J Digit Imaging (2015) 28:272–282
DOI 10.1007/s10278-014-9751-7

Evaluating the Referring Physician's Clinical History and Indication as a Means for Communicating Chronic Conditions That Are Pertinent at the Point of Radiologic Interpretation

Piotr Obara · Merlijn Sevenster · Adam Travis · Yuechen Qian · Charles Westin · Paul J. Chang

Published online: 23 December 2014

Abstract The clinical history and indication (CHI) provided with a radiological examination are critical components of a quality interpretation by the radiologist. A patient's chronic conditions offer the context in which acute symptoms and findings can be interpreted more accurately. Seven pertinent (potentially diagnosis altering) chronic conditions, which are fairly prevalent at our institution, were selected. We analyze if and how in 140 CHIs there was mention of a patient's previously reported chronic condition and if and how the condition was subsequently described in the radiology report using a four-item scheme (Mention/Specialization, Generalization, Common comorbidity, No mention). In 40.7 % of CHIs, the condition was rated Mention/Specialization. Therefore, we reject our first hypothesis that the CHI is a reliable source for obtaining pertinent chronic conditions (≥90.0 %). Non-oncological conditions were significantly more likely rated No mention in the CHI than oncological conditions (58.7 versus 8.3 %, $P<0.0001$). Stat cases were significantly more frequently No mention than non-stat cases (60.0 versus 31.3 %, $P=0.0134$). We accept our second hypothesis that the condition's rating in the CHI is significantly correlated with its rating of the final radiology report (χ^2 test, $P<0.00001$). Our study demonstrates an alarming lack of communication of pertinent medical information to the radiologist, which may negatively impact interpretation quality. Presenting automatically aggregated patient information to the radiologist may be a potential avenue for improving interpretation and adding value of the radiology department to the care chain.

Keywords Clinical history and indication · Radiology workflow · Quality improvement · EMR · CPOE

Introduction

Due to technological innovations, Digital Imaging and Communications in Medicine (DICOM) imaging studies can be distributed throughout the health care enterprise, enabling referring physicians to access imaging data whenever it is most valuable and pertinent to them [1]. Distribution beyond the health care enterprise has led to outsourcing of radiological interpretation to "nighthawk" parties. These trends continue to separate the referring physician from the radiology department and reduce its perceived value to a commodity service provider [2, 3].

In the meantime, governmental regulations pressure health care organizations to transform into accountable care organizations (ACOs) that are reimbursed based on patient outcome and not on, for instance, interventions and imaging studies performed. In an ACO environment, every imaging study becomes a cost center and the radiology department will be expected to eliminate unnecessary and reduce low-yield imaging studies [4, 5].

The technological and regulatory trends combined may synergistically propel the commoditization of radiology and may reduce the need for highly educated and subspecialized radiologists. For radiology to survive in its current form and shape, the radiologist's role of commodity service provider

Electronic supplementary material The online version of this article (doi:10.1007/s10278-014-9751-7) contains supplementary material, which is available to authorized users.

P. Obara (✉) · A. Travis · C. Westin · P. J. Chang
Department of Radiology, University of Chicago, 5841 South Maryland Ave -MC 2026, Chicago, IL 60637, USA
e-mail: probara@gmail.com

M. Sevenster · Y. Qian
Philips Research North America, New York, NY, USA

should be contrasted to that of an expert consultant that is a valued player in the care chain.

Inasmuch as technological developments are commoditizing radiologists, they may also help reverse it, especially if these innovations improve patient outcome in an objective and quantifiable manner. Delivering pertinent patient information is one potential area in which novel technologies could make a difference and increase the perceived value of the radiological interpretation.

In today's health care system, the clinical history and indication (CHI) is oftentimes the only information provided by the referring physician to the radiologist. It has been shown that the CHI can affect the quality of the radiologist's exam interpretation [6, 7]. Pertinent and accurate information regarding the current symptoms and past medical history enable the radiologist to interpret imaging findings in the appropriate clinical context, leading to a more relevant differential diagnosis, a more useful report for the clinician, and ideally a better outcome for the patient.

We performed a retrospective analysis evaluating the quality of the CHI with regard to conveying a known chronic condition pertinent at the point of radiological interpretation. Our results quantify the quality of one (important) stream of clinical information from the referring physician to the radiologist. We test the following two hypotheses: (1) the CHI provided by the referring physician is a reliable source for obtaining potentially diagnosis-altering chronic conditions, i.e., it accurately mentions the known chronic condition in at least 90 % of the cases, and (2) if and how a pertinent chronic condition is mentioned in the clinical history section is correlated with if and how it is mentioned in the clinical indication.

Materials and Methods

Condition and Patient Selection

Seven conditions were selected by two board-certified radiologists and a senior resident that they considered fairly prevalent at the University of Chicago Hospitals and pertinent at the point of radiological interpretation, in the sense that they are potentially diagnosis altering. The selected conditions were as follows: astrocytoma (Ast); cirrhosis (Cir); Crohn's disease (Cro); specific head and neck cancers namely laryngeal, oral, and tongue cancer (HNC); human immunodeficiency virus (HIV); non-Hodgkin lymphoma (nHL); and systemic lupus erythematosus (SLE). Although patients can recover from some of the selected conditions, it was considered that knowing that a patient once suffered from any of the selected conditions was still pertinent at the point of any radiological interpretation at a later point in time. This study was conducted under IRB 11-0193-E and was exempted from institutional review board (IRB) review.

