

Modeled Reductions in Late-stage Cancer with a Multi-Cancer Early Detection Test

Earl Hubbell¹, Christina A. Clarke¹, Alexander M. Aravanis^{1,2}, and Christine D. Berg³

ABSTRACT

Background: Cancer is the second leading cause of death globally, with many cases detected at a late stage when prognosis is poor. New technologies enabling multi-cancer early detection (MCED) may make “universal cancer screening” possible. We extend single-cancer models to understand the potential public health effects of adding a MCED test to usual care.

Methods: We obtained data on stage-specific incidence and survival of all invasive cancers diagnosed in persons aged 50–79 between 2006 and 2015 from the US Surveillance, Epidemiology, and End Results (SEER) program, and combined this with published performance of a MCED test in a state transition model (interception model) to predict diagnostic yield, stage shift, and potential mortality reductions. We model long-term (incident) performance, accounting for constraints on detection due to repeated screening.

Results: The MCED test could intercept 485 cancers per year per 100,000 persons, reducing late-stage (III+IV) incidence by 78% in those intercepted. Accounting for lead time, this could reduce 5-year cancer mortality by 39% in those intercepted, resulting in an absolute reduction of 104 deaths per 100,000, or 26% of all cancer-related deaths. Findings are robust across tumor growth scenarios.

Conclusions: Evaluating the impact of a MCED test that affects multiple cancer types simultaneously requires modeling across all cancer incidence. Assuming MCED test metrics hold in a clinical setting, the aggregate potential to improve public health is significant.

Impact: Modeling performance of a MCED test in a representative population suggests that it could substantially reduce overall cancer mortality if added to usual care.

Introduction

Cancer is the second leading cause of death globally (1). Programs to reduce these deaths through widespread screening have been implemented worldwide. However, no single cancer type is responsible for the majority of deaths (1), and existing programs address a handful of single cancer types individually, at a cost of a cumulative increase in false positives per screening program added (2). Development of new technologies has enabled the simultaneous detection of multiple cancers with high specificity, making “universal cancer screening” (3) a possibility. Blood-based tests leveraging methylation biology or mutations in cell-free DNA (cfDNA; refs. 4, 5), and/or proteomics (6), have demonstrated the ability to detect multiple cancer types, including many without current screening paradigms (e.g., lung cancer in persons not eligible according to current screening guidelines). Altogether, this evidence supports the potential for novel, multi-cancer early detection (MCED) tests that can be applied to population-wide cancer screening (7).

Nearly all prior modeling efforts have focused on stage shift and mortality outcomes of single-cancer screening programs (8), leaving a gap in models addressing the potential impact of MCED in a population. Our recent assessment of data from the Surveillance, Epidemiology, and End

Results (SEER) program showed that stage IV cancers are responsible for 48% of cancer-related deaths within 5 years of diagnosis but only 18% of cancer diagnoses. Reducing this hazard to that of earlier stages (e.g., by early detection) could reduce 15% of total cancer mortality among persons aged 50–79 (9). These simple calculations underscored the burden of stage IV disease, but do not capture more sophisticated aspects of a potential multi-cancer screening strategy, including the published performance of blood tests across cancer types and stages (5), as well as the preclinical sojourn time during which otherwise asymptomatic cancers might be detectable by such testing. Applying insights obtained from prior models of single-cancer screening paradigms (Supplementary Methods and Materials), we constructed a simple, efficient model for predicting diagnostic yield and stage of diagnosis due to screening with a MCED test, and the resulting potential effect on mortality across all cancer incidence (supplemental code/data at https://github.com/grailbio-publications/Hubbell_CEBP_InterceptionModel).

Materials and Methods

Overview

This interception model, a state transition model (Supplementary Methods and Materials), estimates the results of a screening program using a MCED test (5) when added to usual care. Starting with the observed total incidence and stage at diagnosis of cancers, we extend these data with assumptions (Supplementary Methods and Materials) about the duration of cancer in each stage before diagnosis by usual care. Given published data (5) detailing cancer cases that are detectable by a cfDNA-based multi-cancer detection assay, we model (i) the opportunity to intercept cancers before diagnosis by usual care, resulting in earlier stage at diagnosis, and (ii) the potential mortality benefit expected from that earlier stage at diagnosis.

Population cancer incidence and death data

We obtained from the National Cancer Institute’s SEER program crude incidence and 5-year cancer-specific survival (Supplementary

¹GRAIL, Inc., Menlo Park, California. ²llumina, Inc., San Diego, California. ³National Cancer Institute, Bethesda, California.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

A.M. Aravanis was associated with GRAIL, Inc. and Illumina, Inc. at the time of this study.

Corresponding Author: Earl Hubbell, GRAIL, Inc., 1525 O’Brien Dr., Menlo Park, CA 94025. Phone: 650-542-0372; E-mail: ehubbell@grailbio.com

Cancer Epidemiol Biomarkers Prev 2021;30:460–8

doi: 10.1158/1055-9965.EPI-20-1134

©2020 American Association for Cancer Research.

Fig. S1) for all persons aged 50–79 when diagnosed with malignant primary cancer in one of 18 regions from 2006 to 2015 and followed for vital status through December 31, 2017 (10). We chose this age range to minimize competing risks of non-cancer-related deaths among persons aged 80 and older and to overlap with the target population of existing cancer screening recommendations [e.g., United States Preventive Services Task Force (USPSTF) breast, colorectal, and lung]. Additional SEER data specifications are detailed in Supplementary Methods and Materials and Supplementary Table S1, including definitions for ICD-O-3 site and histologic groupings reflecting Liu and colleagues' (5) MCED performance and stage at diagnosis (categorized as I, II, III, IV, and unknown). SEER*Stat software (version 8.3.6) was used for all SEER calculations.

Herein we define *usual care* to represent real-world cancer diagnostic processes (e.g., screening, incidental detection, symptomatic workup) as captured by SEER. This is a snapshot of current practice, including imperfect adherence and limited access to health care. The scenarios assessed herein involve supplementing usual care with a hypothetical screening program involving a MCED test.

