A Clinician-mediated, Longitudinal Tracking System for the Follow-up of Clinical Results

by

Daniel Todd Rosenthal

B.S., Biology (1997)
Duke University

M.D. (2001)
University of South Florida College of Medicine

Submitted to the Health Sciences and Technology Division
and the Division of Biomedical Informatics
in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Biomedical Informatics

at the

Massachusetts Institute of Technology

May 16, 2005 [June 2005]
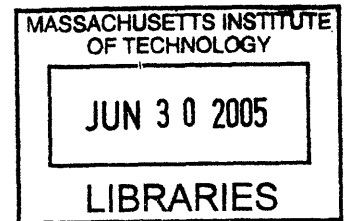
The author hereby grants M.I.T. permission to reproduce and distribute publicly
paper and electronic copies of this thesis and to grant others the right to do so.

Signature of Author ....
Health Sciences and Technology
May 16, 2005

Certified by...
Henry Chueh M.D., M.S.
Assistant Professor of Medicine, Harvard Medical School
Thesis Supervisor

Accepted by ....
Martha L. Gray Ph.D.
Edward Hood Taplin Professor of Medical and Electrical Engineering
Co-director, Harvard-MIT Division of Health Sciences and Technology

A Clinician-mediated, Longitudinal Tracking System for the Follow-up of Clinical Results

by
Daniel Todd Rosenthal

Submitted to the Health Sciences and Technology Division
and the Division of Biomedical Informatics

May 16, 2005

In Partial Fulfillment of the Requirements for the
Degree of Master of Science in Biomedical Informatics at the
Massachusetts Institute of Technology

ABSTRACT

Failure to follow-up on abnormal tests is a common clinical concern comprising the quality of care. Although many clinicians track their patient follow-up by scheduling follow-up visits or by leaving physical reminders, most feel that automated, computerized systems to track abnormal test results would be useful. While existing clinical decision support systems and computerized clinical reminders focus on providing assistance with choosing the appropriate follow-up management, they fail by not tracking that follow-up effectively. We believe that clinicians do not want suggestions how to manage their patients, but instead want help tracking follow-up results once they have decided the management plan. We believe that a well-designed system can successfully track this follow-up and only require a small amount of information and time from the clinician. We have designed and implemented a complete tracking system including 1) an authoring tool to define tracking guidelines, 2) a query tool to search electronic medical records and identify patients without follow-up, and 3) a clinical tool to send reminders to clinicians and allow them to easily choose the follow-up management. Our tracking system has made improvements on previous reminder systems by 1) using our unique risk-management guideline model that more closely mirrors, yet does not attempt to replicate, the clinical decision process, 2) our use of massive population-based queries for tracking all patients simultaneously, and 3) our longitudinal approach that documents all steps in the patient follow-up cycle. With these developments, we are able to track 450 million pieces of clinical data for 1.8 million patients daily.

Keywords: follow-up tracking; reminder system; preventive medicine; computerized medical record system; practice guidelines; clinical decision support system

Thesis Supervisor: Henry Chueh M.D., M.S.
Title: Assistant Professor of Medicine, Harvard Medical School

Table of Contents

1 Introduction

The failure to successfully apply evidenced-based guidelines to the prevention and management of disease has been described as the "quality chasm" by the Institute of Medicine (IOM).[1] Important factors contributing to this "knowledge-performance" gap in delivering quality care include time limitations during the clinical encounter,[2] difficulty managing an increasing burden of clinical data,[3] and lack of appropriate follow-up of abnormal results.[4]

Clinical decision support systems (CDSS) have shown some success improving adherence to guidelines.[5] In addition, computerized clinical reminder (CCR) systems have shown promise; however they are historically "real-time" tools, making recommendations to the provider with the patient present at the point-of-care. Alternatively, other applications have shown success improving guideline adherence by delivering recommendations to the provider "asynchronously" between patient visits.[6]

While these advances aim to improve the quality of care, they also contribute to the increasing volume of data presented to clinicians. In the outpatient setting, the average clinician cares for several hundred patients who often have only one or two office visits each year. With the increasing amount of information to manage, failure to follow-up on abnormal tests is a common clinical concern comprising the quality of care. Furthermore, this failure to follow-up is a significant source of malpractice litigation in outpatient medicine. Currently, providers remember to check follow-up by either a) scheduling a follow-up visit, b) writing a physical note, c) leaving the patient's chart on their desk, or d) manually setting a reminder using an electronic resource such as a calendar program. Providers feel that automated systems to track abnormal test results would be useful.[7]

Our goal was to produce a follow-up tracking system. While existing CDSS and CCR systems focus on providing assistance with choosing the appropriate follow-up management, they fail by not tracking that follow-up. We believe that clinicians do not want suggestions how to manage their patients at the point-of-care, but instead want help tracking follow-up management once the patient has left the office.

We designed our system to reflect this philosophical difference by changing the focus from point-of-care suggestions to follow-up tracking. Our tracking guidelines do not attempt to replicate the decision process, but instead focus on representing all plausible follow-up choices to allow for easy selection for disease follow-up. We are able to catch patients who may fall through the proverbial cracks of follow-up by using a population searching approach rather than tracking patients individually. We search all electronic health records and administrative claims databases to identify groups of patients who need follow-up, send reminders to the clinicians, and then track the follow-up.

Our tracking system provides the solutions to the process of follow-up care (Figure 1) including 1) a unique guideline model created specifically for ease of follow-up tracking, 2) a population-based search engine that identifies sets of patient without follow-up, and 3) a notification tool that clinicians use to view new results and choose follow-up management. Our system then tracks these follow-up plans, continually searches the patient records, and notifies clinicians if follow-up has been missed. We describe the tracking system we created.
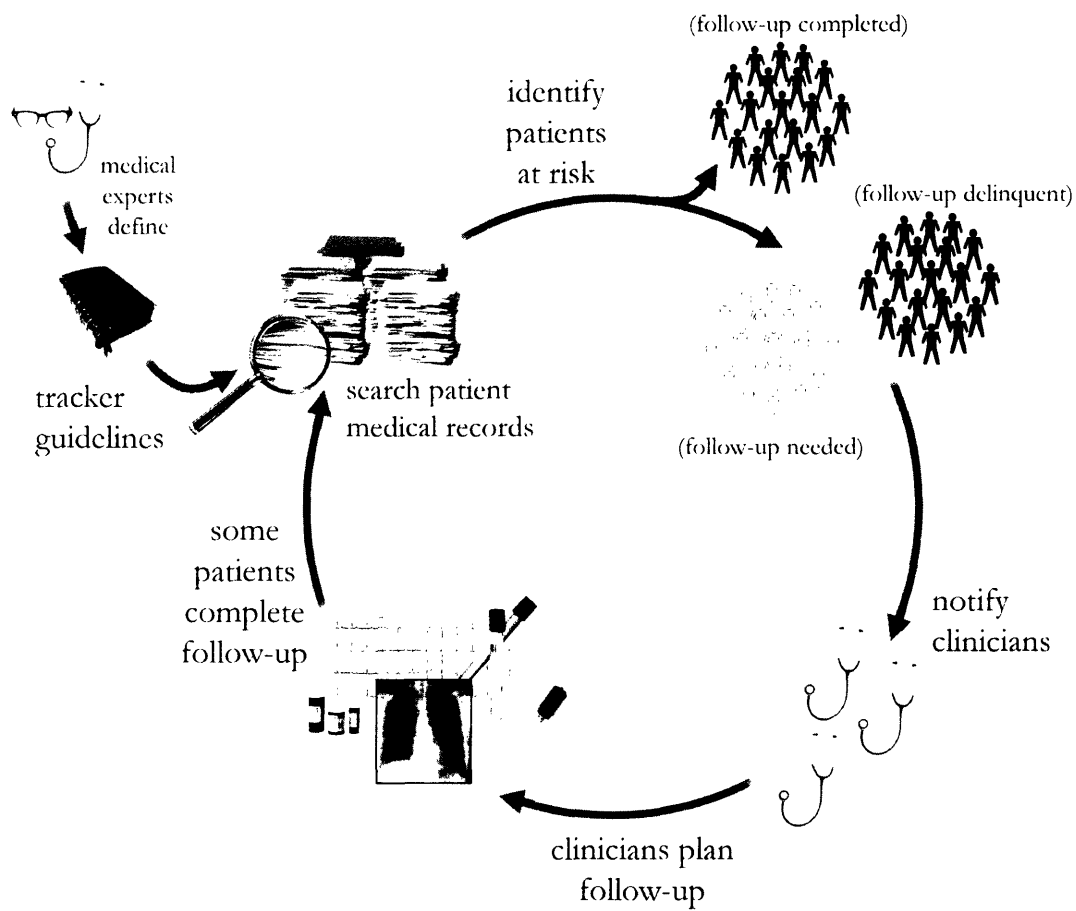
4

Figure 1. Process of Follow-up Care.
To effectively track patients and their follow-up, a system must 1) define what patients require follow-up, 2) search medical records and identify those patients, 3) notify their clinicians, then 4) track their follow-up management. This entire process needs to be repeated frequently and reliably to catch all patients who are missing follow-up.

2 Background

2.1 Preventive Medicine & Appropriate Follow-up

Prevention is increasingly the goal of health care delivery. Prevention strategies differ according to when the preventive action is taken during the disease process, including primary, secondary, and tertiary prevention. Primary prevention occurs before a disease begins to develop. Examples of primary prevention include immunizations and health counseling. Secondary prevention occurs after a disease has begun to develop but prior to symptoms. Most cancer screening programs fall into the category of secondary prevention because they aim to detect existing cancers while the patients are still asymptomatic and unaware of the cancer. In cancer cases this is particularly important as early detection often yields better prognoses. Lastly, tertiary prevention occurs after a disease has developed. While primary prevention may prevent disease and secondary prevention may minimize the impact of an early developing disease, tertiary prevention aims to reduce the negative effects of an existing disease. Chronic disease management is an example of tertiary prevention, such as prevention of end-organ damage from poorly-managed diabetes or mortality from worsening heart disease. Prevention programs that are most successful have a significant public health impact, a long interval of disease progression, an established screening method to identify the target population, and an effective intervention.

However, these preventive efforts are frequently undermined by lack of timely follow-up. This is especially problematic when there is a lack of rapid follow-up in cancer screening. Reducing the burden of morbidity and mortality attributed to cancer have been a national health care priority in the United States for several decades, primarily through prevention.[8, 9] Great strides have been made in improving initial screening rates, but screening alone cannot prevent cancer. For screening to be effective, timely follow-up of abnormal findings is essential. Delays in follow-up and subsequent treatment can decrease survival. Studies demonstrate that many individuals with abnormal screening tests do not receive this timely follow-up. In a study of women with abnormal screening mammograms, inadequate follow-up occurred in 18% of these women.[10] Other studies have suggested that non-white women may experience delays in addressing abnormal mammogram results.[11] In another case series, only 54% of women recommended to have mammographic surveillance after a percutaneous biopsy complied.[12] For cervical cancer screening, studies have shown that between 7 and 49% of women with abnormal Pap smear tests fail to receive appropriate follow-up.[13, 14] In a study of women diagnosed with invasive cervical cancer, 11% had failed to follow-up on a previous abnormal Pap smear.[15] In the Minnesota Colon Cancer Control Study, 5% of individuals with a positive fecal occult blood test (FOBT) did not complete follow-up with a physician.[16] In the Danish FOBT trial, 16% with an abnormal test did not have a complete colonoscopy performed.[17]

2.2 Case Study

2.2.1 Background

Clearly, effective cancer prevention requires screening and appropriate follow-up. For a disease specific needs assessment, we looked at follow-up rates for cervical cancer screening. In 2004, an estimated 10,520 new cases of invasive cervical cancer (ICC) will be diagnosed in the United States, and 3,900 women will die from the disease, accounting for 1.4% of cancer deaths among women.[18] The average latency period from the earliest detectible pre-malignant stage to invasive cervical cancer is between 10 and 20 years.[19] Papanicalou cytology (Pap smear) is the most effective screening test for pre-malignant cervical changes. Abnormal cells detected on Pap smear provide an opportunity for further diagnosis and treatment. Currently, nearly 93% of women in the United States report having at least one Pap smear in their lifetime.[20] While 2% of abnormal pap smears progress to ICC at 24 months,[19] in a 2004 case series 4% of women diagnosed with ICC had a history of an abnormal pap smear without appropriate follow-up in the following 6 months.[21] Thus, the continuing goal of ICC prevention is timely follow-up of abnormal results. Eighteen month follow-up of abnormal results range from 22% to 64%.[22] The goal of this case study is to characterize patterns of follow-up for different Pap smear results over a 4 year period within our own academic health center.

