

Agreement Between Patient-reported Symptoms and Their Documentation in the Medical Record

Serguei V. Pakhomov, PhD; Steven J. Jacobsen, MD, PhD; Christopher G. Chute, MD, DrPH; and Véronique L. Roger, MD, MPH

Objective: To determine the agreement between patient-reported symptoms of chest pain, dyspnea, and cough and the documentation of these symptoms by physicians in the electronic medical record.

Methods: Symptoms reported on patient-provided information forms between January 1, 2006, and June 30, 2006, were compared with those identified by natural language processing of the text of clinical notes from care providers. Terms that represent the 3 symptoms were used to search clinical notes electronically with subsequent manual identification of the context (eg, affirmative, negated, family history) in which they occurred. Results were reported using positive and negative agreement, and kappa statistics.

Results: Symptoms reported by 1119 patients age 18 years or older were compared with the non-negated terms identified in their clinical notes. Positive agreement was 74, 70, and 63 for chest pain, dyspnea, and cough, while negative agreement was 78, 76, and 75, respectively. Kappa statistics were 0.52 (95% confidence interval [CI] = 0.44, 0.60) for chest pain, 0.46 (95% CI = 0.37, 0.54) for dyspnea, and 0.38 (95% CI = 0.28, 0.48) for cough. Positive agreement was higher for older men ($P > .05$), and negative agreement was higher for younger women ($P > .05$).

Conclusions: We found discordance between patient self-report and documentation of symptoms in the medical record. This discordance has important implications for research studies that rely on symptom information for patient identification and may have clinical implications that must be evaluated for potential impact on quality of care, patient safety, and outcomes.

(*Am J Manag Care.* 2008;14(8):530-539)

For author information and disclosures, see end of text.

Physical symptoms account for half of all outpatient visits in the United States¹ and are commonly not diagnosed.² The verbal characterization of the symptoms conveyed by the patient and recorded by care providers is central to the practice of clinical medicine, and increasing importance is attached to patient-centered clinical care.³ With the increasing adoption of the electronic medical record (EMR), free text of the clinical history can now be subjected to automated analysis in ways that are impossible or uneconomic with paper-based records.^{4,6} A common, costly⁷ example of a symptomatic condition in which history taking is central to management is chronic stable angina pectoris,⁸ the number 1 cause of death in the entire Western world.⁹ In addition to its mortality burden, angina is associated with serious morbidity such as myocardial infarction and heart failure. Optimal methods of identifying patients with stable angina remain unclear. Many patients with typical symptoms are not diagnosed with angina,¹⁰ and age, sex, and ethnicity may influence physicians' recommendations for diagnostic testing such as coronary angiography and the resulting *International Classification of Diseases* codes.¹¹⁻¹⁴ Our preliminary findings indicate that natural language processing (NLP) of the free text of the EMR identifies patients with chronic angina pectoris¹⁵ and heart failure¹⁵ who were missed by traditional diagnostic coding approaches. Natural language processing is a range of computational techniques aimed at extracting useful information from unstructured text. In the context of the EMR, NLP offers a promising method to automate the collection of a richer set of information for quality improvement and safety that would otherwise require manual chart abstraction.

Early identification of patients at risk for myocardial infarction is critical to its prevention and improved prognosis,¹⁶⁻¹⁸ and is possible using patient-reportable information.¹⁹ Because the diagnosis of angina pectoris relies on the patient conveying the symptoms to the physician,¹⁰ the ambiguous nature of verbal communication as well as the nature of coronary artery disease (ie, its presentation varies according to patient race and sex) can make the diagnosis challenging. These challenges can lead to inconsistency and incompleteness in both the diagnosis and the information recorded in the EMR.¹¹ Currently, the nature of the processes that lead to the creation of the EMR is poorly understood; however, prior research indicates that these processes have

In this issue

Take-away Points / p538

www.ajmc.com

Full text and PDF

Web exclusive

eAppendix Questionnaire

Patient-reported Symptoms in the EMR

critical implications for clinical care and may have a significant impact on patient outcomes.

The primary objective of the current study is to determine whether there was disagreement between patient-reported symptoms of chest pain, dyspnea, and cough and the documentation of these symptoms by care providers in the EMR. Although in this study we focus on heart disease, which constitutes one of the top national healthcare priority areas,²⁰ our findings have significant implications for any condition whose diagnosis and treatment relies on verbal presentation of symptoms.

METHODS

Sources of Symptom Documentation

For this study we identified 2 primary sources of symptom information: patient-provided information forms and clinical notes. Both are part of the Mayo Clinic EMR.

Electronic Medical Record and Clinical Notes. The Mayo Clinic maintains an EMR for each inpatient and outpatient. A large part of the EMR consists of clinical notes, which represent transcriptions by trained medical transcriptionists of the dictations recorded by care providers after each

contact with the patient. Clinical notes have been in use electronically since 1994; however, all services at the Mayo Clinic switched to electronic notes in 2005. The content of these notes complies with a nationally accepted standard, Health Level 7 Clinical Document Architecture, and consists of standard sections including chief complaint, history of present illness, review of systems, and impression/report/plan, among others. This standard is used by most major EMR vendors in the United States.

Current Visit Information Forms. Many healthcare providers, including the Mayo Clinic, ask their patients to fill out forms detailing their prior health and social history, current symptoms, medications, and allergies. The Mayo Clinic's form is reproduced in the **eAppendix** (available at www.ajmc.com). **Figure 1** shows item 14 of this patient-provided information form, which asks the patient to indicate if he/she has ever experienced a symptom listed on the form. Subsequently, the patient's responses are captured in a structured format by scanning the form into a database.

