



Centre for Reviews
and Dissemination

A Bayesian approach to sharing information on sensitivity of a Multi-Cancer Early Detection test across and within tumour types

Sofia Dias¹, Yiwen Liu¹, Stephen Palmer², Marta O Soares²

¹ Centre for Reviews and Dissemination, University of York, York, UK

² Centre for Health Economics, University of York, York, UK



1



Centre for Reviews
and Dissemination

Disclaimer

Financial support for this study was provided by NHS England.

The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the results.

Authors report no conflict of interest for this work.

Preprint [arXiv:2504.21517](https://arxiv.org/abs/2504.21517)



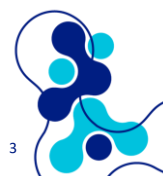
2

Motivation



Centre for Reviews
and Dissemination

- Multi-cancer early detection (MCED) tests are a novel technology that can detect potential signs of cancer before symptoms present
 - based on a single sample of blood
- The use of an MCED test as part of a screening programme could potentially lead to cancers being detected at an earlier stage (stage-shift), provided **sensitivity of the test at early cancer stages is high**.
- The Galleri® (GRAIL) test detects circulating tumour DNA (ctDNA) that signals the presence of cancer then uses a machine learning-based classification algorithm to identify patterns predictive of cancer.
 - The test looks for a signal shared by more than **50** different types of cancer
 - Results are provided as test negative, or test positive, with a suggested “cancer signal of origin”



3

3

Brief note on diagnostic accuracy

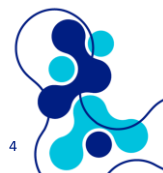


Centre for Reviews
and Dissemination

- Sensitivity (True Positive Rate): The ability of a test to correctly identify those with the disease
 - True positives out of all with disease
- Specificity (True Negative Rate): The ability of a test to correctly identify those without the disease
 - True negatives out of all without disease

$$\text{Sensitivity} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$

$$\text{Specificity} = \frac{\text{True Negatives (TN)}}{\text{True Negatives (TN)} + \text{False Positives (FP)}}$$



4

4

Ongoing RCT and economic evaluation



Centre for Reviews
and Dissemination

- [NHS-Galleri](#) RCT of repeat screening in the general population
 - Funded by GRAIL Bio UK Ltd
 - NHS England is a partner and provides follow-up and care related to the trial
 - Recruited 140,000 asymptomatic volunteers, aged 50-77
 - Expected to report in 2026
 - Powered on stage-shift across **all** cancer types
 - Unlikely to be able to robustly inform stage-shift for **each** cancer type
- An economic model is being developed at York, to provide an initial estimate of the cost-effectiveness of a screening programme using the Galleri test in the UK NHS
 - Will incorporate the primary results of the NHS-Galleri trial.

5



5

Decision-making context



Centre for Reviews
and Dissemination

- Accuracy of the Galleri test to detect each different cancer type at different stages is a key parameter for economic evaluation
- Evidence in the economic model that determines the attribution of the overall stage shift results from the NHS-Galleri trial across cancer types needs to be robust
- Key parameters: **sensitivity of the Galleri test in detecting each different cancer type at different stages**
 - Stage-shift for each cancer type unlikely to be robustly informed from NHS-Galleri trial
 - Stages 1-4 are of interest for all cancers
 - Different treatment options with different effectiveness and costs

6



6

Existing Evidence



Centre for Reviews
and Dissemination

- Information on sensitivity for each cancer type and stage is limited
- A systematic review of accuracy evidence for MCED tests identified, quality appraised and summarised evidence on 13 different technologies.
 - Limited evidence
 - Test sensitivity was found to vary across different technologies, cancer types and stages.
- Three studies evaluated the performance of the Galleri test
 - differed in design and populations
 - none included high-quality evidence in a screening population

[Wade et al. \(2025\)](#) Multi-cancer early detection tests for general population screening: a systematic literature review. *Health Technol Assess*;29(02).

7



7

Evidence: Galleri test



Centre for Reviews
and Dissemination

- PATHFINDER: prospective cohort study, adults aged 50 and over with or without elevated risk of cancer in the US; N=6621 (87 had cancer)
- SYMPLIFY: prospective cohort study, adults aged 18 and over who were referred for urgent investigation of cancer symptoms in the UK; N=5461 (352 had cancer)
- CCGA sub-study 3 (CCGA3): case-control study, adults aged 20 and over with and without cancer in North America (US and Canada); cancer arm: N=2823, non-cancer arm: N=1254
- Diagnostic accuracy of the Galleri test varied:
 - High specificity (98.4% to 99.5%).
 - Sensitivity pooled across all cancer types and stages 20.8% to 66.5 3% (all studies)
- Sensitivity differs within and across studies, by cancer types and stage.
 - lower for detecting earlier stage cancers (Stages 1-2) compared with later stage cancers (Stages 3-4)
- Information on sensitivity for each cancer type and stage is limited

[Wade et al. \(2025\)](#) Multi-cancer early detection tests for general population screening: a systematic literature review. *Health Technol Assess*;29(02).

