

## Weight loss as a predictor of cancer in primary care:

a systematic review and meta-analysis

### Abstract

#### Background

Weight loss is a non-specific cancer symptom for which there are no clinical guidelines about investigation in primary care.

#### Aim

To summarise the available evidence on weight loss as a clinical feature of cancer in patients presenting to primary care.

#### Design and setting

A diagnostic test accuracy review and meta-analysis.

#### Method

Studies reporting 2 × 2 diagnostic accuracy data for weight loss (index test) in adults presenting to primary care and a subsequent diagnosis of cancer (reference standard) were included. QUADAS-2 was used to assess study quality. Sensitivity, specificity, positive likelihood ratios, and positive predictive values were calculated, and a bivariate meta-analysis performed.

#### Results

A total of 25 studies were included, with 23 (92%) using primary care records. Of these, 20 (80%) defined weight loss as a physician's coding of the symptom; the remainder collected data directly. One defined unexplained weight loss using objective measurements. Positive associations between weight loss and cancer were found for 10 cancer sites: prostate, colorectal, lung, gastro-oesophageal, pancreatic, non-Hodgkin's lymphoma, ovarian, myeloma, renal tract, and biliary tree. Sensitivity ranged from 2% to 47%, and specificity from 92% to 99%, across cancer sites. The positive predictive value for cancer in male and female patients with weight loss for all age groups ≥60 years exceeded the 3% risk threshold that current UK guidance proposes for further investigation.

#### Conclusion

A primary care clinician's decision to code for weight loss is highly predictive of cancer. For such patients, urgent referral pathways are justified to investigate for cancer across multiple sites.

#### Keywords

cancer; diagnosis; primary care; systematic review; weight loss.

### INTRODUCTION

Weight loss is a non-specific symptom posing a diagnostic challenge to clinicians in non-specialist settings such as primary care. It can be associated with several cancer and non-cancer conditions.<sup>1</sup> In relation to cancer, two main diagnostic groupings exist: patients with additional clinical features, such as haemoptysis, which can focus diagnostic efforts; and patients without such a pointer in whom, if cancer is suspected, the clinician must consider several possible sites.<sup>2</sup>

There are no international or national clinical guidelines to support primary care physicians in how to respond to or investigate patients who present with weight loss, although several clinical reviews have suggested different approaches, most commonly in older populations.<sup>1,3–6</sup> In the UK, guidance for suspected cancer from the National Institute for Health and Care Excellence (NICE) reported weight loss to be a feature of seven cancers in primary care, citing positive predictive values (PPVs) of 0.4–3.0%, and a cumulative PPV of 7.1% across all cancer sites, a figure considerably above the 3% threshold used by NICE in formulating their urgent investigation guidance.<sup>7</sup> The methods used by NICE will have incorrectly estimated the true PPV for weight loss, because for several cancer sites no evidence could be

found: summary PPVs for each cancer site were added together, meaning any possible contribution to the total PPV from a cancer site as yet unreported was omitted.

People who have lost more weight are more likely to have cancer than those who have lost less,<sup>8</sup> but almost all evidence on this comes from specialist care.<sup>1</sup> Furthermore, weight loss has been previously considered as a feature of advanced cancer only, although reports about weight loss and stage at diagnosis give conflicting results. Studies of colorectal, pancreatic, and lung cancer have reported that even people with early-stage cancer may present with weight loss,<sup>9–11</sup> and yet other data show no relationship between weight loss and stage or mortality from colorectal cancer.<sup>12</sup>

It is also unclear whether weight loss might occur in the period immediately before a diagnosis is established, or be a symptom that occurs well before cancer is manifest. In one study, patients with pancreatic cancer had a similar body mass index (BMI) to controls without cancer, suggesting weight loss has not continued for long.<sup>13</sup> Other evidence, however, suggests that some patients experience symptoms for some time before consulting with insidious weight loss, misattributed to normal weight fluctuations due to changes in diet and exercise.<sup>14,15</sup>

It is possible that non-specialist doctors

**BD Nicholson**, MSc, MRCP, clinical researcher; **J O'Sullivan**, MBBS, AHEA, clinical researcher; **P Aveyard**, PhD, MRCP, FRCGP, FFFPH, professor of behavioural medicine; **FD Hobbs**, FRCGP, FMedSci, head of department and professor of primary care, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford; **NIHR Oxford Biomedical Research Centre**, Oxford University Hospitals NHS Foundation Trust. **W Hamilton**, MD, FRCGP, FRCP, professor of primary care diagnostics, Medical School, University of Exeter, Exeter.

#### Address for correspondence

Brian D Nicholson, Nuffield Department

of Primary Care Health Sciences, Radcliffe Observatory Quarter, University of Oxford, Oxford OX2 6GG, UK.

**Email:** brian.nicholson@phc.ox.ac.uk

**Submitted:** 3 September 2017; **Editor's response:** 13 November 2017; **final acceptance:** 5 December 2017.

#### ©British Journal of General Practice

This is the full-length article (published online 10 Apr 2018) of an abridged version published in print. Cite this version as: **Br J Gen Pract 2018; DOI: <https://doi.org/10.3399/bjgp18X695801>**

## How this fits in

No systematic review and meta-analysis has examined the association between weight loss and cancer in primary care populations. The sensitivity of unexpected weight loss for cancer is low, demonstrating that many primary care patients with cancer do not experience weight loss. Conversely, the specificity is high. Patients aged  $\geq 60$  years with unexpected weight loss have  $>3\%$  chance of having cancer in one of 10 cancer sites. Investment in rapid diagnostic pathways to urgently investigate weight loss across a number of cancer sites is justified.

do not recognise weight loss as a potential symptom of cancer and omit investigation until other symptoms occur. Patients with pancreatic cancer with weight loss as their first symptom had the longest time to diagnosis and poorest survival.<sup>16</sup> This may reflect the lack of guidelines on investigating weight loss, and the lack of a clear pathway to refer patients with isolated weight loss for investigation.

The objectives of this systematic review were, first, to examine the diagnostic value of weight loss — alone and in combination with other clinical features for cancer in primary care patients — and, second, to examine how the predictive value of weight loss varies by cancer type, cancer stage, sex, and age.

## METHOD

### Protocol and registration

A protocol was registered with PROSPERO (review number CRD42016038963).<sup>17</sup> The authors followed standard Cochrane methodology.<sup>18</sup> Reporting followed PRISMA, adapted to suit a systematic review and meta-analysis of diagnostic studies.<sup>19</sup>

### Patient involvement

A survey of patients with experience of a cancer diagnosis informed the choice of methodology.

### Eligibility criteria

Studies were included if  $2 \times 2$  diagnostic accuracy data could be extracted for the symptom of weight loss and cancer in adults ( $\geq 18$  years) presenting to primary care. No restrictions were placed on the definition of weight loss. A cancer diagnosis of any type or stage was included if confirmed by the cancer registry, histology, the general practice electronic record, or by another reliable source judged appropriate by the researchers.<sup>20–24</sup> A

cancer diagnosis within 5 years of the weight loss was permitted, although only one study extended beyond 2 years.

### Information sources and search strategy

The electronic databases EMBASE (OvidSP) (1974 to 23 October 2015), the Science Citation Index & Conference Proceedings Citation Index (Web of Science Core Collection) (1946–present), and MEDLINE (OvidSP) (1946–present) were searched on 25 October 2015. Weight loss terms, general symptom terms, and cancer terms were used, along with primary care setting terms to identify potentially relevant studies (further information available from the authors on request). Duplicates, studies in non-human animals, case reports, letters, comment articles, and editorials were excluded. No language limits were applied. The reference lists of retrieved reviews and all included studies were checked and a 'Related articles' PubMed search conducted. All principal investigators were contacted to request additional relevant data.

### Study selection and data extraction

Two review authors independently screened all titles and abstracts, assessed full-text articles for those deemed relevant based on title and abstract, and where a decision could not be made at screening, and extracted data from eligible studies using an extraction guide refined after initial piloting. A third review author resolved any discrepancies, leading to 59 exclusions following full-text review.

