



Health State Utilities Associated with False-Positive Cancer Screening Results

Louis S. Matza¹ · Timothy A. Howell¹ · Eric T. Fung² · Sam M. Janes³ · Michael Seiden⁴ · Allan Hackshaw³ · Lincoln Nadauld⁵ · Hayley Karn⁶ · Karen C. Chung²

Accepted: 14 September 2023
© The Author(s) 2024

Abstract

Introduction Early cancer detection can significantly improve patient outcomes and reduce mortality rates. Novel cancer screening approaches, including multi-cancer early detection tests, have been developed. Cost-utility analyses will be needed to examine their value, and these models require health state utilities. The purpose of this study was to estimate the disutility (i.e., decrease in health state utility) associated with false-positive cancer screening results.

Methods In composite time trade-off interviews using a 1-year time horizon, UK general population participants valued 10 health state vignettes describing cancer screening with true-negative or false-positive results. Each false-positive vignette described a common diagnostic pathway following a false-positive result suggesting lung, colorectal, breast, or pancreatic cancer. Every pathway ended with a negative result (no cancer detected). The disutility of each false positive was calculated as the difference between the true-negative and each false-positive health state, and because of the 1-year time horizon, each disutility can be interpreted as a quality-adjusted life-year decrement associated with each type of false-positive experience.

Results A total of 203 participants completed interviews (49.8% male; mean age = 42.0 years). The mean (SD) utility for the health state describing a true-negative result was 0.958 (0.065). Utilities for false-positive health states ranged from 0.847 (0.145) to 0.932 (0.059). Disutilities for false positives ranged from -0.031 to -0.111 (-0.041 to -0.111 for lung cancer; -0.079 for colorectal cancer; -0.031 to -0.067 for breast cancer; -0.048 to -0.088 for pancreatic cancer).

Conclusion All false-positive results were associated with a disutility. Greater disutility was associated with more invasive follow-up diagnostic procedures, longer duration of uncertainty regarding the eventual diagnosis, and perceived severity of the suspected cancer type. Utility values estimated in this study would be useful for economic modeling examining the value of cancer screening procedures.

1 Introduction

Early cancer detection and intervention can significantly improve patient outcomes and reduce mortality rates [1–3]. The benefits of early detection and intervention, including reduction in cancer-specific mortality, have been demonstrated across a range of cancer types, such as lung [4, 5], colorectal [6–8], cervical [9, 10], and breast cancer [11, 12]. Cancers for which at-risk patients are regularly screened tend to be detected earlier and have a better prognosis than cancers without available screening procedures [13–16]. Despite these well-known benefits of early cancer

detection, few types of cancer have established screening procedures, and patient adherence to recommendations for cancer screening varies [17].

New cancer screening approaches that may address some of these limitations are currently in development. While some of these novel screening tests focus on identification of one type of cancer [18–20], others are multi-cancer early detection (MCED) tests including blood-based genomic tests that have demonstrated the ability to detect a wide range of cancers via a simple blood draw [21–24]. If a simple MCED policy could be implemented across a broad segment of the general population to complement current cancer screening, the increase in early detection could lead to more people who experience curative treatments, better health outcomes, and decreased mortality [25].

An effective MCED test should detect a variety of cancers across stages, predict cancer signal origin, and have

GRAIL, LLC. is currently held separate from Illumina Inc. under the terms of the Interim Measures Order of the European Commission dated 29 October 2021.

Extended author information available on the last page of the article

Key Points For Decision Makers

A range of novel cancer screening tests are in development, and the utilities estimated in this study will be useful in economic modeling conducted to examine and compare their value.

Results of this study add to the limited body of literature suggesting that there is a disutility (i.e., utility decrease) associated with false-positive cancer screening results.

Greater disutility (i.e., utility decrease quantifying impact on quality of life) of false-positive screenings was associated with more invasive follow-up diagnostic procedures, longer duration of uncertainty regarding the eventual diagnosis, and greater perceived severity of the suspected cancer type.

a very low rate of false-positive results [26]. All screening approaches, including single-cancer and MCED tests, have the potential for false-positive results with a direct and adverse impact on individuals. During the time between the report of a false-positive test result and definitive confirmation that the individual does not have cancer, there are often negative emotional consequences for patients, including worry, distress, and symptoms of anxiety [27–29]. In addition, false-positive screening tests can lead to a range of expensive and possibly invasive follow-up investigations.

As new cancer screening methods are proposed and evaluated, cost-utility analyses (CUAs) will be needed to examine and compare their value, and the impact of false-positive results must be considered as part of these models. Health state utilities representing false positives will be needed. Utilities are values representing the strength of preference for various health states, and they are used to quantify health status and quality of life in CUAs [30].

Previous research suggests that a false-positive screening result is associated with a disutility (i.e., a utility decrease representing a negative impact on quality of life) [31]. However, the few studies estimating disutilities of false positives are limited in their scope and methodology. Published studies generally focus on breast cancer [32–35], with variable findings. For example, Tosteson and colleagues [35] found no utility decrement, possibly because the measure used (3-level EQ-5D) assesses broader domains of general health and may not be sensitive to the specific impact of false-positive test results. Other studies used vignette-based methods (i.e., utility elicitation for health state vignettes in preference-based tasks) [36], but methods for rating the utility of each vignette varied. Some research used visual analog scales, which are not considered to be a true

preference-based utility estimation method [32, 34, 35]. In other studies, vignette descriptions were mapped to the EQ-5D scoring system [33] or valued with direct time trade-off (TTO) methods [34].

In these previous studies, none of the false-positive vignettes included detailed descriptions of follow-up tests or the timeline to resolve the false-positive result. Therefore, available utilities are based largely on perceptions of people with limited insight into the false-positive experience. Furthermore, each study only provided a single utility for a false-positive screening without considering the heterogeneity of patients' false-positive experiences, which vary widely. While some false positives can be resolved quickly following a single diagnostic test (e.g., imaging or biopsy), others precipitate a sequence of multiple tests with varying levels of invasiveness administered over an extended period of uncertainty regarding the cancer diagnosis. In sum, previously published studies do not provide adequate estimates of the utility impact associated with the range of false-positive cancer screenings.

The purpose of this study was to estimate the disutility (i.e., decrease in health state utility) associated with false-positive cancer screenings. This study was conducted using vignette-based methods, which are well-suited for estimating utilities of specific health-related events that are not the focus of generic instruments such as the EQ-5D [36]. To provide disutility values that may be useful in CUAs of cancer screening procedures, vignettes were developed to represent several types of false-positive experiences, varying in type of suspected cancer, extent of follow-up testing, and duration of diagnostic uncertainty.

2 Methods

2.1 Overview of Study Design

Although some health technology assessment (HTA) authorities prefer that preference-based measures such as the EQ-5D are used to generate utilities [37–39], this study was conducted with vignette-based methodology for two reasons. First, generic instruments designed to quantify overall health status may not be sensitive to specific treatment process attributes, such as those described in the current health states (e.g., coping with diagnostic uncertainty, medical testing, perceptions of various cancer types). There is a substantial and growing body of literature suggesting that quality of life and utility may be influenced not only by health status and treatment outcomes, but also by the process of receiving care [36, 40, 41]. Estimation of 'treatment process utilities' is a situation when generic instruments like the EQ-5D are often not considered to be appropriate because

a generic instrument that assesses overall health status may not be sensitive to specific treatment process attributes [36, 40, 41]. Therefore, treatment process utilities are typically estimated using vignette-based methodology [40]. With this method, health state vignettes can describe specific medical treatments so that the resulting utilities represent the impact of these experiences [36].

