The value of linked data for research into surveillance and adverse events

Nick Andrews

PHLS Statistics Unit, 61 Colindale Avenue, London, NW9 5EQ (nandrews@phls.org.uk)

Abstract

Linked data is used by the public health laboratory service (PHLS) in England and Wales for performing enhanced disease surveillance and also assessing the safety of vaccines. Enhanced disease surveillance, which links disease data from different sources is important when introducing a new vaccine in order to assess the need for the vaccine, the impact of the vaccine on disease and to estimate field efficacy. This was particularly important when introducing the new meningococcal C disease conjugate vaccine because it was licensed without any formal efficacy trials. Since the introduction of the vaccine at the end of 1999 field efficacy has been estimated at over 90% in all age groups indicating good short-term efficacy. Another application of data-linkage for surveillance is linking data on deaths due to a disease from different sources and then using the degree of overlap to estimate total deaths due to the disease using capture recapture methodology. This was used to estimate that there were a total of 46 deaths due to pertussis in England from 1994 to 1999.

As diseases disappear focus turns to vaccine safety and it is essential to be able to test hypothesised reactions to vaccination. Record linkage of hospital admissions and vaccine records, along with the self controlled case-series method, has been used many times at the PHLS in the past decade to estimate the relative incidence of adverse events in specified risk periods after vaccination. Examples of such studies include looking at ITP after MMR, convulsions following MMR, bacterial infections after MMR and intussusception after oral polio vaccine.

1. Introduction

This paper examines how disease surveillance and adverse event surveillance is crucial when assessing the impact and safety of vaccination. The necessity of being able to link data in order to carry out enhanced vaccine preventable disease surveillance and also assess safety is highlighted using examples from the PHLS in England and Wales. The recent introduction of the meningococcal group C conjugate vaccine without formal efficacy studies pre-licensure highlighted the need for an enhanced surveillance program to monitor the impact and estimate efficacy of disease post-licensure. Also the increasing publicity over vaccine safety means it is more important than ever to have active post-licensure surveillance of possible adverse events.

In this paper the issues that need to be considered when assessing a vaccine are discussed, next the use of surveillance data and the requirement of data linkage for enhanced surveillance is examined with examples and finally the PHLS record linkage system for adverse event surveillance is presented with examples.

2. Vaccine assessment

Vaccines can protect individuals and the population as a whole against disease and may even enable the elimination of disease. However they may also cause adverse reactions and, if used carelessly, change the disease epidemiology for the worse. Having good systems for vaccine assessment is essential, particularly when introducing a new vaccine. Some of the key questions that need to be answered along with the information required are as follows:

What is the disease incidence, morbidity and mortality? This information is required to decide whether the disease is serious enough to be worth targeting, Surveillance data is required to determine the disease epidemiology and also data on deaths to determine mortality.

Does the vaccine work? This is usually measured by vaccine efficacy and is determined in clinical trials. However field efficacy may differ from that seen in trials. Field efficacy can be estimated using enhanced surveillance where the vaccination status of cases is determined, along with population vaccine coverage data.

How safe is the vaccine? Clinical trials will pick up common reactions, however rare reactions require post-licensure surveillance.

Will the vaccine be cost effective? This can be assessed using mathematical and economic modelling of the likely impact of vaccination on disease, this requires estimates of the cost of the disease and the vaccine as well as vaccine efficacy, coverage and population mixing patterns and also the infectiousness of the disease.

At what age should the vaccine be given? This requires information on the disease epidemiology from surveillance data and also information on the immunological response and efficacy seen at different ages from clinical trials.

What will the effect be on the disease epidemiology? This can be predicted using mathematical modelling based on vaccine efficacy, coverage, and the age at which vaccine is given along with population mixing patterns and also the infectiousness of the disease. Vaccine programmes can be planned to prevent deleterious effects on disease epidemiology by including catch up campaigns or booster doses. Post-vaccination enhanced surveillance will provide information on the effect of the vaccine on disease epidemiology.