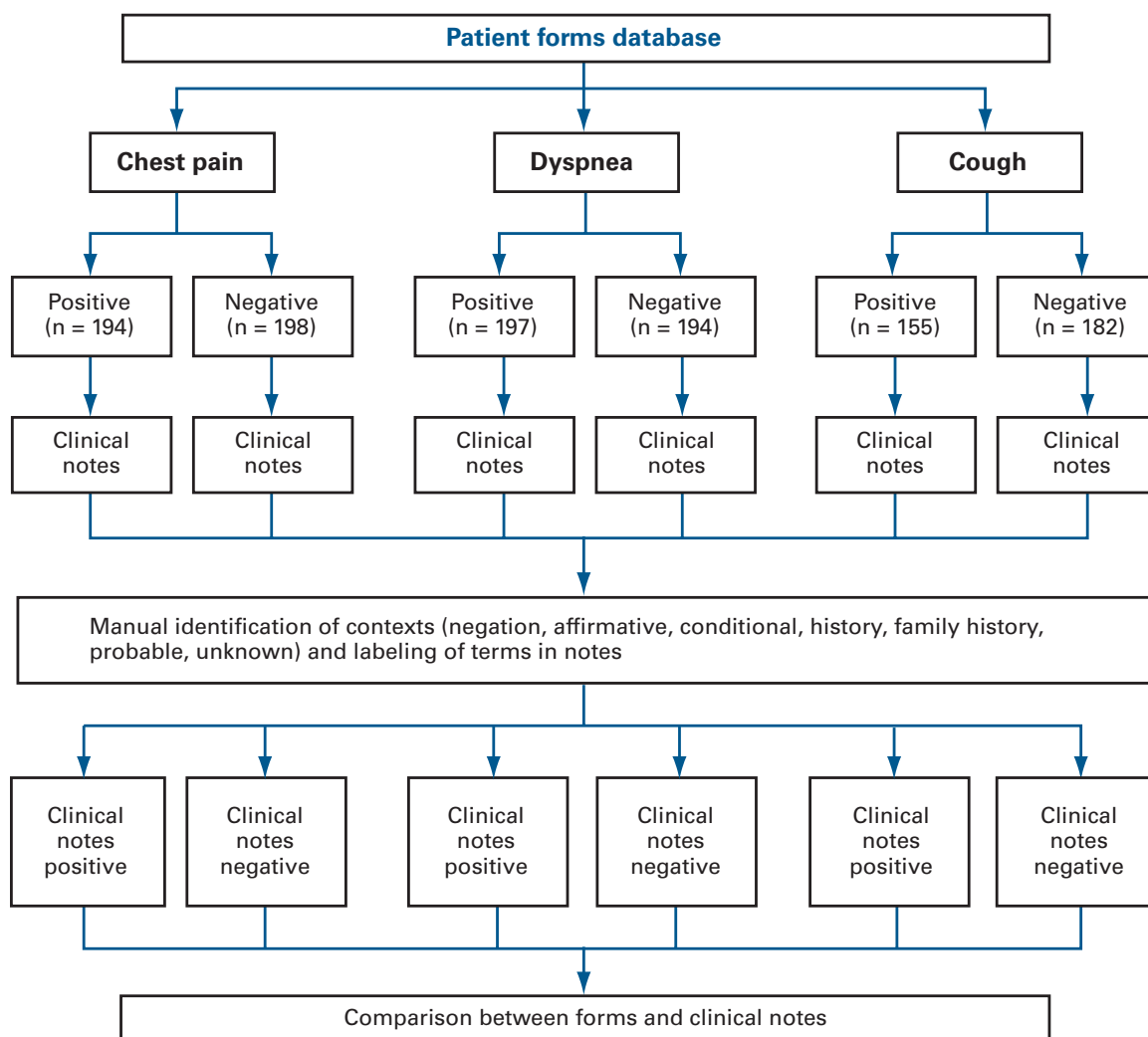
Study Population

We used a subset of the database of patient-provided in-

■ **Figure 1.** "Systems Review" Part of the Mayo Clinic Patient-provided Information Form

