

Software Improves Diagnostic Accuracy with Minimal Training

A. Papier MD, E. Allen, MD, M McDermott PhD, University of Rochester College of Medicine, Rochester NY, P Lawson SUNY Buffalo Medical School, Buffalo, NY

I. Summary

Searchable image databases can provide rapid access to medical pictures, but until now they have not been designed to support the diagnostic evaluation. VisualDx is a JAVA based decision support application focusing on visually diagnosable diseases. Generalist physicians and dermatologists participated in a randomized trial of the software tool in comparison to text-based resources. Accuracy of differential diagnosis was found to be significantly higher among non-dermatologist physicians using the software intervention in comparison to standard textbooks.

II. Background

It has been estimated that 10-20% of all primary care physician visits include at least one skin complaint, as either the chief or secondary complaint. Yet, there is evidence that generalist physicians frequently order the wrong tests and misdiagnose dermatologic and other visually diagnosable problems, increasing the cost of health care delivery and delaying appropriate treatment.

Though photographs of skin disease contained in textbooks and atlases have been used for well over a century in assisting physicians in dermatologic diagnosis, textbooks and atlases have space constraints and indexing limitations. Computer based information resources offer virtually unlimited storage and display of color images. Computer based systems also offer the advantages of multi-axial indexing and searching, as well as the possibility of using images in combination with a knowledge base as a decision support tool.

VisualDx is a JAVA based decision support tool. It was developed as a point of care reference, just as an atlas, or textbook with color photographs would be employed at the point of care, to reference a visual presentation and confirm a diagnosis. A key functionality of the software is facilitation of the task of image matching for the end user, by combining graphical search tools, a computerized knowledge base of relationships between findings and diagnoses, and thousands of digital images. Human skin disease, like the natural world, is defined by variation. Possible diagnoses can number in the hundreds and their related images in the thousands. To facilitate a diagnostic search, images are grouped into diagnostic "stacks". Images within "stacks" are sorted by the user morphologic query, allowing for the viewing of relevant images. Stacking images by diagnosis also condenses the "thumbnail contact sheet" information space, easing user comparison of relevant images and related diagnostic knowledge.

III. Objective

To assess physician diagnostic accuracy when using a visual decision support tool in comparison to referencing textbooks to solve 4 “unknown” cases.

IV. Methods

Case Selection

Case presentations were obtained from the medical literature by conducting a MEDLINE search for cases presenting with fever and a rash in adults.* Four cases were found to have pictures both in the reference text and intervention software, and were of sufficient difficulty to require a physician to seek reference in “solving” the case.

Recruitment

Emergency, internal medicine, family medicine and dermatologist physicians were recruited by telephone, mail and e-mail. A total of 140 physicians were contacted. Fifty physicians participated. Physician participants were granted CME credit approved by the Office of Professional Education at the University of Rochester.

Study Design

The study assistant provided software training using a four-minute script to demonstrate the basic software functionality. Physician self-trial of the software followed. Physicians were randomized to one of four “Case mix groups” (see table). Each case synopsis was presented on a single uniform placard with images. After reading a case, and before using a reference, subjects were asked to write down their preliminary diagnoses ranked by 1st choice, 2nd choice and 3rd choice. Responses were collected, and then depending upon the case mix randomization, the subject used either the textbooks or the software intervention. The participant then provided their 1st choice, 2nd choice and 3rd diagnostic choices as they did prior to using the intervention reference. Depending upon the randomization of the case mixes, subjects either began with 2 computer assisted cases, or with 2 textbook assisted cases.

