

Research Paper ■

Exploring the Degree of Concordance of Coded and Textual Data in Answering Clinical Queries from a Clinical Data Repository

H. DAVID STEIN, MD, PRAKASH NADKARNI, MD, JOSEPH ERDOS, MD, PHD, PERRY L. MILLER, MD, PHD

Abstract **Objective:** To query a clinical data repository (CDR) for answers to clinical questions to determine whether different types of fields (coded and free text) would yield confirmatory, complementary, or conflicting information and to discuss the issues involved in producing the discrepancies between the fields.

Methods: The appropriate data fields in a subset of a CDR (5,135 patient records) were searched for the answers to three questions related to surgical procedures. Each search included at least one coded data field and at least one free-text field. The identified free-text records were then searched manually to ensure correct interpretation. The fields were then compared to determine whether they agreed with each other, were supportive of each other, contained no entry (absence of data), or were contradictory.

Results: The degree of concordance varied greatly according to the field and the question asked. Some fields were not granular enough to answer the question. The free-text fields often gave an answer that was not definitive. Absence of data was most logically interpreted in some cases as lack of completion of data and in others as a negative answer. Even with a question as specific as which side a hernia was on, contradictory data were found in 5 to 8 percent of the records.

Conclusions: Using the data in the CDR to answer clinical questions can yield significantly disparate results depending on the question and which data fields are searched. A database cannot just be queried in automated fashion and the results reported. Both coded and textual fields must be searched to obtain the fullest assessment. This can be expected to result in information that may be confirmatory, complementary, or conflicting. To yield the most accurate information possible, final answers to questions require human judgment and may require the gathering of additional information.

■ JAMIA. 2000;7:42-54.

Affiliation of the authors: Yale University School of Medicine, New Haven, Connecticut.

This work was supported in part by NIH grants T15-LM07056 and G08-LM05583 from the National Library of Medicine.

Correspondence and reprints: H. David Stein, MD, Yale University School of Medicine, Center for Medical Informatics, 333 Decar Street, P.O. Box 208009, New Haven, CT 06520-8009; e-mail: {hdavid.stein@yale.edu}.

Received for publication: 3/31/99; accepted for publication: 8/16/99.

Because of the increasing reliance on medical data to influence both administrative and clinical policies, the data stored in a clinical data repository (CDR) must be available to answer pertinent questions. Such data may be stored in either coded or free-text fields. Coded fields are more efficient for storage and retrieval of data, but clinicians frequently find that the choices available in coded data are too limiting. Moreover, even a vocabulary as large as that of the Unified Medical Language System Metathesaurus was found

wanting in terms when, for example, a problem list was developed, and terms had to be added to satisfy clinicians' needs.¹

Free text is the preferred means of entry for many types of data, such as operative notes and discharge summaries. It allows the clinician to express nuances of information that may be unavailable in coded entries. As a result of the different perceived advantages, most CDRs allow data to be entered in both coded and free-text fields. Since both types of fields are often available, we sought to determine to what extent it is important to look at both types to obtain answers to clinical queries.

We hypothesized that it was important to search both coded data and free text in the CDR. In performing such searches, various issues must be considered. This paper explores these issues in detail in the context of one CDR, looking at several questions relevant to surgical patient management. Data accuracy obviously affects concordance among data in different fields. When data do not agree, it may be difficult to determine which field is correct without questioning the patient or establishing arbitrary standards. Our goal was not to assess the correct value of any particular field or the correct answer to the clinical questions we posed. Rather, the goal of the paper was to determine the extent to which the data stored in different fields and in different types of fields are confirmatory, complementary, or contradictory and whether the answer to a question varies with the type of clinical question asked. We then discuss possible causes for the discrepancies.

Background

Accuracy

Several previous studies have looked at the accuracy of data stored in the patient record and have sought to find the reasons for inaccuracies. One study,² which dealt with recording medications in the patient chart, identified patient error (36.1 percent of errors) as the most common cause of inaccurate data. The most common error identified was caused by the patient neglecting to mention medications. Wilton and Pen-nisi³ looked at the accuracy of data transcribed from handwritten notes and found a 5.9 percent rate of inaccuracy due to transcription. In other studies the accuracy was reported to vary with the type of data and the person entering the data.^{4,5} Error rates vary widely and may be as high as 85 percent (because of failure to enter data) in some fields in anesthesia charts.⁴ Hobbs et al.⁵ found an error rate up to 25 percent in some fields in a study of 158 acute medical hos-