Interpretation of sensitivity and specificity of a multi-cancer early detection test

Reported sensitivities from a MCED test analyzed individuals found by usual care to have cancer (5), reflecting the population fraction at each stage who produced signals detectable by the test. Given that cancer is a progressive disease, we assume that individual cancer cases progress to increasing detection (i.e., when cancer “signal” exceeds the detectable threshold) by stage. Therefore, individuals not detectable at a given stage (or screening event) would not be detectable at any earlier stage (or screening event); conversely, individuals detectable during any given stage, would also be detectable at all later stages.

We apply this assumption to the stage-specific sensitivity estimates for each cancer type from a cfDNA-based MCED test (Supplementary Methods and Materials and Supplementary Fig. S2; ref. 5). This assumption requires that sensitivity estimates are nondecreasing by increasing stage for each cancer type; estimates were adjusted by isotonic regression weighted by the number of observations available per stage. For uncommon cancer types that were not categorized in specific sensitivity groups (Supplementary Methods and Materials) or cancer types that are unstageable/not staged conventionally (e.g., leukemia, myeloma), where stage shift cannot be modeled, we imputed zero sensitivity.

Sensitivity in preclinical stages is based on that observed in the case-control study. We then reconstruct the detectable fraction (the fraction reaching a signal level above threshold) per stage per cancer type to match the observed sensitivity values, implying that individuals not detected at a given stage remain undetectable until their stage progresses (Supplementary Methods and Materials and Supplementary Fig. S3). This reconstruction is therefore a conservative under-estimate of the potential stage at which an individual becomes detectable.

State-flow model and algorithm operation

We devised a simple state-transition model to predict the expected diagnostic yield and stage shift arising from intercepting cancers before discovery by usual care (the “interception model”). We assume the same cancer can only be found once, cancers that are not detectable at some time cannot be found by screening at that time, and in the absence of overdiagnosis, screening only rearranges cancers through time and does not increase average incidence (Supplementary Methods and Materials). Combining these assumptions results in an

algorithm to estimate rates that starts with the incidence under usual care for each cancer type and stage, and then divides that incidence among earlier stages based on the properties of the MCED test, the duration of cancer, and the screening schedule. This algorithm is implemented using the R programming language and the tidyverse (11, 12).

We note that any cancer found in a given year must have been missed by previous screening events and usual care. This splits into three situations (13): (i) invasive cancer had not yet developed at any previous screening; (ii) invasive cancer had developed, but was not detectable at the time of any screening events; (iii) invasive cancer was detectable, but the individual did not have a screening event. The second and third situations apply to every stage before clinical diagnosis. Given the relative rate of such situations, we compute the stage at diagnosis resulting from implementing screening with a MCED test, yielding the incidence at each stage both for those cancers intercepted and those remaining to be found by usual care.

The *prevalence round* of a screening schedule is the round at which an individual is first screened by a test, and *incidence rounds* of a screening schedule are rounds at which individuals have been previously screened. The long-term performance of a screening program is dominated by the performance in the incidence rounds (13). We outline the operation of the algorithm for incidence rounds.

The states and transitions in the model are illustrated in Fig. 1A. The operation of the algorithm is illustrated in Fig. 1B–D. Results are computed independently for each cancer type, and aggregated to summary values as shown in Fig. 1B–D.

Usual care results in the SEER distribution of cancers by stage (lower gray boxes; Fig. 1B). No diagnoses are made using the MCED test and the upper pink boxes contain zero incidence. All diagnosed cancers at a given stage arrive from either a detectable state or a non-detectable state: the ratio here is determined by the MCED test properties by cancer stage and type. Working backward, some detectable cancers at a late stage must not have been detectable at an earlier stage. This illustrates the backward pass of the algorithm in which the fraction of cancers that are detectable by the MCED test at each stage prior to detection by usual care is estimated.

An idealized screening schedule (14), wherein all cancers detectable by screening have a MCED screening event during the earliest stage they are detectable (Fig. 1C), estimates the maximum benefit from screening. From the distribution in Fig. 1B, the algorithm proceeds forward: cancers detectable at each successive stage are intercepted and removed from outgoing flows to later stages. This prevents double-counting of incidence. Because cancers detectable in previous stages have been intercepted, screening by a MCED test finds the newly detectable cancers in each stage. As not all individuals become detectable by the MCED test before being found by usual care, some incidence at each stage is found by usual care. This illustrates the forward pass of the algorithm to compute the yield and stage distribution.

In practice, screening events miss some cancers, and the algorithm must be supplied with the rate of cancer progression (Fig. 1D). Starting with the distribution from Fig. 1B, the algorithm works forwards with positive rates of progression. This computes the stage distribution and yield resulting from a screening schedule interacting with a scenario of tumor growth rates.

These three examples illustrate the fundamental principle: that starting from the incidence at the original stage of clinical presentation, the algorithm back-calculates to find the stages at which a proportion of that incidence become detectable, and then working forward, the

