Title: Potential spillover effects on diagnostic delay for cancer during the NHS-Galleri trial: a quasi-experimental difference-in-differences study

Authors: Sean Mann (0000-0002-5771-1458), Joshua Eagan, Beth Ann Griffin

Sean Mann
RAND, 1200 S Hayes St., Arlington, VA 22202, USA
Policy Analyst

Joshua Eagan
RAND, 1200 S Hayes St., Arlington, VA 22202, USA
Statistical Research Programmer

Beth Ann Griffin
RAND, 1200 S Hayes St., Arlington, VA 22202, USA
Senior Statistician

Correspondence to: S Mann smann@rand.org

Contributors: SM and BAG developed the plan for the study. SM and JE performed the data analysis. SM wrote the first draft of the paper. All authors contributed to the development of the study, interpretation of results, and revision of the manuscript. SM supervised the study and is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This study did not receive any external funding. This work was supported by RAND, which employs the authors. RAND had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/disclosure-of-interest/ and declare: all authors are employed by RAND, which provided support for this work. SM has received research funding from the Patient-Centered Outcomes Research Institute and Centers for Medicare & Medicaid Services. BAG has received research funding and honoraria from the National Institutes of Health and received support for travel to a meeting from Syracuse University. Authors had no other relationships or activities that could appear to have influenced the submitted work.

Ethics approval: Not applicable.

Data availability statement: Data are available in a public, open access repository. The data used in this study are aggregated monthly provider cancer waiting times statistics which are freely available to download from the following website: https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/monthly-prov-cwt/

Transparency: SM affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Abstract

Objective: To examine whether regions participating in the NHS-Galleri trial of a multi-cancer early detection test experienced changes in cancer diagnostic delay rates that could indicate the presence of spillover due to constraints on care delivery.

Design: Quasi-experimental difference-in-differences analysis.

Setting: NHS England, April 2021-September 2023.

Participants: All 21 cancer alliance regions in NHS England, 8 of which participated in the trial and 13 of which did not.

Main outcomes: Rates of cancer diagnostic delays, defined as the percentage of patient referrals for investigation of possible cancer that resulted in diagnostic resolution within 28 days of initial referral. A difference-in-differences analysis was used to compare diagnostic delay rates in regions that participated in the NHS-Galleri trial to regions that did not participate, before and after the start of the trial. Three primary analyses were conducted: for all types of suspected cancer; for the most commonly referred cancers (breast, skin, lower gastrointestinal); and for less commonly referred cancers (all other types).

Results: We observed no significant differences in rates of diagnostic delay for all cancers or for the most commonly referred cancers. For less commonly referred cancers, the percentage of patients experiencing diagnostic delays increased from 32.6% to 35.1% in regions participating in the NHS-Galleri trial from the 6 months before to the 6 months after trial start, compared with a small reduction from 32.7% to 31.7% in regions that did not participate (adjusted difference 3.4 percentage points, 95% confidence interval 1.2 to 5.7). These differences persisted from 6 to 12 months after trial start and receded the following year.

Conclusions: Regional participation in the NHS-Galleri trial was associated with a statistically significant increase in diagnostic delay rates for less commonly referred cancers in the first year of the trial, but not for cancer overall or the most commonly referred cancers. Rollout of the trial intervention may have increased demand for downstream diagnostic services and, in a capacity constrained environment, led to region-wide increases in delayed diagnosis for less commonly referred cancers. If patients in the control group faced delayed diagnosis as a result, this would represent spillover due to constraints on care delivery during the first year and could affect measurement of the trial's primary outcome, incidence of later stage cancer at time of diagnosis. This study underscores the importance of addressing spillover in clinical trials of interventions that affect availability of healthcare services.

Introduction

Detecting treatable cancer early is a public health priority. Randomized controlled trials conducted in real-world health settings are used to determine whether new screening tests,[1] patient navigation services,[2,3] or automated reminder systems[4,5] lead to earlier detection of cancer. Interventions that prove effective in these trials are often integrated into clinical practice guidelines and adopted broadly across health systems.[6-8] Some of these trials may suffer from an overlooked source of bias, however. In real-world health settings, cancer diagnostic resources such as specialist appointments, imaging equipment, or laboratory staff are often limited.[9-11] During a trial, an intervention that increases utilization of a scarce resource by one group of patients can reduce availability of that resource for other patients. This can negatively affect outcomes for patients in the control group and lead trial investigators to overestimate the benefits of an intervention.[12] Economists refer to this problem as "negative spillover" or "crowding out" and some randomized trials of social welfare and public health interventions have accordingly been designed to avoid or test for its presence.[13-15] This threat to trial validity is overlooked in the medical literature and is not addressed in guidance on clinical trial design,[16,17] reporting,[18-20] or risk of bias assessment.[21-23]

The ongoing NHS-Galleri trial is an example of a randomized trial in which spillover due to constraints on care delivery may be a concern. This trial randomized over 140,000 adults (50-77 years) to either receive usual care ("standard-of-care and existing screening modalities") or usual care plus annual screening using a new liquid biopsy multi-cancer early detection (MCED) screening test (Galleri®).[1] The Galleri test relies on next-generation sequencing and artificial intelligence to screen for more than 50 different types of cancer.[24] Participants with suspected cancer—whether identified via usual care or due to a positive result from the Galleri screening test—were referred for follow-up diagnostic assessment according to standard NHS pathways.[1] NHS capacity shortfalls and long wait times for cancer diagnosis have been a long-standing concern[25] and record high referral numbers and rates of delayed follow-up appointments were reported in 2021 as the NHS-Galleri trial was starting.[26] An increase in patient referrals due to rollout of the Galleri screening test may have further strained diagnostic resources within participating cancer alliances (regional NHS partnerships that coordinate cancer care) and led to higher rates of delayed diagnosis, including for patients in the control arm of the trial. This, in turn, could lead to biased overestimation of the Galleri test's ability to achieve the trial's primary objective: reduced incidence of later stage cancer at time of diagnosis.[1] If this type of spillover occurred, screening for cancer using the Galleri test would likely underperform expectations if adopted broadly across the NHS.

We leverage provider-level NHS data on cancer diagnostic delays to examine whether spillover due to constraints on care delivery occurred during the NHS-Galleri trial. This form of spillover is of particular concern for types of suspected cancer detected by the Galleri test for which healthcare systems possess limited spare diagnostic capacity. We

hypothesized that, once the trial started, participating regions would report an increase in diagnostic delays when compared to regions not participating in the trial. Such an increase would suggest that negative spillover, in the form of delayed diagnosis for patients in the control group, could have occurred during the trial.

Methods

Data sources and study population

We obtained monthly provider-level data on cancer diagnostic delays and staffing from three publicly available NHS England datasets. The NHS England Provider-based Cancer Waiting Times dataset reports referral volumes together with the number of referrals for investigation of possible cancer that resulted in diagnostic resolution within 28 days.[27] We linked this dataset using provider codes to two other NHS datasets containing health care staff numbers[28] and absences.[29] The linked data were then aggregated up to the level of cancer alliance regions using NHS provider, integrated care board, and cancer alliance linking files.[30,31] Providers that did not link to cancer alliances during this initial aggregation step were subsequently linked to alliances via manual review of provider websites and locations. We excluded providers that provided services across multiple cancer alliances. The resulting dataset contained monthly counts of total referrals and referrals that breached the 28 day faster diagnosis standard in each cancer alliance. All 21 NHS England cancer alliance regions were included in our analysis, as were all months for which finalized 28-day faster diagnosis data were available (as of May 16, 2024), spanning the period from April 2021 through September 2023. The study period includes a six-month pre-trial period and the first two years of the trial, which remains ongoing.