Second, the goal of this study was to estimate utilities of a broad range of false-positive pathways. For several reasons, it may not be feasible to administer questionnaires to a large enough sample of patients who are currently experiencing the false-positive diagnosis described within each health state. Generic instruments like the EQ-5D are designed to assess health status ‘today,’ and it is not possible to identify a false-positive experience at the time it is occurring because neither the patient nor any clinician knows it is a false positive until after the patient is found to be negative. In addition, several of the cancer types represented in the health states (e.g., pancreatic cancer, specific sub-types of lung cancer) are not specifically detected with currently used screening procedures. Therefore, it is not currently possible to recruit any participants experiencing these false-positive health states. Furthermore, for health states describing false positives that occur with standard screening approaches (e.g., breast cancer, colorectal cancer), a false positive cannot be identified until after the follow-up testing is complete. Therefore, a generic instrument would need to be administered to an extremely large sample in order to identify the relatively small subgroups who experience the nine types of false positives described in the current health states. Finally, because generic measures provide a utility estimate only for the day they are completed, the instrument would have to be administered repeatedly for the weeks following the screening result to capture the disutility of the full false-positive pathway. In contrast, the vignette-based approach avoids these obstacles because vignettes can be drafted based on input from clinicians and valued by members of the general population without requiring a large sample of patients, and the vignette approach allows for estimation of disutility (i.e., quality-adjusted life-year [QALY] decrement) associated with the full false-positive pathway.

Health state vignettes were developed based on published literature, cancer screening guidelines, and expert interviews. Then, the health states were refined in several iterations during a pilot study. The initial health state described a 1-year period that included a negative cancer screening. All other health states described a 1-year period that included a false-positive cancer screening that was eventually resolved following diagnostic testing.

The utility difference between otherwise identical health states with and without the false-positive screening

represents the disutility of the false-positive experience. Because of the 1-year time horizon, this disutility can be considered a QALY decrement. Whereas most vignette-based studies value chronic health states that remain stable over time, the current vignettes can be considered ‘path states,’ which describe a sequence of events. Like other studies involving path states, utilities were estimated for the whole path rather than each part of the path [42–46]. For example, if a health state described an initial cancer screening and two follow-up tests (e.g., an MRI and a PET-CT), the associated utility would represent the entire series of events, and it is not possible to derive a separate disutility for each follow-up test.

Health states were valued in a TTO utility elicitation with a sample of general population participants in the UK. Because of the COVID-19 pandemic, face-to-face interviews were not conducted during the pilot phase (April–June 2021). However, the main valuation study was conducted via individual in-person utility interviews in November 2021. Informed consent was obtained prior to each interview, and the study protocol was approved by an institutional review board (Ethical and Independent Review Services; Study 21017-01).

A variety of time horizons (i.e., the duration of time spent in each health state) can be used in TTO valuations [47, 48]. For valuations of chronic health states, longer time horizons (e.g., 10 or 20 years) are typically used. Because this study valued path states in which the events described were relatively brief (most lasting <1 month as described in the health state text and depicted in a timeline figure at the bottom of each health state), a 1-year time horizon was used. Interviewers explained to participants that the false-positive experience described in each health state occurred for only part of the year that they would value in the TTO task, rather than the whole year. Interviewers used the following language: “Over the course of the year, you will be screened for cancer, learn that you do not have cancer, and then live the remainder of the year in good health. This is shown on the timeline at the bottom of the health state. Please note that the timeline only describes the duration of these events. These events could occur at any time during the year. For the remainder of the year, you will be in good health.”

The 1-year timeline was chosen to ensure that the TTO elicitation method would be sensitive to the impact of the brief false pathways. When estimating utilities associated with brief events, a longer timeline is typically not sensitive to the impact of these events because a brief event is unlikely to have any significant impact on preference for a 10-year period. However, a 1-year time horizon can be sensitive to these short-term differences [42, 44, 45]. The 1-year time horizon also simplifies the interpretation and use of the

results, because the disutility associated with each temporary event can be applied in a CUA as a QALY decrement.

2.2 Health State Development

Health states were developed based on literature review, cancer screening guidelines, and input from six advisors. We focused on cancer screening and diagnostic investigations associated with lung, breast, colorectal, and pancreatic cancer. These four were selected to represent cancers with well-established screening policies, cancers requiring a range of follow-up investigative procedures, and a cancer (pancreatic) that currently has no effective screening test but represents an uncommon cancer with a poor prognosis that can be detected by MCED strategies.

The literature review included peer-reviewed literature [49–58] and official guidelines for cancer screening [59–62]. Google Scholar and PubMed were searched using general terms such as ‘cancer screening procedures’ and specific terms for each type of cancer that was described in the health states (e.g., breast cancer, pancreatic cancer). Articles were selected for review if they provided information describing screening procedures in the US or UK. Websites for cancer research and patient advocacy organizations were also reviewed for cancer screening recommendations and patient-friendly language for describing various tests in the health states [61–70].

To gather further information and refine the health state text, multiple rounds of interviews were conducted with six advisors, including five clinicians experienced in cancer screening methodologies (pulmonologist, radiologist, anesthesiologist, and two oncologists). The other advisor was an academic professor who is a member of the Adult Reference Group of the UK National Screening Committee, which advises the UK government on appropriate population cancer screening policies. The advisors reported an average of over 20 years’ experience with cancer screening, and the five clinicians reported seeing between 10 and 150 patients per month. Two advisors were in London, UK, and others were in the US (Texas, California, Utah, Virginia). In the initial interviews, the advisors were asked to describe patients’ typical experiences with cancer screening and follow-up investigations after an initial positive result. These descriptions were used to develop the first draft of the health states. In subsequent discussions, the advisors reviewed and edited health state drafts to ensure the descriptions were clear and accurate representations of typical patient experiences with false-positive screenings.

For some of the health states, the time required for follow-up investigations was perceived to be longer in the UK than the US. The differences in time required for follow-up reflected longer waiting periods to receive the tests, longer

waiting periods for the tests to be analyzed, and longer periods of time before results were reported to the patient. In these situations, the opinions of the UK advisors were prioritized because the utility elicitation study was planned to be conducted in the UK. All six advisors agreed that the final draft of the health states provided clear and accurate descriptions of typical experiences with false-positive screening results.

Ten health state vignettes were drafted, each describing a temporary cancer screening event. The first health state (health state A) described a true negative result (i.e., a patient is screened for cancer and receives a negative result). The other nine described a false-positive pathway in which a patient is screened for cancer, receives a positive result, completes follow-up procedures, and is then told that no sign of cancer was detected. The procedure of the initial screening test was not provided because these health states were designed to yield utilities that may be applicable to false-positives stemming from any screening test, including currently available single-cancer screening tests (e.g., mammogram, colonoscopy) and MCED tests.

Table 1 presents a list of all health states, with details on the false-positive result, the follow-up procedures, and the number of days of uncertainty prior to learning that the initial screening result was not confirmed by follow-up investigation. These health states were designed to cover a broad range of experiences, from simple to more complex and invasive follow-up procedures, as well as shorter and longer durations of uncertainty.

The health states were presented to respondents on individual cards, each with a series of bullet point descriptions organized into categories with headings intended to help the respondents understand the content: ‘cancer screening,’ ‘screening result,’ the name of each follow-up investigation (e.g., ‘CT scan,’ ‘MRI scan,’ ‘mammogram’), and ‘resolution.’ A timeline of key events within each false-positive pathway was presented at the bottom of each health state. The final health state text is presented in the electronic supplementary material (ESM).

