

Potential Usefulness of Diagnostic Reminder as Web-based Clinical Decision Support System

Keijirou Torigoe¹ and Yasuharu Tokuda²

1. Torigoe Clinic, Ibara, Okayama 715-0025, Japan

2. JCHO Hospitals, Tokyo 108-8583, Japan

Abstract: Diagnostic error is prevalent and there is a need for reducing it for improving patient safety. Electronic resources may be candidates as diagnostic decision support systems to assist physicians in clinics or hospitals. A unique system has been developed by consisting of a disease knowledge database coupled with algorithms designed specifically for clinical reminders during real-time diagnostic processes. This system is currently being used as a diagnostic decision-support tool in a clinic base and its usefulness has been empirically evaluated by applying it to the case reports in the *New England Journal of Medicine*. Further studies are needed to prove its usefulness for reducing diagnostic errors in real clinical practice.

Key words: Computer-aided, decision support, diagnosis reminder, diagnostic error, patient safety.

1. Background

The basic job of primary care physicians is to diagnose and treat common diseases. However, there are sometimes cases of rare diseases that are difficult to diagnose. To timely diagnose these rare diseases without delay, it would be desirable for primary care physicians to remember all possible diseases including rare diseases. However, there seems a limit of human memory in a physician for memorizing diseases.

To assist physicians aiming for an early and timely diagnosis of rare diseases, a computer-aided diagnostic decision support system, Diagnosis Reminder (DR), has been developed. In this paper, brief explanation is provided about the characteristics of DR and the benefits of using DR as a tool to facilitate and expedite diagnosis.

2. Overview of the Diagnostic Decision-Support System

2.1 Disease Knowledge Database (DKDB)

Each disease presents distinct clinical manifestations

Corresponding author: Yasuharu Tokuda, M.D., M.P.H., research field: clinical reasoning.

(CMs) including symptoms, signs, and laboratory results. CMs about approximately 2,000 diseases were collected and these were classified into 2 or 3 grades, according to their importance in achieving an accurate diagnosis. Weight points (WP) were allocated to these grades using models on the basis of the CM frequency (Table 1 and 2). Thus data of each disease have been expressed as distinct Disease Units (Table 3). Consequently the unique database was constructed as Disease Knowledge Database (DKDB). DR has two types of DKDBs: main DKDB and sub DKDB. The main DKDB comprises all Disease Units. The sub DKDB was designed to supplement the main DKDB. CMs in the sub-DKDBs contain supplemental data which could not be stored in the main DKDB.

To establish reminder/diagnostic efficiency, a scoring system is used for CMs in the main DKDB. There are two rules used for point assignment: (1) Higher point values were awarded to CMs that were considered as most important for diagnosis and (2) The overall score of the CMs of one Disease Unit in the main DKDB and/or the average score of each CM of one Disease Unit in the main DKDB were required to fall within a certain range (Table 4, Note 1). The

Table 1 An example of contents in the actual main Disease Knowledge Database (DKDB).

A case for which the "index" (noted below) is "XYZ" (X, Y, and Z are not 0).

R28 Wallenberg syndrome
 HemiplvertigdysphdysarthhiccupssensedysmetadiadHorhoarsgaitnystagECGataxptosmiosspeech
 naussyncopvomitCSFRomhypersalTIAreflex^anisocoughBabMLFdizzfaciaspastrigidtetra 421

1. Explanation of the upper main DKDB contents is as follows:

-R28: disease number

-Wallenberg syndrome: disease name

-ψ hemipl/vertig/ dysph/dysarth/ hiccup/sense/ dysmet/adiad/Hor/

Clinical manifestations (CMs) are divided into 4 blocks. One block is fulfilled by maximum 19 bytes because the memory area was limited approximately 30 years ago when this system initiated the accumulation of disease data. There are 9 CM-codes.

-† hoarse/gait/nystag/ECG/ atax/ptos/mios/speech/

CMs are divided into 2 blocks. There are 8 CM-codes.

-‡ naus/syncop/vomit/CSF/ Rom/hypersal/TIA/ reflex/anisoc/cough/

Bab/MLF/dizz/faccia spast/rigid/tetra/

CMs are divided into 5 blocks. There are 17 CM-codes.

/: division marker of each CM-code

ψ: very important CMs (1 CM): appearance rate is > 0.5

†: moderately important CMs (m CM): appearance rate ranges from 0.2 to 0.49

‡: slightly important CMs (l CM): appearance rate ranges from 0.03 to 0.19

(CMs for which the appearance rate is ≤ 0.03 accumulated in the sub-DKDB files.)

