EDITORIAL

Assessing patients at risk of symptomatic-but-as-yet-undiagnosed cancer in primary care using information from patient records

Evidence arising from primary care electronic health records can help to assess the risk of symptomatic-but-as-yet-undiagnosed cancer. Existing evidence and methodological innovations in this field of study hold further promise for improving the diagnostic process and achieving earlier diagnosis in cancer patients.

MAIN

Most cancer patients are diagnosed after the onset of symptoms relating to their cancer.[1] Timely assessment of patients with symptomatic-but-as-yet-undiagnosed cancer can shorten intervals to diagnosis, helping to improve clinical outcomes and patient experience.[2,3] Sustained research efforts in understanding tumour biology have thus far not delivered the array of accurate, easily useable and cost-effective tests needed to support the diagnostic process in patients with possible cancer symptoms—most of whom will not have cancer. New tests are particularly needed to assist the diagnosis of 'harder-to-suspect' cancers, which most often present with symptoms of relatively low predictive value.[4] These realisations have, increasingly since 2006,[5] focussed research efforts on the understanding of the phenomic signature of symptomatic-but-as-yet-undiagnosed cancer using information included in 'real-world' electronic patient records data.

The *British Journal of Cancer* has recently published two studies of relevance to this expanding field. Nicholson et al.[6] describe the risk of cancer diagnosis following recorded weight loss in primary care consultations; they estimate this risk to be 1.2% by the end of a 3-month follow-up period, increasing to 1.8% by 12 months. Watson et al.[7] examined the risk of cancer following tests for inflammatory markers (such as erythrocyte sedimentation rate, C-reactive protein and plasma viscosity) in primary care; they estimate the risk to be 3.5% in patients with raised values by 12 months. Together with other studies estimating the predictive value of different symptoms and pre-diagnostic features, these estimates can support decision making at the clinical encounter and the production of clinical practice guidelines.[8] Beyond substantially enriching current knowledge, both studies highlight broader methodological issues pertinent to the estimation of the risk of cancer diagnosis based on information included in primary care electronic patient records.

Accurate measurement of exposure is a pre-requisite for any epidemiological study, and variation between doctors in the recording of possible cancer symptoms has long been described.[9] As clinicians may record symptom information in free-text notes differentially between cases and controls, risk estimates, particularly for 'alarm' symptoms with known relatively high predictive value for cancer,[10] may be artefactually inflated.[11] Nicholson et al.[12] should be commended for having previously validated symptom codes associated with unintended weight loss

in Clinical Practice Research Datalink data against objectively measured loss of weight in the same data source. Further, the authors of both commented studies appropriately compared the characteristics of patients with and without recorded weight or recorded inflammatory marker tests, as such analyses can help to assess the generalisability of their findings.[6,7]

Unlike information about the presence/absence of symptoms, information about the results of diagnostic tests could additionally allow for appreciation of 'dose–response' associations (e.g. using exact inflammatory marker,[7] or platelet count,[13] values), which could enable further individualising of risk estimates. Incorporating quantification of unintended weight loss in prediction models might be possible in the future, if the recording of patients' weight as part of primary care encounters was to become uniform and systematic.[14] Additionally, qualifying the nature (e.g. the intensity, persistence and duration) of presenting symptoms (such as weight loss) may be revealing given that they typically represent heterogeneous constructs. Nonetheless, the epidemiology of presenting symptoms in patients subsequently diagnosed with cancer remains understudied, and an area where qualitative inquiries are also needed to illuminate both the nature of symptoms and how they are appraised by patients.[15,16]

It has long been described that the risk of cancer after the first occurrence of relevant pre-diagnosed features is concentrated in the initial period post-presentation, declining rapidly thereafter.[17] Nonetheless, in prior literature predictive value estimates are generally 'averaged out' over relatively lengthy periods between first exposure and cancer diagnosis, of up to 1, 2 or 3 years.[17,18] Faced with this 'still picture vs. video' problem,[19] Nicholson et al.[6] provide time-dependent risk estimates for cancer diagnosis following consultations with unintended weight loss as a presenting symptom. They report that the hazard of cancer diagnosis was highest within the first 3 months of follow-up after this symptom was recorded, waning within 6 months or longer. The authors rightly argue that more accurately appreciation of the evolving levels of risk after the initial recording of a given feature could be useful.

The translation of information on the time-dependent nature of risk is however not straightforward, as the evidence relates to observational data from patient care records. Therefore, the time between symptom onset and diagnosis is influenced by both the natural history of the disease and decisions about investigative

management (the latter also being influenced by the presence/absence of practice recommendations included in clinical guidelines). Most cancer patients who presented with unexplained weight loss were diagnosed during the few months after this symptom was recorded.[6] Consequently, the potential to expedite the diagnosis of these patients may be limited, although it may be possible to improve their diagnostic pathways (e.g. reducing the risk of diagnosis of cancer through an emergency presentation, a diagnostic route associated with worse prognosis[1]). There is the potential that watchful waiting, as discussed by the authors, may lead to use of fewer tests and referrals, and thus longer intervals to diagnosis and potentially worse clinical outcomes in this patient group. Symptoms associated with cancer arising long before a diagnosis is subsequently made present the best opportunity to improve the diagnostic process, but it is not obvious how these could be detected in observational studies when both natural disease progression and investigative management vary widely between patients. Further research is required to understand the degree to which the time-dependent nature of risk reflects the natural history of the disease or historical clinical practice, and how such appreciations can inform improvement efforts.