8



8

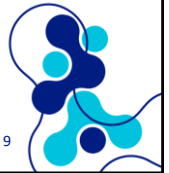
Evidence: Galleri test



Centre for Reviews and Dissemination

- PATHFINDER: prospective cohort study, adults aged 50 and over with or without elevated risk of cancer in the US; N=6621 (87 had cancer)
- SYMPLIFY: prospective cohort study, adults aged 18 and over who were referred for urgent investigation of cancer symptoms in the UK; N=5461 (352 had cancer)
- **CCGA sub-study 3 (CCGA3): case-control study, adults aged 20 and over with and without cancer in North America (US and Canada); cancer arm: N=2823, non-cancer arm: N=1254**
- Diagnostic accuracy of the Galleri test varied:
 - High specificity (98.4% to 99.5%).
 - Sensitivity pooled across all studies 20.8% to 66.5 3% (across all cancer types and stages)
- Sensitivity differs within and across studies, by cancer types and stage.
 - lower for detecting earlier stage cancers (Stages 1-2) compared with later stage cancers (Stages 3-4)
- Information on sensitivity for each cancer type and stage is limited

Wade et al. (2025) Multi-cancer early detection tests for general population screening: a systematic literature review. *Health Technol Assess*;29(02).



9

9

Data: CCGA3

- Overall large number of cancers detected
- But for each cancer type/stage numbers limited
- Many zeros - yellow
- Many 100% (due to small numbers) – green
- Some missing stage data (very few; ignore for now)

Code	Cancer type	Sensitivity by cancer type and stage, CCGA3; s/S (%)					
		Overall	Stage I	Stage II	Stage III	Stage IV	Missing
1	Bladder	8/23 (34.8)	2/6 (33.3)	1/11 (9.1)	3/4 (75.0)	2/2 (100.0)	0/0 (NA)
2	Breast	160/524 (30.5)	7/265 (2.6)	86/181 (47.5)	47/55 (85.5)	20/22 (90.9)	0/1 (0.0)
3	Colon/Rectum	169/206 (82.0)	13/30 (43.3)	34/40 (85.0)	58/66 (87.9)	61/64 (95.3)	3/6 (50.0)
4	Head and neck	90/105 (85.7)	12/19 (63.2)	14/17 (82.4)	16/19 (84.2)	48/50 (96.0)	0/0 (NA)
5	Kidney	18/99 (18.2)	3/61 (4.9)	2/9 (22.2)	1/7 (14.3)	12/22 (54.5)	0/0 (NA)
6	Liver/bile duct	43/46 (93.5)	6/6 (100.0)	7/10 (70.0)	9/9 (100.0)	20/20 (100.0)	1/1 (100.0)
7	Lung	302/404 (74.8)	21/96 (21.9)	35/44 (79.5)	107/118 (90.7)	138/145 (95.2)	1/1 (100.0)
8	Lymphoma	98/174 (56.3)	9/33 (27.3)	28/48 (58.3)	33/46 (71.7)	28/46 (60.9)	0/1 (0.0)
9	Ovary	54/65 (83.1)	5/10 (50.0)	4/5 (80.0)	27/31 (87.1)	18/19 (94.7)	0/0 (NA)
10	Pancreas	113/135 (83.7)	13/21 (61.9)	12/20 (60.0)	18/21 (85.7)	70/73 (95.9)	0/0 (NA)
11	Prostate	47/420 (11.2)	3/95 (3.2)	12/243 (4.9)	7/50 (14.0)	25/30 (83.3)	0/2 (0.0)
12	Sarcoma	18/30 (60.0)	4/10 (40.0)	2/2 (100.0)	5/10 (50.0)	6/7 (85.7)	1/1 (100.0)
13	Thyroid	0/14 (0.0)	0/11 (0.0)	0/1 (0.0)	0/1 (0.0)	0/1 (0.0)	0/0 (NA)
14	Uterus	44/157 (28.0)	20/120 (16.7)	3/10 (30.0)	17/23 (73.9)	4/4 (100.0)	0/0 (NA)
15	Lymphoid leukaemia	21/51 (41.2)					
16	Melanoma	6/13 (46.2)	0/2 (0.0)	0/2 (0.0)	0/3 (0.0)	6/6 (100.0)	0/0 (NA)
17	Plasma cell neoplasm	34/47 (72.3)	11/17 (64.7)	14/16 (87.5)	9/14 (64.3)	0/0 (NA)	0/0 (NA)
18	Anus	18/22 (81.8)	1/4 (25.0)	3/4 (75.0)	13/13 (100.0)	1/1 (100.0)	0/0 (NA)
19	Cervix	20/25 (80.0)	7/12 (58.3)	5/5 (100.0)	7/7 (100.0)	1/1 (100.0)	0/0 (NA)
20	Gallbladder	12/17 (70.6)	0/2 (0.0)	1/3 (33.3)	3/4 (75.0)	8/8 (100.0)	0/0 (NA)
21	Urothelial tract	8/10 (80.0)	0/2 (0.0)	0/0 (NA)	0/0 (NA)	8/8 (100.0)	0/0 (NA)
22	Myeloid neoplasm	2/10 (20.0)					
23	Oesophagus	85/100 (85.0)	1/8 (12.5)	11/17 (64.7)	32/34 (94.1)	40/40 (100.0)	1/1 (100.0)
24	Stomach	20/30 (66.7)	1/6 (16.7)	3/6 (50.0)	4/5 (80.0)	12/12 (100.0)	0/1 (0.0)
25	Other	30/59 (50.9)	2/11 (18.2)	3/3 (100.0)	13/18 (72.2)	11/18 (61.1)	1/3 (33.3)
26	Multiple primaries	16/19 (84.2)	2/2 (100.0)	3/5 (60.0)	6/6 (100.0)	5/6 (83.3)	0/0 (NA)
27	Unknown primary	17/18 (94.4)	0/0 (NA)	1/1 (100.0)	1/2 (50.0)	13/13 (100.0)	2/2 (100.0)
	Total	1453/2823 (51.5)	143/849 (16.8)	284/703 (40.4)	436/566 (77.0)	557/618 (90.1)	10/20 (50.0)

