Gut feeling for the diagnosis of cancer in general practice: a diagnostic accuracy review

Masahiro Yao,1,2 Makoto Kaneko,1,2 Jessica Watson,1,3 Greg Irving1,4

ABSTRACT

Objectives  Diagnostic delay in cancer is a challenge in primary care. Although screening tests are effective in diagnosing some cancers such as breast, colorectal and cervical cancers, symptom-based cancer diagnosis is often difficult due to its low incidence in primary care and the influence of patient anxiety, doctor–patient relationship and psychosocial context. A general practitioner’s gut feeling for cancer may play a role in the early diagnosis of cancer in primary care where diagnostic resources are limited. The aim of this study is to summarise existing evidence about the test accuracy of gut feeling (index test) in symptomatic adult patients presenting to general practice, compared with multidisciplinary team-confirmed diagnosis of cancer (reference standard).

Design  Diagnostic accuracy review following Cochrane methods was performed.

Data sources  MEDLINE, EMBASE, Cochrane Library, the Database of Abstracts of Reviews of Effects and Medion databases.

Eligibility criteria  Cross-sectional, cohort and randomised studies of test accuracy that compared gut feeling (index test) with an appropriate cancer diagnosis (reference standard). No language or publication status restrictions were applied. We included all studies published before 25 March 2022.

Data extraction and synthesis  Methodological quality was appraised, using Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) criteria. Meta-analysis with hierarchical summary receiver operating characteristic (HSROC) models was used.

Results  Of 1286 potentially relevant studies identified, six met the inclusion criteria. For two of the six studies, data could not be extracted despite contacting authors. No studies satisfied all QUADAS-2 criteria. After meta-analysis of data from the remaining studies, the summary point of HSROC had a sensitivity of 0.40 (95% CI: 0.28, 0.53) and a specificity of 0.85 (95% CI: 0.75, 0.92).

Conclusions  Gut feeling for cancer when used in symptomatic adult patients in general practice has a relatively low sensitivity and high specificity. When the prevalence of cancer in the symptomatic population presenting in general practice exceeds 1.15%, the performance of gut feeling reaches the National Institute for Health and Care Excellence 3% positive predictive value threshold for action, which recommends urgent access to specialist care and further investigations. The findings support the continued and expanded use of gut feeling items in referral pathways.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ We summarised existing evidence about general practitioners’ gut feeling for the diagnosis of cancer by conducting a meta-analysis.
⇒ Meta-analysis with hierarchical summary receiver operating characteristic models was used, which calculates summary points of sensitivity and specificity of included studies.
⇒ The number of included studies is small and they have a risk of bias.

INTRODUCTION

As of 2015, the lifetime risk of developing cancer in the UK is estimated to be 50%.1,2 General practitioners (GPs) in the UK newly diagnose 2000 patients with cancer every year. The number is similar to those of diabetes, and twice as frequent as new cases of stroke.1 At present, a GP with 2000 patients has around 70 patients with or surviving cancer, and this number is predicted to double by 2040 in the UK.1,2 A small number of cancers are detected through screening programmes for colorectal, breast and cervical cancers. Around 85% of cancers are diagnosed after symptomatic presentation to a GP.1,3 The clinical symptoms of cancer presentations are diverse. At the same time, a GP must make an accurate and timely assessment of symptoms and signs in a setting where the prevalence of cancer is relatively low compared with other conditions and decide whether to make an urgent referral or arrange further investigation for a suspected cancer diagnosis.

To improve timely access to further investigations and specialists, the National Institute for Health and Care Excellence (NICE) in the UK developed the recommendation of positive predictive value (PPV) 3% as a risk threshold for urgent investigation or referral when symptoms are suspected to be caused by cancers.4 But this risk threshold was not always applied to recommendations relating to tests that are routinely available in primary care settings (including blood tests and...
imaging tests such as chest X-rays). Therefore, GPs are faced with a challenge in making a timely investigation or referral to specialists and an alternative diagnostic method is needed.