As different cancer sites have heterogeneous associations with the same pre-diagnostic feature, both reviewed studies usefully report data on the cancer site case-mix of patients with either unintended loss of weight or positive inflammatory marker testing.[6,7] Nicholson et al.[6] also report that, among patients with unintended weight loss, the greatest cancer site-specific risks relate to pancreatic cancer and cancer of unknown primary, while the risks of either prostate or breast cancer are lower than that in patients without unintended weight loss. Such analyses, of the 'cancer signature' of different pre-diagnostic features, can help to devise cost-efficient investigation strategies, individualising the cascade of specialist diagnostic tests, depending on the presenting symptoms and initial test findings.[13,20] Echoing other recent evidence,[21] a substantial proportion of cancer patients who presented with unintended weight loss were diagnosed in stages other than stage IV.[6]

An issue beyond the optimisation of risk prediction algorithms for symptomatic-but-as-yet-undiagnosed cancer is their appropriate and dependable incorporation into everyday clinical practice, a complex problem in need of implementation research. Artificial intelligence is considered to offer promise in improving the diagnostic process, but while it could improve risk assessment in real time, it is also associated with implementation challenges and potential safety risks.[22] There is currently little conclusive evidence about the effectiveness of decision-support tools embedded in electronic health records, although trials are ongoing.[23,24]

A limitation of current evidence is that, with certain exceptions, studies estimate the risk of cancer without quantification of the risk of other consequential illness such as autoimmune disease or infection.[25,26] Among patients with non-localising symptoms (such as unintended loss of weight), the yield ratio for consequential non-neoplastic illness over cancer diagnosis may be as high as 2:1.[20] Therefore, promptly and efficiently diagnosing symptomatic-but-as-yet-undiagnosed cancer does not simply represent a problem solely for cancer research, rather one for medical research, overall. Broad coalitions of research funders, researchers, clinicians and policy-makers are needed to support research to improve the diagnosis of cancer and other significant diseases.

AUTHOR CONTRIBUTIONS

G.L. first drafted and G.A.A. contributed to subsequent versions of the editorial.

ADDITIONAL INFORMATION

Ethics approval and consent to participate Not applicable.

Georgios Lyratzopoulos[1] and Gary A. Abel[2]
[1]*Epidemiology of Cancer Healthcare and Outcomes (ECHO) Group, Department of Behavioural Science and Health, Institute of Epidemiology and Health Care, University College London, 1-19 Torrington Place, London WC1E 7HB, UK and [2]University of Exeter Medical School, Smeall Building, St. Luke's Campus, Exeter EX1 2 LU, UK*
Correspondence: Georgios Lyratzopoulos (y.lyratzopoulos@ucl.ac.uk)

REFERENCES

1. Elliss-Brookes, L., McPhail, S., Ives, A., Greenslade, M., Shelton, J., Hiom, S. et al. Routes to diagnosis for cancer—determining the patient journey using multiple routine data sets. *Br. J. Cancer* **107**, 1220–1226 (2012).
2. Tørring, M. L., Murchie, P., Hamilton, W., Vedsted, P., Esteva, M., Lautrup, M. et al. Evidence of advanced stage colorectal cancer with longer diagnostic intervals: a pooled analysis of seven primary care cohorts comprising 11 720 patients in five countries. *Br. J. Cancer* **117**, 888–897 (2017).
3. Mendonca, S. C., Abel, G. A., Saunders, C. L., Wardle, J. & Lyratzopoulos, G. Pre-referral general practitioner consultations and subsequent experience of cancer care: evidence from the English Cancer Patient Experience Survey. *Eur. J. Cancer Care (Engl.)* **25**, 478–490 (2016).
4. Lyratzopoulos, G., Neal, R. D., Barbiere, J. M., Rubin, G. P. & Abel, G. A. Variation in number of general practitioner consultations before hospital referral for cancer: findings from the 2010 National Cancer Patient Experience Survey in England. *Lancet Oncol.* **13**, 353–365 (2012).
5. Lawrenson, R., Logie, J. & Marks, C. Risk of colorectal cancer in general practice patients presenting with rectal bleeding, change in bowel habit or anaemia. *Eur. J. Cancer Care (Engl.)* **15**, 267–271 (2006).
6. Nicholson, B. D., Hamilton, W., Koshiaris, C., Oke, J. L., Hobbs, F. D. R. & Aveyard, P. The association between unexpected weight loss and cancer diagnosis in primary care: a matched cohort analysis of 65,000 presentations. *Br. J. Cancer* https://doi.org/10.1038/s41416-020-0829-3 (2020).
7. Watson, J., Salisbury, C., Banks, J., Whiting, P. & Hamilton, W. Predictive value of inflammatory markers for cancer diagnosis in primary care: a prospective cohort study using electronic health records. *Br. J. Cancer* **120**, 1045–1051 (2019).
8. Hamilton, W., Hajioff, S., Graham, J. & Schmidt-Hansen, M. Suspected cancer (part 2—adults): reference tables from updated NICE guidance. *BMJ* **350**, h3044 (2015).
9. Holtedahl, K. A. Inter-observer variation on registration of signals of cancer. *Scand. J. Prim. Health Care* **5**, 133–139 (1987).
10. Delaney, B. C. & Kostopoulou, O. Decision support for diagnosis should become routine in 21st century primary care. *Br. J. Gen. Pract.* **67**, 494–495 (2017).
11. Price, S. J., Stapley, S. A., Shephard, E., Barraclough, K. & Hamilton, W. T. Is omission of free text records a possible source of data loss and bias in Clinical Practice Research Datalink studies? A case–control study. *BMJ Open* **6**, e011664 (2016).
12. Nicholson, B. D., Aveyard, P., Hamilton, W., Bankhead, C. R., Koshiaris, C., Stevens, S. et al. The internal validation of weight and weight change coding using weight measurement data within the UK primary care Electronic Health Record. *Clin. Epidemiol.* **11**, 145–155 (2019).
13. Bailey, S. E., Ukoumunne, O. C., Shephard, E. A. & Hamilton, W. Clinical relevance of thrombocytosis in primary care: a prospective cohort study of cancer incidence