For each condition, 20 patients were selected (see below for selection procedure) who had the following:

- a *baseline study* confirming that the patient has the condition by stating this in the clinical history and/or conclusion section of the study's report; and
- a *follow-up study* of the same modality and body part (e.g., MRI BRAIN or CR CHEST) obtained at least 11 months after the baseline study.

For each follow-up study, we collected the CHI provided by the referring physician as well as the CHR.

Clinical History and Indication

At our institution, the CHI consists of two parts. The first is generated from the computerized order entry implementation (CPOE), at our institution Epic (Epic Systems) which requires at least one ICD-9 code present in the patient's EMR to be associated with the exam during order entry for justifying order appropriateness. This constitutes the first part of the CHI, is typeset in all caps, and always marked by the ICD-9 numerical codes, e.g., "LYMPHOMAS NEC EXTRANODAL/NOS [202.80]." After associating an ICD-9 code(s) with the exam, the ordering provider is able to enter additional symptoms or pertinent history as free text, which constitutes the second CHI component. The free-text fields are prefixed in CPOE by "Clinical question to be answered:" and "Signs and symptoms:", for example "Clinical question to be answered: History of Follicular Non Hodgkin's Lymphoma, Signs and Symptoms: Evaluate for Continued CR." This CPOE process remained unchanged during the study period.

The combination of a textual rendering of the ICD-9 code(s) and any additional free text entered during CPOE defines the CHI provided to the radiologist. Our picture archiving and communication system (PACS) (iSite 3.6 [Philips] with Poeisis worklist wrapper [MedQuist]) displays the CHI above the exam timeline, where it can be consumed directly by the user. Text entered in the "Clinical question to be answered:" section of CPOE is displayed as "Reason for Study:", and text entered in the "Signs and Symptoms:" sections is displayed as "History:" A sample CHI would be displayed as "LYMPHOMAS NEC EXTRANODAL/NOS [202.80]" "Reason for Study: History of Follicular Non Hodgkin's Lymphoma, History: Evaluate for Continued CR."

Radiology reports in our practice are dictated using structured templates that contain section headers. The text appearing under the Clinical History header is regarded as CHR for the report's study. Sample CHIs and CHRs are given in Table 1. Radiology residents are instructed to consume the CHI prior to diagnostic interpretation and provide a CHR, which is primarily based on a transcription of the CHI. If any potentially confusing or uncommonly used acronyms are part

Table 1 The CHI and CHR of seven sample patients including the rates used in the evaluation

Condition	Clinical history and indication (CHI)	Rating	Clinical history section in report (CHR)	Rating
Ast	Thrombocytopenia, unspecified [287.5]/ Malignant neoplasm of other parts of brain [191.8]; Reason for Study: ^Reason: evaluate tumor progression History: L-sided weakness w h/o astrocytoma	Mention/ Specialization	58-year-old man with hypertension and thrombocytopenia. Evaluate for retroperitoneal bleed. History of astrocytoma.	Mention/ Specialization
Cir	Chronic hepatitis C without mention of hepatic coma [070.54]/HCV (hepatitis C virus) [070.70]/renal cyst [753.10]; Reason for Study:	Common comorbidity	Elevate liver lesions, masses, hcc screen	Common comorbidity
Cro	Fever, unspecified [780.60]/Jaundice, unspecified, not of newborn [782.4]; Reason for Study: ^Reason: please evaluate for biliary tree pathology History: pls evaluate for biliary tree pathology	No mention	77-year-old woman, please evaluate for biliary pathology. Fever and jaundice.	No mention
HIV	Unspecified chest pain [786.50]; Reason for Study:	No mention	History of HIV CD4 count 400 with fever to 103. Back pain and chest pain. Status post sternotomy for retrosternal thyroid in 5/09 now with persistent chest pain and dysphagia. Thoracolumbar back pain rule out abscess as cause for back and chest pain.	Mention/ Specialization
HNC (tongue cancer)	Malignant neoplasm of head, face, and neck [195.0]/chemotherapy follow-up examination [V67.2]; Reason for Study:	Generalization	66-year-old man with history of tongue cancer status post chemotherapy and radiation. Check response.	Mention/ Specialization
nHL	Lymphoma [202.80]; Reason for Study: ^Reason: lymphoma, followup not on treatment History: none	Generalization	47-year-old woman with history of lymphoma	Generalization
SLE	Migraines [346.90]; Reason for Study: ^Reason: cva History: ha, hx of cva	No mention	57-year-old woman presents with headache with history of CVA	No mention

of the CHI, residents are encouraged to spell out the full name of the condition in the CHR. Any additional pertinent information discovered from other sources such as prior radiology reports, EMR data, or discussion with the clinician or patient, is to be included in the CHR. Since most cases are dictated by a resident, this structured reporting style influenced most of the CHRs considered in our evaluation.

The patients were selected using a semantically indexed database encompassing the radiology reports written since October 9, 2007 of 17,767 representative patients. The report corpus contained 38,876 reports and was de-identified by means of a home-grown engine driven by an extensive collection of regular expressions geared toward the institute's idiosyncratic reporting style. Dates appearing in the reports or in the database as metadata were shifted using a patient-specific randomly generated offset between 1 and 364 in such a manner that the time interval between any two timestamps is preserved after offsetting. A log was kept by which de-identified patient MRNs and study dates could be mapped onto the original values on a double password-protected and encrypted system that was only accessible by the study supervisor.

