University of Sheffield | School of Medicine & Population Health SCHARR

Cost-effectiveness of Prostate Cancer Screening for Men of Average and High Risk.

Date: 17/11/2025

Prepared for:

National Screening Committee

Project team: Phase 1 and 2

N	Name and surname	Project role
1	Ms Annabel Rayner (AR)	Modelling
2	Dr Jessica Forsyth (JF)	Statistics
3	Dr Maria Hanini (MH)	Modelling
4.	Dr Dan Pollard (DP)	Modelling
5.	Dr Daniel Brady (DB)	Programming support
6.	Mr Mark Clowes (MC)	Information search
7.	Mr Wael Mohammed (WM)	Modelling
8.	Dr Lena Mandrik (LM)	Modelling, project lead

Table of Contents

Abbreviations

ADT	Androgen Deprivation Therapy
BC	Breast Cancer
BRCA	Breast Cancer Gene
CI	Confidence Intervals
CRUK	Cancer Research UK
DRE	Digital Rectal Examination
EQ-5D	EuroQol - 5 Dimensions
GGG	Gleason Grade Group
HRQoL	Health Related Quality of Life
HSE	Health Survey for England
ICER	Incremental Cost Effectiveness Ratio
IMD	Indices of Multiple Deprivation
LYS	Life Years Saved
MP-MRI	Multi-Parametric Magnetic Resonance Imaging
NHD	Natural History Disease
NMB	Net Monetary Benefit
NPCA	National prostate cancer Audit
OC	Ovarian Cancer
ONS	Office for National Statistics
PC	Prostate Cancer
PCPT	PC Prevention Trial
PHE	Public Health England
PSA	Prostate-Specific Antigen
PSSRU	Personal Social Services Research Unit
QALYs	Quality-Adjusted Life Years
RR	Relative Risk
SACT	Systemic Anti-Cancer Therapy
TNM	Tumour, Node and Metastasis
TP	Transperineal

1. Lay Summary

PSA, or Prostate-Specific Antigen, is a protein made by the prostate gland. A PSA test is a simple blood test that measures the level of this protein. While high PSA levels can be a sign of prostate cancer, they can also be caused by other conditions. This is why PSA screening, although a useful tool, involves a complex discussion about its overall benefits and risks.

Screening for prostate cancer can help save lives, but it can also cause harm, such as overdiagnosis. This means finding and treating slow-growing cancers that would never have caused harm during a man's lifetime. Detecting these cancers can lead to unnecessary worry and treatment, as well as taking healthcare resources away from patients who would benefit more. The cost-effectiveness of screening varies across different population groups, depending on their risk and how aggressive their cancer is.

Regardless of the age or risk level of the men being screened, we found that complex screening approaches—such as testing men once and then inviting them to repeat screening after one or more years—did not improve cost-effectiveness. This is because PSA is not a very accurate predictor of a man's risk, and PSA levels naturally change with age.

For the general male population in England, screening may reduce deaths from prostate cancer, but this benefit comes with a high level of overdiagnosis. In this population group, PSA screening is unlikely to be cost-effective overall.

As with the general population, the cost-effectiveness of screening men with any family history of breast, prostate, or ovarian cancers is unlikely be cost-effective and also result in overdiagnosis.

Screening all men and men with a family history (which includes around one third of all men) would require a lot of resources. A large share of those resources would be spent on men who are overdiagnosed and therefore do not benefit from screening. These resources could be spent on providing better diagnostics and better treatment.

It is very uncertain whether screening men of Black ethnicity is cost effective, and it also carries the risk of overdiagnosis. This group would benefit from having

better data on how cancer develops, progresses, and what is expected participation rate in screening would be for this population group. Because this group is relatively small, the overall impact on national healthcare resources would be limited.

Screening men with confirmed BRCA gene mutations is cost-effective and comes with a high level of certainty. This remains true even if screening continues into older ages. While overdiagnosis is still a concern, the benefits of screening are clearer because prostate cancer in this group tends to be more aggressive. It's important to note that this study did not look at whether testing for the BRCA gene itself is cost-effective—that would require separate research and that the study assumed that screening, diagnostic, and treatment data retrieved from general population with cancer would be applicable to BRCA carriers as well.

2. Key Messages

Benefits vs. Harms of PSA Screening

PSA screening may help reduce prostate cancer deaths. However, in 40–50% of screen-detected cases, it identifies slow-growing cancers that would never have caused harm. This leads to unnecessary treatment for patients, may cause side effects and anxiety and also result in substantial costs for the healthcare system.

Comparison of Screening Strategies

One-off PSA screening generally generates less costs and more benefits per screening conducted than repeated screening. Strategies that involve screening once and then re-inviting only those with elevated PSA levels for further screening were not found not to be cost-effective in full incremental analyses.

Methodological and data issues

The natural history of prostate cancer is subject to considerable uncertainty, which strongly influences probabilistic results. There is substantial data uncertainty, specifically around populations of Black ethnicity and younger or older groups. Furthermore, the choice of reference case (i.e., the modelling methods) has a critical impact on cost-effectiveness outcomes and, consequently, on potential implementation decisions

Screening Men at General and Familial Risk

For both men at general risk and those with a family history of prostate cancer, screening is not cost-effective with the decision threshold of £20,000 per quality adjusted life year (QALY). Screening would also involve high resource use and many cases of overdiagnosis.

Screening Men of Black Ethnicity

The cost-effectiveness of screening Black men remains uncertain. This is because data used in the model for this population are either applied from populations at general risk (e.g. survival by age and stage) or are based on the US population (e.g. relative risk of aggressive cancer). Screening in this group may lead to greater reductions in mortality but still carries a risk of overdiagnosis. Its success will depend on high uptake and clear communication of risks. Given the relatively

small population size, the resource impact (and so the impact of screening on population mortality) would be limited. Decisions on screening in this group should be supported by better data.

Screening Men with BRCA Mutations

Screening men with a confirmed BRCA gene mutation is likely to be cost-effective under decision threshold of £20,000 per QALY. These men are at higher risk of developing aggressive cancers, making the benefits of screening clearer.
Although overdiagnosis is still a concern, the healthcare resources required to screen this group are relatively low.

However, this conclusion is reached by applying data on costs, utilities, and cancer survival by stage for BRCA carriers from the general-risk population. This is because there are no data on prostate cancer screening, diagnosis, and treatment in the BRCA population.

3. Background

In November 2023, SCHARR was approached by the UK National Screening Committee (NSC) to reassess the cost-effectiveness of prostate cancer screening for men at both average and high risk. The NSC sought strategic health economics support to advance its aim of reducing prostate cancer mortality by improving diagnosis, treatment, and prevention pathways.

Rapid advancements in prostate cancer detection and treatment have widened the gap between existing evidence and the realities of clinical and policy decision-making. This presents growing challenges for clinicians, commissioners, and policymakers alike.

In response, SCHARR proposed developing a bespoke health economics model to inform strategic decision-making by the NSC. This work leveraged SCHARR's experience with microsimulation models developed for other cancers, which were adapted to simulate the natural history of prostate cancer. The model was designed to generate policy-relevant evidence and support the prioritisation of future research in prostate cancer screening.

The initial work tasks were structured into two phases:

Phase 1: Adaptation and calibration of SCHARR's microsimulation model to reflect the natural history of prostate cancer. This phase focused on evaluating the cost-effectiveness of one-time PSA screening in men at average risk, and separately, a one-time screening for high-risk men, analysed as single-age cohorts.

Phase 2: 1) Assessment of the cost-effectiveness of PSA-based screening in men aged 45–70, with screening intervals stratified by individual risk profiles derived from initial PSA levels, age, ethnicity, and family history; 2) Evaluation of cost-effectiveness of Stokholm3 algorithm for men with the Charlson comorbidity score of below 3.

The Phase 1 report was submitted to the NSC in October 2024. Based on the findings from Phase 1 and feedback from the NSC expert group, the Phase 2 analysis was refined to address identified constraints and improve alignment with policy needs.

This Phase 2 report presents the full modelling methodology and describes the updates and enhancements made following the completion of Phase 1.

In Phase 2, the research team recalibrated the model to improve its fit to epidemiological data. The team then conducted analyses of single PSA screening at different ages and subsequently evaluated the cost-effectiveness of repeat PSA testing at the most cost-effective ages identified.

Next, the team assessed the cost-effectiveness of risk-stratified screening, in which invitations for repeat testing were determined by the PSA results from the initial screening round.

Finally, the research team explored the feasibility of modelling the Stockholm3 algorithm. However, accurate simulation of this intervention required more detailed data than were available in published sources. The team contacted the corresponding authors of the Stockholm3 trials (see the Conceptual Modelling section); however, due to confidentiality of Stockholm3 data and the imposed restrictions on sharing modelling outcomes, the UK NSC and the research team jointly decided to exclude Stockholm3 from the current evaluation.

4. Definitions of modelling outcomes

The model structure represents the natural history of prostate cancer, with additional modules for symptomatic diagnosis and screening diagnosis layered on top.

Parameters describing disease onset and progression are used as model inputs and are calibrated to current epidemiological data (see Calibration section). Stage- and age-specific survival are also included as model inputs (see Survival section).

The model produces outputs including cancer incidence, cancer-related deaths, and total deaths.

4.1. Predictions of benefits of screening

The model is a stage-shift model, meaning that the benefits of screening arise from detecting cancer earlier - before it progresses to more advanced stages with higher stage-specific mortality. In other words, screening leads to earlier diagnosis for some men in the model, resulting in improved survival, since late-stage cancers generally have lower survival rates than early-stage cancers.

Consequently, screening benefits are reflected through:

- A shift in the stage distribution at diagnosis (a higher proportion of early-stage cancers), and
- A reduction in prostate cancer deaths.

The model predicts:

- Life years saved (LYS)
- QALYs
- Distribution by stage and grade at diagnosis
- Number of deaths occurred over a specified time horizon.

4.2. Predictions of screening related harms

The model represents how prostate cancer develops over time and how screening can change its course by detecting cancer earlier - a concept known as a stage shift. Screening benefits some men by identifying cancer before it

progresses to more advanced, harder-to-treat stages. However, not everyone benefits equally. Some cancers grow so slowly that they would never have caused symptoms or death, and older men may die from other causes before cancer becomes life-threatening.

The model also captures the potential harms of screening, focusing on two main types:

- False positives, where the test suggests cancer is present when it isn't.
- Overdiagnosis, where real cancers are found but would never have caused harm during a man's lifetime.

Although the model does not simulate every individual's treatment and follow-up pathway after diagnosis, it assigns average costs based on the cancer stage, age, and time since diagnosis; utilities (i.e. HRQoL values) based on stage at diagnosis, and an annual probability to die from cancer based on age, stage at diagnosis, and time from diagnosis.

Importantly, the impact of cancer diagnosis on HRQoL captures the negative impact of screening and negative impact of cancer diagnosis in general. Men diagnosed with prostate cancer experience a reduction in HRQoL due to both the physical and emotional burden of the disease, as well as treatment-related side effects such as erectile dysfunction, urinary incontinence, fertility loss, pain, bleeding, infection, and other complications that reduce wellbeing.

As a result, some men experience lower quality of life without living any longer - illustrating that screening, while beneficial for some, can also cause harm for others.

False positive diagnosis

Screening may result in false-positive diagnoses, meaning that prostate cancer is diagnosed in men who, in reality, do not have cancer. In the model, men with false-positive diagnoses experience a decrement in HRQoL due to the psychological and physical impact of the diagnostic process, but they cannot die from prostate cancer.

Overdiagnosis

Not all cancers lead to cancer-related death. Some cancers would remain undetected throughout a patient's lifetime in the absence of screening, as they cause no symptoms or harm. These are defined as overdiagnosed cancers.

In the model, overdiagnosis is the model output and is estimated as:

(Incidence in screening arm – incidence in standard care arm)/number of screen -diagnosed cancers

This means that overdiagnosed cases in the model represent the proportion of all screen-detected cancers that would not have caused harm during a patient's lifetime, and they will include "true overdiagnosed cancer cases" and "false positive cases".

HRQoL decrements

Each individual in the model has an assigned HRQoL value derived from HSE data, reflecting age, ethnicity, indices of multiple deprivation, and other relevant factors.

- Men undergoing biopsy experience a temporary decrement in HRQoL to capture the negative impact of this invasive procedure.
- Men diagnosed with prostate cancer experience an additional HRQoL reduction, which is greater for more advanced disease stages.
- All men with diagnosed cancer receive an HRQoL decrement, regardless of whether the cancer is overdiagnosed.

For further details on the HRQoL impacts of screening, see Section 5.10: Utilities.

4.3. Predictions of resource use and costs in screening

The model also estimates:

- Screening costs,
- Diagnostic costs associated with follow-up of screen-positive cases, and
- Treatment costs for all diagnosed cancers.

Resource use predictions include:

- Number of PSA invitations and tests,

- Number of multiparametric magnetic resonance imaging (mpMRI) scans, and
- Number of biopsies required for follow-up of screen-positive cases.

4.4. Predictions of cost effectiveness of screening

Predictions of the cost-effectiveness of screening are primarily based on the incremental net monetary benefit (NMB), calculated as:
Incremental NMB = (Incremental QALYs × Threshold) − Incremental Costs
It is interpreted as follows:
Positive values (above 0): the intervention is cost-effective
Negative values (below 0): the intervention is not cost-effective
The decision thresholds of £20,000 per QALYs was used in the model. The choice of the threshold was based on the previous NICE evaluations of screening technologies with the assumptions that lower thresholds should be used in models with higher uncertainty, such as screening models.

Incremental NMB was selected as the primary health economic outcome instead of the incremental cost-effectiveness ratio (ICER) due to the high modelling uncertainty in the natural history disease models. This approach allows uncertainty to be incorporated not only in the mean values of the summary measures (i.e. mean ICERs and mean NMBs in probabilistic analyses) but also in their credible intervals, which is more straightforward to estimate for NMB than for ICER.

5. Summary of methodological changes in phase 2 work

This section provides a brief overview of the key methodological changes in modelling the natural history of prostate cancer and screening. For a more detailed description of the methods, please refer to the relevant sections of the report.

1. **Distribution of Prostate Cancer by Age and Stage**
   In Phase 1, age-specific incidence rates by stage were sourced from Public Health England, based on data from 2012–2014. In Phase 2, the incidence by stage was assessed by combining more recent data from NHS Digital (2021) and CMA Stage (2013–2021).

2. **Survival by Age and Stage**
   In Phase 1, the assumption of 100% survival for stages 1 and 2 hindered effective extrapolation beyond the 5-year period. In Phase 2, survival estimates for stages 1 and 2 were derived from the ProtecT trial, which provided 15 years of follow-up data. Cancer-specific mortality across all stages was extrapolated over a 70-year horizon.

3. **Gleason Grade Group (GGG) Distribution**
   In Phase 1, the GGG at cancer onset was based on GGG at the time of diagnosis. In Phase 2, data on GGG progression over time were used to recalculate the proportion of patients with GGG1 at the time of cancer onset.

4. **Costs**
   Costs were extrapolated over a lifetime to align with survival and surveillance extrapolations. Palliative care costs were attributed not only to prostate cancer deaths but also to deaths from other causes and were readjusted by age.

5. **Utilities by Stage**
   For calculating utility multipliers, Phase 2 used the recent EEPRU report by Alava (2022) as the reference population, replacing the Health Survey for England - HSE (2018) reference values. This approach aligns with

NHSD utility values reported for 2022. The utility multipliers were separated into those for the first year and subsequent years.

6. **Prostate-specific antigen (PSA) Accuracy**
   Unlike Phase 1, Phase 2 accounted for the age-related increase in PSA values in both non-cancer and cancer populations. In Phase 2, test positivity was determined by whether an individual's PSA level— modelled as a function of age and disease state— exceeded the PSA threshold.

7. **Knowledge of Breast Cancer Gene (BRCA) Status**
   In Phase 1, screening in BRCA1 and BRCA2 carriers assumed their BRCA status was known. Phase 2 introduced a probability model for the likelihood of BRCA status being known by a given age.

8. **Simulated population**
   For the Phase 2 analysis, the simulated population was based on a combined dataset from the HSE 2018 and 2019, rather than using the HSE 2018 population alone. Considering that some prostate cancer cases have a very long sojourn period, the age of the simulated population to reconstruct natural history was lowered from 30 years to 20 years to have a longer burn out period for the model.

6. Modelling methods

6.1. General approach to model development

The model development divided the approach into two parts: conceptual model development and mathematical model development.

6.1.1. Conceptual modelling

The conceptual modelling approach followed the guidance for economic evaluation of complex interventions and public health modelling by Squires (2016)[1] and used problem-structuring methods within a stakeholder participative process. Stakeholders were classified into three groups: people benefiting from the intervention (lay audience and patient representatives, "customers"), people involved in the delivery of the interventions and the system within which the intervention acts (general practitioners [GPs] and diagnostic experts, "actors"), and people with the power to approve or cancel the intervention (the UK National Screening Committee, "owners"). Representatives of all these groups were invited to the stakeholder meetings in May 2023 to conceptualise modelling of the natural history of prostate cancer.

The stakeholders who both accepted the invitations and participated in the stakeholder meetings in the Phase 1 work are listed in Table 1. The minutes from the stakeholders' meetings, along with the outputs and impacts (specifically detailing how stakeholder input was incorporated into the model), are provided in Supplementary I.

For the Phase 2 work, the iterative approach to conceptualisation was undertaken with the stakeholders who provided feedback on the core modelling assumptions indicated in Table 1. This feedback was mainly collected through email exchanges, but also through several one-to-one meetings when required. Additionally, several meetings were held with the developers of the Stockholm3 algorithm—Dr Vigneswaran, Dr Palsdottir, and Dr Govers—to discuss data access and the use of the algorithm in the modelling work. The Stockholm3 algorithm was excluded from the list of interventions due to limitations in the published data, as agreed during discussions with the National Screening Committee.

Table 1: Stakeholders informed the conceptual model for prostate cancer simulation.

