Cancer incidence and competing mortality risk following 15 presenting symptoms in primary care

a population-based cohort study using electronic healthcare records

Matthew Barclay
Senior Research Fellow in Statistics

Epidemiology of Cancer Healthcare & Outcomes
Research Department of Behavioural Science and Health
Institute of Epidemiology and Health Care
University College London

Co-authors:
Cristina Renzi, Hannah Harrison, Ana Torralbo, Becky White, Samantha Ip, Juliet Usher-Smith, Nora Pashayan, Spiros Denaxas, Angela Wood, Antonis Antoniou, Georgios Lyratzopoulos

Conflict of interest statement

- I receive personal fees from GRAIL for membership of an independent data monitoring committee

Most cancers diagnosed symptomatically

Study Aim: Update and extend the evidence on cancer risk in symptomatic patients

Two cohorts drawn from CPRD GOLD

Reference group
(Random sample of CPRD)

~786,000 patients

Index date = randomly chosen date during study period

Symptomatic cohort

~836,000 patients

Index date = date of earliest recorded symptom

Included only if index date was during study period

Excluded if in the reference group

(study period = 1 Jan 2007 to 31 Dec 2017)

Outcomes considered

Follow-up censored at 18 months after index.

Men	Women
Death with no cancer diagnosis	Death with no cancer diagnosis
	Breast cancer
	Gynaecological cancer
Lung cancer	Lung cancer
Upper GI cancer (incl. pancreas)	Upper GI cancer (incl. pancreas)
Lower GI cancer	Lower GI cancer
Urological cancer	Urological cancer
Prostate cancer	
Haematological cancer	Haematological cancer
Other cancers (catch-all, incl. male breast)	Other cancers (catch-all)

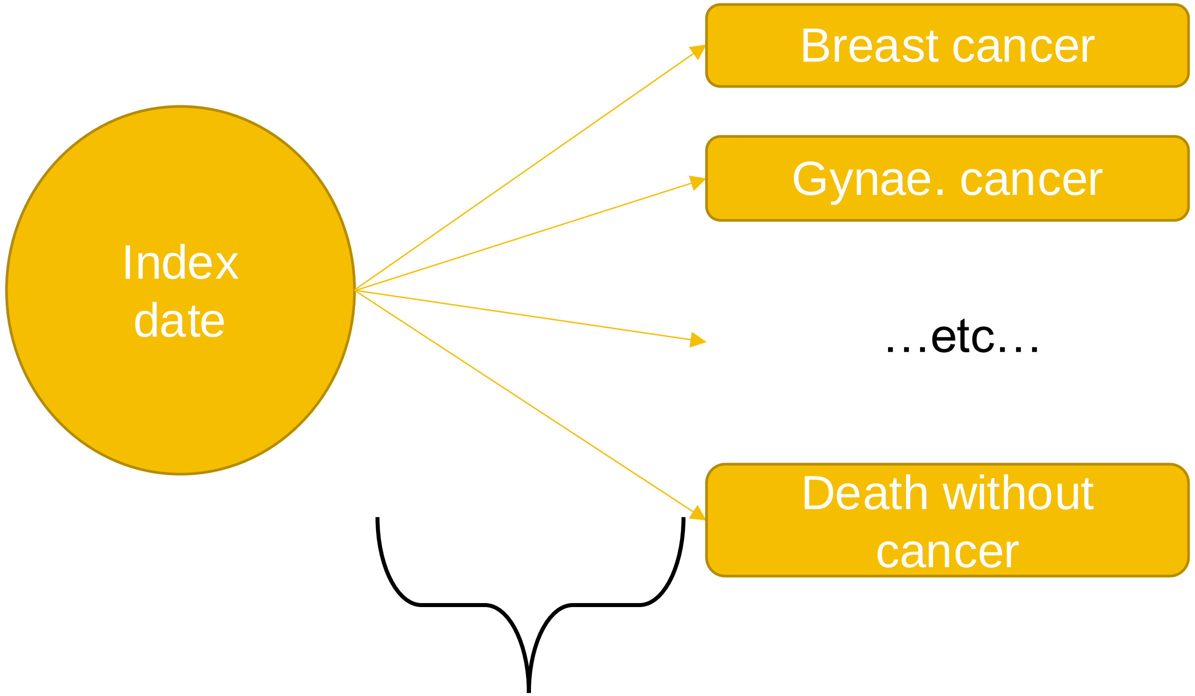
Exposures considered

All patients in the symptomatic cohort had *at least* one symptom, and around 1% had multiple symptoms (same day)

