The ChronoMedIt Temporal Medical Audit Framework: Progress and Agenda

Jim Warren, Thusitha Mabotuwana, Hong Yul Yang
Department of Computer Science / National Institute for Health Innovation
The University of Auckland
Tamaki Campus, Private Bag 92019, Auckland 1142 New Zealand
(jim,thusitha,hongyul}@cs.auckland.ac.nz

Abstract
Chronic condition management presents distinctive challenges for health information systems. ChronoMedIt is a computational framework for the analysis of electronic medical records (EMRs) to support audit of chronic conditional management. This paper describes progress to date, recent developments and research directions related to the extension and application of ChronoMedIt. We address application of the framework to epidemiology, to pay-for-performance schemes, and to support clinical intervention. We also describe enhancement to the ChronoMedIt architecture to make it more flexible.

Keywords: chronic disease management, clinical information systems, decision support, quality improvement, temporal analysis

1 Introduction
According to the World Health Organization, chronic disease was the leading cause of death worldwide in 2005 accounting for around 35 million (or 60%) of the 58 million deaths (WHO 2005). It has been projected that in 2015, 41 million (or 64%) out of all 64 million deaths will be due to chronic disease (Strong et al. 2005). In New Zealand (NZ), chronic disease is the leading cause for hospitalisation and accounts for 70% of health expenditure and 80% of all deaths (National Health Committee 2005). Thus, the burden of chronic disease is a growing concern worldwide and any incremental improvement in the management of chronic disease will have immense population health benefits, as well as cost savings and relief on the demands straining our health workforce.

To focus efforts at improving a process, it is vital to know how one is doing in the first place. A key concept here is that of quality audit indicators which are specific measures used to give an indication of the quality of the patient care delivered (Rosenbrand, Van Croonenborg and Wittenberg 2008). Quality audit indicators may serve to compare the relative merit of practice at one time period to another time period for a given healthcare delivery organisation – i.e., to track the trend in quality over time. For instance, we may hypothesize an upward trend in response to some newly introduced quality improvement effort, such as installation of electronic decision support (EDS) tools. We may also use quality audit indicators to compare between healthcare delivery organisations – e.g., to assess the relative performance of organisations using a particular EDS tool as compared to those not using it.

Chronic disease management, by its nature, introduces technical challenges associated with temporal reasoning, assessing time intervals and their relationships. Moreover, the input data from healthcare information systems is typically in terms of date-stamped events that only imply relevant intervals. For instance, a chronic condition diagnosis – say a diagnosis of diabetes mellitus type 2, or of essential hypertension – is made on a given date, but the implication is that the condition has existed from some (unknown) time earlier and that it will persist more-of-less for the life of the patient. Similarly, the most commonly recorded information about a medication is a prescription (and/or a dispensing of medication to fill a prescription). The prescription implies that the prescriber feels the medication is indicated for the patient. Its dosage instruction indicate how often it will be taken and how much at each administration (the signatura or ‘sigs’, which often include Latin frequency indications such as bid for twice daily). How long the patient is expected to take the medication is often only implied. This might be indicated at one level by the coverage of the prescription in terms of how much medication it can supply. A prescription for a bottle of 60 pills, with a one tablet bid sig and allowing two refills gives 90 days supply. Somewhat less temporal ambiguity is associated with observations such as a blood pressure, which represents a measurement on a patient at a given point in time. Although even here there is the assumption that the reading is representative of the temporal neighbourhood, particularly of all time until a newer reading is made.

In this paper we describe progress to date and ongoing efforts with respect to a computational framework that we have developed to compute quality audit indicators automatically from the data in health information systems (i.e., from electronic medical records, EMRs). The framework is called ChronoMedIt for Chronological (i.e., temporal) Medical Audit. ChronoMedIt is designed to leverage EMRs to support improvement in chronic condition management. The next section introduces the tools itself. The subsequent three sections describe its role in various areas of application. We then describe our efforts to improve the flexibility of the architecture and give a concluding summary.
2 The ChronoMedIt Framework

The ChronoMedIt software was developed as part of the PhD thesis of the second author under the supervision of the first author. It extends earlier work by the first author on the use of EMRs in general practice medicine to provide critique of management of hypertension in terms of a temporal state-transition model. The temporal state is the value at a given time of a set of monitored state variables (basically the set of medications prescribed, but possibly also relevant complicating diagnoses – such as diabetes, which complicates the treatment of hypertension). A temporal state-transition occurs when the state changes – e.g., a new class of medication is prescribed – with the theory that the moment the temporal state changes (transition) is a good moment to critique whether that change fits with clinical practice guidelines. An EDS structure is then formulated as a set of specific clinical alerts that are each attached to particular state transition. It was found that such a framework could give high-specificity (low ‘false alarm’ rate) alerts for antihypertensive prescribing (Gadzhanova et al. 2007).