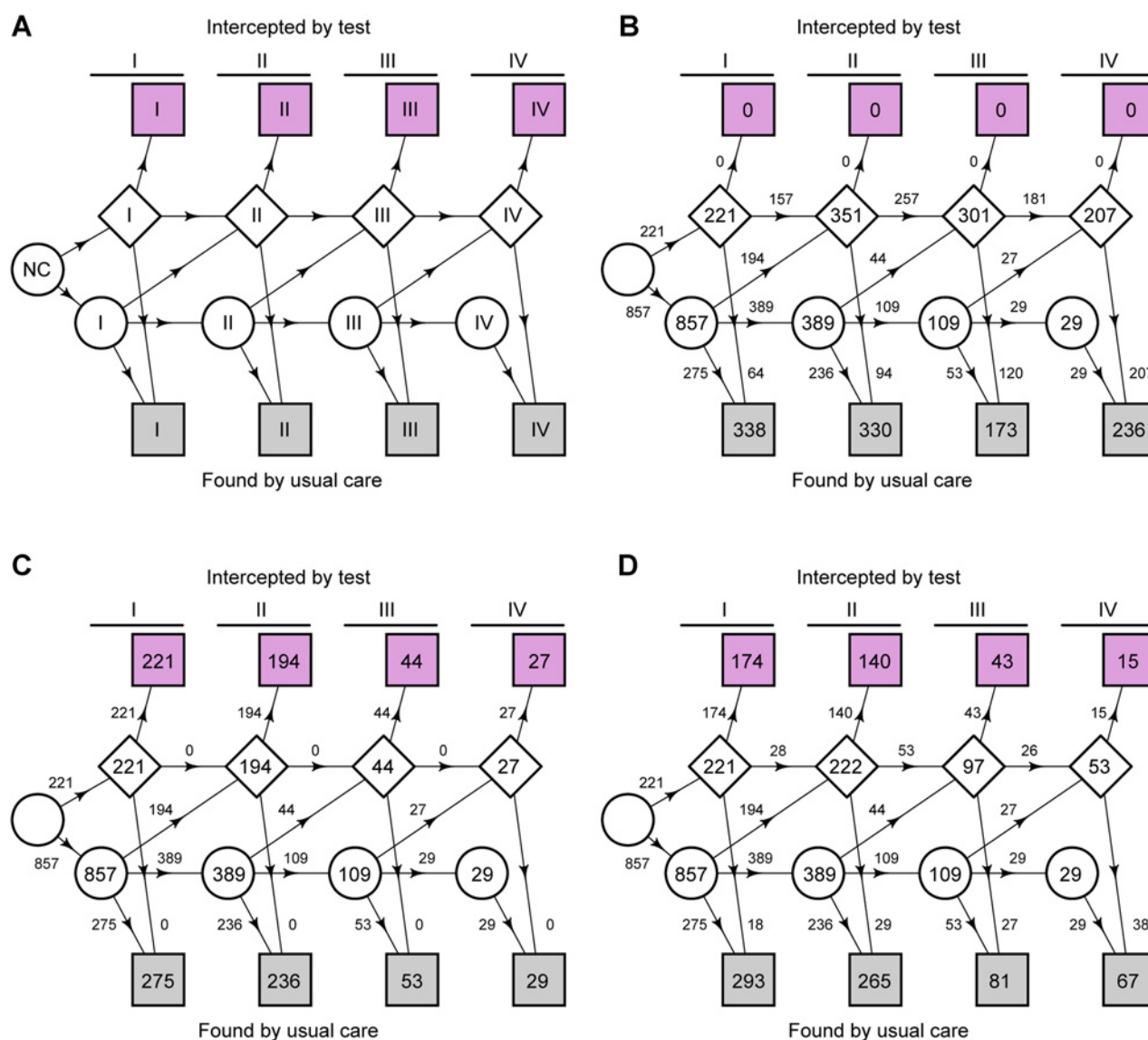


Figure 1. Interception model scenarios. **Figure 1** contains an interception model schematic (**A**) wherein shapes represent states of cancer cases, and arrows between states represent a transition from one state to another. The circle labeled NC (noncancer) represents the precancerous state, from which incidence flows to a cancerous but prediagnosis state (transparent shapes), which terminate at diagnosis (colored boxes). Transparent shapes infer the cancer was either detectable by MCED (diamonds) or not (circles) ahead of the diagnosed state; cancer is diagnosed either by MCED test (interception before usual care; pink boxes) or via usual care (gray boxes). Cancer being a progressive disease is reflected by arrows always pointing to later stages; detectability being progressive (e.g., cancers becoming more detectable by stage) is reflected by arrows going from circles to diamonds. **B–D** illustrate interception model inferences under multiple scenarios: No MCED test screening, Idealized screening schedule, and Fast tumor growth, respectively. Stage shift and incidence are computed for each scenario covering one year of incidence. **B**, All incidence diagnosed by usual care: this results in the current SEER distribution, illustrating potentially detectable cancers by stage. **C**, An idealized screening schedule results in all cancers detectable by MCED screening before usual care found at the earliest possible stage. **D**, Annual screening schedule with a fast tumor growth scenario causes cancer progression to later stages due to cancer evolving faster than the screening interval.

proportion of the detectable incidence that is intercepted due to screening is calculated, and the proportion that is missed by screening progresses in stage and potentially to detection by usual care.

This algorithm computes the yield and stage distribution of cancers that would have been diagnosed by usual care in the span of one year. Because incidence rounds are all similar in yield and stage shift (stationarity, Supplementary Methods and Materials; refs. 15–18), the rates computed here are exactly equal to the rates of those cancers

diagnosed by screening and usual care within the span of a given year. This reflects that incidence of cancer removed from a population by previous screening is replaced by incidence gained from the future by screening.

The expected rate and stage distribution for individuals undergoing their first screening (prevalence round) may be computed (Supplementary Methods and Materials and Supplementary Table S2) but fluctuate based on tumor growth rate scenarios. Performance in

incidence rounds is less affected by tumor growth rate, and we report in the main text incidence results based on their robustness.

Robustness to uncertainty in preclinical sojourn (dwell) time for cancers

The ability of a test to detect cancers is only one component of the performance of a screening program. The threshold sensitivity of the assay determines population fractions of detectable individuals; however, this is further reduced by the *schedule sensitivity* of a given screening schedule. Even an assay with perfect sensitivity cannot intercept a cancer that progresses to diagnosis by usual care between screening events. This affects both the diagnostic yield of cancers as well as the stage at which they are found.

Reliable natural history parameters for cfDNA-detectable cancer cases are not yet available; therefore, we performed a robustness analysis considering plausible durations per cancer type and stage. Preclinical sojourn time reflects the total time before diagnosis by usual care of an invasive cancer and we divide this duration into a dwell time per stage for our model. Three scenarios were chosen varying durations between slow (3–7 years in stage I, depending on cancer type), fast (2–4 years), to aggressively fast (1–2 years). Following multiple Cancer Intervention and Surveillance Modeling Network (CISNET) models (19–21), within each cancer type and stage, we approximated dwell time with an exponential distribution. The stage at which usual care would diagnose a cancer is subject to a competing risk (discovery by usual care) and assigned a shorter dwell time distribution. Later stages were assumed to span less time than earlier stages, reflecting cancer as a progressive disease (Supplementary Methods and Materials and Supplementary Table S3).

Dwell times interact with a periodic screening schedule to produce the rate of progression of cancer missed by the screening. The algorithm requires an estimate of the *slip rate*: the probability per detectable individual that the tumor progresses before a screening event occurs. This is precisely the probability that no screening events happen while an individual is detectable at a given stage (Fig. 2). For the simplest case where slip rates only depend on stage, this quantity can be computed by integration or stochastic approximation given the distribution of dwell times (Supplementary Methods and Methods).

