Information Extraction for Tracking Liver Cancer Patients' Statuses: From Mixture of Clinical Narrative Report Types

Xiao-Ou Ping, MS,[1] Yi-Ju Tseng, BS,[2] Yufang Chung, PhD,[3]
Ya-Lin Wu, MS,[2] Ching-Wei Hsu, MS,[2] Pei-Ming Yang, MD, PhD,[4]
Guan-Tarn Huang, MD, PhD,[4] Feipei Lai, PhD,[1,2,5]
and Ja-Der Liang, MD[4]

[1]Department of Computer Science and Information Engineering,
[5]Department of Electrical Engineering, and [2]Graduate Institute
of Biomedical Electronic and Bioinformatics, National Taiwan
University, Taipei, Taiwan.
[3]Department of Electrical Engineering, Tunghai University, Taipei,
Taiwan.
[4]Department of Internal Medicine, National Taiwan University
Hospital and National Taiwan University College of Medicine,
Taipei, Taiwan.

Abstract

Objective: To provide an efficient way for tracking patients' condition over long periods of time and to facilitate the collection of clinical data from different types of narrative reports, it is critical to develop an efficient method for smoothly analyzing the clinical data accumulated in narrative reports. Materials and Methods: To facilitate liver cancer clinical research, a method was developed for extracting clinical factors from various types of narrative clinical reports, including ultrasound reports, radiology reports, pathology reports, operation notes, admission notes, and discharge summaries. An information extraction (IE) module was developed for tracking disease progression in liver cancer patients over time, and a rule-based classifier was developed for answering whether patients met the clinical research eligibility criteria. The classifier provided the answers and direct/indirect evidence (evidence sentences) for the clinical questions. To evaluate the implemented IE module and the classifier, the gold-standard annotations and answers were developed manually, and the results of the implemented system were compared with the gold standard. Results: The IE model achieved an F-score from 92.40% to 99.59%, and the classifier achieved accuracy from 96.15% to 100%. Conclusions: The application was successfully applied to the various types of narrative clinical reports. It might be applied to the key extraction for other types of cancer patients.

Key words: information management, medical records, e-health, technology

Introduction

According to global cancer statistics in 2002, liver cancer was the sixth most common cancer worldwide and the third most common cause of death.[1] From another study of cancer statistics in 2010 in the United States, the death rate from liver cancer increased in both men and women between 1990–1991 and 2006.[2] Hepatocellular carcinoma (HCC), as the most common primary liver cancer,[3–5] has been the leading cause of cancer death in Taiwan since 1984.[6]

To provide an efficient way for tracking patients' condition over time and to facilitate the collection of clinical features relevant to HCC from narrative reports, it is critical to develop an efficient method for smoothly analyzing the clinical data accumulated in narrative reports, that is, the process of information extraction (IE) is required.

During the past two decades, researchers have successfully applied natural language processing to the medical narrative reports for case finding, case identifying, and IE for specific disease and topic (e.g., heart failure,[7,8] diabetes,[9] hypertension,[10] smoking,[11–13] and cancer[14–17]), for identifying and detecting problems[18] and adverse events,[19,20] for acquiring the relations (e.g., disease manifestation-related symptoms and adverse drug events[21]), and for assessing the quality of care,[22] among others.

The goal of this work was to provide liver cancer patients' statuses from different types of narrative reports for facilitating clinical research in three respects: (1) the single report level identified the major target information and the related attributes from each report, (2) the cross-report temporal level ordered and arranged these extracted results according to temporal order and report type for presenting the variance of each observed target or getting the status before/after specific time point, and (3) the summarized level further provided the answers of whether the patients' statuses met the criteria specified in the clinical research.