The temporal state transition model was conceived as a way to provide alerts, i.e. immediate interaction feedback to cause a physician to reflect on the management of the patient in front of them. We reformulated the use of the model to quality audit more generally as a temporal audit reporting (TAR) model (Gaikwad, Warren and Kenealy 2007) that would indicate which, and how many, patients would receive a particular alert (i.e., failed a particular quality criterion) from among a given patient cohort on a time period of interest (hereafter, the evaluation period). Thus, one quality audit item in a TAR might indicate how many patients in a particular general medical practice were diagnosed with both hypertension and diabetes, and thus indicated for a particular class of antihypertensive called Angiotensin Converting Enzyme inhibitors (ACEi, see relevant clinical practice guidelines; Chobanian et al. 2003) but ran out of supply of it. In terms of the model, these are patients that made a temporal state transition from a having-ACEi state to not-having-ACEi state during the evaluation period.

The genesis of ChronoMedIt was a study wherein we worked closely with a New Zealand general medical practice that served a largely Pacific Islander case load. We developed a set of eight quality audit criteria that suited their needs and demonstrated that these could be derived reasonably accurately (in fact we observed 70% sensitivity and 70% specificity) from their EMR data (Warren et al. 2008). This quality audit report was based on a combination of Structure Query Language (SQL) and procedural code custom to their eight criteria. ChronoMedIt was then formulated to provide a framework for answering any such queries that might arise based on abstraction of the properties of those eight.

The resultant criteria model is shown in figure 1. Any given quality audit criterion is represented as an instance of the criteria model with particular attribute values. The top level of the model is the properties that define the evaluation period and the population of interest (e.g., what ‘classifications’ – basically clinical diagnoses – are characteristic of the group of interest). The model then branches into two major subtypes: 1. lapses in supply of medication indicated for the population of interest; and 2. a set of subtypes related to measurements on the patient. This latter type breaks down into: (a) failure to record the outcome measure of a management activity; (b) presence of a measured value that contraindicates a therapy; and (c) sustained instances of the outcome measure being outside of acceptable bounds. We do not maintain that these are all possible quality audit criteria of interest (see section 6 for our thoughts on additional ones), merely that these are a wide family of criteria that are of interest and that circumscribe those eight that were of interest to the practice with which we were collaborating.

To make use of the criteria model required several system components:

1. A data uptake component to import data from a commercial EMR to the particular concepts of interest to our framework, including diagnostic classifications, time periods covered by medication supply (as indicated by prescriptions, or alternatively by dispensing data), observations (notably blood pressure readings) and laboratory test results. This requires interpreting medication signs and also ‘scraping’ observations, such as blood pressure, out of free-text clinical notes.

2. A domain ontology to represent the values of the above concepts. This is largely a hierarchy with leaves that are the terms used in the EMR for particular diagnoses and medications organised into meaningful subtypes. For instance, sometimes we may be interested in Antihypertensive (AHT) medication generally, or

![Figure 1: Criteria model for chronic condition management quality audit.](image-url)
we may be interested in a sub-types (such as ACEi). As it turns out, ACEi often has the side-effect of a persistent cough, in which case another AHT class called Angiotensin Receptor Blockers (ARBs) can be substituted. Thus, we find it useful to have a mid-level concept, ACEi/ARB, to use in quality audit. Because some drug products are a combination of agents – e.g., a mixed thiazide diuretic (another AHT class) with ACEi – the ontology is not a strict hierarchy. We represent this ontology as a Web Ontology Language (OWL, McGuinness and Harmelen 2004) file and manage it with the Protégé tool (see http://protegewiki.stanford.edu/wiki/Main_Page).

