Word cloud analysis of the BJGP

A ‘word cloud’ is a visual representation of word frequency. The more commonly the term appears within the text being analysed, the larger the word appears in the image generated. Word clouds are increasingly being employed as a simple tool to identify the focus of written material. They have been used in politics, business and education, for example, to visualise the content of political speeches. In the Health Board that I support, word clouds have been applied to analyse the content of Board committee papers to see whether sufficient attention is being given to the core business of the organisation.

Word clouds should be interpreted with certain caveats. They often fail to group words that have the same or similar meaning; for example, ‘GP’ and ‘GPs’ or ‘Research’ and ‘research’. As they tend to focus only on single word frequency, they also do not identify phrases, reducing context.

A word cloud analysis of the entire content of the British Journal of General Practice from 2011, constituting 600 000 words, was conducted using the online programme Wordle (http://www.wordle.net/). A maximum word limit of 100 was set. Common English words were removed. The image generated can be seen on the cover of this Journal.

According to its editorial policy, the BJGP is ‘an international journal publishing articles of interest to primary care clinicians, researchers, and educators worldwide. Priority is given to research articles asking questions of direct relevance to patient care.’

The word cloud generated from the analysis was measured against the stated editorial policy of the Journal above. Firstly, it can be seen that the two most prominent words highlighted are ‘care’ and ‘patients’. This finding does conform to the aim of the Journal to focus on these areas. In terms of the Journal’s concentration on research articles, the words ‘study’ and ‘research’ as major terms both appear, again reflecting well the Journal’s policy. The target audience of the Journal is given as primary care clinicians, researchers and educators worldwide; the terms ‘GP’, ‘primary’, ‘general’, ‘practice’ and ‘clinical’ do emerge in the analysis. However, the word ‘education’ fails to materialise, although ‘training’ does appear, but only as a lower order term. In view of the international aim of the Journal, the analysis seems to suggest that its content is perhaps too weighted towards the home audience, with ‘UK’, ‘London’, and ‘NHS’ being the only geographic terms to appear in the word cloud.

Of interest, the medical conditions that figure most prominently in the analysis are ‘cancer’ and ‘depression’. The full extent of the patient journey is revealed, with ‘symptoms’, ‘diagnosis’, and ‘treatment’ all included. The words ‘time’ and ‘years’ are visible, perhaps reflecting the long-term nature of the GP–patient relationship. ‘Quality’ appears as a minor term, which is a surprise, given its importance in patient care; ‘risk’ appears larger. In terms of my own specialty, public health medicine, the words ‘social’, ‘screening’, ‘population’, and ‘need’ all make an appearance, but only as minor terms. In addition, there is no mention of ‘equity’ or ‘prevention’.

In conclusion, a word cloud analysis of the Journal has shown that it seems to have largely fulfilled its stated aim of ensuring that priority is given to research articles of direct relevance to patient care for primary care clinicians. However, it should, perhaps, reflect on whether it needs to become more international in outlook. Furthermore, the analysis has reflected the very diverse nature of primary care practice.

REFERENCES