All patients in the reference group were coded as having no symptoms

Men	Women
Age (continuous, cubic spline)	Age (continuous, cubic spline)
Smoking (ever vs never)	Smoking (ever vs never)
Symptom presence Abdominal pain Abdominal bloating Rectal bleeding Dyspepsia Dysphagia Jaundice Dyspnoea Haemoptysis Haematuria Fatigue Night sweats Weight loss Breast lump	*Symptom presence* Abdominal pain Abdominal bloating Rectal bleeding Dyspepsia Dysphagia Jaundice Dyspnoea Haemoptysis Haematuria Fatigue Night sweats Weight loss Breast lump Post-menopausal bleeding
Symptom-time varying effect Change in log-hazard ratio for each symptom associated with log(months since index)	*Symptom-time varying effect* Change in log-hazard ratio for each symptom associated with log(months since index)

Statistical modelling



Model coefficients give the transition intensities from index to various end-states

8 (men) or 9 (women) outcome-specific semi-parametric time-to-event models

Hazard ratios and naïve estimators based on these models give associations with risk in world where there is only one of the endpoints

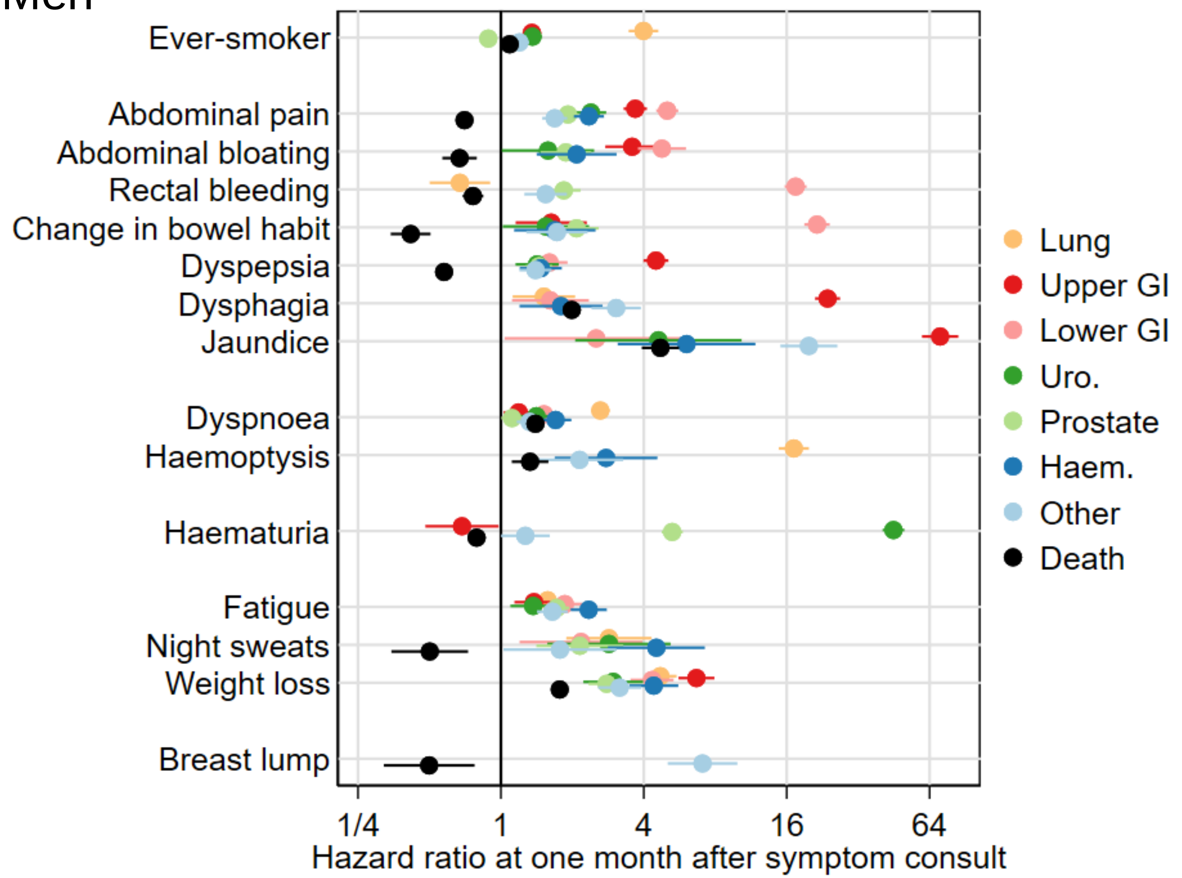
Various methods exist to combine results of several outcome-specific models into overall risk estimates