3. A criteria authoring interface. We have an XML specification for criteria and a graphical user interface that interacts with the ontology to allow a user to select concepts and data value boundaries that define their quality audit requirements.

4. A query processor. The methods to assess which patients in a given EMR extract satisfy the criteria. This uses SQL with supporting stored procedures and has been validated both in terms of real-world face validity and extensive simulation testing based on comparison to results from non-SQL algorithms.

5. Data viewers. This allows the user to see the summary quality levels attained, or to see specific cases for direct follow-up. Details of the cases can be viewed as textual descriptions or as visual timeline graphs (see figure 2).

The ChronoMedIt framework has been described in detail (Mabotuwana and Warren 2010).

3 ChronoMedIt for Epidemiology

One use of ChronoMedIt is as a research tool to understand the patterns of chronic illness and how quality audit indicators align to health outcomes.

We had found there was a great deal of clinical interest in the problem of lapses in medication supply. This led us to take an interest in Medication Possession Ratio (MPR) as a measure of medication adherence (how well patients do what clinical people recommend). MPR is a ‘percent of days covered’ model that assesses how many days in a time period a patient had supply of their medication (Andrade et al. 2006). An MPR<80% is frequently considered as indicating a clinically significant adherence problem.

Because of the configuration of the New Zealand and Australian healthcare systems, general practitioners (GPs) have easy access to their own EMR data. This data indicates prescribing, but does not indicate dispensing. Dispensing on the other hand is best accessed in these healthcare systems via national claims reimbursement databases. These are fairly comprehensive since many medications such as AHT medications are subsidized and thus thoroughly represented in the national data. However, such collections are not designed to be timely or convenient for a local quality improvement effort (as might be initiated by a GP). This led us to investigate how well an MPR from prescribing aligned to one from national claims dispensing.

The answer turns out to be that they align reasonably well, at least for six major long-term medications that we investigated (Mabotuwana, Warren, Harrison and Kenealy 2009). For purposes of this paper, the particular answer is not so much to the point. The point is that the framework can be used to investigate questions about the

Figure 2: Timeline graph for a patient case showing a large gap in 2-agent antihypertensive therapy, an associated gap in blood pressure measurement and persistently uncontrolled blood pressure levels.
alignment of data sources. So if the GP asks, “Do I have to wait for a national e-pharmacy network, or make a request to query the national claims database in order to find patients with low MPRs?” we are able to answer that a reasonably reliable surrogate is available by looking at the EMRs from his (or her) own practice management system (PMS).

A further finding from this same study was about the alarmingly high rates of low MPR, with only half of patients achieving good adherence (MPR=80%) on all of the six sentinel long-term medications that they were prescribed for a 15-month evaluation period and rates of good adherence to individual medications ranging from 55% to 68% (some medications appear to be more amenable to adherence than others). This class of finding is useful for characterising the prevalence of a quality issue.

A further kind of investigation is where we take the cases found through quality audit criteria and compare outcome data from those that are assessed as in violation as compared to those that are compliant. One study we did in this vein combined two of our medication supply criteria – 30 day lapse in supply, and MPR<80% – and took good adherence as having neither of these deficits present during the evaluation period. For a practice population cohort with diabetes and hypertension, good AHT adherence by this measure was associated with triple the odds of successful blood pressure control to recommended levels (Mabotuwana, Warren and Kennelly 2009). With the dangers of poor blood pressure control already well-established, this formulated a link from MPR (and, in fact, MPR based on prescribing, not dispensing) to health outcome.

Studies such as these can help to rationalise the prioritisation of health improvement efforts. Not only are health dollars limited, but the capacity of the health workforce to absorb procedural change and take on additional tasks is limited. Thus we want to establish that a problem is prevalent and important before trying to formulate and implement a relevant quality improvement strategy. In the case of medication adherence, we believe the problem to be under-prioritised and aim to continue to create more results indicating massive medication undersupply and lapse issues and their correlation to poor health outcome.

