes used represent a relatively small sample size and a subset of possible note-types within the chart, and may not capture the full diversity of patient complexity and information transfer needs placed on the medical record. Seventh, there may be additional document quality items not included in our analysis which we did not detect in our literature review or elicit from experts in our original study. Finally, the study was performed on two hospital campuses, but both were part of the same medical school. It is possible that our attending physician perceptions of note quality may differ from those of other institutions.

Trials of the PDQI-9 at other institutions, of different size, using different EHRs, and incorporating additional physician specialties and notes of other healthcare providers are needed to confirm its generalizability.

## 6. Conclusion

We have provided preliminary evidence of the validity and reliability of the PDQI-9 for evaluating inpatient physician notes on two campuses of an urban academic hospital. It complements measurements of inappropriate use of the copy/paste function (e.g., redundancy scores). The PDQI-9 is simple to apply and interpret. It can be used to assess the output quality of EHR note modules, as well as explain the components of a quality note and document areas of improvement for trainees.

## 7. Clinical Relevance Statement

Electronic notes are associated with reductions in-hospital mortality, improvements in ambulatory quality measures, and are a proposed “core” measure for Eligible Providers in Stage 2 Meaningful Use. However, electronic notes suffer from issues of redundancy and readability. The PDQI-9 is a valid, reliable tool for assessing the quality of physician inpatient electronic documentation that will

support further evaluation studies of note input modalities on note quality, and the relationship of note quality to care quality.

**Conflict of Interest**

None

**Human Subjects Protections**

Human subjects were not included in the project, except as users of an instrument to review historical chart notes. The study was performed in compliance with Ethical Principles for Medical Research Involving Human Subjects, and was reviewed by Columbia University Medical Center Institutional Review Board.

**Acknowledgements**

1. Contributors: None
2. Funders
  - National Library of Medicine: K22 LM008805 (PDS)
  - Health Resources and Services Administration: D11HP07346 (SB)
3. Prior Presentations: None

**Table 1** Means and standard deviations for gold standard and test attendings

Note	GS Total PDQI-9		TA Total PDQI-9		GS General Impression		TA General Impression		Redundancy Score‡
	Mean*	(SD)	Mean*	(SD)	Mean†	(SD)	Mean	(SD)	
Admission – Chart C	36.00	5.48	38.17	(5.23)	4.14	0.69	3.96	0.69	0.07
Admission – Chart D	38.14	3.98	37.43	5.26	4.29	0.49	4.13	0.74	0.01
Admission – Chart E	25.00	3.61	28.87	5.34	2.00	0.58	2.64	0.85	0.06
Progress – Chart C	32.43	8.58	37.43	5.62	3.43	0.54	4.00	0.66	0.49
Progress – Chart D	23.00	2.64	28.42	5.55	2.14	0.69	2.57	0.95	0.80
Progress – Chart E	30.71	5.71	31.96	6.519	3.29	0.95	3.22	1.0	0.78
Discharge – Chart C	29.71	6.80	31.92	7.19	2.57	0.98	2.96	1.0	0.07
Discharge – Chart D	24.57	5.71	27.14	6.3	2.43	0.79	2.30	1.06	0.66
Discharge – Chart E	30.71	5.16	29.78	6.46	3.14	1.07	3.00	0.894	0.10

GS= Gold Standard, n = 7

TA= Test Attendings, n = 24

\*Range of Total PDQI-9 score is 9–45. Higher score indicates higher quality.

†Range of General Impression score is 1–5. Higher score indicates higher quality.

‡Redundancy score ranges from 0 to 1. Higher score indicated larger amount of material copied from other notes or reports.