2.2.2 Methods

We analyzed 32,890 Pap smear reports from a cohort of 13,401 women cared for in the Massachusetts General Hospital (MGH) Internal Medicine Associates (IMA) primary care practice between January 1, 2001 and December 31, 2004. All available pathology reports during this period were retrieved from the Clinical Data Repository (CDR) and linked to demographic and administrative claims data from the Research Patient Data Repository (RPDR). We extracted and categorized data from Pap smear reports according to the Bethesda Classification of Cervical Cytology. Abnormal result categories included atypical squamous cells of undetermined significance (ASCUS), atypical squamous cells unable to exclude a high grade lesion (ASC-H), atypical glandular cells (AGC), low grade squamous intraepithelial lesion (LSIL), high grade squamous intraepithelial lesion (HSIL), squamous cell carcinoma, and adenocarcinoma. We limited our data analysis to the 29,863 (91%) reports with cytology quality documented as "satisfactory." We defined the follow-up interval to be the time between the abnormal Pap smear result and a subsequent gynecology office visit or procedure, e.g. repeat Pap smear, human papillomavirus DNA assay, colposcopy, curettage, or biopsy. We chose to examine our data with survival analysis methods which allow censoring of patients lost to follow-up, thus removing them from the study population at the time contact was lost. This methodology identifies the patients lost to follow-up and includes the time up until the loss of contact in our follow-up timing estimates, providing a more accurate view of the actual follow-up trends. For example, if a patient did not receive follow-up for an abnormal Pap smear yet was seen in another MGH clinic 8 months later for an unrelated problem, those 8 months would be included in our analysis, but the patient would be removed from the analysis after that point, regardless of the total time that had passed since the Pap smear was performed. Patients who died before follow-up were censored at their date of death. Patients who did not receive follow-up were censored at the last date that an administrative claim was recorded anywhere within the MGH health care system. We used SAS for the survival analysis. To assess the stability of our data, we created several models by 1) redefining follow-up to include only gynecology procedures, 2) restricting the cases to include only new or progressive disease, and 3) redefining the censoring end date to include only office visits. Since follow-up rates and general trends did not significantly change with any of these alternate models, we chose the most inclusive model for our final analysis.

2.2.3 Results

Of the 32,890 clinical reports, 99.9% were cross-referenced in the administrative claims data, thus validating the use of claims data in this cohort. Of the 29,863 (90.8%) satisfactory specimens, we were able to classify 29,846 (99.9%) yielding 27,753 (92.9%) normal and 2,039 (6.8%) abnormal results (Figure 2).
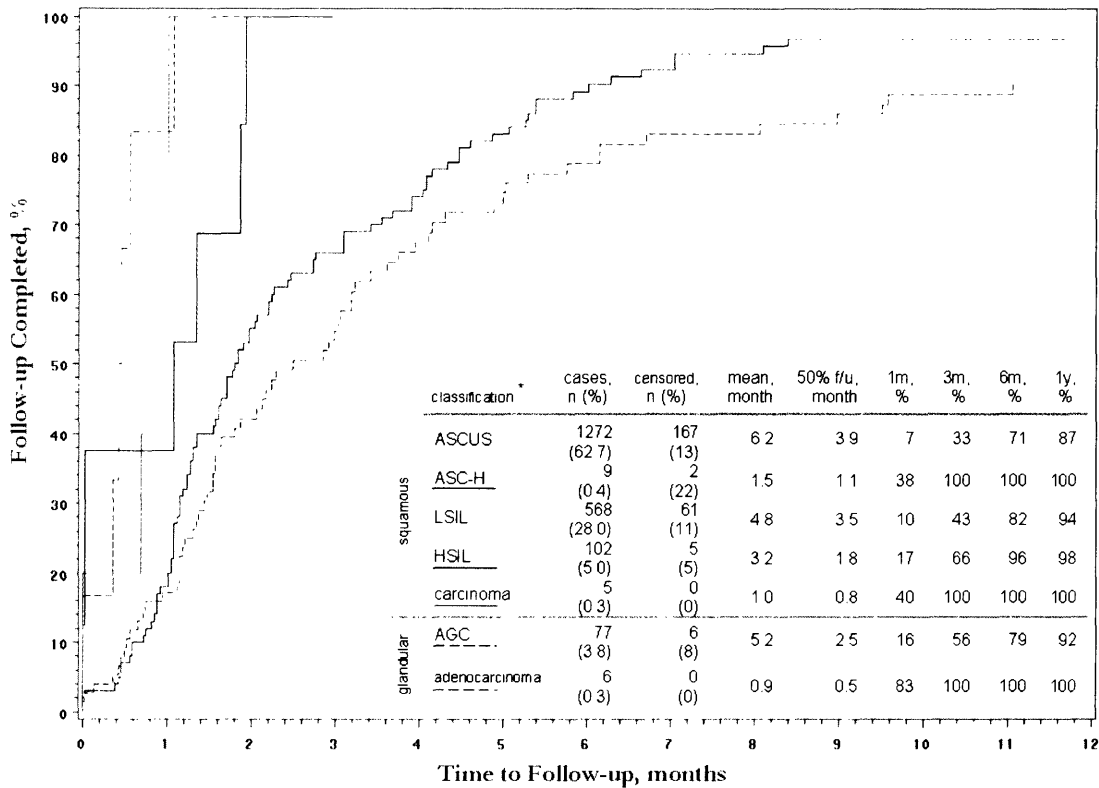

classification *	cases, n (%)	censored, n (%)	mean, month	50% f/u, month	1m, %	3m, %	6m, %	1y, %
ASCUS	1272 (62 7)	167 (13)	6 2	3 9	7	33	71	87
ASC-H	9 (0 4)	2 (22)	1.5	1 1	38	100	100	100
LSIL	568 (28 0)	61 (11)	4 8	3 5	10	43	82	94
HSIL	102 (5 0)	5 (5)	3 2	1 8	17	66	96	98
carcinoma	5 (0 3)	0 (0)	1 0	0.8	40	100	100	100
AGC	77 (3 8)	6 (8)	5 2	2 5	16	56	79	92
adenocarcinoma	6 (0 3)	0 (0)	0.9	0.5	83	100	100	100

Time to Follow-up, months

Figure 2. Time to Follow-up of 2,039 Abnormal Pap Smear Results.
* ASCUS: atypical squamous cells of undetermined significance; ASC-H: atypical squamous cells unable to exclude a high grade lesion; LSIL: low grade squamous intraepithelial lesion; HSIL: high grade squamous intraepithelial lesion; AGC: atypical glandular cells.

Censoring accounted for 239 (11.7%) cases without follow-up and 2 deaths. Of these 239 cases, 55% did not have follow-up at 3 months, 35% at 6 months, 19% at one year, and 14% at 18 months. Of the remaining 1,798 (88.2%) cases with follow-up, ASC-H, carcinoma, and adenocarcinoma received the timeliest follow-up with 100% by 3 months. ASCUS results received the slowest follow-up with 50% at 3.9 months, and 87% at 1 year. AGC, LSIL and HSIL received intermediate follow-up with 50% between 1.8 and 3.5 months, and 92-98% at 1 year, resulting in 122 (16.4%) cases without follow-up at 6 months, and 42 (5.6%) cases at 1 year.

2.2.4 Implications

We identified sub-optimal follow-up of abnormal Pap smear results in a cohort of primary care patients cared for within our academic health system. Our definition of follow-up was chosen to be the most inclusive of all current practice patterns and includes any subsequent evaluation. However, this

preliminary study does not comment on the appropriateness of such follow-up for a given clinical history and Pap smear result. Therefore actual *appropriate* follow-up rates would likely be lower than our estimates, namely for the higher risk lesions of LSIL and HSIL. These empiric data provided justification for designing and implementing our tracking system for result follow-up.

2.3 Identifying Patients without Follow-up

The first task in such a tracking system is to identify at-risk patients without follow-up. This system should monitor these patients along the clinical care pathway (Figure 3).
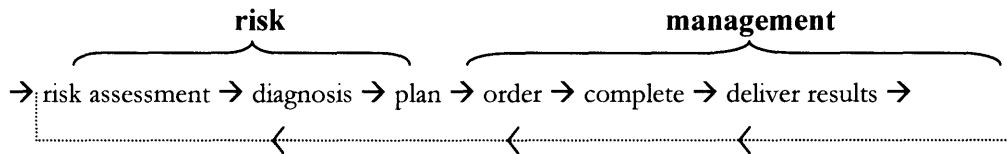

risk   **management**

→ risk assessment → diagnosis → plan → order → complete → deliver results →

Figure 3. Clinical Care Pathway
During a clinical visit, the provider assesses the patient's risk of having a particular disease and decides a preliminary diagnosis. The clinician and patient then decide a management plan. From this point, there are many sequential tasks that need to be completed until these management results return back to the clinician. The process then reiterates with the new information, however with a time delay until this next visit.

To continue with the cervical cancer screening example, during an initial visit (t_0) with a woman, they decide that she is at average, albeit low, risk for developing cervical cancer and they plan to perform cervical cancer screening according to national guidelines. The clinician performs the pelvic exam during the same visit and orders a Pap smear to pathology for evaluation. The remainder of the screening process then occurs without the patient present, asynchronously. At a later date, the pathologist examines the Pap smear and dictates a report which is then sent back to the primary care provider. This process usually takes 1-2 weeks.

The clinical care pathway begins again when the clinician receives the report (t_1). If it is abnormal, the woman's risk of developing cervical cancer increases. At this moment, the clinician processes the abnormal results within the context of the patient's clinical history and usually has a follow-up plan in mind for further diagnostic testing according to national guidelines, thus is a high-yield opportunity to begin tracking this follow-up process. Yet this plan is seldom recorded until the next patient visit. To complete the follow-up, the patient needs to be contacted, schedule a follow-up visit, decide a follow-up plan with the clinician and then execute this management of further testing. Due to the complex nature of asynchronous communications and scheduling, this follow-up cycle (t_1) is prone to system errors and delays. We believe clinicians benefit more from assistance at follow-up (t_1) than during the initial visit at t_0.

We aim to facilitate patient tracking along this care pathway by identifying those at risk for disease who have not had follow-up. This process requires us to identify several sets of patients based on documented risk factors and recorded management actions (Figure 4).

1. R – patients at risk
2. M_c – patients with completed management

As evident in the case study, it is plausible to identify these sets from existing data sources. The reason not all patients in M_c came from R is because sometimes management is completed without explicitly documenting risk factors.

In addition to completed management, it is important to also record management plans. Theoretically, if systems could identify all risks and replicate the decision process from diagnosis to plan, all management could be predicted and successfully tracked. However, this task is complex and not practical to design reliably. Recording the provider plans provides a more reliable prediction of what management results should be performed. Furthermore patients with documented follow-up plans should be tracked differently than patients with no follow-up plan at all. By recording documented follow-up plans, we can identify M_p. Not all patients in M_c are in M_p, because most patients receive management without receiving documented plans.

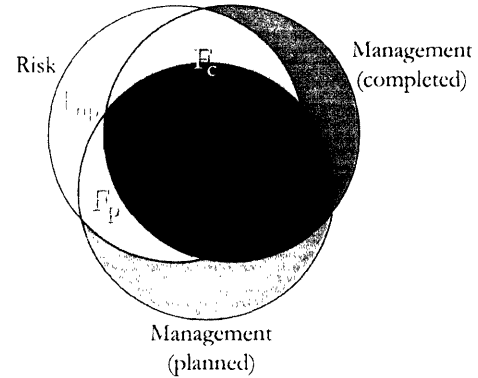3. M_p – patients with documented, planned management

Figure 4. Patient Populations
The large sets are risks (R), management completed (M_c) and management planned (M_p). Patients with completed follow-up are F_c; with planned follow-up are F_p; and no follow-up plan are F_{np}.

From the sets R, M_p, and M_c, we can identify three follow-up sets. Patients at risk who have had completed follow-up, $R \cap M_c$, are F_c. For these patients, it is important to record when the follow-up was completed and *not* send a meaningless reminder to the clinician. Next, patients at risk with no completed management, yet who have a document follow-up plan, $(R - M_c) \cap M_p$, are F_p. It is important to track these patients to ensure their management plan is completed. The remaining at-risk patients, without completed or planned management, $R - M_c - M_c$, are F_{np}. It is important to notify their clinicians of the lack of follow-up.

4. F_c – patients with documented, completed follow-up
5. F_p – patient with documented, planned follow-up
6. F_{np} – patients with no documented plan for follow-up

When a management plan is chosen, those patients move from F_{np} to F_p. As time passes, if the planned management is performed, those patients move from F_p to F_c. However if the management is *not* performed, those patients move from F_p back to F_{np}. It is these patients who move back into F_{np} that become the subjects for delinquent management intervention.