The de-identified reports were segmented in sections and sentences using a home-grown engine. Section headers were normalized with respect to a list of known headers, such as "Clinical History" and "Conclusions." All segmented sentences were individually parsed by MetaMap [9] for Systemized Nomenclature of Medicine—Clinical Terms (SNOMED CT) concepts. Using its innate negation detection module (NegEx [8]), MetaMap checks if a concept appears negated or not. The de-identified reports and derived sentence and concept tables were stored in a MySQL (version 5.2.38, Oracle) database.

The concept table of the database was queried for concepts whose textual representations contain strings denoting one of the seven conditions. For instance, for condition SLE, we used "lupus" and "sle" as denoting strings. The resulting set of SNOMED CT concepts is given in Online supplement A.

For each condition, the database was queried for all reports whose clinical history and/or conclusion section contained a sentence from which one the condition's SNOMED CT concept(s) was extracted. For each report on the resulting list (maximally one per patient), we manually validated that the condition or a history thereof was confirmed. If the condition was raised as a differential diagnosis, the study was discarded. Then, using the mapping file kept by the study supervisor, each report was manually checked in our PACS to confirm that the study met the inclusion criteria (i.e., ≥11 months, same modality and body part). If so, the initial report's study was considered the baseline study and the study retrieved in PACS was considered the follow-up study. The reports on the query result list and follow-up study candidates in PACS were

handled in chronological order. Only one pair of baseline and follow-up studies was selected per patient.

For each of the 140 follow-up studies, three additional properties were marked to refine our analysis.

1. Oncological: Does the study belong to any of the oncological conditions Ast, HNC, or nHL?
2. Stat: Is the study marked as stat in PACS?
3. ICD-9 only: Does the study's CHI only contain automatically generated ICD-9 information?

Rating Scheme

As explained above, for each follow-up study, the CHI and CHR were obtained, resulting in 280 (2×7×20) information items. Two radiology residents independently marked how accurately the condition was described in each of the 280 information items, using the following four-item rating scheme:

- Mention/specialization: The condition or a specialization thereof was mentioned.
- Generalization: A generalization of the condition was mentioned.
- Common comorbidity: A common comorbidity of the condition was mentioned.
- No mention: None of the above.

In the case of condition HNC, it was checked if the specific conditions were accurately described (i.e., laryngeal, oral, and tongue cancer) instead of the container term head and neck cancer. Examples of the rates are given in Table 1.

Conflicting rates were reconciled by the two raters in a face-to-face session. The resulting rating, i.e., the rates on which the raters agreed plus the consensus rates for the conflicting ones, was considered as ground truth.

Evaluation and Metrics

To assess if the rating scheme is well defined and the ground truth creation process is reproducible, we use Cohen's κ to quantify the inter-rater agreement between the two independently obtained ratings of the 280 items. To this end, we compile the two sets of ratings in a 4×4 contingency matrix in which the rows correspond to the four rating items for the first rater and the columns correspond to the four rating items for the second rater. In this manner, the cell in row Mention/Specialization and column Generalization would contain the number of information items that were marked Mention/Specialization by the first rater and Generalization by the second.

To assess correlation between the CHI and the CHR, we compile another 4×4 contingency matrix in which the rows correspond to the four rating items of the CHI in the ground truth and the columns correspond to the four rating items of the CHR in the ground truth, see Table 2. In this manner, the cell in row Generalization and column Mention/Specialization would contain the number of studies in which the condition was described in a generalized manner in the CHI (i.e., rated Generalization) and described accurately (i.e., rated Mention/Specialization) in the CHR.

To assess if the rates of the CHI and CHR correlate, we subject the contingency matrix to a χ^2 analysis, regarding $P<0.05$ as significant.

Table 3 presents the descriptions and formulas of the metrics we compute from the contingency matrix in Table 2. Metrics 2 to 9 assess the rates assigned to the studies' CHI and CHR in isolation. Metric 10 represents the portion of studies in which the pertinent condition was not mentioned in both the CHI and the CHR. Metrics 12 and 13 quantify two senses in which a condition's description is more accurate in the CHR than in the CHI. Metric 11 describes the improvement rate from CHI to CHR.

We determine if the three categories (oncological, stat, ICD-9 only) are statistically correlated with the metrics by means of the χ^2 test. We regard $P<0.05$ as significant. In total, we have 36 statistical tests: 3 conditions×12 metrics. To compensate for multiple testing, we use the Bonferonni-adjusted threshold and regard $P<0.0014$ (0.05/36) as strongly significant.

Hypothesis Testing

We compute the 95 % confidence interval (two-sided binomial test) for the percentage of CHIs rated Mention/Specialization based on the observed percentage. We accept our first hypothesis that the CHI is a reliable source for obtaining potentially diagnosis-altering chronic conditions, if the higher bound of the confidence interval is at least 90 %.

We compute the χ^2 statistic for the contingency matrix between the CHI and CHR rates. We accept our second hypothesis that the two rates are significantly correlated, if $P<0.05$.