An idealized screening schedule (Maximum Interception Scenario; MIS; ref. 14) estimated the maximum yield, stage shift, and mortality benefit from screening with a MCED test by setting the slip rate to zero. This emulates a screening interval sufficiently fast relative to the shortest cancer duration as to intercept cancers in the earliest detectable stage (Supplementary Methods and Materials).

Modeled mortality benefit from stage shift

This model computes stage shifts due to interception by a MCED test, and we estimate improvements in mortality conditional on the stage at interception. Following multiple models (13, 22–24), we divide survival into the lead time until the date of diagnosis by usual care, during which individuals do not die from their cancer, and survival after that time during which cancer-specific mortality can occur. This avoids shortened lifespan due to early detection (25). Differences in mortality then accumulate after the date of original diagnosis. We report the difference in 5-year mortality statistics before and after the stage shift is applied as a simple estimate. This benefit accrues after the dates of original diagnosis and is spread across time after a variable delay. We note that screening particularly affects the fate of individuals who are expected to die within 5 years of their original diagnosis (Individuals Expected to Die, IED), and report metrics associated with such individuals.

Results

Modeled benefits

For the period 2006–2015, annual cancer incidence averaged 1187 cases per 100,000, with 409 of those cancers diagnosed at late (III+IV) stage with 110 cases not stageable for this model (Table 1). Usual care is supplemented with a screening program involving a MCED test under five scenarios: no MCED screening, three scenarios with annual MCED screening but different cancer growth rate assumptions, and the maximum benefit realized from idealized screening. For each, modeled results are recorded in three sections: effect of screening using a MCED test to intercept cancers, the resulting reduction in late-stage cancers, and finally, the effect of stage shift on cancer mortality. Summary statistics are given for both the subset of incidence intercepted by the MCED test, and the resulting incidence that would remain to be found by usual care.

Overall, when modeling for the maximum benefit, 485 (41%) of all cancers would be intercepted by a MCED test (Table 1), whereas the remainder would be found by usual care. Depending on cancer growth rate, this could be adjusted downward by the schedule sensitivity to 312 (26%) in the most aggressive growth rate scenario. In all scenarios, 110 cancers are unstageable and account for 52 deaths within 5 years, which are conservatively assumed to be unaffected by screening as stage shift cannot be modeled for such individuals.

Under maximum benefit, the incidence of cancers found by usual care is heavily weighted toward early-stage cancers, with only 12% of late-stage cancers found in this set. In contrast, 68% of the intercepted

Figure 2.

Influence of screening schedule on stage at diagnosis. **Figure 2** illustrates that the screening schedule affects not only the probability of being intercepted but also the stage at diagnosis. This individual may only be intercepted in stage II, if detectable at stage II. Even if detectable at stage I, this individual cannot be found by this screening schedule before stage II, and “slips” to stage II. Similarly, if this individual is not detectable during stage II, they will only be found by usual care at stage IV. Given distributions of dwell times and a screening schedule, the probability of such events can be calculated and applied in the algorithm.

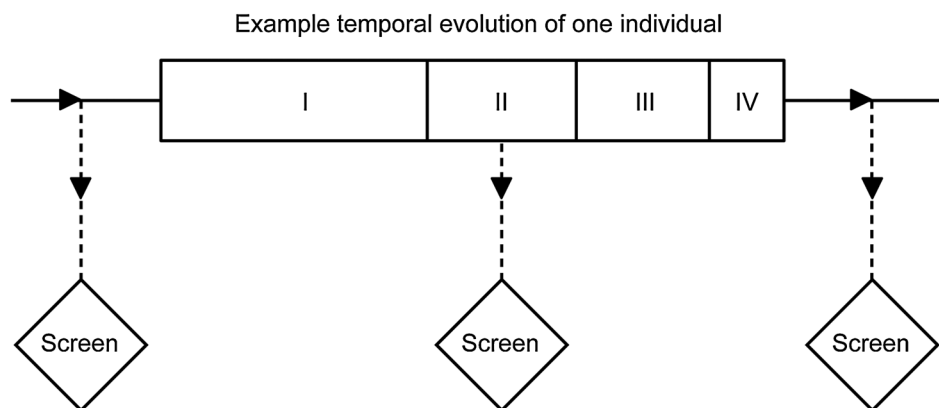


Table 1. Summary interception statistics for one year of incident screening (estimating long-term performance) across cancer growth rate scenarios with and without MCEd test.

	No screening	Cancer growth scenario			Idealized screening
Scenario (years in stage I)	No MCEd	AggFast (1–2)	Fast (2–4)	Slow (3–7)	MIS (NONE)
Screening type	No	Incident	Incident	Incident	Incident
Basic performance					
Total cancer incidence (100K person years), <i>n</i>	1,187	1,187	1,187	1,187	1,187
Found usual care, <i>n</i> (% total cancer)	1,187 (100)	875 (74)	816 (69)	783 (66)	702 (59)
Intercepted, <i>n</i> (% total cancer)	0 (0)	312 (26)	371 (31)	404 (34)	485 (41)
False positives, <i>n</i>	0	692	692	692	692
True positives, <i>n</i> (PPV = % of positives)	0 (NA)	312 (31)	371 (35)	404 (37)	485 (41)
Stage shift					
Total late (III+IV), <i>n</i> (% total cancers)	409 (34)	409 (34)	409 (34)	409 (34)	409 (34)
Final late (III+IV), <i>n</i> (% total cancers)	409 (34)	233 (20)	206 (17)	189 (16)	152 (13)
Reduction late, <i>n</i> (% total late)	0 (0)	–177 (43)	–204 (50)	–220 (54)	–257 (63)
Breakout by mode found					
Stage III+IV usual care, <i>n</i> (% found usual care)	409 (34)	183 (21)	148 (18)	130 (17)	82 (12)
Stage III+IV pre-intercept, <i>n</i> (% intercepted)	0 (NaN)	226 (72)	261 (70)	280 (69)	327 (68)
Stage III+IV post-intercept, <i>n</i> (% intercepted)	0 (NaN)	50 (16)	58 (16)	60 (15)	70 (15)
Reduction (III+IV), <i>n</i> (% late intercepted)	0 (NaN)	–177 (78)	–204 (78)	–220 (79)	–257 (78)
Individuals expected to die (IED) of cancer in 5 years					
Total IED, <i>n</i> (% total cancer)	393 (33)	393 (33)	393 (33)	393 (33)	393 (33)
Final IED, <i>n</i> (% total cancer)	393 (33)	319 (27)	308 (26)	302 (25)	289 (24)
Reduction IED, <i>n</i> (% total IED)	0 (0)	–74 (19)	–84 (21)	–91 (23)	–104 (26)
Breakout by mode found					
IED Usual care, <i>n</i> (% total IED)	393 (100)	212 (54)	183 (47)	166 (42)	126 (32)
IED Pre-intercept, <i>n</i> (% total IED)	0 (0)	181 (46)	210 (53)	227 (58)	267 (68)
IED Post-intercept, <i>n</i> (% IED pre-intercept)	0 (NaN)	107 (59)	126 (60)	136 (60)	163 (61)
Reduction IED, <i>n</i> (% IED pre-intercept)	0 (NaN)	–74 (41)	–84 (40)	–91 (40)	–104 (39)

