BMJ Open

The role of Primary Care Physician factors on diagnostic testing and referral decisions for symptoms of possible cancer: a systematic review

SCHOLARONE™
Manuscripts

BMJ

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

The role of Primary Care Physician factors on diagnostic testing and referral decisions for symptoms of possible cancer: a systematic review

Author and corresponding author

Victoria Hardy, Department of Public Health & Primary Care, University of Cambridge, Cambridge, UK veh29@medschl.cam.ac.uk

Co-authors

Adelaide Yue, Department of Public Health & Primary Care, University of Cambridge, Cambridge, UK

Stephanie Archer, Department of Public Health & Primary Care, University of Cambridge, Cambridge, UK

Samuel W.D. Merriel, College of Medicine and Health, University of Exeter, Exeter, UK

Matthew J. Thompson, Department of Family Medicine, University of Washington, Seattle, Washington, USA

Jon Emery, Centre for Cancer Research and Department of General Practice, University of Melbourne, Victorian Comprehensive Cancer Centre, Melbourne, Victoria, Australia

Juliet Usher-Smith, Department of Public Health & Primary Care, University of Cambridge, Cambridge, UK

Fiona M. Walter, Wolfson Institute of Population Health, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, UK

Key words: Primary care; early cancer diagnosis; decision making; diagnostic error

Word count: 3,996

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT

Background: Missed opportunities for diagnosing cancer cause patients harm and have been attributed to suboptimal use of tests and referral pathways in primary care. Primary care physician (PCP) factors may influence their decisions to investigate cancer, but this is poorly understood.

Objective: To synthesize evidence evaluating the influence of PCP factors on decisions to investigate symptoms of possible cancer.

Methods: We searched MEDLINE, Embase, Scopus, CINAHL and PsycINFO between January 1990 and March 2021 for relevant citations. Studies examining the effect or perceptions and experiences of PCP factors on use of tests and referral decisions for symptomatic patients with any cancer were included. PCP factors comprised personal characteristics and attributes of physicians in clinical practice.

Data extraction and synthesis: Critical appraisal and data extraction were undertaken independently by two authors. Due to study heterogeneity, data could not be statistically pooled. We therefore performed a narrative synthesis.

Results: 29 studies were included. Most studies were conducted in European countries. A total of 11 PCP factors were identified comprising modifiable and non-modifiable factors. Clinical interpretation of symptoms as suspicious or alarm symptoms prompted more investigations than non-alarm symptoms in relation to any cancer. 'Gut feeling' predicted a subsequent cancer diagnosis, facilitated decisions to investigate non-specific symptoms, and was linked to years of experience. Female PCPs investigated cancer more than male PCPs.

The effect of PCP age and years of experience on testing and referral decisions was inconclusive.

Conclusions: Clinical judgement helped PCPs recognize higher-risk presentations and make effective testing and referral decisions leading to a cancer diagnosis. In the absence of alarm symptoms or 'gut feelings' PCPs may not suspect cancer and require strategies to identify patients with non-alarm and non-specific symptoms who need testing or referral.

PROSPERO protocol registration ID: [CRD420191560515]

ARTICLE SUMMARY

Strengths and limitations

- To our knowledge this is the first study that has examined a range of PCP factors affecting testing and referral decisions in a cancer context.
- We identified the role PCP factors play in managing 'alarm' symptoms and make recommendations for improving testing and referral decisions for patients with non-alarm and non-specific symptoms.
- We were unable to perform a meta-analysis due to heterogeneity of outcomes among studies.

INTRODUCTION

Diagnostic error, which encompasses missed, delayed, and incorrect diagnosis, is a major patient safety concern in primary care,(1) and cancer is among the most frequently missed diagnoses in this setting internationally.(2–4) Improving patient safety in primary care has been identified by the World Health Organization (WHO) as a priority.(5) Achieving a timely cancer diagnosis can be challenging due to low cancer incidence among populations presenting to primary care,(6) patient comorbidity,(7,8) and overlapping symptoms between cancers (e.g. abdominal symptoms can herald oesophageal, colorectal, or renal cancer),(9) as well as more commonly occurring benign conditions.(10)

Clinical guidelines, such as the National Institute for Health and Care Excellence (NICE) NG12 (2015) guidelines for suspected cancer which are based on patient risk factors and symptoms most predictive of cancer,(11) are used in several countries to promote appropriate and timely recognition and referral of symptoms warranting investigation.(12) Despite reductions in diagnostic delay across multiple cancer sites since implementation of these guidelines in the early 2000s,(13,14) around one quarter of cancer patients (including those with symptoms of relatively high predictive value; 'alarm' symptoms henceforth) in a recently published analysis of over 17,000 cases from the English National Cancer Diagnosis Audit (NCDA), were assessed by PCPs to have had avoidable diagnostic delay.(15) Such patients often have three or more consultations before referral to secondary care for further investigation,(16) in addition to poor prognosis and patient experiences of health care.(17)

Notwithstanding the inherent difficulty of diagnosing cancer, approximately half of cancer-related diagnostic errors identified from English National Cancer Registries and closed malpractice claims in the United States (US) have been attributed to the PCP, and primarily relate to delay or failure to recommend an appropriate test or referral.(3,7,15) Further, unexplained variation in use of dedicated urgent referral pathways after adjusting for case-mix,(18) patient and practice factors,(19) has suggested the potential involvement of factors related to the PCP.(20) However, the PCP factors contributing to decisions to investigate cancer, and the potential benefits and disadvantages such factors may confer to such decisions are poorly understood. Therefore, in this systematic review, we aimed to identify and determine the influence of PCP factors on testing and referral decisions for symptoms suggestive of cancer.

METHODS

This review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines.(21) A protocol was registered with PROSPERO before conducting the review (CRD42019160515).

Search strategy

We searched Medline, Embase, CINAHL, PsycINFO and Scopus for relevant articles from 1st January 1990 to 31st March 2021. We restricted the search from 1990 to coincide with increased public investment for primary care cancer research.(22) Search queries were developed for Medline (**Supplementary material 1**) and adjusted according to the conventions of each database. Search terms comprised free text words and Medical Subject Headings (MeSH) informed by key words in titles and abstracts of relevant literature known to the authors,(23–25) with input from an information specialist.

Eligibility criteria

Studies were included if they were conducted in a developed country and investigated or described the influence of PCP factors on outcomes of testing and referral decisions at any stage during the diagnostic workup of adult patients (aged ≥18 years) with symptoms indicative of any cancer. We defined PCP factors as attributes that have been the object of interest regarding physician performance and competence, specifically focusing on the characteristics and attributes of physicians in clinical practice settings or the community.(26) For this review, PCPs encompassed General Practitioners and Family Physicians who have first patient contact and assume overall responsibility for coordinating care.(27) We did not limit studies by symptom type or thresholds of cancer risk based on the predictive value of clinical profiles, because we wanted to understand how PCP factors contribute to the management of undifferentiated symptoms characteristic of presentations in primary care. Inferring the effect of an exposure on individual behavior from associations derived in aggregate can be problematic, therefore studies contributing quantitative data were only included if data were reported at the level of the PCP. Systematic reviews were excluded, but where relevant to our review question, were used as a source of additional primary articles. We excluded commentaries, letters and editorials. There were no language restrictions.

Study selection

Citations retrieved by the search were imported into Rayyan QCRI.(28) Following removal of duplicates, titles and abstracts were screened against eligibility criteria by VH. SM independently screened an initial 10% of titles and abstracts to calculate the inter-rater reliability between reviewers using Cohen's Kappa coefficient. The resulting coefficient value of 0.66 reflected substantial agreement and exceeded our minimum *a priori* IRR cutoff of ≥0.61,(29) therefore, the remaining titles/abstracts were screened by VH alone. Full texts of potentially relevant studies were independently assessed by VH and AY to determine final inclusion. Reference lists of included studies and systematic reviews were hand-searched for additional articles not retrieved by the database searches; discrepancies between reviewers were resolved through discussion, and adjudicated by SA/FW when agreement could not be reached.

Data extraction

Separate data extraction spreadsheets for quantitative and qualitative data were developed and piloted in Excel. Three authors (VH and AY/SA) independently extracted data relating to study characteristics, PCP factors, test and referral type, study results/findings, and cancer, using the relevant data extraction spreadsheet. For quantitative studies, VH and AY extracted findings regarding the frequency of PCP use of diagnostic tests and referrals, and subsequent cancer diagnoses. For qualitative studies, VH and SA extracted primary data (i.e., participant quotations and authors' verbatim summaries and interpretation), along with accompanying themes and subthemes to preserve context for data synthesis. Inconsistencies in data extraction between VH and AY/SA were discussed and resolved through discussion.

Critical appraisal

Quality of included studies was assessed using the Joanna Briggs Institute (JBI) Critical Appraisal Tools as they incorporate appraisal checklists for different study designs.(30) VH and AY/SA independently rated each criterion of the relevant checklist for each included study; discrepancies were resolved through discussion. As JBI checklists have a different number of criteria depending on study type, we visually summarize study quality using an approach in a previous systematic review, whereby the overall score for each study was converted to a percentage and classified as 'low' (0-45%), 'fair' (≥46-69%) or 'high' quality (≥70%);(31) studies were not excluded on the basis of low quality.

Data analysis and synthesis

Due to heterogeneity of study outcomes, we were unable to statistically pool data. Therefore, we performed a narrative synthesis. We used a convergent segregated approach whereby quantitative and qualitative data were analysed separately before being thematically synthesized.(32) Quantitative data was initially organized according to the PCP factor label in included studies. Qualitative data was manually coded for PCPs' perceptions of the influence of PCP factors on decisions to investigate cancer that could potentially be grouped into descriptive categories. Definitions of PCP factors from quantitative and qualitative studies were compared to determine the similarity of identified constructs. Data for each construct was iteratively grouped and checked by FW/SA until categories of PCP factors became clear. Final labels for each PCP factor were refined following team consensus meetings **(Table 1)**. Quantitative and qualitative findings for each PCP factor were then separately examined. The direction and magnitude of effect (for quantitative data) and PCPs' perceptions of the influence of PCP factors (for qualitative data) were summarised in textual format for every PCP factor. Textual summaries for PCP factors were juxtaposed and subsequently combined into a new descriptive narrative that encapsulated the findings from each study.(33) Finally, PCP factors were organised into over-arching themes according to the extent to which they were deemed to be modifiable or non-modifiable. Modifiable factors were identified as factors that were susceptible to individual control or being changed with intervention. Non-modifiable factors included factors that were viewed to be outside the purview of individual control and less susceptible to adjustment. Findings for each PCP factor were interpreted in the context of the methodological limitations of each study from the critical appraisal. Only PCP factors for which there were a minimum of two studies were synthesized.

Patient and public involvement

This systematic review was undertaken as part of the CanTest Collaborative research programme funded by Cancer Research UK, which involves close collaboration with a panel of PPI representatives whose views informed the design of this study, data analysis, and data interpretation.

RESULTS

Search results

The search yielded a total of 7,938 studies. After de-duplication, titles and abstracts of 4,135 studies were screened; 3,721 did not meet inclusion criteria and were excluded. We reviewed the full text of 82 studies retrieved from the search plus a further nine identified through snowballing approaches. 29 studies (19 quantitative and 10 qualitative) met the inclusion criteria and were included in the final synthesis (see **Figure 1**).

Study characteristics

The included quantitative studies consisted of: 13 cross-sectional studies(25,34–45) and two experimental studies,(23,46) of which eleven used vignette-based methods;(23,25,35–39,42,44–46) three prospective cohort studies(47–49) and one retrospective cohort study **(Supplementary material 2)**.(49) Qualitative data came from five interview studies,(50–54) two focus group studies,(55,56) two cross-sectional survey studies(57,58) and one retrospective cohort of free-text primary care consultation data.(59) Of these 29 included studies, eleven were conducted in the UK,(23,25,34,37,43,44,50,51,54,55,57) six in Denmark,(40,41,45,47,60,61) two across 20 EU countries,(39,58) two each in Australia,(35,46) USA,(36,38) and Norway,(49,53) and one each in the Netherlands,(59) Sweden,(52) Spain,(56) Australia and the UK (in one study).(42) Together included studies reflected a total of 10,300 PCPs (8,000 from quantitative and 2,300 from qualitative studies) and testing and referral decisions for 15,100 patients.