One area that is promising for further investigation in this regard is statin use – these are medication to control blood cholesterol levels. They serve both to prevent an initial cardiac disease event as well as being indicated for use by everyone who has had a previous event (e.g., a ‘heart attack’) to reduce the odds of another. These two cases divide into ‘primary’ and ‘secondary’ prevention. We plan to use national reimbursement data to characterise the differences in lapse in these two cases. The clinical intuition is that 6 months to a year is a danger time for adherence to statins in secondary use (i.e., it is easy to get people to adherence just after an event, but then they ‘fall off the wagon’). We will use ChronoMedIt to assess the distribution of time until first 30-day-lapse and provide a quantitative basis for where to most urgently target statin adherence promotion interventions.

### 4 ChronoMedIt and Pay-for-Performance

A further domain of application for ChronoMedIt is potentially in performance assessment. One particular management strategy is the notion of ‘pay-for-performance’ – i.e., to provide financial reward for meeting quality audit targets (or, to take a ‘glass half empty’ view, this may be seen as financial penalties for failure to meet targets). This has been done on a massive scale in the UK through the English National Health Service’s Quality and Outcomes Framework (QOF, BMA and NHS Employers 2009).

The QOF was first introduced in April 2004 as part of the General Medical Services contract. It provides a set of clinical indicators across four domains (clinical, organisational, additional services and patient experience) designed around best practice to improve the quality of service provided to patients. Each indicator is allocated a number of points and GP practices are awarded points according to how well they have performed, with associated monetary compensation. The indicators are updated annually.

Most of the widely used quality indicators use the presence of a single point-in-time measurement to determine whether a given indicator is satisfied. For example, an important QOF indicator related to the ongoing management of patients with hypertension, the one with the highest point allocation (57 QOF points) in the ‘clinical’ domain, is ‘BP5’ – “The percentage of patients with hypertension in whom the last blood pressure (measured in the previous nine months) is 150/90 or less.” It should be unsurprising to the reader to hear that the QOF is controversial. Our first attempts to participate in the debate have been to examine point-in-time QOF measures such as BP5 and look at whether patients satisfying BP5 also satisfy quality measures that assess whole time intervals, such as those based on MPR, and are free from measurement violations, such as repeated persistently high blood pressure over a time interval or long periods of missing blood pressure measurement. Using New Zealand data (which – and this could be seen as an advantage or a disadvantage – is not subject to QOF incentives) we find that there are many cases where the last blood pressure satisfied BP5, but the ‘journey’ was suboptimal. We find long periods without measurement to be more the problem than persistently high measures (perhaps if the patient were measured they would have measured high) (Mabotuwana et al. 2010). At any rate, it supports an argument that a system incentivizing just the last measure, say annually, coming good leaves ample room for suboptimal management when defined in more depth over the entire time period.

More broadly speaking, this illustrates use of ChronoMedIt to inform the debate over what performance indicators to incentivize. Further work in this direction could either be look at how ChronoMedIt could be used by practices to identify care process deficits that anticipate failures to meet pay-for-performance targets, or to fuel debate that the incentive systems themselves should be reformed.
5 ChronoMedIt for Clinical Intervention

Up to this point we have mainly spoke of ChronoMedIt as a tool to measure a whole cohort of patients and hence provide an assessment of a health provider setting (whether it is a single general practice or a whole country). The alternative use of the same underlying query results, however, is for ‘case finding.’ That is, rather than being interested in the percentage of cases in violation of a criterion, to be interested in just who those individuals are so as to follow up and improve the situation. The interactive EDS situation is like this for a cohort of size one – if there is a violation, then the user gets an alert. Our technology could be integrated with health information systems for this use, however, we are focusing instead on a process that might be done, say, quarterly where a list of cases is established and then systematically pursued for improvement. It is worth saying that we believe both cases are valid and useful (and we return to this point at the end of this section).

We have investigated the direct clinical use of ChronoMedIt reporting in New Zealand Health Research Council (HRC) Feasibility Study, called Adherence Innovation in Medication use for Health Improvement (AIM-HI). The AIM-HI intervention has involved the identification of a cohort of patients (approximately 200) who have an overall AHT MPR<80% for a 12-month period. These patient are then being actively followed up by practice nurses for a subsequent 12 months with the aim of improving their AHT adherence and their blood pressure control. As a Feasibility Study the size of the experiment is not designed to establish a statistically significant outcome, but rather to determine that the intervention is practical.