**Table 2** Criterion-related validity: Correlations between total PDQI-9 score and general impression score

Variables	All Evaluators		GS		TA	
	R	p	R	p	R	p
Admission –Chart C	0.782	0.000	0.617	0.140	0.878	0.000
Admission – Chart D	0.809	0.000	0.920	0.003	0.795	0.000
Admission – Chart E	0.678	0.000	0.881	0.009	0.610	0.003
Progress – Chart C	0.856	0.000	0.971	0.000	0.839	0.000
Progress – Chart D	0.774	0.000	0.822	0.023	0.776	0.000
Progress – Chart E	0.784	0.000	0.939	0.002	0.756	0.000
Discharge – Chart C	0.813	0.000	0.782	0.038	0.816	0.000
Discharge – Chart D	0.756	0.000	0.864	0.012	0.763	0.000
Discharge – Chart E	0.838	0.000	0.855	0.014	0.850	0.000

GS= Gold Standard, n = 7

TA= Test Attendings, n = 24

## References

- DesRoches CM, Campbell EG, Rao SR, Donelan K, Ferris TG, Jha A, et al. Electronic health records in ambulatory care – a national survey of physicians. *N Engl J Med* 2008; 359(1): 50–60.
- Medicare and Medicaid programs; electronic health record incentive program. Final rule. *Fed Regist* 2010; 75(144): 44313–44588.
- Amarasingham R, Plantinga L, Diener-West M, Gaskin DJ, Powe NR. Clinical information technologies and inpatient outcomes: a multiple hospital study. *Arch Intern Med* 2009; 169(2): 108–114.
- Poon EG, Wright A, Simon SR, Jenter CA, Kaushal R, Volk LA, et al. Relationship between use of electronic health record features and health care quality: results of a statewide survey. *Med Care* 2010; 48(3): 203–209.
- Johnson SB, Bakken S, Dine D, Hyun S, Mendonca E, Morrison F, et al. An electronic health record based on structured narrative. *J Am Med Inform Assoc* 2008; 15(1): 54–64.
- Kirkland LR, Bryan CS. Osler's service: a view of the charts. *J Med Biogr* 2007; 15(Suppl 1): 50–54.
- Weed LL. The problem oriented record as a basic tool in medical education, patient care and clinical research. *Ann Clin Res* 1971; 3(3): 131–134.
- Fries JF. Alternatives in medical record formats. *Med Care* 1974; 12(10): 871–881.
- Reiser SJ. The clinical record in medicine. Part 2: Reforming content and purpose. *Ann Intern Med* 1991; 114(11): 980–985.
- Burnum JF. The misinformation era: the fall of the medical record. *Ann Intern Med* 1989; 110(6): 482–484.
- Embi PJ, Yackel TR, Logan JR, Bowen JL, Cooney TG, Gorman PN. Impacts of computerized physician documentation in a teaching hospital: perceptions of faculty and resident physicians. *J Am Med Inform Assoc* 2004; 11(4): 300–309.
- Hartzband P, Groopman J. Off the record – avoiding the pitfalls of going electronic. *N Engl J Med* 2008; 358(16): 1656–1658.
- Hirschtick RE. A piece of my mind. Copy-and-paste. *Jama* 2006; 295(20): 2335–2336.
- Siegler EL, Adelman R. Copy and paste: a remediable hazard of electronic health records. *Am J Med* 2009; 122(6): 495–496.
- Thielke S, Hammond K, Helbig S. Copying and pasting of examinations within the electronic medical record. *Int J Med Inform*. 2007; 76(Suppl. 1): S122–S128.
- Weir CR, Hurdle JF, Felgar MA, Hoffman JM, Roth B, Nebeker JR. Direct text entry in electronic progress notes. An evaluation of input errors. *Methods Inf Med* 2003; 42(1): 61–67.
- O'Donnell HC, Kaushal R, Barron Y, Callahan MA, Adelman RD, Siegler EL. Physicians' attitudes towards copy and pasting in electronic note writing. *Journal of general internal medicine: official journal of the Society for Research and Education in Primary Care Internal Medicine* 2009; 24(1): 63–68.
- Gelzer R, Hall T, Liette E, Reeves MG, Sundby J, Tegen A, et al. Auditing copy and paste. *J AHIMA*. 2009 Jan;80(1):26–9; quiz 31–2.
- Payne T, Patel R, Beahan S, Zehner J. The Physical Attractiveness of Electronic Physician Notes. *AMIA Annu Symp Proc* 2010: 622–626.
- Institute of Medicine . Committee on Improving the Patient R, Dick RS, Steen EB, Detmer DE. The computer-based patient record : an essential technology for health care. Washington, D.C.: National Academy Press; 1997.
- Romm FJ, Putnam SM. The validity of the medical record. *Med Care* 1981; 19(3): 310–315.
- Tufo HM, Speidel JJ. Problems with Medical Records. *Medical care* 1971; 9(6): 509–517.
- Rosenbloom ST, Crow AN, Blackford JU, Johnson KB. Cognitive factors influencing perceptions of clinical documentation tools. *J Biomed Inform* 2007; 40(2): 106–113.
- Stetson PD, Morrison FP, Bakken S, Johnson SB. Preliminary development of the physician documentation quality instrument. *J Am Med Inform Assoc* 2008; 15(4): 534–541.
- Wrenn JO, Stein DM, Bakken S, Stetson PD. Quantifying clinical narrative redundancy in an electronic health record. *J Am Med Inform Assoc* 2010; 17(1): 49–53.
- Aronsky D, Haug PJ. Assessing the quality of clinical data in a computer-based record for calculating the pneumonia severity index. *J Am Med Inform Assoc* 2000; 7(1): 55–65.
- Coakley FV, Heinze SB, Shadbolt CL, Schwartz LH, Ginsberg MS, Lefkowitz RA, et al. Routine editing of trainee-generated radiology reports: effect on style quality. *Acad Radiol* 2003; 10(3): 289–294.
- Efthimiadis EN, Hammond KW, Laundry R, Thielke SM. Developing an EMR simulator to assess users' perception of document quality. *Proc 43rd Hawaii Int Conf on System Sciences – 2010*. 2010: 1–9.
- Hogan WR, Wagner MM. Accuracy of data in computer-based patient records. *J Am Med Inform Assoc* 1997; 4(5): 342–355.
- Logan JR, Gorman PN, Middleton B. Measuring the quality of medical records: a method for comparing completeness and correctness of clinical encounter data. *Proc AMIA Symp* 2001: 408–412.