In England and Wales much of the information required to answer these questions is provided by the PHLS which reports to the joint committee on vaccines and immunisation (JCVI) which itself makes recommendations to the government.

Clearly the introduction of a new vaccine as well as continuing assessment of vaccines in use requires good quality surveillance data and also the ability to investigate vaccine safety post-licensure. The next two sections highlight the use of linked data at the PHLS in England and Wales to perform surveillance and safety assessment.

3. Data-Linkage for Enhanced disease surveillance

Data on diseases are available from many different sources. In England and Wales, for example, these include disease notifications to the office of national statistics (ONS), the ONS deaths register, hospital episode statistics, routine and reference laboratory confirmations and general practice databases.

These data sources can be used individually to provide information on disease epidemiology, however a complete picture can only be obtained through enhanced surveillance which links together various data sources. In order to link the data sources patient identifiers are required such as NHS number or sex/dob/post-code. Ethical approval is required to use and link the data sources. Some advantages of data linkage for surveillance are listed below:

- Better estimation of disease incidence
- Estimation of correction factors for under-reporting by a single data source.
- Prevention of double counting
- Identification of repeat episodes
- Reduction of missing information.
- Potential use of capture re-capture methods for estimating overall incidence.
- Combination of data only available from different data sources (e.g. deaths, sero-groups, vaccination status).

In the England and Wales enhanced surveillance exists for all diseases for which vaccination is routinely used as well as other diseases for which vaccines may be available in the future.

Two examples follow which demonstrate how data-linkage is used for surveillance. The first shows how capture re-capture methodology can be used to estimate total deaths in England from pertussis. The second gives details of the enhanced surveillance system for meningococcal disease.

3.1 Example: Deaths in England due to Pertussis 1994–1999

In order to improve the estimate of mortality due to pertussis in England and identify reasons for under ascertainment a study was performed which involved linking data from laboratory reports of pertussis, hospital episode statistics (HES) and ONS death registrations [1]. Data were obtained from 1994 to 1999 and a total of 33 deaths were identified. Figure 1 shows that there was overlap between all the sources. Twenty-two deaths were identified from laboratory reports, 18 in ONS mortality data, and 9 in HES.

Using capture-recapture analysis [2], the total deaths from pertussis in the five and a half year period of the study was estimated to be 46 (95% CI 37 to 71), or around nine deaths per year. Further examination of the children who had died from pertussis without mention of pertussis on the death certificate showed that they were more likely to have been certified by coroners than those with mention of pertussis (p=0.005). Data-linkage has therefore shown that less than half the deaths due to pertussis are identified by any single data source.

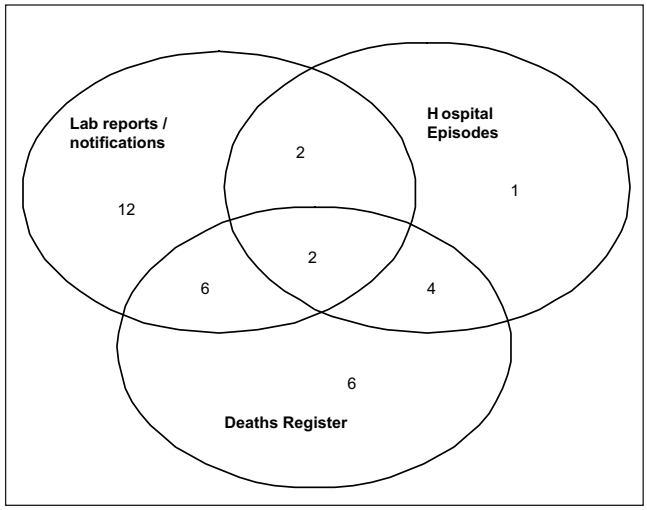

Figure 1 Overlap of deaths due to pertussis from three data-sources

3.2 Example: Enhanced meningococcal disease surveillance

During the 1990s meningococcal group C disease incidence increased steadily with frequent outbreaks and high mortality. Vaccination against group C disease was only available with a polysaccheride vaccine which gave short term protection and it was only recommended for use with close contacts. Also the vaccine did not produce an immune response in infants. Following the success of the HiB conjugate vaccine and conjugate vaccine was developed for Meningococcal group C which gave immune responses in infants and was also likely to give long term protection. The rarity of meningococcal disease made efficacy trials impractical, however the immune response to the vaccine was shown to be good [3] and the vaccine became the first to be licensed without formal efficacy trials. This made the need for enhanced surveillance of the disease even more important in order to monitor the effect and efficacy of the vaccine.

