ORIGINAL PAPER

Consultation analysis: use of free text versus coded text

Pablo Millares Martin[1]

Abstract

General practice in the United Kingdom has been using electronic health records for over two decades, but coding clinical information remains poor. Lack of interest and training are considerable barriers preventing code use levels improvement. Tailored training could be the way forward, to break barriers in the uptake of coding; to do so it is paramount to understand coding use of the particular clinicians, to recognise their needs. It should be possible to easily assess text quantity and quality in medical consultations. A tool to measure these parameters, which could be used to tailor training needs and assess change, is demonstrated. The tool is presented and a preliminary study using a randomised sample of five recent consultations from thirteen different clinicians is used as an example. The tool, based on using a word processor and a spread-sheet, allowed quantitative analysis among clinicians while word clouds permitted a qualitative comparison between coded and free text. The average amount of free text per consultation was 68.2 words, (ranging from 25.4 and 130.2 among clinicians); an average of 6% of the text was coded (ranging from 0 to 13%). Patterns among clinicians could be identified. Using Word cloud, a different text use was demonstrated depending on its purpose. Some free text could be turned into code but nomenclature probably prevented some of the codings, like the expression of time. This proof of concept demonstrated that it is possible to calculate what percentage of consultations are coded and what codes are used. This allowed understanding clinicians' preferences; training needs and gaps in nomenclature.

Keywords Clinical coding · Electronic health record · Family practice · General practice · Records · Systematized Nomenclature of Medicine

1 Background

General practitioners (GPs) in the United Kingdom (UK) have been fully computerized for over two decades [1]. There has been a recent change towards a single coding terminology: The "Systematized Nomenclature of Medicine—Clinical Terms" (SNOMED-CT) from previously used languages aiming to facilitate interoperability [2]. SNOMED-CT is considered to be able to express over 90% of the information entered in medical records as part of the Problem Lists [3], but it will require clinicians to aim to code at these levels. It is clear that "like all computer users, GPs are fundamentally lazy and mildly computer-phobic" [4], furthermore they are poor coders [5], and they lack the time and training they consider needed for it [6]. As a consequence, Electronic health records do not have enough coded entries; this affects the ability to run reports, to facilitate audit and research activities; it can lead to biased results and affect health care.

Much research on coding has been done in hospitals, where clinical coders are used, and considered to be expensive, inefficient and prone to error [7]. In consequence algorithms have been created to reduce costs and automatize the process [8] but taking into account "No free lunch" theorems [9] they cannot be the answer. As stated by Wolpert "For any two learning algorithms A and B […] there are just as many situations (appropriately weighted) in which algorithm A is superior to algorithm B as vice versa" [9].

Secondary sources of information can never be as good as primary sources. It is to be considered also that codes which are infrequent, misspellings and abbreviations –some non-standard- are major limitations [10, 11]. Coding must be embedded in clinical practice; it already happens in primary care in the UK [12] but not enough [13]; code nomenclature is not clearly understood, and time to input and training to feel confident in their use are challenges [6].

✉ Pablo Millares Martin
Pablo.martin@nhs.net

1 Whitehall Surgery, Wortley Beck Health Centre, Leeds LS12 5SG, UK

To improve clinical coding, to engage clinicians coding their consultations it is paramount to understand how much they are coding, what are they coding and what they are not. To determine the level of coding done by particular clinicians or organizations a reliable methodology is needed. So questions like how much coding happens currently in general practice? How much coding language is present when related to non-coded language? How much is prevented by the nomenclature itself? Or are there different purposes identified for coded and free text? Can be answered.

The aim of this paper is to describe a simple assessment tool and to do a preliminary study with a sample of consultations to demonstrate its use and benefits.

2 Methods

A small sample of randomised consultations was envisaged to test a method to answer the research questions. The purpose was demonstrating how compiling consultation data and manipulating it allowed carrying out quantitative and qualitative analysis:

- To calculate the number of words of clinical coding opposed to free text,
- To assess if entries made as free text could be coded using SNOMED-CT.
- To determine most common words present in the consultations, as coded and as free text.