As patients miss more follow-up attempts, more time elapses, diseases progress, and their prognoses worsen. Therefore, these delinquent management patients are the high-yield intervention patients, because the clinician has already decided on management, yet there has been a breakdown in the management pathway. To identify longitudinal errors in this pathway, it is important to maintain a history of how patients move between the three follow-up sets.

2.4 Representing Logic

Before we can begin to identify patients without follow-up, we need a formalism for describing the characteristics of these patient populations. Boolean logic provides such a formalism, using "and," "or," "not" and "if-then" logic represented symbolically by $\wedge$, $\vee$, $\neg$, and $\rightarrow$, respectively. A simple disease

process with only one management plan could be represented in English as "for all patients, *if* they match specific set of risk factors, *then* they should have a specific set of follow-up management."

$$R \rightarrow M$$

The resulting set of patients without follow-up is those with the risk factors *and* who have *not* had management.

$$R \wedge \neg M$$

The risks can be defined as multiple risk groups joined by "or." Each risk group contains one or more clinical elements. For example, a risk group of men > 30 years-old *or* women > 40 years-old could be represented as:

$$(gender="male" \wedge age>"30") \vee (gender="female" \wedge age>"40")$$

Logical clauses may be nested within one another, however most clinical logic is no more than a few layers deep.

$$(gender="male" \wedge (diagnosis="diabetes" \vee diagnosis="heart disease"))$$

Management choices contain similarly structured logic.

$$(procedure="sigmoidoscopy" \vee procedure="colonoscopy")$$

The ultimate structure of a tracker consists of risks and managements.

$$R(riskgroup_1 \vee \ldots riskgroup_n) \rightarrow M(managementgroup_1 \vee \ldots managementgroup_n)$$

Formalizing follow-up tracking as groups of patient sets defined by Boolean logic both mirrors the clinical decision process and is more easily converted into computerized guidelines and database searching programs.

2.5 Guidelines

Clinical care guidelines aim to document and disseminate best-practice patterns to clinicians as we learn more about successful health care delivery. Guidelines are usually consensus statements written by governmental organizations, medical societies, such as the American College of Cardiology, and hospitals. These groups gather data from previous medical studies in a process known as evidence based medicine (EBM), defined as "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients."[23] In the past decade, EBM has helped shape best-practice standards in delivering clinical care.[24]

Dissemination of these guidelines in an environment of rapidly changing clinical evidence has been a challenge for the medical community. As a result, in 1998 the Agency for Healthcare Research and Quality (AHRQ), a branch of the US Department of Health and Human Services, in conjunction with the American Medical Association (AMA), created the National Guideline Clearinghouse (NGC), a website of all published guidelines. Although the clearinghouse currently contains 1,552 guidelines, the NGC does not choose which guidelines are representative for a given disease. Instead, it displays all pertinent guidelines and allows the clinicians to choose which guidelines to follow. Implementing these

guidelines in automated clinical systems can be problematic because they are often written in free-text which requires interpretation to local disease trends and are not structured for easy computer interpretation.

To solve this problem, a longstanding goal of medical artificial intelligence has been to develop computer interpretable guidelines (CIG). Most computational models for representing and interpreting guidelines, including Arden, Asbru, EON, GEM, GLIF, GUIDE, Prodigy, Protégé, and Proforma, use "if-then" logic powered from EMR data.[25] Most models try to represent the full clinical process, all decision steps, supporting patient data, and appropriate management options. The natural extension of these guidelines is to assist with quality healthcare delivery. One mechanism to encourage compliance with guidelines is electronic reminders.

2.6 Computerized Clinical Reminders

Preventive computerized clinical reminder (CCR) systems, working in conjunction with guidelines, have been shown to be effective.[5] Reminders are often delivered by email, alphanumeric pager, printed document, automated phone call, or during computer-based charting or order entry. In primary prevention, outpatient physician reminders increased use of immunization services by an odds ratio of 3.80 (95% CI 3.31-4.37) in a review of 81 studies.[26] In secondary prevention, evaluation of a reminder systems implemented at MGH in 1992 showed long-term changes in compliance with clinical guidelines over 5 years, including cervical cancer screening (64% to 75%) and breast cancer screening (63% to 70%).[27] In Pittsburgh, a Veteran Affairs Hospital reminder system demonstrated an increase in prostate cancer screening (17% to 86%) between 1998 and 2001 after implementing their system.[28] In tertiary prevention, health practice organizations have begun to use clinical reminders to implement a chronic care model at the primary care level, demonstrating significant quality improvement.[29] Despite the success of these systems for basic reminders, little has been done using reminders to track the follow-up process due to significant process barriers.

2.7 Barriers to Guidelines and Resulting Reminders

Although CIG and CCR systems have the potential to improve the quality of care, their actual effectiveness is diminished by both technological and clinician barriers. With so many versatile CIG models and formats, the main technological barrier to their adoption is *not* representing the guideline programmatically; rather it is logistic challenges in gathering supportive evidence to power the guideline logic. Clinical data is represented differently within each electronic health record (EHR). When a guideline is applied to local data, each term must be mapped accordingly. Even if this challenge could be easily overcome, the clinical data is often not recorded in the detail that the guidelines require. Diagnoses and some physical findings may be recorded in the EHR and contained in a problem list, however more specific details are frequently recorded in free-text and not codified. Therefore these guidelines often can not function without requiring the clinician to input further patient data, which interrupts an already busy office visit.

Even if we assume that guidelines can be mapped to local patient resources and all available patient data exists to power the guideline, the larger question remains regarding the usefulness of guidelines to experienced clinicians. While clinicians prefer to have reminders delivered during the clinical encounter, they tend not to interact with reminders that request further information, as it interrupts their clinical workflow.[30] Furthermore, CIG systems often generate reminders that are not specific to the patient due to lack of full electronic data. For example, clinicians may receive reminders that no longer apply to

a patient, because the management had already been performed. In addition, clinicians are often aware of best-practice guidelines and feel CIG systems generate reminders that do not provide them with additional information, but instead overwhelm them with requests that provide minimal additional value to their clinical care. CIG systems fail to deliver their intended outcome due to these technical issues and lack of clinician buy-in.

Previous efforts to develop these systems have focused on what is possible, rather than on what is needed. Simply because it *can* be done, does not mean that it *should* be done. Many guideline and reminder systems attempt to replicate the clinical decision process leaving clinicians with recommendations that they usually already know. We believe the low-hanging fruit of reminder systems is *not* in making management suggestions; instead clinicians need assistance tracking the follow-up management results.

2.8 Goals for Tracking System

With these barriers in mind, we set out to develop a unique, population-based tracking system. Successful guideline and reminder systems need to operate as an *adjunct* to the clinical process as opposed to a *substitute*. We set out to develop a tracker system that mirrors the bare essence of the clinical decision process, assists clinicians at the optimal time for intervention during the clinical care pathway, and can track unlimited patients and disease processes. Our system assists the busy clinician by tracking their intended follow-up, rather than simply suggesting a follow-up plan that the clinician already knows.

2.8.1 Risk-Management Stratification

While most guidelines attempt to assist providers with clinical decisions, we believe that clinicians would benefit more from tracking of follow-up once those decisions have been made. Only a small amount of information is required to begin the tracking process. The process of risk-management stratification requires the aggregation and processing of many risk factors, and then the selection of one or more of many management options. However, no matter how complicated the risk assessment, the clinical information is always distilled to a single risk-management stratum (Figure 5). The follow-up for a given stratum is often limited to a handful of subsequent tests or procedures, e.g. abnormal chest x-ray → repeat x-ray, computerized tomography, or surgical consult; abnormal Pap smear → repeat Pap smear, colposcopy, curettage, biopsy, or gynecology consultation; high LDL → repeat LDL to assess management effectiveness. These limited options are thus easier to codify in guidelines and are usually process-oriented and often available in administrative claims data. Therefore, the most efficient entry point is at the risk-management stratum, for it is the most broad and least granular concept in the clinical decision process, and therefore the easiest to represent. If we allow a clinician to jump right to a risk-management stratum, we can quickly display follow-up choices.
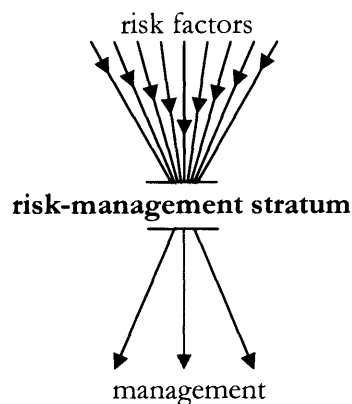

Figure 5. Risk-Management Stratum
A complex collection of risk factors is simplified to a single risk-management stratum, e.g. high risk, through the clinical decision process. This stratum usually an easily definable collection of management follow-up options.

By focusing on the risk-management strata, we avoid attempts to replicate the complex risk assessment process. Instead, we provide *some* basic risk factors but let the clinician perform the majority of the risk-management process by presenting few risk-management strata. From there we do not suggest which management is appropriate. Instead we simply present all management options related to that stratum

and leave the management decisions to the clinician. For these reasons, we built our tracker model around this clinical concept of risk-management stratification.

2.8.2 Timing of Intervention

We believe the optimal moment for guideline intervention is not during risk assessment, rather during results review. To properly interpret the results, the clinician is required to recall the patient history and clinical details and is thus primed to plan for the follow-up. If follow-up is needed, it is decided at this point, often when the patient is not present. Therefore, this asynchronous process is more prone to errors of lost communications and is an opportunity for more successful intervention. We believe that providing guidelines and reminders at results review eliminates the low-yield task of duplicating the clinical decision process and focuses on the high-yield results of process-oriented clinical data tracking. Thus our primary goal of patient tracking is to asynchronously facilitate the follow-up of disease processes.

2.8.3 Scope and Scale

For the greatest clinical impact, we aim to develop a system that can track any disease process, ranging in context from annual screening for diabetes to the specific follow-up of an abnormal Pap smear. From a public health perspective, we aim to be able to track *all* patients for *all* disease processes using *all* available clinical data. Our system is specifically designed on this massive scale.

Most reminder systems were designed to work synchronously during a clinical encounter, therefore were often built to collect and process data on a single patient at a time. This data process is often performed during the visit by accessing multiple data sources and running logical rules one by one for specific disease processes. Alternatively, this individual patient query is performed when a new clinical result is received electronically. This design works well for small a small number of patients (100's - 1000's) and a small number of simple guidelines. However to track patients *not* during a visit would require searching all data resources for all patients every day or by setting hundreds of these triggers for any result related to a disease process. Existing reminder system models simply can not support the volume of clinical data and processing power required to track a large volume of patients asynchronously between visits.

We propose a tracking system designed not for the individual patient, but instead for the entire patient population. Rather than gathering clinical data as needed, we choose to maintain all available clinical data, and process *all* disease trackers for *all* patients using *all* clinical data in large batch processes offline, and simply record the results of these processes. The advantages of this approach allow us to provide:

1. support for a large scale
2. quality assurance on centrally located data
3. instantaneous access to pre-processed results
4. more clinical data from other offline systems such as administrative claims

3 Design

Our goal is to identify patients who have not had follow-up, notify their clinicians, and then track the subsequent follow-up. To accomplish this task, our tracking system searches patient data using defined rules for each disease process. Our tracker guidelines provide the logic for these searching rules. These guidelines automatically search the clinical and administrative data to identify and track the patients without follow-up. When patients fall off track, the system generates reminders to the clinicians, who use the system for both chronic disease management and follow-up tracking. We describe designing the guideline model, identifying patients without follow-up, generating reminders, and lastly the technologies we used in the application architecture.