Results

Inter-Rater Agreement

Comparing the annotations of the two raters, 4.6 % (13/280) of items were in disagreement. On all 280 items, Cohen's κ was 0.931 (95 % confidence interval (CI) 0.893–0.969). On the 140 CHI items alone, Cohen's κ was 0.947 (95 % CI

Table 2 The variables in the contingency matrix of CHI and CHR that are used to define the evaluation metrics

		Clinical history section in report (CHR)				
		Mention/Specialization	Generalization	Common comorbidity	No mention	Sum
Clinical history and indication (CHI)	Mention/Specialization	$N_{1,1}$	$N_{1,2}$	$N_{1,3}$	$N_{1,4}$	$N_{1,*}$
	Generalization	$N_{2,1}$	$N_{2,2}$	$N_{2,3}$	$N_{2,4}$	$N_{2,*}$
	Common comorbidity	$N_{3,1}$	$N_{3,2}$	$N_{3,3}$	$N_{3,4}$	$N_{3,*}$
	No mention	$N_{4,1}$	$N_{4,2}$	$N_{4,3}$	$N_{4,4}$	$N_{4,*}$
	SUM	$N_{*,1}$	$N_{*,2}$	$N_{*,3}$	$N_{*,4}$	$N_{*,*}$

0.900–0.993); on the 140 CHR items alone, it was 0.908 (95 % CI 0.841–0.974). Of the 13 conflicts, 9 were due to obvious oversights and were readily reconciled. For instance, the first rater had erroneously annotated the following CHR as No mention with regard to HIV: "35-year-old male with sickle cell disease and HIV, here with pain crisis […]." The remaining four items were reconciled after very short discussion, taking no more than 1 min per item.

Description of Ground Truth Data

The 140 studies were spread over 30 protocol types. The astrocytoma studies were spread over only three exam types (18 MRI-brain, 1 CT-head, and 1 MRI-cervical spine), having the least variety of exam types relative to the other six conditions. On the other hand, SLE (12 procedures) and HIV (10 procedures) had the least concentrated exam type distribution. Online supplement B details the distribution of the exam types over all conditions. The distribution of the conditions over the various categories of interest is shown in Table 4.

Review of Metrics on All Cases

The contingency matrix of conditions ratings is given in Table 5. Overall, 40.7 % (57/140) of CHIs mentioned the pertinent condition, which corresponds to a 95 % confidence interval of 32.5–49.3 %. Eighty-seven of 140 (62.1 %) CHRs mentioned the pertinent condition.

A condition's rate in the CHI is significantly correlated with its rate in the CHR (χ^2 test, $P<0.00001$). If the condition was mentioned in the CHI, it was mentioned in the CHR in 89.5 % (51/57) of cases.

Table 3 Definition and formula of the evaluation metrics

Name	Description and formula	
1. N	Number of cases	$N_{*,*}$
2. CI—Mention/ Specialization	Percentage of cases with CHI rated Mention/Specialization	$N_{1,*}/N_{*,*}$
3. CI—Generalization	Percentage of cases with CHI rated Generalization	$N_{2,*}/N_{*,*}$
4. CI—Common comorbidity	Percentage of cases with CHI rated Common comorbidity	$N_{3,*}/N_{*,*}$
5. CI—No mention	Percentage of cases with CHI rated No mention	$N_{4,*}/N_{*,*}$
6. Hx—Mention/ Specialization	Percentage of cases with CHR rated Mention/Specialization	$N_{*,1}/N_{*,*}$
7. Hx—Generalization	Percentage of cases with CHR rated Generalization	$N_{*,2}/N_{*,*}$
8. Hx—Common comorbidity	Percentage of cases with CHR rated Common comorbidity	$N_{*,3}/N_{*,*}$
9. Hx—No mention	Percentage of cases with CHR rated No mention	$N_{*,4}/N_{*,*}$
10. No documented evidence	Percentage of cases lacking documented evidence that condition was known, that is, percentage of cases in which condition is neither mentioned in the CHI nor in the CHR (i.e., both No mention)	$N_{4,4}/N_{*,*}$
11. Improvement rate	Percentage of cases in which CHR is more informative than CHI, given that the condition was not mentioned in the CHI (i.e., rated Mention/Specialization)	$(N_{2,1}+N_{3,1}+N_{4,1}+N_{4,2}+N_{4,3})/(N_{2,*}+N_{3,*}+N_{4,*})$
12. Refinement rate	Percentage of cases in which condition is mentioned in CHR, given that it was indirectly mentioned in the CHI	$(N_{2,1}+N_{3,1})/(N_{2,*}+N_{3,*})$
13. Treasure hunt rate	Percentage of cases in which condition is (indirectly) mentioned in CHR, given that it was not mentioned in the CHI (i.e., rated No mention)	$(N_{4,1}+N_{4,2}+N_{4,3})/N_{4,*}$

Table 4 Distribution of the conditions' studies over the three categories of interest and their complement

	Oncological	Non-oncological	Stat	Regular	ICD-9 only	ICD-9 and manual
Ast	20	–	1	19	8	12
Cir	–	20	1	19	13	7
Cro	–	20	5	15	7	13
HIV	–	20	10	10	6	14
HNC	20	–	1	19	9	11
nHL	20	–	–	20	5	15
SLE	–	20	7	13	7	13
SUM	60	80	25	115	55	85

Fifty-one of 140 (36.4 %) CHIs had no mention of the pertinent condition versus 34/140 (24.3 %) CHRs. If the condition was not mentioned in the CHI, it was not mentioned in CHR in 64.7 % (33/51) of cases. In the 33 cases in which the condition was neither mentioned in the CHI nor in the CHR, which corresponds to 23.6 % (33/140) of all cases, we considered that no documented evidence was present to suggest that the radiologist knew the condition at the point of interpretation. This corresponds to the "no documented evidence" metric in Table 6 (under column "All").

Of the 51 cases in which the condition was not mentioned in the CHI, 16 CHRs mentioned the condition, 2 contained a generalizing description thereof, and none referenced a common comorbidity. Thus, the treasure hunt rate was 35.0 % ([16+2+0]/51), see Table 6.