Materials and Methods

The proposed method had two phases for tracking patient disease progression over time and for answering whether the patients met the specific criteria of clinical research from the mixture of narrative reports. The overview of this method is shown in *Figure 1*. In the first phase, the IE module was developed for extracting the concepts related to patient disease from narrative reports. These extracted concepts were further ordered and arranged for grouping the duplicated information and presenting the variances of patient status over time. In the second phase, the rule-based classifier was
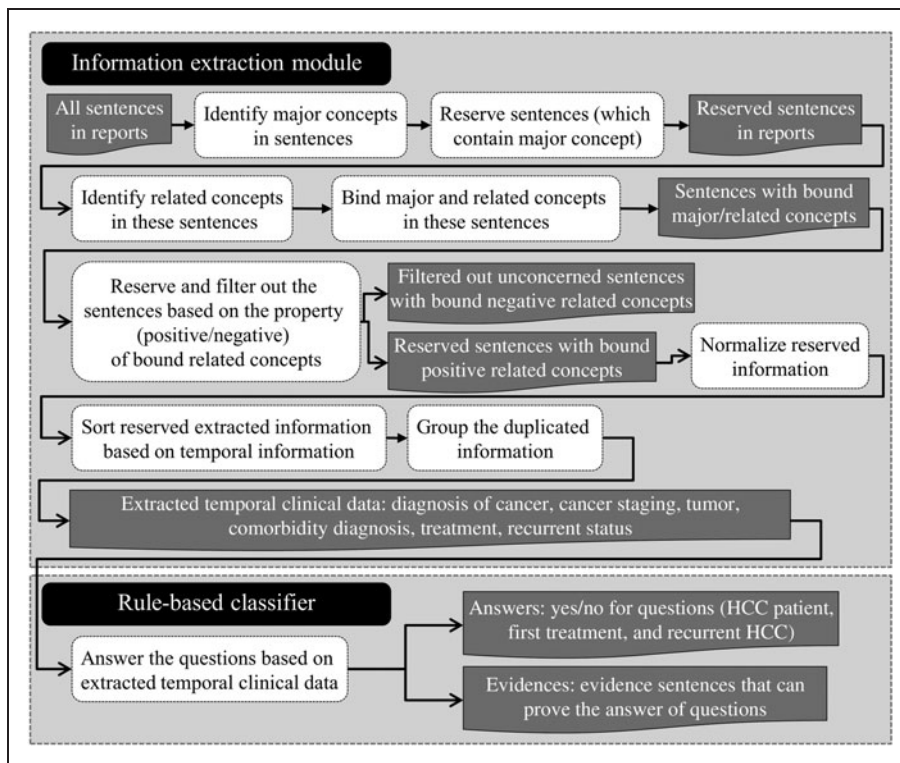
Fig. 1. The overview of this method: the information extraction module phase and the rule-based classifier phase. HCC, hepatocellular carcinoma.

developed for answering three clinical questions according to the arranged information.

SUBJECTS

The narrative reports of 152 patients receiving ultrasound-guided radiofrequency ablation in National Taiwan University Hospital (Taipei, Taiwan) between 2007 and 2009 were involved in the development and evaluation of this method. The development set contained the data of 74 patients who received radiofrequency ablation between 2007 and 2008. This dataset was used to develop the IE module and rule-based classifier. The testing set came from the data of 78 patients who received radiofrequency ablation in 2009. The narrative reports in the testing set were not previously reviewed or analyzed by system developers. This dataset was used to evaluate the proposed method. Different types of reports were produced by different groups of clinicians. For example, radiology reports were produced by about five radiologists specializing in abdomen interpretation. Ultrasound reports were produced by about 15 trained gastroenterology/hepatology specialists. The discharge summaries and admission notes were produced by supervised attending physicians with after-review of trained residents. Operation notes were produced by about five liver surgeons.

The time period for personal reports in the development set ranged from 0.13 years to 10 years. The range was 9.87 years, and the mean ± standard deviation age was 5.63 ± 3.41 years. The average

number of the reports of each patient was 36. The time period for personal reports in the testing set ranged from 0.15 years to 10 years. The range was 9.85 years, and the mean ± standard deviation interval was 5.55 ± 3.58 years. The average number of reports for each patient was 30.

In total, 759 reports were involved in the process of evaluating concept identifications. All patients' personal reports ($n = 2,351$ reports) were involved in the evaluation of the classifications of personal summarized status. The percentage of each report type in different report types are shown as radiology (53%), ultrasound (24%), discharge (14%), pathology (5%), admission (2%), and operation (2%).

