

Short report

Can we trust electronic health records? The smoking test for commission errors

Pablo Millares Martin

Whitehall Surgery, Leeds, UK

ABSTRACT

Cite this article: Millares Martin P. Can we trust electronic health records? The smoking test for commission errors. *J Innov Health Inform.* 2018;25(2):105–108.

<http://dx.doi.org/10.14236/jhi.v25i2.970>

Copyright © 2018 The Author(s). Published by BCS, The Chartered Institute for IT under Creative Commons license <http://creativecommons.org/licenses/by/4.0/>

Author address for correspondence:

Pablo Millares Martin
Whitehall Surgery
Wortley Beck Health Centre
Leeds LS12 5SG, UK
Email: pablo.martin@nhs.net

Accepted May 2018

Background Considerable interest exists on using general practice electronic health records (EHRs) for research and other uses. There is also concern on their quality.

Aim We suggest a simple test to assess errors of commission and subsequently overall EHR data quality that can be done on a periodical basis.

Method Patient records with simultaneous entries of three different stages on smoking were studied. The codes ‘never smoked tobacco’, ‘smoker’ and ‘ex-smoker’ should follow this chronological order. It should then be possible to extrapolate the overall level of errors of commission for the organisation.

Results The smoking test in our sample found errors in 169 patients, with 60 cases where dual errors were discovered. We express it as an estimated error of commission level of 2.6% related to the total population of the practice.

Conclusions Considering the constant and regular entries on smoking status (83.59% of the entries were done over last month), we can conclude smoking entries analysis can serve as a simple test to periodically assess the overall EHR data quality, and any trends.

Keywords: data quality, electronic health records, commission error

INTRODUCTION

General practitioners have been using electronic health records (EHRs) for decades. During this time, records that were only for the use of single organisations have become open to others, such as community services and out of hours services, for contribution. Furthermore, coded data are now used for many other purposes, such as research, health service planning, and epidemiological studies but data quality is not optimal.¹ Many factors could influence it, among them diagnoses being recorded solely in the free text fields,² conditions with clear diagnostic features being better recorded than conditions with more subjective criteria,³ lack of incentives and the effort required to keep records quality,⁴ etc.

Burnum⁵ in 1989 already pointed out the fact that while informatics improve, the information transferred worsens, and it seems we are still in the same conundrum.

Is there a quick way to assess overall quality of records? We considered using a common code, one routinely entered in different specialties and by different staff, which can give an impression on the overall quality of records: smoking status.

Routinely recorded smoking status validity in general practice is considered good,⁶ mainly after 2004 with the increase in its recording secondary to the introduction of quality and outcomes framework and the incentive to code it periodically.⁷ There are many ways of coding the smoking status of a patient, and furthermore, it is a position that can change through time; basically, three stages can be recognised: never smoked, smoker and ex-smoker, although other codes are used and misclassification of ex-smokers as non-smokers is likely.⁶

We consider a simple way to assess the quality of record entries could be to assess the order of these three stages on smoking. It is not about whether an entry exists, but looking into the accuracy of the entry made. There is a clear chronological order on the three stages (codes) on smoking considered. Entering 'never smoked' after 'ex-smoker' or 'smoker' is clearly an error of commission.

In consequence, if there were a significant number of patient records with smoking status errors, the data quality provided by the organisation in question in its electronic records would be poor. Smoking status is probably the most common entry found, and if this type of entry is wrong, it can be expected that other coded diagnoses and entries will also be wrong.

METHODS

A simple audit was conducted among patients registered in our practice. We limited the study to patients with all three codes ('never smoked tobacco', 'smoker' and 'ex-smoker') present on their records. We extracted the dates of the last entries of the three codes in question. Two variables were created, comparing where 'never smoked tobacco' date was before or after 'ex-smoker' and 'smoker' dates. We looked only at the last entries for each code. The purpose of the test is to assess current coding by looking at how entries are

taking place, and in consequence, historical values are less relevant. Furthermore, it is likely there are several entries where the coding is wrong for a patient. Furthermore, it is likely there are several entries where the coding is wrong for a patient. As we are aiming to create a simple and easy-to-use tool to monitor coding quality, looking into all the smoking status entered in a record could prove to be difficult.

We also analysed how many records contained a smoking code, to put into perspective the level of errors among the total number of patients with entries regarding smoking. The test is basically planned to be done on assessing three codes, as a simple tool is intended, and the more it can be simplified, the quicker it can be done and the more likely it will be implemented. Finally, records with the three codes present are the only ones assessed, reducing the number of records that need to be analysed, contributing to the speed of the test, which could be relevant when studying large organisations, and a point to be considered if the aim is to be able to relate different size organisations, as comparable values would be provided.

RESULTS

There are 8813 patients in our sample, 1547 of them under 12 years of age. When looking at individual codes, there are 4669 cases with 'never smoked tobacco', there are 2469 patients with 'ex-smoker' code and finally there are 2432 individuals with 'smoker' code. It is also relevant that 6909 patients have some sort of smoking code, indicating that 357 children have a smoking code in their records.

Records of 390 patients included the three codes we are interested in analysing in more detail (Figure 1). In total, 169 records (43.3%) contained errors: In 141 cases (36.15%), a 'never smoked' entry was made after a 'smoker' entry; in 88 cases (22.56%), it was entered after an 'ex-smoker' code; in 60 cases (15.38%), dual errors were present in patients' records. There are 229 errors in total among 169 patients. Patients' demographics were not considered, as the test is about record entries, and not about health status of the population.

As part of the relevance of the codes to assess continuously quality of records, we checked how many records had an entry of one of these three codes in the last month: 326 cases.