Systems Review		
14	Fill in the circle to the left of each symptom which you wish to call to the attention of your health care provider. Select No Symptoms if you have not experienced any of the listed symptoms. Select Other Symptom(s) if the symptom you wish to report is not listed.	
<input type="radio"/>	fevers	<input type="radio"/>
<input type="radio"/>	enlarged lymph glands	<input type="radio"/>
<input type="radio"/>	nipple discharge	<input type="radio"/>
<input type="radio"/>	breast lump	<input type="radio"/>
<input type="radio"/>	skin rash/skin sores	<input type="radio"/>
<input type="radio"/>	change in sexual drive or performance	<input type="radio"/>
<input type="radio"/>	unusual bruising	<input type="radio"/>
<input type="radio"/>	change in mole or skin spot	<input type="radio"/>
<input type="radio"/>	headaches	<input type="radio"/>
<input type="radio"/>	seizures	<input type="radio"/>
<input type="radio"/>	slurred speech	<input type="radio"/>
<input type="radio"/>	unusual thirst	<input type="radio"/>
<input type="radio"/>	hoarseness	<input type="radio"/>
<input type="radio"/>	double vision	<input type="radio"/>
<input type="radio"/>	sudden loss of vision	<input type="radio"/>
<input type="radio"/>	vision problems	<input type="radio"/>
<input type="radio"/>	shortness of breath	<input type="radio"/>
<input type="radio"/>	coughing	<input type="radio"/>
<input type="radio"/>	wheezing	<input type="radio"/>
<input type="radio"/>	"black outs" or loss of consciousness	<input type="radio"/>
<input type="radio"/>	awakened with shortness of breath	<input type="radio"/>
<input type="radio"/>	hearing loss	<input type="radio"/>
<input type="radio"/>	light-headedness	<input type="radio"/>
<input type="radio"/>	sinus congestion	<input type="radio"/>
<input type="radio"/>	loss of appetite	<input type="radio"/>
<input type="radio"/>	coughing up phlegm	<input type="radio"/>
<input type="radio"/>	coughed up blood	<input type="radio"/>
<input type="radio"/>	swelling in the legs or feet	<input type="radio"/>
<input type="radio"/>	cramping pain in leg muscles when walking	<input type="radio"/>
<input type="radio"/>	chest pain	<input type="radio"/>
<input type="radio"/>	chest pressure	<input type="radio"/>
<input type="radio"/>	rapid or fluttering heart beats	<input type="radio"/>
<input type="radio"/>	difficulty swallowing	<input type="radio"/>
<input type="radio"/>	heartburn	<input type="radio"/>
<input type="radio"/>	nausea and/or vomiting	<input type="radio"/>
<input type="radio"/>	constipation	<input type="radio"/>
<input type="radio"/>	diarrhea	<input type="radio"/>
<input type="radio"/>	blood in stool	<input type="radio"/>
<input type="radio"/>	changes in your stool characteristics	<input type="radio"/>
<input type="radio"/>	frequent urination	<input type="radio"/>
<input type="radio"/>	burning or painful urination	<input type="radio"/>
<input type="radio"/>	difficulty starting urination	<input type="radio"/>
<input type="radio"/>	uncontrolled urge to urinate	<input type="radio"/>
<input type="radio"/>	blood in urine	<input type="radio"/>
<input type="radio"/>	leaking urine	<input type="radio"/>
<input type="radio"/>	fatigue	<input type="radio"/>
<input type="radio"/>	abdominal (belly) pain or cramping	<input type="radio"/>
<input type="radio"/>	pain or stiffness in joints	<input type="radio"/>
<input type="radio"/>	joint swelling	<input type="radio"/>
<input type="radio"/>	muscle pain/stiffness	<input type="radio"/>
<input type="radio"/>	back pain/stiffness	<input type="radio"/>
<input type="radio"/>	weakness in arms or legs	<input type="radio"/>
<input type="radio"/>	numbness or shooting pain in hands, arms, legs or feet	<input type="radio"/>
<input type="radio"/>	noticed tendency to fall easily	<input type="radio"/>
<input type="radio"/>	weight gain of more than 10 pounds	<input type="radio"/>
<input type="radio"/>	weight loss of more than 10 pounds	<input type="radio"/>
<input type="radio"/>	heavy snoring	<input type="radio"/>
<input type="radio"/>	sleep difficulty	<input type="radio"/>
<input type="radio"/>	excessive daytime drowsiness	<input type="radio"/>
<input type="radio"/>	irregular breathing during sleep	<input type="radio"/>
<input type="radio"/>	felt sad most of the time	<input type="radio"/>
<input type="radio"/>	felt restless or irritable	<input type="radio"/>
<input type="radio"/>	felt anxious or nervous	<input type="radio"/>
<input type="radio"/>	had little interest or pleasure in relationships, or activities	<input type="radio"/>
<input type="radio"/>	had difficulty concentrating	<input type="radio"/>
<input type="radio"/>	had recurring thoughts of death or suicide	<input type="radio"/>
<input type="radio"/>	Other symptom(s) not listed	<input type="radio"/>
<input type="radio"/>	No symptoms	<input type="radio"/>

■ **Figure 2.** Diagram Illustrating the Study Design



formation entered during 6 months between January 1, 2006, and June 30, 2006, as the convenience sample consisting of 121,891 patients who filled out the form at least once during the 6-month period. The patients in this sample were not restricted to any geographic location and represent general Mayo Clinic ambulatory and hospitalized populations. We selected records that indicated either a positive or a negative response to questions about chest pain, chest pressure, shortness of breath/dyspnea, and cough. (These codes are internal identifiers provided here for reference and have no relation to any standardized nomenclature of medical concepts.) The analysis presented here is based on a combined symptom for chest pain that includes both the chest pain and chest pressure responses.

Manual Verification of Symptoms

We manually examined clinical notes containing evidence of chest pain, dyspnea, and cough according to the

procedure illustrated in **Figure 2**. For each of the symptoms, we randomly selected 200 patients who marked the symptom on at least 1 of their forms and 200 patients who did not mark the symptom on any of their forms. No restrictions other than age ≥ 18 years were applied. Only records of patients who had 1 or more clinical notes dictated during the study period (January 1, 2006, through June 30, 2006) were used in this study.