In 1999 the surveillance of meningococcal disease was therefore enhanced with information from the data sources shown in figure 2 linked into one database. The vaccine was introduced in November 1999 with priority given to infants and 15–17 year olds in whom death rates were highest. During 2000 all other age groups up to age 17 received the vaccine as it became available.

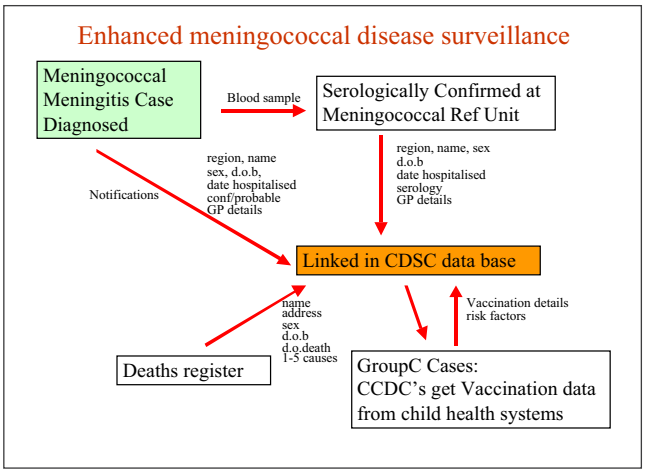

Figure 2 Enhanced Meningococcal disease surveillance

Before the vaccine was introduced the enhanced surveillance system enabled estimates of disease due to group C to be obtained with a correction for unconfirmed cases. Many cases of meningococcal disease are never confirmed because blood samples are not taken or are taken too late, these unconfirmed cases can be assumed to be distributed across sero-groups in the same proportion as seen with confirmed cases. This gave corrected estimates of group C incidence approximately twice as high as those obtained just using confirmed cases. The results showed that the 1999 incidence varied from 0.09 per 10,000 in over 25s to 3.7 per 10,000 in under ones.

Since the vaccine was introduced the enhanced surveillance database has since been used to measure the effect of the vaccine on disease rates and deaths, look at any changes to the sero-epidemiology, identify vaccine failures and estimate vaccine efficacy when used with vaccine coverage data. Analyses comparing data from January to September 1999 with data from January to September 2000 show a decline in deaths of 90% and disease in under 20's of over 80%. Current (unpublished) vaccine efficacy estimates using the screening method [4] are above 90% for all age groups. Sero-epidemiology shows that there is no evidence of common group C strains switching capsules to become other capsular groups.

4. Data Linkage for Vaccine Safety

The ability to detect and investigate possible adverse reactions to vaccination is becoming increasingly important as more vaccine are licensed and as the diseases themselves disappear and focus turns to possible adverse reactions.

Pre-licensure clinical trials can detect and estimate the risk of common reactions to vaccination such as fever and localised rashes and swelling. Rare reactions however may only be detectable and quantifiable in post-licensure surveillance.

Possible reactions may be identified from clinical trial data, passive post-marketing surveillance systems and through the biological mechanism by which the vaccine gives protection. Once a possible reaction has been identified it is important to be able to carry out a formal study of the hypothesis in which bias can be minimised. Possible study designs are cohort studies, case-control studies or self-controlled case-series studies. In these studies the relative incidence (RI) of the reaction in a specified post vaccination risk period compared to the background risk is estimated. If there is an increased risk then the risk per dose attributable to vaccination also needs to be estimated. Ongoing systems that have been developed for this active surveillance include US vaccine safety data-link system [5] which is a cohort method and the UK record-linkage system which use the self controlled case series method [6].