3.1 Guideline Model

Our guideline model (Figure 6) contains multiple top-level trackers for each disease process we want to track ("Cervical Cancer"). Within each tracker are one or more risk-management strata that represent the culmination of the risk assessment process ("routine screening"). Each risk-management stratum contains a single risk set and an accompanying single management set; our goal is to ensure that the patients in the risk set have had follow-up and therefore are also in the management set. The risk set and management set contain one or more groups, each composed of one or more clinical elements, e.g. "male > 40 years old," "diabetes and LDL > 100mg/dl", and optional or-groups with one or more nested clinical elements.

```
<trackers>                                   <managementset>
    <tracker₁>                                   <managementgroup₁>
        <riskmanagementstratum₁>                     <clinicalelementₙ/>
            <riskset>                                and
                <riskgroup₁>                         <orgroupₙ>
                    <clinicalelement₁/>                  <clinicalelementₙ/>
                    and                              </orgroupₙ>
                    <clinicalelementₙ/>          </managementgroup₁>
                    and                          or
                    <orgroup₁>                   <managementgroupₙ>
                        <clinicalelement₂/>          ...
                        or                       </managementgroupₙ>
                        <clinicalelementₙ/>      </managementset>
                    </orgroup₁>                  </riskmanagementstratum₁>
                    and                          or
                    <orgroupₙ>                   <riskmanagementstratumₙ>
                        ...                          ...
                    </orgroupₙ>                  </riskmanagementstratumₙ>
                </riskgroup₁>                    </tracker₁>
                or                               or
                <riskgroupₙ>                     <trackerₙ>
                    ...                              ...
                </riskgroupₙ>                    </trackerₙ>
            </riskset>                       </trackers>
```

Figure 6. Tracker Guideline Model in XML
Each tracker contains risk-management strata, risk/management groups, or-groups and clinical elements

The smallest units in our model are the clinical elements. These elements describe patient clinical data including demographics, diagnoses, procedures, clinical encounters, clinical providers, medications and laboratory tests. A patient's health history is a collection of these clinical elements. Demographic data

are continuous (age) and categorical (race, gender). Diagnoses, procedures, medications, clinical encounters and clinical providers are categorical. Laboratory tests may be either categorical or continuous depending on the test result type. All clinical elements have a unique full name and an abbreviated name. Constraints for each clinical element may include dates ("in the past 1 year") and value limits and their logical operators ("> 100mg/dl"). For example, to identify the set of all patients who had a low density lipoprotein (LDL) level greater than 100mg/dl in the last year, the clinical element would look as follows, represented in XML:

```
<clinicalelement name="LDL" fullname="Lab\Chemistry\LDL"
    operator="greaterThan" lowerValue="100" units="mg/dl" dateLowerLimitNum="1"
    dateLowerLimitUnit="year"/>
```

These clinical elements are used as both risk factors and management choices. The results of today's management choices become the potential risk factors for tomorrow's risk assessment. For example, *if* a person has an LDL > 100mg/dl (today's risk factor) *then* an LDL should be repeated in the next year (tomorrow's management choice). The results of this repeated LDL test will then serve as the risk factor for its subsequent follow-up, and so on.

3.2 Clinical Element Dictionary

The clinical elements are the language we use to understand the patient health records. To successfully search these records, this language must contain a standardized vocabulary of terms, or a lexicon. Numerous medical lexicons have been created to define health records. In an effort too combine all available medical lexicons, the National Library of Medicine created the Unified Medical Language System (UMLS). As a result, this dictionary of dictionaries, called a metathesaurus, contains information about over 5 million biomedical terms from more than 100 lexicons used for patient records. Importantly, the UMLS contains vocabularies and coding systems designated as US standards for the exchange of administrative and clinical data.

For our clinical element dictionary, we chose a lexicon built from UMLS but with additional terms defining local data that are not contained in the UMLS. Our patient data are contained in the Research Patient Data Registry (RPDR).[31] This data warehouse contains patient data from multiple hospital systems within the Partners Healthcare System in Boston, Massachusetts. The RPDR metathesaurus contains 63,000 terms with corresponding UMLS codes, including demographics (age, gender, race), diagnoses (Internal Classification of Diseases, Ninth Revision), laboratory results (Logical Observation Identifiers Names and Codes (LOINC) and local terminology), medications (National Drug Code), procedures (Current Procedural Terminology), and clinical encounters and providers (local terminology). The structure of the RPDR metathesaurus is hierarchical (Figure 7). Broadly defined terms are grouped into categories that contain more specific categories; these categories, in turn, contain even more specific categories, and so on. At the bottom of this hierarchical structure are the individual clinical elements. Each of those groups contains even more specific categories.

The advantages of using the RPDR metathesaurus are that 1) it is derived from the UMLS which allows us to integrate our system with other UMLS-based systems, and 2) its hierarchical structure allows for simplified clinical element terms. To give an example of how using a hierarchical metathesaurus can help simplify guideline terminology, defining the laboratory test "cholesterol" can be simplified to a single term that represents all related cholesterol lab names. The RPDR metathesaurus maps all local lab names and codes from different systems, in different hospitals, from different points in time, all to UMLS terms. The LOINC code for "cholesterol" contained in the UMLS is 2093-3. However, there are 23 different local definitions for this single concept within the Partners Healthcare System, because many systems are older legacy systems and do not take advantage of standard concept coding. To fix this problem, the RPDR metathesaurus represents all of these terms in a single "Labs\Chemistry\Cholesterol" group. If a guideline requires the lab term "cholesterol," the author may include all 23 terms in the guideline within an or-group, or simply include the one RPDR term "Labs\Chemistry\Cholesterol," thus simplifying the data representation. If the guideline wanted to refer to any chemistry lab, it would simply be represented as "Lab\Chemistry" rather than all of its 3,174 individual chemistry terms.
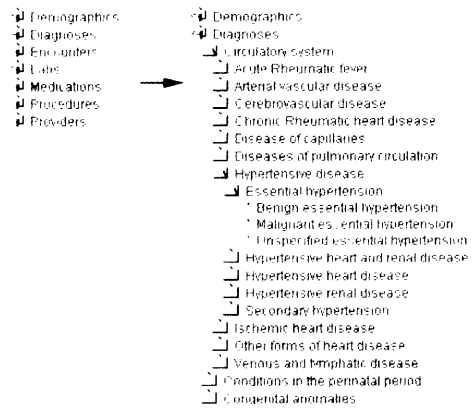

Figure 7. RPDR Metathesaurus Hierarchy
Each category contains either subcategories or individual elements with increasing specificity. In this example, the circulatory diagnoses category contains the subcategories of rheumatic fever, vascular disease, cerebrovascular disease, hypertensive disease, etc.

3.3 Patient Data

Once we define our guidelines with clinical elements, they guide the search through patient data. Our system uses patient data from multiple data warehouses. These clinical repositories contain diagnoses, medications, laboratory test results, procedures, visit information and clinical providers. Information missing in one repository is complimented by data from another. Primarily we use the RPDR data which are stored in a Microsoft SQL 2000 database running on a Windows 2000 server. This data warehouse contains over 450 million clinical elements for 1.8 million patients seen at either Massachusetts General Hospital (MGH) or Brigham & Women's Hospital (BWH) from as far back as 1987. Additional information is extracted from the Partners Clinical Data Repository (CDR).

3.4 Identifying Patient Sets

After we search the patient data using our guidelines, we generate summary scores for all guidelines for all patients. An example question the tracker aims to answer would be, "Have all women with a Pap smear result had timely acknowledgement of the report and follow-up, if abnormal?" The tracker would be "cervical cancer" and the risk-management stratum would be "routine screening." The risk set would contains all women with a Pap smear report, and the management set would be either a repeat Pap smear after some time interval or another gynecologic procedure or office visit.

MRN	tracker	maximum risk set	maximum risk score	maximum management set	maximum management score
000004	annual check-up	(gender="male" ^ age>30)	0.5 (½ match)	(procedure= "chest x-ray" ^ lab="chem7" ^ lab="CBC")	0.33 (1/3 completed)
152342	annual check-up	(gender="female" ^ age>40)	1.0 (all match)	(procedure= "chest x-ray" ^ lab="chem7" ^ lab="CBC")	1.0 (all completed)
000004	heart disease	(lab="LDL>100")	1.0 (all match)	(lab="LDL" ^ medicine="statin")	0.0 (none completed)

Table 1. Summary Patient Scores for All Risk-Management Strata
For example, for the risk set of ((gender="male" ^ age>30) v (gender="female" ^ age>40)), a 25 year-old man would receive 1 out of 2 points (0.5) for matching half of the risk group (gender="male" ^ age>30), but would receive a score of 0 for not matching any of the risk group (gender="female" ^ age>40). Therefore, his risk set score would be his maximum score for those risk groups, 0.5. Alternatively a 65 year-old woman would receive a risk set score of 1.0, for matching both (gender="female") and (age>40). Management scores are also recorded.

To identify these patients, we calculate scores, between 0 and 1, representing how close the patients match the clinical scenario (Table 1). For each tracker risk-management strata, all patients receive both a risk set score and management set score. These scores are the maximum risk group score and maximum management group score, respectively. Each group score is the fraction of clinical elements that match for that patient. If two groups tie for the maximum risk or management set score, both are recorded. After scores for all patients are calculated for all trackers and their risk-management strata, they are recorded in a database.

3.5 Reminder grouping

With these scores, we track patients as they move through the clinical care process and send reminders to clinicians when patients fall off track. To avoid sending a high volume of reminders to clinicians and thus overwhelming them, we built intelligent grouping into the guideline model. These reminder groupings allow us to send results in batches rather than separately. Each management clinical element can have a separate alert type of either none, routine, or priority. These alerts are defined by the minimum time to wait before reporting grouped results to the clinician and the maximum time to wait before reporting missing results (Table 2).

result alerting type	min wait time before reporting results	max wait time before reporting missing results
none	-	-
priority	0 (deliver each immediately as they arrive)	t_{max-p}
routine	t_{min-r} (deliver together but abnormal immediately)	t_{max-r}

Table 2. Grouping Types and Timing

For example, if a clinician chooses to follow-up a patient by ordering management consisting of (lab_1 ^ lab_2 ^ $procedure_1$ ^ $procedure_2$), a clinician may choose *routine* delivery for the labs (t_{min-r} = 3 days, t_{max-r} = 7 days), and *priority* delivery for the procedures (t_{max-p} = 1 month). If lab_1 results arrive to the electronic health record in 1 day, it will not be sent to the clinician until waiting a total of 3 days (t_{min-r}) for lab_2 results to arrive. After 3 days, lab_1 and lab_2 results, if available, will be sent together. If for some reason a lab result is not received, a delinquent alert will be sent to the clinician after the t_{max-r} of 7 days is reached. For the priority procedures, since the minimum wait time is 0, as soon as the result is received, it will be sent to the clinician. The t_{max-p} is set higher, to 1 month, because procedures usually take longer to schedule, perform, interpret and report. Alternatively,

18

if a clinician chooses to check results manually and does not want to be alerted by the system, an alert type of *none* is chosen.

To define these reminder groupings, the following attributes are added to the management clinical elements: alertType {"none", "routine", "priority"}, alertMinNum {number}, alertMinUnit {"day", "week", "month", "year"}, alertMaxNum {number}, alertMaxUnit {"day", "week", "month", "year"}. The above example is represented in XML:

```
<managementgroup>
    <clinicalelement name="lab₁" alertType="routine"
    alertMinNum="3" alertMinUnit="day"
    alertMaxNum="7" alertMaxUnit="day"/>
    <clinicalelement name="lab₂" alertType="routine"
    alertMinNum="3" alertMinUnit="day"
    alertMaxNum="7" alertMaxUnit="day"/>
    <clinicalelement name="procedure₁" alertType="priority"
    alertMinNum="0" alertMinUnit="day"
    alertMaxNum="1" alertMaxUnit="month"/>
    <clinicalelement name="procedure₂" alertType="priority"
    alertMinNum="0" alertMinUnit="day"
    alertMaxNum="1" alertMaxUnit="month"/>
</managementgroup>
```
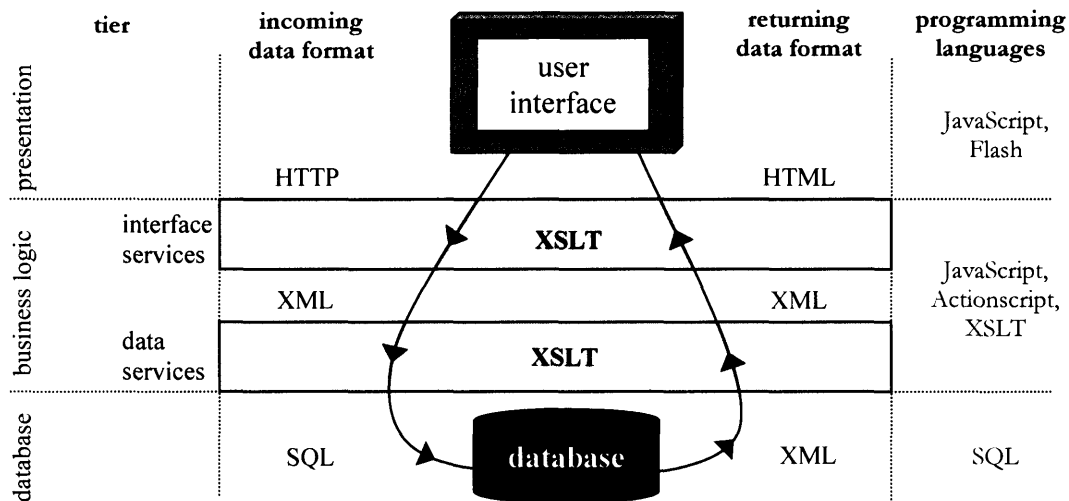
3.6 Application Architecture

Figure 8. Three Tier Application Architecture
As information travels through the application infrastructure, it is processed using different programming languages, converted to other data formats, and transmitted to the next tier.