31. Myers KA, Keely EJ, Dojeiji S, Norman GR. Development of a rating scale to evaluate written communication skills of residents. *Acad Med* 1999; 74(10 Suppl.): S111-S113.
32. Hammond KW, Efthimiadis EN, Weir CR, Embi PJ, Thielke SM, Laundry RM, et al. Initial steps toward validating and measuring the quality of computerized provider documentation. *AMIA Annu Symp Proc* 2010. 2010: 271–275.
33. Bates DW. Getting in step: electronic health records and their role in care coordination. *J Gen Intern Med* 2010; 25(3): 174–176.
34. Bates DW, Bitton A. The future of health information technology in the patient-centered medical home. *Health Aff (Millwood)*. 2010; 29(4): 614–621.
35. O'Malley AS, Grossman JM, Cohen GR, Kemper NM, Pham HH. Are electronic medical records helpful for care coordination? Experiences of physician practices. *J Gen Intern Med* 2010; 25(3): 177–185.
36. Naik AD, Singh H. Electronic health records to coordinate decision making for complex patients: what can we learn from wiki? *Med Decis Making* 2010; 30(6): 722–731.
37. Meaningful Use Workgroup Request for Comments Regarding Meaningful Use Stage 2. Health Information Technology Policy Committee; 2011; Available from: [http://healthit.hhs.gov/media/faca/MU\\_RFC%202011-01-12\\_final.pdf](http://healthit.hhs.gov/media/faca/MU_RFC%202011-01-12_final.pdf).
38. Hammond KW, Helbig ST, Benson CC, Brathwaite-Sketoe BM. Are electronic medical records trustworthy? Observations on copying, pasting and duplication. *AMIA Annu Symp Proc* 2003: 269–273.