Note: All numbers are per 100K people per year. Aggressive: all cancers assumed to have very short dwell times due to aggressive growth; Fast: all cancers have slightly longer dwell times and grow quickly; Slow: all cancers even longer dwell times, slow growth.

No MCEd: usual care results with no MCEd screening.

Abbreviations: MIS: maximum interception (idealized screening), cancers that can be intercepted are found at the earliest possible stage. See dwell time in methods.

cancers would have been late stage under usual care (i.e., “68% of pre-intercept cancers”); postinterception, only 15% would be late stage due to earlier detection, a reduction of 78%. This corresponds to an absolute reduction of 257 cases per 100,000 in late-stage cancers, a 63% reduction of all cancer incidence in late stage. Dependent on the cancer growth rate, the effect of schedule sensitivity would decrease the absolute number of cancers intercepted, but those intercepted have a similar late-stage reduction across scenarios. This change in proportion could be detected in a two-armed experiment with from 10,500 to 26,000 person-years of cancer risk in each arm (power = 0.95, significance level 0.05).

An estimated 393 of the 1,187 individuals diagnosed per year by usual care in the absence of other risks would die of their cancer within 5 years of their original date of diagnosis (Individuals Expected to Die, IED). Under maximum benefit, 267 (68%) of those individuals would be intercepted before usual care. Applying the hazard reduction from the stage at interception when compared with the original stage at diagnosis, only 163 of those individuals would die within 5 years of the original diagnosis date, a 39% reduction in mortality for an absolute reduction of 104 deaths per 100,000, resulting in a relative 26% mortality reduction in all-cancer mortality. Across growth rate scenarios, the effect of schedule sensitivity results in fewer cancers being intercepted, but a stable 40% reduction in mortality among those intercepted. Even the most aggressive growth rate scenario reduces deaths by 74 per 100,000, for a 19% relative mortality reduction in all-cancer mortality. A change of proportion of this magnitude is poten-

tially detectable at a scale of 81,000 to 170,000 person-years of cancer risk per arm (power = 0.95, significance level 0.05) plus associated follow-on time.

Figure 3 illustrates the effect of MCEd screening by individual stage for the MIS scenario on both the intercepted cancers and those remaining to be found by usual care. **Figure 3A** shows the cancer incidence intercepted by screening (485/1187; light gray bars). **Figure 3B** shows the portion of the total cancer incidence by stage that would be found by usual care despite implementation of the MCEd test (702/1187; dark gray bars). For those cancers intercepted, we show both the stage that usual care would have found them (light gray), as well as the stage at which they are intercepted (pink). **Figure 3** (bottom row) extends this to the effect on mortality due to stage shift. The cancer incidence intercepted (485/1,187) includes a high fraction of cancers that would have been diagnosed at a late stage, but once shifted to an earlier stage, improved mortality is observed (**Fig. 3C**). For example, after 5 years from the original date of diagnosis, 45% would have remained alive even if not intercepted, 21% would have survived due to interception, and 34% would have died, whereas the cancer incidence and stage remaining to be found by usual care (702/1,187) consisted of early-stage cancers having relatively low mortality (**Fig. 3D**). For example, after 5 years, 82% would be alive and 18% dead, suggesting that cancers not found by the MCEd test are relatively well controlled by usual care. Barplots illustrating the effect of other tumor growth scenarios on stage can be found in Supplementary Fig. S4.