using English electronic medical records and cancer registry data. *Br. J. Gen. Pract.* **67**, e405–e413 (2017).

14. Nicholson, B. D., Aveyard, P., Bankhead, C. R., Hamilton, W., Hobbs, F. D. R. & Lay-Flurrie, S. Determinants and extent of weight recording in UK primary care: an analysis of 5 million adults' electronic health records from 2000 to 2017. *BMC Med.* **17**, 222 (2019).

15. Koo, M. M., Hamilton, W., Walter, F. M., Rubin, G. P. & Lyratzopoulos, G. Symptom signatures and diagnostic timeliness in cancer patients: a review of current evidence. *Neoplasia* **20**, 165–174 (2018).

16. Mills, K., Birt, L., Emery, J. D., Hall, N., Banks, J., Johnson, M. et al. Understanding symptom appraisal and help-seeking in people with symptoms suggestive of pancreatic cancer: a qualitative study. *BMJ Open* **7**, e015682 (2017).

17. Jones, R., Latinovic, R., Charlton, J. & Gulliford, M. C. Alarm symptoms in early diagnosis of cancer in primary care: cohort study using General Practice Research Database. *BMJ* **334**, 1040 (2007).

18. Hippisley-Cox, J. & Coupland, C. Symptoms and risk factors to identify men with suspected cancer in primary care: derivation and validation of an algorithm. *Br. J. Gen. Pract.* **63**, e1–e10 (2013).

19. Lyratzopoulos, G. Electronic patient records research to aid diagnostic reasoning for possible cancer in primary care. *Br. J. Gen. Pract.* **68**, 408–409 (2018).

20. Forster, A. S., Renzi, C. & Lyratzopoulos, G. Diagnosing cancer in patients with 'non-alarm' symptoms: Learning from diagnostic care innovations in Denmark. *Cancer Epidemiol.* **54**, 101–103 (2018).

21. Koo, M. M., Swann, R., McPhail, S., Abel, G. A., Elliss-Brookes, L., Rubin, G. P. et al. Presenting symptoms of cancer and stage at diagnosis: evidence from a cross-sectional, population-based study. *Lancet Oncol.* **21**, 73–79 (2020).

22. Challen, R., Denny, J., Pitt, M., Gompels, L., Edwards, T. & Tsaneva-Atanasova, K. Artificial intelligence, bias and clinical safety. *BMJ Qual. Saf.* **28**, 231–237 (2019).

23. Moore, H. J., Nixon, C., Tariq, A., Emery, J., Hamilton, W., Hoare, Z. et al. Evaluating a computer aid for assessing stomach symptoms (ECASS): study protocol for a randomised controlled trial. *Trials* **17**, 184 (2016).

24. Electronic Risk Assessment for Cancer for Patients in General Practice. ISRCTN22560297. https://doi.org/10.1186/ISRCTN22560297 (2020).

25. Stapley, S. A., Rubin, G. P., Alsina, D., Shephard, E. A., Rutter, M. D. & Hamilton, W. T. Clinical features of bowel disease in patients aged <50 years in primary care: a large case-control study. *Br. J. Gen. Pract.* **67**, e336–e344 (2017).

26. Watson, J., Jones, H. E., Banks, J., Whiting, P., Salisbury, C. & Hamilton, W. Use of multiple inflammatory marker tests in primary care: using Clinical Practice Research Datalink to evaluate accuracy. *Br. J. Gen. Pract.* **69**, e462–e469 (2019).