10

Data: CCGA3

Code	Cancer type	Sensitivity by cancer type and stage, CCGA3; s/S (%)					
		Overall	Stage I	Stage II	Stage III	Stage IV	Missing
1	Bladder	8/23 (34.8)	2/6 (33.3)	1/11 (9.1)	3/4 (75.0)	2/2 (100.0)	0/0 (NA)
2	Breast	160/524 (30.5)	7/265 (2.6)	86/181 (47.5)	47/55 (85.5)	20/22 (90.9)	0/1 (0.0)
3	Colon/Rectum	169/206 (82.0)	13/30 (43.3)	34/40 (85.0)	58/66 (87.9)	61/64 (95.3)	3/6 (50.0)
4	Head and neck	90/105 (85.7)	12/19 (63.2)	14/17 (82.4)	16/19 (84.2)	48/50 (96.0)	0/0 (NA)
5	Kidney	18/99 (18.2)	3/61 (4.9)	2/9 (22.2)	1/7 (14.3)	12/22 (54.5)	0/0 (NA)
6	Liver/bile duct	43/46 (93.5)	6/6 (100.0)	7/10 (70.0)	9/9 (100.0)	20/20 (100.0)	1/1 (100.0)
7	Lung	302/404 (74.8)	21/96 (21.9)	35/44 (79.5)	107/118 (90.7)	138/145 (95.2)	1/1 (100.0)
8	Lymphoma	98/174 (56.3)	9/33 (27.3)	28/48 (58.3)	33/46 (71.7)	28/46 (60.9)	0/1 (0.0)
9	Ovary	54/65 (83.1)	5/10 (50.0)	4/5 (80.0)	27/31 (87.1)	18/19 (94.7)	0/0 (NA)
10	Pancreas	113/135 (83.7)	13/21 (61.9)	12/20 (60.0)	18/21 (85.7)	70/73 (95.9)	0/0 (NA)
11	Prostate	47/420 (11.2)	3/95 (3.2)	12/243 (4.9)	7/50 (14.0)	25/30 (83.3)	0/2 (0.0)
12	Sarcoma	18/30 (60.0)	4/10 (40.0)	2/2 (100.0)	5/10 (50.0)	6/7 (85.7)	1/1 (100.0)
22	Myeloid neoplasm	2/10 (20.0)					0 (NA)
23	Oesophagus	85/100 (85.0)	1/8 (12.5)	11/17 (64.7)	32/34 (94.1)	40/40 (100.0)	1/1 (100.0)
24	Stomach	20/30 (66.7)	1/6 (16.7)	3/6 (50.0)	4/5 (80.0)	12/12 (100.0)	0/1 (0.0)
25	Other	30/59 (50.9)	2/11 (18.2)	3/3 (100.0)	13/18 (72.2)	11/18 (61.1)	1/3 (33.3)

Breast, prostate, lung, colon/rectum make up a large proportion of the included and detected cancers at all stages.

Breast cancer samples mainly early stage (1-2)

Samples for lung and colon/rectum mostly advanced cancer (3-4).

11

Data: CCGA3

Code	Cancer type	Sensitivity by cancer type and stage, CCGA3; s/S (%)					
		Overall	Stage I	Stage II	Stage III	Stage IV	Missing
1	Bladder	8/23 (34.8)	2/6 (33.3)	1/11 (9.1)	3/4 (75.0)	2/2 (100.0)	0/0 (NA)
2	Breast	160/524 (30.5)	7/265 (2.6)	86/181 (47.5)	47/55 (85.5)	20/22 (90.9)	0/1 (0.0)
3	Colon/Rectum	169/206 (82.0)	13/30 (43.3)	34/40 (85.0)	58/66 (87.9)	61/64 (95.3)	3/6 (50.0)
4	Head and neck	90/105 (85.7)	12/19 (63.2)	14/17 (82.4)	16/19 (84.2)	48/50 (96.0)	0/0 (NA)
5	Kidney	18/99 (18.2)	3/61 (4.9)	2/9 (22.2)	1/7 (14.3)	12/22 (54.5)	0/0 (NA)
6	Liver/bile duct	43/46 (93.5)	6/6 (100.0)	7/10 (70.0)	9/9 (100.0)	20/20 (100.0)	1/1 (100.0)
7							
8							
9							
10							
11	Prostate	47/420 (11.2)	3/95 (3.2)	12/243 (4.9)	7/50 (14.0)	25/30 (83.3)	0/2 (0.0)
12	Sarcoma	18/30 (60.0)	4/10 (40.0)	2/2 (100.0)	5/10 (50.0)	6/7 (85.7)	1/1 (100.0)
13	Thyroid	0/14 (0.0)	0/11 (0.0)	0/1 (0.0)	0/1 (0.0)	0/1 (0.0)	0/0 (NA)
14	Uterus	44/157 (28.0)	20/120 (16.7)	3/10 (30.0)	17/23 (73.9)	4/4 (100.0)	0/0 (NA)
15	Lymphoid leukaemia	21/51 (41.2)					
16	Melanoma	6/13 (46.2)	0/2 (0.0)	0/2 (0.0)	0/3 (0.0)	6/6 (100.0)	0/0 (NA)
17	Plasma cell neoplasm	34/47 (72.3)	11/17 (64.7)	14/16 (87.5)	9/14 (64.3)	0/0 (NA)	0/0 (NA)
18	Anus	18/22 (81.8)	1/4 (25.0)	3/4 (75.0)	13/13 (100.0)	1/1 (100.0)	0/0 (NA)
19	Cervix	20/25 (80.0)	7/12 (58.3)	5/5 (100.0)	7/7 (100.0)	1/1 (100.0)	0/0 (NA)
20	Gallbladder	12/17 (70.6)	0/2 (0.0)	1/3 (33.3)	3/4 (75.0)	8/8 (100.0)	0/0 (NA)
21	Urothelial tract	8/10 (80.0)	0/2 (0.0)	0/0 (NA)	0/0 (NA)	8/8 (100.0)	0/0 (NA)
22	Myeloid neoplasm	2/10 (20.0)					
23	Oesophagus	85/100 (85.0)	1/8 (12.5)	11/17 (64.7)	32/34 (94.1)	40/40 (100.0)	1/1 (100.0)
24	Stomach	20/30 (66.7)	1/6 (16.7)	3/6 (50.0)	4/5 (80.0)	12/12 (100.0)	0/1 (0.0)
25	Other	30/59 (50.9)	2/11 (18.2)	3/3 (100.0)	13/18 (72.2)	11/18 (61.1)	1/3 (33.3)
26	Multiple primaries	16/19 (84.2)	2/2 (100.0)	3/5 (60.0)	6/6 (100.0)	5/6 (83.3)	0/0 (NA)