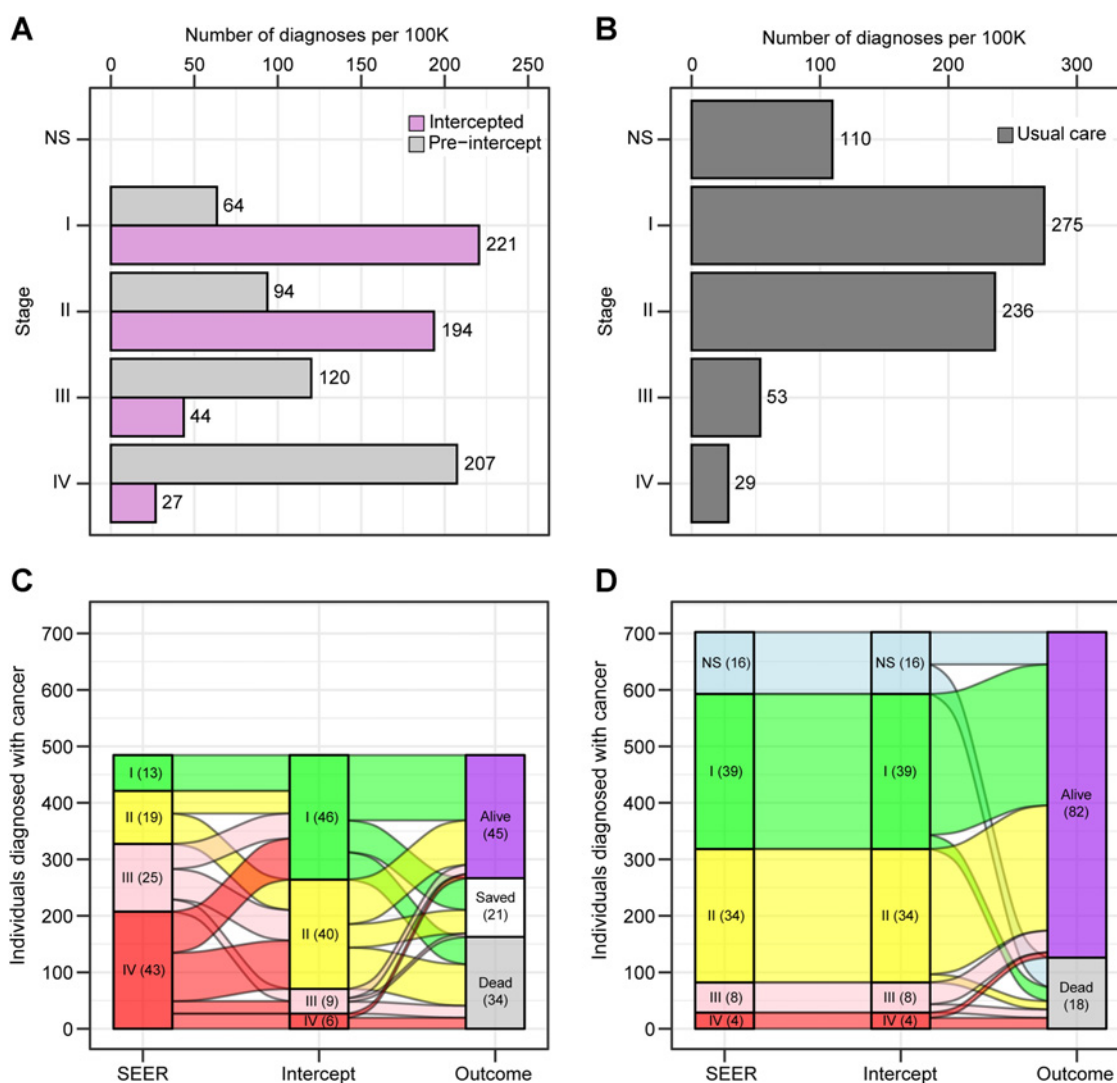


Figure 3. Modeled changes in stage-specific cancer incidence. **Figure 3** contains results from an incidence round under idealized screening (MIS). The top row depicts all-cancer incidence separated into intercepted by MCED, with resulting stage shift showing that before interception most would have been late stage (A), and cancers not intercepted that remain to be found by usual care and are generally already found at early stages providing limited opportunity to benefit from interception (B). The bottom row depicts all-cancer incidence and mortality are separated into intercepted by the MCED test, with corresponding stage shift and mortality reduction shown (C), and cancers not intercepted and remaining to be found by usual care (D). Percentage of individuals in each column shown in parentheses.

Case scenarios of interest

One case of interest assumes the cancer types with USPSTF-recommended screening (breast, cervical, colorectal, and high-risk lung cancer) are already receiving the maximum benefit from any screening, so that supplemental screening does not affect that fraction of cancer incidence (Supplementary Methods and Materials and Supplementary Tables S4 and S5). For incidence rounds at maximum benefit, 320 (27%) cases would be intercepted. A total of 71% would have presented at late stage, but only 15% would be at late stage after interception; this would result in an 80% reduction in late stage in those intercepted (an absolute reduction of 181 per 100,000), or an overall 44% reduction in late-stage cancer incidence. The resultant relative death reduction based on cancers intercepted would be 35% (an absolute death reduction of 68 per 100K), leading to a relative mortality reduction of 17% in all-cancer mortality.

Outcomes following the first year a screening test is introduced (prevalence round) are in Supplementary Table S2. Because the predictions are variable (due to the number intercepted being directly proportional to the dwell times), results are reported only in the supplement.

Modeled harms

Three primary categories of harm occur in a screening program: overdiagnosis, false positives, and direct harms from the test. Overdiagnosis is not explicitly modeled, however, over the 50–79 age range, competing mortality risks are approximately 2% per year (10). If intercepted cancers are shifted from 5 years in the future on average, this would result in an approximate 10% increase in incidence due to diagnosis of cancers in individuals who would otherwise die from non-cancer-specific causes.

False positives are determined by the specificity of the test (reported as 99.3%; ref. 5). This is an overall false-positive rate based only on the noncancer individuals screened and does not accumulate based on the number of cancer types reported. This implies 692 individuals per 100,000 would be falsely positive, and may experience a diagnostic odyssey. We conservatively assume that this rate does not decrease after screening with a MCED test, so that each year new systematic false positives are found at the same rate. This results in a final positive predictive value (PPV) of 41% for the maximum benefit scenario. Schedule sensitivity could reduce this to 31%.

Additional harms are difficult to estimate in the absence of interventional evidence. The test consists of a standard blood draw that is generally considered safe, without the need for immediate invasive procedures. Several previous studies of other cancer screening tests with higher false positive rates have evidenced minimal indirect harms (e.g., anxiety, etc) of receiving test results (26–29).

Discussion

We constructed this interception model to handle a screening test that incorporates cfDNA and, therefore, is sensitive to multiple cancer types. Because the MCED test can potentially detect cancer at any site, the effect of screening aggregates across many sites simultaneously rather than concentrates on a particular organ. In addition, detection of a given cancer by the MCED test relies on cancer biology (e.g., tumor shedding) and not a snapshot of a specific anatomic site (e.g., imaging breast or lung, etc.). Existing models primarily treat screening tests as independent, uncorrelated observations, even when scanning the same individual. In contrast, the threshold sensitivity behavior of the MCED test leads to correlation between screening rounds that must be accounted for in the model. Combined with the fact that incidence rounds have stable incidence, these observations lead to a relatively simple estimator of the total effect of screening including potential stage shift and corresponding mortality benefit.

We examined the robustness of our findings across a broad range of tumor growth rate assumptions and found the majority of the effects of a MCED test were preserved. We demonstrated that a MCED test could produce absolute numbers of potential immediate (within 5 year) deaths averted (77 to 100 per 100,000 per year) larger than existing cancer screening efforts, and is in the same range as other major causes of death, including cerebrovascular disease, among persons aged 50–84 in the United States (10). These potential reductions in mortality represent the cumulative effects of stage shift among multiple cancer types simultaneously with no particular cancer type dominating the effect, and also underscore the problem of contemporary diagnosis of many cancers predominantly at stage IV (9). Furthermore, the very low false positive rate suggests a PPV of 30% to 40% is achievable under these scenarios, comparing favorably to both the cumulative false positive rate and PPVs for existing single cancer screening tests (2, 30–32).