The most common types of investigations evaluated were urgent referrals (37,41–43,51,54,55,57,60) or direct referrals to diagnostic testing centers,(34,40,45,50) imaging,(48) or other specialty services.(35,46) Additional investigations included use of appropriate investigations for the cancer,(23,25,36,38,39) any diagnostic action,(47,49,50,52,53,58) or referrals in general.(56,59) Testing and referral decisions were mainly made in the context of symptoms generally suggestive of the target cancer(s).(35,36,38,41,47,48,55,58–60) Other studies examined decision-making for symptoms at high-,(46,49) low-,(25,34,39,40,50) and mixed-risk of cancer (based on the positive predictive value (PPV) of symptoms in national guidelines),(23,37,42,44) or did not provide details regarding symptomatic context.(43,45,51–54,56,57) Cancers studied were colorectal,(36,37,42,43,52,54,55,59) lung,(23,44,46,55,61) ovarian,(38,60) a mixture of cancers,(25,35,39–41,58) or any cancer.(34,45,47,50,51,53,56,57)

Quality of included studies

Twelve studies were assessed as high quality,(23,25,34,36,37,39,45–47,49,50,53) 11 were fair quality,(35,38,40,41,44,48,51,52,55,59,60) and six were low quality **(Supplementary material 2)**.(42,43,54,56–58) The main quality issues of the cross-sectional and cohort studies related to poor reporting of sample characteristics,(40,41,43,47–49,60) failure to identify or adjust for confounding,(38,42,43,49) and hindsight bias due to knowledge of the diagnosis at the point of assessment of PCP factors.(40,41,49) Qualitative studies were limited by insufficient evidence of researcher reflexivity and positionality,(50–55,57,59) suboptimal presentation of participants' voices,(55,56,59) and failure to indicate whether ethical permissions were obtained or waived.(56) Recall and social desirability bias was also a concern across study types where assessment of the exposure by PCPs was examined retrospectively,(40,41,49) or based on PCPs reflections of their clinical practice.(50–54) Finally, there were concerns about several studies using vignettes which were not adequately realistic of patient cases seen in practice and lacked evidence of validation.(25,35,37,46,57)

Table 1: Description of PCP factors identified from included studies

PCP factor	Description of PCP factors in this review	PCP factor terms in original studies	Definitions of PCP factors in original studies	Assessment methods in original studies
NON-MODIFIABLE FACTORS				
Demographic characteristics				
Sex	Male or female	Sex;(38,45) gender (23,37,39,44,46)	Not specified (23,37,38,44,45,46)	As recorded in the AMA Physician Masterfile;(38) questionnaire;(23,39) not specified(37,44,45,46)
Age	Chronological age	Age (23,38,42,45,48)	Survey year [2008]-year of birth;(38) in years (23,42,45,48)	Questionnaire;(23,35,42) not specified(38,48)
Years of experience	A composite of the number of years since graduation from medical school and completion of PCP training or residency	Years of experience; (35,37,50,52) years in practice or family medicine;(25,38,42, 44) years since qualifying or graduating(23,39)	Survey year [2008]-year graduated from medical school;(38) number of years in clinical practice;(42) not specified(23,25,35,37,39,44,50,52)	PCP selection of: under 10 years/10-19 years/20-29 years/30-39 years/40-49 years or over/or 'I prefer not to say' to the question 'how many years is it since you graduated as a doctor?' in questionnaire;(39) questionnaire;(23,35,42) interviews;(50) not specified(25,37,38,44,52)
MODIFIABLE FACTORS				
Clinical reasoning processes				
Clinical judgement	Clinical judgement regarding the possibility of cancer or other serious illness prompted by PCPs' interpretation of patient's symptoms or other clinical aspects of the presentation	Suspicion of cancer (or other serious disease);(47,49,51, 57,60) symptom interpretation;(41) clinical concern(43)	Interpretation of symptoms as alarm (suggestive of cancer), serious (suggestive of any serious disease) or vague (not directly suggestive of cancer or other serious disease);(41) suspicion of cancer based on patients first	Prospective(49) and retrospective PCP questionnaire completed on basis of medical records;(41,49) surveys,(43,57) or a registration form(47) including questions: 'are you left with the slightest suspicion of cancer or another serious disease (new)?' completed at time of patient presentation,(47) 'how did you interpret the symptoms?'(41) or requesting indication of

			presented symptoms;(60) not specified(43,47,49,51,57)	whether 'cancer possible, follow-up needed' or 'cancer not likely'(49) and presence or absence of clinical concern;(43) interviews(51)
'Gut feeling'	A sense of alarm reflecting an uneasy feeling indicating concern about a possible adverse outcome, even though specific indications are lacking, or a 'sense of reassurance' reflecting a feeling of security about the patient's management, even though the diagnosis may be uncertain;(62,63) typically considered irrational by PCPs(62)	Gut feeling;(40,50,52) hunch/gut instinct in response to symptoms;(55) intuition;(45) clinical suspicion/gut feeling(34,50)	Intuition,(45,53) or an 'intuitive feeling that something is wrong although there are no apparent clinical indications for this/a physician's intuitive feeling that the strategy used in relation to the patient is correct although there is uncertainty about the diagnosis';(40) clinical suspicion/gut feeling of cancer or serious disease;(34,50) features of clinical practice that would raise suspicion of lung cancer;(55) not specified(52)	Questionnaire completed by PCP retrospectively from medical records;(40) and survey including question: 'to what degree do you use intuition in your everyday work as a GP e.g. in relation to referral for investigation, sampling etc.?',(45) primary care referral forms and secondary care datasets;(34) interviews,(50,52,53) and focus groups(55)
First diagnostic impression s	First diagnostic impressions or hypotheses regarding the possible cause of a presentation, typically understood to be based on the most easily recalled information, at the start of a clinical encounter(64)	First impressions (25,59)	Verbalizations of cancer as a possible diagnosis immediately after reading initial description of hypothetical case scenario but before gathering any further information about the 'patients' complaint;(25) not specified(59)	Coding of initial verbalizations as 'cancer mentioned' or 'cancer not mentioned' using audio-recorded think-aloud protocols;(25) thematic analysis of free text in primary care records(59)

Assessment of cancer risk	Assessment of the likelihood or probability of cancer underpinned by the diagnostic value of patient risk factors, symptoms or test results to rule in or rule out cancer as a diagnosis (65)	Estimation of cancer risk at referral;(40) anticipated risk of cancer at referral(45)	Not specified(40,45)	Questionnaire completed retrospectively asking for PCPs estimation of cancer risk based on medical record entries for patients with any of 21 symptoms at time of referral;(40) and survey response to question: "in your judgment, what is the probability that a 50-year-old patient has cancer when you choose to refer the patient to fast track diagnostic services?"(45)
Dealing with uncertainty				
Attitude to uncertainty and risk	Positive or negative psychological responses to unexplained, incongruent, or imperfect diagnostic clues,(66) and/or an inclination to act when the implication of making the wrong decision could be significant	Level of risk;(38) physician risk attitude, tolerance for ambiguity;(45) uncertainty(51,53)	Not specified(38,45,51,53)	Surveys of self-assessed ratings on the 'Tolerance for Ambiguity Scale', and sub-scales of the 'Physician Response to Uncertainty' scale: anxiety due to uncertainty and concern about bad outcomes using six-point Likert scale from 1 (strongly disagree) to 6 (strongly agree);(45) interviews(51,53)
Fear of malpractice	Concern about potential accusations from colleagues or patients of professional negligence, and worry of formal or informal recourse as a consequence of those accusations	Fear of malpractice;(38,50) fear of litigation or complaint(43,56)	Not specified(38,43,50,56)	Surveys(38,43) requesting yes/no response to question: *'what influences your decision to use fast-track referral route?'* for 'fear of litigation';(43) interviews/focus groups(50,56)
Professional role and involvement in continuing medical education				
Attitude to role as gatekeeper	Attitudes towards PCPs perceived role in facilitating or controlling patient access to specialist health care services	Role as gatekeeper (45,50,51)	Not specified(45,50,51)	Surveys including binary (yes/no) responses to two individual statements: "the most important role as gatekeeper is to prevent overuse of secondary health services" and "ensure proper medical guidance and referral"(45) and interviews(50,51)
Participation in Continued	Participation in any professional activity to "maintain, develop, or increase knowledge, skills, professional	Involvement in clinical teaching;(38) value	Not specified(38,51,54,58)	Open-ended survey questions;(58) interviews;(51,54) not specified(38)

| Medical Education | performance and relationships to provide services for patients, the public, or the profession" (67) | of professional education(51,54,58) | BMJ Open | |

Abbreviations: AMA, American Medical Association; PCP, primary care physician

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

PCP factors

A total of 11 PCP factors were identified (**Table 1**). PCP factors included non-modifiable factors, which comprised demographic characteristics, and modifiable factors, including: clinical reasoning processes, management of uncertainty, and professional role and involvement in continuing medical education. The evidence for these 11 PCP factors is summarized below (quantitative and qualitative findings for each factor are presented separately in **Supplementary material 3 and 4**, respectively).

NON-MODIFIABLE FACTORS

Demographic characteristics

Sex (n=7)

Most studies suggested that female PCPs more often investigated possible cancer compared to male PCPs,(37–39,46,48) though findings were predominantly descriptive. Being female was associated with increased odds of urgent specialist review for patients with a lung nodule on CT (OR 1.87 CI 1.36-2.56, p≤0.001),(46) and referral for symptoms suspicious of colorectal cancer.(37) Compared to male PCPs, females more frequently recommended appropriate investigations for ovarian cancer(38) and symptoms of any possible cancer,(39) but these differences were not statistically significant. There was also no observable trend between PCP sex and use of investigations for lung cancer,(23,48) nor was there an association with estimation of cancer risk at urgent referral.(45)

Age (n=5)

The effect of age on decisions to investigate possible cancer was inconclusive. In two studies in Australia/UK and the US, PCPs 30 to 39 years were most likely to recommend specialist referral for colorectal cancer(42) and request appropriate investigations for ovarian cancer.(38) However, in another study, younger PCPs in the UK least frequently used relevant investigations for lung cancer, compared to older PCPs.(23) Odds of urgent referral increased with age, with PCPs 60 to 69 years most likely to recommend urgent referral (OR 15.4 CI 4.4 to 53.8, p≤0.001) compared to those 30 to 39 years (OR 2.8 CI 1.5-5.2, p≤0.001).(42) PCPs 60 years and above tended to estimate cancer risk at referral to be higher than PCPs younger than 45 years, though this was non-significant.(45) There was also no difference in mean age at referral between PCPs using and not using direct CT referral for lung cancer.(61)

Years of experience (n=10)

Evidence regarding the influence of years of experience was also mixed. Compared to those less experienced, PCPs with upwards of ten years of experience were more likely to make a specialist referral for ovarian(35) and colorectal cancer,(42) and initiate diagnostic action in general.(39) PCPs trusted their clinical judgement with increasing experience, and became more willing to investigate when they were suspicious of cancer, regardless of colleagues opinions of their decisions.(50,52) However, one study found PCPs with less than six years of experience were most able to differentiate presentations at risk thresholds of 3% or more requiring urgent referral compared to PCPs with 18 to 36 years of experience.(44) A further three studies found no observable influence of years of experience on decisions to test and refer for possible cancer.(23,25,38)

MODIFIABLE FACTORS

Clinical reasoning processes

Clinical judgement (n=7)

Clinical judgement of symptoms as suspicious or alarming consistently led to more investigations across multiple cancer sites.(41,43,47,49,51,57,60) When PCPs were suspicious of cancer in patients with warning symptoms, they were more likely to recommend imaging investigations (OR 3.95 CI 2.80-5.57) or make a referral (OR 2.56 CI 2.22-2.96), compared to no suspicion.(47) Conversely, symptoms that did not provoke suspicion(60) or were interpreted as non-alarm(41) were less likely to be referred urgently. PCPs were described as having difficulty discerning the appropriateness of using urgent referral pathways when symptoms did not clearly match referral criteria which was perceived not to accommodate individual clinical judgement.(57)