By a division marker (" /"), we know that the CM numbers are 9, 8, and, 17 for iCM, mCM, and l CM, respectively.

-421: "index"

4 indicates the number of blocks of very important CMs.

2 indicates the number of blocks of moderately important CMs.

1 is the number to regulate all CMs total weight points (TWPs).

2. How to determine WP for each CM and how to calculate TWP for each disease.

Now, under above main DKDB definition, the scores of 6, 3, and 1 are given to iCM, mCM, l CM relatively beforehand. And then calculations are performed for each CM's WP. After all, one CM's WP of the iCMs is decided by division; 6/block number. On above example 1.5 point (= 6/4) is given to all CMs of iCM. Similarly calculations are performed for one CM's WP of the mCM. After all it is 1.5 point (= 3/2). The more CMs are, the lower CM's WPs are. The lesser CMs are, the higher CM's WPs are. The "1" that the last number of the "421" is used to calculate one CM's WP of the l CM. On above example 1.0 point (= 1/1) is given to all CMs of l CM. If the last number of the "index" is 2, one CM's WP of the l CM is 0.5 point (= 1/2). If the "index" is "423", one CM's WP of the l CM is 0.333... point (= 1/3). Calculations are performed for the TWP of each disease. On above example, TWP is 42.5 point (1.5 times 9 + 1.5 times 8 + 1.0 times 17). And calculations are performed for mean point of CM. On above case, average point of CM is 1.22 (= 42.5/(9 + 8 + 17)).

Table 2 Another example of contents in the actual main DKDB.

A case for which the "index" is "0XY" (X and Y are not 0).

R29 systemic lupus erythematosus (SLE)
 heartarthanemmurmrGweighcolddyspnskinfevererythmyalESRBtECGLDHbleeddizzches
 pancphotoedemEaRayleucoproteinuliverrenalungpainconvulsivedfatigCRPnausYY
 myastthrombocytomMPPHulceralopcough@lymphoppsycFUlymph 052

1. Explanation of the upper record file contents is as follows:

-R29: disease number

-systemic lupus erythematosus (SLE): disease name

-† heart/arth/anem/murm/rG/ weigh/cold/dyspn/skin/ fever/eryth/myal/ESR/Bt/

ECG/LDH/bleed/dizz/ches/ panc/photo/edem/Ea/Ray

CMs are divided into 5 blocks. One block is fulfilled by maximum 19 bytes, because the memory area was limited approximately 30 years ago when this system initiated the accumulation of disease data. These belong to mCM, because the “index” is “052”.

The CM number is 24.

-‡ leucop/proteinu/liver/ renal/lung/pain/convuls/ lived/fatig/CRP/naus/YY
 myast/thrombocytopenia/MP/ PH/ulcer/alopec/cough/@/ lymphop/psyc/FU/lymph

CMs are divided into 6 blocks. These belong to l CM, because the “index” is “052”. The CM number is 24.

 /: division marker of each CM-code

†: very or moderately important CMs (mCM): appearance rate is ≥ 0.2

‡: slightly important CMs (l CM): appearance rate ranges from 0.03 to 0.19

(CMs for which the appearance rate is ≤ 0.03 are accumulated in the sub-DKDB files.)

 -052: “index”

5 indicates the number of blocks of very or moderately important CMs.

2 is the number to regulate all CMs TWPs.

2. How to determine WP for each CM and how to calculate TWP for each disease.

For this case, under the main DKDB definition given above, the system assigns the scores of 3, and 1 to mCM, and l CM, respectively, beforehand. Then, calculations are performed for WP or TWP in the same manner. However, for calculating WP or TWP, it is necessary to determine a provided value (WP or TWP) 2.5 [= (6 + 3 + 1)/(3 + 1)] times.

Table 3 Schema for the disease unit.

In the main DKDB.
 [Disease name-A: <CMs (WP)>, <CMs (WP)>, <CMs (WP)>]
 [Disease name-B: <CMs (WP)>, <CMs (WP)>, <CMs (WP)>]
 [Disease name-C: <CMs (WP)>, <CMs (WP)>] ψ

In the sub-DKDB.
 [Disease name-A: <CMs>]
 [Disease name-B: <CMs>]
 [Disease name-C: <CMs>]

ψ For some Disease Units in the main DKDB, CMs are classified into only 2 grades.

second rule was set to guarantee equal representation to all Disease Units included in the main DKDB. The sub DKDB had no such constraints and was primarily provided to improve diagnostic accuracy of the DR.