We refined existing single-cancer screening models in several ways, including a focus on incidence rounds. Although prevalence round effects can be computed, results for incidence rounds are much less sensitive to unknown parameters. As most cancer cases will not be found at the prevalence screen (18), incidence round performance best represents the long-term stage shift and mortality reduction (15). Second, the interception model avoids using computationally expensive microsimulations by capturing the major dynamics of stage transitions in a fast model (33–35). Finally, interpreting sensitivity as a detectable fraction of a population is consistent with growth-based models (17). All of these allow

extrapolation of population-scale effects that are robust to uncertainty in the natural history of each cancer type.

Our model has several important limitations stemming from its reliance on cancer registry data. We necessarily utilized historic data (2006–2015) to have sufficient follow-up for cancer mortality that may not reflect current cancer incidence, survival and usual care screening adherence. Furthermore, SEER statistics may not reflect underlying biological differences between cancers that shed cfDNA and cancers that do not. Our model included cancer types for which staging is either unconventional or not available from SEER (e.g., leukemia, myeloma), but considered them to have zero sensitivity, even though published MCED evidence (5) suggests some degree of detectability. Thus, our estimates of performance including the unstageable cancer types are slightly conservative in terms of yield, positive predictive value, and mortality benefit. We focused on persons aged 50–79 but that does not preclude the possibility of impact on persons in other age groups, especially those aged 40–49 for which existing screening recommendations cover (e.g., USPSTF cervical, American Cancer Society breast and colorectal).

Our model has several caveats as well. The duration over which cancers shed detectable biomarkers is not known, which may affect necessary screening intervals. Sensitivity in a case-control study is assumed reflective of performance in a real-world population at preclinical stages. Finally, although Liu and colleagues provide a tissue-of-origin estimate (5) that can guide downstream workup of positive cases, the complexity of diagnosis following an MCED test positive result is not yet known. Resolving some of these limitations would require interventional evidence, such as from a randomized clinical trial or real-world interventional study.

Conclusion

This new ability to efficiently model screening across all cancers allows estimation of the potential aggregate effects of a MCED test based on generally accepted assumptions about cancer natural history. This is timely given recent reports of novel blood-based tests that detect 8 to 50 or more cancer types proposed for population use (5, 36, 37). Further extensions of the model can handle important options: stratifying populations by risk, incorporation of adherence, analysis of screening intervals, and confidence intervals on results given measurement uncertainty. The primary output of the model is stage shift, and the resulting effect on mortality can be refined in future models (25). Refinements may also enable estimation of health economic benefit-cost tradeoffs in the use of a MCED test. The results shown here indicate that the potential benefit from supplementing usual care with a MCED test could include large absolute effects on population burden of late-stage cancer, and improvements in mortality.

Authors' Disclosures

E. Hubbell reports other from GRAIL outside the submitted work; in addition, E. Hubbell has a patent for Multi-assay Prediction Model for Cancer Detection pending, a patent for Machine Learning Variant Source Assignment pending, a patent for Identifying Copy Number Aberrations pending, a patent for systems and Methods for Determining Tumor Fraction in Cell Free Nucleic Acid pending, a patent for Methods for using Fragment Lengths as a Predictor of Cancer pending, a patent for Cancer Tissue Source of Origin Prediction with Multi-Tier Analysis of Small Variants pending, a patent for Endpoint Analysis in Early Cancer Detection pending, and a patent for Identifying Methylation Patterns that Discriminate or Indicate a Cancer Condition pending; and Editorial and submission assistance was also provided by ProEd Communications, Inc. (Beachwood, OH), funded by GRAIL, Inc. E. Hubbell has additional patents issued and pending for various aspects of sequencing

technology and microarrays due to prior work for ThermoFisher, as well as current work for GRAIL, Inc. C.A. Clarke reports other from GRAIL during the conduct of the study. A.M. Aravanis reports personal fees from Myst Therapeutics, Foresite Capital, and personal fees from Illumina outside the submitted work; in addition, A.M. Aravanis has a patent for GRAIL pending. C.D. Berg reports personal fees from GRAIL and personal fees from Mercy BioAnalytics during the conduct of the study. No other disclosures were reported.

Authors' Contributions

E. Hubbell: Conceptualization, resources, data curation, software, formal analysis, validation, investigation, visualization, methodology, writing—original draft, writing—review and editing. **C.A. Clarke:** Conceptualization, resources, data curation, investigation, methodology, writing—original draft, writing—review and editing. **A.M. Aravanis:** Conceptualization, resources, supervision, funding acquisition, writing—original draft, project administration, writing—review and editing. **C.D. Berg:** Conceptualization, writing—original draft, writing—review and editing.

Acknowledgments

This study was funded by GRAIL, Inc. We thank Sarah A. Prins, PhD (GRAIL, Inc.) for medical writing support. Editorial and submission assistance was also provided by ProEd Communications, Inc. (Beachwood, OH), funded by GRAIL, Inc. We gratefully acknowledge Sylvia K. Plevritis, PhD (Stanford University), Ruth Etzioni, PhD (Oregon Health State University), Peter B. Bach, MD, MAPP (Memorial Sloan Kettering Cancer Center), Steven J. Skates (Massachusetts General Hospital), Hormuzd Katai (National Cancer Institute), and Peter Sasieni (King's College London) for their critical input.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 1, 2020; revised October 23, 2020; accepted December 10, 2020; published first December 16, 2020.