pital admissions. Completeness and accuracy of data may also depend on the circumstances of the data entry. This was evident in the study of Hohnloser et al.,⁶ who found that if clinicians were required to look up the codes for diagnoses included in a discharge summary, they tended to leave diagnoses out of the summary, making both the coded information and the free text less accurate. Hogan and Wagner⁷ reviewed 19 studies on the accuracy of data in the computerized patient record. They found a correctness rate (equivalent to the positive predictive value) of 67 to 100 percent and a completeness rate (equivalent to sensitivity) of 30.7 to 100 percent, depending on the study and the data. Iezzoni⁸ argued that administrative databases were of only limited value in answering clinical questions. He found multiple factors that negatively affected the usefulness of the data, including lack of granularity of codes, "code creep" to enhance reimbursement, limitation in the number of allowable diagnoses, variability between hospitals, and lack of specificity of timing. Green and Winfield⁹ looked at the data used to produce "report cards" on hospitals in New York where cardiac surgery was performed. He found a significant increase in risk factors reported to the database by the hospitals after it was clear that the data were being used to evaluate the quality of their surgery. He concluded that knowledge of the eventual use of the data by the person entering the data might alter its content. Young et al.¹⁰ compared the adequacy of treatment of patients with schizophrenia by reviewing patient medical records and interviewing patients. He concluded that more than half the instances of poor patient care would have been missed if the information were based only on a search of the medical record. Hobar and Leahy¹¹ compared a database of information on low-birth-weight infants with a random selection of the original charts and found an error rate that depended on the field, varying from 1.3 percent (date of birth) to 2.1 (gender) and 8.8 percent (date of discharge). Most of these errors were ascribed to transcription or interpretation.

Concordance

Since data analyzed from secondary data sources inherently limit the ability to determine inaccuracies, it would appear logical to search several data items looking for confirmation of data fields to increase accuracy. Few formal studies have explored the concordance of free text and coded data in the same database. Hogan and Wagner¹² found that if clinicians were given the opportunity to supplement the coded data for medication orders with free text, they did so 35 percent of the time. In 81 percent of these instances,

Table 1 ■

Questions Asked and Fields Searched to Determine Answers

Question*	Criteria for Inclusion in Search	Field Searched	Type of Field	Person Entering Data	Time Entered (in Relation to Event)	Search Criteria
Postoperative pulmonary embolism	Presence of dictated discharge summary	Postoperative complication	Coded	Surgical risk coordinator	When event disclosed	"Pulmonary embolism"
		ICD-9-coded discharge diagnosis	Coded	Medical records personnel	After discharge from hospital	Codes for pulmonary embolism (415.1, 415.11, 415.19)
	Operation on this or previous admission prior to embolism	Dictated discharge summary	Free text	Clinician in charge of patient at time of discharge	At or close to time of discharge	Mention of pulmonary embolism, PE, or VQ scan
Wound infection	Presence of dictated discharge summary	Postoperative infection	Coded	Surgical risk coordinator after determination by nurse epidemiologist	When event discovered	"Yes"
		Postoperative complication	Coded	Surgical risk coordinator after determination by nurse epidemiologist	When event discovered	"Wound infection"
		ICD-9-coded discharge diagnosis	Coded	Medical records personnel	After discharge from hospital	Codes for postoperative infection (998.5, 998.51, 998.59)
		Dictated discharge summary	Free text	Clinician in charge of patient at time of discharge	At or close to time of discharge	Mention of "wound infection"
Hernia	Presence of dictated operative note	Postoperative diagnosis	Coded	Surgical team	At completion of operation	"*Hernia*" or "?ih*"
		Primary procedure	Coded	Surgical team	At completion of operation	"*Hernia*" or "?ih*"
		Dictated operative note	Free text	Surgeon performing procedure	Following procedure	Mention of hernia
		ICD-9-coded discharge diagnosis	Coded	Medical records personnel	After discharge from hospital	Codes for abdominal or groin hernia (550 to 553)
		Dictated discharge summary	Free text	Clinician in charge of patient at time of discharge	At or close to time of discharge	Mention of hernia

*Questions are given in their entirety in the text.

the free text altered the meaning of the coded data. They concluded that allowing free-text entries decreased the completeness of the information available to an expert system that interprets only coded fields. Cooper et al.¹³ emphasized the importance of considering both types of data in a study. In trying to locate relevant charts using coded data, they recognized that free-text fields contain information not available in the coded data and may also have to be searched. Jollis et al.¹⁴ looked at the discordance of data used in outcomes research. They compared ICD-9 codes to data entered by a cardiology fellow during cardiac catheterization. They found agreement of 49 to 95 percent, depending on the question asked.

Methods

Data

We used the information stored in the clinical patient record system of the West Haven Veterans Administration (VA) Medical Center, which is written in MUMPS. This is both an administrative and a clinical database. To create a pilot CDR, major portions of the record were downloaded into a relational database. A total of 18,996 patient charts were downloaded; they contained discharge summaries for discharges that occurred from July 1995 to April 1998. A total of 5,135 charts dating from April 1996 to April 1998 contained dictated operative reports. Of these, 2,377 charts contained both an operative report and a discharge summary. In addition to the free-text dictated reports, other pertinent information was available in coded data fields and in two fields that required the textual entry of a simple diagnosis (postoperative diagnosis) or procedure (primary procedure). For purposes of this study, these fields were considered coded fields, since their meaning was easily extracted when they were reviewed and typographic mistakes were disregarded. We then searched the available fields for information to answer the clinical questions posed.