N	Name and position	Phase 1	Phase 2
1	Prof. Hashim Ahmed, Chair in Urology (Clinical) Department of Surgery & Cancer – Faculty of Medicine at Imperial College London	Meeting	X
2	Prof. Jim Catto, NIHR Research Professor, Professor of Surgery at the University of Sheffield and an Honorary Consultant Urological Surgeon at Sheffield Teaching Hospitals NHS Trust.	Meeting	Meetings, emails
3	William Cross, Consultant Urologist, Leeds Teaching Hospitals NHS Trust, in Leeds	Meeting	X
4	Dr Helen Hanson, Joint Lead Consultant Cancer Genetics	Meeting	X
5	Amy Rylance, prostate cancer UK	Meeting	X
6	Natalia Norori, prostate cancer UK	Meeting	X
7	Prof Derek Rosario, Sheffield Teaching Hospital	X	Meetings, emails
8	Prof Adam Brentnall, Queen Mary University London	X	Emails
9	Prof Allan Hackshaw, Professor of Epidemiology & Medical Statistics at University College London	X	Emails
10	Prof. Chris Hyde, The University of Exeter		Meetings, emails
11	Dr Hari Vigneswaran, Karolinska Institute (for Stockholm3 algorithm)	X	Meetings, emails
12	Dr Thorgerdur Palsdottir, Karolinska Institute (for Stockholm3 algorithm)	X	Meetings, emails
13	Dr Tim Govers, CEO Medip Analytics (for Stockholm3 algorithm)	X	Meetings, emails

6.1.2. Mathematical modelling

The model is an individual-level model that simulates people with individual characteristics defining their risks. The model was parametrised by combining different parametrisation strategies: targeted literature searches, reviewing reference lists of the published models, stakeholder inquiries, and calibration (described separately below).

Model parametrisation

The initial targeted literature search for the model parameters in the Phase 1 was conducted by an information specialist and a project modeller (AR).

Following an initial scoping search for relevant clinical guidelines, targeted searches were conducted using MEDLINE and Embase to identify literature around prostate cancer (including synonyms such as neoplasms, tumours, carcinoma, etc.) in relation to each of the relevant risk factors and sub-groups (BRCA2, family history, ethnicity etc.)

Database searches used a combination of controlled vocabulary (MeSH/Emtree) and free text terms, with a focus on specificity over sensitivity; the aim was not to identify every relevant study, but to find the best available data for the purpose of model parameterisation.

Methodological filters were applied to the searches in order to prioritise review-level evidence; however, where no previous reviews were retrieved, or those available were judged to be of questionable quality, primary evidence was considered. No date or geographical limits were applied to the searches, but evidence was given preference on the basis of date, study size, and location.

Initial search results were sifted and selected by the information specialist. Key characteristics of candidate studies for possible inclusion were extracted to a spreadsheet for a final decision to be made by the modellers.

The other parameters (e.g., costs, utilities) were informed by a targeted literature search conducted by AR and verified by LM. For the Phase 2 work, additional data identification and extraction for missing values were carried out. The key parameter sources were presented to the stakeholders (Table 1) and updated based on their feedback.

Model coding

The mathematical model was constructed using R programming language (version 4.3.2). It utilised functions from previous SCHARR cancer models (MiMiC-Bowel and MiMiC-Blaky), supplemented with new code and adaptations to meet the model's specific requirements.

6.2. Model Structure

The model centres on the natural history disease (NHD) modules for prostate cancer, positing that certain risk factors influence cancer onset. Included risk factors are age, Black (higher risk) or Asian (lower risk) ethnicity, familial history of breast, ovarian, or prostate cancer, and being a BRCA1 or BRCA2 carrier.

At the start of each simulated cycle, the individual probability of cancer onset is estimated for each person based on these risk factors. Each person is then assigned to one of three mutually exclusive states: no cancer, cancer (i.e., those who had cancer onset), or non- prostate cancer death (i.e., other-cause mortality).

Individuals diagnosed with cancer are assigned a GGG (1, 2, 3, or 4/5) at the time of onset, influenced by factors such as age, ethnicity, and BRCA status. It is assumed that older age, Black ethnicity, and BRCA mutations are associated with an increased risk of being allocated a higher GGG at cancer onset (validated by the stakeholders during the Phase 1 work).

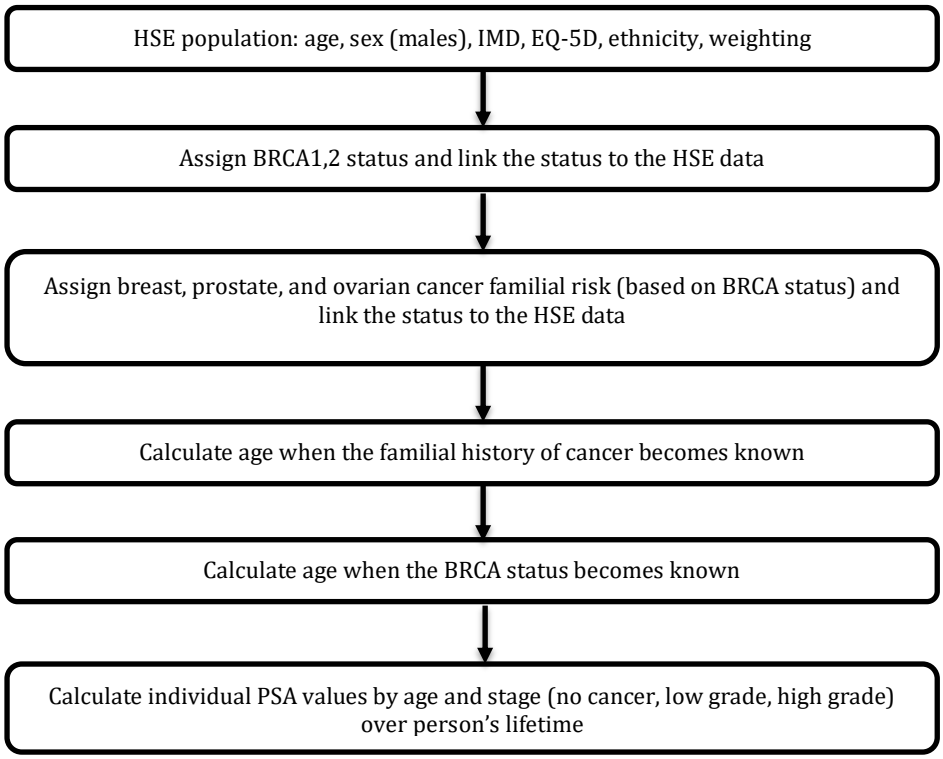
GGG affects the progression time of undiagnosed cancer through stages 1 to 4. The progression time is sampled from a Weibull distribution, with the assumption of perfect correlation—individuals who progress rapidly from stage 1 to 2 will similarly progress quickly through subsequent stages. The impact of GGG on progression time is modelled through multipliers, with values less than 1 for individuals with higher GGG compared to those with GGG1.

The next step is determining the stage of the disease and who receives a cancer diagnosis. In a non-screening scenario, patients receive a clinical diagnosis through symptomatic presentation, with diagnosis probabilities increasing in more advanced cancer stages. Diagnosed patients may die from cancer, and all individuals have a probability of dying from other causes. Death from undiagnosed cancer is assumed only for individuals with stage 4 cancer and aged 70 years or above. Clinically diagnosed patients follow a treatment pathway that includes costs, utility reductions, and reduced survival compared to the general population, with different stages having associated quality of life and costs.

The screening module integrates with the natural history model. Individuals diagnosed through screening follow a diagnostic pathway, and if they test positive, they incur annual treatment expenses and associated disutilities.

Finally, model outcomes are gathered, including costs, life-years saved (LYS), QALYs, resource use, and cancer cases. Half-cycle correction and discounting are applied to costs, QALYs, and LYS, and incremental results are estimated. The model operates in probabilistic and deterministic modes, which can run with different populations varied by demographic characteristics (age, BRCA mutation, ethnicity, and familial risk) to analyse the impact of the PSA test. The functional order of the modelling processes during each cycle is shown in Figure 1, and the NHD part of the model is illustrated in Figure 2.

Figure 1: Functional order of the modelling processes before simulating the population over the lifetime (a) and during the simulation over the lifetime (b)



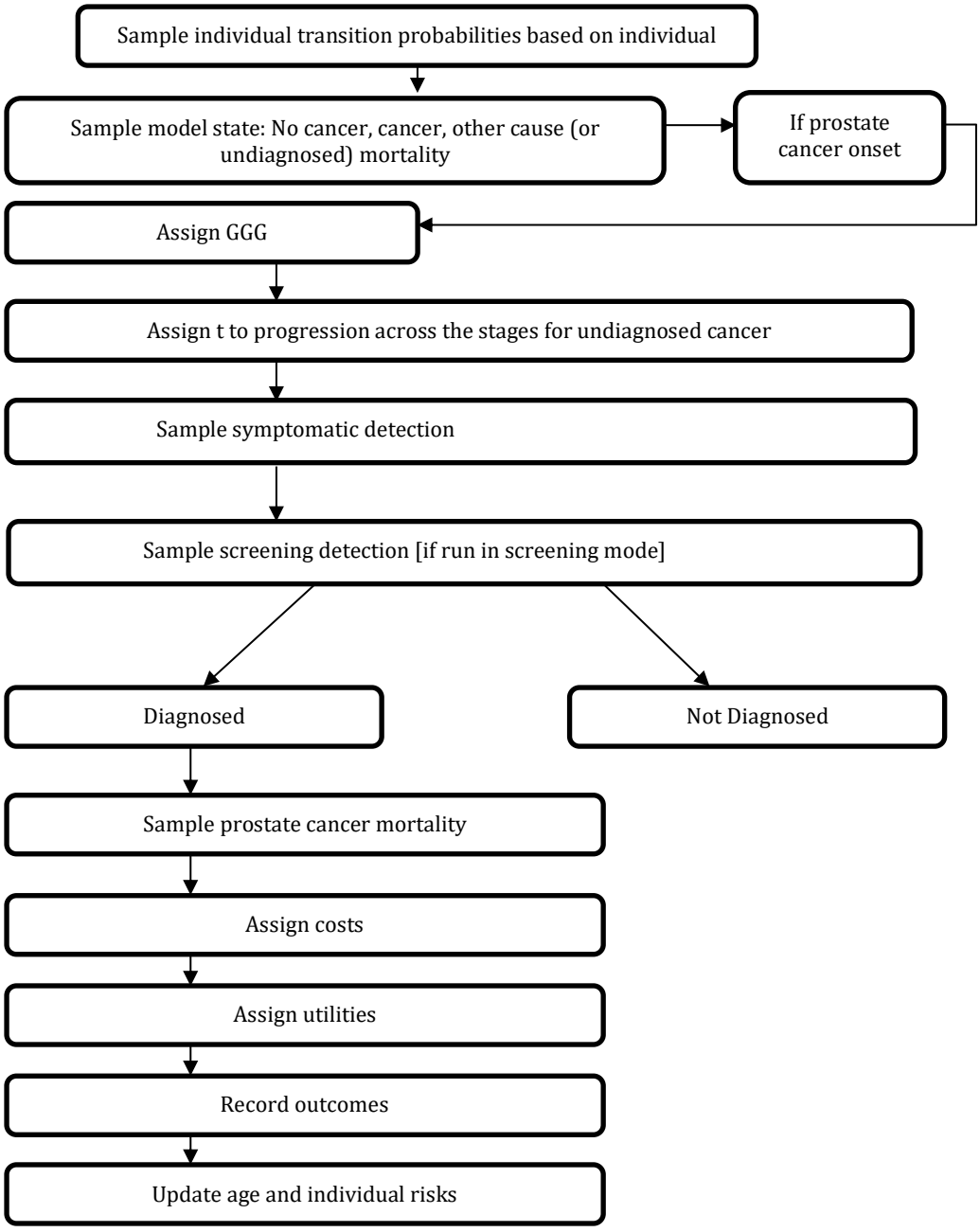
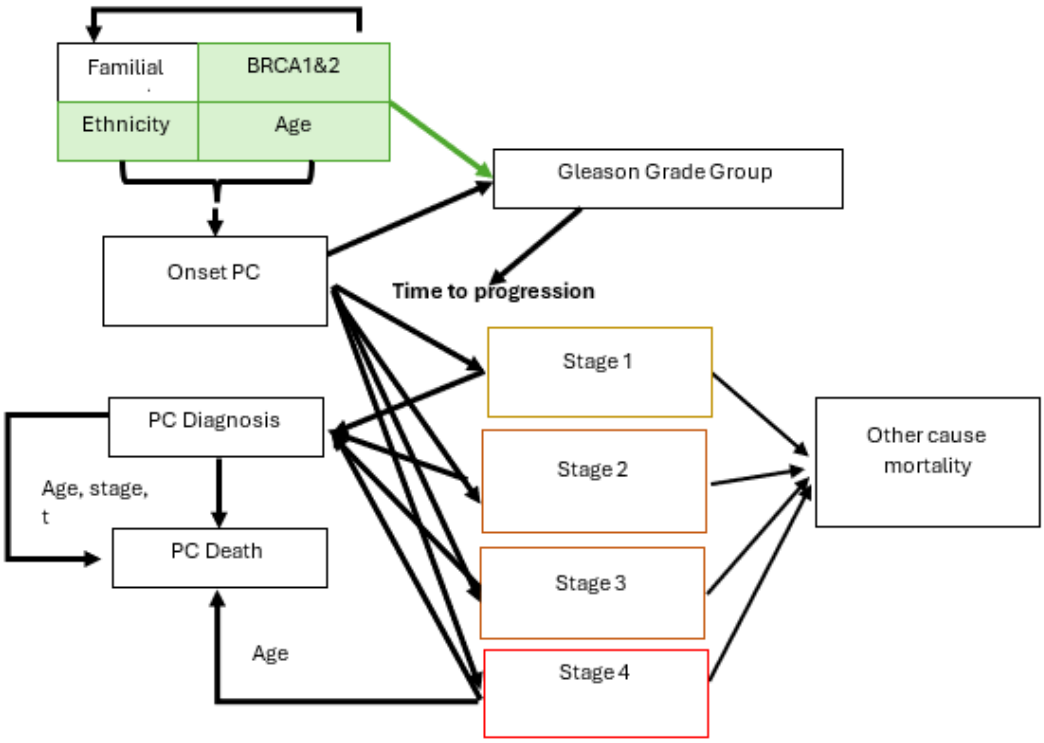

Sample individual transition probabilities based on individual
Sample model state: No cancer, cancer, other cause (or undiagnosed) mortality
If prostate cancer onset
Assign GGG
Assign t to progression across the stages for undiagnosed cancer
Sample symptomatic detection
Sample screening detection [if run in screening mode]
Diagnosed
Not Diagnosed
Sample prostate cancer mortality
Assign costs
Assign utilities
Record outcomes
Update age and individual risks

Figure 2: Structure of natural history disease in prostate cancer model



Legend: PC – Prostate Cancer
t – time since diagnosis
The risk factors in green impact both cancer onset and Gleason Grade Group allocation.

6.3. Model Population

6.3.1. Baseline phenotypic characteristics

The model baseline population is composed of individuals from the HSE 2018 and HSE 2019, an annual survey which is designed to provide a snapshot of the nation's health. Initially, the year 2018 was selected as it is the most recent dataset that included population baseline quality of life values. Individuals aged under 20 were excluded from the model. The individual phenotypic attributes extracted from HSE for use in the model included age, ethnicity, EuroQol – 5 Dimensions (EQ-5D-3L), and indices of multiple deprivation (IMD) quintile (a measurement of socioeconomic deprivation). The survey weights have been calculated by the HSE to enable adjustment of the sample so that it matches national population estimates of age, correcting for non-response, and thereby making the sample more representative of the English population.

6.3.2. Missing data for phenotypic characteristics

In the Phase 2 work, the HSE 2018 data were supplemented with HSE 2019, to provide a larger initial sample. The EQ-5D scores, which were not present in HSE 2019, were imputed from HSE 2018. Rather than imputing each aspect of health individually: mobility, self-care, usual activities, pain and anxiety, instead an individual's EQ-5D score was first derived and then this score was imputed.

Imputation is a method to replace missing data values with an estimate. The process used multiple imputation by chained equations (MICE) with the MICE package in R, using predictive mean matching as the imputation method. It works by predicting a value for the missing data point using a regression model, then identifying a few observed cases with similar predicted values, and randomly selecting one to replace the missing value. The estimates were based on large number of characteristics available in both HSE 2018 and 2019, including body mass index, smoking status, alcohol consumption, and IMD quintile. The mean EQ-5D scores between HSE 2018 and 2019 were compared to ensure that the imputation was realistic.

Values were missing for some of the variables in some individuals. There were a small number of missing values for ethnicity, as well as a number of missing EQ-5D values. These were both addressed in the imputation process. This process used the same HSE characteristics to impute the EQ-5D scores from HSE 2018 to 2019.

In addition to this, HSE 2018 and 2019 reported age in 5-year categories rather than as a continuous variable. To impute continuous age values, we used data from HSE 2014 (the last year with continuous age reporting) as a reference. Instead of randomly sampling from a uniform distribution within each age band, as was done in Phase 1, we used HSE 2014 to inform the imputation process in Phase 2. This process also involved creating a new variable in HSE 2014 of 5-year age categories based on the continuous HSE 2014 age variable to match those in later years. The characteristics used in the imputation were sex, IMD quintile, ethnicity, height, weight, and the categorical age variable. This allowed us to impute continuous ages for HSE 2018 and 2019, with checks to ensure they remained within their respective age categories.

6.3.3. Selecting the target population

All imputations for missing values were conducted before subsetting the HSE 2018 and 2019 population to the target population. Selecting the population by sex (males only) aged over 20 resulted in a sample of 6,217 men.

6.4. Risk Factors for cancer onset

The risk factors incorporated into the model include a range of variables, such as the presence of BRCA1 or BRCA2 mutations, familial predisposition to prostate, breast, or ovarian cancers, ethnicity, and age. The risk of prostate cancer onset at various ages is derived from calibration and is elaborated upon in the corresponding section below.