### Quality assessment

QUADAS-2 was used to assess methodological quality.<sup>25</sup> QUADAS-2 items were weighted to address key potential biases (further information available from the authors on request).

### Analyses

Sensitivity, specificity, positive likelihood ratios (PLRs), and diagnostic odds ratios (DOR) were calculated with corresponding 95% confidence intervals (CIs).<sup>18,26</sup> Review Manager 5.3 was used to produce forest plots to visually show the heterogeneity in sensitivity and specificity across primary studies, and to plot sensitivity and specificity estimates from each study in receiver operating characteristic (ROC) space. A bivariate meta-analysis model was used to calculate pooled estimates for sensitivity, specificity, and positive likelihood ratios (PLRs)<sup>27</sup> for subgroups including four or more studies, using Stata (version 13.1).<sup>28–33</sup>

Positive predictive values (PPVs) were

calculated using Bayes's theorem by multiplying the PLR by the prior odds.<sup>32</sup> Prior odds were derived using the cancer incidence (the pre-test probability) for the age ranges included in each study.<sup>7,34</sup> To give a percentage risk of any cancer, cumulative PPVs were calculated by adding together PPVs for each tumour site, and presented by age group and sex.<sup>7</sup> The highest and lowest cumulative PPVs were calculated by adding the PPVs calculated with the highest and lowest likelihood ratios (LRs) respectively.

The authors conducted a planned bivariate meta-analysis of the predictive value of weight loss by individual cancer site, cancer stage, sex, and age group. Two sensitivity analyses were pre-specified based on QUADAS-2 assessment: one including studies with a low risk of bias, and a second including studies where the risk of bias in only one domain was high. Investigation of publication bias in diagnostic accuracy studies is not recommended by Cochrane, and so this was not assessed.<sup>35-37</sup>

## RESULTS

### Study characteristics

Twenty-five included studies were published between 1994 and 2015. One study was conducted in the US and all others in the UK. The sample size ranged from 398 to 2 140 194 participants, median 43 791 (interquartile range = 8476–1 013 262). Table 1 details each included study.

### Data source

Eight studies used routinely collected electronic health record data from The Health Improvement Network (THIN) database,<sup>8,38-44</sup> seven from the QCancer® database,<sup>22,23,45-49</sup> and four from the Clinical Practice Research Datalink (CPRD).<sup>50-53</sup> These studies used the symptom and diagnosis codes physicians use; uncoded free-text information from the notes was not accessed. Three studies used a combination of written and electronic GP records, coded using the International Classification of Primary Care-2, blinded to the cancer outcome.<sup>54-56</sup> Two used

structured patient questionnaires.<sup>57,58</sup> One included data from three sources: patient questionnaire, patient telephone interview, and the GP record.<sup>59</sup>

### Study design

In all, 13 case-control studies and 12 cohort studies were included. The ratio of cases to controls ranged from 1:1 to 1:10.<sup>8,43,44,50-59</sup> Control patients were matched for age,<sup>8,44,50-58</sup> sex,<sup>8,44,50-58</sup> GP practice,<sup>8,43,44,50-56</sup> and consultation year,<sup>44</sup> by using adjacency on the GP register,<sup>57</sup> using Health Care Financing Administration lists,<sup>58</sup> at random using the electronic record,<sup>8,43,44,50-56</sup> after attending ovarian screening, or being referred for a suspected benign adnexal mass.<sup>59</sup> Cohort entry was defined as the first occurrence of weight loss for the exposed group or, for the unexposed, study entry or 12 months after registering with the practice, if this was later.<sup>22,23,38-42,45-49</sup> One study, at high risk of ascertainment bias, used 5 years.<sup>58</sup>

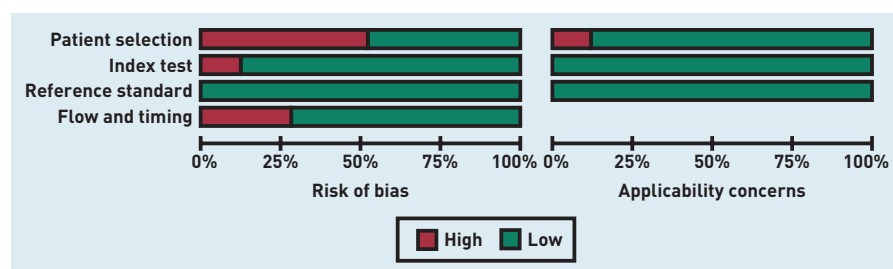
### Weight loss

There was marked heterogeneity in the definition of weight loss. One study used objective weight measurements, calculating the change between the weight measurement nearest to the date of diagnosis and the highest weight in the previous 2 years.<sup>8</sup> All other database studies defined weight loss using a coded entry in the electronic GP record.<sup>22,23,38-53,59</sup> The code lists used to define weight loss were not published, and no study described the method used to discriminate weight loss from planned weight loss [for example, through dieting or exercise]. In three studies 'any mention of weight loss' was coded after review of the entire coded and non-coded GP record.<sup>54-56</sup> Questionnaire studies defined weight loss as 'subjective weight loss',<sup>57</sup> 'weight loss (unplanned)', or the appearance of weight loss,<sup>59</sup> and 'unintended weight loss ≥6 weeks duration'.<sup>58</sup>

### Cancer type and stage

The studies reported weight loss in association with 10 cancer types: colorectal,<sup>8,38,48,55,57</sup> pancreatic,<sup>40,44,49,52,58</sup> gastro-oesophageal,<sup>41,45,53</sup> ovarian,<sup>42,46,59</sup> lung,<sup>43,54</sup> renal tract,<sup>39,47</sup> myeloma,<sup>50</sup> non-Hodgkin's lymphoma,<sup>51</sup> biliary tree,<sup>44</sup> and prostate.<sup>56</sup> Two studies identified all cancer types in each of the two sexes.<sup>22,23</sup> Several methods were used to confirm the diagnosis of cancer. The THIN and CPRD used UK diagnostic Read codes inputted by primary care physicians.<sup>38-44,50-53,59</sup> The QCancer studies used diagnostic codes and Office

**Figure 1. Graphical summary of QUADAS-2 assessment of risk of bias and applicability concerns presented as percentages across all 25 included studies.**