2.3 Participants

Participants were recruited using digital social media marketing (e.g., via Facebook, Twitter, and Google). Interested participants were screened by phone for eligibility. To be eligible for this study, participants were required to be over 18 years of age, a UK resident, able to understand the assessments as judged by the investigator, able and willing to give electronic informed consent, and able to complete the protocol requirements. Because the sample was intended to match the demographics of the UK general population, there were no eligibility criteria for specific clinical characteristics,

Table 1 List of health states presented to participants

Health state ^a	Follow-up procedures ^b	Number of days of uncertainty about cancer diagnosis ^c
A. Cancer screening with negative result		
B1. False positive for lung cancer without head or neck involvement	CT scan	10
B2. False positive for lung cancer with possible head or neck involvement	CT scan; PET-CT	25
B3. False positive for lung cancer with a 6-month follow-up scan	CT scan	185
C. False positive for colorectal cancer	Colonoscopy	14
D1. False positive for breast cancer; no biopsy or MRI	Mammogram/ultrasound (at same visit)	10
D2. False positive for breast cancer; biopsy performed	Mammogram/ultrasound/biopsy (at same visit)	10
D3. False positive for breast cancer; MRI performed	Mammogram/ultrasound (at same visit); MRI	20
E1. False positive for pancreatic cancer; follow-up CT scan	CT scan	9
E2. False positive for pancreatic cancer; follow-up CT scan and PET-CT	CT scan; PET-CT	23

CT computed tomography, MRI magnetic resonance imaging, PET-CT positron emission tomography scan and computed tomography scan

^aThe ‘D health states’ describing false positives for breast cancer were not presented to the male participants. Therefore, men only valued seven health states, whereas women valued all 10 health states

^bEach health state described the typical follow-up procedures that would occur after a false-positive screening result. For some of the health states (e.g., B1, E1), a single follow-up procedure was necessary to confirm that the original result was a false positive. For other health states (e.g., B2, E2), multiple follow-up tests were required

^cEach health state included a timeline illustrating the number of days between the false-positive screening result and the eventual resolution when the result is confirmed to be a false positive. This is the period of uncertainty during which the hypothetical patient does not know whether the positive screening result was accurate

and efforts were made to mirror the UK population with regard to gender, age, racial/ethnic background, and rate of unemployment by implementing recruitment caps on various demographic groups based on UK census information. To maintain the safety of participants and interviewers during the COVID-19 pandemic, participants in the main study (which was conducted in person) were required to provide proof of vaccination and wear a face covering.

2.4 Pilot Study

A pilot study was conducted with 30 participants in the UK (mean [standard deviation (SD)] age: 45.8 [14.5] years; 50.0% female). A three-step approach was used: (1) participants were sent a package with paper copies of all materials required for the interview (i.e., health states, background information, and questionnaires); (2) one-on-one interviews were conducted by Microsoft Teams videoconference; (3) the principal investigator and/or project manager were available to join the interview and assist whenever requested by the interviewer.

Participants completed the virtual TTO valuation and provided feedback on the health states and procedures. The pilot study was conducted in three phases (11 participants in phase 1, eight in phase 2, and 11 in phase 3) to allow for edits to the health states and expert consultation after phases

1 and 2. Participants in all phases reported a good general understanding of the health states. The health states and procedures were edited for clarity and ease of understanding. Data from the 30 pilot study participants were not included in the main analysis sample.

2.5 Utility Interview Procedures and Scoring

The health states finalized in the pilot study were used to assess preference and elicit health state utilities in the main study. Trained interviewers conducted one-on-one in-person interviews in private offices, following a semi-structured interview guide.

Participants were first introduced to the health states. Health state A was presented first followed by the groups of false-positive health states. The four groups of false-positive health states (i.e., the B, C, D, and E health states) were presented in random order. Each group was presented as a set. For example, when presenting the three lung cancer health states in group B, all three were presented on the table simultaneously, but the three were presented in random order. The health states were not presented with the organized lettering/numbering system. Instead, health states were labeled with different letters that would not provide any indication of the organizational structure or order of severity. For example, health states B1 and B2 were labeled as T and

Y. Interviewers reviewed every health state in detail with the participants, who were then given an opportunity to read the materials independently and ask questions. Participants were then asked to rank the health states from most preferable to least preferable.

After completing the ranking, participants valued the health states in a TTO task with a 1-year time horizon. TTO methods have been described extensively in previous publications [30]. For each of the health states, participants were offered a series of choices between spending a 12-month period in the health state versus spending varying amounts of time in full health. Choices were presented in a booklet with bars of varying length illustrating duration of the two choices. Time in full health varied in 1-month increments, alternating between longer and shorter periods of time in full health (i.e., 12 months, 0 months [dead], 11 months, 1 month, 10 months, 2 months...). Each health state received a utility value (u) on a scale with anchors of dead (0) and full health (1) based on the choice in which the respondent was indifferent between 12 months in the health state and x months in full health. The resulting utility estimate (u) is calculated as $u = x/12$.

Given the content of these health states, it was expected that few participants would perceive any of them to be worse than dead. However, when this happened, interviewers switched to the ‘lead-time’ approach, which introduces a lead period in full health. The duration of the lead period can vary but is typically equal to the amount of time spent in the health state (in this case, 1 year) [71, 72]. The combination of the conventional TTO for health states perceived to be better than dead and lead-time TTO for health states perceived to be worse than dead (often called ‘composite TTO’) has frequently been used in recent studies [73–75] and is currently the recommended protocol by the EuroQol group [72].

2.6 Statistical Analysis Procedures

Statistical analyses were conducted with SAS version 9.4. Descriptive statistics were used to summarize demographic data, health state preferences, and utilities. Categorical variables are summarized as frequencies and percentages, while means and standard deviations are reported for continuous variables.

Because health states differed only in their descriptions of the false-positive pathways, any difference in preference and utility among the health states can be attributed entirely to these false positives. The disutility of each false positive was calculated by subtracting the utility of health state A (true negative) from each health state describing a false positive.

Paired t-tests were conducted to examine differences between utility and disutility means (e.g., utility of health state A vs utility of health state B1), and independent t-tests

were used to test for subgroup differences in utilities by age (median split), gender, and employment status (employed vs not employed). Post hoc descriptive analyses were conducted to provide utilities for subgroups of patients categorized based on previous experiences with cancer screening and previous diagnosis of cancer.

3 Results

3.1 Sample Characteristics

A total of 225 participants were scheduled, and 208 participants attended interviews in the main utility valuation study. Five who attended were unable to fully comprehend the task and health states. Therefore, the analyses were conducted with a sample of 203 participants. Demographic characteristics and cancer screening history are reported in Table 2. The most commonly reported health conditions were depression (23.2%), anxiety (20.2%), diabetes mellitus (6.9%), hypertension (5.9%), and arthritis (5.4%).

3.2 Health State Rankings and Preferences

Health state rankings ranged from 1 (most preferable health state) to 10 (least preferable health state) for participants ranking breast cancer health states, and from 1 to 7 for participants who did not rank the breast cancer health states. All three nonbinary participants chose to rank the breast cancer health states. For both groups, almost all participants ranked health state A (true negative) as most preferable (98.0% and 99.0%, respectively). Of the false-positive health states, B1, D1, and E1 were commonly ranked as most or second most preferable. Health states B2, B3, and E2 were most commonly ranked among the least preferable.

After providing their rankings, participants were asked to explain their preferences. In general, reasons for preferences could be divided into three categories: (1) perceived severity of cancer diagnosis (e.g., pancreatic cancer was often perceived to be more severe than the other types of cancer); (2) duration of uncertainty regarding cancer diagnosis (e.g., some false-positive pathways resolve in 10 days while others take 3 weeks or longer); and (3) aspects of the follow-up testing (e.g., many participants viewed colonoscopy as an unpleasant procedure). Examples of quotations from participants describing these preferences are presented in Table 3.