We used the “latent failure time” approach – in essence, one simulates a failure time from every cause-specific model and takes the first failure time as being the ‘true’ one

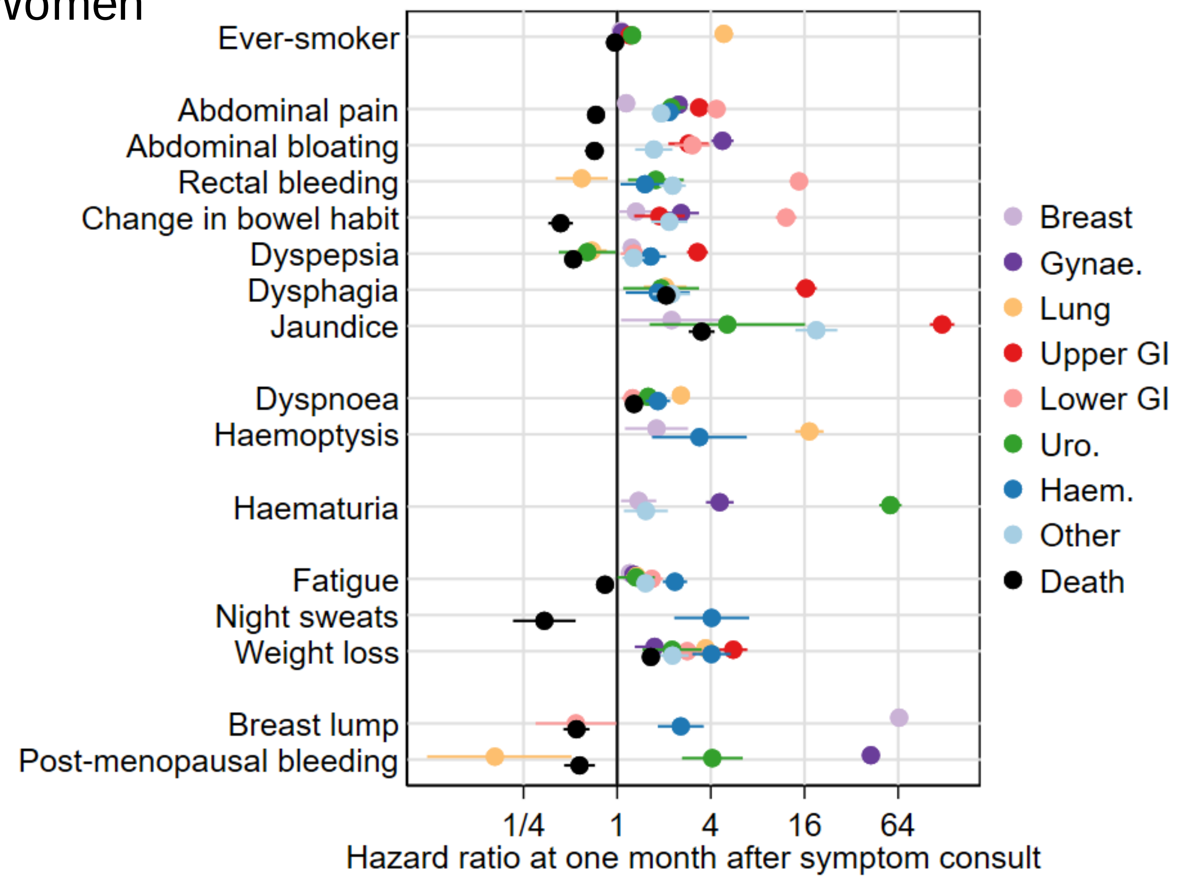
Results – hazard ratios at index

1. Organ-specific symptoms matter for the obvious cancer
2. Non-specific symptoms have non-specific associations
3. Most (but not all) symptoms associated with lower hazard of death
4. Often, cancer-specific hazard ratios for men and women very similar
5. (Not shown) significant decrease in strength of associations as time since consultation elapses (at least where initial HR is very large)

