

Research Paper ■

The Value of Patient Self-report for Disease Surveillance

FLORENCE T. BOURGEOIS, MD, MPH, STEPHEN C. PORTER, MD, MPH, CLARISSA VALIM, ScD, MD, TIFFANY JACKSON, BS, E. FRANCIS COOK, ScD, KENNETH D. MANDL, MD, MPH

Abstract **Objective:** To determine the accuracy of self-reported information from patients and families for use in a disease surveillance system.

Design: Patients and their parents presenting to the emergency department (ED) waiting room of an urban, tertiary care children's hospital were asked to use a Self-Report Tool, which consisted of a questionnaire asking questions related to the subjects' current illness.

Measurements: The sensitivity and specificity of three data sources for assigning patients to disease categories was measured: the ED chief complaint, physician diagnostic coding, and the completed Self-Report Tool. The gold standard metric for comparison was a medical record abstraction.

Results: A total of 936 subjects were enrolled. Compared to ED chief complaints, the Self-Report Tool was more than twice as sensitive in identifying respiratory illnesses (Rate ratio [RR]: 2.10, 95% confidence interval [CI] 1.81–2.44), and dermatological problems (RR: 2.23, 95% CI 1.56–3.17), as well as significantly more sensitive in detecting fever (RR: 1.90, 95% CI 1.67–2.17), gastrointestinal problems (RR: 1.10, 95% CI 1.00–1.20), and injuries (RR: 1.16, 95% CI 1.08–1.24). Sensitivities were also significantly higher when the Self-Report Tool performance was compared to diagnostic codes, with a sensitivity rate ratio of 4.42 (95% CI 3.45–5.68) for fever, 1.70 (95% CI 1.49–1.93) for respiratory problems, 1.15 (95% CI 1.04–1.27) for gastrointestinal problems, 2.02 (95% CI 1.42–2.87) for dermatological problems, and 1.06 (95% CI 1.01–1.11) for injuries.

Conclusions: Disease category assignment based on patient-reported information was significantly more sensitive in correctly identifying a disease category than data currently used by national and regional disease surveillance systems.

■ J Am Med Inform Assoc. 2007;14:765–771. DOI 10.1197/jamia.M2134.

Introduction

A priority of the public health system is early and accurate detection of disease outbreaks. Over the past seven years, there has been an emergence of real time surveillance systems for monitoring illness patterns to enable health officials to identify and localize disease outbreaks.^{1–3} These systems use information routinely collected in patient care

and available at the time of the patient's interaction with his healthcare provider. Data types frequently used by surveillance systems, including BioSense and the Real-time Outbreak and Disease Surveillance (RODS), are emergency department chief complaints and physician diagnostic coding.^{1–5} The majority of emergency departments collect these data for every patient encounter and the data lend themselves well to use in automated surveillance systems since they are frequently electronically and rapidly available.⁵ The surveillance systems use these data to assign patients to disease syndromes of interest, and the patterns of patients presenting with a specific disease type are tracked and compared with historical patterns.

However, studies of systems using chief complaint and diagnoses data have shown that these data types can be limited in disease detection.^{6–12} For example, the sensitivities for the detection of respiratory infections range from 26% to 47% across studies using chief complaint data.^{6,7,12} Proper disease category assignment is critical to outbreak detection and changes in identification methods have a large impact on the modeling and outbreak detection performance of surveillance systems.¹² The use of chief complaints in disease surveillance is limited because they are usually stored as short, unstructured strings of text. They are not collected for the purpose of surveillance and hence used by disease surveillance systems as a secondary data source. While there has been considerable success "reading" these

Affiliations of the authors: Division of Emergency Medicine (FTB, SCP, TJ, KDM), Clinical Research Program (CV), Children's Hospital, Boston, MA; Department of Pediatrics (FTB, SCP, KDM), Harvard Medical School, Boston, MA; Department of Biostatistics (CV), Department of Epidemiology (EFC), Harvard School of Public Health, Boston, MA; Children's Hospital Informatics Program at the Harvard-MIT Division of Health Sciences and Technology (SCP, KDM), Cambridge, MA

This paper was presented in part at the Syndromic Surveillance Conference, Seattle, WA, September 14–16, 2005, and at the Pediatric Academic Societies' Annual Meeting, San Francisco, CA, April 28–May 2, 2006.