The AIM-HI assessment has included patient and health provider focus groups. Some of the qualitative findings include that patients are mostly happy with a telephone call to remind them when it is time to return to the practice. In our group at least, the option of a cell phone reminder (possibly an automated SMS message) was questionable. One should not be lulled by statistics indicating high levels of cell phone ownership – among our participants, among those around age 60, it was raised that they do own a cell phone, but they don’t necessarily keep it charged. Reactions to home visits or postal reminders were variable. A difficult issue was the timing of follow-up, however, as individuals differed in how much warning they wanted. Calling and expecting that they could come in one just a couple days notice was considered rude and disrespectful of how busy their lives were otherwise; but calling a week or more in advance risks heading off the opportunity for them to remember to make the appointment for themselves.

Staff issues included dissatisfaction with identifying low-MPR patients for special follow-up when some other patients may be sicker. Also, there is the general issue of a hard-pressed health workforce. There is a belief that human follow-up is the best way to get to know individual adherence barriers, negotiate around the barriers and influence the patients to do better. But, of course, this requires time that could be used in other activities.

We are pursuing two distinct agenda to follow on from the AIM-HI Feasibility Study. One is around automated technologies that largely avoid the health workforce issue. A particularly relevant automated intervention has been STOMP (Stop Smoking Over Mobile Phone) which demonstrated significant success in smoking quite rates for Maori and non-Maori participants. This inspires the notion that we may be able to engage patients through a programme of SMS messaging to improve their medication adherence behaviour. Obviously this is limited to the cell phone using segment of the cohort. For others it may be that the cornerstone should be something home based that monitors the actual medication administration events. There is also the possibility of using Web-based education, but with a focus on younger family members to influence the patient, rather than aimed directly (or at least exclusively) at the patient themselves.

We are also looking at further development of health workforce based interventions. We are proposing a trial where medical students doing training placements in general practice help patients identified via ChronoMedIt to work through adherence issues. The immediate outcome (even before MPR or blood pressure improvement) would be to see a greater level of Shared Decision Making (Dy 2007) in doctor-patient consults, indicating that the patients were taking the decision be on medication seriously.

These cases have illustrated use of ChronoMedIt reporting directly in clinical care. Health Informatics has focused on great deal on EDS technology. For instance, the ATHENA system is an excellent example of a comprehensive decision support system for hypertension management (Goldstein et al. 2000). Also, in New Zealand, there is substantial uptake of the PREDICT system which assess risk of cardiovascular disease events and then provides tailored recommendations of how to manage down that risk (Kerr et al. 2008). PREDICT has now had over 100,000 sessions of use with patients in New Zealand. We believe that the use of a case finding approach is complementary to these EDS efforts. With a list of cases, all with a common problem, there is the potential to assign a worker to that particular task. That worker will very often not need to be a physician, and may not even need a great deal of education. Furthermore, the two approaches can link in that one action from systematic follow-up of a case by a non-physician could be to tag that case in the practice management system so that the GP is alerted to the issue next time they see the patient.

6 Enhancing ChronoMedIt

We have been working to enhance the ChronoMedIt software for greater flexibility. Key elements of these efforts involve: the data uptake mechanisms; the internal representation of cases; and the family of audit criteria supported.

With respect to the data uptake mechanism, we had initially focused only on the market-leading New Zealand General Practice management system, Medtech32 by Medtech Global, and also national dispensing data as an auxiliary source. Obviously this is limiting, even within New Zealand, but also of course internationally.
To generalise the data uptake mechanism we have made the ChronoMedIt architecture more modular (see figure 3). There will always be some elements that will need to be specific to a given source EMR, but there are a couple components we believe will be quite general. One is notes ‘scraping’ for elements such as blood pressure (which, although they can be represented formally as observation data in Medtech32, are often just typed in as text in-line with other notes). The ways that blood pressures and other observations are typed into free-text notes are likely to be similar irrespective of the software used. Thus it is worth abstracting this module from the system-specific data uptake. We have done this through an XML based description of extraction pattern matching rules and an associated component that applies the rules to the appropriate files as part of the PMS-specific data uptake adapter. A similar generality holds for medication signs – these are essentially universal, not system-specific, and hence the interpreter for them forms a valuable resource that can be reused as we tailor data uptake adapters to particular PMS/EMR solutions.