Gut feeling' (n=7)

The presence of 'gut feeling' did not lead to increased referral rates, but in two studies conducted in the UK and Denmark the odds of a subsequent cancer diagnosis was up to two-fold higher for patients referred along pathways for non-specific and warning symptoms of cancer.(34,40) 'Gut feeling' was perceived to facilitate management of non-specific symptoms, and become more accurate with experience.(55) Justifying referral of non-specific symptoms on the basis of 'gut feeling' was challenging as it was perceived to not be a sufficient indication for referral among hospital specialists.(50,53,57) PCPs' assessment of cancer risk at referral was not influenced by 'gut feeling'.(45)

First diagnostic impressions (n=2)

In a vignette study, when PCPs first diagnostic impressions were of possible cancer, they were more likely to recommend appropriate investigation (OR 1.98 CI 1.10-3.57, p≤0.01).(25) Further, odds of a cancer diagnosis doubled when PCPs verbalized cancer as possible after reading the initial reason for the patients presentation, compared to when they did not mention cancer.(25) Failure to reconsider initial diagnostic hypotheses could herald suboptimal testing strategies and delayed colorectal cancer diagnosis. False reassurance from positive response to medication, intermittent symptoms, misleading test results, and comorbidity were reported to contribute to why PCPs did not consider alternative diagnostic hypotheses.(59)

Assessment of cancer risk (n=2)

The relationship between PCPs' assessment of cancer risk and use of investigations was unclear. In a prospective cross-sectional study, higher estimation of cancer risk did not result in higher referral rates but did lead to a concomitant increase in the likelihood of a cancer diagnosis.(40) Findings from a Danish vignette survey suggested that PCPs over-estimate cancer risk, as approximately one-third of PCPs assessed cancer risk in a hypothetical patient to exceed 50%.(45)

Dealing with uncertainty

Attitude to uncertainty and risk (n=4)

Findings for this PCP factor were inconsistent. Survey responses on the Tolerance for Uncertainty scale suggested that PCPs most tolerant of uncertainty were the least confident in the possibility of an underlying cancer at urgent referral, compared to PCPs with the lowest tolerance for uncertainty.(45) PCP responses to the Physician Risk Attitude scale and self-reported levels of anxiety due to uncertainty and concern about bad outcome, assessed using the Physician Reaction

to Uncertainty scale, was not associated with assessment of cancer risk at referral.(45) Another study found no differences in decisions to test for ovarian cancer by attitude to risk.(38) Diagnostic uncertainty viewed as integral to clinical practice,(53) could lead younger doctors who were less comfortable with uncertainty, to investigate unnecessarily.(51,53)

Fear of malpractice (n=4)

The potential influence of fear of malpractice diverged by study type. In two quantitative studies, PCPs concerns about malpractice did not influence selection of tests for ovarian cancer,(38) nor did PCPs consider fear of malpractice to influence their decisions to use urgent referral pathways.(43) Conversely, qualitative studies suggested that previous experience of complaints from patients or colleagues for suboptimal testing decisions could affect the PCPs future diagnostic approach.(56) Concerned about litigation, defensive testing was considered the only strategy available to PCPs for reassuring patients and protecting themselves medico-legally.(50)

Professional role and involvement in continuing medical education

Attitude to role as gatekeeper (n=3)

Evidence for this PCP factor was predominantly descriptive. PCPs believed their role was to act in the best interests of the patient, advocate for patients healthcare needs with hospital specialists,(51) and make appropriate management decisions.(45) Attitudes were not associated with PCPs' assessment of cancer risk.(45) In their role as gatekeeper, PCPs had mixed encounters with hospital specialists when trying to make a referral on the basis of 'gut feeling', describing productive dialogue with specialists via telephone ahead of a referral, but a reluctance to write 'gut feeling' into a referral letter.(50)

Participation in continued Medical Education (n=4)

In one US vignette study, PCPs involved in clinical teaching were 1.04-fold more likely to use relevant tests for ovarian cancer, compared to PCPs not participating.(38) Across three interview studies, PCPs' wanted more educational opportunities to better differentiate symptoms that could be due to cancer. Frequent training on the latest evidence regarding the predictive value of symptoms for cancer was viewed to be important.(51) Lack of clarity about when to suspect cancer was believed to have contributed to incorrect non-referral of patients in their practice.(54,58)

Discussion

Summary of principal findings

We identified a number of non-modifiable and modifiable PCP factors potentially influencing testing and referral decisions for cancer. Of the PCP factors deemed modifiable, we found most evidence for 'clinical judgement' and 'gut feeling'. PCP judgement of symptoms as suspicious or 'alarm' led to more investigations for possible cancer than symptoms judged to be non-alarm. The presence of 'gut feeling' at referral increased the likelihood of a subsequent cancer diagnosis. PCPs relied on 'gut feeling' to guide decisions to investigate non-specific symptoms. Patients' symptoms could be difficult to reconcile with clinical guidelines, which offered limited scope for PCPs to act on 'gut feeling' or clinical judgment when symptoms were non-specific or did not fit referral criteria. Female PCPs tended to investigate cancer more than male PCPs. The effect of years of experience was inconclusive, but more experience was perceived by PCPs to improve the reliability of 'gut feeling'. The evidence for the remaining PCP factors was insufficient to derive clear conclusions.

Strengths and limitations

To date, research evaluating possible reasons for inconsistencies in PCPs use of suspected cancer referral pathways(20) have primarily focused on the influence of patient,(68) practice,(16,69) and health-system factors.(70,71) By summarizing the available evidence for PCP factors, this review builds on existing knowledge regarding the range of factors affecting PCPs decisions to investigate cancer, and integrates PCPs' perspectives regarding the potential value of those factors in clinical practice. Our search strategy covered a variety of terms for PCP factors increasing the likelihood that all relevant studies were identified. We used a transparent approach to derive PCP factors, which could be useful for developing uniform definitions of these factors that can be applied in future primary care cancer research.

However, this study does have some limitations. While we attempted to synthesize data for PCP factors representing similar constructs, heterogeneity in labelling of PCP factors between studies, which were seldom defined, may have affected construct validity. For example, studies reported sex and gender but did not adequately define these terms or describe how they were assessed, making it difficult to determine which construct was actually measured.(72,73) Additionally, we did not limit evaluation to PCP factors assessed as a primary or secondary outcome, so findings for years of experience, age, and sex, should be interpreted with some caution, as these factors were typically examined *ad hoc* and may not be sufficiently powered to detect a true relationship with testing and referral decisions. Clinical decision-making occurs in the context of the patient agenda

and wider health system.(74,75) A number of included studies did not adjust analyses for confounding from patient or health system factors, potentially obscuring the true relationship of demographic characteristics, 'clinical judgement', attitude to uncertainty/risk, and fear of malpractice with testing and referral decisions. We dichotomized factors as modifiable or non-modifiable which may obfuscate the complex and dynamic mechanisms through which PCP factors influence decision-making, and the extent to which these factors may be modifiable in practice. For example, PCPs' interpretation of symptoms as 'alarm' or non-serious ('clinical judgement') may be due to variations in PCP knowledge of suspected cancer guidelines or application of decisional shortcuts (e.g., availability heuristic, overconfidence). Although there is scope for each of these components to be improved through clinical updates or metacognitive practices that highlight the impact of errors in cognition on patient care,(76–78) knowledge of clinical guidelines may be more susceptible to adjustment than PCP decision-making style which is influenced by personality traits.(79) 'Gut feeling' is associated with empathy,(80) and underpinned by clinical knowledge and experience.(81,82) Thus, the ability of PCPs to access and leverage this feeling in interventions for enhancing empathy may be more instinctive to some PCPs than others.(83) Clinical experience may be more important than the number of years PCPs have accrued in practice. Years of experience is not modifiable but there is interest in the potential for enhancing clinical experience to improve decision-making through simulated diagnostic experiences.(84) Finally, the overall conclusiveness of our findings is limited by methodological weaknesses of studies utilizing retrospective (interview, medical record review) methods that are susceptible to recall, social desirability, and hindsight bias.

Comparison with existing literature

That PCPs were more likely to investigate symptoms judged to be 'alarm' or suspicious of cancer comports with existing literature reporting longer diagnostic intervals for symptoms of lower predictive value not meeting urgent referral criteria across ten cancer sites,(85) and non-investigation or delayed investigation of gynaecological cancers.(86) That PCPs experienced difficulties determining the appropriateness of urgent referral when patients' symptoms did not clearly match referral criteria builds on existing knowledge to understand reasons for reported variations in PCP adherence to clinical guidelines.(87) A systematic review and meta-analysis examining the influence of 'gut feeling' in the diagnosis of cancer similarly found 'gut feeling' to be predictive of cancer (OR 4.24 CI 2.26 to 7.94), and linked with PCP experience.(82) Additionally, 'gut feeling' led to more referrals,(82) which was inconsistent with our findings, and may be due (in

part) to differing interpretation of 'suspicion of cancer' which we labelled clinical judgement as opposed to 'gut feeling'.(47) Qualitative studies of cancer patients symptoms appraisal suggest that patients use vocabulary to communicate symptoms that differ from biomedical symptoms,(88–90) which to some extent may explicate the difficulties PCPs experienced reconciling patients' symptoms with those in referral criteria. A cross-sectional observational study of English practices found that practices with majority male PCPs were less likely to urgently refer for suspected cancer,(91) which accords with our finding that female PCPs were more inclined to investigate.

Unanswered questions and future research

Future research in this area would benefit from clear definition and reporting of PCP factors to ensure findings reflect the intended construct. In the UK, the amalgam "clinical suspicion/gut feeling" is indicated for expedited referral to novel multidisciplinary diagnostic centres (MDCs) for non-specific symptoms, and NICE (2015) suspected cancer guidelines recommend "clinical judgement" is used when making appropriate testing decisions.(11,34) However, the extent to which clinical suspicion/judgement (labelled 'clinical judgement' in this review) and 'gut feeling' reflect distinct or overlapping constructs is unclear. Greater conceptual clarity may enable PCPs to better interpret patients' symptoms and their clinical impression in relation to referral criteria, and help PCPs to more effectively act on clinical recommendations in practice. This could be facilitated by research that elucidates how PCPs appraise symptoms as suspicious or non-suspicious to determine patient eligibility for referral. Consideration should also be given to the subsequent impact on decisions to investigate possible cancer when symptoms are not described in biomedical terms. PCP factors are multi-dimensional and comprise components that may be more modifiable than others. Disentangling the potential involvement of the components of 'clinical judgement' and 'gut feeling' would foster more understanding of these constructs and facilitate identification of the components to target in future interventions. Testing and referral decisions may be driven by aspects of PCP gender as opposed to sex, but empirical evidence for the effect of sociocultural versus biological factors is lacking.(92) Since suboptimal decisions to investigate cancer are likely precipitated by a combination of factors,(93) future research should examine possible interactions between PCP factors (and patient and health system factors). This is pertinent given the potential for reverse causality amongst some PCP factors.

Implications for policy and practice

As the majority of studies in this review were conducted in European countries, our findings have most applicability for countries where PCPs act as gatekeepers to specialty services. Our findings have implications for the management of symptomatic patients with cancer in general, rather than a specific cancer type. This is of clinical import as patients' initial presentation may be at lower risk of an individual cancer but higher risk of an underlying cancer overall.(34,94) The finding that PCPs are able to manage patients with 'alarm' symptoms, but rely on 'gut feeling' for symptoms that are non-specific or do not clearly fit urgent referral criteria suggests that in the absence of 'alarm' symptoms or 'gut feeling' PCPs may not investigate cancer. While non-specific symptom pathways (e.g., SCAN pathway in Oxford(95)) currently in development in the UK will be important for circumventing perceived barriers to acting on clinical judgement or 'gut feeling' for non-specific symptoms, supplementary strategies are needed that support PCPs to recognize and investigate patients with non-alarm and non-specific symptoms that need referral. The challenge for policy makers will be to determine how to enhance the utility of existing guidelines for suspected cancer by more clearly operationalizing 'clinical suspicion/judgement' and 'gut feeling', and ensuring recommendations refrain from reinforcing language siloes between patients and PCPs that may create opportunities for suboptimal testing decisions and diagnostic error.