It is well recognised that gut feeling can play an important role for GPs when faced with uncertainty including whether a patient has cancer or not. The consensus statement of gut feeling by Stolper et al defined it as ‘a physician’s intuitive feeling that something is wrong with the patient, although there are no apparent clinical indications for this, or a physician’s intuitive feeling that the strategy used in relation to the patient is correct, although there is uncertainty about the diagnosis’. Smith et al also conceptualised gut feeling as a rapid summing up of multiple verbal and non-verbal patient cues in the context of the GP’s clinical knowledge and experience.

In terms of the well-known dual process cognitive model, or analytical thinking (slow thinking: system2) and non-analytical thinking (fast thinking: system1), gut feelings emerge as a consequence of non-analytical processing of the available information and knowledge, either reassuring GPs or alarming them. To investigate diagnostic accuracy of gut feeling for the diagnosis of cancer, several studies have been conducted in primary care settings and they have demonstrated a range of results. For example, the PPV for the gut feeling to diagnose cancer has varied widely from 3% to 35% in primary care setting. This review with meta-analysis aims to summarise the evidence about the diagnostic accuracy of gut feeling (index test) in symptomatic adult patients presenting to GPs, compared with confirmed diagnosis of cancer (reference standard), and also assesses whether the PPV meets the NICE threshold of 3%.

**METHODS**

**Design**

We performed a diagnostic accuracy review following the Cochrane method.

**Search strategy**

To identify relevant studies, we searched using the terms “gut feeling”, “gut instinct”, “intuition”, “sense of alarm”, “sense of reassurance”, “primary care” “general practice” and “family medicine” from medical literature databases (MEDLINE, EMBASE, Cochrane Library, the Database of Abstracts of Reviews of Effects and Medion databases). The search details have been previously reported on PROSPERO (CRD42021216211) and the full search strategies are included in the online supplemental material. No language or publication status restrictions were applied. Reference lists from all identified publications were hand searched. We included all studies published before 25 March 2022.

**Inclusion and exclusion criteria**

Studies comparing gut feeling (index test) with the reference standard of confirmed diagnosis of cancer in symptomatic adults presenting to GPs were included. Included study designs were diagnostic cross-sectional, cohort and randomised studies. As with recent diagnostic studies, we considered any cancer rather than focused on specific cancer, because of the very low prevalence of individual cancers in primary care. Because the definition of gut feeling was not consistent and then created heterogeneity in the threshold of the index test, we included studies with a variety of definitions of gut feeling for cancers, for example, ‘possible cancer’ or ‘the slightest suspicion of cancer or another serious disease’. Studies were excluded if participants had an existing diagnosis of a currently active cancer.

**Selection and data extraction**

To determine the eligibility of each potential study, MY and MK independently reviewed the identified studies by title/abstract and by the full text when required. The list of selected studies was then confirmed by a third author (GI) for the final selection decisions. Study quality was assessed using Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2). Any discrepancies were resolved by discussion.

A data extraction form was shared with the researchers to identify study characteristics and results for each included study which were independently extracted by MY and GI. Where possible, data were extracted into 2×2 contingency tables. All authors were contacted if sufficient data were not available in the published report to request further data.

**Statistical analysis and data synthesis**

The hierarchical summary receiver operating characteristic (ROC) model was employed to properly weigh the heterogeneity caused by the threshold of gut feeling (index test) in each primary literature. To analyse the summary points of sensitivities and specificities by the 2×2 contingency tables from each study, we used the command ‘metandi’ (meta-analysis of diagnostic accuracy using hierarchical logistic regression). We used STATA V.16 for statistical analysis. We also assessed heterogeneity visually by forest plot created by RevMan and calculated the I² statistic.

**Patient and public involvement**

This review addresses the James Lind Alliance top priority for detecting cancer early: ‘What simple, non-invasive, painless, cost-effective and convenient tests can be used to detect cancer early?’ No specific patient and public involvement was conducted for this study.

**RESULTS**

As shown in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart (figure 1), we identified 1286 relevant studies after removal of duplication. Two studies were secondary sources from original publications. After title and abstract screening, six studies met...
eligibility criteria. Finally, four studies met the objectives of the review (table 1).

The reason for the exclusion of the two studies was that the data extraction was not possible despite contacting authors. In QUADAS-2 criteria, a high risk of bias was seen in flow and timing at Scheel et al paper.9 All other studies had at least one unclear risk. The I² statistic for heterogeneity for the included studies was 81.7% (high).