3.3 Health State Utilities

Health state utilities are presented in Fig. 1. The utility of the ‘true negative’ health state was 0.96 (SD 0.07). Among the false-positive health states, D1 (false positive for breast

Table 2 Sample characteristics (*N* = 203)

Characteristics	Descriptive statistics (<i>N</i> = 203)
Age, mean (standard deviation) [years]	42.0 (13.8)
Gender, <i>n</i> (%)	
Male	101 (49.8)
Female	99 (48.8)
Nonbinary	3 (1.5)
Ethnicity, <i>n</i> (%)	
White	162 (79.8)
Black	8 (3.9)
Asian	18 (8.9)
Mixed	11 (5.4)
Other	4 (2.0)
Marital status, <i>n</i> (%)	
Single	110 (54.2)
Married	46 (22.7)
Other	47 (23.2)
Employment status, <i>n</i> (%)	
Full-time work	118 (58.1)
Part-time work	32 (15.8)
Other	53 (26.1)
Education level, <i>n</i> (%)	
No university degree	62 (30.5)
University degree or higher	141 (69.5)
Ever been screened for cancer ^a , <i>n</i> (%)	
Lung	2 (1.0)
Colorectal	13 (6.4)
Breast	24 (11.8)
Cervical	36 (17.7)
Pancreatic	5 (2.5)
Other ^b	15 (7.4)
None	134 (66.0)
Ever been diagnosed with cancer, <i>n</i> (%)	
Yes	7 (3.4)
No	196 (96.6)
If yes, type of cancer, <i>n</i> (%)	
Breast	3 (42.9)
Other ^c	4 (57.1)
If yes, time since diagnosis (years)	
Mean (standard deviation)	11.2 (9.2)
Median [range]	11.9 [1.5–23.0]
If yes, current cancer status, <i>n</i> (%)	
In remission	3 (42.9)
Cured (excised/surgically removed)	1 (14.3)
Clear	1 (14.3)
Stage 4/Advanced breast cancer with brain and spine metastases	1 (14.3)
Not reported	1 (14.3)

^aNot mutually exclusive

^bOther cancer screenings include 'brain' (*n* = 1), 'gynecological (uterus)' (*n* = 1), 'lymphoma' (*n* = 1), 'leukemia' (*n* = 1), 'liver' (*n* = 1), 'lymph nodes' (*n* = 1), 'prostate' (*n* = 3), 'skin' (*n* = 4), 'testicular' (*n* = 1), and 'ovarian' (*n* = 1)

^cOther cancer diagnoses include 'chronic myeloid leukemia (CML)' (*n* = 1), 'lymphoma—AIDS related' (*n* = 1), and 'skin' (*n* = 2)

cancer; no biopsy) had the highest (i.e., most favorable) utility (0.93, SD 0.06), suggesting that this was perceived to be the least aversive of the false-positive pathways. B3 (false positive for lung cancer with a follow-up CT scan at 6 months) had the lowest utility (0.85; SD 0.14), suggesting that this tended to be the least preferred false-positive pathway. For all but two participants, utilities of all false-positive health states were less than or equal to the utility of the true negative health state. Both of these participants stated that they would experience a feeling of security from receiving the negative results from the investigations performed to rule out the false positive.

The proportion of participants who did not trade any time in the TTO task (resulting in a utility of 1 [i.e., preference equal to 12 months in full health] or 0.958 [i.e., preference <12 months in full health, but >11 months in full health]) varied by health state, ranging from 89.7% for health state A to 46.3% for health state E2. A total of 77 participants perceived health state A (true-negative cancer screening) to be equal in preference to full health, resulting in a utility of 1. The false-positive health states usually received utilities that were <1 because participants typically preferred full health over the false-positive experiences. The false-positive health states had utilities of 1 for the following numbers of participants: D3, *n* = 0; D2 and E2, *n* = 2; B2, *n* = 3; B3, C, and D1, *n* = 4; E1, *n* = 6; B1, *n* = 7.

No participants valued any health states as equal to dead (i.e., a utility of 0). Only one participant perceived any health state to be worse than dead, and only for health state B2, which described a false positive for lung cancer with possible head/neck involvement. This person reported having a particular sensitivity to this health state because multiple members of her family had died of brain cancer.

The disutility (i.e., utility decrease) associated with each false-positive pathway was calculated by subtracting the utility of health state A (cancer screening without a false positive) from the utility of each health state with a false-positive scenario. Disutilities of false-positive pathways ranged from −0.031 to −0.111 (Table 4). Based on paired t-tests, the utility values for all false-positive health states were significantly lower than the utility of health state A (all *p* < 0.0001). Health states describing different follow-up pathways for the same type of cancer (e.g., the three lung cancer health states) were also compared with each other. Significant differences were found for all comparisons except for health state B2 versus B3 (*t* = −1.0; *p* = 0.299) and health state D2 versus D3 (*t* = −1.6; *p* = 0.123).

Paired t-tests were also conducted to examine factors contributing to differences in preference between the health states. Statistically significant utility differences were found between health states varying by invasiveness of follow-up procedures, duration of uncertainty regarding the eventual diagnosis, and perceived severity of the suspected cancer

Table 3 Participant quotes justifying ranking preferences

Common reasons for preferring some health states over others	Selected quotations
Perceived severity of cancer diagnosis	<p>“Colon cancer is really scary, my grandmother died of it.”</p> <p>“I have a terrible fear of [colorectal] cancer.”</p> <p>“My aunt died of lung cancer; I’d automatically think you’d die.”</p> <p>“Anything pancreatic I’d be very concerned about...from people that you hear who’ve had it and died, it just seems really horrible. The kind of cancer is a big factor for me.”</p> <p>“Breast cancer has the best survival rate.”</p> <p>“Colorectal cancer is far less aggressive than pancreatic cancer.”</p> <p>“My personal knowledge of breast cancer [is that] it’s more treatable, not so life threatening. I know a lot of women who have had breast cancer and survived. But pancreatic cancer is quite an aggressive cancer and might be terminal, and I know a friend who died a couple of months from diagnosis.”</p> <p>“I think lung cancer is one that’s really hard to come back from.”</p> <p>“Pancreatic is one of the worst cancers that someone can have.”</p> <p>“I would be really concerned about pancreatic cancer.”</p>
Duration of uncertainty regarding cancer diagnosis	<p>“It’s better to do the tests all at once rather than spread out.”</p> <p>“The actual testing itself isn’t important, it’s the not knowing, the turnaround time.”</p> <p>“If I have cancer, I don’t mind a painful test, but I want to know as quickly as possible.”</p> <p>“The longer the ordeal goes on, the more serious it feels. Psychologically, you would start to feel like a sick person.”</p> <p>“The longer you wait the more upset you will be.”</p> <p>“You worry when your testing can’t happen right away.”</p> <p>“It was based on the time it takes to get results, the duration of time you are in that limbo of indecision.”</p> <p>“I have been through [a false-positive diagnosis]. I had about three weeks waiting, and it was horrible.”</p> <p>“One of the things is the wait for the results, people are really anxious, my mother had to wait eight months before a lung cancer diagnosis.”</p> <p>“Thinking about how long you’re walking around with uncertainty and anxiety – I would be less worried about the type of procedure.”</p> <p>“I clinically suffer from anxiety and worry about uncertainty. I would be desperately worried until I got some result.”</p>
Aspects of follow-up testing	<p>“The colonoscopy seems quite scary, sounds worse than the CT scan. I’ve had a colonoscopy before, and they are awful.”</p> <p>“I hate needles and blood, so injections are avoided for me.”</p> <p>“Invasiveness of the procedures and the numbers of tests was how I made my decision.”</p> <p>“I know how bad the colonoscopy is.”</p> <p>“I have been through [a breast biopsy] and the biopsy hurt a lot.”</p> <p>“[Factors include] the invasiveness of the procedures and if it’s in a personal area.”</p> <p>“I was worried about exposing myself to radiation in the scans.”</p> <p>“Having a biopsy is a major intrusion.”</p> <p>“I preferred the ones that sounded less painful, less invasive, and relatively short.”</p> <p>“I based it on what I felt was invasive and uninvase.”</p> <p>“The [colonoscopy] prep sounds quite bad and the test itself is quite invasive as well.”</p>