This work was supported by National Research Service Awards 5 T32 HD40128-03 and 5 T32 HD40128-04, by R01LM007970-01 from the National Library of Medicine, National Institutes of Health.

Correspondence: Florence T. Bourgeois MD, MPH, Division of Emergency Medicine, Children's Hospital Boston, 300 Longwood Avenue, Boston, MA 02115; e-mail: <florence.bourgeois@childrens.harvard.edu>.

Received for review: 04/24/06; accepted for publication: 07/14/07.

short text strings with natural language text classifiers, and using them to assign people to illnesses of interest (for example, influenza-like illness), results have been variable.^{8,10,12,13}

An alternative to using routinely collected data is to have hospital staff manually enter information on every patient at the time of the visit. This method may be useful in disease identification, but it is very labor intensive and not sustainable in the long term.^{14,15} This was demonstrated with a system requiring manual data entry by hospital staff which was implemented in New York City following the 2001 terrorist attacks.¹⁵ Even in the short term, completion rates were modest and inconsistent across hospital sites. We propose that patients could be leveraged to provide information directly about their conditions, using interfaces such as electronic personally controlled health records, emergency department-based kiosks, or mobile devices.¹⁶ Although the cost of this process would have to be considered, with the increasing use of patient-centered health information technologies, there is a strong trend towards collecting patient-generated data.¹⁷ The data elements collected could be tailored to specific purposes and patients could participate when they present to a hospital or clinic, or from home, work, and other non-healthcare settings, enabling sampling from a broader population.

To determine whether such a system of data collection is feasible, we must first ascertain whether patients can provide valid medical information. Studies have shown that patients as well as parents of young patients can accurately report information about their own or their child's illness using electronic interfaces.^{18–23} These studies have demonstrated that valid medical data, such as illness symptoms and medication use, can be collected electronically from parents in clinical settings and that socio-demographic or technology-related factors do not have an effect on data quality.^{20,21} More information is needed to describe the types of data that patients can provide and the validity of these data.

Here, we evaluate the use of patient- and parent-produced information for case detection by disease surveillance systems. The objectives were (1) to determine the accuracy of self-reported information as compared with existing data sources in assigning patients to specific disease categories, and (2) to evaluate the usability of the Self-Report Tool, a questionnaire for collecting illness information from parents and patients.

M

Setting and Participants

Participants were recruited in the emergency department (ED) waiting room of an urban, academic tertiary care children's hospital between February 2004 and March 2005. The ED serves approximately 51,000 children per year with an average daily census of 140 children. Research assistants approached all patients and their families waiting to be seen by a physician during selected periods and screened them for eligibility. Recruitment periods included all times of day as well as weekdays and weekend days. The study inclusion criteria were patient age of 22 years or younger, accompaniment by a parent if younger than 18 years of age, and verbalized comfort with reading and writing in English.

During the time periods when a research assistant was recruiting patients, between approximately zero and forty patients were in the waiting room and all patients waiting to be seen by a physician were approached for enrollment. Only very rarely did an eligible patient choose not to participate and information on these patients was not collected. Overall, based on ED and research assistant logs, we estimate that approximately 85% of patients presenting to the ED during recruitment periods were enrolled. After informed consent was obtained from participants, a written Self-Report Tool was administered to the patient and parent. Parents completed surveys for patients up to 13 years of age while patients 14 years of age or older completed the survey themselves. A separate follow up survey was administered to participants after completion of the Self-Report Tool asking them to provide information on their experiences completing the questions. If patients were placed in an examination room prior to finishing the study, they or their parents completed the Self-Report Tool in the room before the patient was seen by a physician, but did not complete the follow-up survey. The Committee on Clinical Investigation at Children's Hospital Boston approved this study.