**Table 1. Characteristics of the included studies**

Design				
Cancer type (Reference) Population	Data source Recruitment Follow-up	Index test — weight loss Reference standard — cancer	Inclusion criteria	Exclusion criteria
<b>Colorectal</b> (Collins and Altman, 2012) <sup>38</sup> Primary care, UK	<b>Cohort</b> THIN database 01/01/00–30/06/08 2 years following entry	Weight loss Read code Relevant cancer Read code	THIN record, starting latest of, study start date, registration, first symptom	Prior colorectal cancer; GP registered <12 months, invalid dates, <30/≥85 years
<b>Colorectal</b> (Curless et al, 1994) <sup>37</sup> Primary and secondary care, North East England	<b>Case-control (1:1)</b> Structured patient questionnaire 01/1989–30/09/90 1 year pre-diagnosis	Patient reported 'subjective weight loss' Histological diagnosis of cancer	Cases: consecutive cases identified within <2 weeks of diagnosis by hospital surgical residents Controls: patients adjacent on GP register for age and sex	Prior colorectal cancer or inflammatory bowel disease, 'non-white', death ( <i>n</i> = 15), outside area ( <i>n</i> = 10), no consent ( <i>n</i> = 2), unclear diagnosis ( <i>n</i> = 1), GP advised against ( <i>n</i> = 4), refused ( <i>n</i> = 43), unable to contact ( <i>n</i> = 6)
<b>Colorectal</b> (Hamilton et al, 2005) <sup>35</sup> Primary care, Exeter, England	<b>Case-control (1:5)</b> GP record (written and electronic) 1998–2002 (diagnosis) 2 years pre-diagnosis	Any mention of weight loss in GP record Cancer registry, histology reports, and specialist case records	Cases: identified by local cancer registry, histology reports, or specialist diagnosis Controls: practice, sex, and age matched to cases, and alive at index diagnosis	Unobtainable records ( <i>n</i> = 96), no entry in the records in 2 years pre-diagnosis ( <i>n</i> = 111), prior colorectal cancer ( <i>n</i> = 32), lived outside Exeter at diagnosis ( <i>n</i> = 22)
<b>Colorectal</b> (Hamilton et al, 2009) <sup>36</sup> Primary care, UK	<b>Case-control (1:7)</b> THIN database 01/2001–07/2006 (diagnosis) 2 years pre-diagnosis	Difference between last recorded weight and highest weight in previous 2 years (≥ 10% or 5–10% weight loss) Cancer recorded in THIN	Cases: THIN code for colorectal cancer Controls: practice, sex, year of birth matched	Prior colorectal cancer; <2 years of records before the index date
<b>Colorectal</b> (Hippisley-Cox and Coupland, 2012) <sup>34</sup> Primary care, UK	<b>Cohort</b> QResearch database 01/01/00–30/09/10 2 years following entry	Weight loss code UK diagnostic code in GP record and ONS cause of death ICD-9 (I53 or I54) or ICD-10 (C18–C21) codes	EMIS record, starting latest of, study start date, 12 months after practice registration, first recorded loss rectal bleeding, loss of appetite, weight loss, or abdominal pain	No Townsend score ( <i>n</i> = 195 305). Prior colorectal cancer ( <i>n</i> = 8414). Rectal bleeding, loss of appetite, weight loss, or abdominal pain in prior 12 months ( <i>n</i> = 89 572)
<b>Gastro-oesophageal</b> (Collins and Altman, 2013) <sup>41</sup> Primary care, UK	<b>Cohort</b> THIN database 01/01/00–30/06/08 2 years following entry	Weight loss Read code Relevant cancer Read code	THIN record, starting latest of, study start date, registration, first red-flag symptom of gastro-oesophageal cancer	Prior gastro-oesophageal cancer, GP registered <12 months, invalid dates, <30/≥85 years
<b>Gastro-oesophageal</b> (Hippisley-Cox and Coupland, 2011) <sup>35</sup> Primary care, UK 2 years following entry	<b>Cohort</b> QResearch database 01/01/00–30/09/10 or ICD-10 (C15 or C16) codes	Weight loss code UK diagnostic code in GP record and ONS cause of death ICD-9 (I50 or I51) or ICD-10 (C15 or C16) codes	EMIS record, starting latest of, study start date, 12 months after practice registration, first recorded dysphagia, haematemesis, appetite loss, weight loss, or abdominal pain	No Townsend score ( <i>n</i> = 195 305). Gastro-oesophageal cancer ( <i>n</i> = 13 777). Dysphagia, haematemesis, loss of appetite, weight loss, and abdominal pain in prior 12 months ( <i>n</i> = 89 572)
<b>Gastro-oesophageal</b> (Stapley et al, 2013) <sup>33</sup> Primary care, UK	<b>Case-control (1:5)</b> CPRD database 01/2000–12/2009 >1 year pre-diagnosis	GP weight loss Read code 24 gastric and 18 oesophageal Read codes	Cases: gastro-oesophageal cancer code Controls: age, sex, and practice matched	No records <1 year before index date ( <i>n</i> = 3335), metastatic cancer ( <i>n</i> = 10), cases with past gastrooesophageal cancer ( <i>n</i> = 28), cases without controls ( <i>n</i> = 17), oesophageal cancer used as gastric ( <i>n</i> = 131), controls are cases (262), controls without cases (809), duplicate control ( <i>n</i> = 427)

... continued

**Table 1 continued. Characteristics of the included studies**

<b>Lung</b> (Hamilton <i>et al</i> , 2005) <sup>154</sup> <i>Primary care, Exeter, England</i>	<b>Case-control (1:5)</b> GP record (written and electronic) 1998–2002 (diagnosis) 2 years pre-diagnosis	Any mention of weight loss in GP record <i>Cancer registry, histology reports, and specialist case records</i>	Cases: identified by local cancer registry, histology reports, or specialist diagnosis Controls: practice, sex, and age matched to cases, and alive at index diagnosis	Unobtainable records ( <i>n</i> = 77), no records in 2 years pre-diagnosis ( <i>n</i> = 98), prior lung cancer ( <i>n</i> = 9), lived outside Exeter at diagnosis ( <i>n</i> = 15), unconfirmed or atypical ( <i>n</i> = 28), metastatic cancer unknown origin ( <i>n</i> = 7)
<b>Lung</b> (Iyen-Omofoman <i>et al</i> , 2013) <sup>143</sup> <i>Primary care, UK</i>	<b>Case-control (1:10)</b> THIN database 01/01/00–28/07/09 4–24 months pre-diagnosis	GP weight loss code <i>Cancer recorded in THIN</i>	Cases: THIN code for lung cancer Controls: practice matched to cases	<1 year of records prior to lung cancer diagnosis, <40 years at index diagnosis ( <i>n</i> = 59)
<b>Multiple, male</b> (Hippisley-Cox and Coupland, 2013) <sup>122</sup> <i>Primary care, UK</i>	<b>Cohort</b> QResearch database 01/01/00–01/04/12 2 years following entry	Weight loss code <i>UK diagnostic code in GP record and ONS cause of death ICD-9 or ICD-10 code</i>	EMIS record for at least a year	No Townsend score ( <i>n</i> = 124 743). Red-flag symptom in the 12 months prior to study entry ( <i>n</i> = 91 819)
<b>Multiple, female</b> (Hippisley-Cox and Coupland, 2013) <sup>123</sup> <i>Primary care, UK</i>	<b>Cohort</b> QResearch database 01/01/00–01/04/12 2 years following entry	Weight loss code <i>UK diagnostic code in GP record and ONS cause of death ICD-9 or ICD-10 code</i>	EMIS record for at least a year	No Townsend score ( <i>n</i> = 132 945). Red-flag symptom in the 12 months prior to study entry ( <i>n</i> = 165 310)
<b>Myeloma</b> (Shephard <i>et al</i> , 2015) <sup>150</sup> <i>Primary care, UK</i>	<b>Case-control (1:5)</b> CPRD database 01/2000–12/2009 >1 year pre-diagnosis	GP weight loss Read code <i>23 myeloma Read codes</i>	Cases: multiple myeloma code Controls: age, sex, and practice matched	No records <1 year before index date ( <i>n</i> = 1201), cases without controls ( <i>n</i> = 1), controls with myeloma ( <i>n</i> = 8), controls with no record after registration ( <i>n</i> = 13), cases with para-proteinaemia >60 days before myeloma code ( <i>n</i> = 26), control of para-proteinaemia case
<b>Non-Hodgkin's lymphoma</b> (Shephard <i>et al</i> , 2015) <sup>151</sup> <i>Primary care, UK</i>	<b>Case-control (1:5)</b> CPRD database 01/2000–12/2009 >1 year pre-diagnosis	GP weight loss Read code <i>106 non-Hodgkin's lymphoma (nHL) Read codes</i>	Cases: nHL code Controls: age, sex, and practice matched	Hodgkin's lymphoma (HL) ( <i>n</i> = 283), control of HL ( <i>n</i> = 1402), case of mycosis fungoides or Sézary syndrome ( <i>n</i> = 147), control of mycosis fungoides or Sézary syndrome ( <i>n</i> = 719); no records <1 year before index date (2018); cases without controls (7); controls with non-HL ( <i>n</i> = 70); controls with no record after registration ( <i>n</i> = 26)
<b>Ovarian</b> (Collins and Altman, 2013) <sup>142</sup> <i>Primary care, UK</i>	<b>Cohort</b> THIN database 01/01/00–30/06/08 2 years following entry	Weight loss Read code <i>Relevant cancer Read code</i>	THIN record, starting latest of study start date, registration, first red-flag symptom of ovarian cancer	Prior ovarian cancer, GP registered <12 months, invalid dates, <30/≥85 years
<b>Ovarian</b> (Hippisley-Cox and Coupland, 2012) <sup>146</sup> <i>Primary care, UK</i>	<b>Cohort</b> QResearch database 01/01/00–30/09/10 2 years following entry	Weight loss code <i>UK diagnostic code in GP record and ONS cause of death ICD-9 codes (183) or ICD-10 (C56) codes</i>	EMIS record, starting latest of, study start date, 12 months after practice registration, first recorded loss of appetite, weight loss, abdominal pain, abdominal distension, rectal bleeding, or post-menopausal bleeding	No Townsend score ( <i>n</i> = 98 260). Prior bilateral oophorectomy ( <i>n</i> = 21 099) or ovarian cancer ( <i>n</i> = 2079). Loss of appetite, weight loss, abdominal pain, abdominal distension, rectal bleeding, or post-menopausal bleeding in prior 12 months ( <i>n</i> = 55 824)