Each clinical note was searched automatically to identify and electronically mark the terms representing each of the symptoms and their orthographic variants and synonyms using search keywords arranged into natural language queries (**Figure 3**). The keywords and methods were similar to those previously reported.¹⁵ For example, a natural language query for “chest pain” identifies portions of clinical note text where 1 of the terms describing PAIN (eg, pressure) either precedes or follows 1 of the terms describing the LOCATION (eg,

■ **Figure 3.** Natural Language Queries Used to Search Clinical Reports

Chest pain	<p>PAIN = (ache, aching, hurt, hurting, pain, pressure, discomfort, tightness, soreness, heaviness, squeezing, crushing, compression, tension, distress)</p> <p>LOCATION = (chest, breast,^a thorax, rib cage, substernal, thoracic)</p> <p>QUERY = (location pain) or (pain location)</p>
Dyspnea	<p>DIFFICULTY = (difficulty, difficult, difficulties, hard, labored, strenuous, problem, gasping, short, shortness, out)</p> <p>BREATH = breathing, breathes, breathe, breath)</p> <p>ONEWORD = breathlessness, breathless, dyspnea, shortness of breath)</p> <p>QUERY = (Difficulty Breath) or (Breath Difficulty) or (Oneword)</p>
Cough	<p>QUERY = (coughing, coughs, cough)</p>

^aWe recognize that, clinically, breast pain is distinct from chest pain; however, this study focuses on the patient presentation of symptoms. Thus, including “breast pain” as a search term is motivated by the desire to capture clinician’s quoting of the patient as in the following real example:

“I suspect that the patient’s breast pain is actually due to significant chest wall pain ...” In our sample, there were only 15 instances of breast pain; therefore, inclusion or exclusion of this term would not significantly affect the conclusions of this study.

chest). Thus, this query will find either “chest pressure” or “pressure in the chest.” Subsequently, the text of the notes was manually examined by 2 nurse abstractors to determine the context in which each term appeared. The range of possible context labels that the abstractors could choose is displayed in column 1 of **Table 1**. The “conditional” context label was selected when the term was mentioned in a conditional context (eg, “I recommended nitroglycerin if he should develop chest pain”). Terms manually identified as “negated,” “family history,” “conditional,” or “unknown” were excluded from subsequent analysis.

In addition to identification of the context in which the query terms appeared, the abstractors manually identified all symptoms of chest pain, dyspnea, and cough in a random sample of 100 clinical notes. These manually identified symptoms were compared with those identified with automated natural language queries to determine their sensitivity.

Angina refers to a complex of symptoms, 1 of which may be chest pain or discomfort. Similarly to dyspnea, orthopnea and paroxysmal nocturnal dyspnea (PND) refer to a specific kind of shortness of breath. The patient-provided information form used for this study contains a question that covers this type of dyspnea (see “awakened with shortness of breath” in Figure 1). For the sensitivity analysis, we added the terms “angina,” “orthopnea,” and “PND” to the natural language queries to determine whether these related terms had a measurable effect on the agreement between symptoms reported by patients and care providers.

Statistical Analysis

For the primary analysis and the symptom-reporting consistency, data were analyzed in terms of positive and negative agreement rather than sensitivity and specificity because neither the patient self-report nor the clinical notes could be considered a perfect criterion standard. Traditional kappa statistics are sensitive to imbalances in the marginal totals of 2 × 2 comparisons.²¹ Positive and negative agreement measures have been proposed as a way to ensure the correct interpretation of kappa values²² and have been used to assess the agreement between patient-reported information and the medical record.²³ The computation of these measures is illustrated in **Table 2**, which represents a 2 × 2 table containing the counts where the patient responses were in concordance or discor-

■ **Table 1.** Distribution of Contexts in Which Chest Pain, Dyspnea, and Cough Occurred

Context	Symptom, No. (%)		
	Chest Pain	Dyspnea	Cough
Affirmative	1178 (62)	1459 (65)	715 (57)
Negated	543 (29)	505 (23)	226 (18)
Probable	5 (<1)	0	0
History	79 (4)	95 (4)	63 (5)
Family history	0	1 (<1)	3 (<1)
Conditional	78 (4)	162 (7)	243 (19)
Unknown	18 (<1)	3 (<1)	2 (<1)
Total	1901 (100)	2225 (100)	1252 (100)

■ **Table 2.** Natural Language Queries Used to Search Clinical Reports

		Patient-reported Symptom		
		Positive	Negative	Marginal total
Clinical Notes Symptom	Positive	<i>a</i>	<i>b</i>	<i>a + b</i>
	Negative	<i>b</i>	<i>d</i>	<i>c + d</i>
	Marginal total	<i>a + c</i>	<i>b + d</i>	<i>N = a + b + c + d</i>

dance with the symptoms found in clinical notes for the same patient. Thus, the positive agreement is a ratio of the concordances in positive responses to the difference between the concordances in positive and negative responses added to the total number of samples according to the following formula: $100 \times (2a/[N + (a - d)])$. The negative agreement is a ratio of the concordances in negative responses to the difference between the concordances in positive and negative responses subtracted from the total number of samples according to the following formula: $100 \times (2d/[N - (a - d)])$.²² In addition we also report on the false-negative rate, which in our case shows the proportion of times when the patient reported a symptom but the symptom did not appear in the physician's note [ie, false-negative rate = $100 \times (c/(a + c))$]. We report standard measures of sensitivity and specificity instead of positive and negative rates for the assessment of reliability of identification of negation by the NLP system because we used a manually created reference standard. Stratified analyses in sex and age subgroups were performed as well.

RESULTS

Of the 121,891 patients who filled out a patient-provided information form during 6 months between January 1, 2006, and June 30, 2006, 6569 patients (5.39%; 95% confidence interval [CI] = 5.26, 5.52) reported chest pain, 6166 patients (5.06%; 95% CI = 4.94, 5.18) reported chest pressure, 13,924 patients (11.42%; 95% CI = 11.24, 11.60) reported dyspnea, and 11,670 patients (9.57%; 95% CI = 9.41, 9.74) reported cough. Combining chest pain and chest pressure as syn-

onymous terms yielded 10,518 patients (8.63%; 95% CI = 8.47, 8.79) who reported either chest pain or pressure.