Case Mix Groups

1 Books with Cases 1 & 2 followed by Software with Cases 3 & 4	2 Software with Cases 1 & 2 followed by Books with Cases 3 & 4
3 Books with Cases 3 & 4 followed by Software with Cases 1 & 2	4 Software with Cases 3 & 4 followed by Books with Cases 1 & 2

Software Application

VisualDx, “Fever and a Rash” JAVA 2 applet/application, Logical Images Inc., Rochester NY, 2001

Texts Used

Andrews' Diseases of the Skin: Clinical Dermatology, 9th ed
Richard B. Odom, William D. James, and Timothy G. Berger, 1135 pp, with 1271 illus,
Philadelphia, Pa, WB Saunders Co, 2000

Fitzpatrick's Dermatology in General Medicine, vols 1-2 Irwin M. Freedberg, Arthur Z. Eisen,
Klaus Wolff, et al, 5th ed, 3129 pp, with 2000 illus,,New York, NY, McGraw-Hill, 1999

Mandell Douglas and Bennett's Principles and Practice of Infectious Diseases, 4th edition,
Churchill Livingstone, approx. 1000 illustrations, 1996

Computer Used

Dell laptop with a 700-MHz Pentium III processor and 15" LCD TFT display

V. Results

Average training time was 4 minutes.

Total Number of correct diagnoses by Intervention
(1st, 2nd and 3rd choices combined)

	Books	Software
Family Medicine	9	19
Internal Medicine	7	16
Emergency Medicine	9	21
All Above Non-Dermatology	25	56
Dermatology	12	21

Diagnostic responses were scored as follows:

- Correct diagnosis as 1st choice = 3 points.
- Correct diagnosis as 2nd choice = 2 points.
- Correct diagnosis as 3rd choice = 1 point.
- Incorrect diagnosis = 0 points.

Average Score by Intervention

	Books	Software
Family Medicine	1.0	1.89
Internal Medicine	.83	1.91
Emergency Medicine	1.125	2.33
All Above Non-Dermatology	.99	2.04
Dermatology	1.33	2.16

Average Score by Intervention
(Improvement Over Baseline)

	Books	Software
Family Medicine	-.03	1.18
Internal Medicine	-.08	.75
Emergency Medicine	.375	1.12
All Above Non-Dermatology	.07	1.02
Dermatology	.25	1.08

Significance of Software Intervention Effect by Group
(Improvement over Baseline)

	Chi-Square	Pr> Chi-Square
Family Medicine	9.1572	<.0025
Internal Medicine	7.8204	<.0052
Emergency Medicine	3.2267	<.0724
All Above Non-Dermatology	20.1888	<.0001
Dermatology	2.8772	<.0898

VI. Discussion

Is there a particular software functionality or design element responsible for the increase in accuracy? Possibilities include the differential diagnosis readily available as a list within the software, the large number and high quality images, and the ability to cross-index by searching on multiple findings. Keystrokes were not recorded, but subjective data suggested the physician subjects favored the morphology and distribution combinatorial approach as well as the quality of the pictures.

If a problem-oriented text existed for the same problem domain could it perform as well as problem-oriented software? Theoretically, a text-based reference would be limited in the number of images, and would have little or no searchability beyond a standard index. No such text exists today to conduct a comparison study.

The data of this pilot study significantly suggests advantages to using a computer-based system in the referencing of complex visual and textual diagnostic knowledge. The comparison to text-based resources modeled clinical practice today, as physicians frequently reach for atlases or texts containing pictures to assist visual diagnostic evaluations. Physicians also preferred the computer-based resource to the texts in the subjective survey conducted during the study. The graphical searching, and the variety, depth and quality of images were all seen as strengths. The search and display capabilities were also preferred by many of the participants.

Future work will include comparing the software with the images removed (the user will view text-based differentials), to a system that contained images associated with diagnoses in lists but without a database. In addition the system will be deployed within an emergency department and

users will be randomized to the software-versus-books again as they evaluate actual patients with fever and a rash.

VII. Conclusion

- **Correct diagnoses more than doubled among the non-dermatologists and was statistically significant.**
- **Physicians subjectively preferred the software system as contrasted to textbooks.**
- **With minimal training, physicians learned to effectively use the tool.**
- **Further study needs to investigate and define the functionality and design elements behind the improvement in accuracy.**

Notes:

This research has been supported by the Dermatology Foundation.

Study conducted by The Center for Future Health at the University of Rochester, Rochester, NY

Thanks To Ken Arndt MD, and the Archives of Dermatology for copyright permission.

ⁱ Clark, RA and Rietschel, RL. The cost of initiating appropriate therapy for skin diseases: a comparison of dermatologists and family physicians. *Journal of the American Academy of Dermatology* 9(5): 787-796, 1983.