The UK record linkage system for active surveillance is used to test hypothesised reactions to vaccines by identifying hospital admissions for the event from ICD codes and linking these records with child health vaccine databases. In the UK ethical approval has been obtained to use hospital episode data from the south east of England and three child-health vaccine databases covering a similar area.

Figure 3 gives details of the record linkage. These data have been used in the past decade to perform many studies and the SCCS method [6,7,8] was developed to analyse the data. This method, which only uses data on cases, has been shown to have similar power to a cohort study and also has the advantage of minimising bias because cases act as there own controls. An example of a record linkage study was the investigation of MMR and ITP.

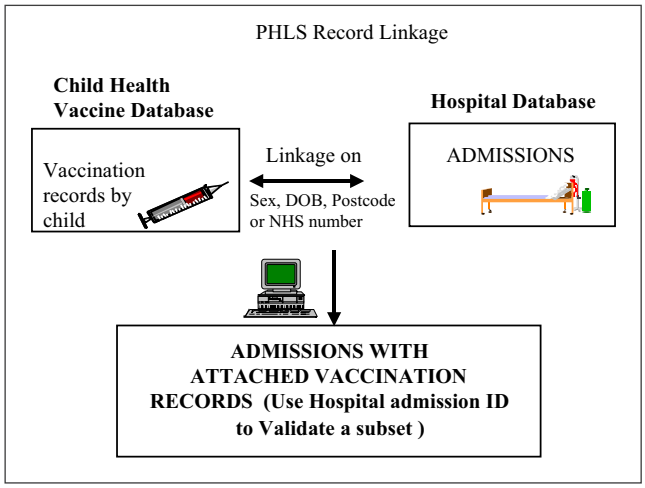

Figure 3 Data-linkage for vaccine safety

4.1 Example: MMR and ITP [Miller et al 9]

Idiopathic thrombocytopenic purpura (ITP) is a blood disorder related to the immune system. Some observational evidence as well as biological plausibility suggested that there may be an increased risk of ITP in the 6 weeks after MMR vaccination. Record linkage was therefore used to link individuals with hospital admission with an ICD code for ITP (ICD 278.3) with vaccine records. At the time of the study data were available between 1991 and 1994.

Figure 4 shows the distribution of events according to the number of weeks from vaccination. The figure suggests there is an increase in the risk period, and this was confirmed in the self controlled case series analysis with a RI of 3.27 and 95% confidence interval 1.49 to 7.16. Based on the estimated number of MMR doses given to the population from which the study cases arose the estimated risk per dose of MMR was 1 in 32000.

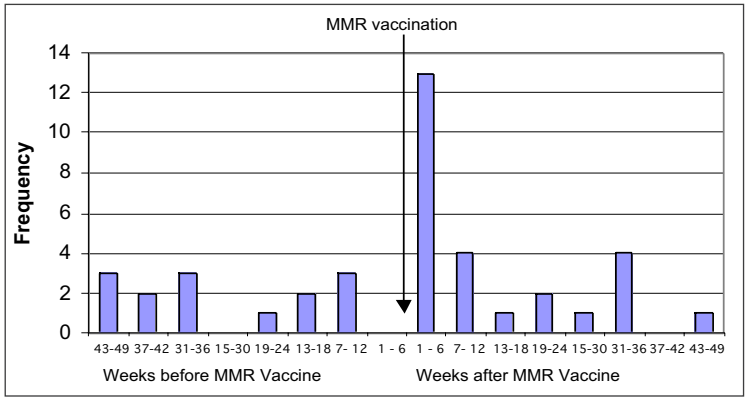

Figure 4 ITP cases by weeks from MMR vaccination to hospital admission

4.2 Other record-linkage studies

Details of some other Record-linkage studies performed by the PHLS are given in table1. Note that the MMR-Autism study differed slightly from the other studies because in this case disability registers and special school were used to identify cases rather than hospital admissions. These studies have enabled quantification of the risks of vaccination and have also shown where hypothesised adverse reactions have been false alarms. The studies have been highly influential in global decisions about the safety of MMR.