In this environment, the computerized patient records include structured data (such as the laboratory test results) and nonstructured data (such as the narrative reports). This study is simply for tracking data from narrative reports but not for structured data sources such as laboratory test results, which could be analyzed statistically.

CONCEPT MODEL

Concepts. Following the approach proposed by Friedman et al.,[23] the liver cancer concept model was identified for assisting in the IE. In the work, the important clinical factors were collected from the existing research, including cancer diagnosis, cancer staging, tumor information, comorbidity diagnosis, treatment, and recurrent status.[24–26] These clinical factors were regarded as the major concepts in the study.

To identify the related information of the major concept, each major concept had a set of related concepts for specifying its relevant information. Temporal information and report types were the related concepts for tracking the disease progression and confirming the examination sources. *Table 1* shows these major concepts and a listing of their related concepts.

Expressions. For identifying the concept from clinical texts, each concept has its corresponding set of regular expressions for matching the different expressions of concepts. The developers previously reviewed the reports in the development set and manually identified the expressions of concepts, including synonyms, typical and atypical abbreviations, common misspellings, etc. *Table 2* shows examples of expressions for the concepts.

IE module (first phase): identification of major and related concepts. The idea of the hot-spotting technique was used for identifying the locations of major interest concepts (e.g., the expressions of tumor: tumor, lesion, nodule, mass, etc.) in the clinical text. The locations were regarded as the bases for searching the related

Table 1. The Major Concepts and Their Related Concepts

MAJOR CONCEPT	RELATED CONCEPTS
Diagnosis of cancer (HCC)	Diagnosis of cancer, diagnosis status (i.e., confirmed, suspected, and no evidence of finding), temporal information, report type
Staging (BCLC)	BCLC staging (e.g., BCLC: class A), diagnosis status, temporal information, report type
Tumor	Tumor major object (e.g., tumor, lesion, mass, and nodule), target location (e.g., liver and segment #7), size (e.g., 1 cm), quantifier (e.g., one, two, three, etc.), temporal information, report type, non-target location (e.g., breast), non-tumor size items (e.g., LeVeen needle, which is used in RFA treatment)
Comorbidity diagnosis (liver cirrhosis)	Comorbidity diagnosis (i.e., liver cirrhosis), diagnosis status, Child–Pugh staging (e.g., Child–Pugh score: class A), temporal information, report type
Treatment	Treatment type (e.g., RFA and TACE), treatment status (e.g., performed and status afterward to confirm the treatment was performed), temporal information, report type
Recurrent status	Recurrent status, diagnosis status, temporal information, report type

BCLC, Barcelona Clinic Liver Cancer; HCC, hepatocellular carcinoma; RFA, radiofrequncy ablation; TACE, transcatheter arterial chemoembolization.

information (e.g., the location and size) from the surrounding text. After the major concepts were located in the text, their related concepts were captured within the predefined window. The regular expressions defined in the concept model were used to capture the texts and values relevant to the concepts and their corresponding concepts.

Binding of matched concepts. For mapping the relationships among multiple extracted concepts in the same document or even the same sentence, the rule-based scheme was built for binding the relationships among these concepts. A set of rules was defined for reserving and filtering out the bound concepts based on the *character* properties of these concepts. The value of *positive* was used for indicating positive correlated information being reserved, and the value of *negative* was used for indicating the extraneous information being filtered out. For temporal IE, this method focused on extracting temporal information explicitly stated in the reports. Different patterns of temporal information were encoded as regular expressions. The temporal information (year, month, and date) might appear together (e.g., "2011/10/27") or separately (e.g., "In 2011,…,10/27"). If only the partial temporal concepts are identified (e.g., "month/date"), the method will search surrounding text for finding other temporal information (e.g., "year") and combine these concepts as one complete temporal concept (e.g., "year/month/date").