If the CHI described the condition's general manner (Generalization), which happened in 23 cases, the CHR mentioned it in 73.9 % (17/23). If the CHI mentioned a common comorbidity of the condition at hand, the CHR mentioned it in 33.3 % (3/9). If the condition was rated Generalization or Common comorbidity in the CHI and Mention/Specialization in the CHR, we considered this a refinement. The overall refinement rate was 63.0 % ([17+3]/[23+9]), see Table 6.

The treasure hunt rate and the refinement rate define two senses in which a condition's description can be improved upon when the radiologist synthesizes a patient's medical history when consuming the CHI and writing the CHR. The improvement rate aggregates the two rates and is 46.0 % (38/83), see Table 6.

Review of Metrics on Oncological Versus Non-Oncological Cases

The non-oncological conditions SLE, HIV, and Crohn's most frequently lacked mention in the CHI: 90.0 % (18/20), 75.0 % (15/20), and 65.0 % (13/20), respectively (Table 4). These conditions were also the top-three conditions not mentioned in the CHR, albeit in a different order.

The oncological conditions (Ast, HNC, nHL) were mentioned more frequently in the CHI than the non-oncological conditions (Cir, Cro, SLE, HIV): 58.3 % (35/60) versus 27.5 % (22/80), respectively (Table 6). This difference is significant at $P=0.0005$. Conversely, the oncological conditions were mentioned less frequently than their complementary conditions (8.3 % (5/60) versus 57.5 % (46/80), respectively; $P<0.0001$).

In only one instance was the oncological condition (HNC) not mentioned in the CHR. In this particular case, the condition was actually mentioned in the CHI. Accordingly, in all oncological cases, documented evidence was present that may have implied the radiologist knew the condition at hand, that is, the no documented evidence rate for oncological cases was 0 % (0/60). This rate was 41.3 % (33/80) for the non-oncological cases ($P<0.0001$).

The improvement rate of oncological conditions was 80.0 % (20/25) and higher ($P=0.0001$) than for non-oncological conditions 31.0 % (18/58).

Review of Metrics on Stat Versus Non-Stat Cases

Stat cases accounted for 17.9 % (25/140) of all studies performed. Among stat cases, HIV (10 studies) and SLE (7 studies) were most prevalent (Table 4).

In all but one of the stat cases, the CHI contained an explicit mention of the condition (Mention/Specialization) or no mention at all (No mention). This skewed distribution was mirrored by the conditions' description in the CHR. In all 25 stat cases, the condition was rated Mention/Specialization or No mention.

In 52.0 % (13/25) of stat cases, the CHR mentioned the pertinent condition. This is lower, but not significant ($P=0.0045$), than the 64.3 % of CHIs that mention the condition. CHIs for stat cases more frequently had no mention of the condition than the non-stat cases (60.0 % (15/25) versus 31.3 % (36/115); $P=0.0134$). Further, the no documented evidence rate of stat cases was higher than that of non-stat cases, but this difference was not significant (48.0 % (12/25) versus 18.3 % (21/115); $P=0.0035$).

Review of Metrics on ICD-9 Only Versus Other Cases

We found that 39.3 % (55/140) of CHIs only contained automatically entered ICD-9 codes and did not have any additional manually entered history. Among the studies whose CHI had only ICD-9 information, Cir (13 studies) and HNC (9 studies) were most prevalent (Table 4).

The CHIs only generated through EMR ICD-9 codes were significantly more likely to provide a generalization of the

Table 5 Contingency table separating the rates of the CHI (row) and CHR (column) in the ground truth subdivided by condition

		Clinical history section in report (CHR)				
		Mention/ Specialization	Generalization	Common comorbidity	No mention	SUM
Clinical history and indication (CHI)	Mention/Specialization	51	3	2	1	57
	Ast	11	2	–	–	13
	Cir	8	–	2	–	10
	Cro	5	–	–	–	5
	HNC	9	–	–	1	10
	HIV	7	–	–	–	7
	nHL	11	1	–	–	12
	SLE	–	–	–	–	–
	Generalization	17	5	1	–	23
	Ast	3	–	–	–	3
	Cir	–	–	–	–	–
	Cro	1	1	1	–	3
	HNC	6	2	–	–	8
	HIV	–	–	–	–	–
	nHL	5	2	–	–	7
	SLE	2	–	–	–	2
	Common comorbidity	3	–	6	0	9
	Ast	1	–	1	–	2
	Cir	2	–	4	0	6
	Cro	–	–	–	–	–
	HNC	–	–	–	–	–
	HIV	–	–	–	–	–
	nHL	–	–	–	–	–
	SLE	–	–	1	–	1
	No mention	16	2	–	33	51
	Ast	2	–	–	–	2
	Cir	–	–	–	4	4
	Cro	3	–	–	9	12
	HNC	1	1	–	–	2
	HIV	4	–	–	9	13
	nHL	1	–	–	–	1
	SLE	5	1	–	11	17
	SUM	87	10	9	34	140
	Ast	17	2	1	–	20
	Cir	10	–	6	4	20
	Cro	9	1	1	9	20
	HNC	16	3	–	1	20
	HIV	11	–	–	9	20
	nHL	17	3	–	–	20
	SLE	7	1	1	11	20

condition compared to when a manual diagnosis was entered in conjunction with the ICD-9 code (30.9 % (17/55) versus 7.1 % (6/85), respectively; P=0.0005). Concordantly, CHIs generated through ICD-9 alone were less likely to specifically mention the condition compared to when there was an additional manually entered diagnosis (20.0 % (11/55) versus 54.1 % (46/85), respectively; P=0.0001).