The assessment tool was the use of current digital capabilities (access to electronic records, a word processor and a word cloud generator) to manually present data in a way that would allow an analysis of the use of coded data opposite to free text during consultations.

When considering the amount of text (free and coded) that is present in a consultation, the word count needs to be unbiased, mainly when expecting codes and abbreviations to be present, and furthermore when grammatical errors are likely to be recorded, affecting what a person could consider a word. There is nevertheless a simple way to remove this bias: Using a word processor. The software will not interpret the content, but assess the word count following its algorithms. For example, entering a measure of blood pressure (BP) "BP 140/80" is two words while "BP 140 / 80" is four. A human is more likely to consider it as three words in both scenarios.

It is possible from the clinical software to select any consultation to be printed, with or without showing the codes in use (SNOMED codes were selected to be shown), automatically creating a word document whose content could be saved and analysed. It will contain free text and coded data (see Fig. 1). Multi-selecting sections, both types

of entries can be separated, and counted. It is of note the software names the different sections of the consultation, and that there are transactions added automatically as code (like prescriptions, sick notes and updates in the system). These instances were not counted (in bold in Fig. 1), as the user cannot opt for free text, neither there is code under SNOMED-CT for drug names.

When considering free text present, it is possible then, through clinical software tools, namely its code browser to look for possible coding solutions, to assess their merit.

By selecting a small number of randomized consultations for each clinician to be assessed, an emerging picture of their coding style is expected, within a short time.

Equally, by using one of the available word cloud tools (For this paper wordart.com was used) it was possible to easily create graphical representations that "allows a viewer to form a quick, intuitive sense" [14] of the consultation text. Due to their nature, comparisons of two word clouds are difficult, and even impossible for simultaneously assessing four or more clouds [15]. The comparison was to be made between the total coded and free text collected, to analyse if different purposes could be identified.

For this exercise, the patient records of our organization -which we open routinely for care provision- were assessed. Thirteen general practitioners were identified among principals, employed, locums and also from our hub (a different organization that provides care for six practices, including ours, during hours we are routinely closed). Authors were excluded from the analysis.

3 Results

The aggregated findings of five consultations per general practitioner were used to define consultation style based in the average of free text and coding as well as where and how in the notes the codes were entered. The exercise quantitative findings are summarized in Table 1.

The average amount of free text per consultation was 68.2 words, ranging from 25.4 and 130.2 depending on clinician. By contrast, the average amount of text in coded format was 3.9 words, ranging from none to 9.6 words. In consequence, among our patients' electronic health records only an average of 6% of text was coded, and ranged among clinicians from 0 to 13%.

When the amount of individual codes was considered it was worrying only 76 codes were used in the 65 consultations analysed, an average of 1.2 codes per consultation; the average by GPs ranged from 0 to 2.

The most common codes were those related to diagnosis, in average of 0.5, so one code for each two consultations, followed by codes added in taking the history (average 0.3) (See Fig. 2). Entries done directly, through templates or

Fig. 1 GP consultation, as extracted, and with separated text and code for word counting

```
DD MMM YYYY HH:MM                          |                  GP3 consultation 1
     History: seen with mum
     looks well in self
     gaining weight along own centile
     lots of colick, wind, discomfort
     on HV mum tried cmomfort milk - started refusing feeds and has
     developed widespread rash
     now back on prev. formula
     no atopy in family
     has had some looser stool with comfort milk - usual tendency to
     cobstipation
     Examination: brigth and and alert
     widespread eczema / dermatitis to abdomen, neck folds
     no infective featrures
     Abdomen examined - NAD (SNOMED: 163133003)
     Diagnosis: Eczema NOS (SNOMED: 43116000)
     Plan: cream all over, steroid to very inflamed affected areas
     discussed amounts and durations
     trial of SMA LF
     seeing HV next week and check progress then
     review sooner with concerns.
     Oilatum cream (Thornton & Ross Ltd) - 150 gram - apply as needed
     Hydrocortisone 1% cream - 30 gram - apply as directed
     ETP FP10: Printed On dow dd mmm yyyy hh:mm By GP3
     Summary Care Record Update
```