Table 2. The Textual Expression Examples of Different Concepts

CATEGORY	EXPRESSION EXAMPLES
Diagnosis of cancer	HCC, hepatocellular carcinoma, hepatoma
Diagnosis status	Suspected, suspicious, impression, probable, definite, no definite, no evidence of, without
BCLC staging	BCLC stage A1, BCLC A1, BCLC clinical stage A2, Barcelona Clinic Liver Cancer stage A4
Tumor major object	Tumor, lesion, mass, nodule, metastasis, HCC, hepatoma
Target location	Liver, hepatic, right superior liver, medial segment of left hepatic lobe, segment 7, S#8
Size	0.6 cm, 1.8 × 1.4 cm, 0.3 × 0.5 × 0.7 cm, 3.7–4 cm, 4.0 * 3.8 cm, 24.0 mm
Quantifier	A, one, single, two, three, four, five, several, multiple
Liver cirrhosis	Liver cirrhosis, cirrhosis of the liver, liver cirrhosis(+)
Child–Pugh staging	Child's B, Child A, Child's class A, Child classification A, Child–Pugh A, Child–Pugh classification A
Treatment type	RFA, radiofrequency ablation, TACE, PEI, PMCT, hepatectomy, liver transplantation
Treatment status	Status post, s/p, perform, receive, evaluate, suggest, arrange, prefer, discuss
Recurrent (HCC)	Tumor recurrence, recurrent, no recurrent HCCs, no definite evidence of tumor recurrence
Temporal information	2007/11/11, 2007-08-04, 2010_03_10, 2010.03.04, 099/10/27, June 10, 2009
Report type	Computed tomography (CT), magnetic resonance imaging (MRI), MR, angiography, echo, sonar, ultrasonography, ultrasound

MR, magnetic resonance; PEI, percutaneous ethanol injection; PMCT, percutaneous microwave coagulation therapy; TACE, transcatheter arterial chemoembolization.

Normalization. The process of normalization was used for the following tasks: (1) normalizing synonyms of the same concepts, (2) normalizing various extracted textual expressions that represented the same concept using full name, typical abbreviations, and atypical abbreviations, (3) standardizing the units for numeric clinical measurement, (4) translating mixture languages (English and Chinese) into English, and (5) standardizing two different types of temporal information,

such as Anno Domini (A.D.) and the "year" used in Taiwan starting from 1911 A.D. (e.g., "100/10/27" being the "year" used in Taiwan standardized as "2011/10/27").

Sorting and grouping. These extracted concepts from the clinical text were sorted according to the temporal information. The group method congregated the clinical findings originally from the same report but appearing and being mentioned in different reports. For example, the status of liver cirrhosis was originally mentioned in the ultrasound report on 2009/5/1, and this finding was also mentioned in the latter admission note and discharge summary. These clinical findings originally from the ultrasound report on 2009/5/1 (relevant to the diagnosis of liver cirrhosis) were grouped together.

Rule-based classifier for summarizing the patient's status (second phase). The following questions could be answered by the classifier:

1. HCC patient. Did the patient have HCC?
2. First treatment. Was the specific treatment the patient's first treatment for HCC?
3. Recurrent HCC: Was the patient recurrent after the specific treatment?

Examples of evidence for these questions are shown in *Table 3*. The classifier gave the answer for each question according to the direct or indirect evidence found in the sorted and grouped extracted results. Furthermore, the report type-oriented properties were used during the processing of classification when concepts were extracted from more than one type of reports. For example, for checking the confirmed diagnosis, the method would check the extracted concepts in the order of pathology, radiology, ultrasound, and other types of reports.

EVALUATION

Report annotation. Gold-standard annotations for reports and classifications for patients' personal summarized statuses were determined by annotators. Normally, each report was annotated by two annotators. For disagreements between two annotators, a third annotator adjudicated the annotations and classifications.