6.4.1. Ethnicity

Ethnicity data are available for the HSE population[2]. In the HSE, ethnicity is self-reported, meaning individuals identify their own ethnic background. The survey uses a standard set of nationally recognised categories, including:

- White (British, Irish, Gypsy or Irish Traveller, Other White)
- Mixed/Multiple ethnic groups (White and Asian, White and Black African, White and Black Caribbean, Other Mixed)
- Asian/Asian British (Indian, Pakistani, Bangladeshi, Other Asian)
- Black/Black British (Caribbean, African, Other Black)
- Other ethnic group (Chinese, Arab, Any other).

In the model, the Black population is defined in alignment with the HSE classification, as the model is based on HSE population characteristics. The model does not explicitly account for differential cancer risk among mixed ethnic groups, as the proportions of specific combinations (e.g., White and Asian, White and Black, or other) are not available.

Research conducted by Perez-Cornago et al. (2017)[3] using UK Biobank data indicates notable disparities in cancer incidence hazard ratios (HR) across ethnic groups. Specifically, compared to individuals of white ethnicity, those of Black ethnicity exhibited a HR of 2.61, 95% CI (2.10-3.24), while individuals of Asian ethnicity demonstrated a HR of 0.62, 95% CI (0.47-0.83)[3]. These HRs were

estimated using minimally adjusted Cox regression models. The ethnicity categories in that study correspond closely to those used in the current model, with the Black population defined separately from Mixed ethnic backgrounds.

The authors also reported an HR of 0.95 (95% CI: 0.55–1.64) for individuals of Mixed ethnicity. However, this subgroup represented only 0.7% of the study population, corresponding to just 13 prostate cancer cases. Given this very small sample size, the resulting uncertainty is substantial, and there is no evidence of elevated prostate cancer risk in this group. Therefore, the Mixed ethnicity category was not modelled separately in the current analysis.

6.4.2. BRCA1 or BRCA2 mutations

The HSE dataset does not contain genetic information.

Actual BRCA1,2 status allocation

Data on the frequency of BRCA1 and BRCA2 mutations were sourced from NHS England estimates, validated by consultant geneticist Dr Helen Hanson. Each individual in the HSE dataset was randomly assigned a binary category for BRCA1 or BRCA2 mutations (1 for having the mutation, 0 for not), based on the prevalence rates (1 in 381 for BRCA1 and 1 in 277 for BRCA2) and assuming independence of risk[4].

Case-control studies included in the review of Nyberg et all (2022), provide odds ratios (OR) for prostate cancer associated with these mutations (the study reports relative risk but as it was clarified with the authors actually provides odds ratios instead): OR of 2.12, 95%CI (1.05-4.30) for BRCA1 and OR of 5.83, 95%CI (3.64-9.32) for BRCA2[5]. For the use in the model, the HR were calibrated to predict the reported OR for the incidence of the PC, as it is reported by Nyberg et al. (2022). These HR (2.42, 95%CI (1.93; 3.05) for BRCA1 and 4.45, 95%CI (4.31; 4.59) for BRCA 2) were used in the model as the assumed "actual" relative risk of prostate cancer in the population. Based on feedback from Dr Hanson (see the stakeholders' section), it was also assumed that having both mutations does not increase cancer risk beyond the highest risk associated with BRCA1 or BRCA2.

The HR for prostate cancer due to BRCA mutations is not modifiable, as it is an inherited genetic risk. However, the impact of age should be considered when

informing this parameter. Prospective studies often report "known" rather than “actual” relative risks due to selection criteria and observation periods. For instance, Nyberg et al. (2022)[5] found higher risks for BRCA mutations in individuals under 65 years compared to those over 65 years, driven by prospective observational studies. This reflects the earlier onset of cancer in individuals with genetic risks.

Known BRCA1,2 status

Knowledge of BRCA1,2 mutation status tends to increase with age. To estimate the probability of individuals knowing their BRCA1,2 status at a given age (assuming they are carriers) we used data from a study by Forde et al. (2020)[6], which examined the uptake of pre-symptomatic genetic testing among relatives of individuals with confirmed BRCA mutations in West England[6]. Age-specific probabilities of BRCA testing (Table 2) were derived from the study and used to calculate cumulative probabilities of BRCA status awareness by age, based on the following assumptions:

- The probability of knowing one’s BRCA status for individuals aged 20–30 was assumed to be equal to those under 29;
- Cumulative probabilities for the 30–39 and 40–49 age groups were set equal to the respective group-level probabilities reported in Forde et al. (2020)[6];
- For individuals aged 50 and above, the cumulative probability of knowing one’s genetic status was assumed to match the testing probability among those aged 50–59.
- The age of knowing one’s BRCA status was sampled using a uniform distribution within the allocated age group, based on cumulative probabilities.

It should be noted that, due to the rapid rise in genetic testing uptake in recent years, these estimates likely underestimate current levels of BRCA1/2 status awareness.

Table 2: Probabilities of knowing their BRCA1/2 status before their age.

Age groups in the study, years	Number of people in the age group	Number of people tested pre-symptomatically	Probability of being tested	Age in the model, years	Cumulative probabilities in the model
<18	101	33	0.32673		
18-28	376	91	0.24202	20-29	0.25996
29-39	519	143	0.27553	30-39	0.27553
40-49	567	185	0.32628	40-49	0.32628
50-59	470	166	0.35319	50 -	0.35319
60+	521	157	0.30134		

6.4.3. Family History of prostate, breast, and ovarian cancers

HSE does not contain any information about family history of prostate, breast, or ovarian cancers. Although family history can be defined in several ways, for the purposes of this model it was defined as having one or more first-degree relatives previously diagnosed with any of these cancers. While familial risk is heterogeneous, no prior studies have evaluated the cost-effectiveness of screening in men with familial risk considered as a single group. Therefore, in this model, familial risk was defined broadly as having a first-degree relative diagnosed with breast, ovarian, or prostate cancer at any age or stage.

Family history of cancers is also correlated with BRCA mutations; this means that family history should not be allocated randomly in the population when BRCA mutations are randomly allocated and that we need to inform the model with the data on lifetime diagnosis of cancers in the presence and absence of BRCA mutations. Family history is correlated with age; however, it would be incorrect to model individual risk levels changing over time due to changes in family history, when in fact it is the knowledge about familial risk that changes, not the risk itself. Instead, two different family history variables were modelled corresponding to true family history and known family history, with true family history used to influence the modelled natural history of prostate cancer and known family history used to calculate prostate cancer risk in risk-stratified screening.

Actual familial risk allocation

True familial risk for all three cancers is allocated the following way. Firstly, we estimate the lifetime risk of cancers in general population (15.3% for breast, 2% for ovarian, 17.9% for prostate cancer) and in population with BRCA mutations (70.26% for breast, 28.37% for ovarian, and 46.95% for prostate cancer based on Petrucelli et al. (updated 2023)[7]). Since part of the risk in general population is still attributed to BRCA mutations, we calibrated the lifetime risk in the non-BRCA population by using a synthetic population with the risk in BRCA-positive people as a deterministic value and predicting the general population values as reported in the review. The calibrated lifetime cancer probability values for non-BRCA carriers are: prostate cancer for men 0.17895; ovarian for women 0.0197; breast cancer for women 0.1525.

Secondly, we calculated in a similar way the probability of a person having a mutation conditional on the BRCA mutation in their relatives. If an index person (a person from HSE population) has a mutation, each of their relatives has a 0.5 probability of also having a mutation. If the index case has no mutation, the probability for their relatives to have a mutation is lower than the average prevalence of the mutation. To estimate this probability, we calibrated it using a fitting approach to predict the total prevalence of BRCA mutations in the population. The calibrated parameter of probability to have a BRCA mutation if one of the first-degree relatives has no mutation is 0.00413478.

Thirdly, we calculated the lifetime probability of having at least 1 relative with cancer conditional on having or not having a BRCA mutation. To do this, we assumed the average number of 1st degree relatives for each HSE person (2.7 first degree relatives, i.e. parents and siblings)[8], half of those are males and half are females.

The probability to have a familial history based on a BRCA status for a person in HSE was calculated in the following way:

$$P(familial\ history) = \left(1 - \left(\frac{1 - P\ of\ cancer}{BRCA}\right)^{N\ of\ relatives}\right) * P\ of\ relatives\ with\ BRCA$$
$$+ \left(1 - \left(\frac{1 - P\ of\ cancer}{no\ BRCA}\right)^{N\ of\ relatives}\right) * P\ of\ relatives\ not\ having\ BRCA$$

Where N of relatives - number of relatives of the relevant sex (1.35 for males and females), P of cancer/BRCA – probability of a relative to have cancer conditional that they have a BRCA mutation, P of cancer/no BRCA – probability of a relative to have cancer conditional that they don't have BRCA mutation.

The calculated probabilities to have familial history for three cancers based on the BRCA status of a person in HSE are reported in Table 3.

Table 3: Calculated probabilities to have familial history for three cancers based on BRCA status of a person in HSE

Cancers	Probability of the lifetime risk of having at least 1 relative with cancer if HSE person has BRCA mutation	Probability of the lifetime risk of having at least 1 relative with cancer if HSE person has no BRCA mutation
Breast	0.5028	0.2027
Ovarian	0.1946	0.0279
Prostate	0.4044	0.2351

Cancer risk based on familial history

Several reviews reported a similar increase in prostate cancer risk associated with a family history of the disease[3, 9, 10]. A more recent systematic review by Perez-Cornago et al. (2017) estimated the HR for individuals with a family history of prostate cancer (compared to those without) as 1.94, 95%CI (1.77–2.13)[11].

For breast cancer, a systematic review and meta-analysis by Ren et al. (2019) found that having a first-degree relative with female breast cancer was associated with a RR of 1.18 (95% CI: 1.12–1.25). To estimate the effect of a family history of breast cancer on the risk of prostate cancer, we used a lognormal distribution calibrated to reflect this risk relationship. The calibrated parameters were log mean = 0.1687, log standard deviation = 0.004797, corresponding to a mean HR of 1.2344[12].

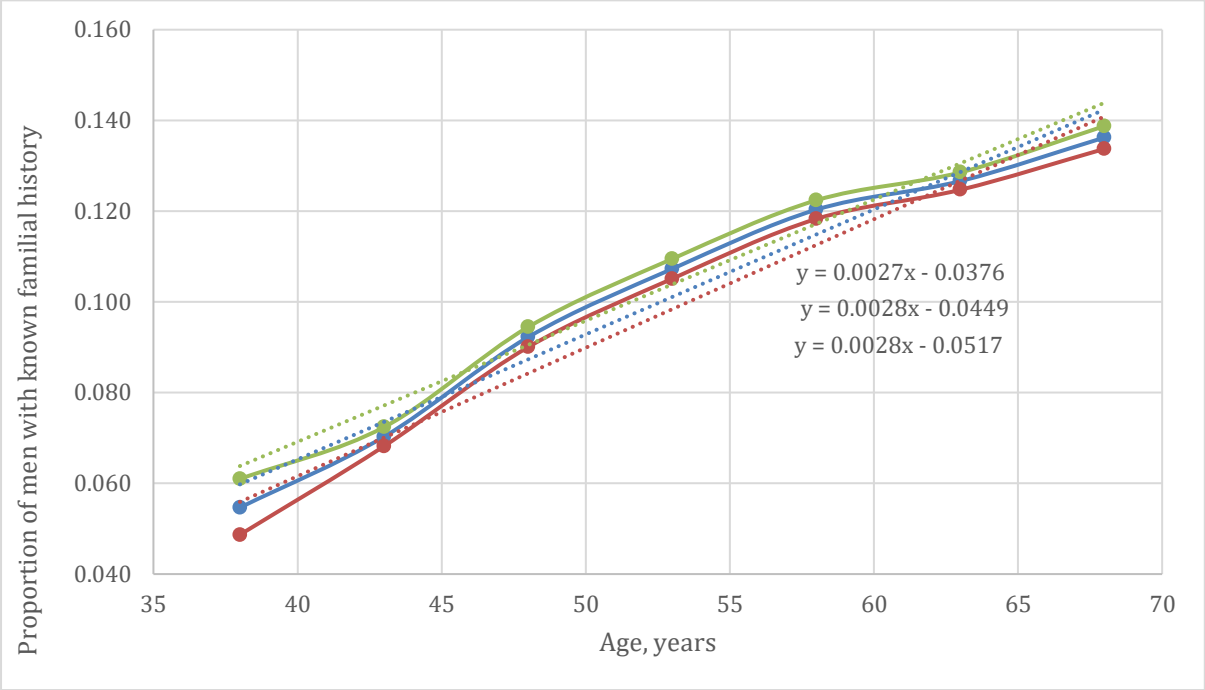
There was no high-quality data available on the familial risk of prostate cancer among relatives of individuals with ovarian cancer. However, Beebe-Dimmer (2020)[13] reported a combined prostate cancer risk estimate for relatives of

individuals with breast or ovarian in a US-based population, with a reported RR of 1.47 (95% CI: 1.43–1.50). Using this, we again applied a lognormal distribution and calibration approach to estimate the risk of prostate cancer conditional on having a relative with ovarian cancer. The calibrated parameters were: log mean = 1.2395, log standard deviation = 0.006221, yielding a mean HR of 3.45401.

Known familial risk

The data on the correspondence of true familial history and known familial history by age is not available. Thus, the assumption was made that the age impact on the known familial history is similar to that for colorectal cancer in the MiMiC-Bowel model, which was obtained from UK Biobank[14]. This does not mean that the familial risk of prostate cancer was based on colorectal cancer data, but that the correlation of having familial history of cancer and knowing about their familial history of cancer for prostate cancer was estimated using the colorectal cancer trends. The impact of age in the model was estimated through fitting a linear model to all data points between the 36-40 and 66-70 age groups from the UK Biobank data (Figure 3). Known family history was assigned to a subset of the individuals with true family history, corresponding to the age at which a risk model might be used to estimate cancer risk (ranging between age 30 and 70, depending upon model user input). It was assumed that by the age 70 years, all people with familial risk were aware of their risk status.

Figure 3: Graph showing increase in family history of colorectal cancer in the MiMiC Bowel model with age, from UK Biobank[14]



6.5. Risk factors for progression of undiagnosed cancer

In the model, it was assumed that characteristics such as age, Black ethnicity, and BRCA 1,2 carrier status influence not only the onset of cancer but also its progression. The impact of these risk factors on the progression of undiagnosed cancer was modelled through the GGG pathway. Specifically, the HR for each factor was used to determine an individual's likelihood of being assigned to a higher GGG category at the time of cancer onset. For patients who developed cancer, a GGG value was sampled at the time of onset, with the assumption that the GGG would remain unchanged thereafter. The HRs were treated as non-cumulative, meaning that if multiple risk factors applied to an individual in the HSE dataset, only the highest associated risk was considered. For example, a Black BRCA carrier would have the RR for BRCA carriers, rather than for Black ethnicity. This assumption was introduced after testing showed that multiplying HRs (as is done for cancer onset) led to unrealistic outcomes. All calculated probabilities were subsequently normalised to ensure they summed to 1.

6.5.1. Allocation of GGG by age

To estimate the GGG distribution at the time of onset was based on data from the NHS Digital (2020) [15] (Table 4).

In the model, data from Table 4 were used to calculate the probabilities of having age-specific GGG. It was assumed that prostate cancer with an onset at age 20 would have the same GGG distribution as at age 35. Since prostate cancer onset among non-BRCA carriers and those without familial risk at age 20 is rare, this assumption does not significantly impact the model. For all other ages, the probabilities were recalculated under the assumption that cancers classified as "unknown" grade had an equal likelihood of being GGG 1-5. The data were also adjusted with the analysis of Sheridan (2008) surveying stage T1c patients[16]. The update was conducted to match GGG distribution at the time of onset and not the diagnosis, since it is likely that there are more lower-grade tumours in undiagnosed cancer. Sheridan et al. (2008) followed up the patients for 5 years and reported that 18.7% of patients in GGG1 progressed to higher groups. The values were adjusted by assuming that 18.7% of cancers diagnosed as high grade were low grade at the time of onset and that unknown GGG cases are equally distributed across the GGG (Table 5).

It was noted by the clinical experts that since these data are from 2020 it could possibly be affected by the Covid pandemic; however, no older data could be identified.

Table 4: Gleason Grade Group distribution by age: original data

Gleason Grade Groups					
Age, lower bound, years	**1**	**2**	**3**	**4 & 5**	**Unknown**
35*	33.3%	38.5%	5.1%	7.7%	15.4%
45	30.7%	39.5%	12.4%	10.8%	6.5%
50	27.0%	41.3%	10.0%	15.8%	5.8%
55	25.5%	36.9%	13.0%	17.9%	6.7%
60	22.2%	33.5%	15.0%	21.9%	7.4%
65	19.4%	34.2%	15.1%	23.0%	8.3%
70	14.5%	28.3%	17.3%	27.8%	12.2%
75	11.5%	25.6%	15.5%	27.6%	19.7%
80	5.4%	10.4%	8.1%	21.6%	54.5%
85	1.7%	3.5%	2.4%	8.5%	83.9%
90	0.5%	0.5%	0.3%	2.5%	96.3%

**This is a combined cohort as counts in these age groups were very low; it is also used for population in the younger age group.*

Table 5: Gleason Grade Group distribution by age: inputs to the model

Gleason Grade Groups				
Age, years	**1**	**2**	**3**	**4 & 5**
20	0.55	0.15	0.06	0.24
35	0.55	0.15	0.06	0.24
40	0.50	0.42	0.02	0.05
45	0.45	0.38	0.09	0.07
50	0.42	0.40	0.06	0.13
55	0.40	0.35	0.10	0.15
60	0.38	0.32	0.12	0.19
65	0.35	0.33	0.12	0.20
70	0.32	0.27	0.15	0.27
75	0.30	0.27	0.14	0.29
80	0.30	0.27	0.14	0.29
85	0.30	0.27	0.14	0.29
90	0.30	0.27	0.14	0.29
100	0.30	0.27	0.14	0.29

6.5.2. Allocation of GGG by ethnicity

In the model, the probability of being in one of the GGG categories is also dependent on Black ethnicity status. The model assumes that the probabilities to be allocated to each GGG for ethnicities other than Black are the same as for White ethnicity. The RR of being in GGG2 vs GGG1 and in GGG3-5 vs GGG1 for the Black population are informed through the US cross-sectional study using Medicare data (Navarro et al. 2022)[17]. This study reported that people of Black ethnicity were 76% more likely to be diagnosed with a Gleason grade score of 7, as opposed to 6 (OR = 1.76, 95% CI = 1.29–2.40), and were 73% more likely to be diagnosed with a Gleason grade score of 8–10 (OR = 1.73, 95% CI = 1.21–2.48). We considered that this corresponds to GGG2 and GGG3-5, respectively.