CT computed tomography

type. For example, health states D1 and D2 described false-positive breast cancer screenings that were identical other than the addition of a biopsy in D2. D2 had a significantly lower utility (0.91 vs 0.93; $p < 0.001$). Health states B1 and B3 described false-positive lung cancer screenings that were resolved with a single CT scan, but differed in duration of uncertainty. B3 with a follow-up scan at 6 months (i.e., 185 days of uncertainty) had a significantly lower utility than health state B1 with a follow-up at 10 days (i.e., 10 days of uncertainty) (0.85 vs 0.92; $p < 0.001$). Health states B1 and E1 were similar in testing procedures and duration of uncertainty, but differed in type of cancer (i.e., lung vs pancreatic cancer). The type of cancer that tended to be perceived as

more severe (i.e., pancreatic) was associated with a lower utility (0.92 vs 0.91; $p = 0.05$).

3.4 Group Comparisons

There were no significant between-group differences in utilities or disutilities by age, gender, or employment status, except for the disutility of health state C (false positive for colorectal cancer), which had a gender difference. For this health state, female participants had a significantly larger disutility (-0.063 [0.098] vs -0.096 [0.115]; $p = 0.028$).

When comparing utilities and disutilities of participants who had previously been screened for cancer versus those

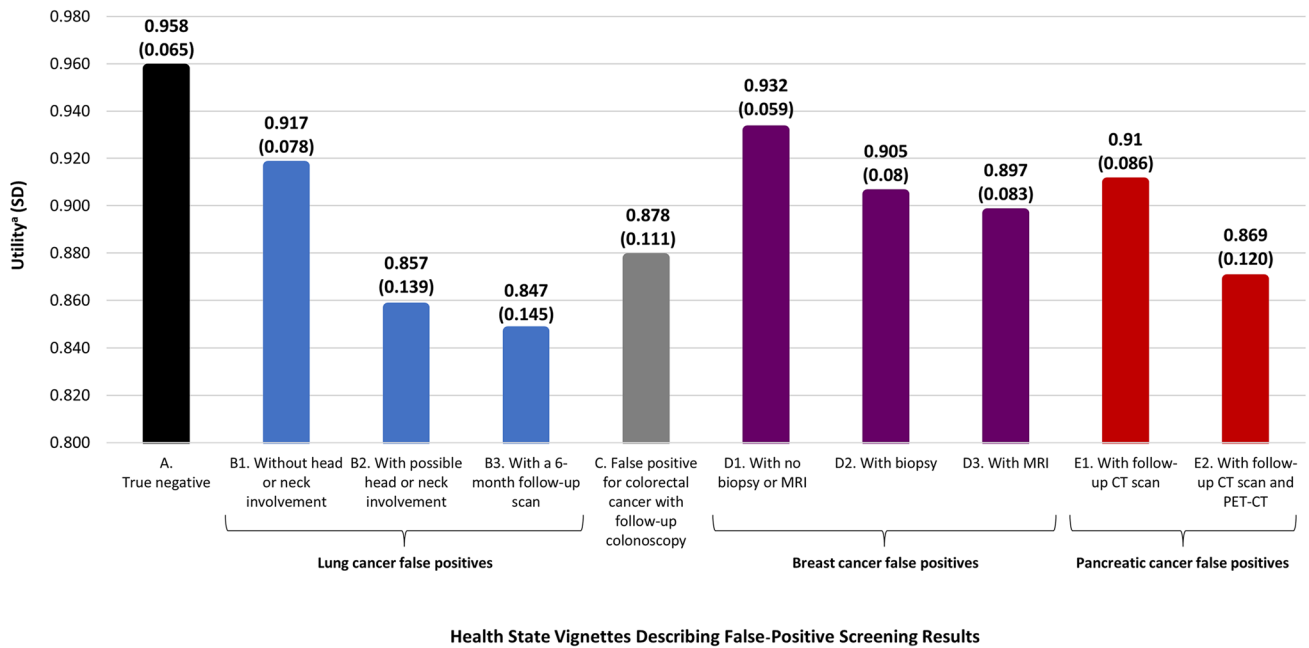


Fig. 1 Health state utilities. ^aTTO scores are on a scale anchored with 0 representing dead and 1 representing full health. *CT* computed tomography, *MRI* magnetic resonance imaging, *PET-CT* posi-

tron emission tomography scan and computed tomography scan, *SD* standard deviation, *TTO* time trade-off

who had not, the only statistically significant difference was for the utility of health state A representing a negative screening, but the between-group difference was relatively small (difference = 0.018; *p* = 0.026). No significant differences in utilities or disutilities were observed between participants who have/had cancer versus those who have not, although only seven individuals had a personal history of cancer.

4 Discussion

Results of this study add to the limited body of literature suggesting that there is a disutility associated with false-positive cancer screening results. Whereas previous studies typically provided a single utility estimate for a breast cancer false positive [32–35], the current study design allows for comparisons between a range of false-positive pathways.

Table 4 Disutilities of false-positive cancer screening^{a,b}

Health state	<i>N</i>	Mean disutility	<i>SD</i>	95% <i>CI</i>
B1. False positive for lung cancer without head or neck involvement	203	– 0.041	0.061	– 0.049 to – 0.032
B2. False positive for lung cancer with possible head or neck involvement	203	– 0.101	0.137	– 0.120 to – 0.082
B3. False positive for lung cancer with a 6-month follow-up scan	203	– 0.111	0.141	– 0.130 to – 0.091
C. False positive for colorectal cancer	203	– 0.079	0.108	– 0.094 to – 0.064
D1. False positive for breast cancer; no biopsy or MRI	102	– 0.031	0.049	– 0.041 to – 0.021
D2. False positive for breast cancer; biopsy performed	102	– 0.058	0.079	– 0.073 to – 0.042
D3. False positive for breast cancer; MRI performed	102	– 0.067	0.083	– 0.083 to – 0.050
E1. False positive for pancreatic cancer; follow-up CT scan	203	– 0.048	0.072	– 0.058 to – 0.038
E2. False positive for pancreatic cancer; follow-up CT scan and PET-CT	203	– 0.088	0.115	– 0.104 to – 0.073

CI confidence interval, *CT* computed tomography, *MRI* magnetic resonance imaging, *PET-CT* positron emission tomography scan and computed tomography scan, *SD* standard deviation, *TTO* time trade-off

^aTTO scores are on a scale anchored with 0 representing dead and 1 representing full health

^bDisutilities (i.e., utility decreases) for each false-positive pathway were calculated by subtracting the mean utility of health state A (true negative screening result) from the utility of each health state with a false-positive screening result

Results suggest that there are differences in preference and utility based on three characteristics of the false-positive screening and follow-up. Greater disutility was associated with more invasive follow-up diagnostic procedures, longer duration of uncertainty regarding the eventual diagnosis, and perceived severity of the suspected cancer type.

