

**METHODOLOGY FOR LINKING AND  
MATCHING HOSPITAL AND  
MORTALITY DATA – SUMMARY**

**Leicester Gill, Alastair Mason and Michael Goldacre**

**National Centre for Health Outcomes Development**

**June 2003**

**UNIT OF HEALTH-CARE EPIDEMIOLOGY**

**UNIVERSITY OF OXFORD**

**REPORT MR1**



**record-based**

**O**

**U**

**T**

**C**

**O**

**M**

**E**

**measures**

**UHCE OXFORD REPORT MR1  
METHODOLOGY FOR LINKING AND MATCHING  
HOSPITAL AND MORTALITY DATA**

**Leicester Gill, Alastair Mason and Michael Goldacre  
National Centre for Health Outcomes Development  
June 2003**

**Summary of objectives**

This report, commissioned by the Department of Health, reviews the methods to be considered in linking and matching hospital episode spells and ONS mortality data.

**Abstract**

Record linkage is simply the bringing together of information from two different records that are believed to belong to the same person. The records may come from a single data file or from multiple data files. They may relate to persons or to families. Where the two records agree on all the variables, and are unlikely to have done so by accident, the level of assurance that the link is correct will be high. Conversely, if most of the variables disagree there will be little doubt that the linkage is wrong. For intermediate situations the record matching methodology must predict whether the link is true or false or indeterminate.

The introduction of errors and variation in a dataset is unavoidable during the transcription and keying of such data items as names, addresses and dates. This makes the task of record linkage far more difficult since it is not possible to be absolutely sure whether two matching records have identical values on the selected variables purely by chance or that the two records genuinely correspond to the same person.

The data files may have been generated through administrative procedures or from surveys or censuses. A file may represent an entire defined population or only a sample from a defined population. The files may have the same or different time references. In many record systems, input files may contain duplicates, and by making linkages both within and between files these duplicates may be removed, also, missing data items may be copied from one record to another.

Automatic record matching and linkage using computer methodology involves striking a balance between the efficiency of the process and the quality of the matched file. While the technology is becoming more powerful and less expensive, enhancements of the mechanics of file organisation and record matching will reduce the time spent in file processing as well as providing an improvement in overall performance and match rate.