These DKDBs included over 2,000 diseases and 630 CMs. The DKDB in itself do not incorporate factors such as the epidemiologic frequency and pattern of disease onset; symptomatic changes in the clinical course, gender, age, race, or region, since these factors should be considered on a case-to-case basis at the time of diagnosis.

2.2 Search Algorithm of Diagnosis Reminder

The search algorithm is presented in Table 4. When a clinical case is presented for diagnosis, all the

patient’s CMs are recorded. The DR then compares, one by one, all the patient’s CMs with all the CMs of one Disease Unit in the main DKDB. This allows the physician to cross-reference any relevant CMs.

If a patient’s CM coincides with a CM associated with the Disease Unit under examination, the DR adds the WP of this CM to create a total WP score for this Disease Unit (Table 5). This allows interested parties to calculate the total score (TS) of one Disease Unit. The primary search allows the DR to apply these processes to all Disease Units in the main DKDB. The program then lists potential diagnoses after sorting them by TS. A scheme of this process is presented in Table 4.

Next, the search process continues with a secondary search, very similar to the primary search, in order to

narrow down the results. This time, only certain CMs are selected on the basis of their diagnostic importance. The selected CMs are then used as Keywords (KWs). In the secondary search, the KWs are compared, on an individual bases, with the CMs of the Disease Units in both the main DKDB and sub DKDB, in the same order as the diseases were listed in the primary search. In this way, the DR narrows the search to diseases documented in the literature using certain KWs (Table 4, Note 2). These primary and secondary searches are expected to enhance DR efficiency and narrow down the choices considered as part of the differential

diagnosis. Future research may provide an algorithm to optimize KWs selection.

2.3 Verifying DR Accuracy

The following example illustrates the validity of DR. A case report from the *NEJM*, "A Jaundiced Eye" [1] was selected. Table 6 shows each CM input used for this search and displays the associated results (The details regarding the information presented in Table 6 can be accessed at the provided website). The ultimate diagnosis in this case was a cytomegalovirus infection. This case empirically verified the utility of DR.

Table 4 Image representation of diagnosis reminders.

For example, CMs of a patient who requires a diagnosis are as follows: a, b, c, d, e, and f. ψ .

Main DKDB†			
[Disease Unit];	<CMs (WP‡)>	<CMs (WP)>	<CMs (WP)>]
	(very important)	(moderately important)	(slightly important)
[Disease-A: <xyaz (2)>		<lmncd (1)>	<epqr (0.5)>]
[Disease-B: <pcdagk (1.5)>		<ehijmnbo (0.75)>	<xyfz (0.3)>]
[Disease-C: <zyxvuts (1.5)>		<qetpr (1)>	<oihad (0.3)>]
[Disease-D: <mn (3)>		<cghijxyp (0.75)>	<abf (0.5)>]
[Disease-E: < () >		<drkefnmsi (1.25)>	<qthalpu (0.6)>] ¶

Image representation of a primary search: total score (TS) for each disease. ||

Rank:

1. TS for Disease-B = $1.5*3$ (a, c, d) + $0.75*2$ (b, e) + $0.3*1$ (f) = 6.3
2. TS for Disease-A = $2*1$ (a) + $1*2$ (c, d) + $0.5*1$ (e) = 4.5
3. TS for Disease-E = $1.25*3$ (d, e, f) + $0.6*1$ (a) = 4.35
4. TS for Disease-D = $0.75*1$ (c) + $0.5*3$ (a, b, f) = 2.25
5. TS for Disease-C = $1*1$ (e) + $0.3*2$ (a, d) = 1.6

Note 1. Overall score (OS) for CMs of each disease. §

The average point (AP) of each CM is in parentheses. §

OS for Disease-A = $4*2 + 5*1 + 4*0.5 = 15.0$ (AP = $15/13 = 1.15$)

OS for Disease-B = $6*1.5 + 8*0.75 + 4*0.3 = 16.2$ (AP = $16.2/18 = 0.9$)

OS for Disease-C = $7*1.5 + 5*1 + 5*0.3 = 17.0$ (AP = $17/17 = 1.0$)

OS for Disease-D = $2*3 + 8*0.75 + 3*0.5 = 13.5$ (AP = $13.5/13 = 1.04$)

OS for Disease-E = $10*1.25 + 7*0.6 = 16.7$ (AP = $16.7/17 = 0.98$)

Note 2. Discussion of a secondary search.

During a secondary search, KWs are compared one by one with CMs of the Disease Units in both the main DKDB and sub-DKDB, in the same order as the diseases listed in the primary search. Thus, DR highlights those diseases that only include KWs (e, g, a, b, and f). In these image figures, sub-DKDB is not shown. However, assuming that the sub-DKDB is empty, Disease-B and Disease-D will be highlighted, because they are the only two diseases that contain the chosen KWs.