Funding

Data access statement

All research data supporting this review are provided within the article.

Contributors

VH, MJT, JE, and FW designed the study; VH developed and performed the search; VH, AY, and SWDM screened the search results; VH, AY and SA extracted the data; VH drafted the manuscript with input from JUS and FW; All authors critically assessed and contributed to manuscript revisions.

Competing interests

None declared.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 1. PRISMA flow diagram

References

1. Singh H, Schiff GD, Graber ML, *et al.* The global burden of diagnostic errors in primary care. *BMJ Qual Saf.* **2017**;26(6):484–94.

2. Fernholm R, Pukk Härenstam K, Wachtler C, *et al.* Diagnostic errors reported in primary healthcare and emergency departments: A retrospective and descriptive cohort study of 4830 reported cases of preventable harm in Sweden. *Eur J Gen Pract.* **2019**;25(3):128–35.

3. Aaronson EL, Quinn GR, Wong CI, *et al.* Missed diagnosis of cancer in primary care: Insights from malpractice claims data. *J Health Risk Manag.* **2019**;39(2):19–29.

4. Kostopoulou O, Delaney BC, Munro CW. Diagnostic difficulty and error in primary care - A systematic review. *Fam Pract.* **2008**;25(6):400–13.

5. Diagnostic Errors: Technical Series on Safer Primary Care. Geneva: World Health Organization. **2016**.

6. Hamilton W. Cancer diagnosis in primary care. *Br J Gen Pract.* **2010**;60(571):121–7.

7. Swann R, McPhail S, Witt J. Diagnosing cancer in primary care: results from the National Cancer Diagnosis Audit. *Br J Gen Pract.* **2018**;68(666):e63-e72.

8. Koo MM, Swann R, McPhail S, *et al.* The prevalence of chronic conditions in patients diagnosed with one of 29 common and rarer cancers: A cross-sectional study using primary care data. *Cancer Epidemiol.* **2020**;69:101845.

9. Koo MM, von Wagner C, Abel GA, *et al.* The nature and frequency of abdominal symptoms in cancer patients and their associations with time to help-seeking: evidence from a national audit of cancer diagnosis. *J Public Health.* **2018**;40(3):e388–95.

10. Stapley S, Peters TJ, Neal RD, *et al.* The risk of oesophago-gastric cancer in symptomatic patients in primary care: A large case-control study using electronic records. *Br J Cancer.* **2013**;108(1):25–31.

11. National Collaborating Centre for Cancer. Suspected cancer: recognition and referral. NICE guideline. Full guideline. **2015**; Available from: https://www.nice.org.uk/guidance/ng12/evidence/full-guideline-pdf-2676000277

12. Funston G, Van Melle M, Baun MLL, *et al.* Variation in the initial assessment and investigation for ovarian cancer in symptomatic women: A systematic review of international guidelines. *BMC Cancer.* **2019**;19(1):1–13.

13. Neal RD, Din NU, Hamilton W, *et al.* Comparison of cancer diagnostic intervals before and after implementation of NICE guidelines: Analysis of data from the UK General Practice Research Database. *Br J Cancer.* **2014**;110(3):584–92.

14. Jensen H, Tørring ML, Olesen F, *et al.* Diagnostic intervals before and after

implementation of cancer patient pathways - a GP survey and registry based comparison of three cohorts of cancer patients. *BMC Cancer.* **2015**;15(1):1–10.

15. Swann R, Lyratzopoulos G, Rubin G, *et al.* The frequency, nature and impact of GP-assessed avoidable delays in a population-based cohort of cancer patients. *Cancer Epidemiol.* **2020**;64:101617.

16. Mendonca SC, Abel GA, Saunders CL, *et al.* Pre-referral general practitioner consultations and subsequent experience of cancer care: Evidence from the English Cancer Patient Experience Survey. *Eur J Cancer Care.* **2016**;25(3):478–90.

17. Black G, Sheringham J, Spencer-Hughes V, *et al.* Patients' experiences of cancer diagnosis as a result of an emergency presentation: A qualitative study. *PLoS One.* **2015**;10(8):1–17.

18. Abel G, Saunders CL, Mendonca SC, *et al.* Variation and statistical reliability of publicly reported primary care diagnostic activity indicators for cancer: A cross-sectional ecological study of routine data. *BMJ Qual Saf.* **2018**;27(1):21–30.

19. Burton C, O'Neill L, Oliver P, *et al.* Contribution of primary carer organisation and specialist care provider to variation in GP referrals for suspected cancer: ecological analysis of national data. *BMJ Qual Saf.* **2019**;29:274–6.

20. Meechan D, Gildea C, Hollingworth L, *et al.* Variation in use of the 2-week referral pathway for suspected cancer: A cross-sectional analysis. *Br J Gen Pract.* **2012**;62(602):590–7.

21. Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ.* **2009**;339(7716):332–6.

22. King's Fund (2005). An Independent Audit of the NHS under Labour (1997-2005). London.

23. Sheringham J, Sequeira R, Myles J, *et al.* Variations in GPs' decisions to investigate suspected lung cancer: A factorial experiment using multimedia vignettes. *BMJ Qual Saf.* **2017**;26(6):449–59.

24. Whiting P, Toerien M, de Salis I, *et al.* A review identifies and classifies reasons for ordering diagnostic tests. *J Clin Epidemiol.* **2007**;60(10):981–9.

25. Kostopoulou O, Sirota M, Round T, *et al.* The role of physicians' first impressions in the diagnosis of possible cancers without alarm symptoms. *Med Decis Mak.* **2017**;37(1):9–16.

26. Wenghofer EF, Williams AP, Klass DJ. Factors affecting physician performance: Implications for performance improvement and governance. *Health Policy.* **2009**;5(2):141–60.

27. Patient care and the general practitioner. *BMJ.* 1994;309:1144.

28. Ouzzani M, Hammady H, Fedorowicz Z, *et al.* Rayyan-a web and mobile app for systematic reviews. *Syst Rev.* 2016;5(1):1–10.

29. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med.* **2012**;22(3):276–82.

30. Joanna Briggs Institute Critical Appraisal Tools [Internet]. **2016**. Available from: https://joannabriggs.org/research/critical-appraisal-tools.html

31. Jones OT, Jurascheck LC, Van Melle MA, *et al.* Dermoscopy for melanoma detection and triage in primary care: A systematic review. *BMJ Open.* **2019**;9(8):e027529.

32. The Joanna Briggs Institute. The Joanna Briggs Institute Reviewers' Manual 2014: the systematic review of economic evaluation evidence since 2014. Available from: https://nursing.lsuhsc.edu/JBI/docs/ReviewersManuals/Economic.pdf

33. Popay J, Roberts H, Sowden A, *et al.* Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. A product from the ESRC Methods Programme. **2006**.

34. Chapman D, Poirier V, Vulkan D, *et al.* First results from five multidisciplinary diagnostic centre (MDC) projects for non-specific but concerning symptoms, possibly indicative of cancer. *Br J Cancer.* **2020**;123:722-729.

35. Ramanathan SA, Baratiny G, Stocks NP, *et al.* General practitioner referral patterns for women with gynaecological symptoms: A randomised incomplete block study design. *Med J Aust.* **2011**;195(10):602–6.

36. Rogers HL, Dumenci L, Epstein RM, *et al.* Impact of Patient Gender and Race and Physician Communication on Colorectal Cancer Diagnostic Visits in Primary Care. *J Women's Heal.* **2019**;28(5):612–20.

37. Kostopoulou O, Nurek M, Delaney BC. Disentangling the Relationship between Physician and Organizational Performance: A Signal Detection Approach. *Med Decis Mak.* **2020**;40(6):746–55.

38. Goff BA, Matthews B, Andrilla CHA, *et al.* How are symptoms of ovarian cancer managed?: A study of primary care physicians. *Cancer.* **2011**;117(19):4414–23.

39. Harris M, Brekke M, Dinant GJ, *et al.* Primary care practitioners' diagnostic action when the patient may have cancer: an exploratory vignette study in 20 European countries. *BMJ Open.* **2020**;10(10):e035678.

40. Ingeman ML, Christensen MB, Bro F, *et al.* 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. *BMC Cancer.* **2015**;15(1):1–11.

41. Jensen H, Tørring ML, Olesen F, *et al.* Cancer suspicion in general practice, urgent referral and time to diagnosis: A population-based GP survey and registry study. *BMC Cancer.* **2014**;14(1).

42. Jiwa M, Gordon M, Arnet H, *et al.* Referring patients to specialists: A structured vignette survey of Australian and British GPs. *BMC Fam Pract.* **2008**;9:1–7.

43. John SKP, Jones OM, Horseman N, *et al.* Inter general practice variability in use of referral guidelines for colorectal cancer. *Color Dis.* **2007**;9(8):731–5.

44. Kostopoulou O, Nurek M, Cantarella S, *et al.* Referral Decision Making of General Practitioners: A Signal Detection Study. *Med Decis Mak.* 2019;39(1):21–31.

45. Pedersen A, Vedsted P. General practitioners' anticipated risk of cancer at referral and their attitude to risk taking and to their role as gatekeeper. *J Heal Serv Res Policy.* 2015;20(4):210–6.

46. Brownell P, Piccolo F, Brims F, *et al.* Does this lung nodule need urgent review? A discrete choice experiment of Australian general practitioners. *BMC Pulm Med.* 2020;20(1):1–8.

47. Hjertholm P, Moth G, Ingeman ML, *et al.* Predictive values of GPs' suspicion of serious disease: a population-based follow-up study. 2014;(June):346–53.

48. Guldbrandt LM, Rasmussen TR, Rasmussen F, *et al.* Implementing direct access to low-dose computed tomography in general practice method, adaption - And outcome. *PLoS One.* 2014;9(11).

49. Scheel BI, Ingebrigtsen SG, Thorsen T, *et al.* Cancer suspicion in general practice: The role of symptoms and patient characteristics, and their association with subsequent cancer. *Br J Gen Pract.* 2013;63(614):627–35.

50. Friedemann Smith C, Kristensen BM, Andersen RS, *et al.* GPs' use of gut feelings when assessing cancer risk in primary care: a qualitative study. *Br J Gen Pr.* 2020;71(706):e356-e363.

51. Green T, Atkin K, Macleod U. Cancer detection in primary care: insights from general practitioners. *Br J Cancer.* 2015;112(s1):S41–9.

52. Högberg C, Samuelsson E, Lilja M, *et al.* Could it be colorectal cancer? General practitioners' use of the faecal occult blood test and decision making - A qualitative study. *BMC Fam Pract.* 2015;16(1):1–8.

53. Johansen ML, Holtedahl KA, Rudebeck CE. How does the thought of cancer arise in a general practice consultation? Interviews with GPs. *Scand J Prim Health Care.* 2012;30(3):135–40.

54. Kidney E, Greenfield S, Berkman L, *et al.* Cancer suspicion in general practice, urgent referral, and time to diagnosis: a population-based GP survey nested within a feasibility study using information technology to flag-up patients with symptoms of colorectal cancer. *BJGP Open.* 2017;1(3): doi.10.3399/bjgpopen17X101109.

55. Wagland R, Brindle L, James E, *et al.* Facilitating early diagnosis of lung cancer amongst primary care patients: The views of GPs. *Eur J Cancer Care.* 2017;26(3):1–8.

56. Harris M, Frey P, Esteva M, *et al.* How health system factors influence referral decisions in patients that may have cancer: European symposium report. *J Cancer Res Ther.* 2016;4(1):7–10.

57. Dodds W, Morgan M, Wolfe C, *et al.* Implementing the 2-week wait rule for cancer referral

in the UK: General practitioners' views and practices. *Eur J Cancer Care.* **2004**;13(1):82–7.

58. Harris M, Thulesius H, Neves AL, Harker S, Koskela T, Petek D, *et al.* How European primary care practitioners think the timeliness of cancer diagnosis can be improved: A thematic analysis. *BMJ Open.* **2019**;9(9):1–10.