... continued

**Table 1 continued. Characteristics of the included studies**

<b>Ovarian</b> (Lim <i>et al</i> , 2012) <sup>39</sup> <i>Ovarian cancer population study, UK</i>	<b>Case-control (1:1.2)</b> Questionnaire, telephone interview, 'GP notes' 16/02/06–28/02/08 15 months pre-diagnosis	Unplanned weight loss, including the appearance of weight loss (questionnaire, telephone interview, or GP code) <i>Independent review of staging, histology, pathology, and case notes by gynaec-oncologist</i>	Cases: 50–79 years, attending preoperative or medical oncology clinic, or admitted for surgery for primary ovarian cancer Controls: attending ovarian screening (post-menopausal) clinic, or surgery for a suspected benign adnexal mass	Differential agreement to take part in questionnaire and telephone studies
<b>Pancreatic</b> (Collins and Altman, 2013) <sup>40</sup> <i>Primary care, UK</i>	<b>Cohort</b> THIN database 01/01/00–30/06/08 2 years following entry	Weight loss Read code <i>Relevant cancer Read code</i>	THIN record starting latest of study start date, registration, first red-flag symptom of pancreatic cancer	Prior pancreatic cancer; GP registered <12 months, invalid dates, <30/≥85 years
<b>Pancreatic</b> (Hippisley-Cox and Coupland, 2012) <sup>49</sup> <i>Primary care, UK</i>	<b>Cohort</b> QResearch database 01/01/00–30/09/10 2 years following entry	Weight loss code <i>UK diagnostic code in GP record and ONS cause of death ICD-9 (157) or ICD-10 (C25) codes</i>	EMS record, starting latest of, study start date, 12 months after practice registration, first recorded dysphagia, loss of appetite, weight loss, abdominal pain, or distension	No Townsend score ( <i>n</i> = 195 305). Prior pancreatic cancer ( <i>n</i> = 257). Dysphagia, loss of appetite, weight loss, abdominal pain, or abdominal distension in prior 12 months ( <i>n</i> = 77 071)
<b>Pancreatic</b> (Holly <i>et al</i> , 2004) <sup>38</sup> <i>San Francisco Bay 'population'</i>	<b>Case-control</b> Two questionnaires by interview 1994–2001 5 years pre-diagnosis	Patient reported weight loss of >6 weeks' duration <i>Surveillance, Epidemiology, and End Results (SEER) database and 18–24 months after diagnosis</i>	Cases: pancreatic cancer identified within 1 month of diagnosis by north California cancer centre rapid case ascertainment Controls: random digit dial age (+/– 5 years) and sex matched (healthcare financing data supplemented recruitment >65 years)	Incomplete second questionnaire, non-English speaker, deceased
<b>Pancreatic and biliary tree</b> (Keane <i>et al</i> , 2014) <sup>44</sup> <i>Primary care, UK</i>	<b>Case-control (1:6)</b> THIN database 01/01/00–31/12/10 2 years pre-diagnosis	GP weight loss code <i>Cancer recorded in THIN</i>	Cases: GP practices with acceptable mortality recording and computer usage Controls: stratified sampling within practice matched by age, sex, year of consultation (controls), or diagnosis (cancer)	<2 years of data prior to the index date
<b>Pancreatic</b> (Stapley <i>et al</i> , 2012) <sup>42</sup> <i>Primary care, UK</i>	<b>Case-control (1:5)</b> CPRD database 01/2000–12/2009 >1 year pre-diagnosis	GP weight loss Read code <i>25 pancreatic cancer Read codes</i>	Cases: pancreatic cancer code Controls: age, sex, and practice matched	Cases without controls ( <i>n</i> = 2), not a pancreatic primary ( <i>n</i> = 10), controls with pancreatic cancer ( <i>n</i> = 64), controls no case ( <i>n</i> = 40), controls with no data in the year before diagnosis the index date ( <i>n</i> = 14 14)
<b>Prostate</b> (Hamilton <i>et al</i> , 2006) <sup>46</sup> <i>Primary care, Exeter, England 1998–2002 (diagnosis)</i>	<b>Case-control (1:5)</b> GP record (written and electronic) 2 years pre-diagnosis	Any mention of weight loss in GP record <i>Cancer registry, histology reports, and specialist case records</i>	Cases: identified by local cancer registry, histology reports, or specialist diagnosis Controls: practice, sex, and age matched, and alive at index diagnosis	Unobtainable records ( <i>n</i> = 82), no entry in the records in 2 years pre-diagnosis ( <i>n</i> = 83), prior prostate cancer ( <i>n</i> = 27), lived outside Exeter at diagnosis ( <i>n</i> = 26), other/unconfirmed ( <i>n</i> = 5)
<b>Renal tract</b> (Collins and Altman, 2013) <sup>39</sup> <i>Primary care, UK</i>	<b>Cohort</b> THIN database 01/01/00–30/06/08 2 years following entry	Weight loss Read code <i>Relevant cancer Read code</i>	THIN record, starting latest of, study start date, registration, first red-flag symptom of renal tract cancer	Prior renal tract cancer; GP registered <12 months, invalid dates, <30/≥85 years
<b>Renal tract</b> (Hippisley-Cox and Coupland, 2012) <sup>47</sup> <i>Primary care, UK</i>	<b>Cohort</b> QResearch database 01/01/00–30/09/10 2 years following entry	Weight loss code <i>UK diagnostic code in GP record and ONS cause of death ICD-9 (188 or 189) or ICD-10 (C64–67) codes</i>	EMS record, starting latest of, study start date, 12 months after practice registration, first recorded loss of haematuria, appetite loss, weight loss, or abdominal pain	No Townsend score ( <i>n</i> = 195 305). Prior renal tract cancer ( <i>n</i> = 4417). Haematuria, loss of appetite, weight loss, or abdominal pain in prior 12 months ( <i>n</i> = 81 332)

CPRD = Clinical Practice Research Database. EMIS = Egton Medical Information Systems. ICD = International Classification of Diseases. nHL = non-Hodgkin's lymphoma. ONS = Office for National Statistics. THIN = The Health Improvement Network.