Questions

The questions we asked were chosen because they were clinical questions of the type that might be asked of a CDR with the expectation that relevant information would be found in both the coded data and the free-text fields, and because they were typical of questions that might be asked in patient care, quality assurance, or outcomes research. The three questions were:

- Did the patient have a postoperative pulmonary embolism?

- Did the patient have a postoperative wound infection?
- Did the patient have an operation for a hernia?

If the patient had had a hernia operation:

- What kind of hernia was it (inguinal, femoral, umbilical, or ventral)?
- If it was a groin hernia, what type (direct, indirect, or recurrent)?
- Which side was the groin hernia on?
- Was a plastic mesh graft used for the repair?

Searching Data

To answer these questions, the coded data fields were searched using relational database query tools. The searches were made broad enough to allow for spelling mistakes by adding wildcard searches and by stemming. The free-text fields were searched for key words using dtSearch, a text retrieval engine that indexes free text and can perform searches in a variety of ways, including Boolean searching and adjacent word matching. When all potentially relevant records were retrieved, each selected field was inspected to determine how the information it contained answered the question.

To answer the three questions, the database was searched for all fields that contained pertinent data. With the exception of the field for the ICD-9-coded discharge diagnosis, the chosen fields were all clinical fields that were utilized in the database to yield clinical, not administrative, information. The ICD-9-coded discharge diagnosis field was chosen, despite previous studies questioning its accuracy, because of its ubiquity and because we wanted to see how it compared with the other coded and free-text fields. Table 1 lists each question, the fields searched, the criteria used, the person who entered the data, and the time data were entered. If a dictated free-text field was required for the search and was not present in a specific chart, that chart was excluded.

Each free-text entry was read in its entirety by an experienced surgeon (H.D.S.) to determine whether the sought-for diagnosis or term was present, or could be interpreted as being present, in the field. Thus, we excluded from the focus of this study the issue of the ability of a textual search engine to find the correct terms in the free text and our ability to arrive at the proper word choice. We did not attempt to determine which data item contained the accurate information. Rather, we were interested in the data themselves and

in reporting their agreement with data in other fields, assuming the best possible retrieval circumstances. We disregarded spelling mistakes and nonstandard abbreviations, since they can be detected by optimal search algorithms.

Database Queries

Pulmonary Embolism

In looking for patients with postoperative embolism (see Table 1), we searched two coded fields (postoperative complication, a choice field completed by the surgical risk coordinator, and the ICD-9-coded discharge diagnosis) for pulmonary embolism and the dictated free-text discharge summary for variations of "pulmonary embolism," "PE," and "VQ scan." We selected those discharge summaries that indicated at least a moderate probability of embolism and those for patients who were subsequently treated for embolism. Any chart that had at least one field positive for pulmonary embolism and indicated an operation on the same visit or within 30 days (the criteria used by the surgical risk coordinator) was selected. To gain insight into the data entry process, the surgical risk coordinator was interviewed about her method of data entry.

Wound Infection

For the questions about wound infections (see Table 1), we searched three coded fields (postoperative infection, postoperative complication, and ICD-9-coded discharge diagnosis) and the free-text discharge summary. The postoperative infection field (in which "yes" was entered if infection was diagnosed) was completed by a nurse epidemiologist, who used the Center for Disease Control definition of wound infection,¹⁵ a nationally accepted standard. There is no ICD-9-discharge code specifically for "wound infection," and most patients were coded under "postoperative infection," which is more general. (This classification is only supportive of wound infection, and it includes such other postoperative infections as urinary tract infections and pneumonia.)

Abdominal Hernias

In answering questions about abdominal hernias (see Table 1), we considered five fields. Three were coded fields (postoperative diagnosis, primary procedure, and ICD-9-coded discharge diagnosis); the other two were free-text fields (the operative note and the discharge summary). Cases of nonabdominal hernias (e.g., vertebral disc herniation or hiatal hernias) were excluded. Hospital policy was that discharge summaries were dictated for inpatients only, not for am-

bulatory surgical patients. Therefore, discharge summaries were not available for the patients who had undergone hernia repair in ambulatory surgery. All hernia operations were included in the study, and the findings from charts with discharge summaries and those without were reported separately. Operating-room personnel were queried about their method of data entry during the operations.

Classification of the Data

The data contained in each field were classified into one of three categories:

- The field might *be mute on a subject* (i.e., the information was *absent*, in that it either did not mention the subject or did not have any code that referred to the subject. We included only those charts in which the pertinent free-text field was present, so that the field could be evaluated in comparison with the coded text. This was done in the knowledge that there might have been charts with coded data and no free text (discharge summary or operative note) available for review.
- The field might *make a statement that supported a fact* but was subject to interpretation. An example is the use of the term "erythema" in the free-text summary to describe the condition of the wound. This does not explicitly state that an infection was present but rather implies that there might have been one. Similarly, a code of "postoperative infection" could support either a wound infection or any other type of postoperative infection (such as pneumonia).
- The field might *contain a definitive statement* as to the presence or absence of a fact.