DATA FREE TEXT 98 words

```
seen with mum
looks well in self
gaining weight along own centile
lots of colick, wind, discomfort
on HV mum tried cmomfort milk - started refusing feeds and has
developed widespread rash
now back on prev. formula
no atopy in family
has had some looser stool with comfort milk - usual tendency to
cobstipation
brigth and and alert
widespread eczema / dermatitis to abdomen, neck folds
no infective featrures
cream all over, steroid to very inflamed affected areas
discussed amounts and durations
trial of SMA LF
seeing HV next week and check progress then
review sooner with concerns.
```

DATA CODED 6 words (Excluding SNOMED code)

```
Abdomen examined - NAD (SNOMED: 163133003)
Eczema NOS(SNOMED: 43116000)
```

searching for the code itself was next, with an average of 0.2. As expected non numeric codes dominated (average 1 versus 0.2 for numeric codes).

A second aspect that could be observed was the amount of text that could have been entered as clinical coding. It was clear that Problems Lists were not commonly used (2 entries among 65 consultations), while in 30 patients diagnoses were coded. An unexpected finding was the fact two clinicians (GP4 and GP10) did not enter any diagnostic text (coded or not) in the section "Diagnosis", furthermore GP9 only used once "Diagnosis" to enter what should be

considered a plan ("*Refer to hearing aid clinic*" (SNOMED: 183,853,009)). By contrast, GP11 entered a coded diagnosis for each consultation, while GP5 and GP12 coded 4 of the 5 diagnoses entered.

Coding does not mean using nomenclature properly, as shown in GP12-patient2 entry: *?"Chronic suppurative otitis media*" (SNOMED: 38,394,007); any report for any purpose will not consider the diagnosis as suspected, but as confirmed.

Differences on codes used for same issue were also apparent, as "*Cystitis*" (SNOMED: 38,822,007),

Table 1 Average Findings among 5 random consultations per clinician

Clinician ID and role		Average Word count (code & text)		Average Point of entry of codes in the electronic health record						Average Number of codes and types			Average Percentage coded
ID	Role	Free-text	Coded text	History codes	Exam codes	Diagnoses codes	Plan codes	Problem codes	Direct/template entries	Not numeric codes	Numeric codes	Number of codes	
1	Hub	25.4	4	0	0.4	0.6	0.2	0	0	1.2	0	1.2	13%
2	Hub	51.4	1.6	0	0	0.2	0	0	0.4	0.6	0	0.6	5%
3	Salaried	79.6	7.4	0.8	0.6	0.6	0	0	0	1.8	0.2	2	9%
4	Locum	77.4	0.8	0	0.2	0	0	0	0	0.2	0	0.2	1%
5	Salaried	93.6	3.4	0	0	1.2	0.4	0	0	1.6	0	1.6	4%
6	Hub	58	3.6	0.2	0	0.4	0	0	0.6	0.8	0.4	1.2	7%
7	Partner	51.8	8	1	0.2	0.2	0	0	0.4	1.2	0.6	1.8	10%
8	Salaried	68.8	3.4	0.8	0	0.4	0	0.2	0.4	1.8	0	1.8	5%
9	Partner	35	1	0	0	0.2	0	0	0	0.2	0	0.2	3%
10	Salaried	130.2	0	0	0	0	0	0	0	0	0	0	0%
11	Locum	86.8	9.6	0	0.2	1.2	0	0	0.6	1.2	0.8	2	8%
12	Hub	58.8	1.6	0	0	0.8	0	0	0	0.8	0	0.8	3%
13	Salaried	69.8	6	1	0	0.2	0.2	0.2	0.2	1.6	0.2	1.8	11%
	Average	68.2	3.9	0.3	0.1	0.5	0.1	0.0	0.2	1.0	0.2	1.2	6%
	Min	25.4	0	0	0	0	0	0	0	0	0	0	0
	Max	130.2	9.6	1	0.6	1.2	0.4	0.2	0.6	1.8	0.8	2	13%