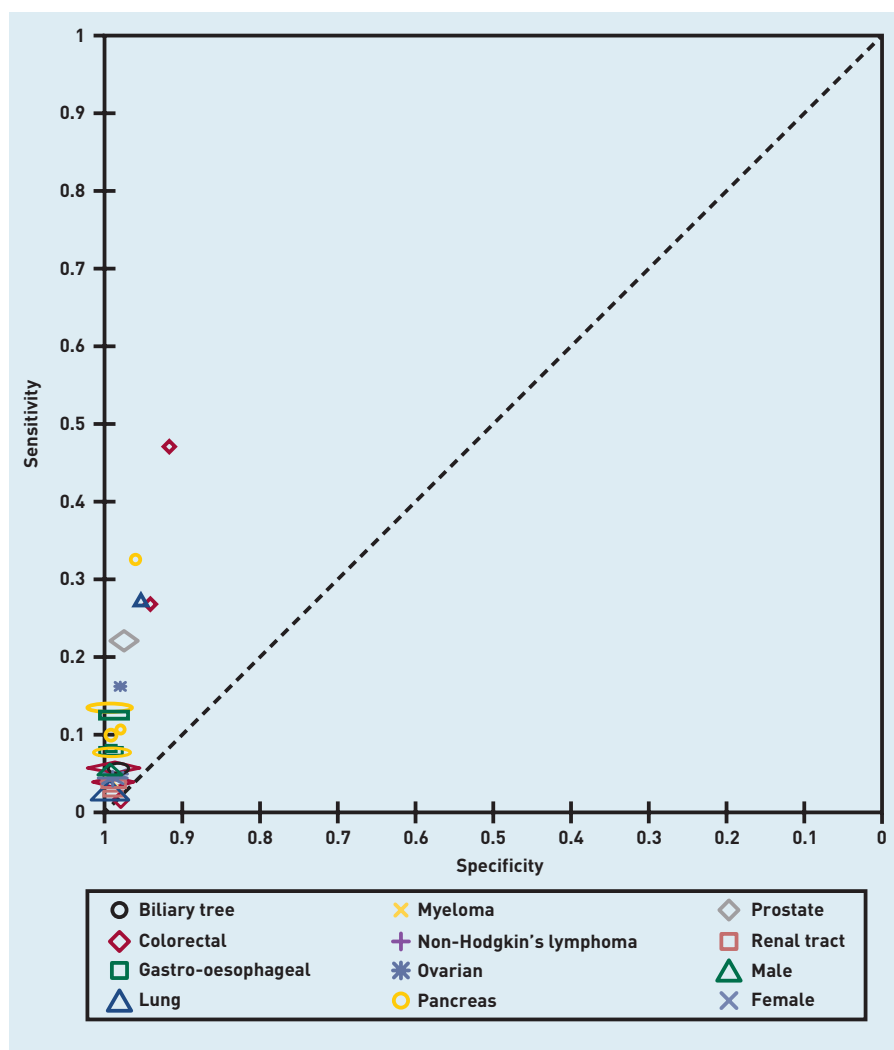


Figure 2. Receiver operating characteristic plot of diagnostic value of unexplained weight loss for different cancers.

for National Statistics (ONS) cause of death codes.<sup>22,23,45–49</sup> Four studies used regional cancer registry and hospital records,<sup>54–56,58</sup> and two only hospital records.<sup>57,59</sup> Three studies, in colorectal, pancreatic, and ovarian cancer, reported cancer stage.<sup>55,58,59</sup>

#### Risk of bias

Figure 1 summarises the QUADAS-2 assessment (further information is also available from the authors on request). In all, 15 studies were classified as high risk of bias in one domain.<sup>8,43–56</sup> Three studies were classified as high risk of bias in more than one domain.<sup>57–59</sup> The main potential sources of bias were:

- applicability concerns when secondary care cases were matched either with primary care controls,<sup>57</sup> using healthcare financing lists,<sup>58</sup> or with outpatients or attendees at screening programmes;<sup>59</sup>
- selection bias in studies using the

case-control design, which exaggerates measures of test accuracy;<sup>8,18,43,44,50–59</sup>

- inclusion bias because eight studies did not describe why all participants were not included in the analysis; and
- recall bias in three studies using a retrospective questionnaire.<sup>57–59</sup>

#### Measures of diagnostic accuracy

Weight loss was a presenting feature of one in 14 cancers overall. This varied by cancer site: one in five prostate; one in 10 colorectal, gastro-oesophageal, and pancreatic; one in 16 biliary tree; one in 25 myeloma, non-Hodgkin's lymphoma, and ovarian; and one in 33 lung and renal tract cancers. Further information on the diagnostic value of weight loss with other clinical and demographic features and forest plots is available from the authors on request. Figure 2 plots each study in ROC space using coloured symbols to denote subgroups of cancer site and sex. Demographic and diagnostic accuracy data for each study are available from the authors on request.

#### Cancer type and stage

Sensitivity ranged from 2% to 47% across cancer types, and specificity from 92% to 99% (further information available from the authors on request). Only colorectal and pancreatic cancer had sufficient studies for separate analyses (further information available from the authors on request). The pooled sensitivity for colorectal cancer was 14% [95% CI = 6 to 30%] and pooled specificity 97% [95% CI = 94 to 99%]. The pooled sensitivity for pancreatic cancer was 13% [95% CI = 8 to 20%] and pooled specificity 99% [95% CI = 98 to 99%]. Removing the studies at high risk of bias in more than one QUADAS-2 domain did not significantly alter the pooled results.

In all, 20 studies adjusted for other covariates. These included potential confounders, such as age, but mainly symptoms known to be predictors of each cancer, for example, rectal bleeding in colorectal cancer (further information available from the authors on request). Adjustment left the diagnostic odds ratios (DORs) unchanged in all but three case-control studies, in which they were significantly reduced.<sup>8,43,55</sup> Hazard ratios (HRs) ranged from 1.6 for ovarian cancer to 12.5 for pancreatic cancer across the cohort studies. In all cases, adjustment for other symptoms of cancer did not abolish the association between weight loss and a diagnosis of cancer. No study reported the diagnostic value of weight loss according to the cancer stage.

**Table 2. Cumulative positive predictive values (PPVs) across all cancer sites<sup>a</sup>**

		Age group, years															
Sex		≥40	≥50	≥60	≥70	≥80	≥90	40–49	50–59	60–69	70–79	80–89	≥90	40–59	60–79	≥80	
Lowest LRs from SR	Female	1.6	2.2	3.0	3.9	4.4	4.0	0.2	0.8	1.9	3.5	4.6	4.0	0.7	2.5	4.4	
	Combined	5.2	7.2	10.0	12.7	13.5	13.9	0.4	2.2	7.2	12.2	13.5	13.9	1.2	9.2	13.5	
	Male	5.4	7.5	10.5	13.7	15.0	16.1	0.4	2.2	7.5	13.0	14.9	16.1	1.3	9.6	15.0	
Highest LRs from SR	Female	3.6	5.0	6.7	8.7	9.9	8.4	0.5	1.7	4.2	7.8	10.5	8.4	1.3	5.7	9.9	
	Combined	7.3	10.1	14.0	18.1	19.6	19.1	0.7	3.2	9.8	17.0	19.9	19.1	1.8	12.6	19.6	
	Male	7.3	10.1	14.2	18.8	21.3	22.3	0.6	3.1	9.8	17.5	21.1	22.3	1.8	12.8	21.3	

<sup>a</sup>Shading: PPV 1–1.9% (yellow) and 2–2.9% (amber); NICE suggests GPs consider primary care testing. ≥3% (red): NICE recommends referral for specialist review or definitive investigation if the investigation is available to the GP. The combined analysis includes: colorectal, gastro-oesophageal, lung, multiple myeloma, non-Hodgkin's lymphoma, ovarian, pancreatic, prostate, and renal tract cancer. Prostate cancer is removed from the female analysis, and ovarian from the male. No shading: <1%, no testing needed. LR = likelihood ratio. NICE = National Institute for Health and Care Excellence. SR = systematic review.

### Sex and age

Five studies examined diagnostic accuracy by sex. In one study, the sensitivity and specificity for any cancer in males was 3% [95% CI = 3 to 4%] and 99% [95% CI = 99 to 99%], respectively.<sup>22</sup> In another, the sensitivity and specificity for any cancer in females was 5% [95% CI = 5 to 6%] and 99% [95% CI = 99 to 99%], respectively.<sup>23</sup> However, there was no significant difference between HRs for males and females after adjustment for age, BMI, and other clinical features (further information available from the authors on request). The authors were not able to calculate summary estimates for age groups because studies reporting the same age range and cancer type were too few.