In the model, these values were used to first calculate the probabilities of being in GGG 2 and higher GGG categories for the non-Black population, using population weights for both non-Black and Black groups in the HSE, as well as a matrix of GGG probabilities by age. For example, the probability of being in GGG 2 by age for the non-Black population was calculated as follows:

$$\mathrm{P}(\mathrm{GGG}\ 2, \mathrm{non}-\mathrm{Black}, \mathrm{age\ x}) = \frac{P\ (GGG\ 2, age\ x)}{\left(RR_{GGG\ 2,Black} \times weighter\ mean\ for\ Black\right) + weighted\ mean\ for\ non-Black}$$

Then, the RR of GGG 2-5 for Black versus non-Black individuals was applied to the matrix of GGG probabilities for the non-Black population to calculate the weighted risk for Black ethnicity. All probabilities were normalised to ensure they summed to 1.

6.5.3. Allocation of GGG by BRCA status

For people with BRCA mutations, the model uses the RR of being in GGG 2-5 vs GGG 1, based on the random effects model in the systematic review of Nyberg et al. (2022)[5]: 1.59, 95%CI (1.02–2.49) and 4.94, 95%CI (3.51–6.96) for BRCA1 and BRCA2 carriers, respectively. It was assumed that the increased risk of a higher GGG for BRCA mutations does not vary with age, as applying the RR already resulted in a near-zero probability of being classified as GGG 1 for BRCA1 carriers. All the transition probabilities were normalised to 1 to avoid implausible (negative) transition probabilities.

6.6. Natural history of prostate cancer

The natural history module of the model relies upon cancer onset, cancer progression, and cancer survival.

6.6.1. Probability of cancer onset

Cancer onset is dependent on age, BRCA mutation, ethnicity, and true familial risk (i.e., the other genetic mutations), as described in the previous sections. For binary characteristics (BRCA1, BRCA2, familial history), the individual value was 1 for an individual possessing those characteristics, and 0 for an individual not possessing the characteristics, with the population mean value representing the proportion of people with the characteristic in the population.

Considering the non-linear relationship between age and the diagnosed incidence of cancers, the relationship between the age, other risk factors, and the risk of cancer onset was modelled using multiplications of the HR (where age impact is informed by the calibration while impact of other factors directly by data as described above). The risk by age is assessed as:

$$risk\ at\ age\ x = age\_20y(^{Age\ x - Age\ 20})$$

Where age_20y – risk of prostate cancer onset for 20-years-olds.

The HR impacting disease onset by demographic factors was multiplied by the probability of cancer onset for a white man without familial history and BRCA mutations.

6.6.2. Time of progression across TNM stages

The allocation of cancer stages according to the TNM categories is conducted in the model probabilistically by sampling the time required to reach each stage from a distribution with a calibrated mean. This time was stochastically drawn from a Weibull distribution, with considerations for the calibrated mean (μ), calibrated shape (k), and computed scale (lambda).

$$Lambda = \frac{\mu}{Gamma(1 + \frac{1}{k})}$$

It was assumed that the mean time of progression across the stages is shorter in higher GGGs. This was implemented in the model by adjusting the mean time for

people in GGG 2-5 with the calibrated multipliers, which are lower than 1 (the multiplier for the GGG1) for GGG 2-5 and are lower in higher GGGs compared to lower GGGs. These assumptions mean that on average population groups with a higher probability of having a high GGG (Black and BRCA 1/2, carriers) will have more aggressive cancer that will progress more quickly across stages 1 to 4.

People of Black ethnicity and BRCA 1/2 carriers, on average, are diagnosed with more advanced cancers, resulting in lower survival rates compared to individuals of other ethnicities or non-carriers of the BRCA 1/2 mutation.

6.6.3. Survival

Net survival is presented in the available datasets and is used as an estimate of the cancer-specific survival of cancer patients compared with the background mortality that patients would have experienced if they had not been diagnosed with cancer. Net survival is therefore the probability of survival solely from the risk of death from cancer. The majority of datasets use the Pohar-Perme estimator of net survival, which accounts for informative censoring bias. Occasionally, net survival greater than 100% can occur if the survival experience in cancer patients due to cancer-related risk is greater than the survival experience of the general population, for instance due to improved health monitoring or lifestyle changes.

Several sources of survival information were available and contained different stratifications of the population and differential follow-up durations, Table 6. The PHE dataset was selected as the most informative data source; this is discussed further in the following section.

Table 6: Databases considered for survival analysis

Database source	Details	Comments
ONS[18]	Adults diagnosed between 2013 and 2017 and followed up to 2018	Survival by age and stage, provided at years 1 and 5. Predicted survival by age only at 10 years.
PHE[19]	Adults diagnosed 2014 to 2018 and followed up to 2019	Survival by age and stage, provided at years 1, 2, 3, 4, 5. Missing data for years 2-5 for stage 1 75-84, stage 2 65-74 and 74-85.
Get Data Out[20]		Staging is not according to TNM stages I-IV. There would be large uncertainty in mapping the stages to the TNM scale.
NHS Digital[21]	Most recent data from NHS digital. Cancer Survival in England, cancers diagnosed 2016 to 2020, followed up in 2021	Survival by age in table 1, survival by stage in table 2. No dual stratification of survival present.
PHE-CRUK[22]	10-year cancer survival by stage for patients diagnosed in the East of England, 2007 to 2017	Difference in population (East of England as opposed to England).

Data availability

The PHE dataset provides the net survival at years 1, 2, 3, 4, and 5 stratified by age and stage, with data summarised in Table 7. This dataset was therefore used for the extrapolation of the net survival for stages 3 and 4. It is noted, however that the data for stages 1 and 2 was incomplete. Additionally, the net survival for stages 1 and 2 was predominantly above 100%, even at later time points for these stages, which made any attempt at net survival extrapolation result in net survival equal to, or greater than, 100%. Therefore, the PHE dataset was not used for the modelling of net survival in stages 1 and 2.

Table 7: PHE net survival (%) for 1-5 years stratified by age and disease stage

Time since diagnosis (years)						
Stage	**Age**	**1**	**2**	**3**	**4**	**5**
1	15-54	99.7	99.8	99.8	99.7	99.3
	55-64	100.2	100.4	100.6	100.6	100.5
	65-74	100.4	100.7	101.1	101.4	101.9
	75-84	101.6	.	.	.	.
	85-99	103.6	105.6	104.2	102.4	97.7
2	15-54	100.1	100	100.1	100.3	100.2
	55-64	100.2	100.3	100.5	100.7	100.6
	65-74	100.6	.	.	.	.
	75-84	101.5	.	.	.	.
	85-99	101.9	103.4	100.5	101.3	92.7
3	15-54	99.8	99.8	99.3	99	97.8
	55-64	100.2	100.1	99.7	99.2	98.4
	65-74	100.4	100.3	99.7	99.1	98.6
	75-84	100.9	100.5	99.8	98.8	97
	85-99	100	97.9	90.7	87.1	78.8
4	15-54	93.2	80.4	70.5	61.9	56.7
	55-64	94.1	83.6	73.4	64.9	58.8
	65-74	91.1	80	70.1	61.5	54.7
	75-84	84	69	58.5	50.6	43.4
	85-99	70	51.6	41.5	33.9	27.1

Legend: Items in red indicate entries with missing data.

The ONS dataset provided 1- and 5-year net survival stratified by age and stage, with data summarised in Table 8. This dataset was used for comparison and to supplement data for the groups with missing net survival for years 1 and 5 in the PHE dataset. The ONS dataset also provided 10 year predicted survival stratified by age (not stage); this was considered for validation.

Table 8: ONS net survival for years 1 and 5 stratified by age and disease stage.

		Time since diagnosis (years)	
Stage	**Age**	**1**	**5**
1	15-54	99.6	99.1
	55-64	100.1	100.3
	65-74	100.4	101.4
	75-84	101.3	101.5
	85-99	103.1	91.6
2	15-54	100.0	100.2
	55-64	100.2	99.8
	65-74	100.5	101.6
	75-84	101.1	102.0
	85-99	100.3	89.7
3	15-54	100.0	96.9
	55-64	100.3	98.6
	65-74	100.4	98.2
	75-84	100.5	95.8
	85-99	99.4	71.2
4	15-54	92.8	54.1
	55-64	93.2	55.4
	65-74	90.7	54.3
	75-84	83.2	42.0
	85-99	67.9	24.8

Survival in stages 1,2

The assumptions on survival in stages 1 and 2 changed in the Phase 2 work based on the consultations with the clinical experts and the NSC feedback. Since the ONS dataset had a maximum of two follow-up times, and resulted in extrapolated net survival close to 100% irrespective of follow-up time since diagnosis (considering 100% survival at the longest follow-up time of 5 years in ONS and PHE datasets), alternative data sources were used for stages 1 and 2. The longest follow-up data were available from the ProtecT trial conducted in the UK in 1999–2009, including 82,429 men aged 50–69 years. In the ProtecT trial, net survival was reported at 10 and 15 years. The reported 15-year survival rates in ProtecT trial were: Stage 1 —98% and Stage 2 —95%[23]. Considering the data limitation, the

modelling assumptions considered that the prostate cancer-specific survival in stages 1 and 2 is age-independent; however, for stages 3 and 4, survival is stratified also by both age and stage.

Extrapolation of survival

Given that the available data for stages 1 and 2, and stages 3 and 4, were only available up to 15 and 5 years, respectively, parametric modelling was employed to extrapolate the survival to the lifetime horizon of the model. In all modelling, net survival above 100% was capped at 100%, considering that in the model cancer-specific mortality and other-cause mortality are simulated separately.

For stages 1 and 2, data were especially limited, with only three data points (time 0, 10, and 15 years). Therefore, an exponential model was used to extrapolate the survival as this is the simplest of the standard survival parametric models, with only one parameter to fit. The exponential is defined as:

$$S(t) = \exp(-\lambda t),$$

Where $S(t)$ is the net survival at time t and λ the rate parameter.

The exponential model assumes the instantaneous risk of death (hazard) is constant. This is a fairly strong assumption, but due to the limited data, alternative models which relax this assumption and have a greater number of parameters could not be used as the associated parameter uncertainty would have been considerable.

For stages 3 and 4, there were an increased number of data points and thus a Weibull model, which permits monotonically increasing or decreasing hazards, was used. The Weibull model can be defined as:

$$S(t) = \exp\left(-\left[\frac{t}{scale}\right]^{shape}\right),$$

Where shape and scale are the model parameters.

In order to fit the above models, a non-linear least squares optimisation was used via the NLS package in R. It is noted that the models were fitted directly to the net survival values, not individual patient data (which were not available); therefore, the uncertainty in the model parameters is associated with the uncertainty in the model fit, not variation in the population.

Point estimates of the model parameter(s) were output for use in the base case. To facilitate the re-sampling of the model parameters in the probabilistic sensitivity analysis, the parameter variance/covariance matrices were also output. Resampling of model parameters for the probabilistic sensitivity analysis was conducted using a multivariate normal distribution, centred at the model parameter point estimates, and variance/covariances output from the model fit. The survival curve was then regenerated according to the exponential or Weibull models accordingly for use within the probabilistic sensitivity analysis.

Considering that survival analysis was based only on five-year data for stages 3,4, several scenario analyses were conducted around survival (see section "Description of screening scenarios".

Probability of death

Based on this, the probability of dying due to cancer was calculated from the survival data as follows:

Cancer_mort(age, stage, year) = 1 – (Cancer_surv(age, stage, year) / Cancer_surv(age, stage, year-1))

It was assumed that the probability of dying from cancer beyond 70 years post-diagnosis was 0. Based on the calibration fit (see the Calibration section), the base case analysis in the model used only 15-year survival data. This approach assumes that men with prostate cancer who do not die from the disease within 15 years of diagnosis are subsequently at risk of dying only from other causes. Lifetime (70 years) extrapolation of survival was explored in scenario analyses.

Mortality from other causes

Cancer mortality was subtracted from all-cause mortality to retrieve other-cause mortality.

6.7. Calibration and validation of the model

6.7.1. Calibration parameters

The following parameters related to NHD were calibrated:

1. Probability of cancer onset at age 20 years for non-Black, no family history and non-BRCA carrier (P.onset).
2. Coefficient defining the impact of age on cancer onset (C.onset.age).
3. Annual probability of symptomatic diagnosis at four different stages (P.sympt.diag).
4. Coefficient for annual decrement in symptomatic presentation after the age of 70 (Symp.decr).
5. Probability to die undiagnosed after age 70 years (P.undiag.dead).
6. Mean time of progression across the stages for undiagnosed cancer and shape of the Weibull distribution for those patients who had GGG1 (e.g., for progression from stage 1 to stage 2: Mean.t.StI.StII and shape.t.StI.StII).
7. Coefficients adjusting the means for progression for GGG2 to GGG 5 compared to GGG1 (k.WB.GGG.2, k.WB.GGG.3, k.WB.GGG.4.5).

6.7.2. Calibration targets

The current care arm parameters of the model were calibrated to the current epidemiological data:

- Incidence of prostate cancer in England by age;
- Incidence of prostate cancer in England by age and stages 1 to 4;
- Mortality by age;
- GGG by age and stage.

This means that clinical diagnosis in the current care arm represents a composite outcome of three groups: men diagnosed through the symptomatic pathway for prostate cancer, those who underwent PSA testing due to suspicion of other diseases, and those identified through opportunistic screening. The model was not calibrated to retrospective data for pragmatic reasons and to avoid potential data conflicts - specifically, inconsistencies that can arise when combining data sources from different time periods, as well as to be able to reconstruct current care

scenario because the uptake with opportunistic screening in England is unknown and fluctuates through years.

Incidence of PC

Age-specific incidences have been sourced from NHS Digital, using the Cancer Registration Statistics for England in 2018[24] (Table 9). As advised by the stakeholders (Prof. H. Ahmed, Prof. J. Catto. and Prof. W. Cross), the incidence by age was used until age 75 years, then the average incidence between 75-100 was used for the 75-100 age group. The model assumed an increase in cancer onset by age and a decrease in clinical diagnosis from age 75 years.

Table 9: Age-specific incidence of prostate cancer in England (per 100,000 population), 2018[24]

Age at diagnosis, lower bound, years	Prostate cancer incidence rate	95% Confidence Interval	
30	0	0	0
35	0.4	0.2	0.9
40	5.5	4.5	6.8
45	22.9	20.8	25.1
50	90.4	86.2	94.7
55	244.2	237.0	251.6
60	406.8	396.6	417.1
65	719.0	704.9	733.4
70	851.0	835.2	867.0
75	949.0	928.5	969.8
80-90	714.0	683.6	744.6

Distribution of prostate cancer by age and stage

The Phase 1 work included age-specific incidence rates by stage sourced from Public Health England from 2012 to 2014. To provide a more recent distribution of stages, that captures the uptake with opportunistic screening, the following two sources were combined (Table 10) in the Phase 2 work based on the feedback from stakeholders and the National Screening Committee:

- NHS Digital (2021): Reports incidence rates for early (stages 1/2) and late (stages 3/4) cancers by age.

- CMA Stage (2013–2021): Used to further refine the distribution by TNM stages.

The missing stage at diagnosis was assumed to be equally distributed across the four TNM stages. This modelling assumption was supported by consultations with clinical experts (see Supplementary I) and is consistent with approaches adopted in other well-established CISNET models [25-27].

The updated data distribution reports a higher proportion of early-stage cancers (stages 1/2) in younger age groups compared to the previously used dataset for t Phase 1, supporting the assumption that the introduction of opportunistic screening may have improved earlier diagnosis of prostate cancer.

Table 10: Distribution of prostate cancer by age and stage used in the model

	NHS digital (2021)			Adjusted by CMA stage (2013/2021)				Proportion of cases by stage			
Age, years	**Stage 1,2**	**Stage 3,4**	**Total N of cases**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**
30 to 34	-	-	-	-	-	-	-				
35 to 39	4	2	6	3	1	1	1	45 %	22 %	18 %	16 %
40 to 44	37	12	49	25	12	6	6	51 %	25 %	13 %	11 %
45 to 49	197	75	272	132	65	40	35	49 %	24 %	15 %	13 %
50 to 54	733	344	1,077	493	240	184	160	46 %	22 %	17 %	15 %
55 to 59	1,808	1,004	2,812	1,215	593	537	467	43 %	21 %	19 %	17 %
60 to 64	2,683	1,772	4,455	1,803	880	947	825	40 %	20 %	21 %	19 %
65 to 69	3,587	2,814	6,401	2,410	1,177	1,504	1,310	38 %	18 %	23 %	20 %
70 to 74	3,947	4,011	7,958	2,652	1,295	2,144	1,867	33 %	16 %	27 %	23 %
75 to 79	3,025	3,588	6,613	2,033	992	1,918	1,670	31 %	15 %	29 %	25 %
80 to 84	946	2,068	3,014	636	310	1,105	963	21 %	10 %	37 %	32 %
85 to 89	303	1,080	1,383	204	99	577	503	15 %	7%	42 %	36 %
90 and over	70	460	530	47	23	246	214	9%	4%	46 %	40 %

Mortality by age

Age-specific mortality rates have been sourced from NHS digital data for 2020 (Table 11).