Random sampling resulted in a study population of 1119 patients, with 373 patients in the chest pain group, 391 patients in the dyspnea group, 337 patients in the cough group, and 18 patients (1.6%) in more than 1 group. Of these 1119 patients, 499 (44.6%) were male, 127 (11.3%) were between age 18 and 34 years, 185 (16.5%) were between age 35 and 49 years, 337 (30.1%) were between age 50 and 64 years, and 470 (42.0%) were age 65 years or older. Twelve percent of the clinical notes for the 1119 patients were marked as a "Hospital Admission" or "Hospital Dismissal" note, thus representing an estimate of the notes from the hospitalized

population. Only 6 patients filled out more than 1 patient-provided information form during the study period. Overall, the clinical notes in this study sample represent a variety of clinical specialties including primary care (20%), cardiovascular services (13%), physical medicine and rehabilitation (5%), hematology (4%), neurology (4%), endocrinology (4%), surgery (4%), psychology/psychiatry (4%), gastroenterology (3%), and emergency medicine (2%), among other less prevalent specialties.

Validation of Natural Language Queries

The distribution of contexts illustrated in Table 1 shows that the terms denoting chest pain, dyspnea, and cough occur in negated contexts between 18% and 29% of the time. These observations are consistent with those of Chapman et al²⁴; in their study, 27% of the findings and diseases were manually identified as negated in the process of creating a reference standard. We examined the precision of the NLP methodology for identification of the context (negated vs affirmative) in which symptoms were mentioned in the free text of clinical reports. Complete results are presented in Table 3. The sensitivity of the NLP negation algorithm was 83% (95% CI = 78, 88), 84% (95% CI = 80, 87), and 88% (95% CI = 84, 90) for cough, dyspnea, and chest pain, respectively. The specificity was 91% (95% CI = 90, 93), 93% (95% CI = 91, 94), and 92% (95% CI = 91, 94), for cough, dyspnea, and chest pain, respectively. The kappa values were 0.69 (95% CI = 0.63, 0.74) for cough, 0.74 (95% CI = 0.71, 0.78) for dyspnea, and 0.78 (95% CI = 0.75, 0.81) for chest pain.

Patient-reported Symptoms in the EMR

In the random samples of 100 notes for each of the 3 symptoms, the abstractors identified 34 mentions of chest pain, 46 mentions of dyspnea, and 24 mentions of cough. Of the 34 chest pain mentions, 31 were also identified by the natural language queries yielding 91% (95% CI = 82, 100) sensitivity. Of the 46 dyspnea mentions, 45 were identified by the queries yielding 98% (95% CI = 94, 100) sensitivity. All 24 mentions of cough also were identified by automatic queries.

Agreement Between Clinical Notes and Patient-provided Information Forms

The results are summarized in **Table 4**. Overall, the positive agreement for chest pain, dyspnea, and cough was 74, 70, and 63, respectively, and the negative agreement was 78, 76, and 75, respectively. The kappa values were 0.52 (95% CI = 0.43, 0.60) for chest pain, 0.46 (95% CI = 0.37, 0.54) for dyspnea, and 0.38 (95% CI = 0.28, 0.48) for cough. Additional analysis was carried out within different sex and age strata to determine whether the concordance and discordance are influenced by these factors. Positive agreement was slightly lower for female patients across all 3 symptoms, and tended to be lower for younger subgroups. Negative agreement varied and tended to be lower for males and older subgroups. Kappa values ranged from 0.44 to 0.55 for chest pain, 0.27 to 0.52 for dyspnea, and 0.31 to 0.56 for cough.

Including nonnegated mentions of angina resulted in identification of 5 additional patients who self-reported chest pain and 3 additional patients who did not, yielding positive agreement of 76 and negative agreement of 75. Including nonnegated mentions of orthopnea and PND resulted in identifying 1 additional patient who did not self-report dyspnea, yielding positive agreement of 70 and negative agreement of 75.

Distribution of Clinical Specialties for Discordant Results

We determined the distributions of clinical services

Table 3. NLP Context (Negation) Identification Algorithm Validation Results

	No. Correctly Identified	Sensitivity, % (95% CI)	No. Correctly Identified	Specificity, % (95% CI)	Kappa (95% CI)
Automatic NLP Annotation	Manual Annotation				
	CHEST PAIN				
	Negated (n = 461)		Affirmative (n = 1188)		
	404	88 (84, 90)	1095	92 (91, 94)	0.78 (0.75, 0.81)
	DYSYPNEA				
	Negated (n = 460)		Affirmative (n = 1517)		
	385	84 (80, 87)	1406	93 (91, 94)	0.74 (0.71, 0.78)
	COUGH				
	Negated (n = 206)		Affirmative (n = 876)		
171	83 (78, 88)	798	91 (90, 93)	0.69 (0.63, 0.74)	
CI indicates confidence interval; NLP, natural language processing.					

that reflect the clinical specialty of care providers for the discordant results for each of the 3 symptoms (**Table 5**). The distributions were computed by counting the number of clinical notes originating in each of the top 10 most prevalent clinical services. Consistent with the overall distribution of clinical services in our sample, cardiovascular services and primary care/family medicine were the top 2 most prevalent specialties for the patients in the discordant groups. The majority of the notes for patients reporting chest pain that was not documented in the EMR were distributed among primary care, cardiovascular services, and endocrinology. A similar distribution was found for dyspnea. For cough, the majority of the notes were distributed between primary care, cardiovascular services, and physical medicine and rehabilitation services. The majority of the notes for patients who had documentation of chest pain in the EMR but did not report it on patient-provided information forms were distributed among primary care, psychology/psychiatry, hematology, and cardiovascular services. A similar distribution was found for cough, with a large proportion of notes from the hematology service. For dyspnea, the majority of the notes originated from primary care and cardiovascular services.