59. van Erp NF, Helsper CW, Olyhoek SM, *et al.* Potential for reducing time to referral for colorectal cancer patients in primary care. *Ann Fam Med.* **2019**;17(5):419–27.

60. Baun MLL, Jensen H, Falborg AZ, *et al.* Ovarian cancer suspicion, urgent referral and time to diagnosis in Danish general practice: A population-based study. *Fam Pract.* **2019**;36(6):751–7.

61. Guldbrandt LM, Fenger-Grøn M, Rasmussen TR, *et al.* The role of general practice in routes to diagnosis of lung cancer in Denmark: A population-based study of general practice involvement, diagnostic activity and diagnostic intervals 21. *BMC Health Serv Res.* **2015**;15(21): doi.1186/s12913-014-0656-4.

62. Stolper E, Van Bokhoven M, Houben P, *et al.* The diagnostic role of gut feelings in general practice A focus group study of the concept and its determinants. *BMC Fam Pract.* **2009**;10:1–9.

63. Pedersen AF, Andersen CM, Ingeman ML, *et al.* Patient-physician relationship and use of gut feeling in cancer diagnosis in primary care: A cross-sectional survey of patients and their general practitioners. *BMJ Open.* **2019**;9(7):1–10.

64. Woolley A, Kostopoulou O. Clinical intuition in Family Medicine: More than First Impressions. *Ann Fam Med.* **2013**;11:60–6.

65. Arroll B, Allan GM, Raina Elley C, *et al.* Diagnosis in primary care: Probabilistic reasoning. *J Prim Health Care.* **2012**;4(2):166–73.

66. Hillen MA, Gutheil CM, Strout TD, *et al.* Tolerance of uncertainty: Conceptual analysis, integrative model, and implications for healthcare. *Soc Sci Med.* **2017**;180:62–75.

67. Van Nieuwenborg L, Goossens M, De Lepeleire J, *et al.* Continuing medical education for general practitioners: A practice format. *Postgrad Med J.* **2016**;92(1086):217–22.

68. Walter FM, Mills K, Mendonça SC, *et al.* Symptoms and patient factors associated with diagnostic intervals for pancreatic cancer (SYMPTOM pancreatic study): a prospective cohort study. *Lancet Gastroenterol Hepatol.* **2016**;1(4):298–306.

69. Joyce K, Zermanos T, Badrinath P. Factors associated with variation in emergency diagnoses of cancer at general practice level in England. *J Public Health.* **2020**;1–8.

70. Rose PW, Rubin G, Perera-Salazar R, *et al.* Explaining variation in cancer survival between 11 jurisdictions in the International Cancer Benchmarking Partnership: a primary care vignette survey. *BMJ Open.* **2015**;5(5):e007212..

71. Harris M, Vedsted P, Esteva M, *et al.* Identifying important health system factors that

influence primary care practitioners' referrals for cancer suspicion: a European cross-sectional survey. *BMJ Open.* **2018**;8(9):e022904.

72. Rich-Edwards JW, Kaiser UB, Chen GL, *et al.* Sex and gender differences research design for basic, clinical, and population studies: Essentials for investigators. *Endocr Rev.* **2018**;39(4):424–39.

73. Tannenbaum C, Greaves L, Graham ID. Why sex and gender matter in implementation research Economic, social, and ethical factors affecting the implementation of research. *BMC Med Res Methodol.* **2016**;16(1):1–9.

74. Balogh EP, Miller BT, Ball JR. Improving diagnosis in health care. Improving Diagnosis in Health Care. **2016**. p1–472.

75. Gillard S, Benson J, Silverman J. Teaching and assessment of explanation and planning in medical schools in the United Kingdom: Cross sectional questionnaire survey. *Med Teach.* **2009**;31(4):328–31.

76. Keijzers G, Fatovich DM, Egerton-Warburton D, *et al.* Deliberate clinical inertia: Using meta-cognition to improve decision-making. *Emerg Med Australas.* 2018;30(4):585–90.

77. Meyer AND, Singh H. Calibrating how doctors think and seek information to minimise errors in diagnosis. *BMJ Qual Saf.* **2017**;26(6):436–8.

78. Whelehan DF, Conlon KC, Ridgway PF. Medicine and heuristics: cognitive biases and medical decision-making. *Ir J Med Sci.* **2020**;189(4):1477–84.

79. del Campo C, Pauser S, Steiner E, *et al.* Decision making styles and the use of heuristics in decision making. *J Bus Econ.* **2016**;86(4):389–412.

80. Pedersen AF, Ingeman ML, Vedsted P. Empathy, burn-out and the use of gut feeling: A cross-sectional survey of Danish general practitioners. *BMJ Open.* **2018**;8(2):1–8.

81. Kristensen BM, Andersen RS, Nicholson BD, *et al.* Cultivating Doctors' Gut Feeling: Experience, Temporality and Politics of Gut Feelings in Family Medicine. *Cult Med Psychiatry.* 2021. doi: 10.1007/s11013-021-09736-3

82. Smith CF, Drew S, Ziebland S, *et al.* Understanding the role of GPs' gut feelings in diagnosing cancer in primary care: a systematic review and meta-analysis of existing evidence. *Br J Gen Pract.* **2020**;70(698):e612–21.

83. Winter R, Issa E, Roberts N, *et al.* Assessing the effect of empathy-enhancing interventions in health education and training: A systematic review of randomised controlled trials. *BMJ Open.* **2020**;10(9):1–12.

84. Zwaan L, El-kareh R, Meyer AND, *et al.* Advancing Diagnostic Safety Research: Results of a Systematic Research Priority Setting Exercise. doi: 10.1007/s11606-020-06428-3

85. Din NU, Ukoumunne OC, Rubin G, *et al.* Age and Gender Variations in Cancer Diagnostic Intervals in 15 Cancers: Analysis of Data from the UK Clinical Practice Research Datalink.

PLoS One. **2015**;10(5).

86. Williams P, Murchie P, Bond C. Patient and primary care delays in the diagnostic pathway of gynaecological cancers: A systematic review of influencing factors. *Br J Gen Pract.* **2019**;69(679):E106–11.

87. Nicholson B, Mant D, Shinkins B, *et al.* International variation in adherence to referral guidelines for suspected cancer: A secondary analysis of survey data. *Br J Gen Pract.* **2016**;66(643):e106–13.

88. Humphrys E, Walter FM, Rubin G, *et al.* Patient symptom experience prior to a diagnosis of oesophageal or gastric cancer: A multi-methods study. *BJGP Open.* **2020**;4(1):1–15.

89. Bankhead CR, Collins C, Stokes-Lampard H, *et al.* Identifying symptoms of ovarian cancer: A qualitative and quantitative study. *BJOG.* **2008**;115(8):1008–14.

90. Mills K, Birt L, Emery JD, *et al.* Understanding symptom appraisal and help-seeking in people with symptoms suggestive of pancreatic cancer: A qualitative study. *BMJ Open.* **2017**;7(9):1–9.

91. Mendonca SC, Abel GA, Gildea C, *et al.* Associations between general practice characteristics with use of urgent referrals for suspected cancer and endoscopies: A cross-sectional ecological study. *Fam Pract.* **2019**;36(5):573–80.

92. Champagne-Langabeer T, Hedges AL. Physician gender as a source of implicit bias affecting clinical decision-making processes: a scoping review. *BMC Med Educ.* **2021**;21(1):1–9.

93. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med.* **2005**;165(13):1493–9.

94. Herbert A, Rafiq M, Pham TM, *et al.* Predictive values for different cancers and inflammatory bowel disease of 6 common abdominal symptoms among more than 1.9 million primary care patients in the UK: A cohort study. *PLOS Med.* **2021**;18(8):e1003708.

95. Nicholson BD, Oke J, Friedemann Smith C, et al. The Suspected CANcer (SCAN) pathway: Protocol for evaluating a new standard of care for patients with non-specific symptoms of cancer. *BMJ Open.* **2018**;8(1):1–8.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

```
                    ┌─────────────────────────┐
                    │  Total records identified │
                    │  through database         │
   Identification   │  searching                │
                    │  n=7,938                   │
                    └─────────────────────────┘
                                │              ┌──────────────────────────┐
                                │─────────────▶│ Duplicates removed        │
                                │              │ manually                  │
                                ▼              │ n=4,135                   │
                    ┌─────────────────────────┐└──────────────────────────┘
                    │  Articles screened after  │
   Screening        │  duplicates removed       │
                    │  n=3,803                   │
                    └─────────────────────────┘
                                │              ┌──────────────────────────┐
                                │─────────────▶│ Articles excluded based   │
                                │              │ on titles & abstracts     │
                                ▼              │ n=3,721                   │
                    ┌─────────────────────────┐└──────────────────────────┘
                    │  Full-text articles       │
                    │  assessed for eligibility │
                    │  n=82                      │
                    └─────────────────────────┘
```

Full-text articles excluded with reasons
n=62

- No GP factors (n=23)
- Not testing or referral (n=15)
- Not cancer (n=4)
- Asymptomatic population (n=3)
- Patient symptom appraisal / help-seeking (n=2)
- Review article (n=2)
- Health service/process evaluation (n=4)
- Conference abstracts /posters (n=10)
- Full paper inaccessible (n=1)

Additional articles identified from reference lists *n*=9
- Included articles (n=8)
- Systematic reviews (n=1)

Final articles included in review
n=29

Articles included in quantitative synthesis
n=19

Articles included in qualitative synthesis
n=10

Identification | Screening | Eligibility | Included

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supplementary material 1 – MEDLINE search strategy

1. primary health care.mp. or exp Primary Health Care/
2. general practice.mp. or exp General Practice/
3. family practice.mp. or exp Family Practice/
4. ((family) adj2 (physician* or doctor* or medicine or practice* or practiti* or healthcare* or health care*)).mp.
5. ((general) adj2 (practice* or practiti*)).mp.
6. gp*.mp.
7. ((primary) adj2 (care* or healthcare*)).mp.
8. exp physicians, family/ or family physician*.mp.
9. **1 or 2 or 3 or 4 or 5 or 6 or 7 or 8**
10. exp bias/ or (bias* or stereotyp* or prejudic* or judg*).mp.
11. first impression*.mp.
12. exp heuristics/ or (heuristic* or mental shortcut*).mp.
13. (knowledge or experience* or competen* or experti?e).mp.
14. (test* adj2 (belie* or attitude* or perce*)).mp.
15. (personality adj2 trait*).mp.
16. (risk adj2 avers*).mp.
17. anxiet*.mp.
18. fear*.mp.
19. psycho*.mp.
20. ((doctor* or practiti* or gp or physician) adj3 confiden*).mp.
21. gut feeling*.mp.
22. gut-feeling*.mp.
23. factor*.mp.
24. characteristic*.mp.
25. satisf*.mp.
26. burnout.mp.
27. stress*.mp.
28. **10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27**
29. (cancer adj3 symptom*).mp.
30. (symptom* adj5 cancer).mp.
31. possib* cancer.mp.
32. suspect* adj3 (cancer or neoplas*).mp.
33. (ambigu* adj3 symptom*).mp.
34. ((low-risk or low risk) adj2 symptom)*.mp.
35. alarm symptom*.mp.
36. (non alarm adj3 symptom*).mp.
37. (non-alarm adj3 symptom*).mp.
38. (suggest* adj3 cancer).mp.
39. **29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38**
40. exp Diagnosis/
41. exp Diagnosis, Differential/
42. manage*.mp.