Different systems (and different jurisdictions) will use different terminology sets for diagnoses and, to a lesser extent, medications. These variations are readily managed through our ontology, which allows multiple sets of terminology to be managed under higher-level relatively universal concepts. For example, New Zealand uses Read Clinical Codes for diagnosis coding in general practices, whereas International Classification of Primary Care (ICPC) codes are more common in Australia; while the details will differ, both systems will have term-code pairs that cluster under a high-level concept like Hypertension.

Ideally, we would like a standards based solution such that EMR systems could produce data ready for use by ChronoMedIt without system-specific adaptation. This level of interoperability is probably infeasible in the near term. We see in the Health Level 7 Clinical Document Architecture (CDA) (Dolin et al. 2006), however, the promise least for a framework to allow such interoperability. We aim to define a CDA input format that provides the data required for ChronoMedIt analysis. With such a format specification, in theory an EMR user with a CDA-compliant system could define a reporting format that met our input requirements directly.

With respect to internal representation, we have unified the concepts of observations and laboratory test results in the internal representation. These items come in very differently in many source EMRs. In Medtech32 the blood pressures are largely recorded in free-text notes. Laboratory test results come in over a third-party network as Health Level 7 messages and are stored in an inbox for the physician who ordered them. Despite the initial differences in source, we have come to realise that the role of each is equivalent in our query criteria – thus, a cholesterol level measured in an external laboratory or a blood pressure measured at the doctor’s office might equally be a measure that is, for instance, not taken as often as appropriate, or is hovering repeatedly at an unacceptable level.

Further complexity is introduced by derivative and composite measures. Blood pressure itself of course is two numbers, the systolic and diastolic’ cholesterol results are typically presented as four numbers. A quality criterion may refer to ranges on multiple numbers (e.g., a blood pressure under 140/90, requiring both systolic and diastolic to be under their respective targets) or as a function of numbers (as in an HDL:LDL, “good cholesterol” to “bad cholesterol” ratio). We have not reached a comprehensive solution for these cases.

An open-ended challenge is with respect to extending the types of queries managed by the system. MPR was an early extension of ChronoMedIt – it was not directly indicated by the initial eight criteria, but rather emerged from the pharmacoepidemiology literature as we researched the medication adherence problems that were so prevalent as initially indicated by simple lapses in medication supply. We have recently identified another query area that may attract clinical interest with respect to therapy that remains static despite failure to meet target. This still fits the larger criterion object model, but will require an extension to the internal data model because ‘static’ therapy implies a lack of dose intensification. So far we have not been concerned with dose levels beyond their role in determining the duration of medication supply.

7 Conclusion

We have developed a framework for analysis of EMRs to provide quality audit reporting in chronic condition management. Our ChronoMedIt framework provides software for data uptake from EMRs, management of relevant concept ontologies, user specification of quality audit criteria details, query processing per se, and reporting, including case timeline visualisation.

In terms of related work by others, our work most closely resembles the IDAN/KNAVE II framework (Boaz and Shahar 2005; Shahar et al. 2006) for temporal abstraction on clinical data. Our main point of distinction is actually with respect to our lack of generality; that is, ChronoMedIt is not designed to allow users to explore arbitrary temporal relationships in clinical data. Rather, it is focused on management of major classes of temporal query that we have found to be relevant for quality audit of chronic condition management.

![Figure 3: ChronoMedIt architecture.](image-url)
We have identified applications of ChronoMedIt in epidemiological research, performance management and clinical intervention. We are working to make the framework more flexible for wider use. We believe we have demonstrated the technical ability to make meaningful measurements of the quality of healthcare processes, and to identify patient cases with care that is suboptimal in well-defined ways. A larger question, however, lies in how best to utilise such capability to enhance healthcare delivery. We are interested in collaboration to apply the framework more broadly, especially where it may link to sustainable health improvement.

8 Acknowledgments

The authors thank their clinical colleagues, particularly Dr John Kennelly, A/Prof. Timothy Kenealy, Dr Jeff Harrison, Dr Leti Chang Wai and Dr Stewart Wells and the medical centres that have worked with us. This work is supported in part by a Partnerships for Excellence endowment from the New Zealand Tertiary Education Commission, a University of Auckland Doctoral Scholarship and a University of Auckland Faculty of Science Research Development Fund grant.

9 References