Inter-annotator agreement. The inter-annotator agreements (IAAs) for IE and classification were calculated between the annotated results of two annotators. In IE, the *F*-score was used for measuring IAA.[27] In the tasks of classification, Cohen's kappa[28–30] was used for measuring IAA because positive and negative cases of patients' classifications could be specified:

$$k = (P_a - P_e)/(1 - Pe) \qquad (1)$$

where P_a was the relatively observed agreement among the annotators and P_e was the expected agreement due to chance.

$$\text{Total cases} = (TP + FP + FN + TN) \qquad (2)$$

$$P_a = (TP + TN)/\text{Total cases} \qquad (3)$$

$$P_e = ([TP + FP]/\text{Total cases})([TP + FN]/\text{Total cases}) + ([FN + TN]/\text{Total cases}) \times ([FP + TN]/\text{Total cases}) \qquad (4)$$

where TP was true-positives, FP false-positives, TN true-negatives, and FN false-negatives.

Evaluation metrics for IE module. For the evaluation of the proposed IE module, the precision, recall, and *F*-score were used. These metrics were frequently used for evaluating the methodology of IE.[31–33]

Table 3. The Evidence Examples of a Patient's Summarization Status	
STATUSES OF SUMMARIZATION	**EXAMPLE EVIDENCE**
Positive HCC patient	Direct evidence: In the pathology report, the evidence sentence was found: "it shows a hepatocellular carcinoma in microtrabecular pattern."
Negative HCC patient	Indirect evidence: There was no HCC diagnosis in all reports.
Positive first treatment	Indirect evidence: There was no other treatments before the specific RFA treatment, and there was no recurrent HCC statements before the specific RFA treatment.
Negative first treatment	Direct evidence: In the operation note, the evidence sentence was found (other treatment was performed before current RFA treatment): "Atypical hepatectomy was performed." Indirect evidence: The description of the recurrent HCC was found before current RFA treatment: "Recurrent HCC at S8."
Positive recurrent HCC	Direct evidence: In the radiology (CT) report, the evidence sentence was found: "Recurrent HCC at S8 status post RFA"
Negative recurrent HCC	Indirect evidence: There was no recurrent HCC information in all reports (only the "non-recurrence" statement: "no evidence of local recurrence is noted").
CT, computed tomography; HCC, hepatocellular carcinoma; RFA, radiofrequency ablation.	

The definitions are listed as follows:

$$Precision = TP/(TP + FP) \quad (5)$$

$$Recall = TP/(TP + FN) \quad (6)$$

$$F-score = (2 \times precision \times recall)/(precision + recall) \quad (7)$$

Evaluation metrics for classification. The following metrics were used for evaluating the rule-based classification results:

$$Sensitivity\ (=Recall) = TP/(TP + FN) \quad (8)$$

$$Specificity = TN/(TN + FP) \quad (9)$$

$$Positive\ predictive\ value\ \ (PPV)\ (=Precision) = TP/(TP + FP) \quad (10)$$

$$Negative\ predictive\ value\ \ (NPV) = TN/(TN + FN) \quad (11)$$

$$Accuracy = (TP + TN)/(TP + FP + FN + TN) \quad (12)$$

Results

Tables 4 and *5* show the IAAs of the concept identification from the total 759 reports and the classification of 78 personal statuses. The human annotators achieved the following IAAs: from 93.26% to 99.04% (*F*-scores) for six major concepts and temporal and report type concepts and from 71.83% to 79.37% (kappa valuess) for classification for 78 patients' personal statuses.

The performance of the IE module achieved an F_1-score from 92.40% to 99.59%, precision from 91.86% to 100%, and recall from 92.94% to 99.19% (*Table 6*). The performance of the rule-based classifier achieved an accuracy from 96.15% to 100%, PPV from 94.12% to 100%, NPV from 82.35% to 100%, sensitivity from 95.31% to 100%, and specificity from 95.56% to 100% (*Table 7*).