Comparison of the Data Fields

The data fields were then compared to determine whether they were in agreement with each other. We recognized four levels of concordance:

- *Data in agreement.* Here data in the two fields led to the same conclusion. An example is that both the ICD-9-coded field and the dictated discharge note stated that the patient had an inguinal hernia.
- *Supportive data.* One field gave a definitive answer, and the other only implied the answer, thus confirming the first. If only the second had been searched, a definitive answer would not have been obtained. One example was that the ICD-9 code for wound infection was actually postoperative infection, which might imply a wound infection but

could also include postoperative pneumonia. Another example is a discharge summary stating "erythema of the wound," which implied an infection but was not a definitive statement.

- *Absent data.* One field contained a finding and the other field did not comment on it. For example, the discharge summary listed "pulmonary embolism" as a diagnosis, but the ICD-9 field was not coded for that diagnosis.
- *Contradictory data.* The two fields disagreed. For example, the coded field for postoperative complication stated that there was a wound infection and the discharge diagnosis stated "wound was clean on discharge with no sign of infection."

Results

Postoperative Pulmonary Embolism

In looking for patients with postoperative pulmonary emboli, we searched the 2,377 charts that contained both the discharge information and the operative information. Although the free-text "operative dictation" field was not searched, the "postoperative complication" field linked to the surgical information was required. Sixteen cases that met the selection criteria were identified (Table 2). In 14 cases, the diagnosis was clear in the free-text dictated discharge summary. In two additional cases, at least one coded field was positive but the discharge summary clearly indicated that pulmonary embolism had been considered and ruled out (in one case because of pneumonia and in the other because of a terminal event that was deemed to be cardiac in origin).

In determining whether a patient had a postoperative pulmonary embolism, we found that the coded fields either indicated its presence or made no comment. There were no uncertain entries in either coded field. Thus, if we had tried to find out how many patients had postoperative pulmonary embolism by looking at a single coded field, we would have found either ten, using the "postoperative complication" field, or eight using the ICD-9-coded field, instead of the 14 identified by the free-text field. The results showed agreement of the free-text and the "postoperative complication" fields eight times and contradiction between them seven times. The free-text and the ICD-9 fields agreed six times and were contradictory ten times.

If the two coded fields were considered together (Table 3), 14 patients would have been identified, but the data were contradicted by the free-text field for two patients (shown in Table 3 under "Contradictory Data.") If all three fields had been searched, all 16

Table 2 ■

Types of Data Identified in Fields Searched to Answer the Question "Did the Patient Have a Postoperative Pulmonary Embolism?" (N = 16 Cases)

Filed (Type)	Type of Data			
	Positive	Supportive	Absent	Contradictory
Postoperative complication (coded)	10	–	6	–
ICD-9-coded discharge diagnosis (coded)	8	–	8	–
Dictated discharge summary (text)	14	–	–	2

Table 3 ■

Levels of Concordance among Three Data Fields Searched to Answer the Question "Did the Patient Have a Postoperative Pulmonary Embolism?" (N = 16 Cases)

Concordance Level	No. Cases (%)
Data in agreement: 3 fields positive	4 (25)
Absent data: 2 fields positive 1 absent	7 (44)
1 field positive 2 absent	
Contradictory data: 2 fields positive 1 contradictory	1 (6)
1 field positive 1 contradictory 1 absent	1 (6)

patients would have been identified, but the results would have required human judgment to resolve the two cases in which the positive diagnosis in the coded fields was contradicted by the free text.

Wound Infection

The 2,377 charts that included both a discharge summary and an operative report were searched to answer the question about wound infection. Of these, 315 charts had at least one field positive for wound infection or an ICD-9 field containing one of the codes

Table 4 ■

Types of Data Identified in Fields Searched to Answer the Question “Did the Patient Have a Postoperative Wound Infection?” (N = 315 Cases)

Field (Type)	Type of Data			
	Positive	Supportive	Absent	Contradictory
Postoperative infection (coded)	136	–	179	–
Postoperative complication (coded)	140	–	175	–
ICD-9-coded discharge diagnosis (coded)	–	198	117	–
Dictated discharge summary (text)	93	37	167	18

that suggested wound infection (Table 4). Of the 315 charts that met the selection criteria for wound infection, 93 to 140 identified infections, depending on the field examined. The coded ICD-9 diagnosis could be described as only supportive, since most infections were coded under “postoperative infection,” which included other sources of infection in addition to wounds. The diagnosis would have been difficult to make if the free-text discharge summary were considered alone. In 93 instances, the physician dictated a definite statement making it clear that a wound infection was present (e.g., “wound infection,” “pus was present in wound,” “positive wound culture for staphylococcus”). In 37 additional dictated summaries, the diagnosis was supported but a definitive statement was not made (e.g., “erythema of wound,” “wound opened”). Eighteen additional charts had terminology indicating the absence of wound infection (e.g., “wound well healed,” “wound clean and dry”).