- Melanoma and urothelial tract had no cancers detected at stages 1-3
- Thyroid had no detections at all

12

Galleri test: Data features

- Variation in test sensitivity across cancer types and stages due to known differences in ctDNA expression
 - Different cancer types will express different levels of ctDNA which is what the Galleri test detects
- For a large proportion of the cancer types, the numbers of observed patients are small
 - significant uncertainty in test sensitivity at each stage (and overall)
- Where the test is expected to have similar sensitivity between two or more tumour types, or across different stages, data on one cancer type at a particular stage, could be used to **share information and strengthen inferences on other cancer types/stages.**

13



13

Aims

- Explore the use of Bayesian information sharing models applied to data on sensitivity of the Galleri MCED test
- Examine support of the data for sharing evidence across cancer types and cancer stages
- Assess improvements in precision
- **Models implemented in JAGS**
<https://github.com/MCED-Galleri-HealthEconomicEval-Program/BayesianModelTestSens-1>

14



14

Modelling the data



Centre for Reviews
and Dissemination

$$s_{ijk} \sim \text{Binomial}(p_{ijk}, S_{ijk})$$

$$\text{logit}(p_{ijk}) = \delta_{ijk} \quad \begin{array}{l} i = 1 \\ j = 1, \dots, 25 \text{ (cancer types)} \\ k = 1, 2, 3, 4 \text{ (cancer stages)} \end{array}$$

$\delta_{ijk} = \mu_{jk}$ common effect (across studies)

or

$\delta_{ijk} \sim \text{Normal}(\mu_{jk}, \sigma^2)$ random effects

s_{ijk} - true positives for cancer type j at stage k , in study i

p_{ijk} - probability of true positive (sensitivity) for cancer type j at stage k , in study i

S_{ijk} - number of individuals with cancer type j at stage k , in study i

δ_{ijk} - log-odds of sensitivity for cancer type j at stage k , in study i

- Will only use evidence from 1 study (CCGA3)
- Methods extend to multiple studies

15



15

Bayesian information sharing models



Centre for Reviews
and Dissemination

- Can be used to synthesise evidence across tumour types and/or stages of disease.
- Do not need to assume that evidence from different sources is estimating the exact same parameter
- Can impose different levels of sharing information based on different assumptions about the relationships across the evidence sources.
- Structural and exchangeability-based relationships can be used and combined with prior information
- However, unexplained heterogeneity can limit precision gains

[Nikolaidis et al. \(2021\)](#) Classifying information-sharing methods. BMC Med. Res. Methodol.;21:107.
[Hemming et al. \(2012\)](#) Pooling systematic reviews of systematic reviews: a Bayesian panoramic meta-analysis. Stat. Med. 2012;31:201-16.

16



16

Clinical determinants of heterogeneity



Centre for Reviews
and Dissemination

- Tumour ctDNA shedding thought to be the main driver of differences in test performance
 - ctDNA independently predicts test performance for different stages
 - Different cancer types shed different levels of ctDNA
- Targeted literature search to identify characteristics that could determine potential homogeneity or heterogeneity of the sensitivity of the Galleri test across different cancer types and stages
- This (indirect) evidence was used to inform plausible sharing assumptions, to be implemented in the Bayesian information sharing models



17

17

Sensitivity increases with stage (regardless of cancer type)



Centre for Reviews
and Dissemination

- There is consistent evidence that ctDNA shedding is **higher** in advanced stages of cancer compared to localised cancer, regardless of the type of cancer.
- The size of the tumour (which typically reflects a more advanced stage) has also been associated with the detectability of ctDNA in blood samples.
- **Cancers at more advanced stages are expected to shed more ctDNA and are more likely to be detected by the Galleri test.**



18

18

Base model: Sensitivity increases with stage



Centre for Reviews
and Dissemination

- Our **base model** will **constrain** the sensitivities to be independent but monotonically increasing (or equal) with increasing stage, for all cancer types

$$\mu_{j1} \leq \mu_{j2} \leq \mu_{j3} \leq \mu_{j4}$$

- Independent, non-informative prior distributions:

$$\mu_{jk} \sim \text{Normal}(0, 100^2)$$

[Owen et al. \(2015\)](#) Network Meta-Analysis: Development of a Three-Level Hierarchical Modeling Approach Incorporating Dose-Related Constraints. Value Health 18:116-26.