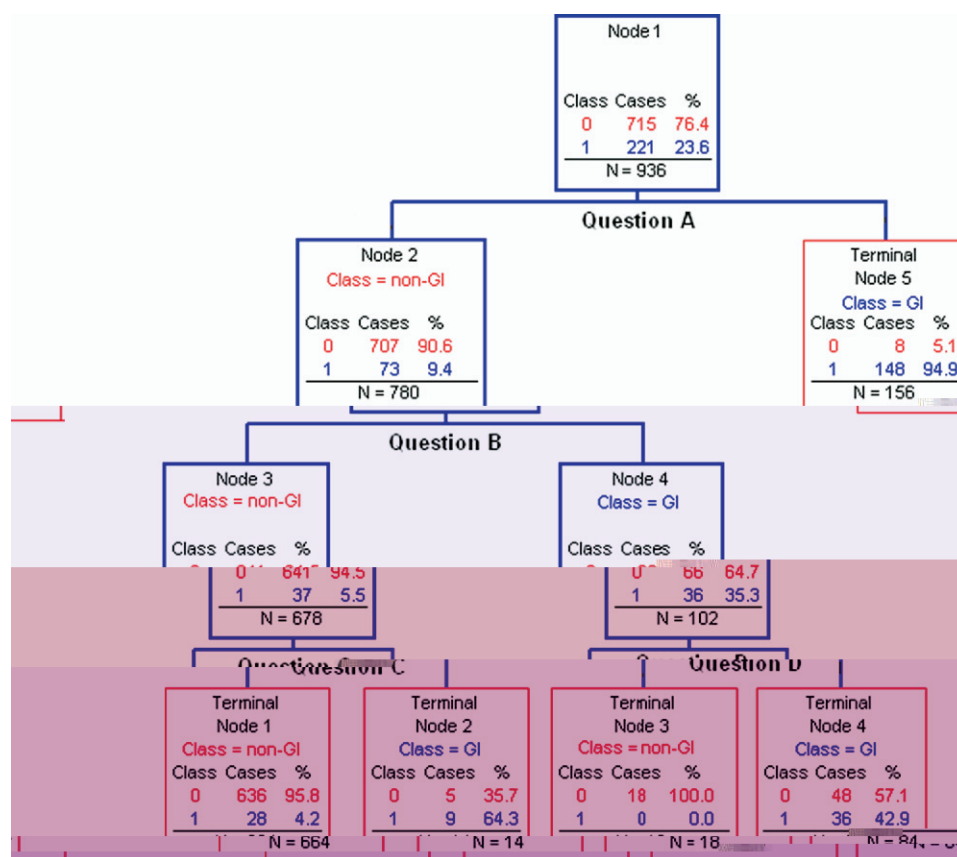
Survey Instrument

The Self-Report Tool consisted of four separate age-specific questionnaires (Appendix A). The age groups were 0 to 1 years of age, 2 to 5 years of age, 6 to 13 years of age, and 14 to 22 years of age. The questionnaires contained a total of 28 to 30 questions pertaining to the patient's current illness and past medical history. Age-specific questions varied in the wording between questionnaires while some questions were questionnaire-specific but corresponded to certain questions in the other questionnaires. Only a few questions were asked just within a specific age group. Responses were used to assign disease categories to patients. Questionnaires were pilot-tested for question clarity and face validity before the final Self-Report Tool was implemented. Revisions were made to the Self-Report Tool based on review of the questions with patients and comparisons between the information in the Self-Report Tool and the patient chart.

Measurements

Patients were assigned disease categories using three different data sources: the ED chief complaint, physician diagnostic coding, and the completed Self-Report Tool. Each of these was assessed against a gold standard medical record review in order to compare their ability to correctly assign disease categories to patients.

Chief complaints are routinely elicited by a triage nurse for every patient upon presentation to our ED. An administrative assistant in the ED subsequently numerically encodes these free-text complaints using a constrained list of 181 possible chief complaint codes and enters them into an electronic log. For this study, two physicians reviewed these codes and classified them into one of the following seven disease categories: respiratory problem, gastrointestinal problem, allergic problem, dermatologic problem, neurological problem, injury, or other. In addition, they identified chief complaint codes for fever. In order for a chief complaint to be included in a specific category, both physicians had to classify the chief complaint in the specific category; otherwise the chief complaint was added to the "other" category.