Creating guidelines, searching and scoring patient data, and generating reminders are all controlled by computer programs within our application architecture. We use a number of technologies and computing languages across three tiers: database, business logic, and presentation (Figure 8).

The database tier consists of both the data storage and the technologies to access these data. We use the structured query language (SQL) to retrieve data from the Microsoft SQL Server relational databases. For data sources that do not allow direct SQL queries, such as the CDR, we use other services to collect those data and store a copy in a local database that we can then query with SQL.

The business logic tier contains the functional programs that communicate between the data and the user interface. Our applications run on Internet servers using Microsoft Active Server Pages (ASP) written in JavaScript. These applications are divided into data services, programs that query and gather information from the database tier, and presentation services, programs that format these data for the user interface. Data is passed between the database and presentation tiers through the business logic tier in the eXtensible Markup Language (XML) standard. These XML data are transformed to the format required by the destination tier using eXtensible Stylesheet Language Transformations (XSLT).

The presentation tier is the display that allows the user to interact with the data. One presentation technology we use is a web browser, displaying data with the Dynamic Hypertext Markup Language (DHTML). Additional presentation logic is programmed within the browser using JavaScript, providing a direct connection to the business logic. This technology enables us to update the webpage dynamically without the need to reload the page.

Another presentation technology we use is Macromedia Flash, a multimedia authoring environment that is presented in the Flash Player, a virtual machine application that runs within a web browser. Flash applications are written in ActiveScript 2.0, an object-oriented language.

4 Implementation

Our tracking system addresses each challenge of the process of follow-up care (Figure 9). The tracking process begins with medical experts defining the guidelines. These guidelines are then used to automatically search patient medical records and identify patients without follow-up. Messages are sent to clinicians who can then view these patients and choose the appropriate follow-up management. If the management occurs on time, those results become part of the patient's medical record, and the next time the patient record is searched, that patient will now have documented follow-up. However, if the management has not occurred, automated reminders are sent to the clinician. This process of searching for patients without follow-up, tracking subsequent follow-up and generating reminders occurs automatically every day as new results are received and new managements are chosen.
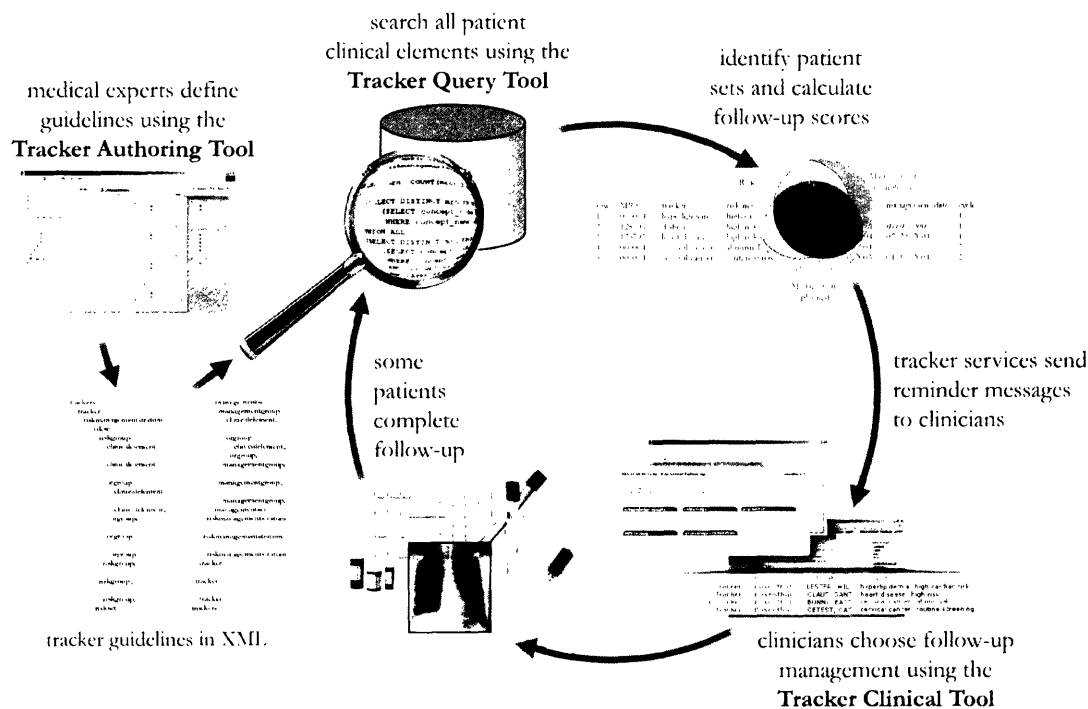

Figure 9. Process of Follow-up Care in the Tracker System
The tracker system encompasses all aspects of the follow-up care process. Medical experts create tracker guidelines using the Tracker Authoring Tool. The Tracker Query Tool automatically searches all patient medical records daily, identifies patients without follow-up and records their status for all trackers in the watchlist. Other services access the watchlist and generate messages to clinicians, such as a new result requires follow-up, or a patient requires annual cancer screening. Third party applications present these messages to clinicians. Clinicians then choose a management option using the Tracker Clinical Tool. Once management is planned, the entire process begins again, and all patients are tracked until all issues requiring follow-up have been resolved.

21

Most of our program interfaces are designed for a rich, web-based environment. Utilizing the Internet allows clinicians to access our tools using a standard Internet browser and removes the need to distribute and maintain client-side software. However with the advantages of a web-based environment come the limitations of a less interactive, sluggish experience using a web-browser. Therefore, to create a smooth, rich experience we employ several web-based techniques including client-side HTTP, hierarchical navigation and drag-n-drop. These technologies make the user interaction more comparable to standard window-based applications.

Client-side HTTP technology utilizes JavaScript objects to directly connect to server-side data and subsequently modify the browser interface with Dynamic HTML. Therefore, the user does not have to submit the webpage to update the data; the result is a smooth interactive experience with minimal observed delay and preservation of the interface and data state. This client-side technology is especially useful for loading large amounts of data in small batches as needed, providing a quick initial load and allowing for hierarchical navigation.

Our interface mirrors the underlying data structure. Tracker guidelines and the RPDR metathesaurus are both structured hierarchically. This organization is intuitive to users and mimics how data is organized in disease-related groups with increasing clinical detail in subgroups. At the top level, the user sees only the most basic groups, but as needed, may expand each group to see nested subgroups and other disease-related items. For example, although the RPDR metathesaurus contains 63,000 terms, initially only the top 7 categories are displayed: Demographics, Diagnoses, Encounters, Labs, Medications, Procedures, and Providers. Expanding the Diagnoses group reveals the main categories of diagnoses. This hierarchical structure prevents the user from information overload and results in a clean graphical user interface.

medical experts
define the
tracker guidelines
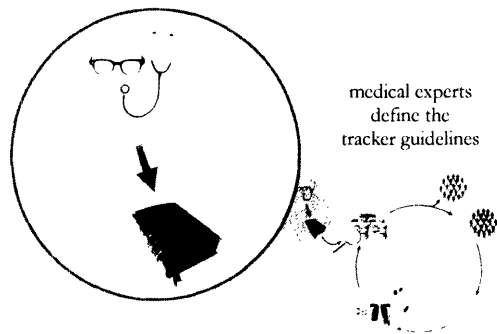
4.1   Tracker Authoring Tool

To begin the tracking process, clinical experts define the tracking guidelines. The Tracker Authoring Tool is a web-based application that enables clinical experts to build tracker guidelines from clinical elements (Figure 10). The user interface displays the tracker navigator on the far left (a), the risk-management composer in the main portion of the screen (b), and system messages and tips in the bottom left (c).

The tracker navigator hierarchy (a) is retrieved using the Tracker Definition Service (TDS), one of many data services that queries the Tracker guideline database in SQL, then transforms the resulting tracker XML guidelines dynamically into HTML for display. The tracker navigator mirrors the structure of the underlying guideline model, such that authors can view trackers, risk-management strata, and clinical elements hierarchically. Using the tracker navigator, the author can create, view and edit trackers and risk-management strata.
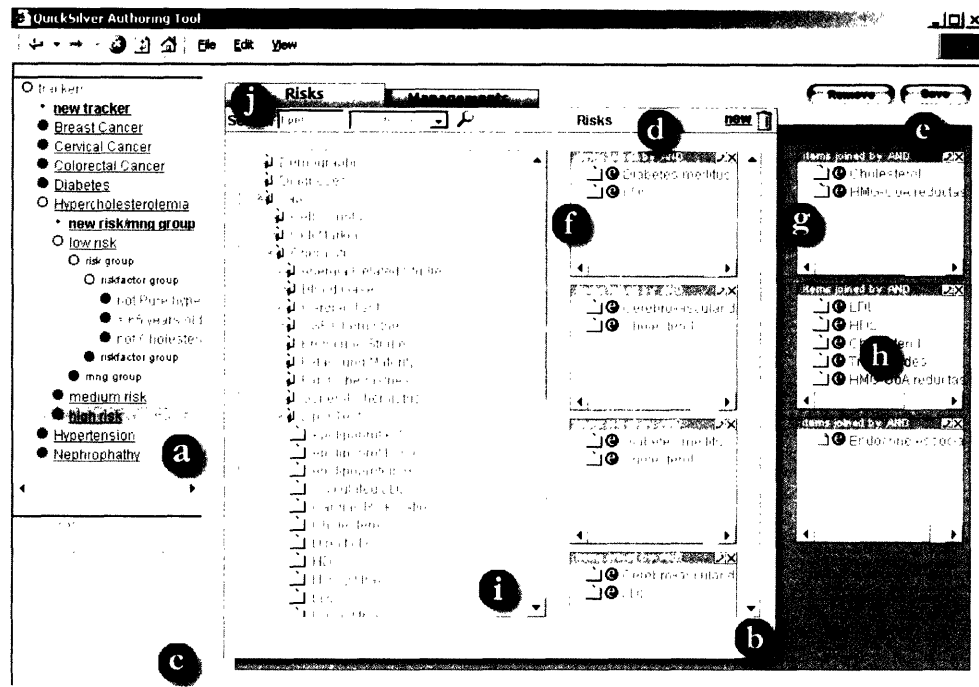
Figure 10. Tracker Authoring Tool

The tracker navigator (a) displays the current tracker hierarchy with embedded action links to author new trackers. The risk-management composer (b) graphically represents the risk-management stratum which includes the risk set (d) and the management set (e), each with several groups (f, g) consisting of clinical elements (h). The composer center space contains the RPDR metathesaurus (i) and searching tool (j). Confirmation messages and tips are shown in the system message window (c).

Once the author selects a risk-management stratum, the remainder of the authoring occurs in the risk-management composer (b). Since the risk set (d) and management set (e) in the tracker model have the same data structure, graphically they appear similarly, side by side. The author selects either the risk or management set by clicking the appropriate tab at the top of the composer; the corresponding set is displayed in the center of the screen (d), and the other set is displayed to the right (e). The author can view both sets of groups simultaneously, for they are intimately related in terms of clinical logic, i.e. R → M.

The author creates a new risk group by clicking "new" in the composer tool bar (j). The author then finds the relevant clinical elements either by manually navigating the metathesaurus hierarchy (i) or by searching for terms explicitly using the search tool (j). The RPDR metathesaurus display data are managed by RPDR Metathesaurus Service. To add the clinical element to the group, the author drags the element from the metathesaurus and drops it in the appropriate group. The author can select individual items (⌐ icon) or an entire RPDR category (⌐ icon) as a clinical element.

The author then defines the attributes of each clinical element. If the author is editing a risk group, risk-oriented attributes are displayed (Figure 11). All RPDR data types have an inclusion date interval, which narrows the dates that the guideline searches for patients with a clinical element. Laboratory risk factors have additional value attributes ("> 100") and normal/abnormal ranges. A mouse-over these areas in this window explains the attribute. In addition, the author may indicate that this clinical element should be the *lack* of evidence, by clicking the NOT logically box.

23

Define LDL (mg/dl)

Figure 11. Risk clinical element definition.

For continuous laboratory terms, the RPDR metathesaurus provides ranges of normal and abnormal values as defined by different laboratories with the Partners Healthcare system, as indicated by the gradient color bar. The red area indicates that results in this range are considered abnormal; orange indicates that *some* laboratories consider these values abnormal; and, green indicates that values in this range are universally considered normal. This clinical element would identify patients with an LDL value greater than 100mg/dl in the past 1 year.