Men



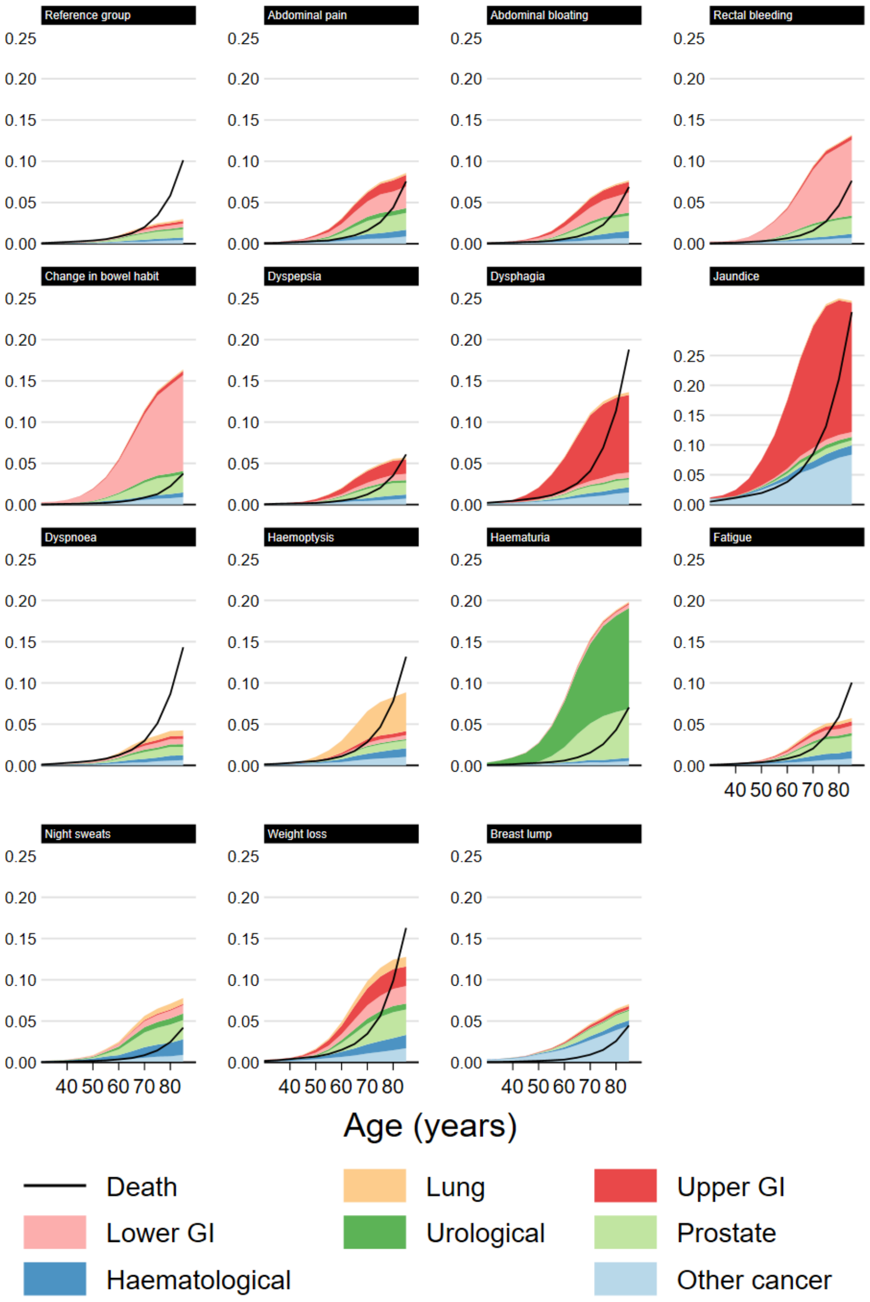
Women



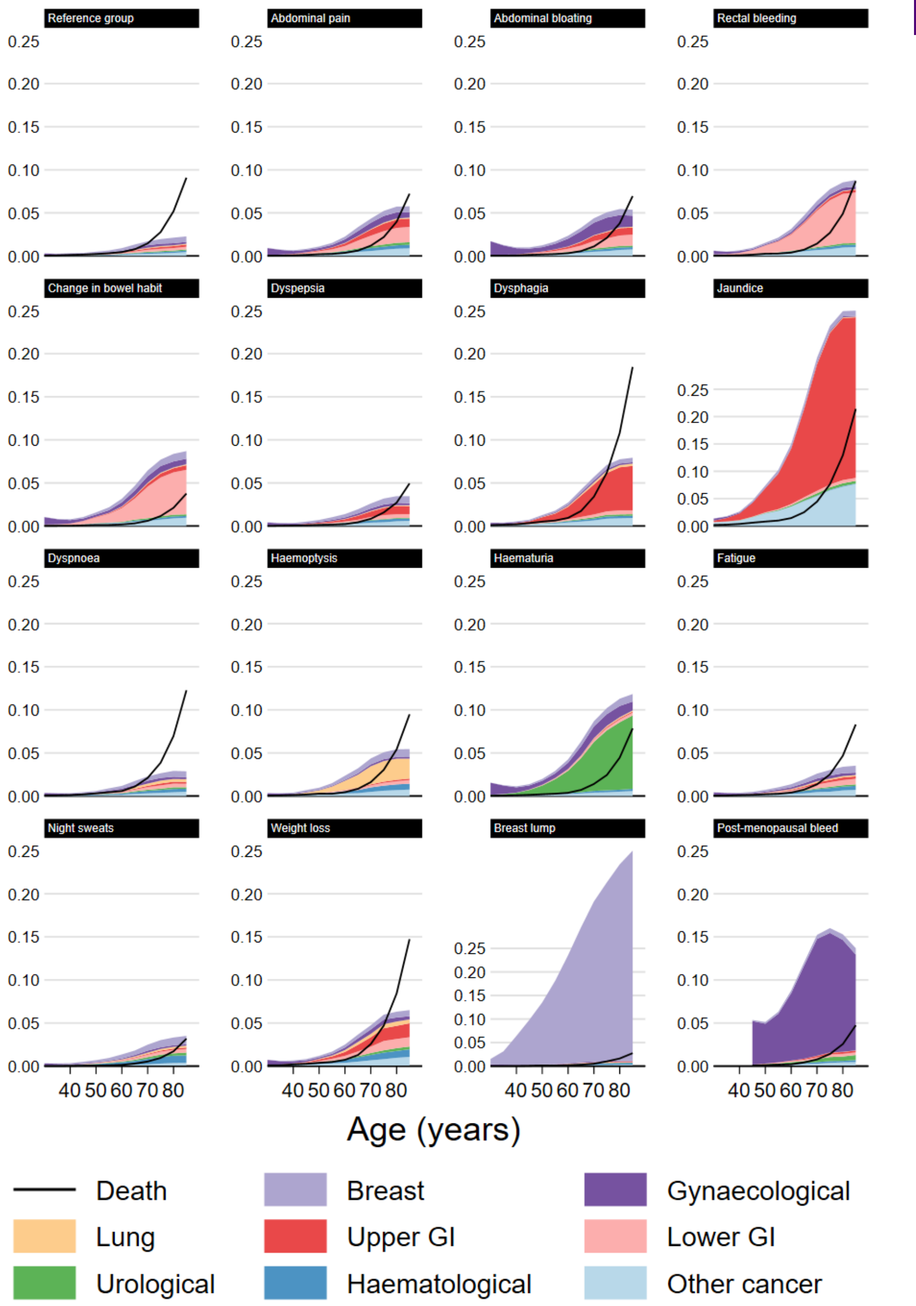
Results – risk of cancer diagnosis within 12 months

Sex-specific differences in baseline hazard of different cancers mean that similar HRs translate to radically different cancer risk