In looking at the concordance of the data fields (Table 5), only 24 charts had agreement in all possible fields (positive in three and supportive in the ICD-9-coded field). In 125 charts at least two fields were positive with no contradictory fields, and in 172 charts at least one field was positive with no contradictory fields. Twelve charts had at least one positive field that was contradicted by the discharge summary. One hundred twenty-four charts had a “supportive” diagnosis in the ICD-9-coded field but no mention of wound infection in another field.

Hernia

The remaining questions concerned patients who had hernia repair. Only the 5,135 charts that contained dictated operative reports were searched. Of 209 patients

who had hernia operations, nine had either hiatus hernia repair or vertebral disc herniation and were excluded. This left 200 patients who had abdominal hernia operations (Table 6). The coded fields completed in the operating room and the dictated operative note were available for all 200, but the discharge summary and coded ICD-9 discharge diagnosis were available for only 95. In the search to determine what type of hernia the patient had (inguinal, femoral, umbilical, or ventral), almost all the fields were found to

Table 5 ■

Levels of Concordance among Four Data Fields Searched to Answer the Question “Did the Patient Have a Postoperative Wound Infection?” (N = 315 Cases)

Concordance Level	No. Cases (%)
Data in agreement:	
3 fields positive 1 supportive	24 (8)
Supportive data:	
2 fields positive 1 absent 1 supportive	36 (11)
2 fields positive 2 supportive	9 (3)
1 field positive 2 absent 1 supportive	24 (8)
2 fields supportive 2 absent	1 (0.3)
1 field supportive 3 absent	124 (39)
Absent data:	
3 fields positive 1 absent	25 (8)
2 fields positive 2 absent	31 (10)
1 field positive 3 absent	23 (7)
Contradictory data:	
2 fields positive 1 supportive 1 contradictory	1 (0.3)
2 fields positive 1 absent 1 contradictory	10 (3)
1 field positive 2 absent 1 contradictory	1 (0.3)
2 fields absent 1 supportive 1 contradictory	6 (2)

Table 6 ■

Types of Data Identified in Fields Searched to Answer the Question “What Kind of Hernia Did the Patient Have (Inguinal, Femoral, Umbilical, Ventral)?”

Field (Type)	Type of Data			
	Positive	Supportive	Absent	Contradictory
All cases (N = 200):				
Postoperative diagnosis (coded)	191	8	1	–
Primary procedure (coded)	193	2	–	5
Dictated operative note (text)	195	–	–	5
Cases in Which Discharge Summary Was Available (n = 95):				
ICD-9-coded discharge diagnosis (coded)	92	2	–	1
Dictated discharge summary (text)	92	–	–	3

NOTE: No discharge summary or ICD-9 data were available for 105 charts.

Table 7 ■

Levels of Concordance among Data Fields Searched to Answer the Question “What Kind of Hernia Did the Patient Have?” (N = 200)

Concordance Level	Cases with Discharge Summary (n = 95)		Cases without Discharge Summary (n = 105)	
	Fields (5 possible)	No. Cases (%)	Fields (3 possible)	No. Cases (%)
Data in agreement:	5 positive	74 (78)	3 positive	99 (94)
Supportive data:	4 positive } 1 supportive }	4 (4)	2 positive } 1 supportive }	1 (1)
	3 positive } 2 supportive }			
Absent data:	4 positive } 1 absent }	2 (2)	2 positive } 1 absent }	1 (1)
Contradictory data:	4 positive } 1 contradictory }	7 (7)	2 positive } 1 contradictory }	4 (4)
	3 positive } 1 supportive } 1 contradictory }	1 (1)		
	3 positive } 2 contradictory }	5 (5)		

NOTE: Percentage totals may not equal 100, because of rounding.

include the data, although 12 records contained only “hernia” rather than “inguinal hernia” or “umbilical hernia” and were, therefore, considered only supportive. In 17 charts, contradictory data were stored in the fields (Table 7). Either a different type of hernia was noted (12 charts) or the surgeon identified a seroma or abscess, rather than a hernia, as the source of the bulge (5 charts). Of the cases in which an ICD-9 diagnosis was specified, that diagnosis was in agreement with the free-text note in 93 percent, with the coded postoperative diagnosis in 86 percent, and with the coded primary procedure in 90 percent.