Discussion

For resolving the problem of information overload for clinicians to review large amounts of patient records, previous investigations put efforts into the creation of concept and problem-oriented viewing of patient records.[34–36]

This proposed method reduced the problem of information overload[34] to three aspects: the first aspect was to retrieve only reports that included concepts relevant to liver cancer; the second aspect was to further group-duplicated clinical findings from the same original report but mentioned in different reports; and the third aspect was to provide answers and evidence sentences of clinical questions using a rule-based classifier. For checking a patient's personal summarized status, the clinicians could only check the brief answers

Table 4. Inter-annotator Agreement for Information Extraction of Concept Entities

CONCEPT IDENTIFICATION	F_1-SCORE (%)
Diagnosis (HCC)	98.65
Staging (BCLC)	97.83
Tumor	97.14
Comorbidity diagnosis (liver cirrhosis)	98.89
Treatment	98.97
Recurrent status	99.04
Temporal information and report type	93.26

BCLC, Barcelona Clinic Liver Cancer; HCC, hepatocellular carcinoma.

and evidence sentences instead of reviewing all extracted and grouped results.

The comparison of time spent between human reviewer and automated method (IE module and rule-based classifier) is shown by the following. The human reviewers took an average of 34 min per patient, and the computer analyzed the reports of on average 6 patients per minute. The automated method is a cost savings compared with reviewing all the considerable quantities of patients' narrative reports and extracting information manually.

For each classification question, each patient only has one classification result. In this study, we only have classification results from a total of 78 patients for each classification question. In this case, perhaps a few errors could largely reduce the kappa scores. That might be the reason that the results of human IAAs are in the range of 71–79% in kappa scores. For example, in the classification of HCC patient, only one disagreement classification occurred in the total of 78 classifications, and the kappa score was 79.37%.

In order to provide more flexibility for processing the grammatical/ungrammatical sentences and narrative/tabular textual formats based on the requirements, the idea of the hot-spotting technique was used for matching the interest concepts, and these relationships were bound among matched concepts. If merely large amounts of gram-

Table 5. Inter-annotator Agreement for Classification

CLASSIFICATION	KAPPA (%)	F_1-SCORE (%)	DISAGREEMENT CLASSIFICATION NUMBER	TOTAL CLASSIFICATION NUMBER
HCC patient	79.37	99.33	1	78
First treatment	71.83	94.40	7	78
Recurrent HCC	74.21	85.29	10	78

HCC, hepatocellular carcinoma.

Table 6. Results for Evaluating the Information Extraction Module

CATEGORY	REPORT	GOLD STANDARD	RECALL (%)	PRECISION (%)	F_1-SCORE (%)
Diagnosis (HCC)	149	331	97.28	97.58	97.43
Staging (BCLC)	156	162	97.53	99.37	98.44
Tumor	156	147	95.92	95.92	95.92
Comorbidity diagnosis (liver cirrhosis)	156	171	98.25	98.82	98.53
Treatment	78	246	99.19	100.0	99.59
Recurrent status	146	157	96.82	97.44	97.12
Temporal information, report type	197	170	92.94	91.86	92.40

BCLC, Barcelona Clinic Liver Cancer; HCC, hepatocellular carcinoma.

Table 7. Results for Evaluating the Rule–Based Classification

CATEGORY	PPV (%)	NPV (%)	SENSITIVITY (%)	SPECIFICITY (%)	ACCURACY (%)
HCC patient	100.0	100.0	100.0	100.0	100.0
First treatment	100.0	82.35	95.31	100.0	96.15
Recurrent HCC	94.12	97.73	96.97	95.56	96.15

HCC, hepatocellular carcinoma; NPV, negative predictive value; PPV, positive predictive value.

ing, among others, partial mixture of languages, synonyms, typical/atypical abbreviations, two mixture types of temporal information, and grammatical/ungrammatical sentences. The application provided the functionality for normalizing these different expressions and further grouping extracted findings from the same sources for reducing the problem of information overload. For future applications, the designed concept might be the tracking items for other cancer patients. As a consequence, it might be applied to the key extraction for other types of cancer patients. For clinical practice, the system may assist clinicians in understanding a patient's status from large amounts of reports in a more effective way. For clinical research, the system may assist researchers in identifying patients who meet the clinical research eligibility criteria from large amounts of patient sets. For the development of clinical applications, the experience of developing this system may be applied to the design of systems relevant to searching desired data from electronic medical records.