F. 1. Recursive partitioning tree for classification of gastrointestinal problems. GI = gastrointestinal problem, Class = disease category assigned by the tree, Case = actual outcome as measured by gold standard, N = total number of subjects in each node of the tree.

Following a patient's ED visit, the ED physician creates an electronic record for the visit and assigns a diagnosis. The diagnosis includes a free-text diagnostic impression and/or an International Classification of Diseases, Ninth Revision (ICD-9) diagnostic code. A billing administrator then reviews all diagnosis and ensures that an ICD-9 diagnostic code is provided for every patient visit. The same two physicians reviewed all ICD-9 codes and chief complaints assigned to patients enrolled in the study, classified them into the seven disease categories of interest, and identified codes for fever. The reviewers were blinded to the patient chief complaints and medical record.

Multivariate analyses with recursive partitioning, also referred to as classification and regression tree analysis (CART 5.0, Salford Systems, San Diego, CA), was used to assign disease categories using the data collected in the Self-Report Tool. Recursive partitioning has the advantage of providing simple to use prediction algorithms based on the process of stratification by identifying important interactions among the predictors in the data set. Data from the four age-specific questionnaires were combined. A separate classification tree was created for each of the eight categories and validated using cross-validation. The validation procedure that CART performed was 10-fold cross-validation where all the data was divided into 10 mutually exclusive subsets. Each subset used 10% of the data to validate a sequence of trees built on the remaining 90% of the data.

Figure 1 shows the tree that this method created for the

detection of gastrointestinal problems. Each branching point seeks to correctly classify the maximum number of subjects as having or not having a gastrointestinal problem based on a specific question in the Self-Report Tool. For example, of the 221 subjects with gastrointestinal problems, 148 were correctly identified based on the subject's response to "Question A" while 8 subjects were incorrectly placed in this category. Further classification based on the results of 2 more questions correctly identified an additional 45 subjects for a total of 193 of 221 total cases with gastrointestinal problems. After the cross-validation procedure, 190 subjects remained in the gastrointestinal category for a final sensitivity of 86% (190/221). This process was repeated for the assignment of the other disease categories.

Two independent reviewers, who were blinded to the ED chief complaint, the diagnostic code, and the Self-Report Tool results, determined the gold standard disease category. They reviewed each subject's ED record and assigned a disease category based on a pre-determined set of criteria drawn from the patient history, physical examination, test results, and treatment.⁷ For every record, the presence or absence of a fever was noted and up to two disease categories assigned. Patients presenting with an injury could only be assigned the category of injury since we sought to identify the illness that brought a patient to the ED. Similarly, patients assigned the category of "other" could only be assigned to that group. A third reviewer examined and

Table 1 ■ Questionnaire Age Group (N=936)

Age group (years)	Percentage of patients
0–1	29
2–5	28
6–13	26
14–22	17

resolved differences in disease category assignments between the two reviewers.

Data Analysis

The sensitivity and specificity of the ED chief complaint, diagnostic code, and the Self-Report Tool in case detection were determined. The case-detection performance of the Self-Report Tool was compared to the performance of the other two measures using sensitivity and specificity rate ratios with 95% confidence intervals. These data analyses were performed using SAS Version 9.0 (SAS Institute, Cary, NC).

R During the study period, 936 subjects were enrolled. There were approximately equal numbers of patients in the younger age groups and fewer patients between 14 to 22 years of age (Table 1). Based on the gold standard chart review, 278 (30%) subjects presented with a fever and the three most common disease categories were respiratory problem (255, 27%), gastrointestinal problem (209, 22%), and injury (196, 21%) (Table 2).

The ED chief complaint was available for 917 (98%) of the 936 patients enrolled and the sensitivities and specificities of this data source in case detection determined (Table 3). Sensitivities ranged from 37% for dermatological problems to 84% for injury. Chief complaints detected most disease categories with sensitivities less than 70%; fever was identified with a sensitivity of 48%. Specificities were high with values of 96% or greater except for the group of other problems, which had a specificity of 87%.