Table 11: Prostate cancer mortality by age (per 100,000 population)

Age at diagnosis, lower bound, years	Rate, mean	Confidence interval	
45	0.3	0.1	0.6
50	2.5	1.9	3.3
55	7.4	6.2	8.7
60	21.0	18.8	23.4
65	51.5	47.7	55.5
70	95.5	90.3	100.8
75	178.4	169.9	187.1
80	322.0	308.3	336.2
85	631.1	605.3	657.8
90	1058.1	1009.7	1108.2

Gleason Grade Group by stage and age

The Get Data Out dataset from the National Cancer Registration and Analysis Service 2020 was used to obtain the incidence rates of GGG by age and stage. However, the staging system used in this dataset was localised, locally advanced, and metastatic. The dataset documentation offered a loose conversion of this staging system to stages I – IV. According to this conversion, localised equates to stages I-II and locally advanced equates to stages III-IV (Table 12 and Table 13), while metastatic equates to stage IV. The Get Data Out dataset did not have a breakdown of GGG for metastatic cancer, so locally advanced has been used for stages III and IV.

Table 12: Localised stages (I-II), incidence rates by GGG per 100,000

Age at diagnosis	Grade Group 1	Grade Group 2	Grade Group 3	Grade Group 4 & 5
Age 00-59	4.3	5.2	1.1	0.5
Age 60-69	63.2	82.1	24.8	18.0
Age 70-74	68.9	106.5	45.6	37.4
Age 75-79	57.1	108.9	50.9	45.8
Age 80+	13.6	18.8	10.3	17.8

Table 13: Locally advanced (stages III-IV), incidence rates by GGG per 100,000

Age at diagnosis	Grade Group 1	Grade Group 2	Grade Group 3	Grade Group 4 & 5
Age 00-59	0.6	1.2	0.9	0.9
Age 60-69	0.6	26.9	21.3	29.3
Age 70-74	4.7	39.3	39.7	58.8
Age 75-79	4.7	36.9	36.3	62.7
Age 80+	4.7	7.6	10.0	20.6

To avoid a conflict with the NHS Digital source for the GGG distributions by stage, we calculated the probability of being in different GGGs for each stage using weighted averages (Table 14). For stages 1 and 2, the average proportion of patients in each GGG at diagnosis in the localised stage was used for age groups 30-59 and 60 and above. For stages 3 and 4, the average proportion of patients in each GGG at diagnosis in the locally advanced stage for ages 60-79 was applied across all ages. For stage 4, it was assumed that no patients have GGG 1. These assumptions were informed and validated by the stakeholders.

Table 14: Probability of being in GGG by TNM stages

	Grade Group 1	Grade Group 2	Grade Group 3	Grade Group 4&5
Stage 1 & 2	0.265354	0.411662	0.169918	0.153066
Stage 3	0.027632	0.285581	0.269377	0.417410
Stage 4	0	0.293696	0.277032	0.429272

Sojourn time

In the model, sojourn time is defined as the period during which cancer is asymptomatic yet detectable through screening, specifically the interval from cancer onset to symptomatic detection. This predicted sojourn time was compared to the secondary analysis of the CAP randomised controlled trial (Martin, 2024)[28]. For the cohort aged 50 to 54 years, the mean sojourn time was reported in this analysis to be 12.1 years (95% CI 12.1-12.2 years), while for those aged 65 to 69 years, it was 15.3 years (95% CI 15.2-15.3 years). This evaluation demonstrates that sojourn time increases with age, implying that cancer progresses more slowly in older individuals. However, stakeholders challenged this assumption, noting that higher grades of GGG are positively correlated with

age and that higher GGG is associated with more rapid cancer progression. Consequently, upon the stakeholders' recommendations, the model was calibrated to a mean sojourn time for 50–69-year-olds reported in the trial with a wider confidence interval of ±10%: 13.4 years (95% CI 12.06-14.74). The stakeholders also indicated that the sojourn time measured in the CAP trial may not be entirely accurate for the population of England, suggesting that prioritising GGG as a calibration target would be more beneficial than relying solely on sojourn time.

6.7.3. Calibration approach

The model was calibrated using the Metropolis-Hastings algorithm (MHA).

Metropolis-Hastings algorithm: general approach

The general calibration approach using the MHA involves a systematic process of estimating the parameters of a model, particularly in Bayesian inference. The process begins with the selection of an initial parameter value, followed by generating a proposal for a new parameter value based on a defined proposal distribution and the step size. The acceptance of this new value is determined by comparing the likelihood of the proposed value to the current value, adjusted for the proposal distribution's characteristics. Specifically, if the proposed value yields a higher likelihood, it is accepted with a probability of 1; if it results in a lower likelihood, it is either rejected or accepted with a lower probability (called “jumps”). This iterative process creates a chain of parameter samples that converge to the target distribution, allowing for the estimation of the parameters' posterior distributions and their associated uncertainties. Calibration is achieved by using these samples to refine model predictions, ensuring that the model reflects the observed data.

The initial step size in the calibration was set to 20% of the parameter values. The step size was reset to 10% when calibration chains broke and were restarted from the last accepted parameters. The maximum step size was also reduced from the original value based on the acceptance rate (after the first 5,000 warming-up runs, each 500 cycles the step was reduced 10% if the acceptance rate was less than 10%).

The proposal parameter set in calibration was always accepted if the proposal parameter set had higher likelihood than the current parameter set. In addition, the calibration used a dynamic probability to accept parameters with lower likelihoods based on a probability to accept such parameters after a warmup period of 0.1 and the difference in likelihoods of less than 10% (i.e., that likelihood did not decrease by more than 10%).

Addressing the limitations of the Metropolis-Hasting algorithm

The MHA has some limitations. Calibration theory says that if the MHA is run for sufficient iterations, it will fully explore the parameter space. This approach, though, is computationally expensive since simulation of the population with rare events (such as cancer) requires substantial running time. With limited processing time, there is a risk that the parameter space is not adequately explored. That is to say that one cannot be sure that if convergence is achieved, there are no other alternative acceptable parameter regions.

To minimise the risk, the parameter space was first explored by running a random calibration with the Latin Hypercube Sample. This method includes determining the number of required samples of parameters (n), assigning a probability density function to each parameter, dividing the distribution into n intervals and randomly drawing from each n^{th} interval for each calibration parameter. The Latin Hypercube Sample with 100,000 random samples generated a near-random sample of parameter values from a multidimensional distribution. In the Phase 2 work, four parameter sets with the highest likelihood from the initial calibration were used as starting points for recalibrating the model, setting up three independent MHA chains. This approach was taken in light of updates to some calibration targets and changes in the population age structure.

Goodness of fit (GOF)

Log-likelihood was used as a goodness-of-fit measure. Since the model is run with a population who are all of the same age, the number of events at each age are dependent and represent events in time. Thus, log-likelihood was calculated as a lifetime outcome. Lower weights were used for all calibration targets (0.3 of the

original value) except for the total incidence and mortality as the data with the assumed highest accuracy.

Exploration of the population size to run in the calibration

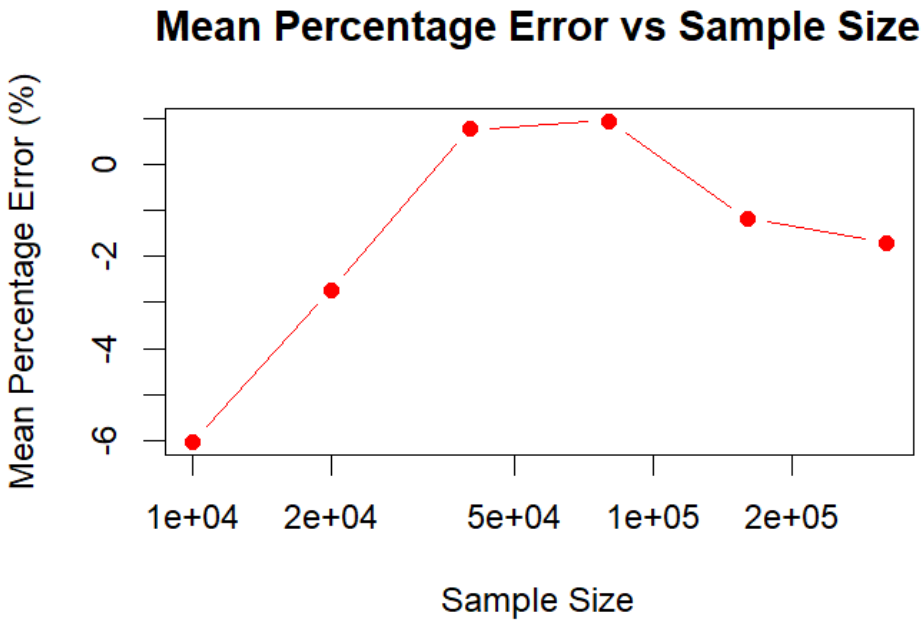
When determining the required sample size, the complexity of models plays a crucial role, as more complex models generally require larger sample sizes. A common rule of thumb for model calibration is to have at least 10 data points per parameter to ensure reasonable accuracy in estimating model parameters (Steyerberg, 2019)[29]. The sample size is also influenced by how well the model fits the data. In Bayesian calibration, smaller sample sizes can suffice if the priors are informative (Jalal, 2017)[30].

In the Prostate.mic model, 17 parameters need to be calibrated. Based on the rule of thumb, a population of 10,000 would provide 170 data points, which should be adequate for estimating incidence, mortality, and incidence by stage (but not by age). However, this sample size may not be sufficient to predict rare events observed in the calibration targets.

To determine the sample size needed to predict rare events, smaller population sizes were compared to a "true" population of 1 million. We calculated the maximum error (standard error, SE) in predicting likelihood for fitting the incidence, incidence by stage, and mortality (i.e., the maximum error for any of the targets), using the 1-million-person simulation as a benchmark. The same simulations were run 100 times to account for the stochastic nature of the process, and the mean errors were calculated (Figure 4).

As illustrated in the figure below, a sample size of over 20,000 yields a mean SE within 5%.

Figure 4: Mean percentage error across all calibration targets with different population sizes



While the sample size of 20,000 people was decided to be appropriate to predict the log-likelihood within the SE of 0.05, additional comparison on the parameter distributions retrieved with the larger population samples (100,000) and smaller population samples (50,000 and 20,000) were conducted. No difference in calibrated distributions of parameters were observed by the sample size. The smaller population size of 50,000 men was used for the model warm up with the final calibration running the population of 200,000 men.

Priors and parameters sampling

The mean time for cancer progression across the stages and the shape parameter of the Weibull distribution were sampled from a normal distribution, with left truncation at zero. All other parameters, including probabilities, were sampled from a beta distribution. Weak priors were set up in the Phase 2 calibration work for the following:

- The symptomatic presentation rate for each more advanced stage must be at least equal to that of the previous stage.
- The probability of symptomatic diagnosis is set at more than 0.2 for Stage 1, between 0.05 and 0.3 for Stage 2, between 0.2 and 0.7 for Stage 3, and less than 0.5 for Stage 4.

- The mean progression time across stages and the shape parameter of the Weibull distribution increase with advancing stages.
- Prostate cancer progresses more rapidly with higher GGG, meaning that the multipliers for higher GGG are smaller for more advanced GGG compared to less advanced groups.

6.7.4. Calibration outcomes

Initially, posterior samples were to be weighted using normalised importance weights, defined as the likelihood of each parameter set divided by the sum of all likelihoods. However, the effective sample size (ESS)—which reflects the number of independent samples effectively contributing to the posterior—was found to be very small when applying tempering. This indicates that only a few samples carried disproportionately high weights.

Given the extremely low ESS, importance weighting was adjusted, as otherwise it would have further reduced sample diversity and potentially introduced bias into the posterior estimates[31, 32]. Instead, we opted to explore the posterior distribution visually, as recommended by Fan and Sisson (2018)[33], examining how the likelihood varied across the parameter space (Supplementary A). Based on this inspection, we calculated probabilities after applying a temper factor, to have a 0.99 probability of having sampled parameters with a likelihood above –2 000 to approximate the posterior distribution. This approach is technically similar to approximate Bayesian computation (ABC) as it uses a likelihood cut-off to keep the parameter draws that are "close enough" to the data[33].

The model predictions fitted better to the calibration targets in the Phase 2 work (with the updated inputs) compared to Phase 1 work. Similar to Phase 1, the model overpredicted mortality; however, this difference was smaller. The fit to data with the random draw of 10 parameters from 100 with the best-fit is presented below (Figures 5a-d), with the sojourn time in the best-fit parameters predicted in the range of 9.4-11.2 years.

Figure 5a. Prostate cancer incidence rate in population (grey —observed, red— predicted)

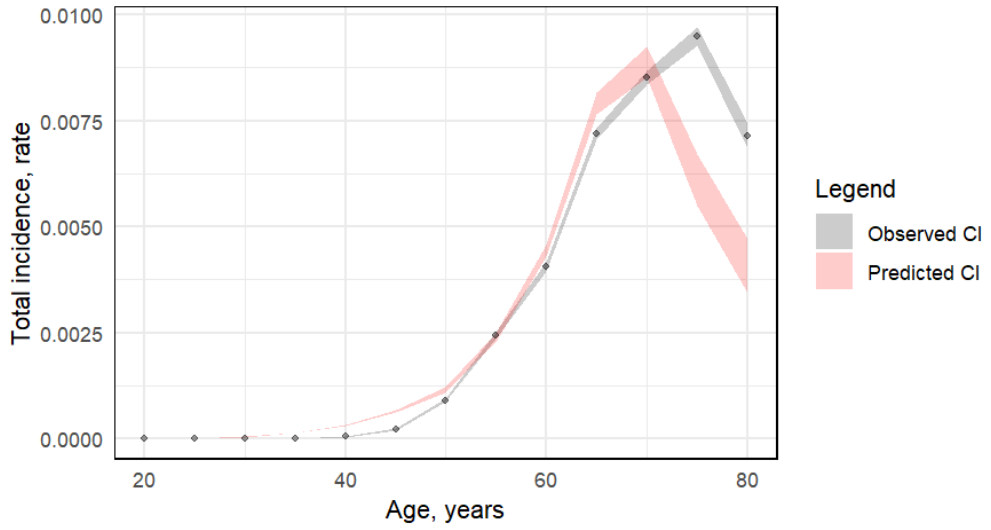


Figure 5b. Prostate cancer mortality rate in population (grey—observed, red—predicted)

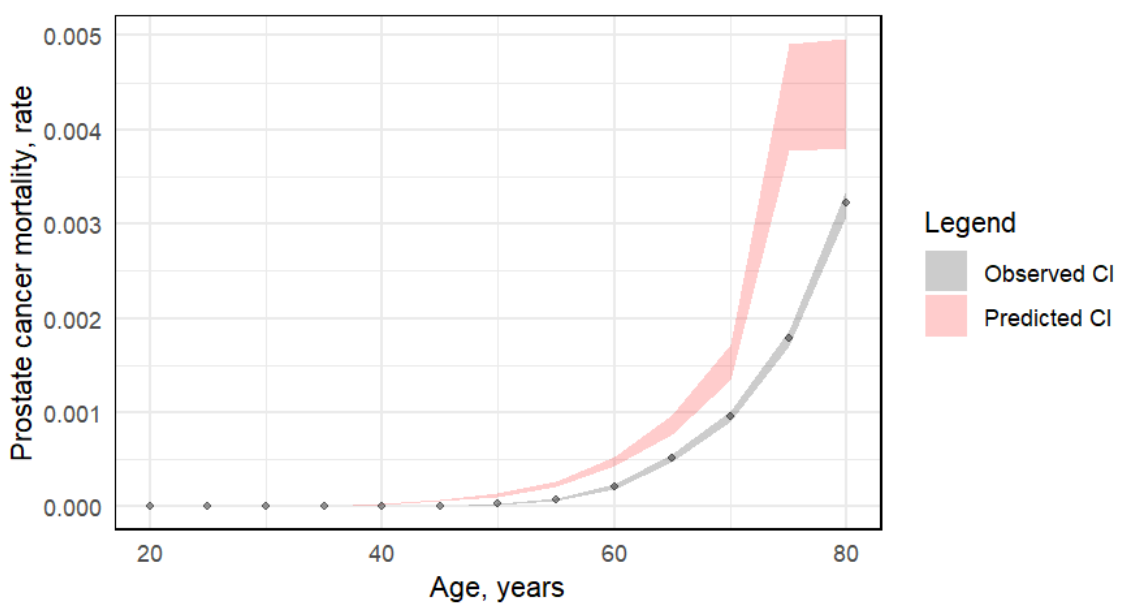


Figure 5c. Stage 1 prostate cancer incidence rate in population (grey – observed, red— predicted)

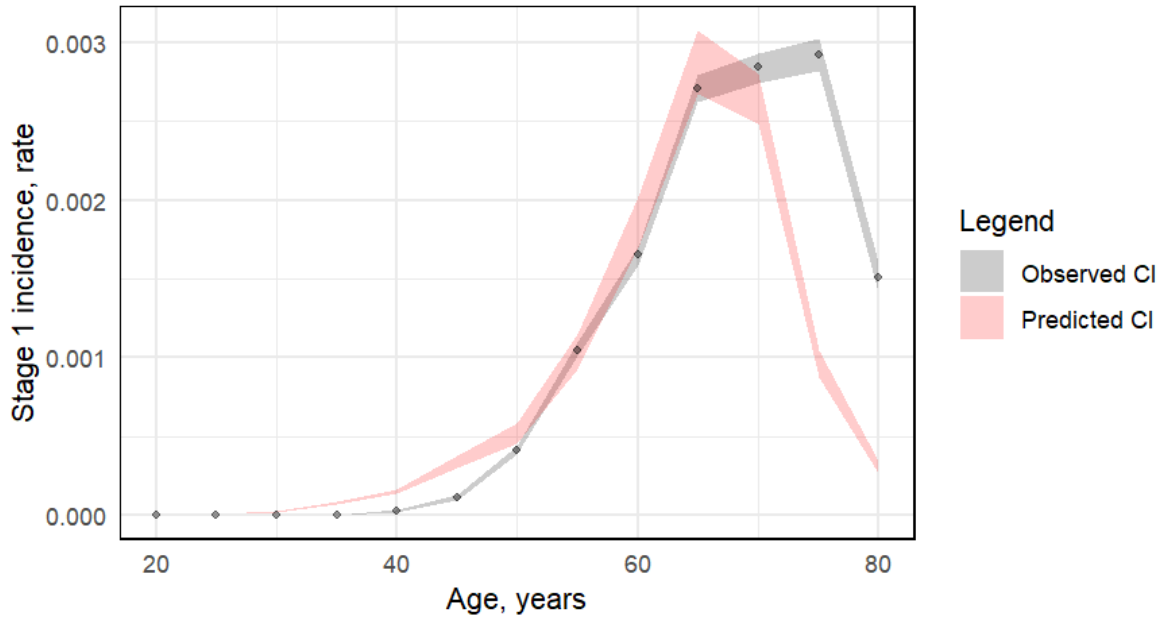


Figure 5d. Stage 2 prostate cancer incidence rate in population (grey —observed, red— predicted)