Exposure to NHS-Galleri trial

The NHS Galleri trial is being conducted in 8 of the 21 cancer alliance regions in England (fig 1). The trial investigators selected participating regions based on factors that included ethnic diversity, socioeconomic status, and relatively higher rates of late-stage cancer diagnosis and cancer mortality.[1]

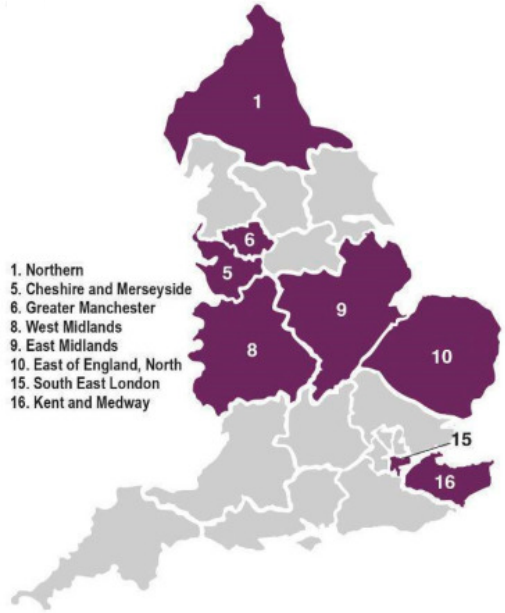

Figure 1. Map of the 21 NHS England cancer alliance regions. The eight cancer alliances participating in the NHS-Galleri trial are colored purple.
Source: Map image reused under CC-BY license from the NHS-Galleri trial protocol.[1]

Trial enrollment began August 31, 2021[1] and participants began providing blood samples for the Galleri test on or after September 13, 2021.[32] Rollout of the Galleri test across the eight participating cancer alliances was phased in over the period from September through November 2021,[33] though public information on exact start dates within each cancer alliance was unavailable. Participants were enrolled over 10 months through July 2022[34] and the trial remains ongoing. The intervention group received Galleri screening tests on an annual basis during the trial, delivered throughout the year. Patients assigned to the intervention group for whom the Galleri test detected a possible cancer signal were usually notified within 30 days following blood draw and referred for follow-up diagnostic testing.[35] General practitioners caring for these patients were also notified of positive test results and instructed to ensure that patients received confirmatory diagnostic procedures for the suspected cancer site following standard clinical guidance.[1]

Given the above timeline, we treated October 2021 as the first month in which participating cancer alliances were exposed to the NHS-Galleri trial. To accommodate uncertainty in timing of exposure across all eight cancer alliance regions we also conducted a sensitivity analysis excluding data from September, October, and November 2021 as a phase-in period. Exposure to the trial continued through the end of the study period in September 2023.

Outcomes and covariates

Our outcome measure of interest was the monthly rate of cancer diagnostic delays, aggregated up to the cancer alliance level. We followed the NHS 28 day faster diagnosis standard in defining rate of cancer diagnostic delay as the percentage of patient referrals for investigation of possible cancer that resulted in diagnostic resolution within 28 days following initial referral.[36] NHS providers are asked to ensure that no more than 25% of patients receive a delayed diagnosis according to this standard.[37]

We examined three primary outcomes in our analyses: diagnostic delay rates for all cancers, for the 'most commonly referred' cancers, and for all other 'less commonly referred' cancers. Our concern over potential spillover from the NHS-Galleri trial was particularly focused on types of suspected cancers detected by the Galleri test for which there was limited existing diagnostic capacity. In the absence of publicly available data on diagnostic capacity by type of suspected cancer, we relied on the volume of referrals that resulted in diagnostic resolution as an indicator of capacity. Referral volume was then used to stratify referrals into two groups according to whether they were for one of the 'most commonly referred' cancers or for any other 'less commonly referred' cancer.

Within NHS England, frequency of diagnostic referral varies greatly by type of suspected cancer (fig 2). Referrals for suspected breast, skin, and lower gastrointestinal cancers were much higher than for other cancers and together accounted for 58.7% of all cancer diagnostic referrals. We grouped these three types of cancer together as the 'most commonly referred' cancers. All other types of cancer were grouped together as 'less commonly referred' cancers. We included referrals for which specific cancer type was marked "missing or invalid" or "suspected cancer – non-specific symptoms" in our analysis of all cancers but not for other analyses. Secondary analyses examined diagnostic delays for each specific type of suspected cancer. We excluded referrals marked as 'Exhibited (non-cancer) breast symptoms - cancer not initially suspected' from all analyses.

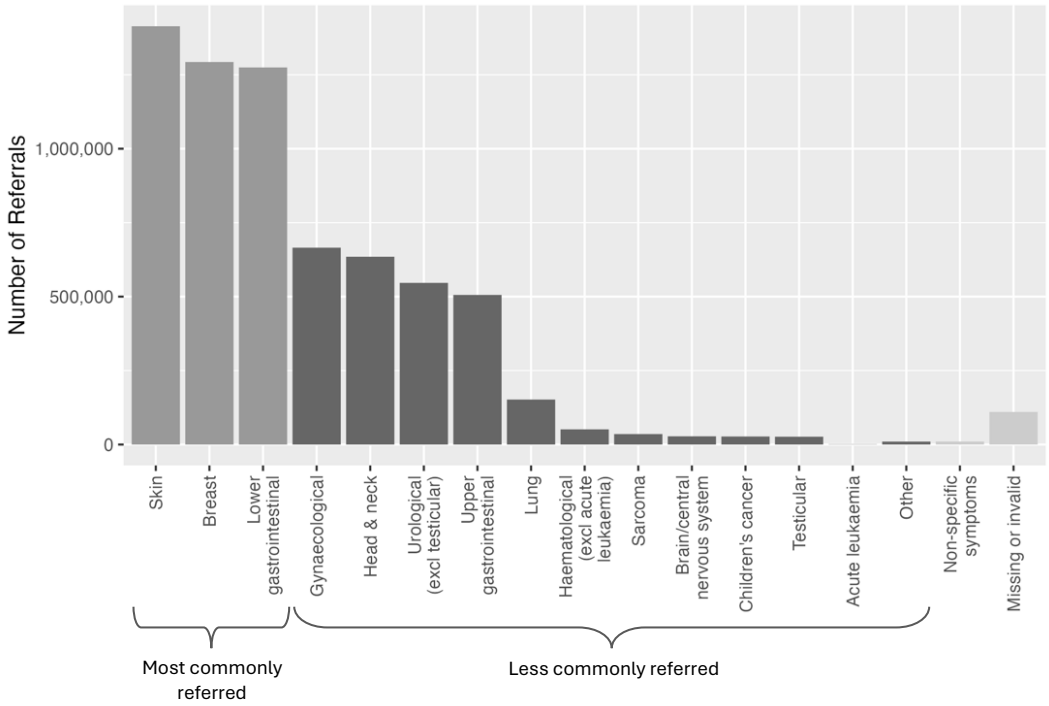

Figure 2. NHS England diagnostic referrals from April 2021-September 2023, by suspected cancer type.