### Positive predictive values

PPVs for weight loss ranged from 0.0% for biliary tract cancer to 3.3% for prostate cancer (further information available from the authors on request). One study reported higher PPVs for colorectal cancer in people aged ≥70 years old (1.7%) compared with <70 years old (0.4%).<sup>57</sup> Six studies reported the PPV for weight loss paired with another clinical feature ranging from 0.1% for chest pain for myeloma to 31.6% for jaundice in pancreatic cancer (further information available from the authors on request). Weight loss with another clinical feature yielded a PPV greater than weight loss alone, the paired clinical feature alone, or by adding together the individual PPVs.

The authors calculated cumulative PPVs by age group (Table 2). The more conservative estimates using the lowest LRs had a cumulative PPV of 3.0% in females ≥60 years, and 10.5% in males ≥60 years, reaching the threshold for investigation (3%) used in the 2015 NICE guidelines.<sup>7</sup> The more liberal estimate using

the highest reported LRs for each cancer site resulted in females (6.7%) and males (14.2%) ≥60 years. The excess in males was largely the contribution of prostate cancer due to the relatively high LR and incidence. After removing prostate cancer from the analysis, the male PPV ≥60 years (7.8%) still reached the threshold for investigation (further information available from the authors on request).

## DISCUSSION

### Summary

Patients presenting to primary care with weight loss are at higher risk of having cancer than patients without recorded weight loss. The 10 cancers were prostate, colorectal, lung, gastro-oesophageal, pancreatic, non-Hodgkin's lymphoma, ovarian, multiple myeloma, renal tract, and biliary tree. Across these cancer sites, sensitivity ranged from 2% to 47%, and specificity from 92% to 99%. The PPV of cancer in male and female patients with weight loss aged 60–79 years and ≥80 years exceeds the risk threshold for urgent investigation set out in NICE guidelines.<sup>7</sup> The risk with weight loss increases when it presents alongside another clinical feature suggesting an individual cancer site, and with increasing age.

### Strengths and limitations

To the authors' knowledge, this is the only review of weight loss as a feature of cancer in non-specialist settings. The authors used a broad search strategy to interrogate large electronic databases. They have calculated cumulative PPVs for the first time for weight loss across 10 cancer sites by using high-quality national cancer incidence data.

However, there are important limitations. Foremost, the authors can provide no clarification about how much weight loss



is required to make cancer a possibility. Most studies defined weight loss using coding, but no studies reported the code lists used by the authors to define weight loss, so the authors are unable to report with confidence whether weight loss was unexpected or expected (following diet or exercise). It seems likely that a weight loss code represents unexpected weight loss, as a person who achieves weight loss intentionally is unlikely to report this to their GP. Similarly, a GP is unlikely to think that intentional weight loss is diagnostically relevant enough to code: GP coding has been shown to be prone to recording bias, which occurs when GPs preferentially code the 'alarm' features they consider to support their working diagnosis.<sup>60,61</sup> For example, the inclusion of free-text data in one study reduced the PPV of jaundice for pancreatic cancer from 12.8% [95% CI = 7.3 to 21.6] to 6.3% [95% CI = 4.5 to 8.7], and visible haematuria for bladder cancer from 4.0% [95% CI = 3.5 to 4.6] to 2.9% [95% CI = 2.6 to 3.2].<sup>60</sup> GPs commonly use free text to amplify the clinical narrative after coding 'alarm' symptoms, meaning that for some patients the entire details of a clinical feature remain hidden, invisible to analyses solely using symptom codes.<sup>60</sup> Low-risk-but-not-no-risk vague symptoms of cancer, such as abdominal pain or cough, have high prevalence in primary care, are coded less, and patients with them are less likely to prompt clinicians to refer patients with these symptoms for urgent investigation.<sup>60,62</sup> However, the authors have found no reports investigating the influence of recording bias on the prevalence of weight loss. From the colorectal cancer studies in this review, the prevalence of weight loss was highest when defined as 'any mention of weight loss' in the GP record (8.9%), intermediate when objective weight measurements were used (4.8%), and lowest when defined by Read coding alone (1.1%).<sup>8,48,55</sup> Only one study used objective measurements to define weight loss, leaving studies at risk of misclassification due to missing data on weight.<sup>8,63</sup> Furthermore, there were insufficient data to describe the grade or stage of cancer diagnosed in patients with weight loss and how this compares with cancer in patients presenting without weight loss.

Finally, heterogeneity in diagnostic accuracy relates to heterogeneity in study quality: sensitivity was higher in studies at risk of recall bias; PPVs were higher in case-control studies compared with cohort studies reporting the same tumour site; and the calculation and interpretation of pooled

estimates were limited by this heterogeneity. One study, for example, showed that weight loss was reported more frequently when data were collected by questionnaire or telephone interview compared with using only a coded entry in the electronic record.<sup>59</sup> Collecting data on unexpected weight loss by questionnaire increased the sensitivity and decreased the specificity compared with using only coded GP data. Ascertainment bias is also likely, despite most studies using an index test and reference standard within a 2-year period. Two studies have reported that within a 2-year window 'serendipitous' asymptomatic cancers are diagnosed when symptoms (not caused by cancer) are investigated.<sup>9,64</sup>

### Comparison with existing literature

Others have reviewed the symptoms of single cancer sites.<sup>65-70</sup> The authors have included all of their studies, except where concerns existed over study population<sup>65,66,69</sup> or inadequate reporting of 2 × 2 data,<sup>70</sup> and they included an additional study post-dating their searches.<sup>43</sup> The authors identified no previous review of symptoms presenting to primary care in association with myeloma, non-Hodgkin's lymphoma, ovarian, or prostate cancer.

### Implications for research and practice

This review suggests that patients aged ≥60 years presenting to primary care settings with weight loss that prompts a clinical record entry warrant rapid investigation for possible cancer, potentially across multiple possible sites. The findings will therefore be of interest to policymakers developing multidisciplinary symptom-based referral pathways to complement existing site-specific routes.<sup>71,72</sup> As a symptom of cancer, recorded weight loss is the second highest risk after rectal bleeding in colorectal cancer,<sup>65</sup> haemoptysis in lung,<sup>54</sup> rib pain in myeloma,<sup>50</sup> jaundice in pancreatic,<sup>69</sup> and haematuria in renal tract;<sup>39</sup> and third highest in non-Hodgkin's lymphoma,<sup>51</sup> and gastro-oesophageal.<sup>41,45,53</sup> Additional symptoms in combination with weight loss strengthen these associations.

The low sensitivity for weight loss reflects the clinical reality for patients that some cancers lead to weight loss while many more do not. Understanding the site distribution of cancer in these patients may inform the sequencing of investigation in this patient group. The low sensitivity also demonstrates that the absence of weight loss cannot be used to rule out cancer. The extremely high specificity for weight loss has a number of potential and overlapping explanations:

- patients may only be attending primary care once weight loss is marked;
- clinicians may only code the symptom after conducting simple investigations to rule out other causes of weight loss; and
- the number of false positives is artificially low in comparison with the extremely high number of true negatives that represent the low prevalence of cancer in primary care populations.

It is probable that when GPs choose to code weight loss this decision reflects that the clinician believes the symptom to be important. This belief may be based on additional symptoms or physical findings that may themselves represent an increased risk for cancer. Thus, some of the risk of cancer associated with coded weight loss represents a broader assessment by the clinician. Arguably, this is helpful. These results clearly show that coded weight loss warrants investigative action. If clinical systems were designed to prompt the clinician to consider cancer whenever they enter a weight loss code, this prompting may lead to an expedited diagnosis.

To overcome the limitations outlined above, the ideal methodology to ascertain the diagnostic value of weight loss (or any other clinical feature relevant to multiple cancer types) in primary care would be a prospective cohort study. The prospective design overcomes biases in the studies in this review. As the prevalence of cancer (and weight loss) is low in primary care, a study powered to investigate all relevant

symptoms to all cancers would have to be prohibitively large. The CANcer Diagnosis Decision Rules (CANDID) study is underway in UK primary care to investigate the clinical features of colorectal and lung cancer including weight loss, but the prevalence of weight loss in this cohort is still likely to be low.<sup>73</sup>

There is more to be gained from historical cohort studies using routinely collected electronic record datasets<sup>22,23,74–76</sup> that:

- ascertain a prevalence for weight loss based on coded, non-coded, and numeric data;
- confirm cancer utilising the clinical record (coded and non-coded), and linkages to national cancer registries;
- report 2 × 2 data for weight loss in combination with other clinical features; and
- report diagnostic intervals and longer-term outcomes.