The diagnostic codes assigned to patient visits were available for 820 (88%) of the 936 subjects. The sensitivities of these were similar to the ones for the chief complaints, with a few exceptions (Table 3). The classification by diagnostic codes had a sensitivity of 21% for fever and of zero for allergic problems. The specificities were high, all reaching 99% except for the categories of injury (97%) and other problems (79%).

For the Self-Report Tool, the sensitivities were greater than 70% for the detection of all categories, with the lowest values found for other problems (71%) and neurological problems (77%) (Table 3). The Self-Report Tool detected fever with a sensitivity of 91% and respiratory and gastrointestinal problems with sensitivities of 88% and 86%, respectively. The specificities were lower than with the chief complaints and diagnostic codes, but none lower than 84%.

Sensitivity and specificity ratios were calculated to further compare the results of the Self-Report Tool to those of the ED chief complaints and diagnostic codes (Table 4). Compared to ED chief complaints, the Self-Report Tool

was significantly more sensitive in identifying fever, with a rate ratio of 1.90 (95% confidence interval [CI] 1.67–2.17), indicating that there was a 90% increase in the detection of fever using the Self-Report Tool as compared to the chief complaints. The Self-Report Tool was more than twice as sensitive in identifying respiratory problems (Rate ratio [RR]: 2.10, 95% CI 1.81–2.44), and dermatological problems (RR: 2.23, 95% CI 1.56–3.17), as well as significantly more sensitive in detecting gastrointestinal problems (RR: 1.10, 95% CI 1.00–1.20) and injuries (RR: 1.16, 95% CI 1.08–1.24). Sensitivities were also significantly higher when compared to diagnostic codes, with a rate ratio of 4.42 (95% CI 3.45–5.68) for fever, 1.70 (95% CI 1.49–1.93) for respiratory problems, 1.15 (95% CI 1.04–1.27) for gastrointestinal problems, 2.02 (95% CI 1.42–2.87) for dermatologic problems, and 1.06 (95% CI 1.01–1.11) for injuries.

Specificities were slightly lower for almost all categories for the Self-Report Tool compared with the ED chief complaints and diagnostic codes, but only a few of the specificity ratios were less than 0.90. These included respiratory problems (0.86, 95% CI 0.83–0.90) and dermatologic problems (0.87, 95% CI 0.85–0.90) when compared with ED chief complaints, and fever (0.88, 95% CI 0.86–0.91), respiratory problems (0.85, 95% CI 0.82–0.88), and dermatologic problems (0.87, 95% CI 0.85–0.90) when compared with diagnostic codes.

A total of 610 (65%) of the subjects completed the follow up survey on their experience using the Self-Report Tool (Table 5). Subjects' responses to the eight questions are shown in Table 5. Subjects' responses to eight questions on a 5-point scale indicated that both parents and patients did not find the questions difficult to answer and would for the most part be willing to use a similar tool again.

Discussion

Using patient and family self-reported information, we assigned cases to disease categories with sensitivities much higher than conventional data sources currently used by disease surveillance systems. Respiratory and gastrointestinal problems were identified using the Self-Report Tool with sensitivities of 88% and 86%, respectively, using patient reported information. By contrast,

Table 2 ■ Proportion of Patients by Disease Categories based on Gold Standard Assignments*

Disease category	First disease category, %	Second disease category, %	First or second disease category, %
Respiratory	27	49	28
Gastrointestinal	22	31	23
Dermatologic	6	20	6
Neurological	5	3	4
Allergic	1	0	1
Injury	21	N/A	20
Other	18	N/A	18
TOTAL	936	39	975

*Subjects could be assigned to up to 2 different disease categories, except if they were classified in the Injury or Other categories.