Fig. 2 Average number of codes in the different sections of consultation
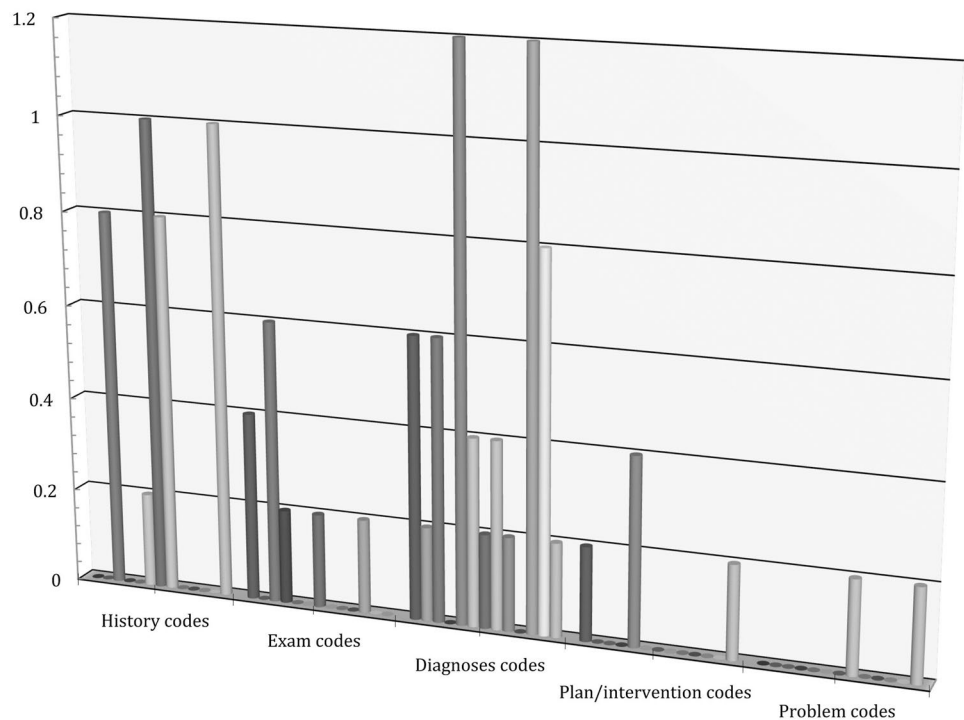
"*Urinary tract infectious disease*" (SNOMED: 68,566,005) and "*Suspected UTI (Urinary tract infection)*" (SNOMED: 314,940,005) were used for three different patients' diagnoses. It is of note the parents for these entries are also distinct, and in consequence reporting on this matter could be also affected if not carefully planned.

Coding nomenclature available for all data entered by free text in other sections of the consultation was not assessed. This paper is focused on the methodology, and examples above sufficed to describe benefits of a fast assessment tool showing code use and even possible shortcomings of the nomenclature and software.

It should be expected that SNOMED-CT links among different codes, like a "*Suspected*" (SNOMED: 415,684,004) added to a diagnosis to clarify the matter, would be handed adequately by the clinical software but it is not the case.

Another matter explored was why there was not a code for every diagnosis; considering the limitation above described, it could explain eight diagnoses entered as free text as they were all suspected, and with no single code to express it, opposite to three cases where a "*Suspected COVID-19*" (SNOMED: 1,240,761,000,000,102), "*Suspected infectious disease*" (SNOMED: 473,130,003) or "*Suspected mumps*" (SNOMED: 1,087,741,000,000,101)" could have been used. In a further instance, there was no single code available to indicate "convalescence from chest infection"; finally the diagnosis for the free text "anaemia from regular venesection", which could be perhaps "iatrogenic anaemia" was neither available.

Finally, two word clouds were created, one based on all free text collected (see Fig. 3), and the other from all the words included in the codes used (see Fig. 4). Numbers were excluded from the analysis. It is clear that different concepts are in each cloud. Free text has as top hit "Week", a concept of time (linked to the duration of symptoms), which is not possible to code in the current nomenclature. Following it, the concepts of "pain", a common symptom that is possible to code easily, and with multiple options available, and "left", an indication of lateralisation in most instances, and as well part of the history section of the consultation. They are followed by "Advise" and "See", words indicating a plan of action. In contrast, the coded top five words are "infection", "Nos" (indicating "Not Otherwise Specified"), "Respiratory" and "tract", words of a diagnosis, and "kg", indicating weight measurement (kilograms).