Disclosure Statement

No competing financial interests exist.

mar rules and syntactic rules were used for parsing all sentences, the ungrammatical sentences and tabular textual formats might not be handled well. Therefore, the method in this study used a more flexible way to identify major concepts in sentences and then parsed the surrounding text based on major concepts for identifying their related concepts. Although the flexibility of this method may produce errors in processing some cases, this method has reported good extraction scores and classification scores in this study.

For the reports from a homogeneous group of patients, the concepts relevant to liver cancer were collected by clinicians and could be identified with these methods. For other patient groups, specific concepts might be collected from other specialized clinicians. The architecture of this system could be reused for the reports from other patient groups by replacing the components of this system such as regular expressions and rule classification. Therefore, not only the liver disease patients could be approached by these methods, but also patients with other diseases.

Conclusions

The application was successfully applied to the mixture types of narrative clinical reports with the following characteristics, includ-

REFERENCES

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* **2005;**55:74–108.

2. Jemal A, Siegel R, Xu JQ, Ward E. Cancer statistics, 2010. *CA Cancer J Clin* **2010;**60:277–300.

3. Capocaccia R, Sant M, Berrino F, Simonetti A, Santi V, Trevisani F. Hepatocellular carcinoma: Trends of incidence and survival in Europe and the United States at the end of the 20th century. *Am J Gastroenterol* **2007;**102:1661–1670; quiz 1660, 1671.

4. Yu MC, Yuan JM. Environmental factors and risk for hepatocellular carcinoma. *Gastroenterology* **2004;**127(5 Suppl 1):S72–S78.

5. Lee LT, Huang HY, Huang KC, Chen CY, Lee WC. Age-period-cohort analysis of hepatocellular carcinoma mortality in Taiwan, 1976–2005. *Ann Epidemiol* **2009;**19:323–328.

6. Chen CH, Chen DS. Hepatocellular carcinoma: 30 years' experience in Taiwan [in Japanese]. *J Formos Med Assoc* **1992;**91:187–202.

7. Pakhomov SV, Buntrock J, Chute CG. Prospective recruitment of patients with congestive heart failure using an ad-hoc binary classifier. *J Biomed Inform* **2005;**38:145–153.

8. Pakhomov S, Weston SA, Jacobsen SJ, Chute CG, Meverden R, Roger VL. Electronic medical records for clinical research: Application to the identification of heart failure. *Am J Manag Care* **2007;**13:281–288.

9. Turchin A, Kohane IS, Pendergrass ML. Identification of patients with diabetes from the text of physician notes in the electronic medical record. *Diabetes Care* **2005;**28:1794–1795.

10. Turchin A, Kolatkar NS, Grant RW, Makhni EC, Pendergrass ML, Einbinder JS. Using regular expressions to abstract blood pressure and treatment intensification information from the text of physician notes. *J Am Med Inform Assoc* **2006;**13:691–695.

11. Wiceintowski R, Sydes MR. Using implicit information to identify smoking status in smoke-blind medical discharge summaries. *J Am Med Inform Assoc* **2008;**15:29–31.

12. Uzuner O, Goldstein I, Luo Y, Kohane I. Identifying patient smoking status from medical discharge records. *J Am Med Inform Assoc* **2008;**15:14–24.

13. Cohen AM. Five-way smoking status classification using text hot-spot identification and error-correcting output codes. *J Am Med Inform Assoc* **2008;**15:32–35.

14. Denny JC, Choma NN, Peterson JF, Miller RA, Bastarache L, Li M, Peterson NB. Natural language processing improves identification of colorectal cancer testing in the electronic medical record. *Med Decis Making* **2012;**32:188–197.

15. Jain NL, Friedman C. Identification of findings suspicious for breast cancer based on natural language processing of mammogram reports. *Proc AMIA Annu Fall Symp* **1997;**829–833.

16. Gruschkus SK, Hoverman JR, Muehlenbein CE, Forsyth M, Chen C, Lopez W, Lawson AH, Pohl G. Evaluation of the reliability of electronic medical record data in identifying comorbid conditions among patients with advanced non-small cell lung cancer (Nsclc). *J Cancer Epidemiol* **2011;**2011:983271.