The improvement rate for cases for which CHI consists solely of automatically generated ICD-9 codes is 65.9 %. This is significantly higher than cases with CHIs that also contain

Table 6 Outcomes of the metrics with *P* values of statistical tests

	All	Oncological	Non-oncological	Significant difference?	*P* value
N	100.0 (140/140)	42.9 (60/140)	57.1 (80/140)		
CI—Mention/Specialization	40.7 (57/140)	58.3 (35/60)	27.5 (22/80)	Yes[b]	0.0005[a]
CI—Generalization	16.4 (23/140)	30.0 (18/60)	6.3 (5/80)	Yes[b]	0.0004[a]
CI—Common comorbidity	6.4 (9/140)	3.3 (2/60)	8.8 (7/80)	No[b]	0.3446[a]
CI—No mention	36.4 (51/140)	8.3 (5/60)	57.5 (46/80)	Yes[b]	<0.0001[a]
Hx—Mention/Specialization	62.1 (87/140)	83.3 (50/60)	46.3 (37/80)	Yes[b]	<0.0001[a]
Hx—Generalization	7.1 (10/140)	13.3 (8/60)	2.5 (2/80)	Yes[b]	0.0330[a]
Hx—Common comorbidity	6.4 (9/140)	1.7 (1/60)	10.0 (8/80)	No[b]	0.1007[a]
Hx—No mention	24.3 (34/140)	1.7 (1/60)	41.3 (33/80)	Yes[b]	<0.0001[a]
No documented evidence	24.0 (33/140)	0.0 (0/60)	41.3 (33/80)	Yes[b]	<0.0001[a]
Improvement rate	46.0 (38/83)	80.0 (20/25)	31.0 (18/58)	Yes[b]	0.0001[a]
Refinement rate	63.0 (20/32)	75.0 (15/20)	41.7 (5/12)	No[b]	0.1314[a]
Treasure hunt rate	35.0 (18/51)	100.0 (5/5)	28.3 (13/46)	Yes[b]	0.0070[a]

	Stat	Regular	Significant difference?	P value	ICD-9 only		ICD-9 and manual		Significant difference?	P value
N	17.9 (25/140)	82.1 (115/140)			39.3	(55/140)	60.7	(85/140)		
CI—Mention/Specialization	36.0 (9/25)	41.7 (48/115)	No[b]	0.7605[a]	20.0	(11/55)	54.1	(46/85)	Yes[b]	0.0001[a]
CI—Generalization	0.0 (0/25)	20.0 (23/115)	Yes[b]	0.0317[a]	30.9	(17/55)	7.1	(6/85)	Yes[b]	0.0005[a]
CI—Common comorbidity	4.0 (1/25)	7.0 (8/115)	No[b]	0.9232[a]	9.1	(5/55)	4.7	(4/85)	No[b]	0.4963[a]
CI—No mention	60.0 (15/25)	31.3 (36/115)	Yes[b]	0.0134[a]	40.0	(22/55)	34.1	(29/85)	No[b]	0.5985[a]
Hx—Mention/Specialization	52.0 (13/25)	64.3 (74/115)	Yes[b]	0.0045[a]	69.1	(38/55)	57.6	(49/85)	No[b]	0.2360[a]
Hx—Generalization	0.0 (0/25)	8.7 (10/115)	No[b]	0.2706[a]	5.5	(3/55)	8.2	(7/85)	No[b]	0.7734[a]
Hx—Common comorbidity	0.0 (0/25)	7.8 (9/115)	No[b]	0.3192[a]	7.3	(4/55)	5.9	(5/85)	No[b]	0.9799[a]
Hx—No mention	48.0 (12/25)	19.1 (22/115)	Yes[b]	0.0052[a]	18.2	(10/55)	28.2	(24/85)	No[b]	0.2489[a]
No documented evidence	48.0 (12/25)	18.3 (21/115)	Yes[b]	0.0035[a]	18.2	(10/55)	27.1	(23/85)	No[b]	0.3150[a]
Improvement rate	25.0 (4/16)	50.7 (34/67)	No[b]	0.1146[a]	65.9	(29/44)	23.1	(9/39)	Yes[b]	0.0002[a]
Refinement rate	100.0 (1/1)	61.3 (19/31)	No[b]	1.0000[a]	77.3	(17/22)	30.0	(3/10)	Yes[b]	0.0303[a]
Treasure hunt rate	20.0 (3/15)	41.7 (15/36)	No[b]	0.2486[a]	54.5	(12/22)	20.7	(6/29)	Yes[b]	0.0271[a]

[a] Chi-squared test result not reliable

[b] Significant with respect to Bonferonni-adjusted α=0.001388889 (0.05/[3×12])

manual input (23.1 % (9/39); $P=0.0002$). Note that both the refinement rate and the treasure hunt rate for ICD-9 cases is higher than for their complementary cases, but that neither difference is significant ($P=0.0303$ and $P=0.3177$, respectively).

Discussion

Lack of appropriate clinical history is a known detriment to radiologic exam interpretation [6, 7]. For example, ground-glass opacities on a chest CT are generally not a specific finding; however, in immunecompromised patients they are suspicious of an atypical infection. Similarly, enlarged axillary lymph nodes on a mammogram could be an ominous finding unless there is history of an inflammatory condition such as lupus. Given the established importance of past medical history in radiological exam interpretation, we evaluated the rates at which chronic conditions are mentioned in the CHI and in the CHR.

