

Coding of medical records via restrictive semantic topic tracking

HINZ Conference & Exhibition 2006

Primary Care and Beyond

Building the e-Bridge to Integrated Care

9-11 August 2006, Aotea Centre, Auckland, New Zealand

Alexei Tretiakov¹, Inga Hunter¹,
Dick Whidett¹ & Erkki Sutinen²

¹Department of Information Systems,
Massey University, Palmerston North, New Zealand.

²Department of Computer Science,
University of Joensuu, Finland

Email: I.Hunter@massey.ac.nz

Outline

- *introduction*
 - *why code medical records*
- *methods for coding medical records*
 - *manual coding*
 - *computer assisted coding*
- *automatic coding module to support
Practice Management System (PMS).*
 - *proposed system to automatically assign codes to medical records in general practice based on the free text entered by the practitioner*

Introduction

- richest data store of routinely collected personal health information.
- team-based primary health care
- focus on chronic disease management
- demonstrate improved health outcomes for the investment in primary care



increased need to share and to report on its contents

Sharing data

- Problems
 - quality of the data
 - need to be able to communicate and process the data efficiently.
- Solution
 - computer coding of medical records with controlled vocabulary
 - use of standardised coding systems

Coding problems

- neither standardised nor used consistently

Coding problems

- neither standardised nor used consistently
 - multiple coding systems
 - READ, ICD-9, ICD-10, ICPC, SNOMED, SNOMED-CT, locally developed
 - consultation diagnoses coded in 64.7% of practices
 - 24.6% of these practices only 'sometimes' or 'occasionally' used codes or only coded for ACC purposes.
 - 94.6% of these practices recorded screening information or keep disease registers on their PMS, what they recorded varied widely, ranged from cervical screening (97.8%) to prostatic specific antigen (PSA) (2.7%)

Coding problems

- neither standardised nor used consistently
- data may be only partially coded

Coding problems

- neither standardised nor used consistently
- data may be only partially coded
- code assigned to an event may be inaccurate

Coding problems

- code assigned to an event may be inaccurate
 - selection of the wrong code
 - range of potential codes may be wide and not standardised
 - correct code may not actually exist
 - sensitivity of use of diagnostic code for chronic obstructive pulmonary disease (COPD) was 79% with a positive predictive value of 75.3%, and patients with COPD had to be identified from additional searches
 - 3803 patients with a pre-existing ischaemic heart disease code (IHD) Read code, 15% no evidence of IHD.

Implications of poor coding

- accuracy and quality of the subsequent analyses, reporting and health planning based on inaccurate or incomplete coding.

How to increase the accuracy and quality of the coding applied to medical records

??????????

Current situation

- manual coding
 - medical practitioners
 - trained coders

- clinical coding by computer
 - semiautomatic
 - from free text

Manual coding by practitioner

- manual coding by practitioner seeing patient
 - best position to assign a relevant code

- But
 - the focus is on caring for the patient; assigning codes to a medical record is a low priority task
 - the time of medical practitioners is expensive.
 - additional and unattractive workload reducing time available for direct patient care

Practitioner coding time

- a range of 53-227 hours/month spent by general practices recording data
 - about 1,230,000 hours would be spent every month collecting data in general practice in the UK.
- GPs spend an average of 2 minutes manually coding 'reason for referral' using ICPC
 - systematic coding of the problems managed for all general practice encounters in France = about \$640 million each year.

