Prompting Physicians For Cost-Effective Test Ordering in The Low Prevalence Conditions of Family Medicine

Robert M. Bernstein, Ph.D., M.D., C.M., C.C.F.P.
Gary R. Hollingworth, M.D., F.C.F.P.
William E. Wood, M.D., C.M., C.C.F.P.

Medical Informatics Research Group, Clinical Epidemiology Unit,
Department of Family Medicine, University of Ottawa

We have developed a computerized prompting system for test ordering which we feel will decrease the cost of investigations and at the same time promote an evidence based learning approach to test ordering. Prompting systems have been shown to be cost-effective but suffer from many disadvantages in the family practice setting. They tend to be difficult to modify by the user and contingent on an inflexible rule based structure. Many suggestions are ignored implying that they are not relevant. In family practice most conditions are of low prevalence. Prompting for test ordering where the pre-test likelihood of disease is small will result in a large number of false positives and many unnecessary repeat or confirmatory investigations and attendant anxiety unless the prompting system is specifically designed to be used in a low prevalence environment. PROMPTOR-FM (PRObabilistic Method of Prompting for Test ORdering in Family Medicine) was developed to overcome these perceived difficulties. It allows the physician to rapidly calculate the positive and negative predictive values of a test being considered based on the clinical index of suspicion. The physician is able to repeat the calculations and compare the results with previous calculations. By using PROMPTOR-FM repetitively, the clinician can learn to balance the risk of "missing" a rare but serious condition against the risk of falsely identifying disease with its downstream hazards and costs of further investigation. Prompting for test ordering is therefore uniquely tailored to each patient's situation.

I. INTRODUCTION

Governments are trying to control health care costs. The College of Family Physicians of Canada, and other professional organizations are developing practice guidelines and quality assurance programs. The success of these initiatives will depend on developing ways the practicing physician can initiate change in the course of his/her usual practice. An electronic record with integrated guidance is one such method.

We have developed a computerized prompting system for test ordering (PROMPTOR-FM - PRObabilistic Method of Prompting for Test ORdering in Family Medicine) which we feel will decrease the cost of physician test ordering and at the same time promote an evidence based learning approach to test ordering.

Prompting systems have been used and studied since the late 70's. They have been shown to be cost effective [1-4], however, they suffer from major disadvantages in the family practice setting. They are either poorly integrated with the process of usual care or they are rule based, inflexible, difficult to modify or update without programming, and result in many suggestions being ignored.[5] They do not distinguish situations of low prevalence of disease (as in much of family practice) from situations where the prior probability of disease is high, thereby potentially leading to inappropriately high test ordering. In addition, no learning seems to result from their use. (physicians revert back to old habits once the system is no longer available for prompting. [6]

Family practice is a low prevalence domain. The most common reason for visit in our clinic is hypertension which accounts for less than 5% of the visits. Many studies of the content of family or general practice confirm this frequency distribution of diagnoses [7]. Only 18 conditions have a visit frequency of over 1%. All the rest of primary care medicine has a visit frequency of less than one percent. Prompting for test ordering in these situations where the prior probability is low will result in a large number of false positives and many unnecessary repeat or confirmatory investigations

0195-4210/94/$5.00 © 1994 AMIA, Inc.

unless care is taken to tailor the prompting system to the environment in which it is used.

Family and general practice residents are trained throughout most of medical school and much of residency by specialists in hospitals. Specialty medicine is high prevalence, and in hospital specialty wards the prevalence of the diseases in the domain of that specialty is high. Therefore trainees for general practice are taught and develop their attitudes and practices in conditions which do not mimic those which they will encounter after graduation in office based ambulatory primary care. Since the probability that a patient who presents to a cardiologist with chest pain has angina is high, and the probability of chest pain presenting in general practice being angina is low, therefore the approach to diagnostic testing must follow different paths.

II. OBJECTIVES OF PROMPTOR-FM

PROMPTOR-FM was developed to overcome the perceived difficulties with previous systems. The OBJECTIVES are: (a) to allow the physician to maintain his/her autonomy when ordering tests; (b) to allow the physician to become an active participant in the decision process when ordering tests, thus promoting active learning; (c) to allow the physician to understand the process of ordering tests in a low prevalence setting; (d) to promote an approach to test ordering based on the principles of evidence based learning;[8,9] and (e) to be a learning tool for trainees and a practical application of an evidence based curriculum.

There are 3 main reasons why tests should be ordered:

1. Case finding (suspicion of disease),
2. Screening for disease, and
3. Following or monitoring a known clinical condition.

PROMPTOR-FM is intended to be used for the first two. Its utility in case finding is clear, as it allows the physician to rapidly calculate the positive and negative predictive values of the test he/she is considering ordering based on his/her clinical index of suspicion. The physician is able to repeat the calculations and compare the results with previous calculations. The utility in screening is less obvious, since usually screening procedures appear to have been established by a consensus authority and are not usually questioned. However, many suggestions for screening in the Canadian Task Force [10-12] are discretionary and depend on the circumstances and risks of the individual patient. In these situations PROMPTOR-FM provides a guide to the usefulness of the screening procedure.