Alternatively, management clinical elements contain alert grouping parameters (Figure 12). For example, the date interval here is the time from when the patient initially became at-risk up to the time at which this management should have occurred. The alert grouping displays the default values for that RPDR clinical element.

Define LDL (mg/dl)

Figure 12. Define management clinical element.

This management definition identifies patients who have not had an LDL result within 3 months from when they became at-risk ("LDL > 100mg/dl"). The alerting is set to "routine," and results will be held up to 1 week until other grouped results are received. If no results have been received after 1 month, the system sends a delinquent alert.

24

After the clinical elements are chosen, the author can modify each group by using the 1) title, 2) or-group, 3) date-group options (Figure 13). A title can be defined by clicking the pencil (⌐ icon).The author may save this group to be used in other trackers (⌐ icon). This saved group is especially useful for defining panels, such as representing a CBC as a panel of WBC *and* Hct *and* HgB *and* platelets. Similarly the author may create a nested or-group (⌐ icon) which is useful for simplifying the graphical representation, e.g. [(a ^ b) v (a ^ c) v (a ^ d)] can be simplified to [a ^ (b v c v d)]. Or-groups can be saved (⌐ icon) which is useful for defining a synonym group of terms that have similar implications (colonoscopy *or* sigmoidoscopy *or* barium enema are similar means for colon cancer screening). Additionally, terms in the RPDR that may not be grouped semantically as the author would like, can be regrouped using the or-group. Both panels and synonym definitions are then available to the author as additions to the RPDR metathesaurus. Lastly date-groups (⌐ icon) allow the author to temporally constrain clinical elements, relative to each other. For example, the author may define a risk group as individuals with a particular lab result in the setting of a certain procedure. Although each term date interval may be set for the previous 2 years, using the date-group, the author may specify that the lab and procedure have occurred within 3 days of each other.
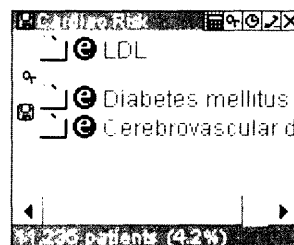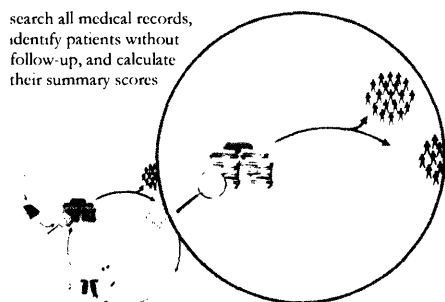

Figure 13. Group Window
Within each group, the author may define additional or-groups, date-groups, and define synonyms and panels. This example matches patients with a high LDL *and* either diabetes *or* cerebrovascular disease.

For real-time feedback, the author can view how many patients match the group. When the author has completed the group, clicking the calculator (⊞ icon) counts the number of patients who match the group logic. This is done with the Tracker Query Tool on a representative sample of the RPDR clinical data. The results are displayed at the bottom of the group window. The author can use this real-time feedback to modify the group if no patients match.

The author defines all risk groups and management groups. Clinical elements can be dragged and duplicated between group windows, using standard keyboard shortcuts. Finally, the risk-management stratum is completed by clicking the "save" button in the upper right corner of the composer. The new definitions are updated and displayed in the Tracker Navigator. During each step in the process, confirmatory messages and tips are displayed in the system message window. At any time, the author may return to this tool, modify existing trackers or create new trackers.

search all medical records, identify patients without follow-up, and calculate their summary scores

4.2 Tracker Query Tool

Once the guidelines are created, they are used repeatedly to identify patients without follow-up. The Tracker Query Tool uses the guidelines to search the warehouse of patient data. Run daily, it begins by gathering patient data from different databases and converting the data to one common format. The query tool then reads all guidelines, converts them to SQL, searches the data and generates patient scores for all trackers. These results are then used by other programs to generate reminder messages.

4.2.1 Gather Patient Data

Before the tool begins its query, it gathers data from three different sources: the RPDR, Clinical Data Repository Database (CRDB) and the Tracker Management Database (TMDB). All data is formatted to include a Medical Record Number (MRN), which is an Enterprise Master Patient Index (EMPI) patient identifier, the concept identifier for the clinical element, the time and date when the clinical element was recorded, and any associated value. The majority of the patient data is extracted directly from the RPDR. However, to compensate for the delay in updating of RPDR procedures by up to one month, the tool extracts data from the Clinical Data Repository (CDR).

All CDR data are replicated into the SQL database CRDB. The CDR data model was designed for clinical use rather than for research, therefore not all of the data contains consistent identifiers. For example, Pap smear reports are titled "Cytopathology" and the details that the report is actually a Pap smear are contained somewhere within the report text. To consistently identify reports types, we created the Tracker Concept Mapping Service (TCMS) that uses a combination of values and regular expressions from the report title, report service, and report text to map the report to an RPDR concept identifier. TCMS also buffers reports that are not successfully mapped. For example a report title may be changed by a new hospital system. A programmer reviews these unmapped reports and creates a new entry in the TCMS pattern definition database. Once added, these unmapped reports can then be mapped accordingly using the new TCMS definition. These RPDR and CRDB are used to identify patients in the risk set, R, and the completed management set, M_c.

The query tool then identifies patients who have management planned, M_p. This is important because these patients should not be considered as missing follow-up yet, in an effort to not send an additional reminder during this waiting period. To identify these patients with planned management, we include management plans from the Tracker Management Database. Clinicians define these plans after getting a reminder that a patient is missing follow-up. These data are structure like RPDR data, but in addition to the date the management is planned to occur, they also contain an expiration date. After the planned management passes the expiration date, if it has not been completed, it is removed from the management database, and the patient is no longer considered to have planned management.

4.2.2 Construct SQL From Tracker Logic

After the query tool gathers patient data, the search process begins by reading the logic directly from the guidelines. The Tracker Definition Service (TDS) requests all guideline definitions in XML. For ease of human interpretability, this XML is transformed into a more logically interpretable format (Figure 14) using XSL. Clinical elements that are customized *synonyms* or *panels* are replaced with the nested group of clinical elements represented in the synonym or panel, joined within *or* or *and* logic respectively.

```
<trackers>                                    <trackers>
 <tracker₁>                                    <tracker₁>
  <riskmanagementstratum₁>                      <riskmanagementstratum₁>
   <riskset>                                      <max category="risk">
    <riskgroup₁>                                   <and₁>
     <clinicalelement₁/>                            <clinicalelement₁/>
          and
          <clinicalelement_n/>       →             <clinicalelement_n/>
          and
          <orgroup₁>                                <or₁>
           <clinicalelement₂/>       →              <clinicalelement₂/>
               or
               <clinicalelement_n/>                 <clinicalelement_n/>
          </orgroup₁>                 →           </or₁>
         </riskgroup₁>                            </and₁>
          or
         <riskgroup_n>                            <and_n>

          ...                                      ...
         </riskgroup_n>                           </and_n>
        </riskset>                               </max>
   </riskmanagementstratum₁>                   </riskmanagementstratum₁>
```

Figure 14. Tracker Transformation to Logic Format

The tracker guidelines are converted into a logical format. The risk and management sets search for the maximum group score, thus renamed to *max*. The risk and management groups calculate the sum of clinical elements, thus renamed to *and*. Or-groups are renamed to *or*.

This logical XML is then transformed again using XSL into a large SQL statement (Figure 15). Microsoft SQL Server optimizes the SQL prior to execution; therefore we can repeat SQL segments where needed in this transformation without compromising performance. This is especially useful for it allows us to transform our XML directly into SQL without having to create numerous subqueries.

The query begins by simplifying all clinical elements (CE) to either *true* or *false* (0 or 1) based on their value or date constraints, e.g. either a patient has an LDL > 100 in the past 1 year or not. Each clinical element is actually a collection of all similar concepts. For example, LDL is represented as "Labs\Chemistry\LDL", but this actually represents all 23 variants of LDL coding used in different hospital labs within our healthcare system. Taking advantage of the hierarchical structure of the RPDR metathesaurus, SQL identifies this set of 23 LDL concept codes (C_{LDL}) whose names all begin with "Labs\Chemistry\LDL%" using the "%" wildcard. Therefore, the patients in the "LDL" group all have any one of the 23 concept codes in C_{LDL} with a value > 100mg/dl. This hierarchical search method is one of the key advantages to using the RPDR metadata and data. Furthermore, if a new LDL code is added to a hospital laboratory system, the RPDR updates their metadata to include the new code name as "*Labs\Chemistry\LDL*\newLDLname" and then records the lab results with the new code. Although there is a new name for the LDL lab test, the tracker definition of "*Labs\Chemistry\LDL*" will not need to be changed, because it will still match the beginning of the new code name.

```
SELECT mrn, MAX(score) AS maxscore, type='risk_actual'
group, riskmanagementstratum='Cardiac Risk'
(
  SELECT mrn, COUNT(mrn)/3 AS score, group='Cardiac Labs' FROM
  (
    (SELECT DISTINCT mrn FROM rpdr_data WHERE concept_code IN
      (SELECT concept_code FROM rpdr_metathesaurus
       WHERE concept_name LIKE 'Demographics\Gender\Male%') )
    UNION ALL
    (SELECT DISTINCT mrn FROM rpdr_data WHERE concept_code IN
      (SELECT concept_code FROM rpdr_metathesaurus
       WHERE concept_name LIKE 'Labs\Chemistry\LDL%')
       AND concept_value > 100
       AND DATEDIFF(month, concept_date, '05/16/2005') <= 12 )
    UNION ALL
    (SELECT DISTINCT mrn FROM rpdr_data WHERE concept_code IN
      (SELECT concept_code FROM rpdr_metathesaurus
       WHERE concept_name LIKE 'Labs\Chemistry\HDL%')
       AND concept_value < 30
       AND DATEDIFF(month, concept_date, '05/16/2005') <= 12 )
  )
  UNION ALL
  SELECT mrn, COUNT(mrn)/1 AS score, group='Cardiac Disease' FROM
  (
    SELECT DISTINCT mrn FROM
    (
      (SELECT DISTINCT mrn FROM rpdr_data WHERE concept_code IN
        (SELECT concept_code FROM rpdr_metathesaurus
         WHERE concept_name LIKE
         'Diagnoses\Circulatory\Ischemic heart disease%') )
      UNION ALL
      (SELECT DISTINCT mrn FROM rpdr_data WHERE concept_code IN
        (SELECT concept_code FROM rpdr_metathesaurus
         WHERE concept_name LIKE
         'Diagnoses\Metabolic\Disorders of lipoid metabolism%'))
    )
  )
)
```

Figure 15. SQL Query to Identify and Score Patients

This SQL statement calculates the risk set within a single risk-management stratum. It identifies "Cardiac Risk" patients due to either abnormal "Cardiac Labs" or "Cardiac Disease." All clinical elements (CE) are resolved to *true* or *false*, with a score of 0 or 1. CE_1 is *true* for all men. CE_2 is *true* for all patients with an LDL > 100mg/dl in the past 12 months. $Group_1$ represents *and* logic by counting CE_1, CE_2, and CE_3 and dividing by 3. $Group_2$ represents *or* logic by counting if a patient has either CE_4 or CE_5 by using the DISTINCT command. Lastly, for this risk set of "Cardiac Risk," the maximum score is taken from $group_1$ and $group_2$. When this query is completed, every patient will have a recorded maximum score for this risk set, R. Scores are then calculated for the management set, M (not shown), and the two sets are compared to identify patients without follow-up.

After identifying patient sets who match each clinical element, the query tool calculates risk-management scores (Table 1). For each tracker and risk-management stratum, the tool calculates the maximum scores of the risk set and management set using the SQL command MAX on the group scores within each set. To calculate these group scores, the query tool counts the number of *matching* clinical elements divided by the number of *total* clinical elements, e.g. 2 out of 4 matches yields a group score of 0.50 for that patient. Within or-groups the tool utilizes UNION ALL which combines all clinical elements into one set; then it extracts the set of unique patients. Therefore, for all patients with CE_1 *or* CE_2, extracting the unique patients from this set yields a single value for each patient, consistent with *or* logic. We use similar nested SQL commands iteratively to represent nested *and* or *or* groups.