Men, 12 months after symptom presentation, non-smokers

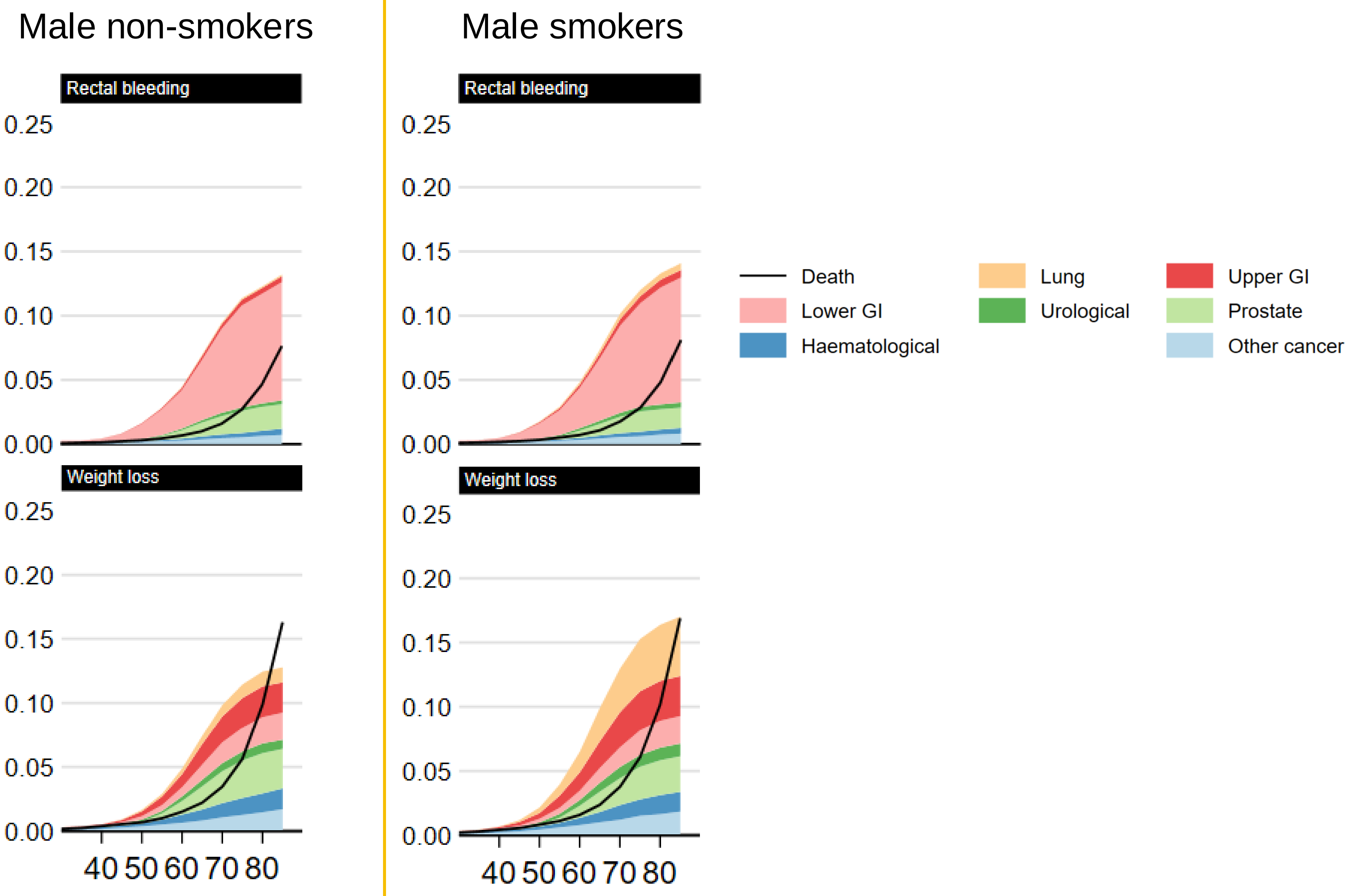


Women, 12 months after symptom presentation, non-smokers



Results – risk of cancer diagnosis within 12 months

Smoking highly informative – but mainly when patients have non-specific or respiratory symptoms