Due to the limitations of previous research and variations in methodology, it is not straightforward to compare current results with previously reported disutilities of false positives. Still, it is possible to consider the results of a systematic review that found utilities of false positives only in the context of breast cancer [31]. Across five studies reported in this review, disutilities varied from 0 to 0.26. The current study found disutilities ranging from 0.031 to 0.111 across all cancer types and 0.031 to 0.067 for breast cancer false positives. Values from the current study are within the range reported by the review.

Between the nine false-positive pathways examined in this study, the greatest disutility was for the health state where a follow-up CT scan was required months after the initial screening. Although the health state vignette specified that the spot on the lung detected during initial screening was “unlikely to be cancer,” many respondents reported that they would be uncomfortable living with the uncertainty for 6 months. There was also a substantial disutility associated with suspected pancreatic cancer because this type of cancer was perceived to be particularly serious. Follow-up investigations perceived to be aversive or invasive also resulted in notable utility decreases, including colonoscopy and biopsy. Quotations in Table 3 demonstrate respondents’ perceptions underlying these utility decreases.

The mean disutilities listed in Table 4 may be useful for representing the impact of false-positive screenings in cost-utility modeling that is conducted to assess the value of a range of screening methods, including single-cancer screening and MCED tests. When using these values, modelers must consider the time horizon of the current TTO elicitation. Each false-positive health state (i.e., the B, C, D, and E health states) described a year in which a false-positive pathway occurred. The true negative health state (i.e., A) to which they were compared described a year without a false-positive pathway. Therefore, the disutilities in Table 4 represent the impact of the false-positive pathway on a year, and each disutility can be used in a model as a QALY decrement. For example, to model a hypothetical patient who had a false-positive breast cancer screening with no biopsy or MRI, the disutility of -0.031 for health state D1 can be applied as a one-time QALY decrement.

One inherent limitation of all vignette-based utility elicitation studies is that the resulting utilities are based on perceptions of the health states rather than personal experience [36]. However, it may not be feasible to use more standardized methods, like the EQ-5D or other generic

instruments, to estimate the disutility of patients experiencing a false-positive pathway. These are relatively brief events, and patients do not know the diagnosis is a false positive until after follow-up testing has determined the screening result was not correct. Therefore, it would be challenging to identify and access these patients at the time the events are occurring. Furthermore, the false-positive pathways do not consist of a chronic unchanging experience. These pathways change over time, beginning with the impact of learning of the suspected cancer, followed by a varied set of follow-up procedures, leading to the eventual resolution. It does not seem feasible to assess the utility of a patient at each of these time points. In contrast, with the path state approach in the current study, it was possible to estimate the utility impact of the entire false-positive pathway. In cost-utility models, the resulting values would need to be used to represent the path, rather than each individual component of the path.

Another limitation of the vignette-based approach is that the health states cannot possibly represent the wide range of experiences with false-positive screenings. For example, there is variability in the amount of time patients are required to wait for their scheduled follow-up investigations, and more extended time periods are likely to be associated with greater anxiety. It is likely that some patients experience follow-up durations that are substantially different from the timelines presented in the health states. Therefore, these health states and resulting utilities may not be applicable to all patients who have false-positive screenings. The timelines were developed with the intention of representing a typical patient experience based on published literature, cancer screening guidelines, and the input of cancer screening experts. Still, because patient experiences vary in ways that cannot be captured in a vignette, results should be interpreted with appropriate caution.

A possible limitation of the health state development process should also be considered. The vignette content was developed primarily based on the input of clinical advisors rather than patients. Clinical advisors were consulted for this information because the health states were designed to describe objective medical information (e.g., testing procedures, typical duration of uncertainty about diagnosis). Study respondents valued the health states based on their expectations of how they might feel in these situations. Unlike most vignette-based utility studies with health states focusing on disease severity, the current health states do not describe subjective patient experiences such as symptoms. Therefore, people who had experienced false-positive screening results were not recruited to provide input on their subjective experience. However, it is possible that people who had received a false-positive diagnosis may have highlighted aspects of the false-positive experience that were not reported by the clinicians.

The COVID-19 pandemic resulted in some challenges with data collection. Due to the pandemic, the study was conducted in a single location to minimize the possibility of COVID exposure for the study team while traveling to additional locations. Also, to minimize risk of COVID transmission, interviewers and participants were required to have been vaccinated and to wear masks during interviews. This could have resulted in a biased sample selection because the sample does not include people who chose not to be vaccinated or refused to wear a mask. It is unknown whether this exclusion would have provided systematically different health state utilities.

It should be noted that this general population sample reported an unexpectedly high rate of depression (23.2%) and anxiety (20.2%). It is possible that levels of these symptoms were elevated due to the ongoing pandemic. Because these conditions were reported via checkboxes on a participant-completed form rather than via a thorough assessment, it is not known whether participants were truly reporting clinically relevant depression or anxiety. It is possible that they interpreted these terms in a more colloquial way, rather than indicating that they met criteria for a formal diagnosis of depression or anxiety.

Overall, the current study is an important step forward for economic analysis of cancer screening procedures. A wide range of novel cancer screening tests are in development, and CUAs will be needed to examine and compare their value. The utilities estimated in this study will be useful for representing the impact of false-positive screening results in these analyses.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s41669-023-00443-w>.

Acknowledgements The authors would like to thank Carly Brown, Dave Watkins, and Stephanie Foy of Liberating Research for assistance with participant recruitment; Matthew Laudie and Jafi Lipson for clinical consultation during health state development; Kristen Deger, Sonya Stanczyk, Katie Stewart, Nashmel Sargalo, Marissa Stefan, Gabriela Fernandez, Natalia Pigłowska, and Sasha Suleyman for assistance with data collection; Robyn Cyr for statistical programming; and Amara Tiebout and Fritz Hamme for editorial support.

Declarations

Data Availability Data will be shared upon reasonable request.

Ethics Approval The study protocol was approved by an institutional review board (Ethical and Independent Review Services; Study 21017-01).

Consent to Participate Informed consent was obtained prior to each interview.

Consent for Publication Informed consent included a statement regarding publication of de-identified data in journals.

Funding This study was funded by GRAIL, LLC, a subsidiary of Illumina Inc. currently held separate from Illumina Inc. under the terms of the Interim Measures Order of the European Commission dated 29 October 2021.

Competing Interests LM, TH, and HK are employed by Evidera, a company that received funding from GRAIL for time spent conducting this research. KC and ETF are employees of GRAIL, LLC, a subsidiary of Illumina, Inc., with equity in Illumina, Inc. MS is a paid consultant to GRAIL. AH is co-investigator for an academic study (SUMMIT) sponsored by UCL, which is funded by GRAIL, has received one honorarium for an advisory board meeting for GRAIL and an honorarium from Evidera Inc (for this current project), and previously owned shares in Illumina. LN has stock options in Culmination Bio. SJ has received fees for advisory board membership in the last three years from Astra-Zeneca, Bard1 Lifescience, and Johnson and Johnson. He has received grant income from Owlstone and GRAIL Inc. He is an unpaid member of a GRAIL advisory board. He has received lecture fees for academic meetings from Cheisi and Astra Zeneca. His wife works for Astra Zeneca.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were directed by LM and TH. The first draft of the manuscript was written by LM and TH. The other authors provided edits and comments on the first draft. All authors read and approved the final manuscript.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.


References

1. Hiom SC. Diagnosing cancer earlier: reviewing the evidence for improving cancer survival. *Br J Cancer*. 2015;31(112 Suppl 1):S1-5.
2. McPhail S, Johnson S, Greenberg D, Peake M, Rous B. Stage at diagnosis and early mortality from cancer in England. *Br J Cancer*. 2015;31(112 Suppl 1):S108-15.
3. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(3):145-64.
4. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, National Lung Screening Trial Research Team, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011;365(5):395-409.
5. Krist AH, Davidson KW, Mangione CM, Barry MJ, Cabana M, US Preventive Services Task Force, et al. Screening for lung cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2021;325(10):962-70.