III. THEORETICAL PERSPECTIVE: PROMPTING IN LOW PREVALENCE CONDITIONS

The predictive value of a test is dependent on three variables: a) the sensitivity, b) the specificity and c) the pre-test likelihood, or prior probability of disease. For most practical situations sensitivity and specificity are fixed properties of a given test.[13] The pre-test likelihood however is a reflection of the clinician's judgment.

For example, the sensitivity of the mammogram for breast cancer is about 80 (i.e. the test will correctly detect 80% of individuals known to have breast cancer) and the specificity 90% (i.e. it will correctly identify 90% of those who do NOT have the disease)[14]. If a 50 year old woman presents with a small mobile painless breast lump we can make a judgment that although the presentation is atypical there is a certain probability of breast cancer and that probability may lead us to order a mammogram. We know that mammography is not totally reliable, so what confidence can we have in the result? The answer depends entirely on our judgment of the likelihood of cancer in this woman. If we had 100 patients JUST LIKE HER, we might suspect that 25% had breast cancer.[14] That clinical judgment is the pre-test likelihood. If these numbers are plugged into the PROMPTOR-FM program we will know that a positive test has a 73% probability of being correct and a negative test has a 93% probability of being correct. Both these estimates leaves us much more certain then the original 25:75, and give us reason to follow up a positive test, and some assurance that the benign nature of the exam is confirmed by a negative test.

However suppose this same woman had presented with a 3 cm fixed irregular mass and a family history of breast cancer in a first degree relative we might estimate that of 100 patients of her age, family history, physical findings etc. (i.e. JUST LIKE HER) 70% would have breast cancer. Would a mammogram help us to identify those? Putting these figures into PROMPTOR we discover that the positive predictive value is 95% and negative

predictive value is 66%. We have confidence in a positive result, but a negative test is wrong 34% of the time. Since a positive result leads inevitably to biopsy and a negative test, having a significant risk of missing serious and life threatening pathology, will also lead to a biopsy, it calls into question whether we should order a mammogram in this setting at all.

What about screening a 40 year old woman who has read some of the latest articles in popular magazines and requests a mammogram, having no symptoms, signs or risks. Her prior probability of breast cancer is about 0.5% [15] and therefore a positive test is false 96% of the time. Of a thousand patients there would be 100 FALSE positives for only 4 true positives - clearly a waste of time and money, and worse if the false positives are followed up with more invasive and risky procedures. Screening at this age is not cost effective.

What is clear as the program is run repetitively is that the lower the prior probability of disease the less confidence we have in a positive result. In fact, even changing the sensitivity or specificity just to see what happens leads to little gain in clinical acumen unless we choose exceptionally good sensitivities or specificities. For a test with an excellent specificity the proportion of false positives falls making the test more useful in low prevalence or screening situations. A test with a 80% sensitivity, a 99% specificity used for a patient with a prior probability of just 5% yields a positive predictive value of 81%.

As the sensitivity rises we gain more confidence that we have not "missed" anyone. The clinician therefore can learn to balance the need to identify rare but serious conditions with the risk of false identification of disease and its downstream hazards and cost of further investigation. Each patient is unique and each situation is uniquely analysed. The investigations and be able to counsel the patient based on the facts of the case and the probabilities using a mathematical estimate of the consequences as a tool. The results of the calculations can be stored in the electronic record and/or printed for reference and patient education.

IV. DEVELOPMENT OF PROMPTOR-FM

A "BASIC" program which allows the user to determine the predictive value of a test based on the sensitivity, specificity and pre-test likelihood of disease was developed by Bernstein in 1989.[16] In addition, Hollingworth Wood and Bernstein developed a computerized electronic record for use in a family medicine musculo-skeletal clinic [17]. The two were then integrated.

The PROMPTOR-FM program is called by a "help" button at a point in the record when the physician is ordering diagnostic tests. Its use by the physician is intended to be discretionary although it can be made mandatory for teaching purposes. The program supplies the sensitivities and specificities of the common tests used in an orthopedic medicine clinic and allows the user to input others in the course of using the program.

Pre-test likelihood is always input by the physician as a reflection of his/her clinical judgment. The program asks the purpose of testing (case finding, screening or follow-up) and provides some guidance for each. Results of the calculations are displayed in a manner which "guides and teaches"[18] based on a 2 X 2 table. Suggestions are permissive. The physician can repeat the calculations with different clinical judgment and display the results of current and previous calculations for comparison.