The individual use of these free-texts could then be assessed based on clinicians. For example, regarding plans, GPs tended to write a safety net for patients to return: GP1 preferred "See SOS", GP3 used "See back" or "See sooner" while GP4, GP6, GP8 and GP13 tended to use "Review". GP11 wrote a sentence beginning with "If" and detailing action to take then. GP 12 favoured "Seek review" and "r.v.". All those entries could be recorded more quickly coded with "Advice to return if problem persists or deteriorates" (SNOMED: 927,621,000,000,101).

Word	Occurrences
Week	27
Pain	26
Left	24
Advise	24
See	22
Review	22
Throat	20
Mum	20
Infected	19
Now	19
Cough	19
Ear	18
Still	17
Back	17
Day	17
Last	16
Time	16
Feels	16
Use	16

Fig. 3 Word cloud of free text, with list of most common words encountered

4 Discussion

The paper suggests a new approach, a new tool to assess electronic health records content (quantity and quality of coding). It is expected this novel approach should facilitate focused training and eventually an improvement on coding use, a necessity not just for clarity of records (although it could do if abbreviations used by clinicians -like "r.v"- are changed to coded text) but more importantly for their secondary purpose: auditing and research.

4.1 Summary

Assessing clinician style on using the clinical software could help to tailor the training needs of different individuals and it could ultimately increase the uptake of clinical coding in an organization. The tool described has shown how a small random sample of consultations could clearly differentiate ways clinicians enter their notes in the electronic health records (EHR), and also on their use of the coding nomenclature. Studies tend to focus on attitudes of clinicians towards

electronic health records (EHR) and clinical coding [6, 16–21] as well as on training needs [6, 10, 17, 18, 22] but analysis of the consultations themselves are needed, and for that a common method of assessment is required. By comparing the amount of text present in relation to the amount of code it should be possible to evaluate changes happening as consequence of additional training been provided for example.

A second aspect is to allow finding reasons for text to be preferred instead of coding. SNOMED-CT could be able to describe almost every diagnosis [3] but to do so at times the answer is a top code with little clarity, like "*Disorder of menstruation*" (SNOMED: 386,804,004); or not clearly separating a diagnosis from a symptom, like "*Hearing loss*" (SNOMED: 151,880,001) as it could be a temporary symptom or a long term condition. Furthermore, clinical software restrictions prevent clinicians using linked codes ("*Suspected*" (SNOMED: 415,684,004) added to any diagnosis for example) and in consequence Problem Lists get negatively affected. It seemed clinicians failed even to look for those conditions were "suspected" formed part of a single code (like "*Suspected COVID-19*" (1,240,761,000,000,102) and written in free text instead ("Suspected COVID").

Word	Occurrences
Infection	7
NOS	7
Respiratory	5
Tract	5
Kg	5
Upper	4
New	4
Review	4
Issued	4
Fit	4
Work	4
MmHg	4
Weight	4
Pain	3
Depression	3
Emed	3
Statement	3
Observation	3
St	3

Fig. 4 Word cloud of coded text, with list of most common words encountered

If steps are taking encouraging clinicians to use codes (with additional training as needed), starting with diagnoses and Problem Lists, progressing towards coded symptoms and examination (improving nomenclature and software as required) EHR will eventually benefit from the power of informatics (To have alerts and other assisted diagnostic tools, to have data quality that supports adequately research); there is a need for clinical coding to be not just embedded in practice but owned by clinicians.

The assessment tool described is simple, and the example in a small organization, focused in diagnoses is limited, but in the same way it is easily reproducible for larger institutions and more thorough assessment of codes versus free text is also possible. In these circumstances, it should be considered the possibility of a "Plan, do, study and act" (PDSA) cycle, similar to other PDSA processes used to improve clinical coding [23].

In the end, a major limitation to improve code use is the software itself. When a clinician writes text and the software suggests different codes replacing it only an adequate and responsive natural language recognition would work. If the clinician needs to access a code browser for every item to be included in the record, the process can become a barrier itself. As in the examples suggested above a conversation among users could trigger sharing some appropriate codes to replace free text, but their access needs to be straightforward or risk getting forgotten. Furthermore, to improve coding use, code suggestions to entered text could follow a Whale Optimization Algorithm (WOA) [24] or a Particle Swarm Optimization (PSO) [25] protocol, but these software developments are beyond the aims of this paper.