DISCUSSION

Electronic medical records offer a way to make improvements in the quality and safety of patient care by addressing the information challenge of organizing and making patient charts accessible and interoperable across healthcare providers.²⁵ The primary and most important use of the EMR is to

■ **Table 4.** Positive and Negative Agreement Results as Well as Kappa Values for the Comparison Between Patient Self-report and Care Providers' Notes

Analysis	Chest Pain (n = 373)			Dyspnea (n = 391)			Cough (n = 337)		
	Ppos	Pneg	Kappa (95% CI)	Ppos	Pneg	Kappa (95% CI)	Ppos	Pneg	Kappa (95% CI)
Unstratified	74	78	0.52 (0.44, 0.60)	70	76	0.46 (0.37, 0.54)	63	75	0.38 (0.28, 0.48)
Stratified by sex									
Male	76	74	0.51 (0.38, 0.64)	75	76	0.52 (0.40, 0.65)	63	74	0.38 (0.24, 0.53)
Female	72	76	0.48 (0.37, 0.60)	64	75	0.40 (0.28, 0.52)	61	75	0.37 (0.24, 0.51)
Age, y									
18-34	65	79	0.44 (0.17, 0.70)	62	85	0.48 (0.21, 0.75)	74	81	0.56 (0.27, 0.84)
35-49	78	76	0.55 (0.37, 0.73)	49	76	0.27 (0.02, 0.52)	45	77	0.23 (0.04, 0.50)
50-64	76	74	0.51 (0.35, 0.66)	69	75	0.45 (0.28, 0.61)	50	79	0.31 (0.12, 0.49)
≥65	71	76	0.47 (0.32, 0.62)	75	72	0.47 (0.34, 0.60)	69	66	0.37 (0.23, 0.52)

CI indicates confidence interval; Pneg, negative agreement; Ppos, positive agreement.

facilitate and streamline care delivery; however, its secondary uses for clinical research and quality and safety assurance also are important.²⁶ From the standpoint of the primary use of the EMR, we currently do not have the data to determine whether the discordances between the symptoms reported by patients and those documented by care providers had any significant clinical consequences. However, we do present evidence that effective secondary uses of the EMR require examination of the communication approaches of care providers and patients.

Our findings indicate substantial discordance between patient reporting and care provider documentation of the symptoms. The 2 sources of the symptoms may complement each other and have implications for clinical studies and quality measurements that rely on the medical record for identification of symptoms. For example, to ensure completeness of identifying and recruiting participants with specific symptoms for clinical studies, it may be necessary to use the information reported by patients on self-entry forms in addition to other sources such as the clinical notes. It also may be necessary to use both sources of information for applications such as postmarketing medication safety surveillance, where it may be necessary to detect alarming trends in symptoms to prevent more serious adverse drug reactions. The current study is the first step toward understanding the implications of symptom documentation practices for both primary and secondary uses of the EMR.

Validation of Natural Language Queries

The results in Table 3 indicate that NLP is a valid tool for

identification of symptoms that occur in negated contexts in the free text of clinical reports. These results are promising because they show that the NLP methods used in this pilot study may be applied retrospectively to an existing cohort of patients with known outcomes to determine whether there was an association between the outcomes and symptom documentation. We also found that the methods used to identify the mentions of chest pain, dyspnea, and cough are highly sensitive. Automated natural language queries missed chest “tightness,” “sense of bruising in the lateral and posterior chest wall,” “anginal symptoms,” and “dyspneic.” Capturing these types of cases would require only minor modifications to the algorithms used for term identification. Inclusion of the terms “angina,” “orthopnea,” and “PND” resulted in very minor changes in the positive and negative agreement, not affecting the conclusions of the primary analysis. It is likely that these more specialized terms are used in conjunction with more generic terms such as “chest pain” and “dyspnea” and thus do not yield additional patients.

Agreement Between Clinical Notes and Patient-provided Information Forms

In a previous study of the Mayo Clinic EMR, St. Sauver et al showed that positive reports of cardiovascular disease risk factor information (blood pressure; triglycerides; cholesterol; history of heart rhythm, heart valve, and arterial problems) by patients are largely inaccurate, while negative reports were unlikely to be noted on the medical record by care providers, thus showing a low level of agreement between patient-reported risk factors and physician documentation.²³ The study

Patient-reported Symptoms in the EMR

Table 5. Proportion of Notes Generated by Various Services for Patients With Symptoms Noted Either on Self-report and/or in the Text of the EMR^a

Result	Clinical Specialty, %										Total
	PSY	PMR	GI	PC	HEM	NEURO	CV	EDPT	SURG	ENDO	
Chest Pain											
TP	5	18	3	29	6	4	25	3	4	3	100
FN	2	4	5	29	2	8	24	1	8	17	100
TN	6	8	9	39	3	6	5	2	14	9	100
FP	23	1	0	24	18	5	16	2	7	3	100
Dyspnea											
TP	2	2	2	29	6	5	41	4	4	6	100
FN	12	6	8	23	3	13	13	2	3	17	100
TN	7	7	5	42	1	9	6	3	15	4	100
FP	1	4	2	46	2	1	27	5	4	8	100
Cough											
TP	3	3	6	35	8	3	31	6	1	4	100
FN	7	21	3	31	2	12	19	1	1	3	100
TN	8	10	5	37	3	8	12	2	8	7	100
FP	8	3	1	21	35	3	18	4	7	1	100

CV indicates cardiovascular; EDPT, emergency medicine; EMR, electronic medical record; ENDO, endocrinology; FN, false negative (patient reported symptom but not documented); FP, false positive (patient did not report a symptom but the symptom was found in documentation); GI, gastroenterology; HEM, hematology; NEURO, neurology; PC, primary care; PMR, physical medicine & rehabilitation; PSY, psychology/psychiatry; SURG, surgery; TN, true negative (patient did not report a symptom and no documentation found); TP, true positive (patient reported symptom and it was documented).