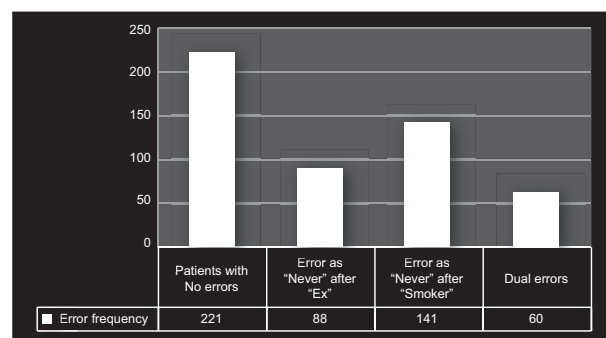


Figure 1 Smoking test results

DISCUSSION

A simple methodology, as demonstrated, can prove to the organisation or to external assessors of the institution if the data quality is good compared to others. This is a simple test to get the numbers needed for a snapshot of data quality. It has its limitations, as more complex patterns are concealed and the error level could potentially be somehow higher. Nevertheless, it is the first simple tool to our knowledge to assess commission error that can be repeated on regular basis, as smoking status is one of the pieces of information most regularly entered in records, demonstrated by the fact that in our sample 83.59% of the entries were done over the previous month.

It seems there is considerable error from the above numbers, but 229 inconsistencies represent just a small proportion of the practice population. We could express it as the number of patients affected in absolute values, 169, but we could also argue that considering the number of inconsistencies is larger, the patients with dual errors should also be counted twice. In consequence, we can state that commission errors occur in 2.6% of cases as we consider the level of errors encountered related to the whole organisation's population.

Other ways of measuring the percent could be considered. We could actually argue that the commission error occurred in 3.62% of patients with the code of 'never smoking' (4669 individuals), but there will be many patients who never had that code as the first time an entry was made they were already in a different category. We could also describe it at the most basic level, stating that among the 390 patients with the three codes, 43.33% contained errors. It could also be misleading; suggesting a high level of error when actually a very small sample of the practice contained all three codes, and it does suggest the expectation is to have the three codes in the majority of patients; actually large populations are 'never smoker' and remain as such.

We could also consider the fact we limited our audit to last entries of each of the codes, and that multiple errors could have occurred in the same patient entry. It will be very complex to carry on such an analysis of each smoking entry on each EHR. We accept it is a limitation on the study, but the

simplicity of the analysis can be considered a strength as it can be done easily by researchers interested in a quick assessment of the overall quality of the records.

An assessment on why the data might be inconsistent is beyond the purpose of the test. It is for the organisation to determine the roots; we presume the most likely reason is human error, miscommunication between patient and clinician, not being able to automatically crosscheck with past entries on the matter.

We can consider the level of commission error to be low, and considering the high level of entries on these particular codes over the last month (83.59%), we can also conclude smoking status can serve as a barometer of the quality of records in general practice. It is a test that can be repeated in a few months and results can easily be different because last entries are likely to be different to the ones assessed on previous occasion.

CONCLUSIONS

We can conclude the smoking test is a simple and effective way to extrapolate from smoking entries analysis the overall data quality of EHR. It can also show any trends, as smoking status entries are constantly entered on records, as we demonstrated in our sample.

We suggest using a percent value as the final result, linking the percent of commission error noted to the total population rather than specific subgroups analysed as a way to facilitate its interpretation.

Funding

None.

Ethical approval

Not required as internal clinical audit used anonymised data.

Competing Interests

None.

REFERENCES

1. Liaw ST, Powell-Davies G, Pearce C, Britt H, McGlynn L and Harris MF. Optimising the use of observational electronic health record data: current issues, evolving opportunities, strategies and scope for collaboration. *Australian Family Physician* 2016;45(3):153. PMID:27052055.
2. Coleman N, Halas G, Peeler W, Casaclang N, Williamson T and Katz A. From patient care to research: a validation study examining the factors contributing to data quality in a primary care electronic medical record database. *BMC Family Practice* 2015;16(1):11. Available from: <https://doi.org/10.1186/s12875-015-0223-z>. PMID:25649201; PMCid:PMC4324413.
3. Khan NF, Harrison SE and Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. *British Journal of General Practice* 2010;60(572):e128–36. Available from: <https://doi.org/10.3399/bjgp10X483562>. PMID:20202356; PMCid:PMC2828861.
4. Orueta JF, Nuño-Solinis R, Mateos M, Vergara I, Grandes G and Esnaola S. Monitoring the prevalence of chronic conditions: which data should we use? *BMC Health Services Research* 2012;12(1):365. Available from: <https://doi.org/10.1186/1472-6963-12-365>. PMID:23088761; PMCid:PMC3529101.

5. Burnum JF. The misinformation era: the fall of the medical record. *Annals of Internal Medicine* 1989;110(6):482–4. Available from: <https://doi.org/10.7326/0003-4819-110-6-482>.
6. Marston L, Carpenter JR, Walters KR, Morris RW, Nazareth I, White IR, *et al*. Smoker, ex-smoker or non-smoker? The validity of routinely recorded smoking status in UK primary care: a cross-sectional study. *BMJ Open* 2014;4:e004958. doi: 10.1136/bmjopen-2014-004958. Available from: <https://doi.org/10.1136/bmjopen-2014-004958>.
7. Taggar JS, Coleman T, Lewis S and Szatkowski L. The impact of the Quality and Outcomes Framework (QOF) on the recording of smoking targets in primary care medical records: cross-sectional analyses from The Health Improvement Network (THIN) database. *BMC Public Health* 2012;12:329. Available from: <https://doi.org/10.1186/1471-2458-12-329>. PMID:22559290; PMCID:PMC4104830.