We collected monthly data on the total number of full-time equivalent NHS health care staff working in each cancer alliance as well as the number of NHS health care staff absent from work due to sickness or self-isolation related to COVID-19 or any other health condition. These data were used to calculate the average percentage of the NHS workforce absent each month in each cancer alliance region, which we included as a covariate in our analysis.

Statistical analysis

We analyzed changes in the rate of cancer diagnostic delays in cancer alliance regions that participated in the trial compared to regions that did not participate using a difference-in-differences framework with an event study design. Multivariate linear regression was used to analyze the association of trial participation with changes in diagnostic delay rates for each month during the study period. These changes were estimated using a set of variables that captured the interaction between each monthly time period and participation in the NHS-Galleri trial. Coefficient estimates for these variables represent the mean difference in diagnostic delay rates between participating and non-participating regions for each month during the study period. Months were labeled numerically in relation to the start of the trial in October 2021, which was labeled month 0. The first month available in our dataset (April 2021, month -6) served as the reference period. These monthly regression models included month and cancer alliance fixed effects to control for time invariant differences between alliances as well as England-wide changes in diagnostic delays over time. Regressions were weighted by

the monthly volume of cancer diagnostic referrals in each cancer alliance region. All models included clustered standard errors at the cancer alliance level.

We used the same regression approach in a separate set of biannual models that estimated differences between participating and non-participating cancer alliances for each six-month time period in our dataset. In these models, the six-month period prior to the start of the trial (from April to September 2021, months -6 to -1) served as the reference period.

Bonferroni correction for multiple comparisons was used to assess significance of findings from our three primary analyses. Significance is not assessed, and p-values are not reported, for secondary analyses of specific cancers; point estimates and 95% confidence intervals are provided instead. Analyses were performed in R version 4.3.3 (R Foundation for Statistical Computing).

We used alternative model specifications to examine the robustness of our findings. These included analyses that excluded September to November 2021 as a phase-in period and excluded the staff absence covariate.

Finally, our quasi-experimental difference-in-differences event study design relies on the assumption that, in the absence of exposure to the NHS-Galleri trial, trends in outcomes would be similar for participating and non-participating regions. We visually examine potential deviations from this assumption by examining diagnostic delay rates in the six months preceding the start of the trial (fig 3). Unadjusted rates of diagnostic delays and their trajectories over time were somewhat similar during the pre-trial period, though rates in participating regions were slightly higher for all cancers and the most commonly referred cancers. Diagnostic delay rates and their month-to-month trajectory before the trial were most similar for less commonly referred cancers.

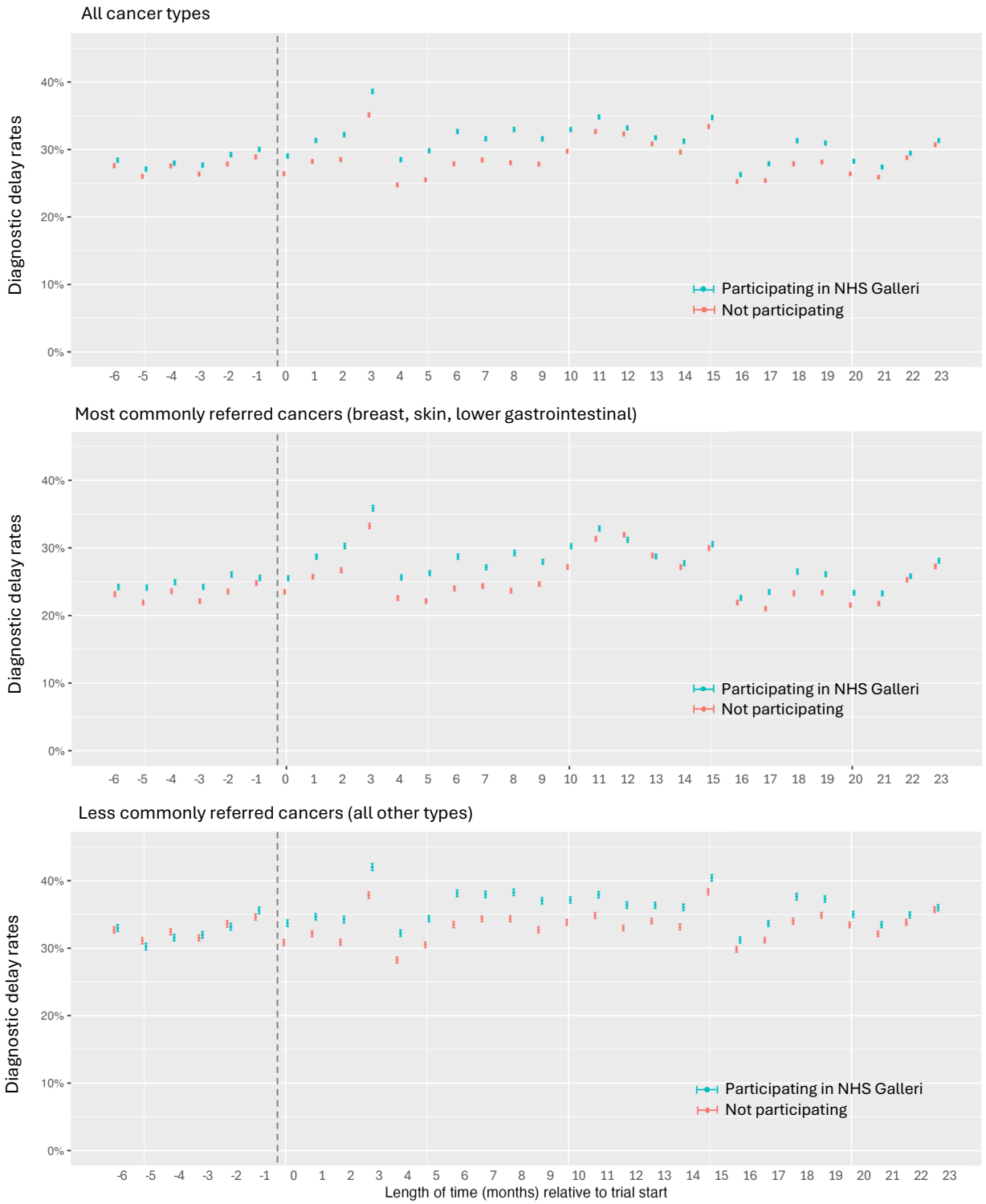

Figure 3. Unadjusted rates of diagnostic delays in regions participating in the NHS-Galleri trial and regions that are not participating. The vertical dashed line represents the start of the trial, with the area to the left depicting diagnostic delay rates in the six months (months -6 to -1) during the pre-trial baseline period. Months 0 to 23 depict the period of participating regions' exposure to the trial. Diagnostic delay rates are shown as point estimates and 95% confidence intervals.

Patient and public involvement

No patients or members of the public were involved in developing, conducting, or writing up the results of this study. We did not use individual-level data in our study, which focused instead on data on provider-level referral counts that NHS England has made publicly available for download.

