

**eHID Project:
Lessons for making international
comparisons using electronic medical
records**

Douglas Fleming

Birmingham Research Unit RCGP

RCGP annual meeting Edinburgh Oct 2007

Objectives of eHID

- To examine recording practices in existing primary care networks
- To make site visits and observe recording first hand
- To make recommendations on the use of emr for epidemiology thus enabling comparison between networks

Session Agenda

- To use the findings of our project to highlight the limitations inherent in using routine data contained in electronic medical records when making international comparisons.
- To suggest ways in which the use of differing recording systems can be used to provide comparable information.

The EMR is a filing cabinet containing medical information



- Everything put in it can be retrieved
- But if we want to retrieve information readily we must put in in an orderly fashion.
- Links between files are essential

A Computer is a filing cabinet

- 'Garbage in and get garbage out'
- But medical information is not garbage
- It is incomplete information and inconsistent entry which are the chief causes of garbage out
- Free text information is un-filed data. It cannot be used in an effective way for epidemiology

Coded information correctly filed produces reliable information.

Classification systems are the filing clerks but they must be used properly.

The links between files are the software. Just because the information is recorded in the computer does not mean it can be used for epidemiology

Asthma: Free text records are not an orderly record nor readily useable

- Recent hospital admission for asthma
- Had asthma as a child.
- Never had asthma.
- Reversibility test for asthma
- Worried about son with serious asthma.
- Father died of asthma.
- Occupational asthma.
- Asthma attack.
- Asthma review.

Key issue 1: Denominator

- Age gender specific denominators are essential for epidemiological analysis
- Formal nomination of a central record for all patient specific information provides an alternative
- Systematic estimation of the denominator from the pattern of patients consulting is possible

Individual long term patient registration provides the most useful denominator

Key issue 2: Sample or not?

- Total data has strength of numbers but loss of quality.
Networks are needed in order to test for data quality
- Sample of people, patients or doctors?
Doctors or practices most feasible
- Representativeness of selected samples by age and gender; region; ethnicity; social class etc
Representativeness desirable but not essential if key data are retained whereby representativeness can be estimated
- Continuity of sample.
Desirable for monitoring trends

Key issue 3: Episode construction

- We need to distinguish new episodes of illness from ongoing consultations and ideally first episodes from recurrent or new
- Episode construction can be made by the doctor at the time of recording or by the analyst using an algorithm
- Episodes distinguish data about prevalence from incidence and workload

Episode construction to be made routinely by the doctor at point of data entry

Key issue 4: Health system structure

- Coverage for all or for some? Balance between state and insurance cover.
- The GP as gatekeeper
- Is specialist provided information entered in the notes in a usable manner?
- Is the record complete or is health related information consistently lost?
- What happens to information about emergency care?

No recommendation on national choice of healthcare structure, but it is highly desirable to ensure that all primary healthcare data are stored in a single place

Key issue 5: Data quality

- Completeness of data; how do you deal with missing data?
- How do you recognize that data are missing?
- Do you know if the patient is still alive?
- Are other data sets available which can be used for cross validation?
- Should you edit according to recording quality?

Without individual patient linkage you cannot audit the completeness or quality of data. Studies of internal consistency are desirable. It is important to have dedicated networks in order to check the validity in more detail and to substantiate the findings from total data sets.

Key issue 6: What is contained within the record?

- Reason for encounter?
- Doctor assessment ?
- Proven diagnoses?
- Can you use a patient summary record?

Doctor assessment is the cost driver of health services and therefore the preferred piece of information if only one is to be used in an automated way. However, the basis of the diagnosis must always be clear. If reason for encounter is to be recorded it needs to be an additional item. Summaries are not sufficiently consistent.

Key Issue 7: Disease Classifications

Several classifications were in use (ICPC – The Netherlands: ICPC-2 – Belgium, Spain and Malta: ICD-9 – Italy, ICD-10 – Spain: DCR – France: Read codes – UK)

Networks wished to retain their existing classification system

The classification is not the primary issue. The emr needs to preserve diagnostic information in a coded form at the most precise level of diagnostic detail as this is essential for management. For integration with secondary care, classification linkage is essential if different systems are in use in differing sectors of healthcare

Key issue 8: Coding

- The networks were accessing the emr and coding information in differing ways; some using coding manuals; others a Thesaurus approach; and others using Automated methods via user friendly software
- Data storage for primary record needs to be in maximum detail however rare.

We recommend automated coding via a user friendly Thesaurus to be applied at every consultation. Free text is not amenable to routine epidemiological investigation. Access routes to classification systems need to be borne in mind.

Key issue 9: Data links

- The electronic record needs to be
 - an integrated record which can be used by other doctors in both primary and secondary care
 - capable of absorbing linked data from laboratory sources
 - linked to prescribing information
 - secured with an audit trail
 - retain diagnostic links for all data entries

We recommend that software systems ensure that all data entries are linked directly to the problem or diagnosis to which they relate

Key issue 10: Definitions

- Age group definition;
recommended breakpoints at 5, 15, 45, 65 and 75 years but greater detail an advantage
- Case definition:
needs to be appropriate for the routine activities of healthcare delivery.
- Incidence or episode incidence?
First incidence as well as episode incidence but the latter essential
- Point, annual, or cumulative prevalence?
Annual prevalence preferred as a routine
- Comparability across classifications.
Important issue but we cannot make recommendation as to how this is achieved

Key issue 11: Confidentiality issues

- Data string with name allowing retrospective linkage
- Data string with encryption of name
- Data string with no possibility of returning to source
- Data extraction after analysis
- Patient opt-out issues
- Need the concept of custody of individual files & links

Recommend availability of data for epidemiology to be based on fully anonymised (untraceable) data but preferably with unique linkages retained.