19

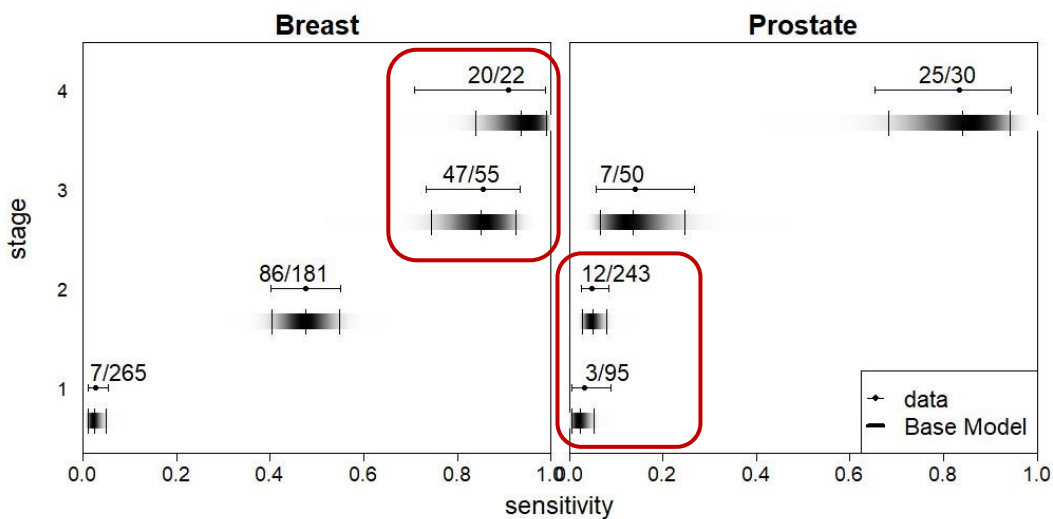


19

Data and base model: Example results 1



Centre for Reviews
and Dissemination



20

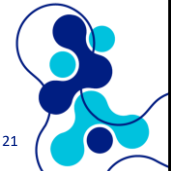
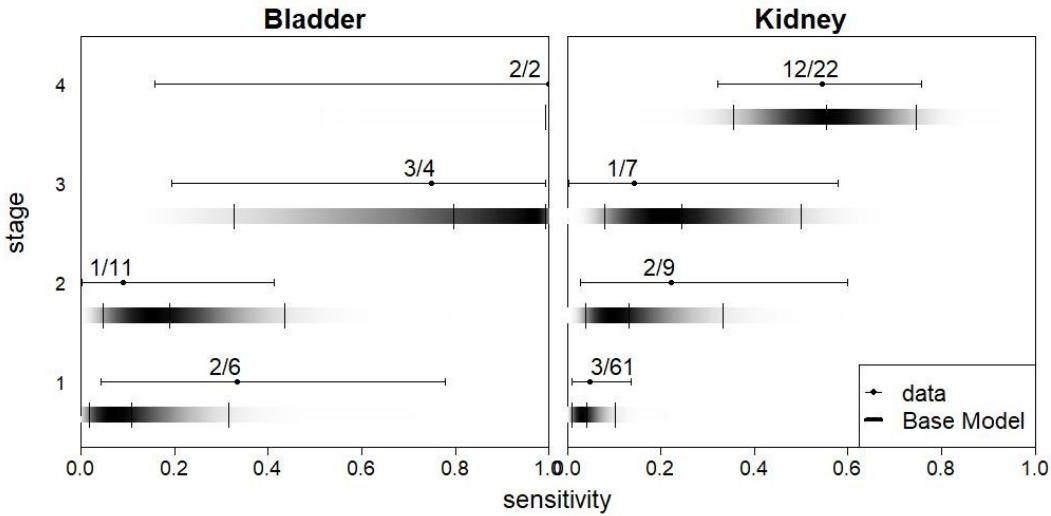