6. Bosetti C, Levi F, Rosato V, Bertuccio P, Lucchini F, Negri E, et al. Recent trends in colorectal cancer mortality in Europe. *Int J Cancer*. 2011;129(1):180–91.
7. Lin JS, Perdue LA, Henrikson NB, Bean SI, Blasi PR. Screening for colorectal cancer: updated evidence report and systematic review for the US Preventive Services Task Force. *JAMA*. 2021;325(19):1978–98.
8. Davidson KW, Barry MJ, Mangione CM, Cabana M, Caughey AB, US Preventive Services Task Force, et al. Screening for colorectal cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2021;325(19):1965–77.
9. Nahand JS, Taghizadeh-Boroujeni S, Karimzadeh M, Borran S, Pourhanifeh MH, Moghoofei M, et al. microRNAs: new prognostic, diagnostic, and therapeutic biomarkers in cervical cancer. *J Cell Physiol*. 2019;234(10):17064–99.
10. Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, US Preventive Services Task Force, et al. Screening for cervical cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2018;320(7):674–86.
11. Siu AL. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med*. 2016;164(4):279–96.
12. van den Ende C, Oordt-Speets AM, Vroiling H, van Agt HME. Benefits and harms of breast cancer screening with mammography in women aged 40–49 years: a systematic review. *Int J Cancer*. 2017;141(7):1295–306.
13. Chen M, Zhao H. Next-generation sequencing in liquid biopsy: cancer screening and early detection. *Hum Genom*. 2019;13(1):34.
14. Tanos R, Thierry AR. Clinical relevance of liquid biopsy for cancer screening. *Transl Cancer Res*. 2018;7:S105–29.
15. World Health Organization. Cancer. 2022 [cited 2022 June 14]; <https://www.who.int/news-room/fact-sheets/detail/cancer>. Accessed Jun 15, 2022.
16. Yassin MM, Peel AL, Thompson WD, Patton J, Ashton V, Leaper DJ. Does screen-detected breast cancer have better survival than symptomatic breast cancer? *Asian J Surg*. 2003;26(2):101–7.
17. Hall IJ, Tangka FKL, Sabatino SA, Thompson TD, Graubard BI, Breen N. Patterns and trends in cancer screening in the United States. *Prev Chronic Dis*. 2018;26(15):E97.
18. di Pietro M, Modolell I, O'Donovan M, Price C, Pilonis ND, DeBiram-Beecham I, et al. Use of Cytosponge as a triaging tool to upper gastrointestinal endoscopy during the COVID-19 pandemic. *Lancet Gastroenterol Hepatol*. 2020;5(9):805–6.
19. Peled N, Fuchs V, Kestenbaum EH, Oscar E, Bitran R. An update on the use of exhaled breath analysis for the early detection of lung cancer. *Lung Cancer (Auckl)*. 2021;12:81–92.
20. Yang HY, Wang YC, Peng HY, Huang CH. Breath biopsy of breast cancer using sensor array signals and machine learning analysis. *Sci Rep*. 2021;11(1):103.
21. Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature*. 2019;570(7761):385–9.
22. Klein EA, Richards D, Cohn A, Tummala M, Lapham R, Cosgrove D, et al. Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set. *Ann Oncol*. 2021;32(9):1167–77.
23. 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(6499):1–11.
24. Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV, Consortium C. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. *Ann Oncol*. 2020;31(6):745–59.
25. Hackshaw A, Cohen SS, Reichert H, Kansal AR, Chung KC, Ofman JJ. Estimating the population health impact of a multi-cancer early detection genomic blood test to complement existing screening in the US and UK. *Br J Cancer*. 2021;125(10):1432–42.
26. Aravanis AM, Lee M, Klausner RD. Next-generation sequencing of circulating tumor DNA for early cancer detection. *Cell*. 2017;168(4):571–4.
27. Kim A, Chung KC, Keir C, Patrick DL. Patient-reported outcomes associated with cancer screening: a systematic review. *BMC Cancer*. 2022;22(1):223.
28. McGovern PM, Gross CR, Krueger RA, Engelhard DA, Cordes JE, Church TR. False-positive cancer screens and health-related quality of life. *Cancer Nurs*. 2004;27(5):347–52.
29. Rasmussen JF, Siersma V, Malmqvist J, Brodersen J. Psychosocial consequences of false positives in the Danish Lung Cancer CT Screening Trial: a nested matched cohort study. *BMJ Open*. 2020;10(6):e034682.
30. Brazier J, Ara R, Rowen D, Chevrou-Severac H. A review of generic preference-based measures for use in cost-effectiveness models. *Pharmacoeconomics*. 2017;35(Suppl 1):21–31.
31. Li L, Severens JLH, Mandrik O. Disutility associated with cancer screening programs: a systematic review. *PLoS ONE*. 2019;14(7):e0220148.
32. Bonomi AE, Boudreau DM, Fishman PA, Ludman E, Mohelnitzky A, Cannon EA, et al. Quality of life valuations of mammography screening. *Qual Life Res*. 2008;17(5):801–14.
33. Gerard K, Johnston K, Brown J. The role of a pre-scored multi-attribute health classification measure in validating condition-specific health state descriptions. *Health Econ*. 1999;8(8):685–99.
34. Johnston K, Brown J, Gerard K, O'Hanlon M, Morton A. Valuing temporary and chronic health states associated with breast screening. *Soc Sci Med*. 1998;47(2):213–22.
35. Tosteson AN, Fryback DG, Hammond CS, Hanna LG, Grove MR, Brown M, et al. Consequences of false-positive screening mammograms. *JAMA Intern Med*. 2014;174(6):954–61.
36. Matza LS, Stewart KD, Lloyd AJ, Rowen D, Brazier JE. Vignette-based utilities: usefulness, limitations, and methodological recommendations. *Value Health*. 2021;24(6):812–21.
37. Canadian Agency for Drugs and Technologies in Health. Guidelines for the Economic Evaluation of Health Technologies: Canada. Version 1.0. 4th ed. Ottawa, Canada; 2017.
38. National Institute for Health and Care Excellence. Process and methods guides: Guide to the methods of technology appraisal 2013 (PMG9). London, UK; 2013.
39. Pharmaceutical Benefits Advisory Committee. Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee. Australian Government DoH ed. Canberra, ACT; 2016.
40. Brennan VK, Dixon S. Incorporating process utility into quality adjusted life years: a systematic review of empirical studies. *Pharmacoeconomics*. 2013;31(8):677–91.
41. Higgins A, Barnett J, Meads C, Singh J, Longworth L. Does convenience matter in health care delivery? A systematic review of convenience-based aspects of process utility. *Value Health*. 2014;17(8):877–87.
42. Howell TA, Matza LS, Jun MP, Garcia J, Powers A, Maloney DG. Health state utilities for adverse events associated with chimeric antigen receptor T-cell therapy in large B-cell lymphoma. *Pharmacoecon Open*. 2022;6(3):367–76.
43. MacKeigan LD, O'Brien BJ, Oh PI. Holistic versus composite preferences for lifetime treatment sequences for type 2 diabetes. *Med Decis Making*. 1999;19(2):113–21.
44. Matza LS, Kim KJ, Yu H, Belden KA, Chen AF, Kurd M, et al. Health state utilities associated with post-surgical *Staphylococcus aureus* infections. *Eur J Health Econ*. 2019;20(6):819–27.