Figure 2 shows the sensitivities and specificities in each study. After integration and meta-analysis, the sensitivity of gut feeling for the final diagnosis of cancer was 0.40 (SE 0.06, 95% CI 0.28 to 0.53) and the specificity was 0.85 (SE 0.04, 95% CI 0.75 to 0.92) (figure 3 and table 2).

Assuming 1–4% prevalence of cancer in adults with any symptoms in general practice, the PPV was 2.6–5.16% (online supplemental figure 1). The results of these studies indicate that in theory, if the gut feeling were to be used in primary care setting in a group of 1000 people where x (p%) have cancer, an estimated (true positive (TP)+false positive (FP)) would have (gut feeling) result indicating (cancer) is present and of these (FP) (<100×FP/TP+FP>) would be incorrectly classified as having (cancer); of the (true negative (TN)+false negative (FN)) people with a result indicating that (cancer) is not present, (FN) (<100×FN/TN+FN>) would be incorrectly classified as not having (cancer).

**DISCUSSION**

This study found that gut feeling for cancer when used in symptomatic adults in general practice has a relatively low sensitivity and high specificity in diagnosing cancer. When the prevalence of cancer in the symptomatic population presenting in general practice exceeds 1.15%, the performance of gut feeling reaches the NICE 3% PPV threshold for action, which recommends urgent access to specialist care and further investigations.4 15 In the UK, the prevalence of symptomatic and surviving cancer is ~3.5% with the prevalence appearing to be increasing over time.1

Smith et al reported a systematic review which mainly assessed how useful gut feeling is and how doctors used it in general practice both by qualitative and quantitative syntheses.7 Although their included primary studies on quantity were the same with ours and they also performed a meta-analysis, they reported a pooled OR, rather than summary of meta-analysed ROC points and calculated PPVs that could be with the NICE risk threshold for urgent action. The Scheel et al study had a large amount of missing data (952 cases: 15.1%).9 Smith et al appear to have used the confirmed numbers of cancer diagnosis and the total number only rather than explore and make use of the missing data.7 After contacting the authors and

<table>
<thead>
<tr>
<th>Author and publication year</th>
<th>Title</th>
<th>Country</th>
<th>Study design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holtedahl et al 201710</td>
<td>Abdominal symptoms in general practice: frequency, cancer suspicions raised, and action taken by GPs in six European countries. Cohort study with prospective registration of cancer</td>
<td>The Netherlands, Denmark, Norway, Sweden, Belgium, Scotland</td>
<td>Prospective cohort</td>
</tr>
<tr>
<td>Ingeman et al 201518</td>
<td>The Danish cancer pathway for patients with serious non-specific symptoms and signs of cancer—a cross-sectional study of patient characteristics and cancer probability</td>
<td>Denmark</td>
<td>Cross-sectional</td>
</tr>
<tr>
<td>Hjertholm et al 20144</td>
<td>Predictive values of GPs’ suspicion of serious disease: a population-based follow-up study</td>
<td>Denmark</td>
<td>Prospective cohort</td>
</tr>
<tr>
<td>Scheel et al 20139</td>
<td>Cancer suspicion in general practice: the role of symptoms and patient characteristics, and their association with subsequent cancer</td>
<td>Norway</td>
<td>Prospective cohort</td>
</tr>
</tbody>
</table>

GPs, general practitioners.
discussing, we treated these missing data as gut feeling (sense of alarm (SA)) was not recorded. Our justification for this links to the work of Oliva-Fanlo et al and Barais et al, where they found GPs’ gut feeling exists in almost all encounters and the prevalence of sense of reassurance (SR) is very high rather than SA (SR: 75.3%, SA: 21.3%). It follows that if SA was observed, the GP would register and record this.16 17

In this study, we included studies with a variety of definitions of gut feeling since no study used a standardised scale of it. Two of the included studies set the range of gut feeling into categories such as slight, medium and strong.10 18 Consequently, the stronger gut feeling for cancer was, the higher the specificities were (specificity of 0.97 and sensitivity of 0.19, calculated in the literature, Holtedahl et al10). That indicates if the range of gut feeling could be quantified or categorised in detail, future studies would qualify its specificity more. In 2017, for the growing need for a clearer definition of gut feeling, Stolper et al developed the Gut Feeling Questionnaire (GFQ) which determines the presence of gut feeling in the context of GPs’ diagnostic reasoning.19 GFQ assesses both SA and SR. In the future diagnostic accuracy research, both senses will be assessed with measurable scale, which may improve quality of each specificity. And also, if studies on SR might have high specificity, that would be related to reducing unnecessary investigation or referrals.