ψ In this schema chart, 1 CM is shown by 1 character although 1 CM is an encoded character string in an actual file.

† The main DKDB is the union of each Disease Unit.

‡ These points are previously decided in a unique manner.

¶ CMs of Disease-E are classified only into 2 grades.

|| TS for 1 disease is a sum total of the point obtained by multiplying the number of coinciding CMs and each corresponding WP. The coinciding CMs are displayed in parentheses. § OS is the total of that point obtained by multiplying the number of each CM and each corresponding WP. AP for 1 CM is that value for which OS is divided by the number of all CMs.

Table 5 Step by step explanation for patient’s CMs that coincide with the Diagnosis reminders CMs for Disease-A to create WP TS in this Disease Unit.

Patient’s CMs are as follows: a, b, c, d, e, and f.

2. Disease Unit.
 [Disease name: <CMs (WP)>, <CMs (WP)>, <CMs (WP)>]
 (very important) (moderately important) (slightly important)

[Disease-A : <xyaz (2)>, <lmncd (1)>, <epqr (0.5)>]

3. Coinciding disease CMs and their WPs by patient and disease unit.
 A patient’s CMs: a, c, d, e
 [Disease-A : <xyaz (2)>, <lmncd (1)>, <epqr (0.5)>]
 [Coinciding : <a (2)>, <c, d (1) >, <e (0.5)>]

4. Total score (TS) = number of coinciding CMs multiplied by their WPs.
 TS of Disease-A = 2*1 (= a) + 1*2 (= c, d) + 0.5*1 (= e) = 2 + 2 + 0.5 = 4.5.

Table 6 DR applied to 1 “Clinical problem-solving” case [1].

A. First search and its results, CMs, for this search (all information was collected during the first consultation):
 Sore throat
 Fatigue (“run down” feeling)
 Cough
 Fever
 Abdominal pain
 Conjunctivitis (red-eye, itchy eye and thick discharge)
 Dark-colored urine
 Results:
 1. Cytomegalovirus infection scored 8.95 points and ranked 61st in this list.
 2. After inputting all CMs as KWs and narrowing down the number of diseases, cytomegalovirus infection was listed as 5th among a total of 16 suspected diseases.

B. Secondary search and its results, CMs, for this search (next 11 items are added to the original search items):
 Jaundice
 Lymph node swelling
 Hepatosplenomegaly
 Nausea
 Vomiting
 Elevated serum alanine aminotransferase level
 Tachycardia
 Skin lesion
 Dyspnea
 Leukocytosis
 Pulmonary infiltration (abnormal lung shadow on chest x-ray photo)
 Results:
 1. Cytomegalovirus infection scored 9.72 points and ranked 20th in this list.
 2. Among the CMs and inputs used in the primary search, some CMs were excluded as KWs because “Dark-colored urine” has the same meaning as “Jaundice”, “Nausea” has nearly the same meaning as “Vomiting”, and “Tachycardia” is a nonspecific symptom. After inputting another 15 CMs as KWs and narrowing down the diseases, cytomegalovirus infection was listed as 4th on the ranking scale of a total of 15 diseases.

Table 7 Results of the primary search for NEJM paper cases.

1. The disease database sufficiency rate for the main DKDB was 95.3%. (Of all 470 cases, only 22 diseases were not included within the main DKDB).
 2. When we considered those cases for which the ultimate diagnosis (UD) was included within the main DKDB, the percentage of UD listed in the top 20 was 50.0%.
 3. When we considered those cases for which UD was within the main DKDB, the percentage of UD listed in the top 50 was 73.9%.
 4. When we considered those cases for which UD was within the main DKDB, the percentage of UD found anywhere in the list was 95.3%.

(1) DR efficiency increased as more CMs were incorporated in the primary search.

(2) The secondary search further improved DR efficiency—if the KWs were selected carefully.

(3) The diagnostic decision support tool has been proven useful when the DR is prepared appropriately and used accurately.

The system was tested once again using “*Case Records of the Massachusetts General Hospital*” and has produced significant results since 2013. These results are shown in Table 7. The secondary search idea was developed and has been applied since January 2004. Table 7 shows the results of the primary search, including 470 cases cited in *NEJM 2000; Vol. 343* and *NEJM 2013; Vol. 369*. (In cases in which more than complicated diseases were considered in the differential diagnosis, only the main disease was the object of analysis). The accuracy of the diagnoses based on the DR was slightly superior to that achieved using previously developed conventional approaches [2].