20

Data and base model: Example results 2



Centre for Reviews and Dissemination



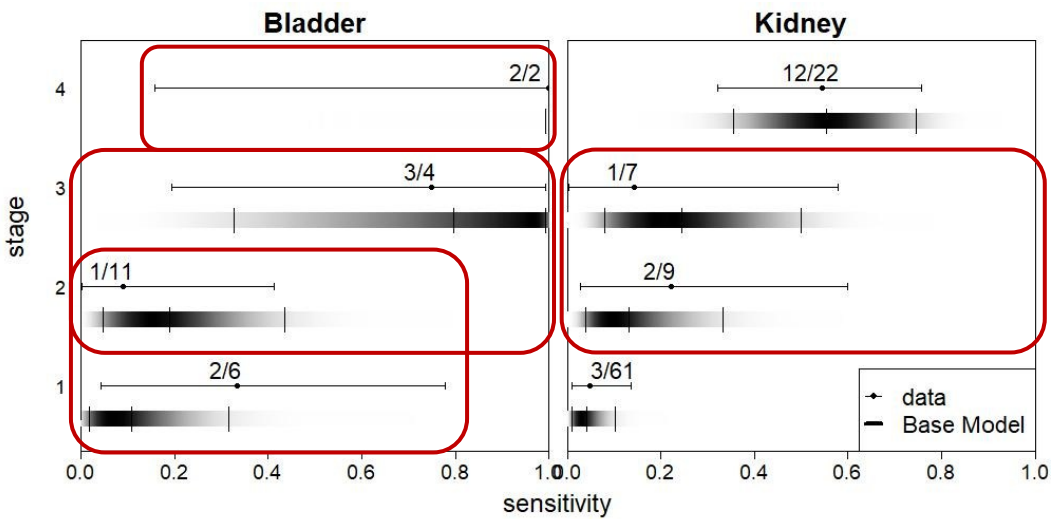
21

21

Data and base model: Example results 2



Centre for Reviews and Dissemination



22

22

Base model: results

- Including constraints implies **information sharing across adjacent cancer stages within-cancer type**
- Leads to increase in precision of estimates across stages for each cancer type compared to a model with no constraints.
- Reflects typical assumptions in evidence synthesis in this area, and mimics previous analyses of these same data
 - Mean estimated sensitivities agree with those used in a previous modelling study^[1] although uncertainty measures were not presented
- Model will serve as a benchmark against which to measure precision increases and changes to point estimates for models that share information across cancer types and stages more explicitly.

[1] [Sasieni et al. \(2023\)](#) Modelled mortality benefits of multi-cancer early detection screening in England. Br J Cancer 2023;129:72-80.

23



23

Additional sharing assumptions

- Assumptions on similarities (or differences) between sensitivities for different cancer types and stages will reflect what is known about the heterogeneity in ctDNA shedding of each cancer type at each stage.
 - Based on review of ctDNA literature
- Cancer types described as “Other”, “Unknown” or “Multiple” are not included in any sharing assumptions.
- Only 22 cancers affected by sharing models

24



24

Sharing models



Centre for Reviews
and Dissemination

- Base model extended to allow different levels of sharing of information between cancer types
- Implement alternative sharing assumptions in addition to the increasing sensitivity constraints based on 3 types of Bayesian models:
 1. Exchangeability models
 2. Mixture models
 3. Class models



25

25

Exchangeability models



Centre for Reviews
and Dissemination

- Exchangeability models allow the probabilities of detecting cancer at particular stages to be assumed to be similar (exchangeable) across some cancer types whilst allowing for heterogeneity between the different cancer types.

$$\mu_{jk} \sim \text{Normal}(m_k, \tau_k^2) \quad m_k \sim \text{Normal}(0, 100^2), \tau_k \sim \text{Uniform}(0, 5)$$

- If exchangeability is only appropriate across certain cancer types, these can be specified in a set V sharing assumed only for j in set V
- Models can be adapted to allow for common means or heterogeneity across only some stages k



26

26

Mixture models



Centre for Reviews
and Dissemination

- Mixture models explicitly consider how much each individual cancer type contributes to the sharing element, using a (mixture) probability parameter which moderates how much, for a given stage, each particular cancer type contributes to the sharing element.
 - Low mixture probability \rightarrow evidence strongly indicates that a cancer type differs from others which are more similar between themselves (i.e. for which there is no strong evidence of difference)
 - High mixture probability \rightarrow no evidence of a difference between a particular cancer type and the others, so information can be shared.
- Mixture models can prevent too much sharing from “extreme” cancer types and stages (i.e. where sensitivities are very different)
 - reduce the contribution of extreme indications to the heterogeneity, and strengthen sharing within the cancer types that are more similar

[Neuenschwander et al. \(2016\)](#) Robust exchangeability designs for early phase clinical trials with multiple strata. Pharm. Stat. 15:123-34.
[Singh et al. \(2025\)](#) Multi-indication Evidence Synthesis in Oncology Health Technology Assessment: Meta-analysis Methods and Their Application to a Case Study of Bevacizumab. MDM 45:17-33.

27



27

Mixture model



Centre for Reviews
and Dissemination

- Variable X determines whether the sensitivity for a particular cancer type and stage is exchangeable with the rest according to a probability π which is given a Beta prior distribution:

$$X_{jk} \sim \text{Bernoulli}(\pi_{jk})$$

$$\pi_{jk} \sim \text{Beta}(a_{jk}, b_{jk})$$

Prior distributions:
non-informative ($a=b=1$), or informed by external evidence
(e.g. elicitation of expert opinion)

$$\mu_{jk} \sim \begin{cases} \text{Normal}(m_k, \tau_k^2) & \text{if } X_{jk} = 1 \quad (\text{sharing}) \\ \text{Normal}(0, 100^2) & \text{if } X_{jk} = 0 \quad (\text{no sharing}) \end{cases}$$

28



28

Class models

- Explicitly include external evidence suggesting that some cancer types are more similar to certain others
- Define groups (classes) of cancers within which sharing stage-specific sensitivity is appropriate
 - but sharing **across** classes is not allowed
- Log-odds of sensitivities for each stage within cancer classes indexed by a vector D , assumed to come from a common distribution:

$$\mu_{jk} \sim \text{Normal}\left(m_{D_j,k}, \tau_{D_j,k}^2\right) \quad m_{ck} \sim \text{Normal}\left(0, 100^2\right), \quad c = 1, \dots, nClass; \quad k = 1, \dots, 4$$

$$\tau_{ck} \sim \text{Uniform}(0, 5)$$

[Owen et al. \(2015\)](#) Network Meta-Analysis: Development of a Three-Level Hierarchical Modeling Approach Incorporating Dose-Related Constraints. Value Health 18:116-26.