28

The management set scores are calculated with similar SQL methodology. The key difference is that while *risk* clinical elements are identified relative to the current date, the *management* clinical elements are identified relative to the most recent date a patient became a member of the associated risk set. This is done in SQL by a slight modification to the DATEDIFF function. For example the *risk* clinical element to identify patients with a concept within 12 months of 05/16/2005 would be coded as:

```
... DATEDIFF(month, concept_date, '05/16/2005') <= 12
```

and stored in temporary memory. Using these results, we can identify patients with a *management* clinical element that occurred within 3 months of this *risk* clinical element by:

```
... DATEDIFF(month, concept_date,
  (SELECT concept_date FROM riskset WHERE mrn = rpdr_data.mrn)
) <= 3
```

4.2.3 Planned Versus Completed Follow-up Scores

Once we have identified the different sets of patients for all trackers and risk-management strata, we create summary scores for patients in need of follow-up. Using SQL aggregate functions, we gather the set of patients with maximum risk names and scores (R), as well as the set of patients with maximum management names and scores (M_c), for all trackers and all risk-management strata within the trackers. These sets are generated from *actual* patient data. The same SQL query process is now repeated with the original risk set of patients (R), but now using both the *actual* management data (M_c) and *planned* management data (M_p). From these 3 sets (R, M_c, M_p), we generate the three follow-up sets F_c (R ∩ M_c), F_p ([R - M_c] ∩ M_p), and F_{np} (R - M_c - M_p).

The most important set of patients to track are in F_{np}; these are the patients who are at-risk, yet have no completed management nor planned management. In terms of scores, these patients have a risk set score of 1.0 (have 100% of the risk factors) and do *not* have a management set score of 1.0 (have *not* had 100% of the management completed). We can emulate fuzzy logic by adjusting the risk threshold score to less than 1.0. For example, we may choose to increase the sensitivity of our patient capture by considering patients at-risk as having greater than 75% of the risk factors.

row	MRN	tracker	risk-management stratum	risk date	management date	cycle
1	000004	hyperlipidemia	high cardiac risk	10/11/2005	--	1
2	112358	diabetes	high risk	11/07/2004	*05/07/2005*	1
3	112358	heart disease	high risk	02/09/2004	02/29/2004	1
4	000004	cervical cancer	abnormal	04/12/2004	--	2
5	000004	cervical cancer	routine screening	03/05/2003	04/12/2004	2

Table 3. Tracker Watchlist

The culmination of this entire query process is summarized in the Tracker Watchlist (Table 3). When a patient is first identified as have no follow-up plan (F_{np}), this information is recorded without a management date (row 1). As clinicians choose management plans using the Tracker Clinical Tool, the patients move from F_{np} to F_p, and they are recorded with a pending management date in italics (row 2). If the planned management is completed, the patient moves from F_p to F_c, and the actual management completion date is recorded (row 3). However, if that planned management is *not* completed, the pending management will eventually expire, and the patient will move from F_p back to F_{np}. This is reflected by the cleared management date and the cycle incremented by 1 (row 4). If these patients with planned, yet missed follow-up finally do receive their planned management, the date of completion will be recorded (row 5). Otherwise, the planned management date will expire again, the cycle will increment again by 1, and the process begins again.

It is significant to note that this query process produces *all* scores of *all* trackers using *all* 450 million clinical elements for *all* 1.8 million patients in the RPDR in single batch query. This process is run daily and takes on the order of minutes to hours, depending on the number of tracker guidelines.

reminders are sent to clinicians who choose follow-up management
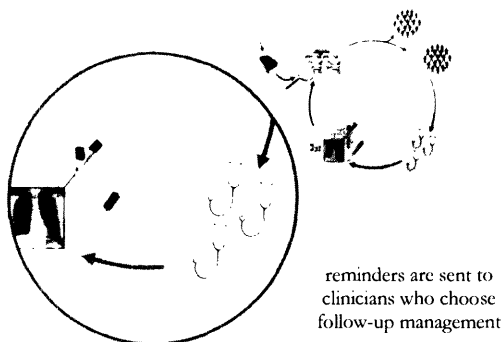
4.3 Tracker Clinical Tools

After the query tool searches and summarizes the patients' clinical records, the Tracker Clinical Tools (TCT) sends this information to the clinicians, who can then act upon it. The typical format of such a message is "Dr. D: patient P is at risk for disease D yet has not had follow-up" or "Dr. D: you had planned follow-up for patient P regarding disease D, yet after M months, it has not happened."

The goals for the clinical tools are to 1) present to providers their patients without follow-up, 2) enable simple selection of follow-up management, 3) track this follow-up management, and 4) remind clinicians if their patients miss follow-up. Third party programs use the results of the tracking query tool to send and display reminder messages. Multiple interfaces can utilize our common Application Programming Interface (API) to access these tracking data. We describe one such interface developed for the Partners OnCall Electronic Health Record used within the Massachusetts General Hospital Internal Medicine Associates outpatient clinic.

4.3.1 Clinical Services

Three Internet-based clinical information services access the tracking data: 1) the Tracker Watchlist Service, 2) the Tracker Definition Service, and 3) the Tracker Management Service. These services have a simple command-line interface consisting of a Universal Resource Locator (URL) that receives input parameters and returns tracking data in XML.

The Tracker Watchlist Service (TWS) provides either individual or group information from the watchlist. For example, an application used during a clinic visit may request all trackers without complete follow-up for an individual patient (Figure 16). Alternatively, a population manager application may request trackers for all patients under the care of a given clinician or with a certain disease process.

TrackerWatchlistService.asp?mrn=000004&followup=none

```
<trackerwatchlist mrn="000004" followup="none">
  <tracker name="hyperlipidemia" riskmanagementstratum="high cardiac risk"
    riskdate="10/11/2005" managementdate="" cycle="1" id="274"/>
  <tracker name="cervical cancer" riskmanagementstratum="abnormal"
    riskdate="04/12/2004" managementdate="" cycle="2" id="214"/>
</trackerwatchlist>
```

Figure 16. Tracker Watchlist Service
This example provides watchlist information from the patient with MRN of 000004 without follow-up: 1) no follow-up for hyperlipidemia and 2) missed follow-up once (cycle="2") for cervical cancer.

30

The Tracker Definition Service (TDS) is used to request additional tracker guideline definitions. This information can be displayed with watchlist data, to help the clinician understand the details of the individual trackers. Using the relational database structure, all elements within the tracker model have an associated ID that can be used to gather additional information. For example, the third party may display the risk logic associated with that tracker watchlist entry, or perhaps display the follow-up management options available for that tracker (Figure 17).

TrackerDefinitionService.asp?level=riskmanagementstatum&id=214

```
<trackerdefinition>
  <tracker name="cervical cancer">
    <riskmanagementstratum name="abnormal" id="214">
      <riskset>
        <riskgroup>
          <clinicalelement name="pap smear"
          fullname="Procedures\Female GU\Screening Papanicolaou Smear"/>
        </riskgroup>
      </riskset>
      <managementset>
        <managementgroup>
          <orgroup>
            <clinicalelement name="pap smear"
            fullname="Procedures\Female GU\Screening Papanicolaou Smear"/>
            <clinicalelement name="colposcopy"
            fullname="Procedures\Female GU\Colposcopy"/>
            <clinicalelement name="HPV" fullname="Laboratory\DNA\HPV"/>
          </orgroup>
        </managementgroup>
        <managementgroup name="gynecology consultation">
          <clinicalelement name="consultation"
          fullname="Procedures\Evaluation\Consultations"/>
          <orgroup>
            <clinicalelement name="Gynecology" fullname="Encounter\Gynecology"/>
            <clinicalelement name="Gyn Oncology" fullname="Encounter\Gyn Onc"/>
            <clinicalelement name="Ob-Gyn" fullname="Encounter\Ob-Gyn"/>
          </orgroup>
        </managementgroup>
      </managementset>
    </riskmanagementstratum>
  </tracker>
</trackerdefinition>
```

Figure 17. Tracker Definition Service

This example displays the guideline risk factors and management options for cervical cancer. In this example, patient 000004 became at-risk by having a Pap smear without documented follow-up. The clinician chooses from the management options shown: repeat Pap smear, HPV test, colposcopy, or gynecology consultation.

Combining data from the watchlist and definition services, the clinician can view those patients without follow-up and choose the appropriate follow-up. These management choices are recorded into the Tracker Management Database using the Tracker Management Service (Figure 18).

```
<trackermanagement mrn="000004" provider="0112358">
  <clinicalelement name="HPV" fullname="Laboratory\DNA\HPV"
  alertType="priority" alertMaxNum="7" alertMaxUnit="day"/>
</trackermanagement>
```

Figure 18. Tracker Management Service

Provider 0112358 decided that the Pap result for patient 000004 was abnormal and chose to follow-up with an HPV lab test. If the result had been normal, the provider would likely have chosen a routine screening Pap smear to be completed within 1 to 3 years, depending on other patient risk factors. .
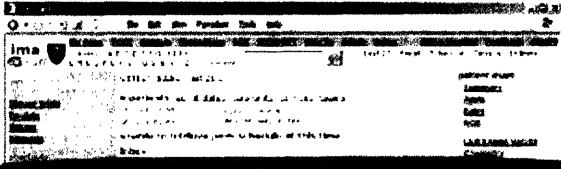
These three services serve as the Application Programming Interface (API) that gives developers the tools to utilize the Tracker System.

4.3.2 Inbox Application

The Inbox Instant Messaging Application (Inbox) is a third party program used to communicate messages between clinicians using the OnCall EHR. Utilizing these three services of the Tracker API, we developed the Tracker Clinical Tool (TCT) that displays reminder messages within the Inbox. The Tracker Clinical Tool is just one example of an interface. The TCT uses the watchlist service to send reminder messages to clinicians of patients who have not had follow-up. The messages are displayed to the clinician in OnCall in DHTML. When a clinician clicks on a tracker message, the management options are displayed in a separate window using Flash. The clinician then chooses the management plan, and the results tracking process begins.

4.3.2.1 DHTML Component

The clinical tool displays the tracker watchlist within the OnCall EHR clinician summary page (Figure 19). Utilizing JavaScript functionality, the clinician can sort these messages dynamically by clicking on the column title. By presenting this watchlist on the clinician summary page, a clinician does not need to open a specific patient's electronic chart in order to view a tracker message. This design choice facilitates asynchronous care not dependent on a patient visit, but rather on a new or delinquent clinical result found by the tracker query tool. When a clinician opens an individual patient's electronic chart, only tracker messages that pertain to that patient are displayed.

Figure 19. Tracker Messages Within The OnCall Health Record
The clinician views all patients on the watchlist who need follow-up. Each message contains the tracker name and the respective risk-management stratum for each patient. An exclamation point indicates patients who had planned follow-up but never completed the full management plan, thus "lost" to follow-up.

4.3.2.2 Flash Component

To plan management for a patient who needs follow-up, the clinician clicks the tracker message in the DHTML OnCall component. The tracker is opened in separate window using the Inbox Flash object (Figure 20). The actual message itself contains minimal information: patient name, MRN, provider, date, and a tracker watchlist reference number. The Flash component uses this message information to gather supporting data to display to the clinician: 1) the tracker definition service for the tracker guideline details and 2) the RPDR service to gather the actual patient clinical data for this tracker.
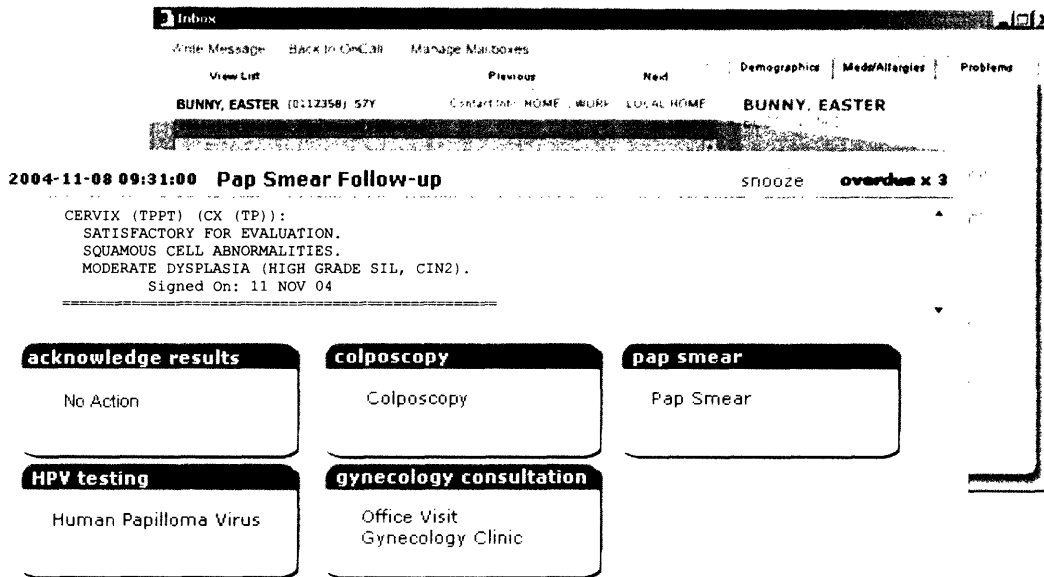
Figure 20. Tracker Management in Inbox Flash
For this cervical cancer screening tracker message, the current abnormal Pap smear report and previous reports are displayed in the risk results window in red. The patient's follow-up status, in the upper right, indicates that the patient has missed 3 attempts at follow-up. Below these reports, the menu displays the management groups. In this example, the clinician chose to follow-up the abnormal Pap smear with HPV testing. The clinician may have also chosen to "snooze" this message, if the results required follow-up at a later date, like repeating the Pap smear in 3 months. Other patient information, such as demographics, medications, allergies, and the problem list are displayed as a separate part of the Inbox Application on the right.