To improve coding use, several actions are needed. As mentioned above, the software could provide better suggestions when free text was added, but it could also facilitate review of entries already in place, and whether creating word clouds or showing repeated free-text sections, it would assist clinicians with training or even self-training. Medical nomenclature is in itself a different language, and clinicians need to be proficient in it, to express their consultations in a way that is beneficial to others.

4.2 Strengths and limitations

An easy and fast way of analysis of consultations is possible using the technique described, it allows assessing not only differences among clinicians, but even within the organization as a whole, facilitating the finding of gaps on nomenclature and its use.

The technique has its limitations. Although it can provide a quick assessment, it is not exhaustive. It does not provide information on the reasons behind the type of text of preference found, nor is can link free text to possible codes.

This analysis was done manually, and the amount of manpower needed to run it in a larger organization could be considerable. In consequence, its benefit could be limited unless software allowing automated data mining and processing is used, so the process is more efficient.

Considering the poor use of code among clinicians [5] and that important information for clinical decision-making is only recorded in free text [26], it is still a way to reflect, and to monitor improvements. It is vital to improving clinical coding accuracy, which is paramount for audit, benchmarking and surveillance [27], and more tools are needed to help in this purpose.

5 Conclusions

It is possible to quickly assess the amount of clinical coding in electronic health records and to determine patterns among clinicians using the described tools. It is possible to target training and to assess progress in the use of clinical coding by repeating the test. It is also possible to find limitations from nomenclature and clinical software to be addressed.

Clinical coding is poor and action is needed to improve it. A pathway has been shown to facilitate the quantitative and qualitative analysis of electronic health records, but from the use of the tool, it has been suggested some software changes that could make the process of scrutiny better.

The small sample used in this preliminary study is not valid to conclude determinants in clinical coding in primary care, but it shows a technique that could be used to target training as well as to assess larger organisations on the way clinical coding is used in a methodical approach.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval No ethical approval required as only patients under our care were accessed.

References

1. Benson T. Why general practitioners use computers and hospital doctors do not–Part 1: incentives. BMJ. 2002;325(7372):1086–9. https://doi.org/10.1136/bmj.325.7372.1086.
2. NHS Digital, Snomed CT. 2019. Retrieved April 22nd 2020 from https://digital.nhs.uk/services/terminology-and-classifications/snomed-ct.
3. Elkin PL, Brown SH, Husser CS, Bauer BA, Wahner-Roedler D, Rosenbloom ST, Speroff T. Evaluation of the Content Coverage of SNOMED CT: Ability of SNOMED Clinical Terms to Represent Clinical Problem Lists. Mayo Clin Proc. 2006;81(6):741–8. https://doi.org/10.4065/81.6.741.
4. Benson T. The history of the Read codes: the inaugural James Read Memorial Lecture 2011. J Innov Health Inform. 2011;19(3):173–82. https://doi.org/10.14236/jhi.v19i3.811.
5. Liaw ST, Chen HY, Maneze D, Taggart J, Dennis S, Vagholkar S, Bunker J. Health reform: is routinely collected electronic information fit for purpose? Emerg Med Australas. 2012;24(1):57–63. https://doi.org/10.1111/j.1742-6723.2011.01486.x.
6. Millares Martin P, Sbaffi L. Electronic health records (EHR) and problem lists in Leeds, UK. Variability of general practitioners' views. Health Inform J. 2019. https://doi.org/10.1177/1460458219895184.
7. Catling F, Spithourakis GP, Riedel S. Towards automated clinical coding. Int J Med Informatics. 2018;120:50–61. https://doi.org/10.1016/j.ijmedinf.2018.09.021.
8. Kaur R. Distributed Knowledge Based Clinical Auto-Coding System. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop. 2019. https://doi.org/10.18653/v1/p19-2001.
9. Wolpert DH. The Lack of A Priori Distinctions Between Learning Algorithms. Neural Comput. 1996;8(7):1341–90. https://doi.org/10.1162/neco.1996.8.7.1341.
10. Mullenbach J, Wiegreffe S, Duke J, Sun J, Eisenstein J. Explainable Prediction of Medical Codes from Clinical Text. Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). 2018. https://doi.org/10.18653/v1/n18-1100.
11. Xie P, Xing EA. Neural Architecture for Automated ICD Coding. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Volume 1 (Long Papers). 2018. https://doi.org/10.18653/v1/p18-1098.
12. Pearce C, Mcleod A, Patrick J, Ferrigi J, Bainbridge MM, Rinehart N, Fragkoudi A. Coding and classifying GP data: the POLAR project. BMJ Health Care Inform. 2019;26(1). https://doi.org/10.1136/bmjhci-2019-100009.
13. Tate AR, Martin AGR, Ali A, Cassell JA. Using free text information to explore how and when GPs code a diagnosis of ovarian cancer: an observational study using primary care records of patients with ovarian cancer. BMJ Open. 2011. https://doi.org/10.1136/bmjopen-2010-000025.
14. DePaolo CA, Wilkinson K. Get your head into the clouds: Using word clouds for analyzing qualitative assessment data. TechTrends. 2014;58(3):38–44.
15. Silva e Silva LG, Assunção RM. COWORDS: a probabilistic model for multiple word clouds. J Appl Stat. 2018;45(15):2697–717.
16. Cabitza F. On the attitudes of GPs toward novel features of their next EPRs. Qual Life Through Qual Inf. 2012;180:911–91.
17. Doyle RJ, Wang N, Anthony D, Borkan J, Shield RR, Goldman RE. Computers in the examination room and the electronic health record: physicians' perceived impact on clinical encounters before and after full installation and implementation. Fam Pract. 2012;29(5):601–8. https://doi.org/10.1093/fampra/cms015.