3. Discussion

Basic research on a computer-aided diagnostic decision support system produced interest among scientific community in the United States from 1970s to 80s [3-5]. However, the practical use of such a system remains to be developed because of technological limitations as well as logistic difficulty in complicated collection and analysis about medical data [6, 7]. The system described here may advance the field [8-13].

Since the beginning of 1980s, new methods have been sought for the analysis of patient-specific clinical evidence. Electronic medical records allow physicians to access abundant information with just a few key strokes. The system in this report is currently used at the clinical institution as an integral component of routine clinical examinations [14, 15].

If the comprehensive development of the DKDB is achieved and KWs are carefully selected, it can

provide a wealth of information related to diagnostic possibilities that would be useful to physicians. This system is available at the public domain at the URL shown below [16-18]. Future studies are needed to validate its usefulness in using such a diagnostic reminder as web-based clinical decision support system.

References

- [1] John, A. K., Henry, R., Chadd, S., Findlay, W., and Sanjay, S. 2006. “A Jaundiced Eye.” *NEJM* 354: 1516-20.
- [2] Eta, B. S., George, W. D., Alwyn, S. A., James, J. R., James, A., Alfred, B. L., Eugene, B. V., Glenn, C., Vincent, D. W., Eugene, F. P., Leonard, H. D., Elliot, M. L., Charles, R. E., and David T. O. 1994. “Performance of Four Computer-Based Diagnostic System.” *NEJM* 330: 1792-6.
- [3] Weed, L. L. 1964. “Medical Records, Patient Care and Medical Education.” *Ir. J. Med. Sc.*: 271-82.
- [4] Walton, L. P., Holland, R. R., and Wolf, L. L. 1979. “Medical Guidance and PROMIS.” *Computer* (November): 19-27.
- [5] Miller, R. A., Pople, H. E. Jr., and Myers, J. D. 1982. “INTERNIST-1, an Experimental Computer-Based Diagnostic Consultant for General Internal Medicine.” *NEJM* 307: 468-76.
- [6] Peter, S., Ramesh, P. S., and Williams, S. B. 1988. “Artificial Intelligence in Medical Diagnosis.” *Ann. Int. Med.* 108: 80-7.
- [7] Millar, R. A. 1994. “Medical Diagnostic Decision Support System—Past, Present and Future: a Threaded Bibliography and Brief Commentary.” *J. American Med. Inf. Ass.* 1: 8-27.
- [8] Tierney, W. M., Miller, M. E., Overhage, J. M., and McDonald, C. J. 1993. “Physician Inpatient Order Writing on Microcomputer Workstations; Effects on Resource Utilization.” *JAMA* 269: 329-33.
- [9] Cimino, J. J., Socratous, S. A., and Clayton, P. D. 1995. “Internet as Clinical Information System: Application Development Using the World Wide Web.” *J. American Med. Inf. Ass.* 2: 273-84.
- [10] Evans, R. S., Pestotnik, S. L., Classen, D. C., Clemmer, T. P., Weaver, L. K., Orme, J. F. Jr., and Burke, J. P. 1998. “A Computer Assisted Management Program for Antibiotics and Other Antiinfective Agents.” *NEJM* 338: 232-8.
- [11] Suptendra, N. S. 2004. “Automated Diagnostic System.” *Ind. J. Med. Inf.* 1: 25-8.

- [12] Barnett, G. O., Cimino, J. J., Hupp, J. A., and Hoffer, E. P. 1987. "DXplain: An Evolving Diagnostic Decision-Support System." *JAMA* 258: 67-74.
- [13] Barness, L. A., Tunnessen, W. W. Jr., Worley, W. E., Simmon, T. L., and Ringe, T. B. K. Jr. 1974. "Computer-Assisted Diagnosis in Pediatrics." *Am. J. Dis. Child* 127: 852-8.
- [14] Torigoe, K., and Arai, T. 1984. "Trial to Computer-Aided Clinical Diagnosis in Internal Medicine." *Japan Medical Journal* 3131: 29-31 (In Japanese).
- [15] Torigoe, K., Katou, G., and Oota, Y. 2003. "Computer-Aided Clinical Diagnostic Decision-Support Tool." *Japan Medical Journal* 4120: 24-32 (In Japanese).
- [16] http://www.openclinical.org/dm_cacddst.html.
- [17] <http://60.32.120.74/examples/en/syoujou1.jsp>.
- [18] http://mith.akira.ne.jp/test/diagnosis_en/syoujou1.jsp.