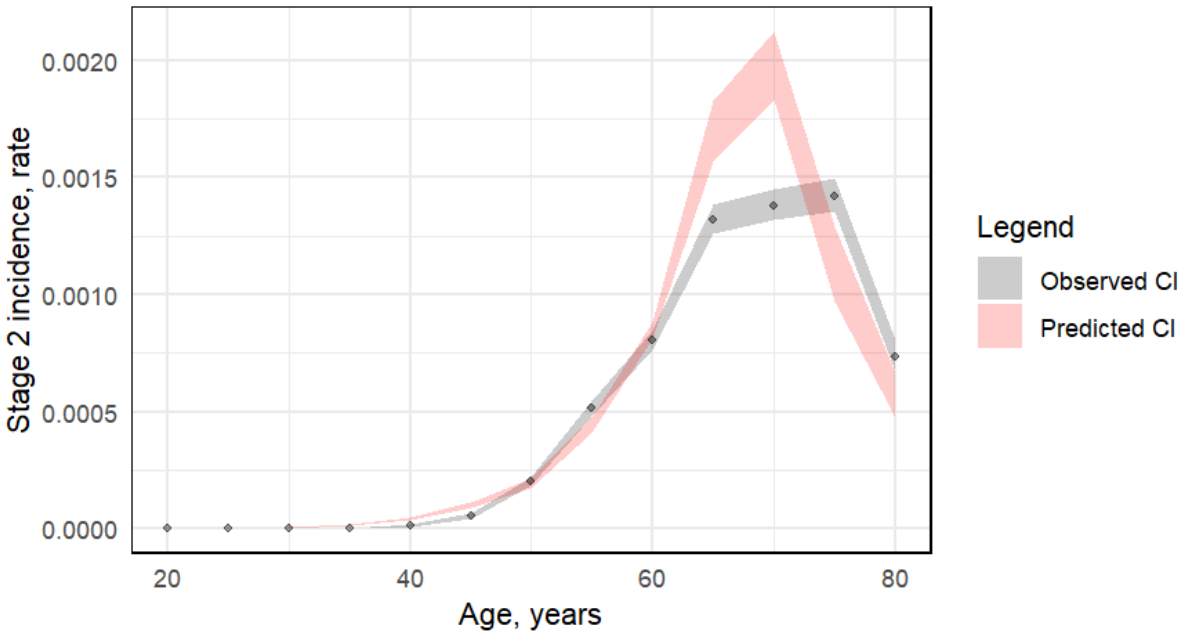


Figure 5e. Stage 3 prostate cancer incidence rate in population (grey —observed, red— predicted)

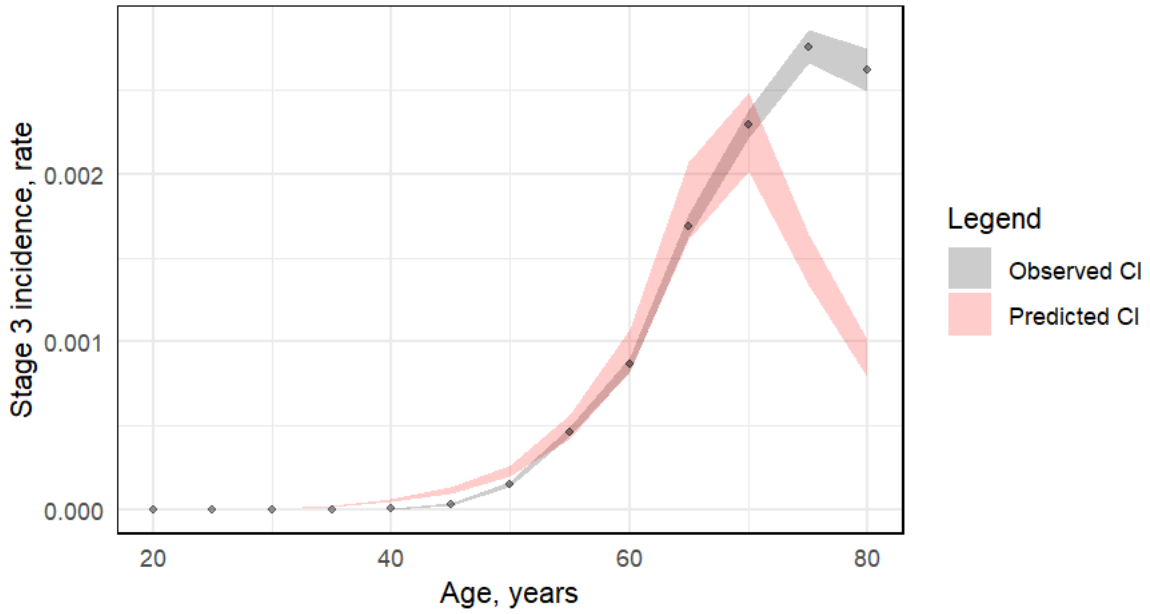


Figure 5f. Stage 4 prostate cancer incidence rate in population (grey —observed, red— predicted)

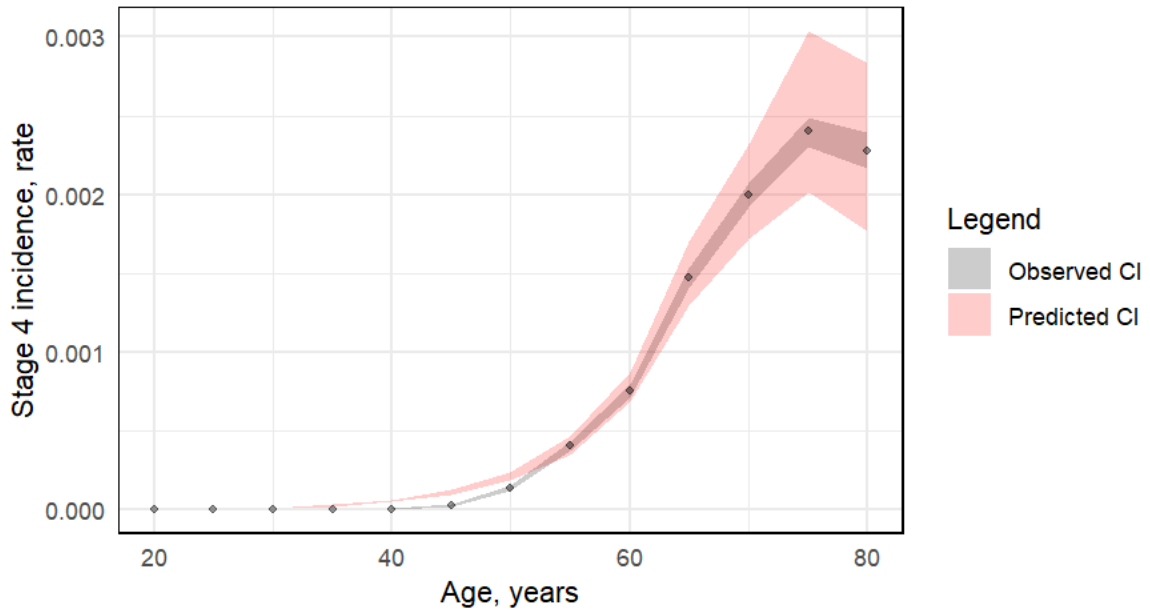


Figure 5g Distribution by GGG for Stage 1-2 prostate cancer (black —observed, red— predicted)

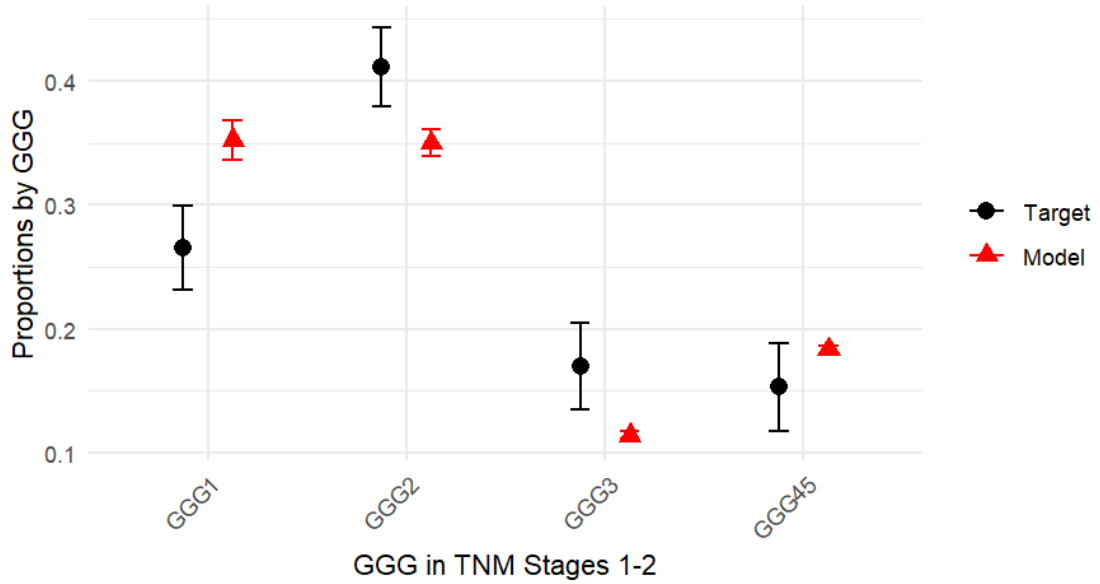


Figure 5h. Distribution by GGG for Stage 3 prostate cancer (black —observed, red— predicted)

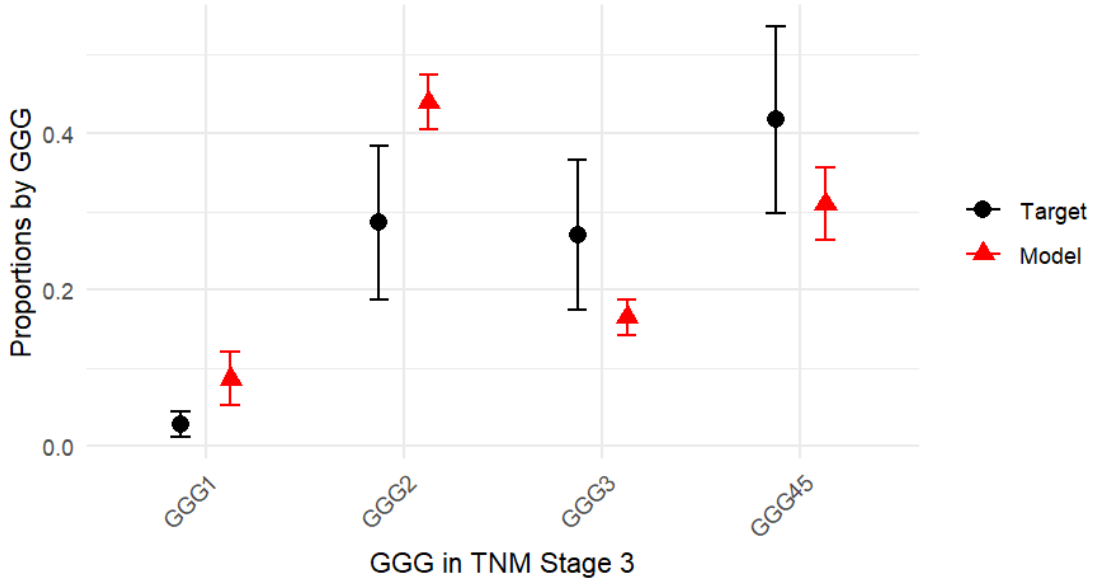
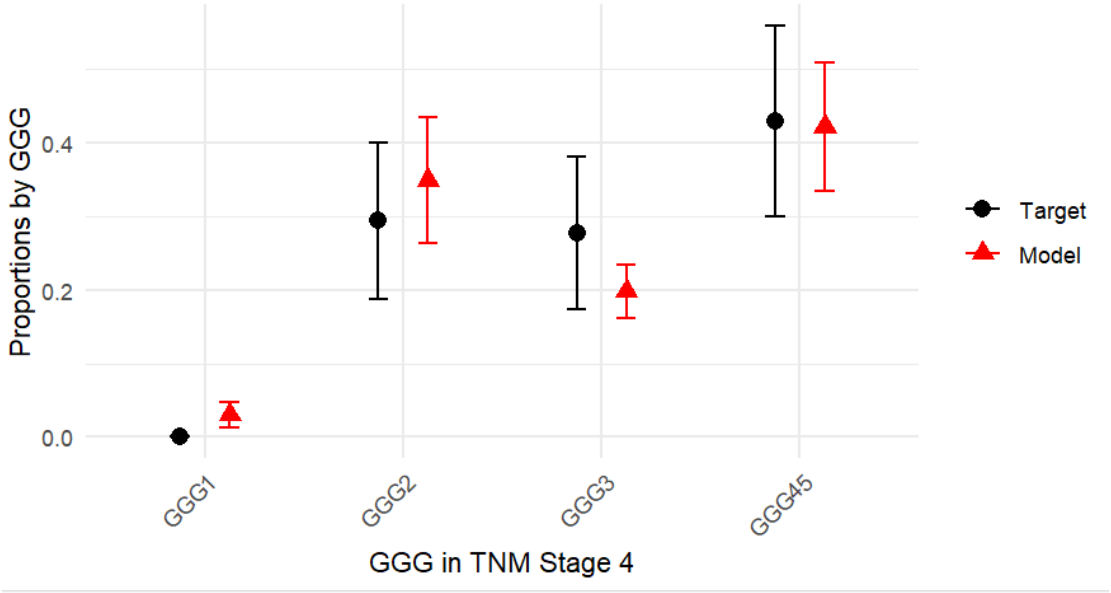


Figure 5i. Distribution by GGG for Stage 4 prostate cancer (black —observed, red—predicted)



6.7.5. Calibrated parameters

The two best-fit calibrated parameters with the similar likelihood (and so the largest probabilities to be sampled in the probabilistic analysis) are reported in Table 15.

Table 15: Best-fit calibrated NHD parameters

Parameter	Parameter Name	Best fit 1, p=0.34	Best fit 2, p=0.16
Probability of cancer onset at age 30, white, no BRCA carrier and no familial history	P.onset	0.00011	0.00016
Coefficient in the equation on probability of prostate cancer onset by age (i.e., a risk at age X = P.onset *C.onset.age^ (age X-20)	C.onset.age	1.148	1.148
Annual probability of being symptomatically or opportunistically diagnosed at stage 1	P.sympt.diag_St1	0.008	0.007
Annual probability of being of being symptomatically or opportunistically diagnosed at stage 2	P.sympt.diag_St2	0.055	0.083
Annual probability of being of being symptomatically or opportunistically diagnosed at stage 3	P.sympt.diag_St3	0.514	0.250
Annual probability of being of being symptomatically or opportunistically diagnosed at stage 4	P.sympt.diag_St4	0.833	0.608
Probability to die undiagnosed after age 70 years	P.ungiag.dead	0.058	0.123
Coefficient for annual decrement in clinical diagnosis rate after age of 70 (multiplier)	Symp.decr	0.799	0.803
Mean time of progression from stage 1 to stage 2 (for those with GGG 1)	Mean.t.StI.StII	27.549	40.253
Mean shape for Weibull distribution time from stage 1 to stage 2 (for those with GGG 1)	shape.t.StI.StII	2.640	18.837
Mean time of progression from stage 2 to stage 3 (for those with GGG 1)	Mean.t.StII.StIII	14.337	14.566
Mean shape for Weibull distribution time from stage 2 to stage 3 (for those with GGG 1)	shape.t.StII.StIII	3.250	3.311
Mean time of progression from stage 3 to stage 4 (for those with GGG 1)	Mean.t.StIII.StIV	4.634	6.401
Mean shape for Weibull distribution time of progression from stage 3 to stage 4 (for those with GGG 1)	shape.t.StIII.StIV	8.133	7.601

Coefficient adjusting the means for progression for GGG2 compared to GGG1	k.WB.GGG.2	0.543	0.526
Coefficient adjusting the means for progression for GGG3 compared to GGG1	k.WB.GGG.3	0.511	0.355
Coefficient adjusting the means for progression for GGG4_5 compared to GGG1	k.WB.GGG.4.5	0.506	0.350

6.8. Modelling screening pathway

The pathway to screening diagnosis is presented in Figure 6. The model assumes that individuals in the target population are invited to participate in screening through mailed invitations. Those who accept the invitation attend screening at their local GP practice, where the procedure is carried out by a nurse.