18. Holanda A, e Sá HLDC, Vieira APGF, Catrib AMF. Use and satisfaction with electronic health record by primary care physicians in a health district in Brazil. J Med Syst. 2012;36(5):3141–3149. https://doi.org/10.1007/s10916-011-9801-3.

19. Makam AN, Lanham HJ, Batchelor K, Moran B, Howell-Stampley T, Kirk L, Halm EA. The good, the bad and the early adopters: providers attitudes about a common, commercial EHR. J Eval Clin Pract. 2013;20(1):36–42. https://doi.org/10.1111/jep.12076.

20. Secginli S, Erdogan S, Monsen KA. Attitudes of health professionals towards electronic health records in primary health care settings: a questionnaire survey. Inform Health Soc Care. 2014;39(1):15–32. https://doi.org/10.3109/17538157.2013.834342.

21. Bishop TF, Ryan MS, Mccullough CM, Shih SC, Casalino LP, Ryan AM. Do provider attitudes about electronic health records predict future electronic health record use? Healthc. 2015;3(1):5–11. https://doi.org/10.1016/j.hjdsi.2014.04.002.

22. Price M, Singer A, Kim J. Adopting electronic medical records: Are they just electronic paper records? Can Fam Physician. 2013;59:e322-329.

23. Abdulla S, Simon N, Woodhams K, et al. Improving the quality of clinical coding and payments through student doctor–coder collaborationin a tertiary haematology department. BMJ Open Qual. 2020;9:e000723. https://doi.org/10.1136/bmjoq-2019-000723.

24. Mirjalili S, Lewis A. The whale optimization algorithm. Adv Eng Softw. 2016;95:51–67.

25. Kennedy J, Eberhart R. Particle swarm optimization. In Proceedings of ICNN'95-International Conference on Neural Networks. 1995;4:1942–1948. IEEE.

26. Shah AD, Quinn NJ, Chaudhry A, et al. Recording problems and diagnoses in clinical care: developing guidance for healthcare professionals and system designers. BMJ Health Care Inform. 2019;26:e100106. https://doi.org/10.1136/bmjhci-2019-100106.

27. Priyatilake H, Kumar L, Appleton S. Implementing a cost-effective method of improving coding accuracy. Bull R Coll Surg Engl. 2019;101(6):238–41.