The primary literature in this study was from European countries with healthcare systems similar to that of the UK, so the results could be applicable to a UK setting, as long as it deals with symptom-based judgement on further investigations or specialist referrals in general practice. Furthermore, it might be possible to compare outcomes under other healthcare situations or systems in specific circumstances, such as home care and nursing facility care settings, where diagnostic resources are limited.

Implications of the study
To our best knowledge, this is the first diagnostic accuracy review with meta-analysed ROC points about GPs’ gut feeling. This study suggests that gut feeling for cancers should be used where appropriate to strengthen guidelines and cancer care pathways in the UK and internationally. Gut feeling is a non-invasive, inexpensive tool that can help to reduce the primary care interval if used correctly, getting the patient to the appropriate specialist in the shortest possible time without relying on unnecessary investigation. In practical terms, when a GP has a ‘gut feeling’ that a patient may have cancer, this means that it is highly likely that the patient does have cancer. There are fewer FPs—instances where the GP suspects cancer but the patient does not actually have it. However, given the diagnostic performance of gut feeling in this setting, it also means that the GP’s gut feeling does not accurately identify all cases of cancer. There may be

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Summary points for diagnostic accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary point</td>
<td>Coefficient</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.4</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.85</td>
</tr>
</tbody>
</table>
instances where the GP does not suspect cancer, but the patient actually has it. Therefore, relying solely on the GP’s gut feeling may result in missed or delayed diagnoses. More research is needed to develop standardised definitions of gut feeling and also explore which cancers may benefit most from this approach.

LIMITATIONS
First, the number of integrated studies is small (four), and there is a possibility of publication bias. All studies had at least one unclear risk and the study by Scheel et al. was judged as high risk in the flow and timing items in QUADAS-2 criteria. There is also a high level of heterogeneity. The reasons for this are likely multifactorial but are likely related to the range of definitions used for determining gut feeling and the range of population studies. Also, because the range of 95% CIs and prediction regions are wide, the results need to be carefully interpreted. In the future research, patient characteristics (age, mental disorder, smoking status, symptoms, etc) as well as GPs’ level of experience, sex, doctor–patient relationships (well-known or not) and clinical setting (first contact or continuity) should be assessed using a clearer definition of gut feeling such as GFOQ. Second, although this study adopted any cancer as a reference standard, cancer types could have a high heterogeneity, both within individual studies and between studies. Some cancers, especially myeloma and stomach cancer, are difficult to diagnose earlier, compared with breast cancer, testicular cancer and melanoma. To improve diagnostic delay in each cancer, future research also needs to assess cancer types.

CONCLUSIONS
Gut feeling is a useful cognitive tool that can potentially help to reduce the primary care interval for cancer diagnosis in a non-invasive and inexpensive way if used correctly. In the field of primary care, where medical resources are limited, GPs’ gut feeling for cancer has a relatively low sensitivity and high specificity. Assuming a prevalence of symptomatic cancer in the general practice population of greater than 1.15%, the performance of gut feeling reaches the NICE 3% threshold for action, which recommends urgent access to specialist care and further investigations. The findings support the continued and expanded use of gut feeling items in UK cancer referral pathways.

Twitter Jessica Watson @drjessicawatson and Greg Irving @gregjirving

Contributors MY and MK designed the study and served as guarantors. MY and MK analysed the data. MY drafted the manuscript. MK, JW and GI contributed to the design of the study and critically reviewed the manuscript. All authors had full access to the data and took responsibility for the integrity of the data and the accuracy of the analysis.

Funding No funding was received for this study.

References