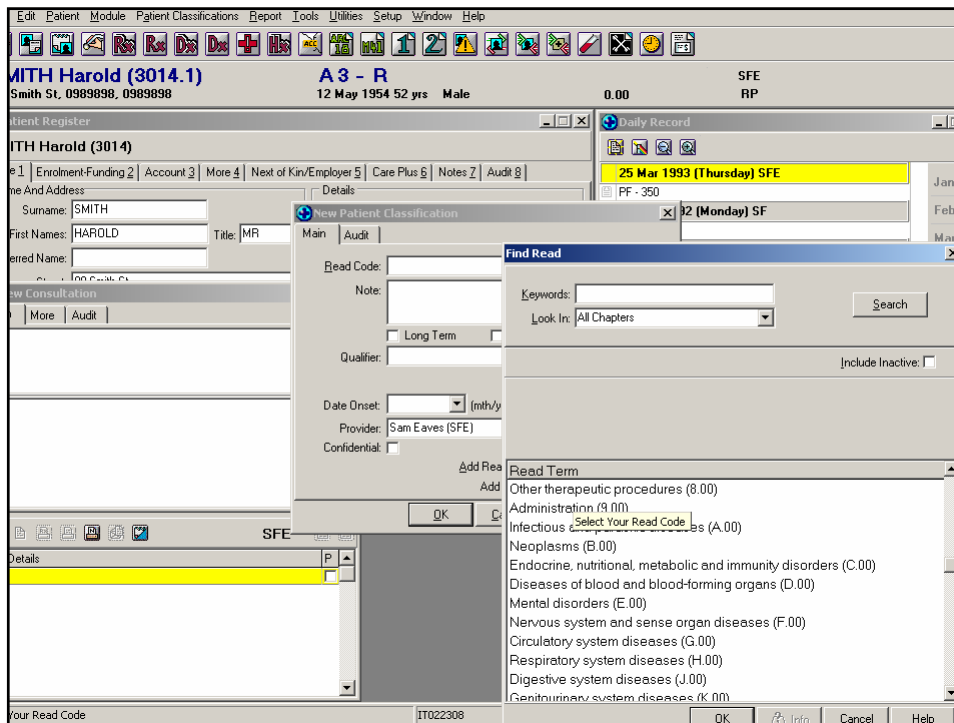
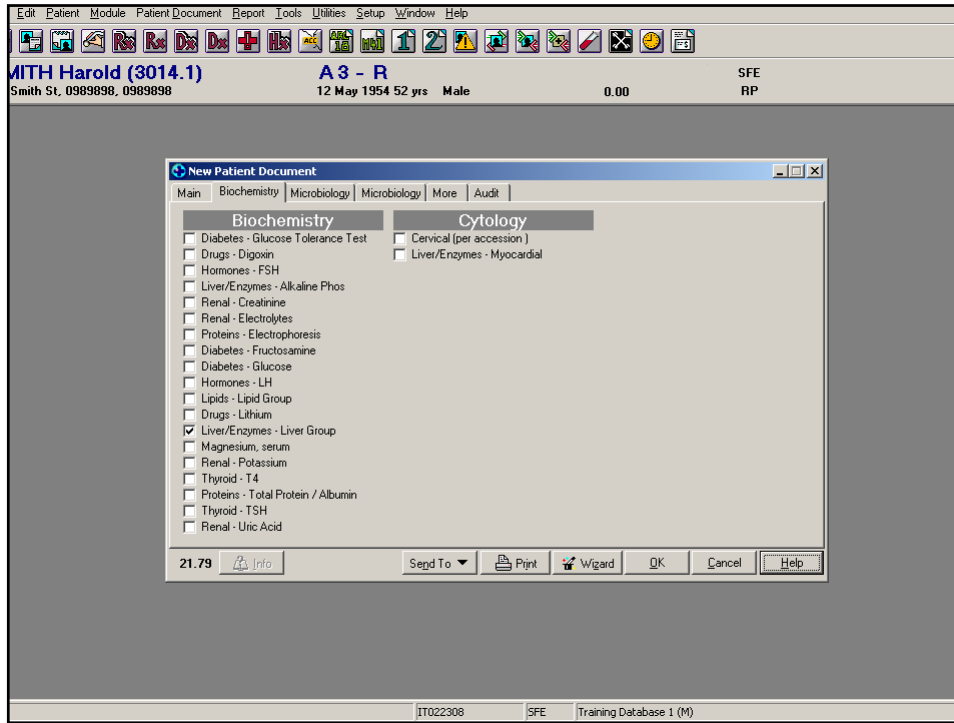
Manual coding by coders

- manual coding by trained coders
 - secondary/tertiary care
 - health maintenance organisations (HMOs) in USA
- on-line web service

Clinical coding by computer

- to improve accuracy with no additional workload for general practitioners (GPs)
 - is increasingly being used in medical records.
 - semi-automatic coding systems operating at the time of data entry (e.g. 'popup' or 'drop-down' menus of standard clinical terms)
 - being extended to computer assisted coding (CAC) using structured input,
 - the selection of menu item results in narrative text being produced and entered in to the medical record, such as that used in MedTech 32 Advanced Forms.