29



29

Sharing assumptions: exchangeability models

- Assume that **sensitivity for each stage is similar across all cancer types**
→ **Model 1**

- Sensitivity is shared across all cancer types for each stage

ctDNA literature suggests:

- **Higher sensitivity for stage 4 cancers** than for cancers at stages 1-3 → **Model 2**
 - Sensitivity for stage 4 is shared across all cancer types; others independent
- **Bladder, kidney and thyroid cancer have a lower probability** of being detected at all stages. Information from these cancers not shared with the other cancer types → **Model 3**
 - Sensitivity for each stage shared across cancer types, except for bladder, kidney and thyroid
- **Breast, lung and colorectal cancer have higher sensitivity at stage 4** than other cancer types → **Model 4**
 - Sensitivity for stage 4 is shared across cancer types, **except** breast, lung and colon/rectum; others independent

30



30

Sharing assumptions: other models



Centre for Reviews
and Dissemination

- Mixture models
 - Mixture probabilities depend on cancer type and stage → **Model 5**
 - Single mixture probability across all stages of the same cancer → **Model 6**
 - non-informative (Beta(1,1)) prior distributions for all mixture probabilities
- Cancer Groups (classes) → **Model 7**
 - group 1 = bladder, kidney, prostate, thyroid, melanoma (low sensitivity at all stages)
 - group 2 = remaining cancer types

31



31

Model selection



Centre for Reviews
and Dissemination

- Models with different information sharing assumptions are compared based on
 - measures of model fit
 - **magnitude of estimated heterogeneity** across stages or cancer types (not shown)
 - and **clinical plausibility of model predictions** given by the (shrunk) estimates for each cancer type and stage.
- Models with adequate fit (based on residual deviance) and predictions (based on the shrunk estimates), compared by looking at differences in Deviance Information Criteria (DIC).
- Models with similar residual deviance but lower DIC were preferred.

32



32

Results: model fit



Centre for Reviews
and Dissemination

- Base model has best fit overall
 - Prefer model that includes plausible sharing assumptions to improve precision whilst still fitting the data well
- Models 2 and 4 also fit well
- Fit of Model 2 comparable to Model 4
 - relaxing the assumption of sharing over “high sensitivity” cancers (model 4) was not considered necessary
 - Model 2 is preferred over Model 4
- Model 2 does not allow sharing over early-stages of cancer, where precision gains are likely to have greater value.
- Model 3 has more useful sharing assumptions
 - selected as alternative plausible model (even though fit is slightly worse)

	Residual deviance (100 data points)	Deviance	pD	DIC
Base model	92.7	287.0	74.3	361.4
Model 1	117.8	312.1	88.3	400.3
Model 2	98.6	292.9	76.8	369.7
Model 3	109.7	304.0	80.1	384.0
Model 4	97.1	291.4	77.0	368.4
Model 5	103.6	297.9	90.9	388.8
Model 6	111.5	305.8	86.5	392.3
Model 7	122.4	316.7	85.1	401.8

33

33

Comments on results



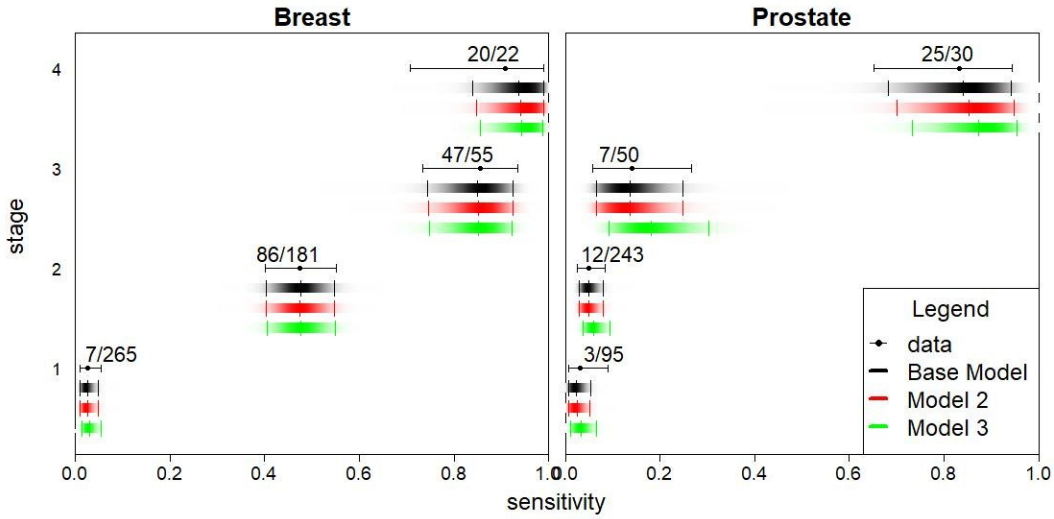
Centre for Reviews
and Dissemination

- Support was strongest for the assumption that sensitivity of the Galleri test can be shared across stage 4 for all cancer types (model 2)
- However, movement from detection of cancers at stage 4 to detection at earlier stages is the most important determinant of value of this type of test when used for screening, so this model is not very useful.
- Model 3 also considered acceptable and more useful.
- There was significant heterogeneity within stages 1-3
 - Our explorations explained some of the heterogeneity but remaining unexplained heterogeneity was still high
 - **Limits sharing**

