Primary Care Respiratory Journal (2011); **20**(1): 4-5

Primary Care
RESPIRATORY JOURNAL
www.thepcrj.org

EDITORIAL

Approaches to enhancing the validity of coded data in electronic medical records

See linked article by Juhn *et al.* on page 79

*Dipak Kalra[a]
Bernard Fernando[b]

[a] Clinical Professor of Health Informatics Centre for Health Informatics and Multiprofessional Education, University College London
[b] General Practitioner and Honorary Research Fellow, eHealth Research Group, Centre for Population Health Sciences: GP Section, The University of Edinburgh

*Correspondence:
Dr Dipak Kalra
Clinical Professor of Health Informatics
Centre for Health Informatics and Multiprofessional Education
University College London
Holborn Union Building,
Highgate Hill, London N19 5LW
Tel: +44 (0)20 7288 5966
E-mail: d.kalra@ucl.ac.uk

Received 30/10/10

In the linked paper entitled "Characterisation of children's asthma status by ICD-9 code and criteria-based medical record review",[1] the authors have observed a discrepancy between the presence of clinical features compatible with the diagnosis of asthma in children's hospital electronic health records and the tagging of the record with a corresponding classification code. The authors' assertion that both patients and the hospital may be disadvantaged by failure to document the diagnosis formally raises some interesting questions.

The planning of health care services, reimbursement and public health measures undoubtedly depend upon robust metrics on the incidence, prevalence and distribution of disease. Some health care services, such as screening programmes, depend upon accurate disease registers. A formal clinical diagnosis is sometimes the trigger to invoke a guideline or care pathway, a specific medical intervention, or to engage a broader multidisciplinary clinical team. It is known that at times a diagnostic label is helpful to patients, providing a reassuring recognition of their illness and a label that they can use to describe their health situation to others.[2] Perhaps the most compelling benefit from accurate diagnoses is to improve the safety of prescribing through facilitating decision support checks for the inadvertent prescription of contra-indicated drugs.

However, many diagnostic labels span a wide spectrum of disease expressions, severities and management strategies. Medicine is, it has been argued, more keen to label patients than they are themselves to be labelled: patients sometimes feel that labelling them with a diagnosis feels like the start of factory processing, and that their individuality or personal story then goes unrecognised.[3,4] Given an increase in interest (and potential value) in personalised medicine, the promotion of coarse-grained diagnostic labelling (coding) needs to be sensitive to the benefits it genuinely offers. It is especially important to distinguish the value to patients of clinical coding from the benefits that an organisation may gain from more management-oriented coding systems – such as the NHS data dictionary used for Hospital Episode Statistics.

The other side of the coin is the benefits that coding offers research.[5,6] Juhn *et al.*'s concluding assertion is that ascertainment of asthma by physician diagnosis (i.e., ICD codes) alone would underestimate the prevalence of asthma in an epidemiological study.[1] This questions the validity of ICD codes in the asthma records examined and the completeness of the coding process. A valid ICD code essentially means that it accurately represents what is expected by the user of the code – and in this case, all people with true asthma are tagged with expected ICD codes and no one else.[7,8] Although valid codes are always desired, they are often difficult to achieve for many reasons. Mechanics of clinical coding show us three stages of the coding process where the validity of codes can be compromised. First is the 'recognition stage', when the problem is recognised. Accuracy of this process is compromised when there is no consensus on diagnosis or when an error in diagnosis is made. Second is the 'naming stage', when the diagnosis is named (labelled) accurately. Finally, in the 'matching stage', the label is matched with a term in a computerised pick list. A term which matches exactly might be unavailable or difficult to find in the pick list. The approaches to enhancing the validity of coded data, therefore, could be as wide ranging as reaching consensus on diagnosis, promoting accurate diagnostic practices and use of correct terminology, establishing record keeping standards, and the design of user friendly coding tools.

Above all, establishing an ongoing process to review records (for completeness), recheck codes and correct errors and omissions (for validity) will have a greater short term impact than addressing the above mentioned approaches. This is especially the case when coding is incentivised. This was seen when the NHS financial reforms were introduced by way of "Payment by Results" (PbR) in hospitals[9] and the "Quality and Outcomes Framework" (QOF) in family practice. This also has been the case in some research databases such as the GPRD where coding is incentivised.[10] There is, of course, a downside to incentives, but in practice there is as yet no good evidence for "gaming" and "upcoding".[9]

Acknowledgements

We acknowledge the support of the NHS Connecting for Health Evaluation Programme (NHS CFHEP) and for the grant received for an evaluation of different levels of structuring within the clinical record (NHS CFHEP 009).

Conflict of interest declaration

None declared.

References

1. Juhn Y, Kung Amiinah, Voigt R, Johnson S. Characterisation of children's asthma status by ICD-9 code and criteria-based medical record review. *Prim Care Resp J* 2011;**20**(1):79-83. http://dx.doi.org/10.4104/pcrj.2010.00076
2. Wright A, Jorm AF, Harris MG, McGorry PD. What's in a name? Is accurate recognition and labelling of mental disorders by young people associated with better help-seeking and treatment preferences. *Soc Psychiatry Psychiatr Epidemiol* 2007;**42**(3):244-50. http://dx.doi.org/10.1007/s00127-006-0156-x
3. Hurwitz B. Narrative and the practice of medicine. *The Lancet* 2000;**356**(9247): 2086-9. http://dx.doi.org/10.1016/S0140-6736(00)03412-7
4. Edwards M. Doctors 'over-diagnose'. The Daily Telegraph, 20 September 2010. Available from http://www.telegraph.co.uk/health/8008608/Doctors-over-diagnose.html (Last accessed 10 Oct 2010)
5. Simpson CR, Newton J, Hippisley-Cox J, Sheikh A. Incidence and prevalence of multiple allergic disorders recorded in a national primary care database. *J R Soc Med* 2008;**101**(11):558-63. http://dx.doi.org/10.1258/jrsm.2008.080196
6. Hippisley-Cox J, Coupland C,Robson R, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore *BMJ*. 2009;**338**:b880. Published online 2009 March 16. http://dx.doi.org/10.1136/bmj.b880.
7. Thiru K, Hassey A, Sullivan F. Systematic review of scope and quality of electronic patient data in primary care. *BMJ* 2003;**326**(7398):1070.
8. Jordan K, Porcheret M, Croft P. Quality of morbidity coding in general practice computerised medical records: a systematic review. *Family Practice* 2004;**21**: 396-412. http://dx.doi.org/10.1093/fampra/cmh409
9. Improving data quality in the NHS. Annual report on the PbR assurance programme. Health 2010. Audit Commission 2010.
10. Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. *Br J Gen Pract* 2010;**60**(572): e128-36. http://dx.doi.org/10.3399/bjgp10X483562