In determining whether the inguinal hernia was direct, indirect, or recurrent, few of the coded fields were found to include this information (Table 8). Only the operative note reliably identified the type of inguinal hernia. The coded fields recorded only whether a recurrent hernia was present and this only sporadically. In two cases (both of recurrent hernia), there was agreement in all fields (Table 9). The operative note identified a recurrent hernia in nine other cases. The operative note was worded in such a way that the answer could be determined in all but nine cases. In 28 cases the operative note had to be interpreted

Table 8 ■

Types of Data Identified in Fields Searched to Answer the Question "What Type of Inguinal Hernia Did the Patient Have (Direct, Indirect, or Recurrent)?"

Field (Type)	Type of Data			
	Positive	Supportive	Absent	Contradictory
All cases (<i>N</i> = 151):				
Postoperative diagnosis (coded)	4	–	147	–
Primary procedure (coded)	2	–	149	–
Dictated operative note (text)	114	28	9	–
Cases in Which Discharge Summary Was Available (<i>n</i> = 64):				
ICD-9-coded discharge diagnosis (coded)	6	–	58	–
Dictated discharge summary (text)	19	–	45	–

NOTE: No discharge summary or ICD-9 data were available for 87 charts.

Table 9 ■

Levels of Concordance among Data Fields Searched to Answer the Question "What Kind of Inguinal Hernia Did the Patient Have (Direct, Indirect, or Recurrent)?" (*N* = 151)

Concordance Level	Cases with Discharge Summary (<i>n</i> = 64)		Cases without Discharge Summary (<i>n</i> = 87)	
	Fields (5 possible)	No. Cases (%)	Fields (3 possible)	No. Cases (%)
Data in agreement:	5 positive	1 (2)	3 positive	1 (1)
Supportive data:	1 positive 1 supportive 3 absent	1 (2)	1 supportive 2 absent	16 (18)
	1 supportive 4 absent		11 (17)	
Absent data:	3 positive 2 absent	2 (3)	2 positive 1 absent	2 (2)
	2 positive 3 absent	14 (22)	1 positive 2 absent	63 (72)
Contradictory data:	1 positive 4 absent	29 (45)	3 absent	5 (6)
	5 absent	4 (6)		
Contradictory data:	2 positive 2 absent 1 contradictory	1 (2)		
	1 positive 3 absent 1 contradictory	1 (2)		

NOTE: Percentage totals may not equal 100, because of rounding.

using surgical judgment, since the type of hernia was not specified. For example, the statement "hernia sac was located on the anteromedial surface of the cord" was considered supportive of an indirect hernia.

The side of the inguinal hernia was included in the postoperative diagnosis and primary procedure fields

as well as the dictated operative note. It was surprising that the operative note omitted the side in 13 reports (Tables 10 and 11). The coded ICD-9 discharge diagnosis has no code for side, so the data in this field agreed with the others only in the records of patients with bilateral hernia (which is a coded diagnosis). The code in the field for postoperative diagnosis contra-

Table 10 ■

Types of Data Identified in Fields Searched to Answer the Question "Which Side Was the Hernia On?"

Field (Type)	Type of Data			
	Positive	Supportive	Absent	Contradictory
All cases (N = 155):				
Postoperative diagnosis (coded)	143	–	12	–
Primary procedure (coded)	147	–	8	–
Dictated operative note (text)	142	–	13	–
Cases in Which Discharge Summary Was Available (n = 65):				
ICD-9-coded discharge diagnosis (coded)	5	–	60	–
Dictated discharge summary (text)	64	–	1	–

NOTE: No discharge summary or ICD-9 data were available for 90 charts.

Table 11 ■

Levels of Concordance among Data Fields Searched to Answer the Question "Which Side Was the Hernia On?" (N = 155)

Concordance Level	Cases with Discharge Summary (n = 65)		Cases without Discharge Summary (n = 90)	
	Fields (5 possible)	No. Cases (%)	Fields (3 possible)	No. Cases (%)
Data in agreement:	5 positive	31 (48)	3 positive	73 (81)
Absent data:	4 positive } 1 absent }	17 (26)	2 positive } 1 absent }	10 (11)
	3 positive } 2 absent }	9 (14)	1 positive } 2 absent }	1 (1)
	2 positive } 3 absent }	4 (6)		
	1 positive } 4 absent }	1 (2)		
Contradictory data:			2 positive } 1 contradictory }	5 (6)
	3 positive } 1 absent } 1 contradictory }	1 (2)		
	2 positive } 1 absent } 2 contradictory }	2 (3)		
			3 contradictory	1 (1)

NOTE: Percentage totals may not equal 100, because of rounding.

dicted the dictated operative report in nine cases; thus, there was an explicit disagreement in 5 percent of cases on whether the hernia was on the left or right side!

Only the operative report reliably noted whether plastic mesh was used in the repair (Tables 12 and 13), even though the field for primary procedure allowed for its

inclusion and actually did include it in three cases. The dictated discharge summary noted this fact in 25 cases, and the primary procedure noted it in three.