43. test* adj3 decision*.mp.
44. (diagnos* or detect* or different*).mp.
45. ((order* or request* or refer* or initiat*) adj5 (test* or imaging or radiolog* or investigation*)).mp.
46. triag*.mp.
47. 40 or 41 or 42 or 43 or 44 or 45 or 46
48. 9 AND 39 AND (28 OR 47)

Supplementary material 2. Characteristics of included studies and critical appraisal

Study	Study type	Country	Data collection period	Cancer	Symptoms	Sample N	Methods	PCP Factors	Tests/ referrals	JBI summary score, %
QUANTITATIVE STUDIES										
Baun *et al.* 2019	Retrospective cohort	Denmark	2010-2016	Ovarian	Symptoms recorded in medical records at index consultation	285 patients	PCP questionnaire based on medical record data for OC patients; completed within 2-5 months of identification from national registry	Suspicion of cancer	Urgent CPP referral	59%
Brownell *et al.* 2020	Discrete choice experiment	Australia	NS	Lung	Symptoms with/without a lung nodule on CT with a PanCan Risk of >10%	152 PCPs	Eight randomly selected vignettes completed via online survey platform	Gender	Urgent specialist review for LC	70%
Chapman *et al.* 2020	Cross-sectional	England	2016-2018	Any/serious disease	Non-specific symptoms not meeting NICE referral criteria	2961 patients	Primary care referrals linked with secondary care datasets from ten MDC pilot sites	Clinical suspicion ('gut feeling')	MDC referral	88%
Goff *et al.* 2011	Cross-sectional	USA	NS	Ovarian	Persistent gastrointestinal or genitourinary symptoms consistent with OC	1532 PCPs	Random assignment of a single vignette in a 12-page mailed survey booklet	Age, sex, specialty, years in practice, non-professional experience of cancer, level of risk taking, fear of malpractice, involved in clinical teaching, board certified	Any OC test (CA125, transvaginal ultrasound, pelvic CT)	50%
Guldbrandt *et al.* 2014	Prospective cohort	Denmark	2011-2013	Lung	Respiratory symptoms	133 PCPs 648 patients	PCP referrals and medical records of completed LDCT scans linked with national registry data	Gender, age	Direct referral to LDCT	46%
Hjertholm *et al.* 2014	Prospective cohort	Denmark	2008-2009	Any/serious disease	Any symptoms	404 PCPs 4518 patients	PCP report of patient presentations linked with national registry data of new diagnoses identified within 6 months of index consultation	Suspicion of cancer or serious disease	Diagnostic action taken (tests, referrals, follow-up consultation)	90%

| Study	Study type	Country	Data collection period	Cancer	Symptoms	Sample *N*	Methods	PCP Factors	Tests/ referrals	JBI summary score, %
Harris *et al.* 2020	Cross-sectional	20 European countries	2015-2016	Lung, ovarian, breast, colorectal	Symptoms at low but significant risk of cancer	2086 PCPs	Four vignettes at varied cancer risk completed via online survey platform	Gender, years since graduation	Immediate selection of appropriate tests or referrals	94%
Ingeman *et al.* 2015	Cross-sectional	Denmark	2012-2013	Lung, colorectal, pancreatic, hematopoietic tissue cancer	Presence/absence of 21 symptoms at time of referral	1278 patients	PCP questionnaire of patients referred to NSSC-CPP linked with national registry data	'Gut feeling', estimation of cancer risk at referral	Referral to a diagnostic centre	63%
Jensen *et al.* 2014	Cross-sectional	Denmark	NS	Lung, colorectal, melanoma, breast, prostate	Symptoms recorded in medical records at first presentation	3823 patients	PCP questionnaire based on medical record data for cancer patients identified in national registry	Symptom interpretation	Urgent CPP referral	63%
Jiwa *et al.* 2008	Cross-sectional	Australia and UK	NS	Colorectal	High and low-risk bowel symptoms based on NICE guidelines	133 PCPs	Questionnaire with nine vignettes at varying cancer risk presented as a PCP referral letter	Age, years in practice	Immediate referral; urgent 2WW referral	31%
John *et al.* 2006	Cross-sectional	UK	2004-2005	Colorectal	NS	202 PCPs 175 patients	Postal survey and colorectal cancer database	Fear of litigation, clinical concern	Urgent 2WW referral	33%
Kostopoulou *et al.* 2017	Cross-sectional	UK	2013-2014	Colorectal, Lung, myeloma	Patients with subtle and non-alarm symptoms	90 PCPs	Six interactive vignettes including three cancer cases, completed via online web tool guided over the telephone by study researcher	Years of experience, first diagnostic impressions	Appropriate specialist referral	75%
Kostopoulou *et al.* 2019	Cross-sectional	England	NS	Lung	Persistent or worsening symptoms from: cough, fatigue, appetite/weight loss, raised platelets	216 PCPs	44 interactive vignettes at varying cancer risk completed in two phases over 24hrs via online web tool; order randomized	Gender, years in practice	Appropriate urgent referral	50%
Kostopoulou *et al.* 2020	Cross-sectional	England	NS	Colorectal	Symptoms associated with colorectal cancer	252 PCPs	48 vignettes (half >3% and half <3% cancer risk) delivered in two	Gender, years of experience	Urgent referral	75%

Study	Study type	Country	Data collection period	Cancer	Symptoms	Sample *N*	Methods	PCP Factors	Tests/ referrals	JBI summary score, %
					based on national guidelines		counterbalanced groups 24 hours apart via online survey platform			
Pedersen & Vedsted 2015	Cross-sectional	Denmark	2012	Any	NS	568 PCPs	Questionnaire including a brief description of a patient with possible cancer	Sex, age, anticipated risk of cancer, 'gut feeling', attitude to uncertainty and risk, attitude to role as gatekeeper	Urgent referral to cancer diagnostic services	63%
Ramanathan *et al.* 2011	Cross-sectional	Australia	2009	Cervical, endometrial, ovarian	Symptoms consistent with gynaecological cancers	1402 PCPs	Survey including random allocation of 12 vignettes completed online or by return-post	Years of experience	Specialty referral to gynaecologist/ oncologist	50%
Rogers *et al.* 2019	Cross-sectional	USA	NS	Colorectal	Symptoms consistent with colorectal cancer	207 PCPs 367 patients	Two in-person unannounced standardized patient visits	Relational communication	Appropriate lower GI tests or specialty referral	75%
Scheel *et al.* 2013	Prospective cohort	Norway	2006	Any	Patients with focal or general warning symptoms of cancer	396 PCPs 6321 patients	PCP questionnaire at index consultation for patients with ≥1 warning symptom	Suspicion of cancer	Diagnostic action taken (laboratory tests, imaging, or referral)	73%
Sheringham *et al.* 2016	Factorial experiment	England	2012-2013	Lung	Low, medium and high-risk symptoms concordant with NICE guidelines	227 PCPs	Six vignettes of filmed patient actors completed via interactive online application	Gender, age, years since qualifying, ethnicity, job role	Any LC investigation	75%
QUALITATIVE STUDIES										
Dodds *et al.* 2004	Cross-sectional	UK	NS	Any	NS	331 PCPs	Short postal survey with open-ended questions	Suspicion of cancer	Urgent 2WW referral	35%
Green *et al.* 2015	Interview	England	2012- 2013	Any	NS	55 PCPs	Face-to-face in-depth semi-structured interviews	Suspicion of cancer, continued medical education, attitude to role as gatekeeper	Any investigation, urgent 2WW referral	55%

Study	Study type	Country	Data collection period	Cancer	Symptoms	Sample N	Methods	PCP Factors	Tests/ referrals	JBI summary score, %
Friedemann Smith *et al.* 2020	Interview	England	2019-2020	Any	Non-specific symptoms suspicious of cancer	19 PCPs	Semi-structured interviews conducted face-to-face or via telephone	Years of experience, gut feeling, fear of malpractice, attitude to role as gatekeeper	MDC referral	70%
Harris *et al.* 2016	Consensus building focus group	Spain	2014	Any	NS	8 PCPs	Brainstorming and freewheeling techniques facilitated by mediator during symposium	Fear of litigation	Referral	25%
Harris *et al.* 2019	Cross-sectional	20* European countries	2015- 2016	Lung, Colorectal, ovarian, breast	Symptoms suggestive of lung, colorectal, ovarian, or breast cancer	1833 PCPs	Analysis of open-ended question in online survey	Continued medical education	Any investigation	45%
Högberg *et al.* 2015	Interview	Sweden	NS	Colorectal	NS	9 PCPs	Semi-structured interviews conducted face-to-face or via telephone	Years of experience	Any investigation; iFOBT	65%
Johansen *et al.* 2012	Interview	Norway	2010	Any	NS	11 PCPs	Semi-structured interviews	Suspicion of cancer, 'gut feeling', tolerance for uncertainty, attitude to risk	Any investigation; CT referral	85%
Kidney *et al.* 2017	Interview nested within a feasibility study	England	2014	Colorectal	NS	18 PCPs	Face-to-face semi-structured individual or group interviews	Continued medical education	Urgent 2WW referral	45%
Van Erp *et al.* 2019	Retrospective cohort†	Netherlands	2007-2011	Colorectal	Patients with symptoms directly or indirectly suggestive of colorectal cancer	31 patients	Review of free text/coded primary care data of symptomatic colorectal cancer patients' data from consultations 5-years before diagnosis	First diagnostic impressions	Referral	50%
Wagland *et al.* 2017	Focus group	England	2014	Lung	Nine symptoms experienced by patients with lung cancer listed in the	16 PCPs	Focus groups	Suspicion of cancer, 'gut feeling'	Urgent 2WW referral	55%

Study	Study type	Country	Data collection period	Cancer	Symptoms	Sample *N*	Methods	PCP Factors	Tests/ referrals	JBI summary score, %
					IPCARD study questionnaire					

Abbreviations: *CPP*, Cancer Patient Pathways; *CT*, computed tomography; *CME*, Continuing Medical Education; *GPs*, general practitioners; *iFOBT*=Immunochemical Faecal Occult Blood Test; *IPCARD*=Identify symptoms that Predict Chest And Respiratory Disease study; *JBI*, Joanna Briggs Institute Critical Appraisal; *LDCT*, Low-Density Computerised Tomography; *MDC,* Multi-disciplinary Diagnostic Centre; *NICE*, National Institute for Clinical Excellence; *NS*, not specified; *NSSC-CPP*, Non-specific Symptoms and Signs of Cancer-Cancer Patient Pathway; *PCPs*, Primary Care Physician(s); *USP*, Unannounced Standardised Patient visits; *WSC*, Warning signs of cancer; *2WW*, 2-week wait (fast-track) pathway

*Bulgaria, Croatia, Denmark, England, Finland, France, Germany, Greece, Israel, Italy, The Netherlands, Norway, Poland, Portugal, Romania, Scotland, Slovenia, Spain, Sweden and Switzerland

†Mixed-methods study, but only qualitative data included

Key to JBI score: ≥70% 46-69% 0-45%
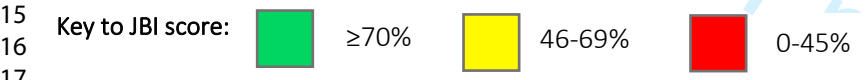
Supplementary material 3. Quantitative findings for the influence of GP factors on testing and referral decisions

GP factor	Study	Cancer	Type of investigation	Subgroup	PCPs N, % (median: IQR)	Testing and referral n, % (mean [SD])	OR/[RR]/PRR (95% CI)	Other, as indicated (95% CI)	Diagnosis Cancer n, %	OR/PRR (95% CI)
NON-MODIFABLE FACTORS										
Demographic characteristics										
Sex	Brownell *et al.* 2020	Lung	Urgent specialist review	F *vs* M			1.87 (1.36-2.56)$^\Psi$			
	Goff *et al.* 2011	Ovarian	Any OC test	F *vs* M		91.7 > 88.1$^{n.s}$				
	Guldbrandt *et al.* 2014	Lung	Direct CT referral vs no referral	F *vs* M		72.3 > 63.7$^{n.s}$				
	Harris *et al.* 2020	Any	Immediate diagnostic action	F	61.1			r=63.6 (58.9 to 68.3)$^{n.s}$		
				M	37.9			r=62.7 (57.9 to 67.6)$^{n.s}$		
	Kostopoulou *et al.* 2020	Colorectal	Urgent referral	F *vs* M				b = 0.18 [0.06 to 0.30]$^\xi$		
	Pedersen and Vedsted 2015	Any	Urgent referral	F *vs* M	47 < 52.6			0.06$^{n.s}$		
	Sheringham *et al.* 2017	Lung	Any LC investigation	F *vs* M		70.48 < 77.12				
Age	Goff *et al.* 2011	Ovarian	Any OC test	30-39 *vs* 40-49; 50-59		92.3 > 87.9; 89.3$^{n.s}$				
	Guldbrandt *et al.* 2014	Lung	Direct CT referral *vs* no referral	54.2 (38-66) > 53.4 (35-68)						
	Jiwa *et al.* 2008	Colorectal	Immediate referral	30-39			5.3 (1.9-14.5)$^\Psi$			
				40-49			2.3 (1.1-5.0)†			
			Urgent referral	30-39			2.8 (1.5-5.2)$^\Psi$			
				40-49			3.2 (1.1-9.3)†			
				50-59			4.6 (1.9-11.4)$^\Psi$			
				60-69			15.4 (4.4-53.8)$^\Psi$			
	Pedersen and Vedsted 2015	Any	Urgent referral	<45 *vs* ≥45 to <60	21.1 < 53.5			$X^{2,a}$: 3.22$^{n.s}$		
				<45 *vs* ≥60	21.1 < 22.5			$X^{2,a}$: 3.68$^{n.s}$		
	Sheringham *et al.* 2017	Lung	Any LC investigation	25-34		70.06				
				35-44		72.89				
				45-54		78.69				