We evaluated the trustworthiness of the CHI as a means to communicate pertinent chronic conditions to the radiologist. By pre-selecting patients whom were known to suffer from seven chronic conditions, we could track the rate by which these conditions were successfully communicated. This unique study design sets it apart from earlier studies [6, 10] in which the CHI proper was evaluated without structural reference to the patient's complete clinical history.

Quality of CHI

We set out to evaluate the hypothesis that the CHI is a reliable source for obtaining a potentially diagnosis-altering chronic condition, which had been confirmed in the patients' prior radiology reports. We found that 40.7 % (57/140) of CHIs were rated Mention/Specialization, with 95 % CI 32.5–49.3 %. Based on these results, since the interval's upper bound does not exceed 90 %, we reject our first hypothesis. If we consider CHIs in which the condition is mentioned indirectly (Generalization) or implicitly (Common comorbidity) as reliable sources, in addition to the CHIs rated Mention/Specialization, the success rate is 63.6 % (89/140), with 95 % CI 55.0–71.5 %. Even under this more liberal definition, we reject our first hypothesis.

We found that non-oncologic conditions were much more likely not to be mentioned compared to oncologic conditions; this may be because the oncologic conditions in question are often the patient's most significant clinical problem and many imaging exams are ordered specifically for the purpose of cancer follow-up. On the other hand, chronic non-oncological conditions such as HIV or SLE may not appear to be directly implicated in a patient's acute presentation and, for this reason, may be omitted from the indication for imaging exam.

One cause for failing to mention chronic conditions may be that the ordering physician is not aware of a patient's chronic condition, especially for patients who see multiple different subspecialists, visit various hospital networks, or present to the ER with symptoms seemingly unrelated to their chronic condition. Additionally, clinical personnel other than the referring physician, such as nurses, often place orders for imaging studies and may have insufficient understanding as to what information is relevant to the interpreting radiologist and/or have had minimal exposure to the radiological workflow. This root cause was confirmed by an exam-ordering oncologist at our institution.

Time pressure may constitute another root cause for inferior CHI, which is supported by our finding that stat cases were more likely to have no mention of a chronic condition than routine cases. In an acute setting such as the emergency room, there may be insufficient time to gather a thorough history or perhaps a known chronic condition may be seen as irrelevant to the interpretation of a stat exam.

As discussed above, the CHI provided to the radiologist is composed of one or more ICD-9 diagnosis codes associated with the order and additional free text history typed by the ordering clinician. While the ICD-9 code may be sufficient for the exam to be ordered, we have shown that alone, it is more likely to generalize a condition compared to a CHI which also includes additional, manually entered history. Conversely, we found that CHIs including manually entered histories were significantly more likely to make specific mention of the condition. The lack of time in a clinician's busy practice as well as delegation of order placement to support staff such as nurses may again be potential sources for these inferior CHIs.

Nationwide, there is an increasing focus on ensuring the appropriateness of imaging examinations and reducing unnecessary studies. Enhancing imaging appropriateness has been a driver for CPOE systems and not surprisingly, the order-entry workflow orchestrated by CPOE systems aims to create an appropriateness-centric experience. As an unanticipated consequence, as ordering physicians increasingly focus on exam appropriateness, they may be inclined to submit the minimal set of information required for approval of the exam, feeling that any additional information is either unnecessary or perhaps may flag the order as inappropriate. The sparseness of patient information may, therefore, be an unintended negative consequence of CPOE systems.

Correlation Between CHI and CHR

We observed that if and how a chronic condition is mentioned in the CHI is significantly correlated with if and how it is mentioned in the CHR ($P<0.0001$), which confirms our second hypothesis.

We found that the majority of CHIs (59.3 % (83/140)) contain generalized, implicit (Common comorbidity), or no mention of the targeted condition. These mentions can be improved upon by the radiologist who needs to consult other sources such as a prior radiology report, the EHR, or through interaction with the clinician. These steps are presumably time consuming and certainly disruptive to workflow. The overall improvement rate in our study was 46.0 % (38/83). The improvement rate was even higher for ICD-9 code only CHIs (65.9 % [29/44]), which tended to have a generalization of the history that may not have been adequate for interpretation.

If the CHI contains no mention of the target condition, for it to be known at the time of interpretation by a radiologist, it would need to have been retrieved by "accident" or by a radiologist who conscientiously aggregates patient information from other sources, a so called "treasure hunt." The overall treasure hunt rate was 35.0 % (18/51). The treasure hunt rate of CHIs containing manually entered information (20.7 % (6/29)) was lower than the overall treasure hunt rate, which may reflect a false sense of being sufficiently informed by information entered manually by a clinician. By contrast, the treasure hunt rate of ICD-9 only CHIs was significantly higher at 54.5 % (12/22). This may be due a hypothesized effect that ICD-9 code only CHIs flag the radiologist that no human provided background information. The treasure hunt rate of stat cases (20.0 % (3/15)) was lower yet, potentially due to the fact that such cases are read under increased time pressure.

The no documented evidence rate gives the portion of cases in which the target condition is not mentioned (No mention) in both the CHI and the CHR. Hypothetically, it is possible that the radiologist learned about the condition in a treasure hunt and did not report it in the CHR. Therefore, the reported no documented evidence rates should be interpreted as upper bounds on the real portion of exams that were interpreted without the radiologist's understanding of the target condition.

We found that no documented evidence was present for about a quarter of cases (24.3 % (34/140)). The rate was higher for non-oncological conditions such as SLE, HIV, and Crohn's disease (41.3 % (33/80)). Given the pertinence of these conditions, as was laid out above, it is conceivable that these high rates contain individual cases whose finalized radiology read is inaccurate in the light of the chronic condition.