It appears that the dictated operative report most reliably recorded information about the hernia operation. The coded fields completed in the operating room appeared less complete.

Table 12 ■

Types of Data Identified in Fields Searched to Answer the Question "Was Mesh Used for Hernia Repair?"

Field (Type)	Type of Data			
	Positive	Supportive	Absent	Contradictory
All cases (<i>N</i> = 200):				
Postoperative diagnosis (coded)	–	–	200	–
Primary procedure (coded)	3	–	197	–
Dictated operative note (text)	140	–	60	–
Cases in Which Discharge Summary Was Available (<i>n</i> = 95):				
ICD-9-coded discharge diagnosis (coded)	–	–	95	–
Dictated discharge summary (text)	25	–	70	–

NOTE: No discharge summary or ICD-9 data were available for 105 charts.

Table 13 ■

Levels of Concordance among Data Fields Searched to Answer the Question "Was Mesh Used for Hernia Repair?" (*N* = 200)

Concordance Level	Cases with Discharge Summary (<i>n</i> = 95)		Cases without Discharge Summary (<i>n</i> = 105)	
	Fields (3 possible)	No. Cases (%)	Fields (2 possible)	No. Cases (%)
Data in agreement:	3 positive	2 (2)	2 positive	1 (1)
	3 negative (absent)	36 (38)	2 negative (absent)	24 (23)
Absent data:	2 positive } 1 absent }	24 (25)		
	1 positive } 2 absent }	33 (35)	1 positive } 1 absent }	80 (76)

Discussion

Contemporary medicine requires that an increasing quantity of clinical information be stored and readily available for retrieval in CDRs. In most CDRs more than one field can be expected to contain the data to answer a given question. Some of this information may be stored in coded data fields and some in free-text data fields.

Our study had some inherent limitations. We looked at only one database, three broad questions, and a limited number of charts that dealt with these questions. In interpreting free text and those fields with short-text entry (postoperative diagnosis and primary procedure), we made an effort to be inclusive by searching for all likely possibilities and spellings of key words. Also, only one expert interpreted the free text.

In this study we were interested in the relationship between coded and textual data and in whether the

information in some data fields conflicted with or confirmed the information in other data fields, rather than in the accuracy of information in individual data fields. Accuracy is difficult to determine. For example, if one data field states that the patient had a right inguinal hernia and the other indicates a left inguinal hernia, it might be impossible to determine which is correct by just reading the chart.

Several papers have documented the inaccuracy of information stored in the electronic patient record, but little has been written about conflicting information in different types of fields. Hogan and Wagner¹² showed that if a clinician were allowed to supplement the coded fields with free text, the results would frequently be conflicting. In that study, the main reason given by clinicians for entering free text was their opinion that a situation could not be adequately expressed by coded text.

Our study considered the concordance of the coded and free-text fields in a database in which different

people completed the data fields at different times, independent of one another. We found that, depending on the question asked, there was concordance between at least two queried data fields 43 to 100 percent of the time and contradiction between at least two fields 4 to 13 percent of the time. The study excluded all charts lacking pertinent free text. Including them would have made concordance less because of the absent data.

Jollis et al.,¹⁴ in a paper on discordance of ICD-9 and clinical data, wrote that “different purposes of the two data sources best explain their disagreement.” They also identified data entry errors and timing (event occurring after entry of data) as sources of discordance.

Because of the previously described discrepancies between the administrative ICD-9-coded data and clinically derived data, we looked at the agreement of our ICD-9-coded field with the free-text fields and compared this with the agreement of other clinical coded fields (“postoperative complication” in answering the question about pulmonary embolism and “postoperative diagnosis” and “primary procedure” in answering the question about the type of hernia). These questions were chosen because the pertinent ICD-9 codes had sufficient granularity to answer them. Our results indicate that when the ICD-9 code is present and has sufficient granularity, its agreement with the free text is similar to that of other coded fields. This raises the question whether the previously noted discordance of the ICD-9-coded field, which is administrative in nature, with data fields in the clinical data set reflects a broader problem, namely, the general level of discordance between many data elements in a clinical database.

Previous authors have identified several factors that affect the accuracy of the data, such as transcription error¹¹ and patient error.² After searching the data fields, reading the free text, and speaking to some of the personnel involved in entering the data, we developed several hypotheses about why the information in data fields conflicts and is not always complete and accurate. We concluded that the completeness and accuracy of the data in any field depend on the type of field, the type of information contained in the field, the person entering the data, and the timing of the entry in relation to the event or diagnosis. We identified several issues that may influence the quality of the data:

- The motivation and expertise of the person entering the data vary widely, but not always in the same direction. For example, surgeons are experts on the data being entered but have little incentive to re-

cord wound infections accurately. They are, however, unlikely to omit a major clinical event like a pulmonary embolism. On the other hand, the clinical significance of a complication may be unknown to a medical record coder.