Results

Study Sample and Baseline Characteristics

Prior to data cleaning and linking, the NHS England Provider-based Cancer Waiting Times dataset reported that 7,117,316 patient referrals for suspected cancer resulted in diagnostic resolution during the study period. Of these, we excluded 332,281 referrals labeled “Exhibited (non-cancer) breast symptoms - cancer not initially suspected” from our analysis, which left 6,785,035 referrals in the dataset. Of the 152 health organization providers included in the NHS England dataset, 133 were linked to cancer alliance regions using provider, integrated care board, and cancer alliance attribution files. We manually reviewed provider websites and locations to link the 19 remaining health providers to cancer alliances (appendix table 1). This manual review resulted in the exclusion of 14,085 referrals associated with 6 providers that provide services across multiple cancer alliance regions. Following data cleaning and linking, our study sample included 6,770,950 cancer referrals associated with 146 providers (appendix fig 1). 3,967,783 of these referrals were for more commonly referred cancers and 2,683,017 were for less commonly referred cancers. Another 120,150 referrals were marked as for “missing or invalid” cancer type or for “suspected cancer – non-specific symptoms.”

The eight cancer alliance regions participating in the NHS-Galleri trial, considered as a group, contained slightly fewer health organization providers and health care staff than the thirteen non-participating cancer alliances (table 1). This corresponds to the somewhat smaller patient population served by participating cancer alliances as compared to non-participating alliances. Population age structure, sex ratio, cancer prevalence, and diagnostic referral rates for cancer were very similar across the two groups.

Table 1. Characteristics of participating and non-participating cancer alliance regions prior to the start of the NHS-Galleri trial.

	Participating (n=8)	Not participating (n=13)
Total no. of providers	69	75
Total no. of staff	425,957	453,353
Total population (1000s)	25,648	30,902
Age, y		
<50	61.8%	62.8%
50-64	19.3%	19.0%
65-74	10.2%	9.7%
75-85	6.3%	6.0%

85+	2.5%	2.5%
Sex		
Female	50.6%	50.5%
Male	49.4%	49.5%
Cancer prevalence (per 100,000)	3,986	3,964
Diagnostic referrals (per 100,000)		
All types of suspected cancer	2,195	2,180
Most commonly referred types	1,272	1,263
Less commonly referred types	834	840
Missing/invalid type	89	77

Note: Number of providers and referral rates are for the six months prior to NHS Galleri start (April – September 2021) from NHS England, Monthly Provider Based Data and Summaries.[27] Two providers included in the data covering the full study period were not recorded as referring any patients during this pre-trial period. Number of staff associated with providers are averaged across the six months prior to the trial and are from NHS Workforce Statistics.[28] Population, age, sex, and cancer prevalence statistics reflect data from December 31, 2020 and are from NHS England, Cancer Prevalence Statistics England 2020.[38]

Diagnostic Delay Rates

In the first six months after the start of the NHS-Galleri trial, the unadjusted diagnostic delay rate for all cancer referrals in cancer alliance regions participating in the trial increased from 28.4% to 31.5%, compared to an increase in non-participating regions from 27.4% to 28.0% (table 2; adjusted difference 2.31 percentage points, 95% CI -1.06 to 5.68, p=0.169, relative increase of 8.1% from the pre-trial baseline 28.4). In this same period, the unadjusted diagnostic delay rate for referrals for the most commonly referred cancers increased from 24.9% to 28.6% in participating regions and from 23.2% to 25.5% in non-participating regions (adjusted difference 1.25 percentage points, 95% CI -3.61 to 6.11, p=0.597, relative increase of 5.8% from baseline 24.9). No statistically significant changes in diagnostic delay rates for all cancer referrals or the most commonly referred cancers were observed during the study period, in both the biannual (table 2) and monthly (fig 4) regression analyses.

The unadjusted diagnostic delay rate for referrals for less commonly referred cancers increased from 32.6% to 35.1% in participating regions in the first six months following trial start, while slightly decreasing in non-participating regions from 32.7% to 31.7% (table 2, adjusted difference 3.42 percentage points, 95% CI 1.17 to 5.67, p=0.005, significant at Bonferroni-adjusted threshold of $p < 0.016$, relative increase of 10.5% from baseline 32.6). The diagnostic delay rate for this subset of cancer referrals also remained differentially higher over the next six-month period (adjusted difference 3.68 percentage points, 95% CI 0.67 to 6.68, p=0.019, relative increase of 11.3% from baseline 32.6). These adjusted differences correspond to an estimated additional 8,133 referrals experiencing diagnostic delay in the first six-month period and an additional 9,163 referrals experiencing diagnostic delay in the second six-month period during the trial, for an estimated total of 17,296 additional referrals for less commonly referred

cancers experiencing diagnostic delay over the trial's first year in participating regions (appendix table 4).

The estimated effect of the trial on diagnostic delay rates for less commonly referred cancers diminished over the trial's second year and was not statistically significant in either the first half (table 2; adjusted difference 2.30, 95% CI -0.75 to 5.35, p=0.132, relative increase 7.0% from baseline 32.6) or the second half of that year (adjusted difference 1.51, 95% CI -1.78 to 4.80, p=0.350, relative increase 4.6% from baseline 32.6). A similar pattern, of significantly higher diagnostic delays in the first year of the Galleri trial but not in the second year, was observed in the monthly regression analysis (fig 4).

Our secondary analyses of specific types of suspected cancer found that diagnostic delay rates increased in participating regions during the first six months of the trial for upper gastrointestinal cancer (table 2; adjusted difference 4.87 percentage points [95% CI 1.10 to 8.64], relative increase 14.2% from baseline 34.4), urological malignancies excluding testicular cancer (adjusted difference 4.15 percentage points [95% CI 0.40 to 7.90], relative increase 10.1% from baseline 41.3), and lung cancer (adjusted difference 2.15 percentage points [95% CI 0.02 to 4.28], relative increase 8.7% from baseline 24.8). In the second six-month period of the trial, delay rates in participating regions were higher for upper gastrointestinal cancer (adjusted difference 5.78 percentage points [95% CI 1.53 to 10.04], relative increase 16.8% from baseline 34.4) and head and neck cancer (adjusted difference 3.62 percentage points [95% CI 0.87 to 6.38], relative increase 13.0% from baseline 27.8). Delay rates for other specific cancer types and over the second year of the trial were generally similar across participating and non-participating regions.

Our sensitivity analyses included regressions that excluded September to November 2021 (months -1 to 1) as a phase-in or washout period at the start of the trial (appendix table S1 and appendix fig S1) and excluded the staff absence covariate (appendix table S2 and appendix fig S2). Results of these models were similar to the results of our main analysis.

Table 2. Adjusted differential change in diagnostic delay rates for cancer alliance regions participating in the trial compared to non-participating regions.