34

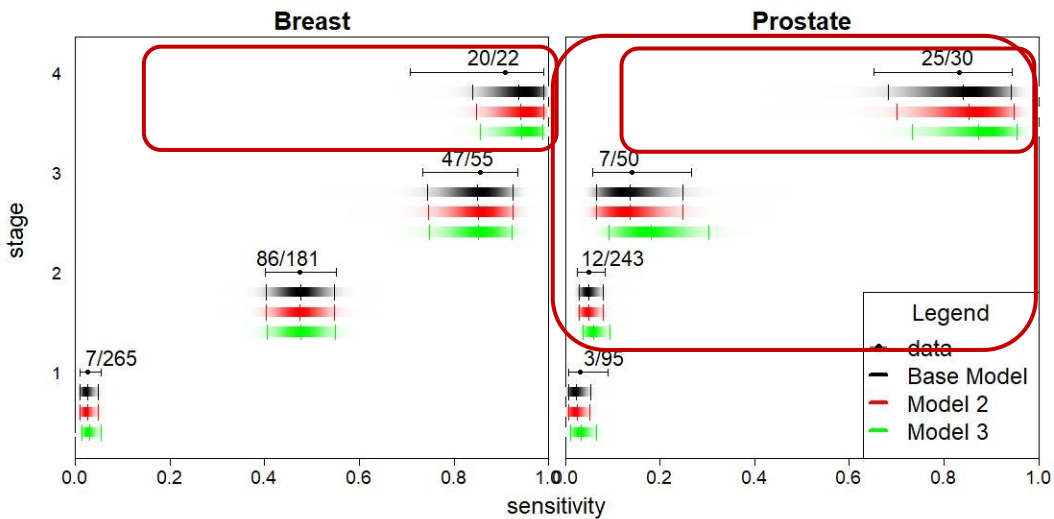
34

Selected models: Example results 1



35

Selected models: Example results 1

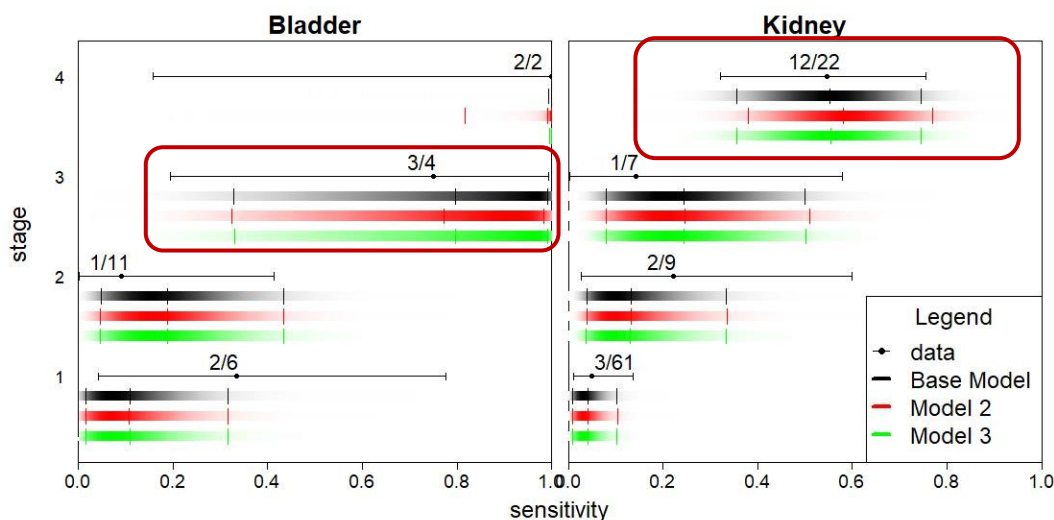


36

Selected models: Example results 2



Centre for Reviews
and Dissemination



37

Discussion



Centre for Reviews
and Dissemination

- Conducted exploratory analysis of determinants of heterogeneity across tumour types and stages.
- Sharing models can strengthen the evidence on sensitivity across tumour types and stages, whilst quantifying heterogeneity between cancer types.
- Found substantial heterogeneity of test sensitivity across cancer types and stages, particularly for detection of early-stage cancers
 - Suggests there are some cancer types that the Galleri test is more likely to find early than others.
- Models share information only weakly but can provide some increases in precision
- Gains in precision just by constraining values to increase with stage for each tumour type
- Additional sharing assumptions implemented using flexible sharing models are supported by the data
 - But limited sharing due to heterogeneity

38

38

Discussion (cont.)



Centre for Reviews
and Dissemination

- Explorations of heterogeneity based on current evidence on the determinants of ctDNA shedding and detection in a blood test
 - Evidence focussed on a few of the most common cancers and in comparing late-stage vs early-stage rather than providing stage-specific evidence.
- Further understanding of the heterogeneity would allow models with stronger sharing assumptions to be implemented, and evidence to be used more efficiently.
- More data needed to identify which cancer types are more likely to be detected **early** by the Galleri test.
- Until more is known, sharing of evidence using flexible models will be required to accommodate uncertainty.
- Further research will seek expert opinion to support more complex, but also more realistic, sharing assumptions.
 - E.g. eliciting support for groupings of cancer types or eliciting prior distributions for sharing probabilities in the mixture models.

39



39

NHS-Galleri trial evidence



Centre for Reviews
and Dissemination

- We used only one study because we expected **overall** sensitivity to differ with different study designs
- But, conditional on stage and tumour type, it is possible that sensitivities may be similar across different study designs
 - Additional analyses pooling the 3 existing Galleri studies did not change results
 - Few cases in other 2 studies, not much additional information
- We will explore incorporating tumour and stage-specific data from NHS-Galleri trial with the existing Galleri evidence in sharing models

40



40



Centre for Reviews
and Dissemination



THANK YOU!
sofia.dias@york.ac.uk

Preprint [arXiv:2504.21517](https://arxiv.org/abs/2504.21517)
<https://github.com/MCED-Galleri-HealthEconomicEval-Program>