V. EXAMPLE

In the example of the 40 year old woman above the program displays the prompting suggestion as follows:

		DISEASE		
		Present	Absent	
TEST	pos	4	100	104
	neg	1	895	896
		5	995	1000

For the mammogram

with a sensitivity of	80 %
a specificity of	90 %
and a pre test likelihood of	.5 %

the POSITIVE predictive value is	3.9 %
the NEGATIVE predictive value is	99.9 %

This means that if your patient has a positive test this result will be true 3.9% of the time.

Of 1000 patients **just like yours**

you will find 4 **true** positive patients
to 100 **false** positives.

You will **miss** 1 patient who has breast cancer.

DO YOU WANT TO DO THIS TEST?

VI. CONCLUSIONS

A. Low prevalence conditions call for different approaches to prompting for test ordering than when prevalence is high.

B. Prompting systems should guide and teach in the course of usual record keeping.

C. Cost effective test ordering depends on an understanding of the impact of the prior probability of disease on the predictive values of a test.

D. The most important variable in the calculation of the predictive values of a test is clinical judgment of the pre-test likelihood.

E. The downstream cost implications of inappropriate testing where the prior probability is low are enormous, both financial and in terms of the risk to patients from invasive follow up investigations or the unnecessary anxiety generated.

F. PROMPTOR-FM is a prompting tool which can provide the clinician with guidance to appropriate test ordering based on probability of disease, not just suggestions for testing based on diagnosis with a high probability of being ignored.

G. PROMPTOR-FM is designed to promote active learning in the process of guiding the clinician.

VII. REFERENCES

[1]. McDowell I, Newell C, Rosser WW. A Randomized Trial of Computerized Reminders for Blood Pressure Screening in Primary Care. Med. Care 1989;27:297-305.

[2]. McDowell I, Newell C, Rosser WW. Computerized Reminders To Encourage Cervical Screening in Family Practice. J.Fam. Pract. 1989 28:420-424.

[3].Tierney WM, McDonald CJ, Martin DK, et al: Computerized Display of Past Test Results, Effect on Outpatient Testing. Ann. Int. Med. 1987;107(4):569-574.

[4]. Tierney WM, McDonald CJ, Hui SL, et al: Computer Predictions of Abnormal Test Results, Effect on Outpatient Testing. JAMA, 1987;259(8):1194-1198.

[5]. McDonald CJ, Hui SL, Smith DM. et. al. Reminders to Physicians from an Introspective Computer Medical Record. A Two-Year Randomized Trial Ann. Intern. Med. 1984;100:130-138.

[6]. McDonald Clement J. Protocol-Based Computer Reminders, The Quality of Care and the Non-Perfectibility of Man. NEJM 1976;295(24):1351-1355.

[7]. Tailoring Family Medicine Residency Programs to Meet Community Needs, Department of Family Medicine, University of Ottawa, Can. Med. Assoc .J. 1984; 131:1205-1206

[8]. Elmslie, TJ. Integrating Research Findings into the Clinical Setting. in Tools for Primary Care Research Volume 6: Ways to Disseminate Research Findings and Have an Impact on Practice. Ed. Dunn E, Norton P, Stewart M, Tudiver F, and Bass M. Sage Publications 1994 in press.

[9]. Evidence Based Medicine Working Group. Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine. JAMA 1992; 268:2420-2425.

[10]. Canadian Task Force on the Periodic Health Examination: The Periodic Health Examination. Can. Med. Assoc. J. 1979;121:1193-1254

[11]. Canadian Task Force on the Periodic Health Examination: Periodic Health Examination Monograph: Report of the Task Force to the Conference of Deputy Ministers of Health (cat H39-3/1980E), Health Services and Promotion Branch, Dept. of National Health and Welfare, Ottawa, 1980.

[12]. Canadian Task Force on the Periodic Health Examination: The Periodic Health Examination: 1. Introduction. Can. Med. Assoc. J. 1989;141:205-207

[13]. Department of Clinical Epidemiology and Biostatistics, McMaster University. How to Read Clinical Journals: II. To Learn About a Diagnostic Test. Can. Med. Assoc. J. 1981;124:703-710

[14]. Diagnostic Strategies for Common Medical Problems. ed. Panzer RJ, Black ER, and Griner PF. American College of Physicians 1991.

[15]. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 1. Breast Cancer Detection and Death Rates Among Women Aged 40 to 49 Years. Can. Med. Assoc. J. 1992; 147:1459-1476

[16]. Bernstein, RM. A "Basic" Program: Determining the Predictive Value of Laboratory Tests.Canadian Family Physician, 38:1357-1360, June 1992.

[17]. G. Hollingworth, WE. Wood, R. Bernstein. A Self Entry Database for a Musculoskeletal Clinic in Family Medicine. Department of Family Medicine Research Inquiry and Opinion Conference Ottawa May 27, 1993.

[18]. Weed LL. Medical records that guide and teach. New. Engl. J. Med. 1968; 278(11):593-600