GP factor	Study	Cancer	Type of investigation	Subgroup	PCPs N, % (median: IQR)	Testing and referral n, % (mean [SD])	OR/[RR]/PRR (95% CI)	Other, as indicated (95% CI)	Diagnosis Cancer n, %	OR/PRR (95% CI)
				55-64		75				
				65+		66.67				
Years of experience	Goff et al. 2011	Ovarian	Any OC test	0-10; 11-20; ≥21		90.8 > 89.1; 89.4[n.s]				
	Harris et al. 2020	Any	Appropriate diagnostic action	<10	15.5			r=56.1 (50.5 to 61.6)*		
				10-19	26.9			r=64.3 (59.2 to 69.3)[n.s]		
				20-29	29.2			r=63.6 (58.5 to 68.6)[n.s]		
				30-39	23.9			r=66.1 (61.0 to 71.3)[n.s]		
				≥40	3.6			r=66.3 (58.1 to 74.6)[n.s]		
	Jiwa et al. 2008	Colorectal	Immediate referral	20+			3 (1.3-6.8)[ξ]			
	Kostopoulou et al. 2017	Lung, colorectal, myeloma	Appropriate referral	Increasing experience			0.99 (0.97-1.01)[n.s]			
	Kostopoulou et al. 2019	Lung	Appropriate urgent referral (≥3% PPV)	0-6 (vs 18-36)				b=0.40 (0.22- 0.59)[Ψ]		
				7-10 (vs 18-36)				b=0.21 (-0.02 − 0.40)[n.s]		
				11-17 (vs 18-36)				b=0.15 (-0.05 − 0.34)[n.s]		
			Appropriate urgent referral (≥6.43% PPV)	0-6 (vs 18-36)		(.83 [.35])		b=0.17 (0.03–0.30)[†]		
				7-10 (vs 18-36)		(.77 [.34])		b=0.10(-0.04-0.24)[n.s]		
				11-17 (vs 18-36)		(.83 [.38])		b=0.16 (0.02-0.30)[†]		
	Kostopoulou et al. 2020	Colorectal	Urgent referral	Years of experience	(13: 8-24)			b=-0.0003 (-0.007-0.006)[n.s]		
	Ramanathan et al. 2011	Ovarian	Specialty referral	16-30 vs 1-15			1.1 (0.9-1.4)[Ψ]			
	Sheringham et al. 2017	Lung	Any lung cancer investigation	0-2		71.43				
				2-5		69.14				
				5-10		73.75				
				10-20		77.58				
				20+		76.4				

MODIFIABLE FACTORS

Clinical reasoning processes

Clinical judgement	Baun et al. 2019	Ovarian	CPP (vs No CPP)	Suspicion		65.7 > 31.4				
			CPP (vs No CPP)	No suspicion		34.3 < 72.8	0.35 (0.20-0.60)[Ψ]			
	Hjertholm et al. 2014	Any	Tests in PC clinic	Suspicion vs no suspicion		54.7 > 40.9	1.29 (1.16-1.44)*			
			PC specialist			10.9 > 5	2.35 (1.65-3.33)*			

GP factor	Study	Cancer	Type of investigation	Subgroup	PCPs N, % (median: IQR)	Testing and referral n, % (mean [SD])	OR/[RR]/PRR (95% CI)	Other, as indicated (95% CI)	Diagnosis Cancer n, %	OR/PRR (95% CI)
			Outpatient clinic			15.6 > 4.9	3.27 (2.34-4.56)*			
			Imaging			16.8 > 4.2	3.95 (2.80-5.57)*			
			All referrals			52.3 > 20.6	2.56 (2.22-2.96)*			
	Jensen et al. 2014	Colorectal	Colorectal CPP	Serious vs alarm		13.4 < 54.7	0.27 (0.17-0.41)†			
		Lung	Lung CPP			32.1 < 50.0	0.64 (0.49-0.85)†			
		Breast	Breast CPP			18.5 < 72.8	0.23 (0.09-0.56)†			
		Prostate	Prostate CPP			33.9 < 45.8	0.71 (0.47-1.06)n.s			
		Other	Other CPP referral			15.2 < 41.0	0.34(0.26-0.45)†			
		All cancers	All referral types			20.0 < 52.0	0.4 (0.34-0.48)†			
		Colorectal	Colorectal CPP	Vague vs alarm		22.5 < 54.7	0.4 (0.30-9.54)			
		Lung	Lung CPP			40.9 < 50.0	0.83 (0.65-1.05)n.s			
		Melanoma	Melanoma CPP			32.0 < 42.1	0.76 (0.52-1.13)n.s			
		Breast	Breast CPP			19.4 < 72.8	0.27 (0.17-0.43)†			
		Prostate	Prostate CPP			33.9 < 45.8	0.72 (0.58-0.90)†			
		Other	Other CPP referral			18.0 < 41.0	0.44(0.35-0.55)†			
		All cancers	All referrals			25.9 < 52.0	0.53 (0.48-0.60)†			
	John et al. 2006	Colorectal	Urgent referral	Clinical concern vs no concern	73.6 > 26.4					
	Scheel et al. 2013	Any	Lab tests	Cancer possible vs not likely		30.8 < 57.7			63.5 > 23.1	
			Imaging			40.4 < 49.2			60.6 > 21.2	
			Referral			41.4 < 45.8			73.2 > 9.8	
			1 action			26.6 < 60.7			59.5 > 23.8	
			2 actions			44.7=44.7			72.2 > 13.9	
			3 actions			54.8 > 36.3			50 > 25	
Gut feeling	Chapman et al. 2020	Any	Referral to MDC	As only indication		4.3				
				With symptoms		31.7				
				All occurrences		36.1			Y: 11.2 N: 6.2	1.88 (1.20–2.94)ξ
	Ingeman et al. 2015	Any	Referral	A little vs no		19.2 < 24.6			11.2 < 16	0.65 (0.38-1.10)
				Some vs no		36.4 > 24.6			14.8 < 16	0.86 (0.56-1.31)
				Much vs no		15.6 < 24.6			23.6 > 16	1.55 (0.97-2.48)

GP factor	Study	Cancer	Type of investigation	Subgroup	PCPs N, % (median: IQR)	Testing and referral			Diagnosis Cancer n, %	OR/PRR (95% CI)
						n, % (mean [SD])	OR/[RR]/PRR (95% CI)	Other, as indicated (95% CI)		
				Very much vs no		4.3 < 24.6			34 > 16	2.57 (1.31-5.05)*
	Pedersen and Vedsted 2015	Any	Urgent referral	Use of gut feeling vs no use	52.8 > 47.0			$X^{2,a}$: 0.07 (-0.16-0.30)$^{n.s}$		
First diagnostic impressions	Kostopoulou $et\ al.$ 2017		Appropriate referral	Cancer mentioned vs not mentioned	51 vs 49 instances	1.98 (1.10-3.57)$^{\xi}$			62 vs 25	4.9 (2.72-8.84)$^{\Psi}$
Assessment of cancer risk	Ingeman $et\ al.$ 2015	Any	Referral	21-40% vs 0-20%		16 < 36.8			12.3 > 8	1.43 (0.88-2.33)
				41-60% vs 0-20%		25.8 < 36.8			15 > 8	1.69 (1.12-2.56)*
				61-80% vs 0-20%		12.6 < 36.8			26.5 > 8	2.96 (1.96-4.48)*
				81-100% vs 0-20%		8.6 < 36.8			50 > 8	5.3 (3.62-7.76)*
	Pedersen and Vedsted 2015	Any	Urgent referral	1-14%	20.3					
				15-24%	17.4					
				25-49%	20.6					
				50-74%	28.2					
				75-100%	13.6					
Dealing with uncertainty										
Attitude to uncertainty and risk	Goff $et\ al.$ 2011	Ovarian	Any OC test	Low vs medium; high		90.1; 89.1; 90.2$^{n.s}$				
	Pedersen and Vedsted 2015	Any	Urgent referral	Risk attitude	20.3 < 25.7			-0.07 (-0.38-0.24)$^{n.s}$	(highest vs lowest quartiles)	
				Anxiety due to uncertainty	24.5 < 35.9			$X^{2,a}$: 0.12 (-0.18-0.43)$^{n.s}$		
				Tolerance for ambiguity	23.6 < 26.9			$X^{2,a}$: -0.31 (-0.70 to -0.06)†		
				Concern about bad outcome	21.1 < 37.0			$X^{2,a}$: 0.09 (-0.24-0.43)$^{n.s}$		
Fear of malpractice	Goff $et\ al.$ 2011	Ovarian	Any OC test	Low vs medium; high		90.3; 87.0; 90.7$^{n.s}$				
	John $et\ al.$ 2006	Colorectal	Urgent referral	Fear of litigation vs no fear	8.5 < 91.5					

| GP factor | Study | Cancer | Type of investigation | Subgroup | PCPs N, % (median: IQR) | Testing and referral | | | Diagnosis | |
						n, % (mean [SD])	OR/[RR]/PRR (95% CI)	Other, as indicated (95% CI)	Cancer n, %	OR/PRR (95% CI)
Professional role and involvement in continuing medical education										
Attitude to role as gatekeeper	Pedersen and Vedsted 2015	Any	Urgent referral	Prevent overuse of secondary care *vs* no/neutral	33.6 < 65.1			$X^{2,a}$: 0.14 (-0.11-0.38)**n.s**	(highest *vs* lowest quartiles)	
			Urgent referral	Proper medical guidance & referral *vs* no/neutral	91.7 > 7.9			$X^{2,a}$: 0.23 (-0.15-0.62)**n.s**		
Participation in continued Medical Education (CME)	Goff *et al.* 2011	Ovarian	Any OC test	Involvement *vs* no involvement	92.9 > 87.7		[1.04 (1.01-1.08)]*			

Abbreviations: CA125, cancer antigen 125; CI, confidence interval; CPP, Cancer Patient Pathways; CT, computed tomography; GI, gastrointestinal; GP, general practitioner; GU, genitourinary; F, female; LC, lung cancer; M, male; N, no; OC, ovarian cancer; OR, odds ratio; PR, prevalence ratio; PRR, prevalence rate ratio; ref, reference value; RR, relative risk; SE, standard error; TVU, transvaginal ultrasound, Y, yes

Bold, denotes statistical significance: †$p<0.05$; ξ$p≤0.01$; Ψ$p≤0.001$; *statistically significant, p-value not given

n.s., statistically, non-significant

r, Pearson's correlation coefficient

b, Probit regression coefficient

$X^{2,a}$, Chi-square distribution

t=Student's t-test

Supplementary file 4. Qualitative findings for the influence of PCP factors on testing and referral decisions