Table 3 ■ ED Chief Complaint, Diagnostic Code, and Self-Report Tool Classification*

Disease category	ED Chief Complaint		Diagnostic Code		Self-Report Tool	
	Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %
Fever	48 (130/272)	98 (634/645)	21 (50/243)	99 (576/577)	91 (253/278)	88 (579/658)
Respiratory	42 (112/269)	97 (628/648)	52 (130/252)	99 (561/568)	88 (239/273)	84 (555/663)
Gastrointestinal	78 (171/218)	96 (670/699)	75 (129/172)	99 (640/648)	86 (190/221)	92 (656/715)
Dermatologic	37 (22/60)	99 (853/857)	41 (21/52)	99 (765/768)	82 (49/60)	87 (762/876)
Neurological	60 (26/43)	99 (867/874)	67 (20/30)	99 (788/790)	77 (33/43)	90 (801/893)
Allergic	50 (5/10)	99 (904/907)	0 (0/8)	99 (811/812)	90 (9/10)	99 (916/926)
Injury	84 (161/192)	98 (712/725)	91 (172/188)	97 (615/632)	97 (190/196)	96 (708/740)
Other	62 (102/164)	87 (652/753)	75 (112/150)	79 (531/670)	71 (122/172)	84 (644/764)

*ED chief complaint available for 917/936 (98%) of subjects; diagnostic code available for 820/936 (88%) of subjects; Self-Report Tool available for 936 (100%) of subjects.

using chief complaint data, the sensitivities were 42% for respiratory problems and 78% for gastrointestinal problems. Similar results were obtained using diagnostic codes, indicating that patient- and parent-produced data can effectively be used to identify disease cases. The specificity for case-detection using the Self-Report Tool ranged from 84% to 99%, with the majority at 90% or greater. These specificities were generally lower than those found using traditional chief complaints, representing a trade-off for the higher sensitivities. Further work is needed to refine the questionnaires and maximize the performance of the Self-Report Tool. Results of the follow up survey indicated that both parents and patients did not find the questions in the Self-Report Tool difficult to answer and would for the most part be willing to use a similar tool again in the future.

Patient self-reported information lends itself well to use in disease surveillance for several reasons. First, the collection of data can potentially be performed electronically with participants providing information using a computerized version of the Self-Report Tool. Previous studies have shown that valid information on patients' illnesses can be collected from patients and parents of children using electronic interfaces.^{19–22} In these studies, detailed information on illness symptoms, medical history, and medication use was obtained using an electronic interface in the ED setting. The information was found to be accurate when validated against criterion interviews and medical record reviews, demonstrating the feasibility of electronically collecting medical information from patients in clinical settings. For certain data elements, the

parent's report was in fact found to be more accurate than the information obtained by nurses and physicians.²⁰ In addition, socio-demographic and technology-related factors were examined and found to be unrelated to data quality. Using the Self-Report Tool we developed, information could be collected either in the physician's office or from home through personally controlled health records and other consumer-facing applications, enabling a much larger and geographically diverse population to be under surveillance. The information could be rapidly processed and counts of different disease categories monitored in near real-time.

Secondly, the information collected is richer, including more data elements on individual patients than conventional surveillance data and providing the potential to examine and monitor more precise disease categories. Many illnesses, naturally occurring or maliciously introduced, present with a group of symptoms which has led to the use of syndromes to describe presentations such as sepsis, botulism, or hemorrhagic illnesses.^{13,14} Patient self-reported information, as collected using a set of questions in the Self-Report Tool, contains multiple data elements which can be combined to define syndromes and refine disease identification. Additional disease categories or syndromes could be derived as needed using the available information or by expanding the data collected by the Self-Report Tool.

There are several limitations to this study which warrant discussion. Two reviewers assigned the gold standard disease category with a third party to resolve differences.