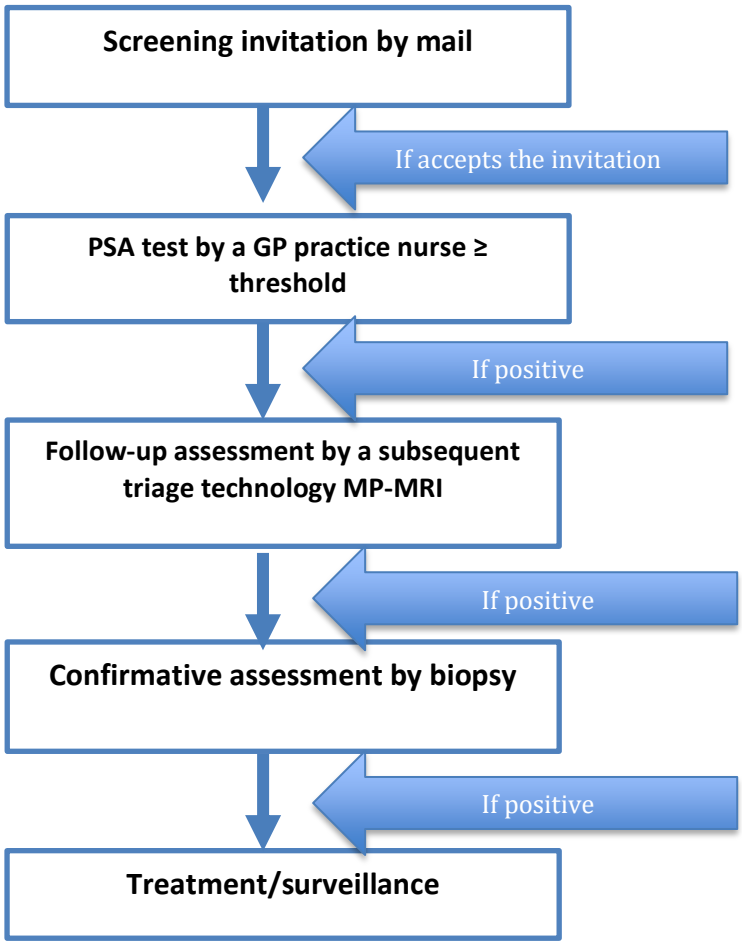
If the PSA test result is positive, the individual is referred for a multi-parametric MRI (mpMRI). If the mpMRI result is positive, the individual is then referred for a biopsy. A positive biopsy result leads to treatment and surveillance.

The cohort of individuals with positive biopsy results consists of:

- True positives, including both those with life-threatening prostate cancer and those who are overdiagnosed (i.e. cancers that would not have become clinically significant), and
- False positives, referring to individuals with benign conditions or diseases other than prostate cancer who are incorrectly diagnosed as having the disease.

The treatment and surveillance components of the model are not simulated explicitly. Instead, each individual diagnosed with cancer is assigned corresponding costs, survival estimates (by age, stage, and years since diagnosis), and HRQoL decrements. The latter are represented as stage-specific multipliers applied to baseline HRQoL values from the no-cancer population. The model assumes no difference in treatment or post-diagnosis surveillance costs based on the route to diagnosis. Differences between screen-detected and symptomatically detected cases arise from the higher proportion of early-stage cancers among screen-detected cases, which are less costly to treat and associated with higher survival rates.

Figure 6. Screening pathway in the model



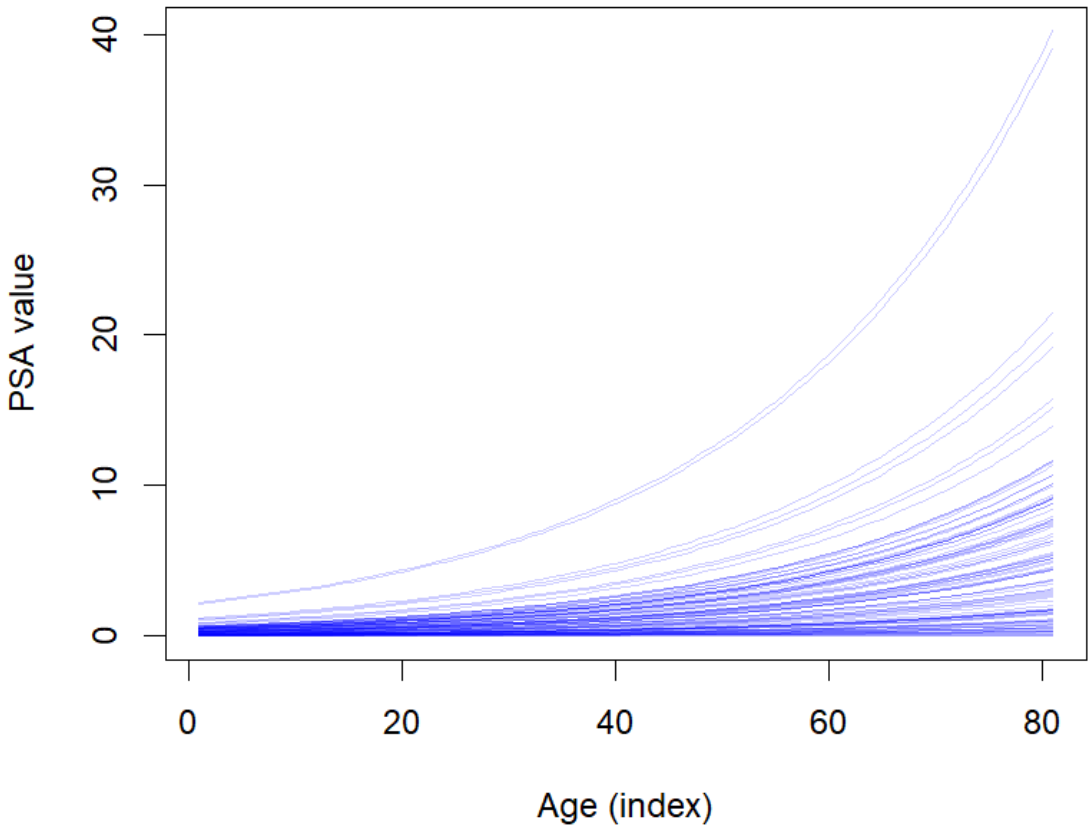
6.8.1. Sensitivity and specificity of the PSA test

Most of the literature on this topic focuses on the accuracy of the PSA test in symptomatic populations. Consequently, despite its age, data from the PCPT trial [34] were utilised, as the results from symptomatic and screening populations differ significantly. For instance, a systematic review and meta-analysis of symptomatic individuals found a sensitivity of 0.93 (95% CI 0.88–0.98) at a PSA threshold of ≥ 3ng/mL[35]. In contrast, Ankerst et al. (2014) reported calculated thresholds of 0.184 and 0.374 for low-grade and high-grade prostate cancer, respectively[34]. The PCPT trial is considered more reliable because PSA testing was conducted in healthy men, and biopsies were performed at the study's conclusion, enabling a more accurate comparison.

The PCPT report by Ankerst et al. (2014)[34] does not provide sensitivity estimates stratified by age, nor does it offer information on how individual PSA values change with age in men with or without prostate cancer. As a result, applying fixed test accuracy values directly in the model would overlook the potential impact of age on test sensitivity. Moreover, using flat sensitivity values would not adequately inform a model simulating risk-stratified screening.

To address this, the base case analysis used direct data from the PCPT trial on the distribution of men without confirmed cancer, with low-grade cancer, and with high-grade cancer by PSA level (Table 16). Since the mean age in the PCPT trial was 62, and test sensitivity varies with age, individual-level PSA values at age 62 were extrapolated to younger and older ages. This was done by applying an annual rate of change in PSA levels derived from Reza et al. (2020), which reported an increase in PSA from 1.47 at age 62 to 3.05 over a 20-year period[36]. The change in the PSA levels by age in 100 randomly selected men is reported on the Figure 7. Based on expert input from the clinical expert Dr Derek Rosario, this age-related trend was assumed to apply to men without cancer, as well as those with low- and high-grade cancer in the base case analysis. A scenario analysis was also conducted under the assumption that the age-related PSA trend applies only to men without cancer.

Figure 7. Change in the PSA by age



For stage 4 prostate cancer, the test was assumed to be significantly more sensitive, based on the premise that most patients at this stage would develop symptoms within a year and could therefore be considered symptomatic. To reflect this, data from Zheng et al. (2020) were used, which reported a test sensitivity of 0.973 for stage 4 cancer at a PSA threshold of 4 ng/ml)[37]. It was assumed that, by this stage, few patients would still have PSA levels between 3 ng/ml and 4 ng/ml, making the 0.973 estimate a reasonable proxy.

Table 16: Proportion of men with different PSA results in three health states in the PCPT trial[34]

Threshold	No cancer state	Low-grade cancer	High-grade cancer
0.9	0.409	0.271	0.169
1.9	0.318	0.392	0.228
2.9	0.125	0.123	0.154
3.9	0.029	0.031	0.075
4.9	0.065	0.113	0.154
5.9	0.029	0.047	0.118
6.9	0.012	0.010	0.051
10	0.015	0.015	0.051
Total	1	1	1

6.8.2. Sensitivity and specificity of the mpMRI

The sensitivity and specificity for the mpMRI and biopsy were based on a paper by Ahmed et al. (2017)[38]. This can be seen in Table 17.

Table 17: Accuracy of diagnostic mpMRI

Test	Sensitivity	Specificity
mpMRI	0.88 (0.84-0.91)	0.45 (0.39-0.51)

Legend: MP-MRI -Multi-Parametric Magnetic Resonance Imaging

6.8.3. Sensitivity and specificity of the LATP Biopsy

In the original modelling plan, TRUS biopsy was assumed to be the method used for following up screen-positive and symptomatic cases. However, after consultations with stakeholders, who indicated that LATP biopsy is increasingly replacing TRUS biopsy in the UK, and considering the data presented in the National prostate cancer Audit [39], it was decided to use LATP biopsy in the model instead.

We used two systematic reviews by Goldberg et al. (2020)[40] and Kanagarajah et al., (2023)[41] to extract the detection rate for "any prostate cancer" and "Clinically significant prostate cancer (Gleason≥3 + 4)" reported in TRUS and LATP studies, respectively. The "Clinically significant prostate cancer (Gleason≥3 + 4)" is prostate cancer in GGG2+. We used the Ahmed et al. (2017) PROMIS trial[38] to assess the sensitivity of TRUS biopsy for these cancers. We adjusted

the sensitivity of LATP biopsy using the proportional difference in the detection rates and assumed that the higher detection rate is only related to different sensitivities. This assumption is fixed as we have no data to inform otherwise. We assume that the specificity of the LATP biopsy is equal to specificity of TRUS biopsy (Table 18).

Table 18: Accuracy of LATP biopsy

Test	Sensitivity GGG1	Sensitivity GGG2+	Specificity
LATP Biopsy	0.52 (0.46-0.60)	0.85 (0.78-0.93)	0.98 (0.96-1)

6.8.4. Screening uptake

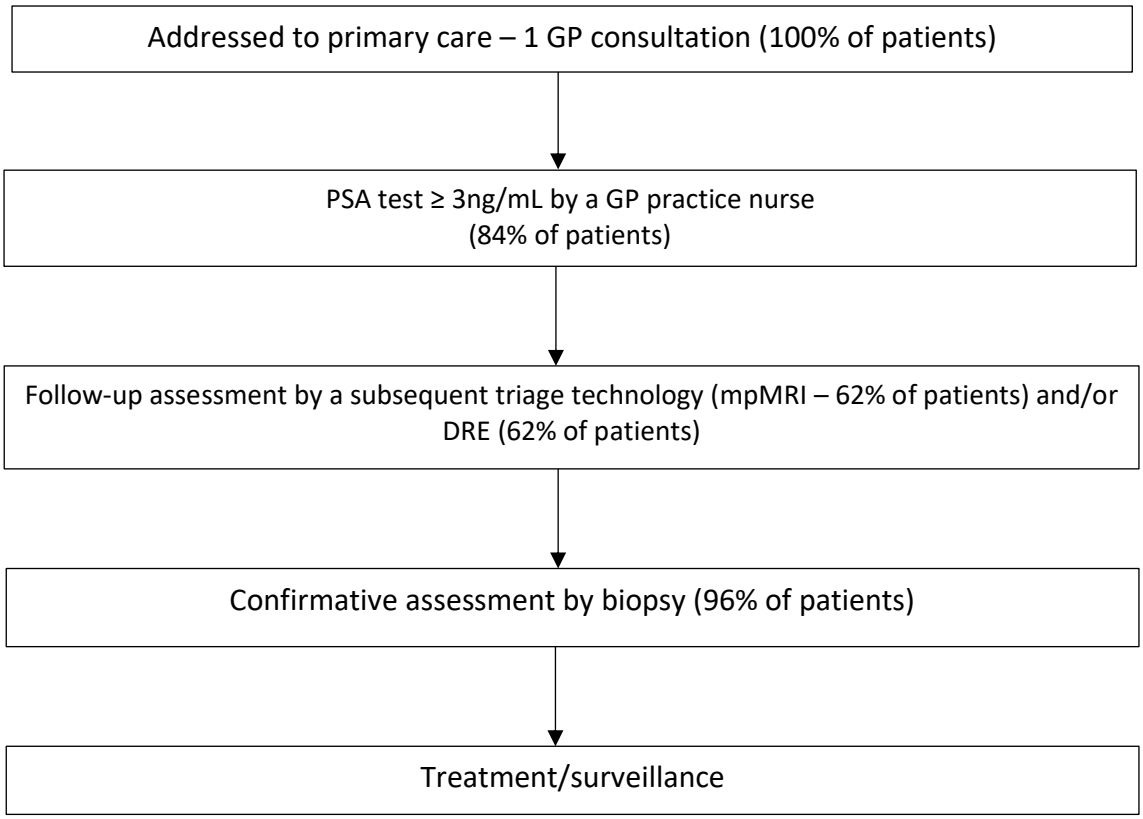
The base case screening uptake was used from the CAP trial: 36% for the PSA test and 85% for the follow up biopsy[28]. It was assumed that mpMRI will have the same uptake as the biopsy. When evaluation considered the net benefit of screening, 100% uptake was considered as well as 100% follow-up.

6.9. Modelling diagnostic pathway for patients diagnosed outside of organised screening

The diagnostic pathway for symptomatic and opportunistic cases was informed by previous decision models on prostate cancer screening (Keeney et al., 2022[42], NICE Guidelines on PC, 2021[43]).

In the model, the diagnostic process for patients diagnosed outside the organised screening programme was not modelled explicitly. Instead, data on average resource use (Figure 8) were used to calculate and assign the mean diagnostic costs for all patients diagnosed through symptomatic or opportunistic pathways. These diagnostic costs are reported in the section on Diagnosis Costs.

Figure 8: Prostate cancer diagnostic pathway outside of organised screening programme



Legend: DRE – digital rectal examination, GP – general practitioner, mpMRI - multiparametric magnetic resonance imaging.

6.10. Utilities

The methodology for calculating individual utility values and the decrement in utility due to age was detailed in the "Modelling Changes in Phenotypic Characteristics by Age" section.

6.10.1. Utility multipliers

In determining the utility multipliers associated with prostate cancer diagnoses, we referred to the 2024 NHS Cancer Quality of Life Survey[44]. This entailed examining EQ-5D scores across stages 1-4 for prostate cancer (48,063 patients), and a collective group of cancers (190,464 patients), collected at 18 months since diagnosis. The weighted average utilities were computed as follows: 77.83, 95% CI (77.42- 78.26) for prostate cancer.

Utility multipliers for prostate cancer by stage were calculated starting with evaluation of year 1 utilities, reflecting the assumption that these are lower than in subsequent years and that 18-months utilities will represent utilities at year 2 in the model. Specifically, the first-year multiplier was derived from the difference in QALYs between the first and second years (0.98), as reported by Noble et al. (2020)[45]. Then, utilities for 70-year-old men from the EEPRU report by Alava et al. (2022)[46] (0.841) were used as reference values for the general population without prostate cancer. Using this reference value, utility multipliers for stages 1 to 4 were calculated separately for the first year and the following years. Based on consultations with the clinical expert Derek Rosario, who suggested that utilities in the year 1 may not be different for patients with prostate cancer compared to the following years, we applied flat utility values by stage in the base case scenario, and utilities adjusted by year since diagnosis in the scenario analyses.

This approach updates the methodology from the Phase 1 report, where utilities from the HSE 2018 dataset were used as reference values, leading to multipliers above one for early prostate cancer. Recognising potential population-wide changes in utility values, particularly post-COVID-19, and the improbability of prostate cancer diagnosis increasing population utilities, we adopted the more recent EEPRU reference values.

The derived utility values are presented in Table 19.

Table 19: Utility multipliers

Stages	Reported Utilities, NHS Cancer Quality of Life Survey (2024)		Calculated multipliers, year 1 (*scenario only*)	Calculated multipliers, years 2 and further
	All cancers	Prostate	Prostate	Prostate
1	76.18 (76.01;76.37)	80.51	0.94 (0.94;0.95)	0.96 (0.95;0.96)
2	73.97 (73.75;74.20)	80.56	0.94 (0.94;0.95)	0.96 (0.95;0.96)
3	73.74 (73.50;73.99)	77.59	0.91 (0.90; 0.91)	0.92 (0.92; 0.93)
4	69.02 (68.66;69.39)	69.72	0.82 (0.81; 0.82)	0.83 (0.82; 0.84)

The impacts of disease on utilities are assumed to last over the patients' lifetime, while the patient receives surveillance and cancer impacts survival.

6.10.2. Utility decrements

Those who undergo a biopsy experience a small temporary utility decrement. This utility decrement is 0.013 based on Kasivisnathan et al. (2018)[47]. This is a comparison between baseline and 24 hours after the procedure. By 30 days post-intervention, the HRQoL has more than fully recovered. On the other hand, Li et al. (2019)[48] a systematic review on the disutility associated with cancer screening programmes, found that there was no disutility of prostate cancer screening, based on two previous studies. Taking this into account and that the longest side effects related to biopsy reported in the ProtecT trial lasted 5 days, the utility decrement period of biopsy was 1 week. The utility decrement was applied for both screen-diagnosed and symptomatically diagnosed patients in the first year of their diagnosis.

6.11. Costs estimation

6.11.1. Diagnosis costs

The unit costs for screening and symptomatic diagnosis and surveillance costs are available in Table 20 and, where relevant, inflated to 2022/23 costs.

PSA test cost

The PSA test cost was based on Mowatt et al., (2013)[49], which is the study used by the NICE guidelines (2019)[50]. This cost was inflated from 2009/10 price year to the 2022/23 price year using PSSRU.

MP-MRI cost

The costs were based on Mowatt et al. (2013) [49] and inflated to 2022/23 costs. The total mpMRI cost was calculated as £316.01.