	Adjusted difference, percentage points (95% CI)			
Suspected Cancer Type	**Months 0 to 5**	**Months 6 to 11**	**Months 12 to 17**	**Months 18 to 23**
All Cancers	2.31 (-1.06-5.68, p=0.169)	2.33 (-1.35-6.01, p=0.201)	-0.00 (-3.36-3.35, p=0.999)	0.29 (-3.76-4.33, p=0.884)
Most Commonly Referred Cancers	1.25 (-3.61-6.11, p=0.597)	1.43 (-3.22-6.08, p=0.529)	-1.57 (-5.84-2.70, p=0.451)	-0.55 (-6.19-5.09, p=0.840)
Suspected breast cancer	0.24 (-6.75-7.22)	-1.58 (-9.49-6.34)	-0.97 (-5.70-3.75)	0.64 (-4.68-5.95)
Suspected lower gastrointestinal cancer	4.47 (-0.73-9.67)	3.42 (-2.74-9.59)	-0.45 (-7.66-6.77)	1.17 (-8.48-10.83)
Suspected skin cancer	-1.29 (-7.42-4.84)	0.10 (-6.02-6.23)	-3.05 (-10.02-3.92)	-3.22 (-11.34-4.90)
Less Commonly Referred Cancers	3.42 (1.17-5.67, p=0.005)	3.68 (0.67-6.68, p=0.019)	2.30 (-0.75-5.35, p=0.132)	1.51 (-1.78-4.80, p=0.350)
Suspected acute leukaemia	-1.61 (-12.51-9.30)	-8.16 (-29.68-13.36)	13.16 (-17.33-43.65)	18.45 (-3.49-40.38)
Suspected brain/central nervous system tumours	0.94 (-7.59-9.48)	-0.67 (-8.14-6.79)	-0.60 (-10.88-9.67)	-5.57 (-17.37-6.22)
Suspected children's cancer	0.98 (-2.55-4.51)	0.33 (-3.98-4.64)	1.61 (-3.58-6.79)	-0.47 (-4.33-3.40)
Suspected gynaecological cancer	2.70 (-0.50-5.91)	2.02 (-3.57-7.62)	3.95 (-2.12-10.02)	2.08 (-5.10-9.26)
Suspected haematological malignancies (excluding acute leukaemia)	0.11 (-5.83-6.05)	-1.56 (-7.73-4.61)	-3.62 (-11.24-4.01)	-3.42 (-13.45-6.60)
Suspected head & neck cancer	2.00 (-0.66-4.67)	3.62 (0.87-6.38)	2.19 (-0.55-4.94)	1.33 (-1.53-4.18)
Suspected lung cancer	2.15 (0.02-4.28)	4.25 (-0.25-8.74)	0.63 (-3.94-5.19)	0.11 (-4.79-5.01)
Suspected other cancer	8.18 (-11.06-27.41)	3.25 (-22.71-29.22)	3.73 (-25.72-33.18)	6.33 (-18.01-30.67)
Suspected sarcoma	0.15 (-8.78-9.09)	4.54 (-3.80-12.88)	8.08 (-3.50-19.66)	6.84 (-4.11-17.79)
Suspected testicular cancer	2.33 (-3.85-8.51)	2.38 (-4.18-8.94)	0.16 (-6.40-6.71)	1.63 (-4.52-7.78)
Suspected upper gastrointestinal cancer	4.87 (1.10-8.64)	5.78 (1.53-10.04)	3.10 (-1.93-8.13)	0.60 (-6.39-7.59)
Suspected urological malignancies (excluding testicular)	4.15 (0.40-7.90)	2.87 (-2.19-7.93)	-0.07 (-5.36-5.21)	1.39 (-2.62-5.40)

Note: Adjusted differential changes are estimated for the four six-month periods following the start of the trial in month 0. Estimates are calculated at the level of six-month periods relative to the period prior to the trial start (April 2021 to September 2021, months -6 to -1) using the difference-in-differences model described in methods. The model includes time period and cancer alliance fixed effects as well as the average percentage of health care staff absent in each cancer alliance over each time period. Regressions are weighted by referral volume and include clustered standard errors at the cancer alliance level.

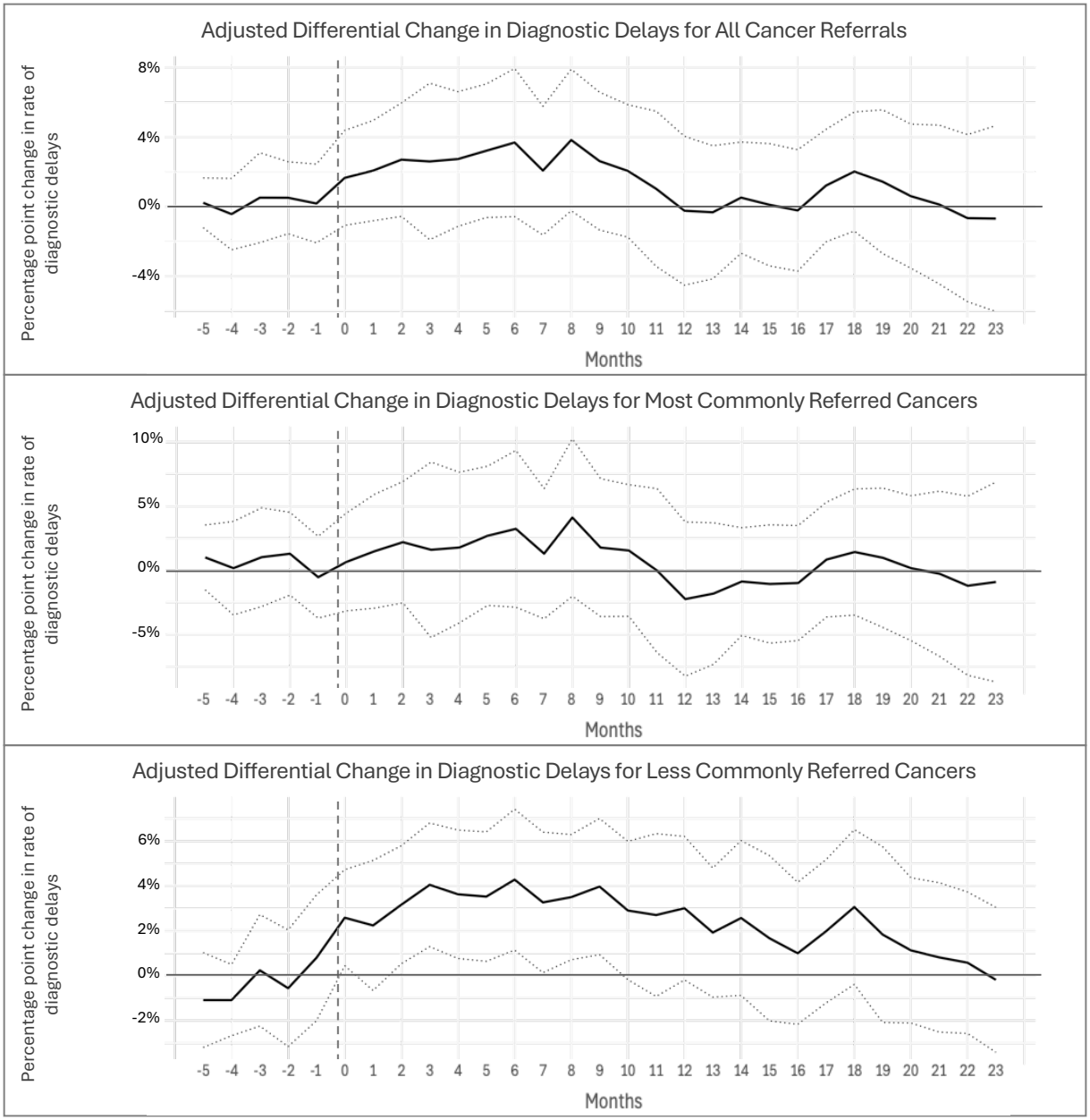

Figure 4. Adjusted differential change in diagnostic delay rates for cancer alliance regions participating in the trial compared to non-participating regions, by month. The vertical dashed line represents the start of the NHS-Galleri trial, with month 0 representing the first month (October 2021) following trial start. The dotted outer lines represent the upper and lower bounds of the 95% CI for each monthly estimate. Monthly estimates are calculated relative to the first month available in our dataset (April 2021, month -6) which served as the reference period. The model includes month and cancer alliance fixed effects as well as the average percentage of health care staff absent in each cancer alliance in each month. Regressions are weighted by referral volume and include clustered standard errors at the cancer alliance level.