References

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* 2020;70:7–30.
- Croswell JM, Kramer BS, Kreimer AR, Prorok PC, Xu J-L, Baker SG, et al. Cumulative incidence of false-positive results in repeated, multimodal cancer screening. *Ann Fam Med* 2009;7:212–22.
- Ahlquist DA. Universal cancer screening: revolutionary, rational, and realizable. *NPJ Precis Oncol* 2018;2:23.
- Phallen J, Sausen M, Adleff V, Leal A, Hruban C, White J, et al. Direct detection of early-stage cancers using circulating tumor DNA. *Sci Transl Med* 2017;9:eaan2415.
- Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV, Cummings SR, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. *Ann Oncol* 2020;31:745–59.
- Cohen JD, Li Lu, Wang Y, Thoburn C, Afari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science* 2018;359:926–30.
- Srivastava S, Hanash S. Pan-cancer early detection: hype or hope? *Cancer Cell* 2020;38:23–24.
- CISNET Model Registry. National Cancer Institute Cancer Intervention and Surveillance Modeling Network, CISNET Model Registry; 2020. Available from: <https://resources.cisnet.cancer.gov/registry/packages/>.
- Clarke CA, Hubbell E, Kurian AW, Colditz GA, Hartman A-R, Gomez SL. Projected reductions in absolute cancer-related deaths from diagnosing cancers before metastasis, 2006–2015. *Cancer Epidemiol Biomarkers Prev* 2020;29:895–902.
- Surveillance, Epidemiology, and End Results (SEER) Program. Surveillance, Epidemiology, and End Results (SEER) Program SEER*Stat Database: Mortality - All COD, Aggregated With State, Total U.S. (1969–2016) <Katrina/Rita Population Adjustment>, National Cancer Institute, DCCPS, Surveillance Research Program, released December 2018. Underlying mortality data provided by NCHS (www.cdc.gov/nchs). Surveillance, Epidemiology, and End Results (SEER) Program. Available from: www.seer.cancer.gov.
- R Core Team. Core R: A Language and Environment for Statistical Computing, Version 3.5.3. Vienna: R Foundation for Statistical Computing; 2019. Available from: <https://www.R-project.org/>.
- Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the tidyverse. *J Open Source Softw* 2019;4:1686.
- Gulati R, Feuer EJ, Etzioni R. Conditions for valid empirical estimates of cancer overdiagnosis in randomized trials and population studies. *Am J Epidemiol* 2016;184:140–7.
- Lee SJ, Zelen M. Mortality modeling of early detection programs. *Biometrics* 2008;64:386–95.
- Liu ZA, Hanley JA, Saarela O, Dendukuri N. A conditional approach to measure mortality reductions due to cancer screening: measuring mortality reductions due to screening. *Int Stat Rev* 2015;83:493–510.
- Michalopoulos D, Duffy SW. Estimation of overdiagnosis using short-term trends and lead time estimates uncontaminated by overdiagnosed cases: results from the norwegian breast screening programme. *J Med Screen* 2016;23:192–202.
- Skates SJ, Singer DE. Quantifying the potential benefit of CA 125 screening for ovarian cancer. *J Clin Epidemiol* 1991;44:365–80.
- Shen Y. Robust modeling in screening studies: estimation of sensitivity and preclinical sojourn time distribution. *Biostatistics* 2005;6:604–14.
- Plevritis SK, Sigal BM, Salzman P, Rosenberg J, Chapter GP. 12: A stochastic simulation model of U.S. breast cancer mortality trends from 1975 to 2000. *JNCI Monogr* 2006;2006:86–95.
- Han SS, Erdogan SA, Toumazis I, Leung A, Plevritis SK. Evaluating the impact of varied compliance to lung cancer screening recommendations using a microsimulation model. *Cancer Causes Control* 2017;28:947–58.
- CISNET Fred Hutchinson Cancer Research Center (PSAPC) Prostate Cancer Model Readers Guide; 2009. Available from: https://cisnet.flexkb.net/mp/pub/cisnet_prostate_fhrc_psapc.pdf.
- CISNET Erasmus MC. Prostate Cancer Model Readers Guide; 2009. Available from: https://cisnet.flexkb.net/mp/pub/CISNET_ModelProfile_PROSTATE_ERASMUS_001_12152009_69754.pdf.
- Gulati R, Gore JL, Etzioni R. Comparative effectiveness of alternative prostate-specific antigen-based prostate cancer screening strategies: model estimates of potential benefits and harms. *Ann Intern Med* 2013;158:145.
- Etzioni R, Tsodikov A, Mariotto A, Szabo A, Falcon S, Wegelin J, et al. Quantifying the role of PSA screening in the US prostate cancer mortality decline. *Cancer Causes Control* 2008;19:175–81.
- Wever EM, Draisma G, Heijnsdijk EAM, de Koning HJ. How does early detection by screening affect disease progression?: modeling estimated benefits in prostate cancer screening. *Med Decis Making* 2011;31:550–8.
- Kirkøen B, Berstad P, Botteri E, Åvitsland TL, Ossum AM, de Lange T, et al. Do no harm: no psychological harm from colorectal cancer screening. *Br J Cancer* 2016;114:497–504.
- Wu GX, Raz DJ, Brown L, Sun V. Psychological burden associated with lung cancer screening: a systematic review. *Clin Lung Cancer* 2016;17:315–24.
- Taksler GB, Keating NL, Rothberg MB. Implications of false-positive results for future cancer screenings: implications of false-positive results. *Cancer* 2018;124:2390–8.
- Andrykowski M. Psychological and behavioral impact of participation in ovarian cancer screening. *Diagnostics* 2017;7:15.
- Lin JS, Piper MA, Perdue LA, Rutter CM, Webber EM, O'Connor E, et al. Screening for colorectal cancer: updated evidence report and systematic review for the us preventive services task force. *JAMA* 2016;315:2576.
- Pinsky PF, Gierada DS, Black W, Munden R, Nath H, Aberle D, et al. Performance of lung-RADS in the national lung screening trial: a retrospective assessment. *Ann Intern Med* 2015;162:485–91.
- Lehman CD, Arao RF, Sprague BL, Lee JM, Buist DSM, Kerlikowske K, et al. National performance benchmarks for modern screening digital

- mammography: update from the breast cancer surveillance consortium. *Radiology* 2017;283:49–58.
33. Vänskä S, Bogaards JA, Auranen K, Lehtinen M, Berkhof J. Fast approximate computation of cervical cancer screening outcomes by a deterministic multiple-type HPV progression model. *Math Biosci* 2019; 309:92–106.
 34. de Carvalho TM, Heijnsdijk EAM, Coffeng L, de Koning HJ. Evaluating parameter uncertainty in a simulation model of cancer using emulators. *Med Decis Making* 2019;39:405–13.
 35. Vilaprinyo E, Forné C, Carles M, Sala M, Pla R, Castells X, et al. Cost-effectiveness and harm-benefit analyses of risk-based screening strategies for breast cancer. *PLoS One* 2014;9:e86858.
 36. Shen SYi, Singhanian R, Fehringer G, Chakravarthy A, Roehrl MHA, Chadwick D, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature* 2018;563:579–83.
 37. Lennon AM, Buchanan AH, Kinde I, Warren A, Honushefsky A, Cohain AT, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. *Science* 2020;369:eabb9601.