Table 4 ■ Sensitivity and Specificity Rate Ratios: Self-Report Tool vs. ED Chief Complaints and Diagnostic Codes

Disease category	Sensitivities		Specificities	
	Ratio (95% CI) for Self-Report Tool and ED chief complaint	Ratio (95% CI) for Self-Report Tool and diagnostic code	Ratio (95% CI) for Self-Report Tool and ED chief complaint	Ratio (95% CI) for Self-Report Tool and diagnostic code
Fever	1.90 (1.67–2.17)	4.42 (3.45–5.68)	0.90 (0.87–0.92)	0.88 (0.86–0.91)
Respiratory	2.10 (1.81–2.44)	1.70 (1.49–1.93)	0.86 (0.83–0.90)	0.85 (0.82–0.88)
Gastrointestinal	1.10 (1.00–1.20)	1.15 (1.04–1.27)	0.96 (0.93–0.98)	0.93 (0.91–0.95)
Dermatologic	2.23 (1.56–3.17)	2.02 (1.42–2.87)	0.87 (0.85–0.90)	0.87 (0.85–0.90)
Neurological	1.27 (0.95–1.70)	1.15 (0.85–1.56)	0.90 (0.88–0.93)	0.90 (0.88–0.92)
Allergic	1.80 (0.94–3.46)	N/A	0.99 (0.98–1.00)	0.99 (0.98–1.00)
Injury	1.16 (1.08–1.24)	1.06 (1.01–1.11)	0.97 (0.96–0.99)	0.98 (0.96–1.00)
Other	1.14 (0.98–1.33)	0.95 (0.83–1.09)	0.97 (0.93–1.01)	1.06 (1.01–1.12)

Table 5 ■ Patient and Parent Feedback on Use of the Self-Report Tool

Survey items (1=strongly agree, 5=strongly disagree)	Response (mean \pm SD)
"I liked using this kind of questionnaire to give information about myself/my child."	2.2 \pm 0.9
"It was easy using the questionnaire to give information about myself/my child."	1.7 \pm 0.9
"I thought the questions about my/my child's illness were too private."	3.7 \pm 1.4
"I thought the questions asking for addresses and locations were too private."	2.6 \pm 1.4
"I would be willing to give information about myself/my child again using this type of questionnaire."	2.4 \pm 1.1
"I was annoyed that the questionnaire asked medical questions that did not apply to me/my child."	3.8 \pm 1.2
"I did not like answering questions about addresses and places."	2.8 \pm 1.4
"I thought the questions asking about places were too difficult to answer."	4.0 \pm 1.1

Although this is a rigorous approach, some of the cases may still have been misclassified. It is also possible that in assigning only two disease categories (or only one for subjects with injuries) we failed to identify additional disease categories that may have applied to the patient. In addition, we used a convenience sample of patients who were English speaking, accompanied by an adult if younger than 18 years of age, not so ill as to require immediate placement in a treatment room, and willing to participate, which may have resulted in less generalizable conclusions than if all patients presenting to the ED had been enrolled. However, a research assistant approached all patients present in the waiting room during recruitment periods with very few patients refusing to participate and approximately 85% of patients presenting to the ED were enrolled during recruitment periods. Furthermore, we have no reason to believe that patients who did not participate had medical problems more difficult to identify using the Self-Report Tool than those who did participate. A further potential limitation is that the Self-Report Tool we employed was administered in a paper format as opposed to an electronic version. To be used in surveillance, patient self-reported information would have to be converted to an electronic format or provided electronically in order to enable rapid analysis and reporting. Given the promising results of previous studies examining the validity of electronically collected patient information, it is unlikely that this format change would affect the information provided by participants, but we cannot be certain of this without testing an electronic version. Finally, the recursive partitioning trees were validated using cross-validation instead of using a separate validation sample. Prospective study of patient self-reports for surveillance is warranted.



Patient and parent self-reported data were more sensitive in identifying disease categories than either chief complaint or diagnostic data, which represent information currently used

by many national and regional surveillance systems. Involving patients in the disease category assignment process may be an effective method of improving the accuracy and augmenting the capabilities of real time disease surveillance systems.