- The allowance of default entries may make it easier to omit corrections, as appeared to be the case with the postoperative diagnosis in the operating room.
- Absence of data could represent either negative data or incomplete entry of data into a field. How the absence should be interpreted depends on the field and its contents. If the side on which a hernia is present is not included, it is obviously an omission of data. The reason for the failure to record a complication is not always clear.
- Granularity of the clinical terms can be an issue. Some of the coded data are too general to answer questions definitively and can only suggest an answer. An example of this is the ICD-9 code used for wound infection, which is too general to be useful. The terminology used in free text may also be too nonspecific to allow for a definitive answer to a question.
- It is logical to expect that the sooner after an event a field is completed, the better the chance the data in the chart will be accurate. Events occurring after data entry may also render the data inaccurate, when a wound infection is detected after discharge.
- In computing the incidence of an entity, the numerator is usually clear (if not always accurate), but there may be a question about which denominator to use, an observation also noted by Iezzoni.⁸ For example, the denominator may be an operation, a discharge, or a patient. Wound infections and other postoperative complications are usually calculated using the number of operations as the denominator. It therefore makes sense that the “postoperative complication” and the “postoperative infection” fields are linked to a specific operation. The discharge summary, which may also record wound infection, may cover many operations and is linked to a hospital discharge, as is the ICD-9-coded discharge diagnosis.

Summary

The goal of this study was to investigate the discordance between data items, especially between free text and coded text items, and to illustrate and define some problems involved in answering questions of the CDR. Searching the database revealed many fields that contained data that might be helpful in answer-

ing any one specific clinical question. Some of these data were stored in coded data fields and some in free-text fields. When the data were compared, poor concordance between the fields was often found. This analysis suggests several conclusions:

- The database cannot just be queried in an automated fashion and the results reported.
- Multiple fields must be queried to get an initial assessment. This can be expected to provide information that may be confirmatory, complementary, or conflicting.
- Both coded and textual fields must be searched to obtain the fullest assessment.
- Final answers to questions require human judgment and integration of the information as a whole, and additional information gathering (e.g., chart review, contact with providers or patients) may be required to obtain the most accurate information possible.
- Since free text is an invaluable part of the database, algorithms must be established to make it possible to search the information in the free text, as well as the coded data, when answering questions.

References ■

1. Payne TH, Martin DR. How useful is the UMLS metathesaurus in developing a controlled vocabulary for an automated problem list? *Proc 17th Annu Symp Comput Appl Med Care*. 1993:705-9.
2. Wagner MM, Hogan WR. The accuracy of medication data in an outpatient electronic medical record. *J Am Med Inform Assoc*. 1996;3(3):234-44.
3. Wilton R, Pennisi AJ. Evaluating the accuracy of transcribed clinical data. *Proc 17th Annu Symp Comput Appl Med Care*. 1993:279-83.
4. Yeakel AE. Evaluation of accuracy in a clinical data-processing system. *Anesth Analg*. 1967;46(1):18-24.
5. Hobbs FD, Parle JV, Kenkre JE. Accuracy of routinely collected clinical data on acute medical admissions to one hospital. *Br J Gen Pract*. 1997;47(420):439-40.
6. Hohnloser JH, Puerner F, Soltanian H. Improving coded data entry by an electronic patient record system. *Methods Inf Meth*. 1996;35(2):108-11.
7. Hogan WR, Wagner MM. Accuracy of data in computer-based patient records. *J Am Med Inform Assoc*. 1997;4(5):342-55.
8. Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med*. 1997;127(8 pt 2):666-74.
9. Green J, Wintfeld N. Report cards on cardiac surgeons: assessing New York State's approach. *N Engl J Med*. 1995;332(18):1229-32.
10. Young AS, Sullivan G, Burnam MA, Brook RH. Measuring the quality of outpatient treatment for schizophrenia. *Arch Gen Psychiatry*. 1998;55(7):611-7.
11. Horbar JD, Leahy KA. An assessment of data quality in the Vermont-Oxford Trials Network database. *Control Clin Trials*. 1995;16(1):51-61.
12. Hogan WR, Wagner MM. Free-text fields change the meaning of coded data. *Proc AMIA Annu Fall Symp*. 1996:517-21.
13. Cooper F, Buchanan BG, Kayaalp M, Saul M, Vries JK. Using computer modeling to help identify patient subgroups in clinical data repositories. *Proc AMIA Annu Fall Symp*. 1998:180-4.
14. Jollis JG, Ancukiewicz M, DeLong ER, Pryor DB, Muhlbaier LH, Mark DB. Discordance of databases designed for claims payment versus clinical information systems: implications for outcomes research. *Ann Intern Med*. 1993;119(8):844-50.
15. Horan TC, Gaynes RP, Martone WJ, Jarvis WR, Emori TG. CDC definitions of nosocomial surgical site infections, 1992: a modification of CDC definitions of surgical wound infections. *Am J Infect Control*. 1992;20(5):271-4.