The quality of the clinical indications has been analyzed in several studies, some with alarming outcomes. One study found that the current diagnosis was missing in more than 28 % of the cases and that less than 31 % of the cases had an appropriate indication [11]. Another study showed that the clinical indication was inadequate or incomplete in 24 % of the cases [12]. In yet another study, less than 7 % of the clinical order indications were marked as reasonable or excellent [13]. A survey of radiologists showed that the majority (72 %) felt in need of more clinical information about their patients, and 87 % stated that additional information would have significant impact on interpretation [14].

As a means to improve quality, prior studies have suggested educating clinicians as to what constitutes an adequate clinical indication and to engage technologists in the completion of clinical background [11]. The referenced study reports positive short-term effects on the information content of the CHI, but also observed that in a 3-month window, the "rates decayed back toward baseline." Several studies also showed positive impact of CPOE on the quality of clinical indications as compared to paper-based orders [12–15]. Technological innovation that automatically synthesizes patient background information and presents it to the radiologist in an intuitive manner is yet another route for resolving the clinical information communication gap between referring clinicians and radiologists laid bare by our study [16].

Our study has a number of limitations. First, we established that there was a correlation between the condition ratings in the CHI and the CHR; however, this correlation does not establish causation. Second, the conditions were selected based on radiologists' subjective opinion as to what constitutes a pertinent condition during exam interpretation. We did not evaluate if a radiologist's awareness of the condition actually affected the interpretation or, perhaps more importantly, patient care. Third, we did not record which clinicians actually placed exam orders in our study. Therefore, we were not able to analyze whether there was any difference in quality between CHIs completed by, for example, a nurse, resident, or attending physician. It is possible that variations in CPOE profiles among these users may have impacted workflow and resulted in CHI quality differences. Finally, our study was conducted at one institution and it is hard to predict how our results generalize to other institutions. For instance, the selected conditions may be less prevalent elsewhere.

In spite of the aforementioned limitations, our study demonstrates an alarming lack of communication of pertinent medical information to the radiologist at the time of exam interpretation. Given the recent changes in health care regulations, including the formation of ACOs, the yield and appropriateness of radiologic studies will likely now be more heavily scrutinized than even before. Our study suggests that one possible avenue to improve the value of radiologic exams would be to improve the communication of vital past medical history to radiologists at the time of exam interpretation.

References

1. Bidgood WD and Horii SC: "Introduction to the ACR-NEMA DICOM Standard," *Radiographics*, vol. 12, no. May, pp. 345–355, 1992
2. Borgstede JP: "Radiology: commodity or specialty," *Radiology* vol. 247, no. 3, pp. 613–616, 2008

3. Boland GW: "Teleradiology for auction: the radiologist commoditized and how to prevent it," *J Am Coll Radiol*, vol. 6, no. 3, pp. 137–138, 2009
4. Abramson RG, Berger PE and Brant-Zawadzki MN: "Accountable care organizations and radiology: threat or opportunity?," *J Am Coll Radiol* vol. 9, no. 12, pp. 900–6, Dec. 2012
5. Lexa FJ and Berlin JW: "Accountable care organizations for neuroradiologists: threats and opportunities," *Neuroimaging Clin N Am.*, vol. 22, no. 3, pp. 437–41, Aug. 2012
6. Leslie A, Jones AJ and Goddard PR: "The influence of clinical information on the reporting of CT by radiologists," *Br J Radiol* vol. 73, no. 874, pp. 1052–5, Oct. 2000
7. Cohen MD: "Accuracy of information on imaging requisitions: does it matter?," *J Am Coll Radiol* vol. 4, no. 9, pp. 617–21, Sep. 2007
8. 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* vol. 34, no. 5, pp. 301–10, 2001
9. Aronson AR and Lang FM: "An overview of MetaMap: historical perspective and recent advances," *J Am Med Inform Assoc* vol. 17, no. 3, pp. 229–236, 2010
10. Loy CT and Irwig L: "Accuracy of Diagnostic Tests Read With and Without Clinical Information-A Systematic Review," *JAMA* vol. 292, no. 13, pp. 1602–1609, 2004
11. Gunderman RB, Phillips MD and Cohen MD: "Improving clinical histories on radiology requisitions.," *Acad Radiol*, vol. 8, no. 4, pp. 299–303, Apr. 2001
12. Cohen MD, Curtin S and Lee R: "Evaluation of the quality of radiology requisitions for intensive care unit patients," *Acad Radiol*, vol. 13, no. 2, pp. 236–40, Feb. 2006
13. Schneider E, Franz W, Spitznagel R, Bascom DA and Obuchowski NA: "Effect of Computerized Physician Order Entry on Radiologic Examination Order Indication Quality," *Arch Intern Med* vol. 171, no. 11, pp. 1036–1037, 2011
14. Boonn WW and Langlotz CP: "Radiologist use of and perceived need for patient data access.," *J Digit Imaging*, vol. 22, no. 4, pp. 357–62, Aug. 2009
15. Alkasab TK, Alkasab JR, Abujudeh HH: "Effects of a computerized provider order entry system on clinical histories provided in emergency department radiology requisitions.," J Am Coll Radiol vol. 6, no. 3, pp. 194–200, Mar. 2009
16. Travis AR, Sevenster M, Qian Y, Mankovich G, Buurman J, Chang PJ: "Enhancing provided patient clinical information by automated review of prior radiology reports using the Clinical Context Indicator (CCI): a NLP based data extraction and presentation PACS-integrated tool," in RSNA Annu Symp Proc., Dec. 2013