Discussion

Principal Findings

We found no statistically significant differential changes in diagnostic delay rates between regions participating in the Galleri trial and non-participating regions for all cancers as well as the most commonly referred cancers. In contrast, we found evidence of a modest but statistically significant increase in diagnostic delays for less commonly referred cancers within participating regions during the first year of the Galleri trial (table 2). By the end of the second year of the trial, diagnostic delay rates in participating regions were nearly identical to rates in non-participating regions (fig 3 and fig 4).

Policy Implications

We do not know the exact mechanism behind the observed increase in diagnostic delay rates for less commonly referred cancers during the first year of the trial. A likely explanation, however, is that an increase in diagnostic referrals for intervention group patients who receive positive Galleri test results led to greater strain on diagnostic resources in regions participating in the trial. It is also likely that higher delay rates affected patients broadly across these regions, including non-participants and members of the control group, given that all received care via the same NHS referral pathways and relied on a shared set of diagnostic service providers.

Delays in diagnostic resolution can lead to stage migration (diagnosis at more advanced stage) and lower survival rates.[39] Missed opportunities for diagnosis are also a concern and could have a greater impact on cancer stage at time of diagnosis; in a recent study, researchers found that general practitioners in NHS England made fewer cancer diagnostic referrals when diagnostic delay rates were higher.[40] Delayed or missed diagnosis could negatively affect control group outcomes and thus lead to overestimation of the Galleri test's effect on the primary trial objective: a reduction in the incidence of later stage cancer at time of diagnosis.

Healthcare decision-makers need a measure of MCED test effectiveness that accurately reflects the extent to which adopting the test within a health system would benefit the patient population as compared to no adoption. Estimates from the NHS-Galleri trial may not accurately reflect this ideal measure of effectiveness, however. First, if spillover led to a system-wide increase in diagnostic delay rates, this would depress control group outcomes during the trial compared to outcomes from the ideal control group represented by the counterfactual case in which no one received MCED testing. This would lead to biased overestimation of intervention benefit during the trial. Second, the intervention group could also face delayed diagnosis due to spillover during the trial, but not to the full extent that they would if MCED testing was adopted widely across the health system and thus led to even greater capacity strain. This would improve intervention group outcomes during the trial compared to the counterfactual case in which all eligible patients receive MCED testing and further lead to biased overestimation of MCED benefit by trial investigators.

The trajectory of diagnostic delay rates we observed over the course of the NHS-Galleri trial further complicates the causal inference problem posed by spillover. Our analysis suggests that spillover effects appeared rapidly at the start of the trial but eventually diminished over time. This is consistent with an economic equilibrium model in which participating regions experienced an initial demand shock as the trial began but then gradually increased diagnostic services' supply to meet the new higher level of demand.[41] If diagnostic capacity expands in response to increased demand, over an acceptable period and at an acceptable cost, then this changes the decision-making calculus over whether to adopt MCED testing across the NHS. In this case, estimates of MCED test effectiveness calculated by comparing intervention and control group outcomes during the trial, even though affected by spillover, might still be an accurate reflection of expected benefits if the test was widely adopted, given that additional strain would be offset by additional capacity.

Comparison to other studies

Our study presents empirical evidence of a potential spillover effect on region-wide rates of diagnostic delays for less commonly referred cancers during the first year of the NHS-Galleri trial. The quasi-experimental difference-in-differences event study design that we use is similar to that used by other researchers to analyze the potential for spillover due to constraints on service delivery in trials of social welfare and public health interventions.[13,15] Our analysis complements existing literature examining capacity constraints and delays in cancer diagnosis prevalent in NHS England and other health systems.[11,42,43]

Spillover as a potential source of bias and risk to patients has not been addressed in other commentaries on this trial, which have largely focused on the costs of the Galleri test, its accuracy, and use of cancer stage rather than mortality as the primary trial endpoint.[44,45] The NHS-Galleri study protocol, which considers other sources of bias that could affect trial results such as contamination or differential attrition, does not address the potential for spillover due to constraints on care delivery.[1] This reflects a broader omission of this form of spillover in the clinical trials literature, including in guidance on trial design,[16,17] reporting,[18-20] ethics,[46,47] and risk of bias assessment.[21-23] Recent commentaries have raised the possibility that spillover due to constraints on care delivery could affect clinical trials of diverse interventions—including elective induction of labor, AI-based alert systems, and patient navigation for cancer diagnosis—but have not provided new empirical evidence to support this claim.[12,48] Our findings begin to address this gap and underscore the importance of considering such spillovers in the design of trials evaluating interventions, such as new MCED screening tests, that affect patient utilization of limited healthcare resources downstream of intervention delivery.

Limitations

Our study has several limitations. First, our unit of analysis is regional cancer alliances, and these were deliberately, rather than randomly, selected by trial investigators for participation in the trial which raises the possibility of selection bias. In addition, dynamic factors including both the COVID-19 pandemic and changing institutional factors such as nationwide adoption of the 28-day faster diagnosis standard for cancer in October 2021 at the start of the trial might have affected regions differently during the study period. As a result, we urge caution in interpreting our results given the possibility of residual confounding. Second, while our study was motivated by the concern that the Galleri test would increase intervention group utilization of diagnostic services and thus strain diagnostic capacity, the NHS-Galleri trial could also have increased diagnostic delay rates via other mechanisms, such as by diverting resources to directly support Galleri test delivery or by broadly increasing demand for cancer diagnostic services within participating regions as a result of local media coverage of the trial. Finally, we relied solely on publicly available data, aggregated up to the provider-level, that covered a limited timeframe and contained only a narrow set of information. Further analyses would benefit from access to individual-level data on patient characteristics, trial participation, and outcomes as well as data on NHS provider characteristics and diagnostic service availability over time.

Future research could help address these limitations. First, NHS-Galleri trial investigators and researchers conducting an accompanying economic evaluation could further investigate the potential presence of spillovers due to capacity constraints during the trial. These researchers will have access to additional data, including on the timing and volume of referrals prompted by the Galleri test as well as on control group outcomes and access to care, to support estimation of spillover effects. Disease-specific microsimulation models developed to support power calculations prior to the trial[49] could be adapted to examine the extent to which increased diagnostic delays and forgone referrals during the first two years of the trial could have affected intervention and control group outcomes. Additional analyses could use patient outcomes in non-participating regions as an external control. These and other post-hoc analysis could inform robustness checks examining the extent to which spillover due to constraints on care delivery might have affected estimates of Galleri test effectiveness derived from the trial.

