Validity of using Hospital Episode Statistics data on monitoring disease trends

We read with interest the article by Koshy et al. The findings are important in documenting changes in admission rates of childhood pneumonia and empyema since the introduction of heptavalent pneumococcal conjugate vaccine (PCV7). We are concerned that undue emphasis has been placed on Hospital Episode Statistics (HES) data to define the aetiology of childhood pneumonia, particularly ‘bacterial pneumonia’.

Given the magnitude of the case numbers reported, it would appear that the analyses are based on all pneumonia codes collectively. This would also (although it is not clear from the article) include ‘unspecified pneumonia’, which describes pneumonia of any aetiology. Our analysis of national HES data on childhood pneumonia (1997–2006) showed that 91% of cases were coded as unspecified pneumonia. This may be of significance given that much unspecified pneumonia in children is likely to be viral; in routine clinical practice it can be difficult to differentiate between viral and bacterial pneumonia.

The authors also assert that ‘PCV7 offers protection against the most common serotypes accounting for most of the bacterial pneumonias in children’. The references provided do not support this statement. There are international variations in serotype distributions of laboratory-confirmed pneumococcal disease. There are no published data on the serotype distribution of pneumococcal pneumonia for UK children.

We have evaluated the accuracy of HES data for paediatric pneumonia in the North East of England. The incidence was previously established in a prospective study, and we repeated it prospectively between 2008 and 2009. Of 50 subjects identified during prospective recruitment (14 (28%) had misattributed codes and were not identified in the coding list. These patients were coded, for example, as unspecified acute upper respiratory tract infection (I06.9), dyspnœa (R06.0) and cough (R05), despite a clinical diagnosis of pneumonia. Among those identified by HES codes, pneumonia (N=5) and lower respiratory tract infection (N=2) were coded as secondary diagnoses. These figures suggest that reliance on primary diagnostic codes on the basis of HES data could underestimate the levels of pneumonia. There are no reasons to think that levels of miscoding has changed over time.

This article does not describe trends in bacterial pneumonia as stated throughout the paper but all causes of pneumonia. We suggest that use of HES data should be limited to analysis of changes in the overall incidence of pneumonia.

Authors’ response

Elemraid and colleagues raise important points about potential diagnostic misclassification and under-ascertainment using the Hospital Episodes Statistics (HES) database, as well as the absence of national pneumococcal serotype data for children.

The HES database covers all NHS hospital activity in England and has been widely used to report disease trends. It also provides the opportunity to estimate the clinical impact of major clinical policies on disease burden, but, with any large epidemiological dataset, has inherent weaknesses at the individual level.

Our study aimed to focus on common community-acquired bacterial pneumonia trends to evaluate the impact of the hepta-valent pneumococcal conjugate vaccine (PCV7). We agree that pneumonia is a clinical diagnosis and that it is difficult to differentiate between bacterial and viral causes. Furthermore, a significant proportion of cases (5–40%) of community-acquired pneumonia may be of mixed aetiology. Hence, we applied broad pneumonia definitions. We aimed to focus on bacterial pneumonias and so excluded specified pneumonias (eg ‘viral pneumonia, not elsewhere classified’—all J12 codes). The codes we searched are listed in the Appendix.

The authors highlight a useful point that some children diagnosed with pneumonia may have symptoms and/or signs recorded in the primary diagnosis field. Hence, we acknowledge under-ascertainment is possible for some pneumonia admissions. HES coding is dependent on the recording of the ‘reason for admission’ by clinicians and the subsequent coding by the trained staff, and we included this as a potential limitation in our discussion. We agree that such levels of miscoding are unlikely to have significantly changed over time. Therefore, this would suggest the pneumonia admission trends that we observed are likely to represent real changes.

We used the Health Protection Agency cumulative weekly incidence reports of PCV7 and non-PCV7 isolates for children under 5 years, together with the national serotype surveillance for all ages, as the best available source of information on pneumococcal serotypes causing invasive pneumococcal disease. Admittedly, this covers a broader spectrum of invasive diseases. In