The risk results window displays information about the result that need follow-up along with the current follow-up status, initially "new." However, if follow-up has been planned and missed, and a delinquent message alert is sent, the status becomes "overdue" along with the number of times follow-up has been missed. The background color of the message changes with the urgency of the message: a new result is white, planned follow-up is yellow, overdue follow-up is red, and completed follow-up is green.

To begin follow-up tracking, the clinician selects from the menu of management groups. Groups may contain one or more management options. Although the groups are designed for easy one-click management, the clinician may choose multiple actions from any group. Default alert settings (routine, priority, or none) are used unless the clinician chooses to define different alert timing. In addition to these management choices, all tracker management options also include an additional "acknowledge results" group with one management clinical element of "no action". If the clinician decides that the results do not need follow-up tracking, then "no action" removes this patient tracker from the watchlist by including "no action" as the completed management. These management choices are stored in the Tracker Management Database using the Tracker Management Service.

However, not all results need immediate follow-up. For example, if an abnormal LDL requires a repeat LDL in 6 months, or an ASCUS Pap smear requires a repeat in 3 months, the clinician may choose "snooze" and select the snooze time. This process will record "no action" as the follow-up management, but with an expiration date equal to the snooze time. The only difference is that when the patient moves from F_p to F_{np} when the "no action" snooze management expires, the cycle is not incremented by 1, because this does not represent delinquent follow-up. Rather the exact message is resent. At that time later time, the clinician then chooses the appropriate follow-up management. Choosing follow-up restarts the tracking loop, and the process begins again.

33

5 Discussion

Our goal was to identify and track patients without follow-up. Our design and implementation have made improvements on previous reminder systems, namely in our risk-management guideline model, our use of population-based queries, and our longitudinal approach that records patient movements through the follow-up cycle.

The risk-management model allows clinicians to bypass redundant clinical decision support and enter the guideline directly at the risk-management level, usually a choice between two or three groups. This saves the clinician time interacting with the guideline and keeps the clinician in full control of the decisions. Once the risk-management stratum is chosen, we can then present the handful of possible follow-up choices. We do not suggest which follow-up is appropriate, but instead make all choices available and let the clinician decide. We present context-specific management choices that pertain to both a) the results and b) the patient's other risk factors.

Population-based queries support a tracking system on a massive scale. We are able to search 450 million pieces of clinical data for 1.8 million patients daily. When new clinical information becomes available, rather than modifying our guidelines, we simply add those new term definitions to our metathesaurus and the new results to our database. Furthermore, other health care systems can utilize our tracking system, by mapping their clinical data to UMLS terms and storing them in a SQL database.

Lastly, our longitudinal approach tracks patients and their follow-up trends over time. Not only do we know which patients are currently without follow-up, but we also have documented all information regarding steps to ensure their follow-up, such as when they became at risk, whether follow-up plans were documented, and how many times the patient has missed follow-up.

Although our tracking system is robust, it does have several limitations. Identifying the patients without follow-up will not necessarily change the follow-up rates. It is important to note that patients can be lost to follow-up due to a number of complicated issues. In the case of patient non-compliance, our tracking system is of little use to encourage compliance. However, we are able to identify those patients. This serves three purposes. First, although reminders are sent to clinicians, they can also be sent to case managers, social workers, or even directly to the patients via a letter. Secondly, although the clinician may not be able to control the non-compliance issues, it is important to document this non-compliance in the chart for medical-legal purposes. Lastly, if a patient receives follow-up care outside of our system, this patient will appear to have not had follow-up, because the follow-up results will not be in our system. Although the clinician will then receive an erroneous reminder, this is also an opportunity to document that follow-up care was received elsewhere, and close the follow-up loop.

Another limitation to our tracking system is our reliance on clinical data. Although we have demonstrated that clinicians can enter a guideline at the risk-management stratum and view context specific follow-up options, an intelligent system operates more smoothly with a large supply of patient data. For example, without clinical data for risk factors, the tracker would not identify patients at-risk, therefore would not perceive these patients to be lacking follow-up. Furthermore without clinical data for management follow-up, the tracker would not recognize when follow-up has occurred once a tracker has been started. Both of these scenarios limit the intelligence of our tracking system.

However, we believe our same tracking system can function in a clinical system with sparse patient data. This would require the clinician to be a more active participant with the tracker. For example, if a woman was being seen for her annual Pap smear, the clinician could click on "cervical cancer screening" from a list of all available trackers, then jump right to the risk-management stratum "routine screening" with one

34

click. Clicking on "Pap smear" would set the tracking system listening for the Pap smear results. Therefore we are able to display and track all follow-up options with very little information provided by the clinician. However, without clinical results coming into our system, the clinician would only get a reminder a month later that the patient "missed" follow-up, when, in fact, follow-up may have already occurred, just not reported electronically to our tracking system. At this point, the clinician would open the reminder message and click "follow-up complete, no further action" or alternatively, an additional follow-up test. In this scenario, although not powered by clinical data, our system would provide an opportunity for clinicians to easily set trackers and be reminded to follow-up on the results between patient visits.

Two other data constraints limit our system. First, despite being able to track follow-up plans, our tracking system does not actually order the follow-up; rather, after the clinician selects the follow-up, orders need to be placed. This redundancy of action may discourage clinicians from using the tracking system. However, we feel the peace of mind provided by our tracking will outweigh this work-flow inconvenience. Selecting a management option is a one-click process done during results review, a time already requiring the clinician's attention. Furthermore, reminders are not sent to patients who receive timely follow-up. Therefore, the reminders sent would be specific to missed follow-up and would be perceived as less intrusive.

Lastly, the majority of our data is read from the RPDR, an analytic research database that is updated once daily for labs, but the lag can be up to a month for procedures. While this lag prevents our system from being real-time, we are comfortable with the lag for outpatient care. For example, when tracking labs in the outpatient setting, clinicians usually do not need reminders of missing labs one the granularity of less than one day. Therefore the one day lag for labs is acceptable. Similarly, procedure results are usually not expected within an interval of less than one month. We feel clinicians will tolerate a larger lag for procedures. However, to compensate for this lag, we supplement the RPDR data with real-time CDR data for the most common procedures. Since the CDR is limited and does not contain administrative claims data like the RPDR, the combination of the two provides the most thorough clinical data for closing the follow-up loop.

6 References

1. Committee on Quality of Health Care in America, Institute of Medicine, *Crossing the Quality Chasm: A New Health System for the 21st Century Health Care Services*. 2001: Washington, DC.
2. Kottke, T.E., M.L. Brekke, and L.I. Solberg, *Making "time" for preventive services*. Mayo Clin Proc, 1993. **68**(8): p. 785-91.
3. Hibble, A., et al., *Guidelines in general practice: the new Tower of Babel?* Bmj, 1998. **317**(7162): p. 862-3.
4. Yabroff, K.R., et al., *Is the promise of cancer-screening programs being compromised? Quality of follow-up care after abnormal screening results*. Med Care Res Rev, 2003. **60**(3): p. 294-331.
5. Garg, A.X., et al., *Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review*. Jama, 2005. **293**(10): p. 1223-38.
6. Lester, W.T., et al., *Facilitated lipid management using interactive e-mail: preliminary results of a randomized controlled trial*. Medinfo, 2004. **11**(Pt 1): p. 232-6.
7. Murff, H.J., et al., *Primary care physician attitudes concerning follow-up of abnormal test results and ambulatory decision support systems*. Int J Med Inform, 2003. **71**(2-3): p. 137-49.
8. National Center for Health Statistics, *Healthy People 2000 Final Review*. 2001, Hyattsville, Maryland: Public Health Service.
9. U.S. Department of Health and Human Services, *Healthy People 2010: Understanding and Improving Health*. 2nd ed. 2000, Washington, DC: U.S. Government Printing Office.
10. McCarthy, B.D., et al., *Inadequate follow-up of abnormal mammograms*. Am J Prev Med, 1996. **12**(4): p. 282-8.
11. Chang, S.W., et al., *Racial differences in timeliness of follow-up after abnormal screening mammography*. Cancer, 1996. **78**(7): p. 1395-402.
12. Goodman, K.A., R.L. Birdwell, and D.M. Ikeda, *Compliance with recommended follow-up after percutaneous breast core biopsy*. AJR Am J Roentgenol, 1998. **170**(1): p. 89-92.
13. Michielutte, R., et al., *Noncompliance in screening follow-up among family planning clinic patients with cervical dysplasia*. Prev Med, 1985. **14**(2): p. 248-58.
14. Marcus, A.C., et al., *Improving adherence to screening follow-up among women with abnormal Pap smears: results from a large clinic-based trial of three intervention strategies*. Med Care, 1992. **30**(3): p. 216-30.
15. Janerich, D.T., et al., *The screening histories of women with invasive cervical cancer, Connecticut*. Am J Public Health, 1995. **85**(6): p. 791-4.
16. Mandel, J.S., et al., *Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study*. N Engl J Med, 1993. **328**(19): p. 1365-71.
17. Kronborg, O., et al., *Randomised study of screening for colorectal cancer with faecal-occult-blood test*. Lancet, 1996. **348**(9040): p. 1467-71.
18. Jemal, A., et al., *Cancer statistics, 2004*. CA Cancer J Clin, 2004. **54**(1): p. 8-29.
19. Melnikow J, et al., *Natural history of cervical squamous intraepithelial lesions: a meta-analysis*. Obstet Gynecol, 1998. **92**(4 Pt 2): p. 727-35.
20. Sirovich, B.E. and H.G. Welch, *The frequency of Pap smear screening in the United States*. J Gen Intern Med, 2004. **19**(3): p. 243-50.
21. Sung, H.Y., et al., *Papanicolaou smear history and diagnosis of invasive cervical carcinoma among members of a large prepaid health plan*. Cancer, 2000. **88**(10): p. 2283-9.
22. Bastani, R., *Interventions to Improve Follow-Up of Abnormal Findings in Cancer Screening*. Cancer, 2004. **101**(5 Suppl): p. 1188-1200.
23. Sackett, D.L., et al., *Evidence based medicine: what it is and what it isn't*. Bmj, 1996. **312**(7023): p. 71-2.

24. Peleg, M., et al., *The InterMed approach to sharable computer-interpretable guidelines: a review Human factors barriers to the effective use of ten HIV clinical reminders.* J Am Med Inform Assoc, 2004. **11**(1): p. 1-10.

25. Peleg, M., et al., *Comparing computer-interpretable guideline models: a case-study approach.* J Am Med Inform Assoc, 2003. **10**(1): p. 52-68.

26. Stone, E.G., et al., *Interventions that increase use of adult immunization and cancer screening services: a meta-analysis.* Ann Intern Med, 2002. **136**(9): p. 641-51.

27. Morgan, M.M., J. Goodson, and G.O. Barnett, *Long-term changes in compliance with clinical guidelines through computer-based reminders.* Proc AMIA Symp, 1998: p. 493-7.

28. Calabrisi, R.R., T. Czarnecki, and C. Blank, *The impact of clinical reminders and alerts on health screenings. The VA Pittsburgh Healthcare System achieves notable results by enhancing an automated clinical reminder system within its CPR--and has the data to prove it.* Health Manag Technol, 2002. **23**(12): p. 32-4.

29. Bodenheimer, T., E.H. Wagner, and K. Grumbach, *Improving primary care for patients with chronic illness.* JAMA, 2002. **288**(14): p. 1775-9.

30. Maviglia, S.M., et al., *Automating complex guidelines for chronic disease: lessons learned.* J Am Med Inform Assoc, 2003. **10**(2): p. 154-65.

31. Murphy, S.N., G.O. Barnett, and H.C. Chueh, *Visual query tool for finding patient cohorts from a clinical data warehouse of the partners HealthCare system.* Proc AMIA Symp, 2000: p. 1174.