Clearly it is important to be able to investigate reactions as quickly as possible to prevent public concern causing a drop in coverage and the re-emergence of the disease.

Study [Reference]	Result
MMR and Autism with regression [12]	RI = 0.85 (0.45 –1 .60) within 6 months of MMR
MMR and Convulsions [6]	RI=3.0 6-11 days post MMR Risk per dose 1 in 3000
DTP and Convulsions [6]	RI = 3.0 0-3 days post DTP Risk per dose 1 in 12,500
MMR and Aseptic Meningitis [6]	RI = 38.1 15-35 days post MMR Risk per dose 1 in 16,000 (Urabe mumps strain only)
MMR and Bacterial Infections [11]	No increased risk 6 weeks post vaccination
OPV and Intussusception [10]	No increased risk 6 weeks post vaccination
Men C Vaccine and Purpura	In Progress
Men C Vaccine and Convulsions	In Progress

Table 1 PHLS Record Linkage Studies

5. Conclusions

Linked data contributes to enhanced disease surveillance which is essential for assessing the impact of vaccination as well as other health interventions. The linkage of hospital admissions data and vaccine data along with the SCCS analysis method provides a rapid, cost effective and powerful tool for investigating hypothesised adverse reactions to vaccination. This system is currently only used in the UK but may provide other Countries with a relatively inexpensive method for assessing vaccine safety. There are clearly difficulties with ethical approval, however investigating adverse events will be increasing important as diseases disappear and attention is turned to these adverse events.

References

1. NS Crowcroft, N Andrews, C Rooney, M Brisson, E Miller. Deaths from pertussis are underestimated in England. 2002, Arch Dis Child. In Press

2. International working group for disease monitoring and forecasting. Capture-recapture and multiple-record systems estimation II: Applications in human diseases. American J Epidemiol 1995;142:1059–1068

3. Richmond PC, Miller E, Borrow R, Clark S, Sadler F, Fox AJ, Begg NT, Morris R, Cartwright KAV. Meningococcal serogroup C conjugate vaccine is immunogenic in infancy and primes for memory. J Infect Dis 1999;179:1569–72.

4. Farrington C.P. (1993). Estimation of vaccine effectiveness using the screening method. International Journal of Epidemiology, 22: 742–746.

5. Chen RT, Glasser, JW, Rhodes PH, Davis RL, Barlow WE et al. Vaccine Safety Datalink Project: A New Tool for Improving Vaccine Safety Monitoring in the United States. Pediatrics 1997;99:765–773.

6. Farrington P, Pugh S, Colville, A, Flower A, Nash J, Morgan-Capner P, Rush M, Miller E. A new method for active surveillance of adverse events from diphtheria/tetanus/pertussis and measles/mumps rubella vaccine. Lancet 1995; 345:567–69.

7. Farrington CP. Relative Incidence Estimation From Case Series for Vaccine Safety Evaluation. Biometrics 1995;51:228–235.

8. Farrington CP, Nash J, Miller E. Case Series Analysis of Adverse Reactions to Vaccines: A comparative Evaluation. Am J Epidemiol 1996;143: 1165–1173.

9. Miller E, Waight P, Farrington P, Andrews N, Stowe J, Taylor B. Idiopathic thrombocytopenic purpura and MMR vaccine. Arch Dis Child 2001; 84:227–229.

10. Andrews N, Miller E, Waight P, Farrington P, Crowcroft N, Stowe J, Taylor B. Does Oral Polio Cause Intussusception? Evidence from a sequence of three self controlled case series studies in the UK. 2002 European Journal of Epidemiology. In Press

11. E Miller, N Andrews, P Waight, B Taylor. Bacterial infections, immune overload and MMR vaccine. Arch Dis Child. In Press.

12. Taylor B, Miller E, Farrington CP, Petropoulos M-C, Favot-Mayaud I, Li J, Waight PA. Autism and measles, mumps and rubella vaccine: no epidemiological evidence for a causal association. Lancet 1999; 353: 2026–2029.