References ■

1. Mandl KD, Overhage JM, Wagner MM, et al. Implementing syndromic surveillance: a practical guide informed by the early experience. *J Am Med Inform Assoc*. 2004;11(2):141–150. [Pre-Print published Nov 121, 2003; as doi:2010.1197/jamia.M1356].
2. Lombardo J, Burkom H, Elbert E, et al. A systems overview of the Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE II). *J Urban Health Jun* 2003;80(2 Suppl 1):i32–42.
3. Loonsk JW. BioSense—a national initiative for early detection and quantification of public health emergencies. *MMWR Morb Mortal Wkly Rep Sep* 24 2004;53 Suppl:53–55.
4. Tsui FC, Espino JU, Dato VM, Gesteland PH, Hutman J, Wagner MM. Technical description of RODS: a real-time public health surveillance system. *J Am Med Inform Assoc* 2003;10(5):399–408.
5. Lober WB, Karras BT, Wagner MM, et al. Roundtable on bioterrorism detection: information system-based surveillance. *J Am Med Inform Assoc* 2002;9(2):105–115.
6. Espino JU, Wagner MM. Accuracy of ICD-9-coded chief complaints and diagnoses for the detection of acute respiratory illness. *Proc AMIA Symp* 2001:164–168.
7. Beitel AJ, Olson KL, Reis BY, Mandl KD. Use of emergency department chief complaint and diagnostic codes for identifying respiratory illness in a pediatric population. *Pediatr Emerg Care* 2004;20(6):355–60.
8. Ivanov O, Wagner MM, Chapman WW, Olszewski RT. Accuracy of three classifiers of acute gastrointestinal syndrome for syndromic surveillance. *Proc AMIA Symp* 2002:345–349.
9. Tsui FC, Wagner MM, Dato V, Chang CC. Value of ICD-9 coded chief complaints for detection of epidemics. *Proc AMIA Symp* 2001 711–5.
10. Chapman WW, Dowling JN, Wagner MM. Fever detection from free-text clinical records for biosurveillance. *J Biomed Inform* 2004;37(2):120–7.
11. Chapman WW, Dowling JN, Wagner MM. Classification of emergency department chief complaints into 7 syndromes: a retrospective analysis of 527,228 patients. *Ann Emerg Med* 2005;46(5):445–55.
12. Reis BY, Mandl KD. Syndromic surveillance: the effects of syndrome grouping on model accuracy and outbreak detection. *Ann Emerg Med* 2004;44(3):235–41.
13. Chapman WW, Christensen LM, Wagner MM, et al. Classifying free-text triage chief complaints into syndromic categories with natural language processing. *Artif Intell Med* 2005; 33(1):31–40.
14. Fleischauer AT, Silk BJ, Schumacher M, et al. The validity of chief complaint and discharge diagnosis in emergency department-based syndromic surveillance. *Acad Emerg Med* 2004; 11(12):1262–7.
15. Das D, Weiss D, Mostashari F, et al. Enhanced drop-in syndromic surveillance in New York City following September 11, 2001. *J Urban Health Jun* 2003;80(2 Suppl 1):i76–88.
16. Porter SC. The Asthma Kiosk: A Patient-centered Technology for Collaborative Decision Support in the Emergency Department. *J Am Med Inform Assoc* 2004;(11):458–67.
17. Mandl KD, Szolovits P, Kohane IS. Public standards and patients' control: how to keep electronic medical records accessible but private. *BMJ* 2001;322(7281):283–7.
18. Pless CE, Pless IB. How well they remember. The accuracy of parent reports. *Arch Pediatr Adolesc Med* 1995;149(5):553–8.

19. Porter SC, Fleisher GR, Kohane IS, Mandl KD. The value of parental report for diagnosis and management of dehydration in the emergency department. *Ann Emerg Med* 2003; 41(2):196-205.
20. Porter SC, Kohane IS, Goldmann DA. Parents as partners in obtaining the medication history. *J Am Med Inform Assoc* 2005;12(3):299-305.
21. Porter SC, Silvia MT, Fleisher GR, Kohane IS, Homer CJ, Mandl KD. Parents as direct contributors to the medical record: validation of their electronic input. *Ann Emerg Med* 2000;35(4): 346-52.
22. Porter SC, Mandl KD. Data quality and the electronic medical record: a role for direct parental data entry. *Proc AMIA Symp* 1999:354-8.
23. Lutner RE, Roizen MF, Stocking CB, et al. The automated interview versus the personal interview. Do patient responses to preoperative health questions differ? *Anesthesiology* 1991;75(3): 394-400.