Research is required to understand the drivers of weight recording in primary care, and the extent of weight loss that prompts a GP to code it. Without these studies, we will remain unable to answer fundamental questions of 'How much weight loss should I worry about?'

In the meantime, once a physician considers that weight loss in a patient ≥60 years warrants a clinical entry, these data indicate that investigation to identify a cancer is then necessary.

### Funding

Paul Aveyard was funded by the NIHR Oxford Biomedical Research Centre. This manuscript presents independent research funded as part of the National Institute for Health Research (NIHR) Doctoral Research Fellowship number (DRF-2015-08-18). FDR Hobbs is supported by the NIHR Oxford Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

### Provenance

Freely submitted; externally peer reviewed.

### Competing interests

The authors have declared no competing interests.

### Acknowledgements

The authors wish to thank Emeritus Professor David Mant for his expertise and invaluable guidance when developing the fellowship proposal underpinning this research, and Nia Roberts for her expertise in developing the search strategy.

### Discuss this article

Contribute and read comments about this article: [bjgp.org/letters](http://bjgp.org/letters)

## REFERENCES

- Wong CJ. Involuntary weight loss. *Med Clin North Am* 2014; **98**(3): 625–643.
- Hamilton W. Five misconceptions in cancer diagnosis. *Br J Gen Pract* 2009; DOI: <https://doi.org/10.3399/bjgp09X420860>.
- McMinn J, Steel C, Bowman A. Investigation and management of unintentional weight loss in older adults. *BMJ* 2011; **342**: d1732.
- Rolland Y, Kim MJ, Gammack JK, *et al*. Office management of weight loss in older persons. *Am J Med* 2006; **119**(12): 1019–1026.
- Alibhai SM, Greenwood C, Payette H. An approach to the management of unintentional weight loss in elderly people. *CMAJ* 2005; **172**(6): 773–780.
- Gaddey HL, Holder K. Unintentional weight loss in older adults. *Am Fam Physician* 2014; **89**(9): 718–722.
- National Institute for Health and Care Excellence. *Suspected cancer: recognition and referral*. NG12. London: NICE, 2015. <https://www.nice.org.uk/guidance/ng12> [accessed 29 Mar 2018].
- Hamilton W, Lancashire R, Sharp D, *et al*. The risk of colorectal cancer with symptoms at different ages and between the sexes: a case-control study. *BMC Med* 2009; **7**: 17.
- Ades AE, Biswas M, Welton NJ, Hamilton W. Symptom lead time distribution in lung cancer: natural history and prospects for early diagnosis. *Int J Epidemiol* 2014; **43**(6): 1865–1873.
- Olson SH, Xu Y, Herzog K, *et al*. Weight loss, diabetes, fatigue, and depression preceding pancreatic cancer. *Pancreas* 2016; **45**(7): 986–991.
- Ewing M, Naredi P, Zhang C, Månsson J. Identification of patients with non-metastatic colorectal cancer in primary care: a case-control study. *Br J Gen Pract* 2016; DOI: <https://doi.org/10.3399/bjgp16X687985>.
- Stapley S, Peters TJ, Sharp D, Hamilton W. The mortality of colorectal cancer in relation to the initial symptom at presentation to primary care and to the duration of symptoms: a cohort study using medical records. *Br J Cancer* 2006; **95**(10): 1321–1325.
- Pannala R, Leibson CL, Rabe KG, *et al*. Temporal association of changes in fasting blood glucose and body mass index with diagnosis of pancreatic cancer. *Am J Gastroenterol* 2009; **104**(9): 2318–2325.
- Wallace JL, Schwartz RS. Epidemiology of weight loss in humans with special reference to wasting in the elderly. *Int J Cardiol* 2002; **85**(1): 15–21.
- 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.
- Gobbi PG, Bergonzi M, Comelli M, *et al*. The prognostic role of time to diagnosis and presenting symptoms in patients with pancreatic cancer. *Cancer Epidemiol* 2013; **37**(2): 186–190.
- Nicholson B, O'Sullivan J, Aveyard P, *et al*. Unexpected weight loss as a predictor of cancer in primary care. PROSPERO 2016 CRD42016038963. [http://www.crd.york.ac.uk/PROSPERO/display\\_record.asp?ID=CRD42016038963](http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016038963) [accessed 21 Mar 2018].
- Macakill P, Gatsonis C, Deeks J, *et al*. Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C, eds. *Cochrane handbook for systematic reviews of diagnostic test accuracy*. <http://methods.cochrane.org/sdt/handbook-dta-reviews> [accessed 15 Mar 2018].
- Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; **339**: b2700.
- Boggon R, van Staa TP, Chapman M, *et al*. Cancer recording and mortality in the General Practice Research Database and linked cancer registries. *Pharmacoepidemiol and Drug Saf* 2013; **22**(2): 168–175.
- Dregan A, Moller H, Murray-Thomas T, Gulliford MC. Validity of cancer diagnosis in a primary care database compared with linked cancer registrations in England. Population-based cohort study. *Cancer Epidemiol* 2012; **36**(5): 425–429.
- Hippisley-Cox J, Coupland C. Symptoms and risk factors to identify men with suspected cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2013; DOI: <https://doi.org/10.3399/bjgp13X660724>.
- Hippisley-Cox J, Coupland C. Symptoms and risk factors to identify women with suspected cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2013; DOI: <https://doi.org/10.3399/bjgp13X660733>.
- Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. *Br J Gen Pract* 2010; DOI: <https://doi.org/10.3399/bjgp10X483562>.
- Whiting PF, Rutjes AW, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Int Med* 2011; **155**(8): 529–536.
- Knottnerus J, Buntinx F. *The evidence base of clinical diagnosis: theory and methods of diagnostic research*. 2nd edn. Oxford: Blackwell Publishing, 2009.
- Reitsma JB, Glas AS, Rutjes AW, *et al*. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; **58**(10): 982–990.
- Harbord RM, Whiting P. metandi: meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata Journal* 2009; **9**(2): 211–229.
- Takwoingi Y, Guo B, Riley RD, *et al*. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Stat Methods Med Res* 2015; DOI: [10.1177/0962280215592269](https://doi.org/10.1177/0962280215592269).
- Harbord RM, Deeks JJ, Egger M, *et al*. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostat* 2007; **8**(2): 239–251.
- Riley RD, Abrams KR, Sutton AJ, *et al*. Bivariate random-effects meta-analysis and the estimation of between-study correlation. *BMC Med Res Methodol* 2007; **7**: 3.
- Riley RD, Ahmed I, Debray TPA, *et al*. Summarising and validating test accuracy results across multiple studies for use in clinical practice. *Statist Med* 2015; **34**(13): 2081–2103.
- Whiting PF, Smidt N, Sterne JC, *et al*. Systematic review: accuracy of anti-citrullinated peptide antibodies for diagnosing rheumatoid arthritis. *Ann Int Med* 2010; **152**(7): 456–464.
- Cancer Research UK. Cancer statistics for the UK. 2016. <http://www.cancerresearchuk.org/health-professional/cancer-statistics> [accessed 15 Mar 2018].
- Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005; **58**(9): 882–893.
- Leeflang MM, Deeks JJ, Gatsonis C, *et al*. Systematic reviews of diagnostic test accuracy. *Ann Int Med* 2008; **149**(12): 889–897.
- Song F, Khan KS, Dinnes J, Sutton AJ. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 2002; **31**(1): 88–95.
- Collins GS, Altman DG. Identifying patients with undetected colorectal cancer: an independent validation of QCancer® (Colorectal). *Br J Cancer* 2012; **107**(2): 260–265.
- Collins GS, Altman DG. Identifying patients with undetected renal tract cancer in primary care: an independent and external validation of QCancer® (renal) prediction model. *Cancer Epidemiol* 2013; **37**(2): 115–120.
- Collins GS, Altman DG. Identifying patients with undetected pancreatic cancer in primary care: an independent and external validation of QCancer® (pancreas). *Br J Gen Pract* 2013; DOI: <https://doi.org/10.3399/bjgp13X671623>.
- Collins GS, Altman DG. Identifying patients with undetected gastro-oesophageal cancer in primary care: external validation of QCancer® (gastro-oesophageal). *Eur J Cancer Care* 2013; **49**(5): 1040–1048.
- Collins GS, Altman DG. Identifying women with undetected ovarian cancer: independent and external validation of QCancer® (ovarian) prediction model. *Eur J Cancer Care* 2013; **22**(4): 423–429.
- Iyen-Omofoman B, Tata LJ, Baldwin DR, *et al*. Using sociodemographic and early clinical features in general practice to identify people with lung cancer earlier. *Thorax* 2013; **68**(5): 451–459.
- Keane MG, Horsfall L, Rait G, Pereira SP. A case-control study comparing the incidence of early symptoms in pancreatic and biliary tract cancer. *BMJ Open* 2014; **4**(11): e005720.
- Hippisley-Cox J, Coupland C. Identifying patients with suspected gastro-oesophageal cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2011; DOI: <https://doi.org/10.3399/bjgp11X606609>.
- Hippisley-Cox J, Coupland C. Identifying women with suspected ovarian cancer in primary care: derivation and validation of algorithm. *BMJ* 2012; **344**: d8009.
- Hippisley-Cox J, Coupland C. Identifying patients with suspected renal tract cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2012; DOI: <https://doi.org/10.3399/bjgp12X636074>.
- Hippisley-Cox J, Coupland C. Identifying patients with suspected colorectal cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2012; DOI: <https://doi.org/10.3399/bjgp12X616346>.
- Hippisley-Cox J, Coupland C. Identifying patients with suspected pancreatic cancer in primary care: derivation and validation of an algorithm. *Br J Gen*