^aThe percentages highlighted in blue indicate the top 2 services with highest proportion of notes for discordant results.

by St. Sauver et al, however, did not address symptom reporting and documentation, as that was not in the scope of their research questions. Although patient self-report is a secondary source of information on lab results and past medical diagnoses, with physician-documented medical record being the primary source (thus perhaps contributing to the low agreement found by St. Sauver et al²³), the situation is reversed for patient symptoms in cases where the patient is the primary source and the care provider is the secondary source. The results of the current study are consistent with those reported by St. Sauver et al and contribute to constructing a more complete picture of clinical process and documentation.

Although we found discordance between patient-reported and physician-reported information, the causes of such discordances can only be hypothesized. Indeed, discordance may reflect the nature of the patient-physician interaction or other factors such as clinical specialty. For example, for discordances on chest pain reporting and documentation, one might expect to find the majority of clinical notes originating from noncardiovascular services. Our data shows that although this is true, a relatively large proportion of the notes (24%) originated from cardiovascular services. The results

are similar for dyspnea and cough. The majority of the notes for patients who reported chest pain not documented in the EMR originated in 3 services: cardiovascular, endocrinology, and primary care. Direct examination of the patient-physician interaction would be necessary to determine the reasons for the discordance; however, the reasons may differ by specialty. Furthermore, the discordance also may reflect the specifics of the documentation system. The Mayo Clinic uses an integrated EMR where all notes and forms (including the patient-provided information forms) from all services for a given patient are available to any clinician working with the patient. Clinicians may decide to avoid repeating a symptom in their clinical note that is already documented elsewhere. Thus, the nature and the consequences of the interaction between the individual characteristics of the documentation systems, the care providers, and the patients warrant further investigation within the framework of patient-centered care.

Of interest are those instances where the patient reported a symptom not documented in the record (ie, false negatives); 31% of patients with chest pain, 38% with dyspnea, and 45% with cough did not have a positive mention of the symptom in the clinical notes. We could hypothesize

Take-away Points

Three symptoms (chest pain, dyspnea, and cough) reported by patients were compared with symptoms identified by natural language processing of the text of clinical notes from care providers.

- The electronic medical record documentation of symptoms differed from the symptoms self-reported by patients.
- Research studies that rely on using symptom information must take this discordance into account.
- This discordance also may have a potential impact on quality of care, patient safety, and outcomes.

The use of the state-of-the-art EMR environment at the Mayo Clinic is a distinct strength of this study, enabling an investigation of the discrepancies between patient-reported symptoms and provider documentation. This study presents a novel approach that relies on automated and thus easily scalable investigation of the free text contained in provider documentation; this approach can be

used in future large-scale studies of the interaction between the patient and the care provider. Despite the fact that some natural language term variants (particularly those due to misspellings and nonstandard descriptions) are likely to be missed by the automated natural language queries used in this project, our methodology is highly sensitive for identification of chest pain, dyspnea, and cough.

that the outcomes of such patients may differ from those for whom symptoms were documented in the EMR. Conversely, when the patient does not report a symptom documented in the EMR, such discordance may reflect differential elicitation of symptoms by care providers by specialty. As these considerations remain speculative, further study is needed to examine the reasons for discordances and to determine whether these affect clinical outcomes. The present study shows that the discordances exist, which has implications for clinical research that relies on the EMR to ascertain symptoms.

Limitations and Strengths

The generalizability of the study depends on the availability of the EMR; however, the adoption of EMRs across the United States is growing.²⁷ Many medical centers already have their patient notes in electronic form, and systems for accessing the information in the notes can be constructed even without a fully integrated EMR system. Although the data sources used in this study are specific to the Mayo Clinic, collecting patient-reported information before a visit and physician documentation of the visit are standard healthcare practices, and similar resources are in use at other institutions across the United States. Thus, our results are generalizable to other institutions that are equipped with an EMR and can guide the improvement of existing EMR systems by suggesting new avenues for improved elicitation and capture of patient-specific information central to patient care and research.

In this study, we did not look for an association between the literacy level or proficiency in the English language. We did not have these data available for this study, but we recognize that these are important variables to consider because the forms require a certain level of knowledge of the English language. Care providers' demographic characteristics (as well as the interaction of these characteristics with patient characteristics) also may be important predictors of concordances and discordances in documentation. These characteristics were not available for the current study but will be considered in future work.

used in future large-scale studies of the interaction between the patient and the care provider. Despite the fact that some natural language term variants (particularly those due to misspellings and nonstandard descriptions) are likely to be missed by the automated natural language queries used in this project, our methodology is highly sensitive for identification of chest pain, dyspnea, and cough.

Acknowledgments

We thank Ellen Koespell and Diane Batzel for their dedicated efforts in annotating clinical notes.

Author Affiliations: From the Department of Pharmaceutical Care and Health Systems, University of Minnesota (SVP), Minneapolis; the Department of Research and Evaluation, Kaiser Permanente (SJJ), Pasadena, CA; and the Department of Health Sciences Research (CGC) and the Division of Cardiovascular Diseases (VLR), Mayo Clinic, Rochester, MN.

Funding Source: This work was supported by Public Health Service grants RO1-72435, GM14321, and AR30582, and the NIH Roadmap Multidisciplinary Clinical Research Career Development Award Grant (K12/NICHD-HD49078).