Patient opt-out is a serious threat to information systems for the management of public health.

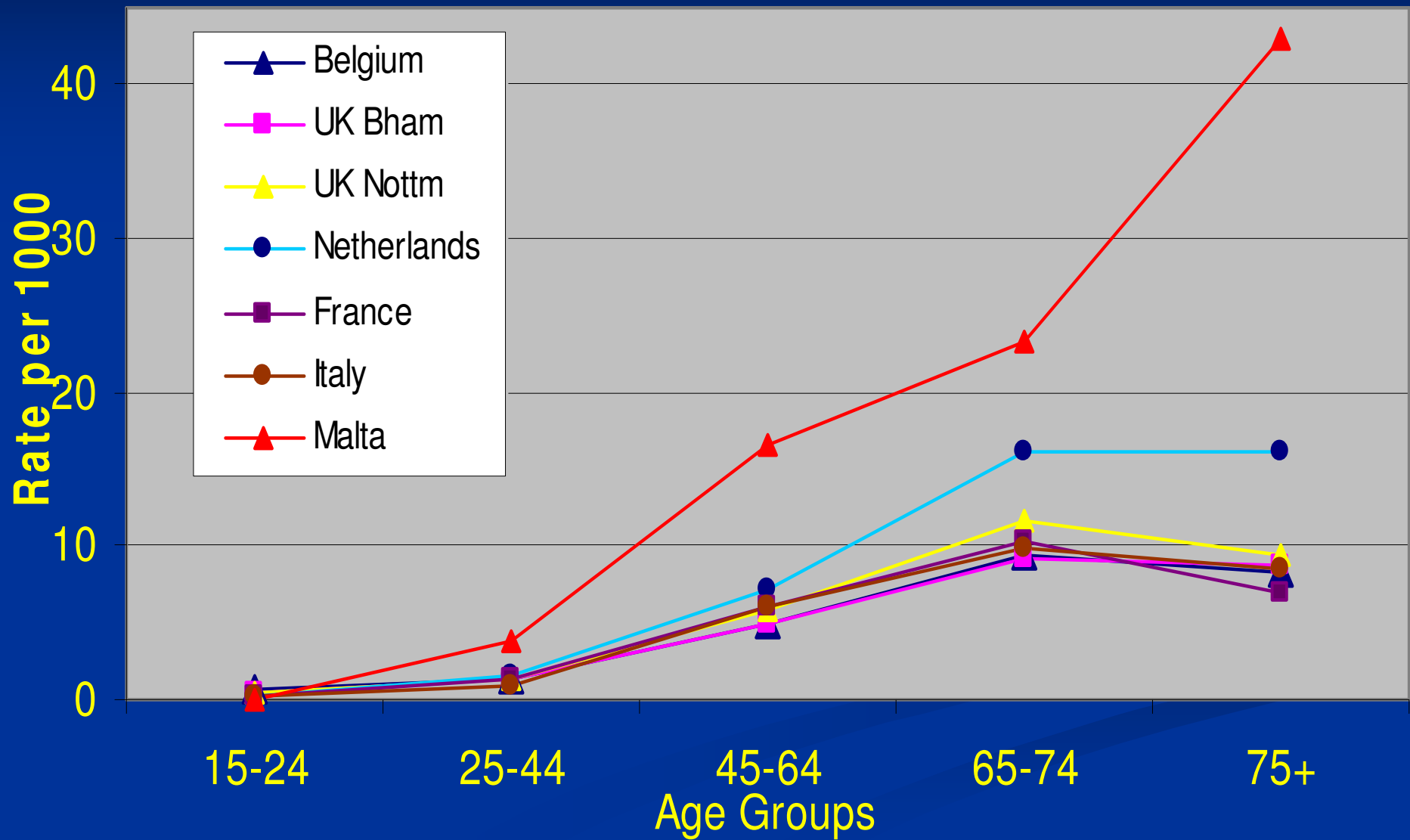
All file boxes need to be filled in a structured patient electronic record

- *Essential for an accurate patient record.*
- *Free text is important. Negative and qualifying details need to be stored but not in a way that confuses analysis*
- *Sometimes you can bring data together from different consultations in order to fill every box*
- *Many episodes of illness involve only one consultation, therefore complete the assessment box at each consultation.*

Summary

- The electronic medical record has arrived
- It will provide good epidemiological data only if we keep the epidemiological purpose at the top of the agenda
- There are problems to ensure comparability but most are surmountable
- Crude number counts are no good without interpretation in the context of the health system from which they come
- Issues of data quality cannot be ignored
- Confidentiality can be secured, but individual patient opt-out cannot be accommodated.

Diabetes Incidence Adults 2004 non standardised



Diabetes Incidence 2005 Adults

non standardised