45. Matza LS, Paramore LC, Stewart KD, Karn H, Jobanputra M, Dietz AC. Health state utilities associated with treatment for transfusion-dependent beta-thalassemia. *Eur J Health Econ.* 2020;21(3):397–407.
46. Kuppermann M, Shiboski S, Feeny D, Elkin EP, Washington AE. Can preference scores for discrete states be used to derive preference scores for an entire path of events? An application to prenatal diagnosis. *Med Decis Making.* 1997;17(1):42–55.
47. Attema AE, Brouwer WB. On the (not so) constant proportional trade-off in TTO. *Qual Life Res.* 2010;19(4):489–97.
48. Matza LS, Boye KS, Feeny DH, Bowman L, Johnston JA, Stewart KD, et al. The time horizon matters: results of an exploratory study varying the timeframe in time trade-off and standard gamble utility elicitation. *Eur J Health Econ.* 2016;17(8):979–90.
49. Bevers TB, Anderson BO, Bonaccio E, Buys S, Daly MB, Dempsey PJ, et al. NCCN clinical practice guidelines in oncology: breast cancer screening and diagnosis. *J Natl Compr Canc Netw.* 2009;7(10):1060–96.
50. Canadian Task Force on Preventive Health Care. Recommendations on screening for lung cancer. *CMAJ.* 2016;188(6):425–32.
51. Canadian Task Force on Preventive Health Care. Recommendations on screening for colorectal cancer in primary care. *CMAJ.* 2016;188(5):340–8.
52. Fletcher SW, Elmore JG. Clinical practice. Mammographic screening for breast cancer. *N Engl J Med.* 2003;348(17):1672–80.
53. Klapman J, Malafa MP. Early detection of pancreatic cancer: why, who, and how to screen. *Cancer Control.* 2008;15(4):280–7.
54. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition—summary document. *Ann Oncol.* 2008;19(4):614–22.
55. Roberts H, Walker-Dilks C, Sivjee K, Ung Y, Yasufuku K, Hey A, et al. Screening high-risk populations for lung cancer: guideline recommendations. *J Thorac Oncol.* 2013;8(10):1232–7.
56. Winawer SJ, Fletcher RH, Miller L, Godlee F, Stolar MH, Mulrow CD, et al. Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterology.* 1997;112(2):594–642.
57. Wolf AMD, Fontham ETH, Church TR, Flowers CR, Guerra CE, LaMonte SJ, et al. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. *CA Cancer J Clin.* 2018;68(4):250–81.
58. Zhang L, Sanagapalli S, Stoita A. Challenges in diagnosis of pancreatic cancer. *World J Gastroenterol.* 2018;24(19):2047–60.
59. American College of Radiology. Lung-RADS® Version 1.1. 2019a [cited 2022 June 15]; <https://www.acr.org/-/media/ACR/Files/RADS/Lung-RADS/LungRADSAssessmentCategoriesv1-1.pdf>. Accessed Jun 15, 2022.
60. American College of Radiology. Lung CT Screening Reporting & Data System (Lung-RADS). 2019b [cited 2022 June 15]; <https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Lung-Rads>. Accessed Jun 15, 2022.
61. National Comprehensive Cancer Network. NCCN Guidelines for Patients-Lung Cancer Screening. 2020 [cited 2022 June 15]; https://www.nccn.org/patients/guidelines/content/PDF/lung_screening-patient.pdf. Accessed Jun 15, 2022.
62. National Comprehensive Cancer Network. NCCN Guidelines for Patients - Pancreatic Cancer. 2021 [cited 2022 June 15]; <https://www.nccn.org/patients/guidelines/content/PDF/pancreatic-patient.pdf>. Accessed Jun 15, 2022.
63. American Cancer Society. Breast Ultrasound. 2022a [cited 2022 June 15]; <https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/breast-ultrasound.html>. Accessed Jun 15, 2022.
64. American Cancer Society. Breast MRI. 2022b [cited 2022 June 15]; <https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/breast-mri-scans.html>. Accessed Jun 15, 2022.
65. American Cancer Society. Tips for Getting a Mammogram. 2022c [cited 2022 June 15]; <https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/mammograms/mammograms-what-to-know-before-you-go.html>. Accessed Jun 15, 2022.
66. American Cancer Society. Mammogram Basics. 2022d [cited 2022 June 15]; <https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/mammograms/mammogram-basics.html>. Accessed Jun 15, 2022.
67. Harvard Health Publishing, Harvard Medical School. Preparing for a colonoscopy. 2020 [cited 2022 June 15]; <https://www.health.harvard.edu/diseases-and-conditions/preparing-for-a-colonoscopy>. Accessed Jun 15, 2022.
68. National Breast Cancer Foundation. Clinical Breast Exam. 2022 [cited 2022 June 15]; <https://www.nationalbreastcancer.org/clinical-breast-exam>. Accessed Jun 15, 2022.
69. Pancreatic Cancer Action Network. Endoscopic Ultrasound (EUS). 2021 [cited 2022 June 15]; <https://www.pancan.org/facing-pancreatic-cancer/diagnosis/endoscopic-ultrasound-eus/>. Accessed Jun 15, 2022.
70. Pancreatic Cancer Action Network. Computed Axial Tomography (CAT or CT) Scan. 2022 [cited 2022 June 15]; <https://www.pancan.org/facing-pancreatic-cancer/diagnosis/computed-tomography-ct-scan/>. Accessed Jun 15, 2022.
71. Janssen BM, Oppe M, Versteegh MM, Stolk EA. Introducing the composite time trade-off: a test of feasibility and face validity. *Eur J Health Econ.* 2013;14(Suppl 1):S5–13.
72. Oppe M, Rand-Hendriksen K, Shah K, Ramos-Goni JM, Luo N. EuroQol protocols for time trade-off valuation of health outcomes. *Pharmacoeconomics.* 2016;34(10):993–1004.
73. Jiang R, Shaw J, Muhlbacher A, Lee TA, Walton S, Kohlmann T, et al. Comparison of online and face-to-face valuation of the EQ-5D-5L using composite time trade-off. *Qual Life Res.* 2021;30(5):1433–44.
74. Joshi N, Hensen M, Patel S, Xu W, Lasch K, Stolk E. Health state utilities for acute myeloid leukaemia: a time trade-off study. *Pharmacoeconomics.* 2019;37(1):85–92.
75. Xie F, Pullenayegum E, Gaebel K, Bansback N, Bryan S, Ohinmaa A, et al. A time trade-off-derived value set of the EQ-5D-5L for Canada. *Med Care.* 2016;54(1):98–105.

Authors and Affiliations

Louis S. Matza¹  · Timothy A. Howell¹ · Eric T. Fung² · Sam M. Janes³ · Michael Seiden⁴ · Allan Hackshaw³ · Lincoln Nadauld⁵ · Hayley Karn⁶ · Karen C. Chung²

✉ Louis S. Matza
louis.matza@evidera.com

Timothy A. Howell
Timothy.Howell@evidera.com

Eric T. Fung
efung@grailbio.com

Sam M. Janes
s.janes@ucl.ac.uk

Michael Seiden
mvseiden@gmail.com

Allan Hackshaw
a.hackshaw@ucl.ac.uk

Lincoln Nadauld
Lincoln.Nadauld@imail.or

Hayley Karn
Hayley.karn@evidera.com

Karen C. Chung
kchung@grailbio.com

- ¹ Evidera, Bethesda, MD, USA
- ² GRAIL, LLC., a subsidiary of Illumina Inc., Menlo Park, CA, USA
- ³ UCL Respiratory, University College London, London, UK
- ⁴ Physician in Residence, GRAIL, LLC., Menlo Park, CA, USA
- ⁵ Intermountain Healthcare, Salt Lake City, UT, USA
- ⁶ Evidera, London, UK