PCP factor	Study	Cancer	Author's interpretation	Source quotes
NON-MODIFIABLE FACTORS				
Demographic characteristics				
Sex			*No studies identified*	-
Age			*No studies identified*	-
Years of experience	Friedemann Smith *et al.* 2020	Any	By presenting gut feelings as grounded in clinical knowledge, GPs challenged the notion that gut feelings are 'unscientific' and emphasised the importance of amassing broad clinical experience before gut feelings could be considered reliable	"…the more and more exposures you have to similar cases and different cases, the more basis you have for your gut feeling. And the more informed it is, so I suspect that sort of more experienced clinicians' gut feeling is more refined than more junior clinicians.' (GP10, F, 1-year qualified)
			Some GPs contrasted their current experience with earlier stages of their career or training, expressing growing confidence in making decisions based on their gut feelings. Some more recently qualified GPs anticipated that, like a skill, their gut feelings would become more accurate and their confidence would grow with increasing experience and use	"…my impression is that one becomes more trusting of one's gut feeling as you get more experienced, I think.' (GP09, Male [M], 4-years qualified)
	Hogberg *et al.* 2015	Colorectal	With increasing experience, GPs described being more confident in making decisions, but also becoming more humble. They described feeling more secure about not examining every symptom in detail and learning to harbour uncertainty and live with the fact nothing was certain	"…deciding whether to investigate or not, for example. And, how to follow-up. I've gotten used to it over the years, and it's not a big problem. I can make a decision pretty well and then let it go without it bothering me"
			GPs noted that growing knowledge and experience did not always lead to greater certainty. They described becoming more cautious with a greater awareness of the risks or pitfalls, and perhaps being more generous with referrals for bowel imaging	"you don't always feel more confident just because you have more information…I've become more uncertain about things like iFOBT, for example… I used to think those tests were a lot more help than I do these days"
			Gut-feelings were considered to be based on experience and so they changed over time. With greater experience, it could also be easier to see whether patients diminished their problems	"perhaps you understand people better with time, you understand that some perhaps play down their symptoms because they're scared of what it could be, and that you somehow see through this better with time"
			GPs described becoming less concerned about what others thought about their referral decisions and also more humble. The patients were their focus.	"I've worked for quite a few years … there's no work prestige involved. … I don't really care if there is someone at the other end that laughs at my referral … it's not my problem."
MODIFIABLE FACTORS				
Clinical reasoning processes				
Clinical judgement	Dodds *et al.* 2004	Any	One difficulty expressed by GPs when deciding to refer was interpreting and applying guidelines and deciding when patient's symptoms were suspicious enough to refer, and which specialty to refer to. Problems arose when patients did not fall within the guidelines but for whom the GPs instinct was to be suspicious of cancer. GPs commented that the system was inflexible and the guidelines rigid as they did not offer scope for GPs' own judgement and experience	*None identified*

	Green *et al.* 2015	Any	GPs were aware that many symptoms patients presented with that were indicative of cancer would prove to be non-malignant but at the same time felt under pressure not to 'miss' potential cancer symptoms. Data show the uncertainty that surrounds non-specific symptoms and the skill needed filtering out patients who might be at risk of cancer from those with self-limiting problems	"You don't want to miss anything and being the gatekeeper is great, erm, but sometimes you tend to find that opening that door is easier than keeping it closed because you're always fearful of missing something"
Gut feeling	Dodds *et al.* 2004	Any	One difficulty expressed by GPs when deciding to refer was interpreting and applying guidelines and deciding when patient's symptoms were suspicious enough to refer, and which specialty to refer to. Problems arose when patients did not fall within the guidelines but for whom the GPs instinct was to be suspicious of cancer. GPs commented that the system was inflexible and the guidelines rigid as they did not offer scope for GPs' own judgement and experience	*None identified*
	Friedemann Smith *et al.* 2020	Any	Several GPs said it was unlikely they would ignore a gut feeling. The few examples of when they would act counter to gut feeling were when the gut feeling was reassuring. In this instance, despite their gut feeling, they might still order some tests as the consequences of missing a diagnosis were worse than those of investigating the patient unnecessarily	I feel, my confidence grows in being able to listen to gut feelings that tell me, "Look there's nothing going on here", you don't need to investigate them to the […] nth degree. You can do what seems sensible, and if those things are normal, there is nothing going on here.' (GP14, F, 2-years qualified)
			[…] using their gut feeling to negotiate investigations to rule out disease could also be useful, acknowledging that some patients for whom they had experienced a gut feeling were not diagnosed with cancer, but that this in itself could be valuable. None of the GPs described instances of an incorrect gut feeling for cancer that they felt had been harmful. This was qualified by many who stressed that, while this was the case, it was still necessary to avoid overburdening the system, causing the patient anxiety, and to be 'mindful of not over-investigating people' (GP09, M, 4-years qualified), ensuring that if gut feeling is used, a thorough assessment of the patient is still carried out.	*Embedded in authors' interpretation*
			Several strategies were described to bolster gut feelings, primarily so that GPs' decisions or requests for further investigations would be accepted and avoid 'being led up the garden path' (GP09, M, 4-years qualified). Strategies included building an evidence base through further questioning about symptoms and examining referral guidelines to see how the patient could be fitted to the criteria. Supporting evidence was also sometimes sought by ordering additional tests if the test results on record were unrelated to the current clinical presentation…"	'… we will attempt to put it [gut feeling] in some kind of framework that we think will be recognisable to a specialist nurse or a junior doctor, who's reading the referral in clinic, because we don't want the referral to be dismissed […] We want people to take it seriously.' (GP17, M, 9-years qualified)
	Johansen *et al.* 2012	Any	Suspicion of cancer was sometimes linked to what some GPs called gut feeling or intuition. This mode of knowing was about quickly grasping the essence of the presentation: is it serious or not?...[…] For some, it was the sum of pieces of evidence that together contributed to the feeling that something did not add up.	"it is the sum of all your knowledge, the sum of all your experience … all your knowing from reading updates, attending courses, all the patients you have had whom you have investigated, referred and received feedback about…"

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

			Some GPs recounted negative experiences of referring patients on for biopsies or CT scans which were rejected because of "lack of medical indication" and later revealed as cancers. Suspicion often involved tacit knowing, which was not written into the referral.	"sometimes there... is slightly more behind than...you might say a feeling, an intuition […] yes, maybe a little change of weight and that, nothing big"
	Wagland *et al.* 2017	Lung	Several GPs believed that in the absence of definitive symptoms presented by patients, their gut instinct was their most valuable tool when deciding to investigate for lung cancer.	*None identified*
First diagnostic impressions	Van Erp *et al.* 2019	Colorectal	A subtheme was "the GP omitting to reconsider the initial diagnosis" […] the GP did no further investigating anaemia or rectal blood loss, even after the initial explanatory cause had resolved…[..] The factor most often explaining an alternative working hypothesis was the subtheme 'presence of an explanatory concomitant decision', either pre-existing or detected during the consultation [....]. Other subthemes causing the GP to stick to the original hypothesis were "good symptomatic response to initial therapy" […] "misleading results from additional testing" […], and "intermittent characteristics of the complaints".	*None identified*
Assessment of cancer risk			*No studies identified*	
Dealing with uncertainty				
Attitude to uncertainty and risk	Green *et al.* 2015	Colorectal	[GPs] […] felt under pressure not to 'miss' potential cancer symptoms. Data show the uncertainty that surrounds non-specific symptoms and the skill needed in filtering out patients who might be at risk of cancer from those presenting with self-limiting problems	"You don't want to miss anything and being the gatekeeper is great, erm, but sometimes you tend to find that opening that door is easier than keeping it closed because you're always fearful of missing something"
	Johansen *et al.* 2012	Any	The younger doctors, in particular, felt that living with this uncertainty was not easy. Some doctors mentioned that their precaution could lead to "unnecessary" investigations.	*None identified*
			Consequently, a lot of the GPs' daily work was about "excluding" cancer, and questions of probability and risk.	
Fear of malpractice	Friedemann Smith *et al.* 2020	Any	Several described how primary care is becoming increasingly risk averse and litigious, with investigations often being the only way to provide patients with adequate reassurance. As such, they had become 'fearful' of receiving a complaint and so more inclined to practice defensively	I think my level of tendency to investigate people is probably a bit higher now than it used to be, which is ironic, because I'm more experienced. So, you might think that it had gone the other way, but I am fearful about you know a complaint or so forth […]" (GP08, F, 30-years qualified)
	Harris *et al.* 2016	Any	Fear of litigation, or complaint over a failure to refer, is a significant factor in many countries (Croatia, Slovenia, Spain, Switzerland, UK). Some GPs may be influenced in their decision-making by previous experience of criticism from patients or colleagues when a serious diagnosis was delayed due to late referral. The opposite experience is also possible, with criticism from patients or colleagues who think that the GP should have managed a presenting problem without referral (UK).	*None identified*
Professional role and involvement in continuing medical practice				

Attitude to role as gatekeeper	Friedemann Smith *et al.* 2020	Any	When GPs felt the need to make a clinical decision based on a gut feeling, they often described having discussions directly with the secondary care colleagues to whom they were hoping to refer their patients. Many of the descriptions of this interaction resulted in the consultant agreeing to see the patient or suggesting a more appropriate referral route. Success stories of using this strategy tended to be told by GPs with greater experience:	"If I say to a more senior surgeon or physician, this patient's not well and I'm just not happy managing them in the community, in a way it doesn't matter what the parameters are […] if I'm not happy then they're not happy either, and will take it.'(GP11, F, 26-years qualified) "I have referred a few people in like that [on a gut feeling] before, and I've not had the best response […] I might be being over simplistic but, I would never write, "I've got a gut feeling" on a referral letter." (GP12, F, 2-years qualified)
	Green *et al.* 2015	Colorectal	GPs valued their gatekeeper role and perceived the skill was to identify patients in need of further investigation from those who could be managed within primary care. The ability to perform this role adequately was perceived to be dependent on the GP's role as patients' advocate […].	"our role is gatekeeper, go between, between patients and the hospital, we're there as an intermediary […] guide them in the right direction" "Secondary care relies on proficient general practice to make sure that which needs to get through gets through and hopefully we filter out that which doesn't". "We are the gatekeepers. We are that very first step, so that decision we made right at the beginning can have massive consequences".
Participation in continued Medical Education (CME)	Green *et al.* 2015	Any	Participants were aware that many symptoms patients presented with that were indicative of cancer would prove to be non-malignant but at the same time felt under pressure not to 'miss' potential cancer symptoms. Data show the uncertainty that surrounds non-specific symptoms and the skill needed in filtering out patients who might be at risk of cancer from those presenting with self-limiting problems	"We're trained in a differential diagnosis approach, that means that when somebody comes in with symptoms we think ah, this could be a range of things, it could be something that's self-limiting, something that progresses, it could be something that's extremely serious, it could be something that's serious but is, is treatable, so we're thinking of a range of things and what we do is we, we use time, we use wait, you know, we wait, we see can this resolve on its own".
	Harris *et al.* 2019	Lung Colorectal Ovary Breast	GPs had suggestions on how to improve their own knowledge […] by increasing the amount of training and improving the teaching content. Improving the training of PCPs […] was considered by many respondents to be important	"By educating HCPs… When to suspect and when to do further tests?" "CME on early signs and up-to-date investigation processes"
			Primary care doctors also need to be aware of the relevant clinical pathways	"Training of GPs…taking into account the entire pathway from symptom to diagnosis"
	Kidney *et al.* 2017	Colorectal	In some cases, non-referral of patients with symptoms was explained by GPs' own poor knowledge of colorectal referral guidelines…	"I'll have to look at it and familiarise…. Persistent diarrhoea, what does persistent [mean], you know, how long?"

Abbreviations: CME, Continuing Medical Education; CXRs, chest x-rays; CT, computed tomography; GP, general practitioner; iFOBT, immunochemical Fecal Occult Blood Test; HCPs, Healthcare Professionals; NICE, National Institute for Clinical Excellence; PCP, primary care physician; 2WW, 2-week wait (fast-track) referral pathway […], text truncated

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	2
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	2
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	2
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	3
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	2
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	2
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	3
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	3
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	3
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	3
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	N/A
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	4

PRISMA 2009 Checklist

Page 1 of 2

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	3, 25
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	Not performed
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9-12
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	16-24
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	13-15
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	8
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Not performed
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	25
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	25
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	26
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	26

For more information, visit: **www.prisma-statement.org**.