MedTech-32 Patient Register software interface showing a form for entering patient details. The form includes fields for Name (NEXT), Address, and various medical details like Date of Birth, GMS, Gender, Registered status, Provider, Ac Group, Chart No, and Ethnicity. It also includes a section for Cards (CS Card, HU Card) and Start/Exp dates. The interface is titled 'MEDTECH-32' and 'Patient Register'.



Computer assisted coding (semi-automatic) - problems

- constraints on the practitioner to use some kind of structured language or enter data in pre-defined manner, thus affecting the way in which they enter clinical data in to the EMR.
- inability to cope with the level of detail required by clinical practice.
- important clinical data may be recorded in free-text and be invisible to coding by traditional standardised coding systems
- thus currently only limited information in health records is in a structured format

Computer assisted coding (CAC) - from free text

- increase the completeness of any analysis from medical records
- allow GPs to continue to enter data in a flexible manner
- “ideal of a (coding) system that combines the advantages of structured records with the richness of free text” (Lewis, 2002)
- limited application to general practice consultations to date

CAC technology

- underlying automatic reasoning engines
 - combining soft computing and rule-based reasoning.
 - Soft computing approaches require the use of learning data sets. They may be applied to capture and to re-use tacit knowledge.
 - Rule-based reasoning provides the the reasoning engine with the ability to justify its decisions

A proposal for patient management system enhancement

To improve the coding process, a PMS incorporating an EHR system should be able to:

1. assess the quality of coding by an individual practitioner or by an organisation or a division in an organisation;
2. generate codes automatically when codes are not assigned by the practitioner;
3. assist the practitioner during data entry by suggesting codes based on the free text information entered by the practitioner.

methodology

- document clustering techniques
- tree or as a directed acyclic graph (DAG), with nodes/subnodes of the tree or DAG corresponding to individual codes
- for each layer, a topic tracking algorithm is applied to place each record entered by a given practitioner

restrictive clustering

- topic tracking algorithm makes best effort to assign each record to a number of codes at different level of detail
 - a value corresponding to the confidence level of each assignment is calculated
 - the code corresponding to the lowest level of granularity, for which the confidence level is better than an a priori taken acceptable confidence level is associated with the record as its "soft code".
 - A record that can not be assigned to any of the codes with sufficient confidence remains unassigned.

‘Soft codes’

- The soft code is used to fulfil the three requirements stated earlier
 - The difference between the soft code and the "hard code" entered by the practitioner (if entered at all) is taken as an indicative measure of the coding quality.
 - When codes are not assigned by the practitioner, soft codes can be used as surrogates of hard codes in situations when a code is required (e.g., for reporting purposes).
 - the PMS may assign a soft code to a record that has just been entered, and prompt the practitioner whether s/he would like to accept the resulting soft code as a hard code (a decision, for which, ultimately, the practitioner will be responsible).

Conclusion

- computer assisted coding from free text (CAC)
 - improves completeness of analysis from medical records,
 - accommodates flexible manner of data entry,
 - increases the quality and completeness of the coding of clinical encounters in general practice
 - without placing any additional workload on the practitioner.
- automatic CAC module enhance PMS systems:
 - providing functionality to assess quality of coding by individual practitioner/organisation/division in an organisation,
 - generate codes automatically when codes not assigned by practitioner,
 - assist practitioner during data entry by suggesting codes based on free text information entered by practitioner.

The challenge

- Can it be done in practice?
- One further question -

The challenge

- Can it be done in practice?
- One further question -

Should it be done?
or is semi-structured CAC sufficient?