Conclusion

The clinical utility of an MCED test depends not just on its ability to accurately detect early signs of cancer, but on its effectiveness in improving patient outcomes when integrated into the broader continuum of cancer screening, diagnosis, and treatment. This understanding motivates the NHS-Galleri trial and has led government regulators and research funders in the United States to call for similar pragmatic trials of MCED tests in real-world settings. But even as test effectiveness depends on the healthcare system in which it is delivered, the test's delivery also can change that system, including by increasing utilization of downstream diagnostic resources and thus reducing their

availability. As a result of such spillovers, MCED test effectiveness could differ according to the scale of testing and whether tests are delivered together with a corresponding increase in downstream diagnostic resources.

There is limited awareness of the potential for spillover to affect patient outcomes during pragmatic clinical trials. As researchers begin to analyze the results of the NHS-Galleri trial, they should carefully consider, and if possible correct for, the possibility that spillover has affected their results. Future trials evaluating MCED tests, such as the Vanguard study in the United States,[50] should be designed from the outset to avoid or detect spillover. The possibility that spillover will negatively affect control group and non-participant outcomes should be assessed by health providers and institutional review boards as they weigh risks associated with a proposed trial. Trials of other interventions that affect utilization of scarce healthcare resources should also address the potential for spillover to affect patient outcomes and trial results.

----------Summary Box----------

What is already known on this topic

- The NHS-Galleri trial randomized over 140,000 adults to receive usual cancer screening or usual screening plus a new multi-cancer early detection screening test in 8 out of 21 cancer alliance regions in England.

- During the trial, increased utilization of scarce cancer diagnostic services by the intervention group could have led to delayed or reduced availability of services for members of the control group and patients not enrolled in the trial.

- The potential for this form of spillover to bias trial results and affect patient access to care is not addressed in the medical literature, guidance on clinical trial methods, or the NHS-Galleri trial protocol.

What this study adds

- Participation in the trial led to an increase in region-wide diagnostic delay rates for less commonly referred cancers (all except breast, skin, and lower gastrointestinal cancer) in the first year but no significant change in delays for more commonly referred cancers or for cancer overall.

- Increased diagnostic delays could contribute to stage shift for some control group patients and lead to overestimation of the Galleri test's effect on early detection of less commonly referred cancers.

- Future trials of interventions that affect utilization of scarce healthcare resources should address the potential for spillover due to constraints on care delivery to affect patient outcomes and trial results.

---

1. Neal RD, Johnson P, Clarke CA, et al. Cell-Free DNA-Based Multi-Cancer Early Detection Test in an Asymptomatic Screening Population (NHS-Galleri): Design of a Pragmatic, Prospective Randomised Controlled Trial. *Cancers (Basel)*. Oct 1 2022;14(19)doi:10.3390/cancers14194818
2. Ferrante JM, Chen PH, Kim S. The effect of patient navigation on time to diagnosis, anxiety, and satisfaction in urban minority women with abnormal mammograms: a randomized controlled trial. *J Urban Health*. Jan 2008;85(1):114-24. doi:10.1007/s11524-007-9228-9
3. Raich PC, Whitley EM, Thorland W, Valverde P, Fairclough D. Patient navigation improves cancer diagnostic resolution: an individually randomized clinical trial in an underserved population. *Cancer Epidemiol Biomarkers Prev*. Oct 2012;21(10):1629-38. doi:10.1158/1055-9965.Epi-12-0513
4. Murphy DR, Wu L, Thomas EJ, Forjuoh SN, Meyer AND, Singh H. Electronic Trigger-Based Intervention to Reduce Delays in Diagnostic Evaluation for Cancer: A Cluster Randomized Controlled Trial. *Journal of Clinical Oncology*. 2015/11/01 2015;33(31):3560-3567. doi:10.1200/JCO.2015.61.1301
5. Atlas SJ, Tosteson ANA, Wright A, et al. A Multilevel Primary Care Intervention to Improve Follow-Up of Overdue Abnormal Cancer Screening Test Results: A Cluster Randomized Clinical Trial. *JAMA*. 2023;330(14):1348-1358. doi:10.1001/jama.2023.18755
6. Wolf AMD, Fontham ETH, Church TR, et al. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. *CA: A Cancer Journal for Clinicians*. 2018/07/01 2018;68(4):250-281. doi:https://doi.org/10.3322/caac.21457
7. Cardoso F, Costa A, Norton L, et al. ESO-ESMO 2nd international consensus guidelines for advanced breast cancer (ABC2). *The Breast*. 2014/10/01/ 2014;23(5):489-502. doi:https://doi.org/10.1016/j.breast.2014.08.009
8. Department of Veterans Affairs VHA. *VHA Directive 1015: Colorectal Cancer Screening*. 2020. April 3, 2020. Accessed February 26, 2024. https://www.va.gov/vhapublications/ViewPublication.asp?pub_ID=8815
9. Richards M, Maskell G, Halliday K, Allen M. Diagnostics: a major priority for the NHS. *Future Healthc J*. Jul 2022;9(2):133-137. doi:10.7861/fhj.2022-0052
10. Weller D, Coleman D, Robertson R, et al. The UK colorectal cancer screening pilot: results of the second round of screening in England. *Br J Cancer*. Dec 17 2007;97(12):1601-5. doi:10.1038/sj.bjc.6604089
11. Hubers J, Sonnenberg A, Gopal D, Weiss J, Holobyn T, Soni A. Trends in Wait Time for Colorectal Cancer Screening and Diagnosis 2013-2016. *Clin Transl Gastroenterol*. Jan 2020;11(1):e00113. doi:10.14309/ctg.0000000000000113
12. Mann S. Negative Spillover: A Potential Source of Bias in Pragmatic Clinical Trials. *arXiv*. March 26, 2024 2024;doi:arXiv:2309.10978
13. Giannella E, Homonoff T, Rino G, Somerville J. Administrative Burden and Procedural Denials: Experimental Evidence from SNAP. *National Bureau of Economic Research Working Paper Series*. 2023;No. 31239doi:10.3386/w31239
14. Mody A, Sikazwe I, Czaicki NL, et al. Estimating the real-world effects of expanding antiretroviral treatment eligibility: Evidence from a regression discontinuity analysis in Zambia. *PLoS Med*. Jun 2018;15(6):e1002574. doi:10.1371/journal.pmed.1002574