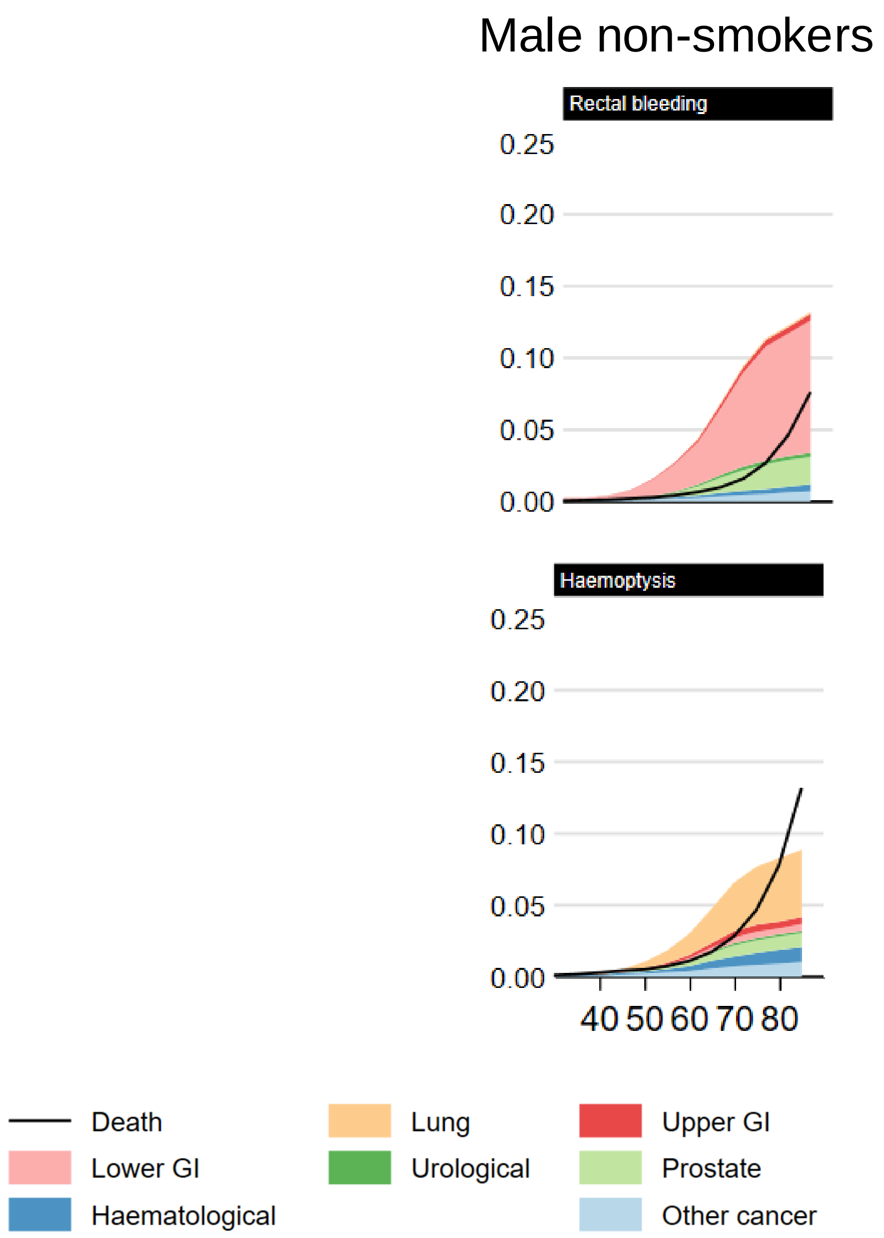


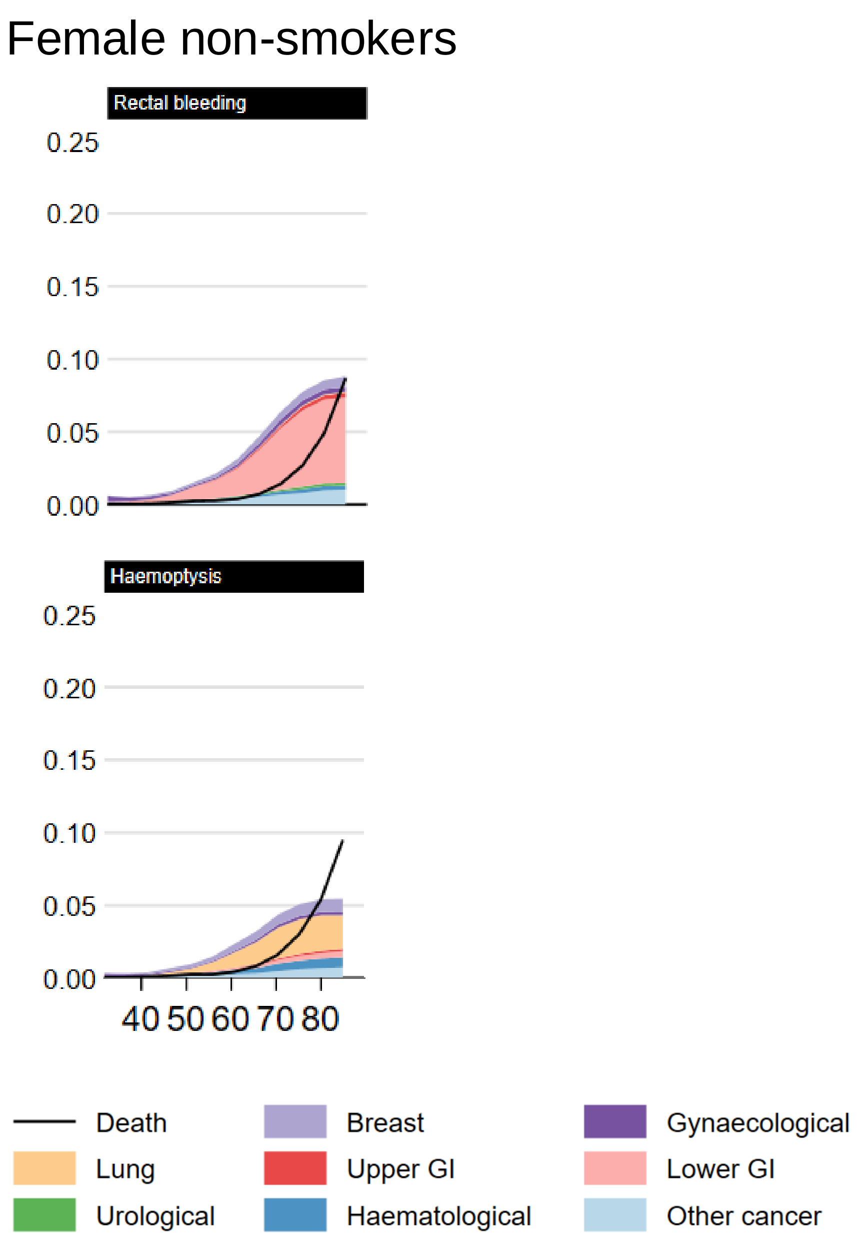
Age from which (cancer-specific) risk at 12 months exceeds 3%

Cohort	Symptom	Any	Lung	Upper GI	Lower GI	Uro.	Prostate	Haem.	Other
Male non-smokers	**Reference group**	**90**							
	Abdominal pain	60							
	Rectal bleeding	60			65				
	Haemoptysis	**60**	**70**						
	Haematuria	55				55	65		
	Weight loss	60					80		
Male smokers	**Reference group**	**75**							
	Abdominal pain	60							
	Rectal bleeding	60			60				
	Haemoptysis	**55**	**55**						
	Haematuria	50				55	70		
	Weight loss	55	70	75					

Results – risk of cancer diagnosis within 12 months

Sex-specific differences in baseline hazard of different cancers mean that similar HRs translate to radically different cancer risk





Summary 1 – main results

Symptoms are highly informative about short-term risk of cancer diagnosis – and non-cancer mortality

We provide (age-sex-smoking status-specific) estimates of risk of cancer diagnosis, by site, for 15 symptoms - updating and improving existing evidence

Policy questions

Should all (symptom-age-sex-smoking status) groups with an estimated any-cancer risk of above 3% should be added to NICE referral guidelines?

How should risk of non-cancer mortality affect referral decision-making? (Should it at all?)

Do symptomatic referral pathways need further adaptation to reflect the range of cancers associated with different symptoms?

Thanks!

For further details, see medRxiv preprint:

https://doi.org/10.1101/2024.05.21.24307662

Summary 2 – limitations

Reliant on completeness of coding of CPRD symptoms – one reason for ‘random’ reference group rather than ‘symptom free’

Only captures (15) symptoms and age, sex, smoking-status – know many other clinical features matter!

Index date = first symptom only, failing to make full use of longitudinal nature of the underlying dataset