- Pract 2012; DOI: <https://doi.org/10.3399/bjgp12X616355>.
50. Shephard EA, Neal RD, Rose P, *et al*. Quantifying the risk of multiple myeloma from symptoms reported in primary care patients: a large case-control study using electronic records. *Br J Gen Pract* 2015; DOI: <https://doi.org/10.3399/bjgp15X683545>.
  51. Shephard EA, Neal RD, Rose PW, *et al*. Quantifying the risk of non-Hodgkin's lymphoma in symptomatic primary care patients aged >40 years: a large case-control study using electronic records. *Br J Gen Pract* 2015; DOI: <https://doi.org/10.3399/bjgp15X684805>.
  52. Stapley S, Peters TJ, Neal RD, *et al*. The risk of pancreatic cancer in symptomatic patients in primary care: a large case-control study using electronic records. *Br J Cancer* 2012; **106**(12): 1940–1944.
  53. Stapley S, Peters TJ, Neal RD, *et al*. The risk of oesophagogastric cancer in symptomatic patients in primary care: a large case-control study using electronic records. *Br J Cancer* 2013; **108**(1): 25–31.
  54. Hamilton W, Peters TJ, Round A, Sharp D. What are the clinical features of lung cancer before the diagnosis is made? A population based case-control study. *Thorax* 2005; **60**(12): 1059–1065.
  55. Hamilton W, Round A, Sharp D, Peters TJ. Clinical features of colorectal cancer before diagnosis: a population-based case-control study. *Br J Cancer* 2005; **93**(4): 399–405.
  56. Hamilton W, Sharp DJ, Peters TJ, Round AP. Clinical features of prostate cancer before diagnosis: a population-based, case-control study. *Br J Gen Pract* 2006; **56**(531): 756–762.
  57. Curless R, French J, Williams GV, James O. Comparison of gastrointestinal symptoms in colorectal carcinoma patients and community controls with respect to age. *Gut* 1994; **35**(9): 1267–1270.
  58. Holly EA, Chaliha I, Bracci PM, Gautam M. Signs and symptoms of pancreatic cancer: a population-based case-control study in the San Francisco Bay area. *Clin Gastroenterol Hepatol* 2004; **2**(6): 510–517.
  59. Lim AWW, Mesher D, Gentry-Maharaj A, *et al*. Predictive value of symptoms for ovarian cancer: comparison of symptoms reported by questionnaire, interview, and general practitioner notes. *J Natl Cancer Inst* 2012; **104**(2): 114–124.
  60. Price SJ, Stapley SA, Shephard E, *et al*. Is omission of free text records a possible source of data loss and bias in Clinical Practice Research Datalink studies? A case-control study. *BMJ Open* 2016; **6**(5): e011664.
  61. Shephard E, Stapley S, Hamilton W. The use of electronic databases in primary care research. *Fam Pract* 2011; **28**(4): 352–354.
  62. Jensen H, Topping 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**: 636.
  63. Bhaskaran K, Forbes HJ, Douglas I, *et al*. Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD). *BMJ Open* 2013; **3**(9): e003389.
  64. Biswas M, Ades AE, Hamilton W. Symptom lead times in lung and colorectal cancers: what are the benefits of symptom-based approaches to early diagnosis? *Br J Cancer* 2015; **112**(2): 271–277.
  65. Astin M, Griffin T, Neal RD, *et al*. The diagnostic value of symptoms for colorectal cancer in primary care: a systematic review. *Br J Gen Pract* 2011; DOI: <https://doi.org/10.3399/bjgp11X572427>.
  66. Astin MP, Martins T, Welton N, *et al*. Diagnostic value of symptoms of oesophagogastric cancers in primary care: a systematic review and meta-analysis. *Br J Gen Pract* 2015; DOI: <https://doi.org/10.3399/bjgp15X686941>.
  67. Jellema P, van der Windt DA, Bruinvels DJ, *et al*. Value of symptoms and additional diagnostic tests for colorectal cancer in primary care: systematic review and meta-analysis. *BMJ* 2010; **340**: c1269.
  68. Schmidt-Hansen M, Berendse S, Hamilton W. The association between symptoms and bladder or renal tract cancer in primary care: a systematic review. *Br J Gen Pract* 2015; DOI: <https://doi.org/10.3399/bjgp15X687421>.
  69. Schmidt-Hansen M, Berendse S, Hamilton W. Symptoms of pancreatic cancer in primary care: a systematic review. *Pancreas* 2016; **45**(6): 814–818.
  70. Shim J, Brindle L, Simon M, George S. A systematic review of symptomatic diagnosis of lung cancer. *Fam Pract* 2014; **31**(2): 137–148.
  71. Vedsted P, Olesen F. A differentiated approach to referrals from general practice to support early cancer diagnosis — the Danish three-legged strategy. *Br J Cancer* 2015; **112**(Suppl 1): S65–S69.
  72. 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**: 421.
  73. Mansell G, Shapley M, van der Windt D, *et al*. Critical items for assessing risk of lung and colorectal cancer in primary care: a Delphi study. *Br J Gen Pract* 2014; DOI: <https://doi.org/10.3399/bjgp14X681001>.
  74. Hamilton F, Carroll R, Hamilton W, Salisbury C. The risk of cancer in primary care patients with hypercalcaemia: a cohort study using electronic records. *Br J Cancer* 2014; **111**(7): 1410–1412.
  75. Merriel SWD, Carroll R, Hamilton F, Hamilton W. Association between unexplained hypoalbuminaemia and new cancer diagnoses in UK primary care patients. *Fam Pract* 2016; **33**(5): 449–452.
  76. Taylor A, Stapley S, Hamilton W. Jaundice in primary care: a cohort study of adults aged >45 years using electronic medical records. *Fam Pract* 2012; **29**(4): 416–420.