15. Okeke EN, Abubakar IS. Healthcare at the Beginning of Life and Child Survival: Evidence from a Cash Transfer Experiment in Nigeria. *J Dev Econ*. Mar 2020;143doi:10.1016/j.jdeveco.2019.102426
16. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ : British Medical Journal*. 2015;350:h2147. doi:10.1136/bmj.h2147
17. Heagerty P. *Experimental Designs and Randomization Schemes*. 2023. *Rethinking Clinical Trials: A Living Textbook of Pragmatic Clinical Trials*. Accessed November 5, 2024. https://rethinkingclinicaltrials.org/chapters/design/experimental-designs-and-randomization-schemes/experimental-designs-introduction/
18. Schulz KF, Altman DG, Moher D, the CG. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Medicine*. 2010/03/24 2010;8(1):18. doi:10.1186/1741-7015-8-18
19. Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P. CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments: A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. *Ann Intern Med*. Jul 4 2017;167(1):40-47. doi:10.7326/m17-0046
20. Daniela RJ, Liliane Z, Susan G, et al. CONSORT Harms 2022 statement, explanation, and elaboration: updated guideline for the reporting of harms in randomised trials. *BMJ*. 2023;381:e073725. doi:10.1136/bmj-2022-073725
21. Higgins J, Savović J, Page M, Elbers R, Sterne J. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins J, Thomas J, Chandler J, et al, eds. *Cochrane Handbook for Systematic Reviews of Interventions version 63 (updated February 2022)*. Cochrane; 2022.
22. U.S. Preventive Services Task Force. Appendix VI. Criteria for Assessing Internal Validity of Individual Studies. Accessed August 22, 2023, https://www.uspreventiveservicestaskforce.org/uspstf/about-uspstf/methods-and-processes/procedure-manual/procedure-manual-appendix-vi-criteria-assessing-internal-validity-individual-studies
23. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *Journal of Clinical Epidemiology*. 2011;64(4):407-415. doi:10.1016/j.jclinepi.2010.07.017
24. Nicholson BD, Oke J, Virdee PS, et al. Multi-cancer early detection test in symptomatic patients referred for cancer investigation in England and Wales (SYMPLIFY): a large-scale, observational cohort study. *The Lancet Oncology*. 2023;24(7):733-743. doi:10.1016/S1470-2045(23)00277-2
25. National Health Service. *The NHS Cancer Plan*. 2000. September 2000. Accessed May 24, 2024. https://webarchive.nationalarchives.gov.uk/ukgwa/20130222181549mp_/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4014513.pdf
26. NHS England. *Cancer Waiting Times: National Time Series - Provider based, October 2009 to September 2023*. 2023. Accessed May 24, 2024. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2024/01/Cancer-Waiting-Times-National-Time-Series-Oct-2009-Sep-2023-with-Revisions.xlsx

27. NHS England. *Monthly Provider Based Data and Summaries*. Accessed May 16, 2024. https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/monthly-prov-cwt/
28. NHS England. *NHS Workforce Statistics*. 2024. April 25, 2024. Accessed May 20, 2024. https://digital.nhs.uk/data-and-information/publications/statistical/nhs-workforce-statistics/january-2024
29. NHS England. *COVID-19 Hospital Activity*. Accessed May 20, 2024. https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-hospital-activity/
30. NHS England. Attribution from A&E Providers to ICBs. Accessed May 16, 2024, 2024. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2023/02/Trust-ICB-Attribution-File.xls
31. NHS National Cancer Patient Experience Survey. 2022 ICB to Alliance mapping document. Accessed May 16, 2024, 2024. https://www.ncpes.co.uk/wp-content/uploads/2023/07/Sub-ICB-to-Alliance-mapping-2022_v1.4.csv
32. NHS launches world first trial for new cancer test. September 13, 2021, 2021. https://www.england.nhs.uk/2021/09/nhs-launches-world-first-trial-for-new-cancer-test/
33. Kumar H. Accelerated Access Collaborative: GRAIL-NHS England Partnership. Accessed May 20, 2024, 2024. https://www.england.nhs.uk/aac/wp-content/uploads/sites/50/2022/03/Galleri-test-trial-launch.pdf
34. GRAIL and National Health Service (NHS) England Complete Enrollment of 140,000 Participants in Largest Study of Multi-Cancer Early Detection Test. July 18, 2022, https://grail.com/press-releases/grail-and-national-health-service-nhs-england-complete-enrollment-of-140000-participants-in-largest-study-of-multi-cancer-early-detection-test/
35. Grail. NHS Galleri Trial: Participant Information Sheet. Accessed July 13, 2024, 2024. https://www.nhs-galleri.org/assets/pdf-resources/Participant_information_sheet__EN.pdf
36. NHS. *NHS Cancer Programme: Faster Diagnosis Framework*. 2022. November 28, 2022. Accessed July 13, 2024. https://www.england.nhs.uk/wp-content/uploads/2019/07/B1332-NHS-Cancer-Programme-Faster-Diagnosis-Framework-v5.pdf
37. NHS England. *National Cancer Waiting Times Monitoring Dataset Guidance, version 12.0*. 2023. Accessed July 13, 2024. https://www.england.nhs.uk/wp-content/uploads/2023/08/PRN00654-national-cancer-waiting-times-monitoring-dataset-guidance-v12.pdf
38. NHS England. *Cancer Prevalence Statistics England, 2020*. 2023. Accessed July 22, 2024. https://digital.nhs.uk/binaries/content/assets/website-assets/national-disease-registration-service/ncras/data-outputs/cancer-prevalence-statistics-england-2020_download.xlsx
39. Maringe C, Spicer J, Morris M, et al. The impact of the COVID-19 pandemic on cancer deaths due to delays in diagnosis in England, UK: a national, population-based, modelling study. *Lancet Oncol*. Aug 2020;21(8):1023-1034. doi:10.1016/s1470-2045(20)30388-0
40. Hayes H, Meacock R, Stokes J, Sutton M. How do family doctors respond to reduced waiting times for cancer diagnosis in secondary care? *The European Journal of Health Economics*. 2024/07/01 2024;25(5):813-828. doi:10.1007/s10198-023-01626-2

41. Duflo E, Glennerster R, Kremer M. Chapter 61 Using Randomization in Development Economics Research: A Toolkit. In: Schultz TP, Strauss JA, eds. *Handbook of Development Economics*. Elsevier; 2007:3895-3962.
42. Sud A, Torr B, Jones ME, et al. Effect of delays in the 2-week-wait cancer referral pathway during the COVID-19 pandemic on cancer survival in the UK: a modelling study. *The Lancet Oncology*. 2020/08/01/ 2020;21(8):1035-1044. doi:10.1016/S1470-2045(20)30392-2
43. Aggarwal A, Simcock R, Price P, et al. NHS cancer services and systems: ten pressure points a UK cancer control plan needs to address. *The Lancet Oncology*. 2024;25(8):e363-e373. doi:10.1016/S1470-2045(24)00345-0
44. Turnbull C, Wald N, Sullivan R, et al. GRAIL-Galleri: why the special treatment? *The Lancet*. 2024;403(10425):431-432. doi:10.1016/S0140-6736(23)02830-1
45. McCartney M, Cohen D. Galleri promises to detect multiple cancers—but new evidence casts doubt on this much hyped blood test. *BMJ*. 2024;386:q1706. doi:10.1136/bmj.q1706
46. World Medical Association. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Participants. *JAMA*. 2024;doi:10.1001/jama.2024.21972
47. Califf RM, Sugarman J. Exploring the ethical and regulatory issues in pragmatic clinical trials. *Clin Trials*. Oct 2015;12(5):436-41. doi:10.1177/1740774515598334
48. Mann S, James KF. Elective Induction of Labor May Have Negative Effects at the Hospital Level. *Journal of Obstetric, Gynecologic & Neonatal Nursing*. October 10, 2024 doi:10.1016/j.jogn.2024.09.003
49. Dai JY, Zhang J, Braun JV, Simon N, Hubbell E, Zhang N. Clinical performance and utility: A microsimulation model to inform the design of screening trials for a multi-cancer early detection test. *J Med Screen*. Sep 2024;31(3):140-149. doi:10.1177/09691413241228041
50. Minasian L., Castle P. E. Cancer Screening Research Network / Multi-Cancer Early Detection Evaluation. National Cancer Institute. Accessed October 9, 2024, 2024. https://prevention.cancer.gov/sites/default/files/2023-02/Cancer-Screening-Research-Network-MCED-20220615.pdf