Development of a method to link infection and mortality data

N. Potz¹, D. Powell³, R. Pebody⁴, R. Lamagni⁵, D. Bridger⁶, G. Duckworth⁶
¹Health Protection Agency, Cfl, 61 Colindale Avenue, London NW9 5EQ ²The Office for National Statistics, 1 Drummond Gate, London, SW1V 2QQ

Introduction

Probabilistic Record Linkage Methodology

Probabilistic record linkage is a method developed specifically for date and fact based containing errors and omissions, and is used to link records that do not have universally available and unique identifiers.

The infection records linked were:
- A random sample of 1247 infections due to Streptococcus pneumoniae
- Blood and/or CSF from meningococcal infections
- From England
- From Scotland in the same period of 1st July 2003 - 30th June 2004

The mortality records linked were:
- All death registrations (N = 924,938)
- From 01 July to 31st March 2004
- From England and Wales

The variables available for linkages in the infection and mortality data sets were:
- NHR number
- Source
- Day, month and year of birth
- Sex
- Forename initial
- Postcode prefix (e.g. SW, ME)

NHS Central Register Tracing Service

The same 1243 invasive S. pneumoniae infection records from 1st July 2003 to 30th June 2004 were submitted to the NHSIR for Tracer Extracts (ie. all deaths). Only deaths occurring prior to 31st March 2005 were included in the analysis to allow comparison with the probabilistic record linkage results.

- 413 deaths were identified through the deterministic "Matcher" service
- 67 further deaths were identified through manual Operator Matching

480/1243 (38.6%) infection records matched to a death record

Probabilistic Record Linkage Evaluation

By comparison with the conventional "golden standard", we can see that for any infection record positively linked to a death record by probabilistic record linkage, there is a 99% probability that that person has in fact died. This is very similar to the figure observed for the NHSIR Tracing.

Probabilistic Record Linkage Results

The subsequent record linkage method developed during this study has succeeded in establishing an accurate link between infection and mortality records. Many infection records do not contain summary data, only Sources, resulting in the ability to link the records with the most accuracy. This linkage between probabilistic record linkage will therefore be utilised during the main phase of the study to analyse mortality post-invasive Meningococcal infection through:
- estimation of case fatality ratios at different time intervals between diagnosis and death, stratified by age, sex and documented underlying cause of death.
- examination of the relationship between laboratory confirmed infection prior to death and routine reporting of infection on death certificates, and factors associated with reporting such as age, sex, underlying cause of death and time interval between laboratory diagnosis and death.

This work has the potential to substantially improve our understanding of mortality following infection, and is providing a platform for further epidemiological study.

Conclusion

References

1. Health Protection Agency, Newcastle upon Tyne Institute of Virology and AIDS. "Introduction to linkage and methodology:
2. D. Bridge, R. Pebody, T. Lamagni, I. Wilson, D. Pollard, N. Potts, N. Hymas, G. Duckworth. Methodology for the National Confidential Study of Chal...