17. Wilson RA, Chapman WW, Defries SJ, Becich MJ, Chapman BE. Automated ancillary cancer history classification for mesothelioma patients from free-text clinical reports. *J Pathol Inform* **2010;**1:24.

18. Meystre S, Haug PJ. Automation of a problem list using natural language processing. *BMC Med Inform Decis Mak* **2005;**5:30.

19. Reichley RM, Henderson KE, Currie AM, Dunagan WC, Bailey TC. Natural language processing to identify venous thromboembolic events. *AMIA Annu Symp Proc* **2007;**1089.

20. Melton CB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. *J Am Med Inform Assoc* **2005;**12:448–457.

21. Wang X, Chase H, Markatou M, Hripcsak G, Friedman C. Selecting information in electronic health records for knowledge acquisition. *J Biomed Inform* **2010;**43:595–601.

22. Chiang JH, Lin JW, Yang CW. Automated evaluation of electronic discharge notes to assess quality of care for cardiovascular diseases using Medical Language Extraction and Encoding System (MedLEE). *J Am Med Inform Assoc* **2010;**17:245–252.

23. Friedman C, Huff SM, Hersh WR, Pattison-Gordon E, Cimino JJ. The Canon Group's effort: Working toward a merged model. *J Am Med Inform Assoc* **1995;**2:4–18.

24. Benson AB, Abrams TA, Josef EB, Bloomston PM, Botha JF, Clary BM, Covey AM. *NCCN Clinical Practice Guidelines in Oncology Hepatobiliary Cancers.* Fort Washington, PA: National Comprehensive Cancer Network, **2011.**

25. El-Serag HB. Hepatocellular carcinoma. *N Engl J Med* **2011;**365:1118–1127.

26. Izumi N. Diagnostic and treatment algorithm of the Japanese Society of Hepatology: A consensus-based practice guideline. *Oncology* **2010;** 78(Suppl 1):78–86.

27. Hripcsak G, Rothschild AS. Agreement, the F-measure, and reliability in information retrieval. *J Am Med Inform Assoc* **2005;**12:296–298.

28. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* **1960;**20:37–46.

29. Carletta J. Assessing agreement on classification tasks: The kappa statistic. *Comput Linguist* **1996;**22:249–254.

30. Hripcsak G, Heitjan DF. Measuring agreement in medical informatics reliability studies. *J Biomed Inform* **2002;**35:99–110.

31. Nassif H, Woodsz R, Burnsidey E, Ayvaci M, Shavlik J, Page D. Information extraction for clinical data mining: A mammography case study. *Proc IEEE Int Conf Data Min* **2009;**37–42.

32. Mykowiecka A, Marciniak M, Kupsc A. Rule-based information extraction from patients' clinical data. *J Biomed Inform* **2009;**42:923–936.

33. Coden A, Savova G, Sominsky I, Tanenblatt M, Masanz J, Schuler K, Cooper J, Guan W, de Groen PC. Automatically extracting cancer disease characteristics from pathology reports into a disease knowledge representation model. *J Biomed Inform* **2009;**42:937–949.

34. Zeng Q, Cimino JJ, Zou KH. Providing concept-oriented views for clinical data using a knowledge-based system: An evaluation. *J Am Med Inform Assoc* **2002;**9:294–305.

35. Zeng Q, Cimino JJ. A knowledge-based, concept-oriented view generation system for clinical data. *J Biomed Inform* **2001;**34:112–128.

36. Bashyam V, Hsu W, Watt E, Bui AAT, Kangarloo H, Taira RK. Informatics in radiology problem-centric organization and visualization of patient imaging and clinical data. *Radiographics* **2009;**29:331–344.

Address correspondence to:
Ja-Der Liang, MD
Department of Internal Medicine
National Taiwan University Hospital and
National Taiwan University College of Medicine
No. 7, Chung-Shan South Road
Taipei, 100
Taiwan

E-mail: jdliang@ntuh.gov.tw