Author Disclosure: The authors (SVP, SJJ, CGC, VLR) report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (SVP, SJJ, CGC); acquisition of data (SVP, VLR); analysis and interpretation of data (SVP, SJJ, VLR); drafting of the manuscript (SVP, VLR); critical revision of the manuscript for important intellectual content (SVP, SJJ, CGC); statistical analysis (SVP, VLR); provision of study materials or patients (SVP); obtaining funding (SVP); administrative, technical, or logistic support (SVP, CGC); and supervision (SVP, CGC, VLR).

Address correspondence to: Serguei V. Pakhomov, PhD, Department of Pharmaceutical Care and Health Systems, University of Minnesota, 308 Harvard St, SE, 7-125F Weaver-Densford Hall, Minneapolis, MN 55401. E-mail: pakh0002@umn.edu.

REFERENCES

1. Hing E, Chery DK, Woodwell DA. National Ambulatory Medical Care Survey: 2004 summary. *Adv Data*. 2006 Jun 23;(374):1-33.
2. Lamberg L. New mind/body tactics target medically unexplained physical symptoms and fears. *JAMA*. 2005;294(17):2152-2154.
3. Committee on Quality of Health Care in America, Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: Institute of Medicine; 2001.
4. Sager N, Lyman M, Bucknall C, Nhan N, Tick LJ. Natural language processing and the representation of clinical data. *J Am Med Inform Assoc*. 1994;1(2):142-160.
5. Friedman C. A broad-coverage natural language processing system. In: Overhage JM, ed. *Proceedings of the 2000 AMIA Annual Symposium*. Bethesda, MD: American Medical Informatics Association; 2000:270-274.

6. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. *J Am Med Inform Assoc.* 2005;12(4):448-457.
7. Javitz HS, Ward MM, Watson JB, Jaana M. Cost of illness of chronic angina. *Am J Manag Care.* 2004;10(11 suppl):S358-S369.
8. Hemingway H, McCallum A, Shipley M, Manderbacka K, Martikainen P, Keskimaki I. Incidence and prognostic implications of stable angina pectoris among women and men. *JAMA.* 2006;295(12):1404-1411.
9. Thom T, Haase N, Rosamond W, et al; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2006 update. A report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation.* 2006;113(6):e85-e151. Epub 2006 January 11.
10. Philpott S, Boynton PM, Feder G, Hemingway H. Gender differences in descriptions of angina symptoms and health problems immediately prior to angiography: the ACRE study. Appropriateness of Coronary Revascularisation study. *Soc Sci Med.* 2001;52(10):1565-1575.
11. Schulman KA, Berlin JA, Harless W, et al. The effect of race and sex on physicians' recommendations for cardiac catheterization. *N Engl J Med.* 1999;340(8):618-626.
12. Hemingway H, Shipley M, Britton A, Page M, Macfarlane P, Marmot M. Prognosis of angina with and without a diagnosis: 11 year follow up in the Whitehall II prospective cohort study. *BMJ.* 2003;327(7420):895.
13. Pakhomov SV, Coden A, Chute CG. Developing a corpus of clinical notes manually annotated for part-of-speech. *Int J Med Inform.* 2006;75(6):418-429. Epub 2005 September 19.
14. Ford EW, Menachemi N, Phillips MT. Predicting the adoption of electronic health records by physicians: when will health care be paperless? *J Am Med Inform Assoc.* 2006;13(1):106-112.
15. Pakhomov SS, Hemingway H, Weston SA, Jacobsen SJ, Rodeheffer R, Roger VL. Epidemiology of angina pectoris: role of natural language processing of the medical record. *Am Heart J.* 2007;153(4):666-673.
16. An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. GUSTO investigators. *N Engl J Med.* 1993;329(10):673-682.
17. Randomized trial of intravenous streptokinase, oral aspirin, both or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. Second International Study of Infarct Survival Collaborative Group. *Lancet.* 1988;2(8607):349-360.
18. Arciero TJ, Jacobsen SJ, Reeder GS, et al. Temporal trends in the incidence of coronary disease. *Am J Med.* 2004;117(4):228-233.
19. Wang SJ, Ohno-Machado L, Fraser HS, Kennedy RL. Using patient-reportable clinical history factors to predict myocardial infarction. *Comput Biol Med.* 2001;31(1):1-13.
20. Adams K, Corrigan JM, eds, Committee on Identifying Priority Areas for Quality Improvement. *Priority Areas for National Action: Transforming Health Care Quality.* Washington, DC: Institute of Medicine; 2003.
21. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol.* 1990;43(6):543-549.
22. Cicchetti DV, Feinstein AR. High agreement but low kappa:II. Resolving the paradoxes. *J Clin Epidemiol.* 1990;43(6):551-558.
23. St Sauver JL, Hagen PT, Cha SS, et al. Agreement between patient reports of cardiovascular disease and patient medical records. *Mayo Clinic Proc.* 2005;80(2):203-210.
24. Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. *J Biomed Inform.* 2001;34(5):301-310.
25. Dove JT. President's page: achieving the heights of quality with electronic health records. *J Am Coll Cardiol.* 2007;49(13):1498-1500.
26. Hersh WR. Adding value to the electronic health record through secondary use of data for quality assurance, research, and surveillance. *Am J Manag Care.* 2007;13(6, part 1):277-278.
27. Medical Group Management Association. Crunching numbers: research finds low EHR adoption rates for physician groups. 2005. <http://www.mgma.com/article.aspx?id=2180>. Accessed April 25, 2008. ■