Transperineal biopsy cost

The cost of transperineal biopsy was based on the calculation used in Nicholson et al. (2015)[51], combining a biopsy and histopathology cost. The transperineal biopsy cost was based on outpatient procedures for a LATP from the NHS reference costs of £1,138 (2022/2023). The histopathology part of this cost was based on an estimate from a laboratory manager at the Department of Cellular

Pathology at the North Bristol NHS Trust (NCCC, 2014)[52]. This was then inflated to 2022/2023 costs.

Table 20: Diagnosis Unit Costs

Items	Unit Costs (uninflated), GBP	Inflated Costs, GBP
PSA Test Cost	5.91	7.81
Practice Nurse Consultation	12.00	15.86
PSA Total	**17.91**	**23.67**
MP-MRI Total	**239.06**	**316.01**
Transperineal Biopsy Cost	1,138	1,138
Histopathology Cost	112.79	137.89
Biopsy Total	**1,250.79**	**1,275.89**
DRE Total	**49.00**	**49.00**
GP Consultation Cost Total	**49.00**	**49.00**

The Legend: DRE - Digital Rectal Examination; GP – general practitioner; MP-MRI - Multi-Parametric Magnetic Resonance Imaging

6.11.2. Symptomatic diagnosis costs

The symptomatic diagnosis costs are a weighted average of diagnostic procedures that people receive. The proportions are in Table 21. When using the NPCA, the most recent annual report in which the data were available was used[39].

The cost of primary care diagnosis for symptomatic patients was based on resource use from NPCA and Merriel et al. (2024)[39, 53]. The NPCA was used as the main source, as it is based on more recent data. Merriel et al.(2024)[53] relies on older sources, so it was only used in cases where the NPCA did not report, or had low completeness of data. Combining these proportions with the unit costs from Table 20 gives an estimated cost of symptomatic diagnosis of £1,520.

Table 21: Symptomatic Diagnosis Proportions

Diagnostic Test	Proportion	Source
PSA Test	0.84	Merriel et al., (2024)[53]
mpMRI	0.62	NPCA Annual Report (2019)
Biopsy	0.96	NPCA Annual Report (2020)[54]
GP Visit	1	Assumption
DRE	0.62	Merriel et al., (2024)[53]

The Legend: DRE - digital rectal examination; GP – general practitioner; mpMRI - Multi-Parametric Magnetic Resonance Imaging

6.11.3. Additional screening costs

In addition to the diagnostic screening costs, there is also a cost of an invite to the PSA test, which was estimated as £9.17 based on the invitation costs for bowel cancer screening[55]. These costs included invitation letters, reminder letters for non-responders, helpline services, postage, staff time, and overheads. Given that establishing an organised screening programme entails substantial fixed and operational costs, and that NHS overheads are generally estimated to be comparable to direct costs, this assumption was considered reasonable.

6.11.4. Surveillance costs

Surveillance costs involve monitoring for those who have already received radical treatment for their prostate cancer. NICE guidelines (2021)[43] recommend that PSA levels are checked for all people with prostate cancer who are having radical treatment no earlier than 6 weeks after treatment, at least every 6 months for the first 2 years, and then at least once a year after that[43]. So, these costs were calculated based on PSA test cost. Average surveillance costs were based on those who receive treatment at each cancer stage from Wills et al., (2024)[56]. The surveillance costs were added from year 2, based on the proportion of patients who received treatment. These costs remain the same from year 4 onwards (Table 22).

Table 22: Surveillance Costs by Year and Stage

Year	Stage 1 (£)	Stage 2 (£)	Stage 3 (£)	Stage 4 (£)
Year 1	0	0	0	0
Year 2	17.90	32.48	36.27	23.67
Year 3	17.90	32.48	36.27	23.67
Year 4	8.95	16.24	18.13	11.84

6.11.5. Active surveillance costs

Active surveillance involves monitoring those with cancer who do not receive immediate treatment, based on the NICE guidelines (2021) protocol for active surveillance[43]. In the first year, this involves a PSA test every 3 to 4 months, a DRE at 12 months and an mpMRI at 12 to 18 months. In year 2 and afterwards, a PSA test is conducted every 6 months, and a DRE is conducted every 12 months. The cost of a DRE is considered to be the cost of a GP appointment. Average active surveillance costs were based on those who do not receive treatment from Wills et al., (2024)[56]. The active surveillance will be costed up to the patient's death from other causes or the year of cancer death (and so the assumed cancer progression or relapse). The costs remain the same from year 3 onwards, Table 23.

Table 23: Active surveillance costs by year and stage

Year	Stage 1 (£)	Stage 2 (£)	Stage 3 (£)	Stage 4 (£)
Year 1	82.02	41.40	30.86	65.93
Year 2	256.49	129.48	96.49	206.18
Year 3	59.93	30.25	22.55	48.17

6.11.6. Treatment costs

To estimate the stage costs, Wills et al. (2024) was used[56]. The paper estimates the costs of initial cancer treatment based on stage at diagnosis. Considering that the model does not explicitly simulate different follow-up strategies, we used the average cost for the whole cohort, regardless of whether they received treatment

or not. The costs in this paper were assigned using NHS reference costs from 2017/18. Therefore, the costs were inflated to 2022/23 figures using the PSSRU (Table 24).

Table 24: Prostate cancer stage costs, not including updated SACT costs

Stage	Cost (£) (2017/2018)	Inflated costs (£) (2022/2023)
1	2,216	2,482
2	4,335	5,062
3	5,179	6,047
4	2,629	3,070
Unknown	2,216	2,482

These costs reported by Wills et al. (2024)[56] were only applied for the first year of treatment. Therefore, we used Noble et al. (2020)[45], which is a study based on the ProtecT trial, to extrapolate the costs from the first year to year 10. It was assumed that beyond year 10, the annual costs are equal to those of year 10 for all living patients. We assume that stages 1 and 2 follow the same trend over time.

However, based on recommendations from clinicians, it was suggested that stage 3 and 4 treatment costs have changed dramatically in recent years due to new innovative therapies. This change in systemic anti-cancer therapies (SACTs) is supported by the National prostate cancer Audit (NPCA)[39, 54], in which apalutamide and enzalutamide now take up a much larger proportion of cancer treatment compared to just a couple of years ago.

Based on this, the SACT costs were recalculated for stages 3 and 4. These costs were based on a Resource Impact Template (RIT) of Enzalutamide from NICE, NPCA, Wills et al.(2023)[39, 54, 56, 57] and the British National Formulary (BNF). SACTs were assumed to all be supplied alongside ADTs (Table 25).

Table 25: Cost per treatment, cycle length and source

Treatment	Treatment Cost (£)	Treatment Distribution	Cycle length	Source
ADTs	944	100%	As long as receiving treatment	NICE RIT[57]
Docetaxel + G-CSF	1613	37%	18 months	NICE RIT, BNF[57, 58]
Enzalutamide	35,551	37%	36 months	BNF, NICE RIT[57, 59]
Apalutamide	35,555	26%	20.5 months	BNF, Agarwal et al., (2019)[60, 61]

Legend: BNF - British National Formulary; RIT - Resource Impact Template.

These calculated costs were then weighted by the percentage of people who receive treatment in stages 3 and 4. This was based on Wills et al. (2024) of 1.8% and 28% for stages 3 and 4, respectively[56].

Table 26 shows the estimated cost of SACTs for the average person in stage 3 and 4 by year.

Table 26: Calculated SACT costs for stages 3 and 4

Stage	Year 1 (£s)	Year 2 (£s)	Year 3 (£s)
Stage 3	431	372	254
Stage 4	6703	5781	3947

As the initial Wills et al. (2024)[56] costs being used already included SACT costs, these had to be subtracted from the initial costs for the first year. In the case of stage 3, so few people received SACTs as treatment that these costs did not affect the overall costs. However, a portion of stage 4 costs were subtracted to account for the existing stage 4 costs. This portion was based on an estimated reading from Figure 3 in Wills et al. (2024) of the SACTs for all tumours in stage 4.

Figures 9 & 10 show costs over time for each stage. These costs were then combined with the average surveillance and active surveillance costs in Table 27.

Figure 9: Extrapolated costs for stages 1 & 2

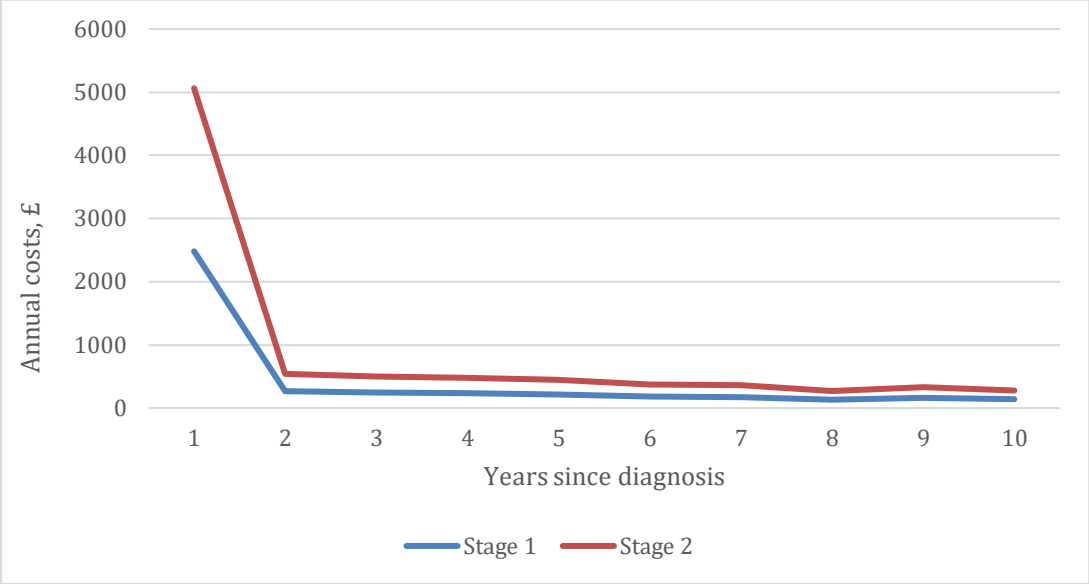


Figure 10: Extrapolated costs for stages 3 & 4, including updated SACTs

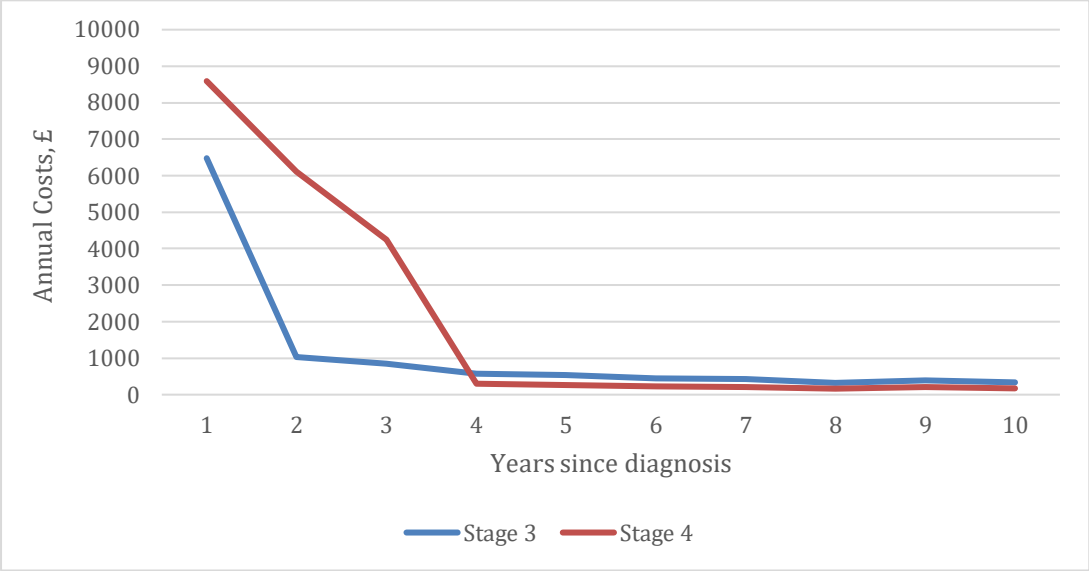


Table 27: Annual costs applied from time of diagnosis

Year	Stage 1	Stage 2	Stage 3	Stage 4
Year 1	2564.02	5103.40	6508.75	8655.69
Year 2	541.94	707.64	1156.25	6341.48
Year 3	323.60	563.98	911.35	4323.08
Year 4	305.02	528.10	616.00	352.00
Year 5	287.28	491.93	572.79	330.07
Year 6	249.28	414.41	480.19	283.08
Year 7	244.21	404.08	467.85	276.81
Year 8	201.14	316.23	362.91	223.55
Year 9	230.02	375.14	433.28	259.27
Year 10 onwards	207.22	328.64	377.72	231.07

6.11.7. Palliative care costs

In addition to this, patients were also assigned palliative care costs in the year of death. This is because Wills et al. (2024)[56] does not include palliative care costs in their cost estimations. The palliative costs in the Phase 1 report were only assigned to prostate cancer deaths and were based on Round et al. (2015)[62], which models the estimated costs of end-of-life care for various cancers, including prostate cancer in 2013/14 prices. However, in the Phase 2 work, palliative care costs were assigned to both prostate cancer and non-prostate cancer deaths based on data from Diernberger et al. (2023)[63]. These costs were inflated from 2017 to 2022/2023 figures (Table 28). These costs were assumed to equate to roughly 12 months of care and also assumed to encompass any cancer recurrence.

Table 28: Palliative care costs by age and mortality cause

	Original palliative care costs, £		Palliative care costs inflated to 2022/2023, £	
Age	Non-PC death	PC Death	Non-PC death	PC Death
60-64	12,490	15,915	15,072	19,206
65-69	12,160	13,342	14,675	16,101
70-74	10,901	12,835	13,155	15,489
75-79	10,783	12,720	13,012	15,350
80-84	10,028	10,920	12,101	13,178
85-89	8,993	9,039	10,853	10,908
90+	8,374	8,187	10,105	9,880

Legend: PC – prostate cancer

7. Model Validation and technical characteristics

7.1. Internal validation

Internal validation included code verification, data verification, and validation of model predictions against expected outcomes. Code verification for the NHD model was conducted independently by LM, AR, and MH. Data verification was performed by MH, with the checks detailed in Supplementary B. Internal validation of model predictions was carried out by DP and is reported in Supplementary C and D.

7.2. External validation

The model was validated to the intervention and control arms from the CAP[64] and ERSPC [65] trials, by running the simulation with deterministic parameters. To address stochastic uncertainty in the model, the simulation was run with a population of 10 million people.

7.2.1. CAP trial

The model's screening arm was designed to reflect the screening protocol used in the CAP trial. In the CAP trial, men aged 50 to 69 years from participating general practices were randomised, with screening offered to men in the intervention arm. The median age at screening among attendees was 59 years, so the model simulated a single screening event at age 59 to align with this.

Men with PSA levels ≥3.0 ng/mL were offered a standard 10-core transrectal ultrasound–guided biopsy. As mpMRI was not part of the CAP trial protocol, the model similarly assumed that all men with elevated PSA levels underwent biopsy directly. PSA and biopsy uptake rates were based on CAP trial data: 36% and 85%, respectively.

Although the CAP trial used TRUS, the current model uses LATP biopsy. To reconcile this difference, biopsy accuracy in the model was adjusted based on data from Ahmed et al. (2017), assuming 30% sensitivity for low-grade and 70% sensitivity for high-grade cancers (except for those with stage 4 cancer)[38].

7.2.2. ERSPC

The ERSPC represents a pooled analysis of several screening programmes, each with different designs, making it challenging to replicate precisely in a model. Most centres invited men aged 55–69 years for PSA screening, with a mean age at first screening of 62 years. The typical protocol involved two screening rounds with a 4-year interval, using a PSA threshold of 3.0 ng/ml.

However, screening intervals, age ranges, and diagnostic follow-up procedures varied across countries. Following a positive PSA test, participants typically underwent DRE, TRUS, and systematic prostate biopsies. Although mpMRI was not part of the ERSPC protocol, the model assumed the combined diagnostic accuracy of DRE and TRUS to be equivalent to mpMRI to simplify the modelling (with 100% uptake, considering no data on uptake for these tests were reported).

Biopsy procedures also evolved during the trial—from sextant biopsies to 10–12 core protocols. Biopsy accuracy was assumed to be similar to that in the current model. PSA and biopsy uptake rates were assumed to be 64% and 86%, respectively, based on average uptake rates reported across ERSPC centres. Where specific uptake data for follow-up procedures was unavailable, the same uptake rate as for biopsy was assumed.

7.2.3. Model predictions

The model predicted prostate cancer deaths as a proportion of those diagnosed in the non-screening arm, higher than in the CAP data[64]. The model however accurately predicted the impact of screening on all-cause mortality (calculated as the difference in average life years between the no-screening and screening arms for ages 59 to 74 years) (Table 29).

Table 29. Comparison to the CAP trial outcomes [64]

Reporting from	T1/T2 cancers in no screening arm, % of all cancers	Prostate cancer deaths in no screening, % from diagnosed (in 14 years of follow up)	Difference in all-cause mortality in 14 years in Screening vs no-screening arms, per 1,000 people
Model prediction, deterministic parameters	54.7%	15%	0.24
CAP trial[64]	60.0%	10.7%	0.23, 95%CI (0;0.46)

Table 30. Comparison to the ERSPC trial outcomes at 23 years of the follow up [66]

Outcome	Model prediction, deterministic parameters	ERSPC data[65]
Cumulative incidence of prostate cancer at 23 years of the follow up, rate ratio	1.31	1.30; 95% CI, 1.26 to 1.33
Rate ratio of prostate cancer mortality reduction with screening at 23 years of follow up	0.87	0.87; 95% CI, 0.80 to 0.95

The predicted rate ratio for cumulative incidence of prostate cancer cases in the model was similar to one reported in the ERSPC data with 23 years of the follow up (Table 30)[65, 66]. Similarly, and despite differences in predictions of prostate cancer deaths compared to the CAP trial, the model's prediction on the relative difference in prostate cancer deaths was identical to the ESRPC trial with 23 years of the follow up.

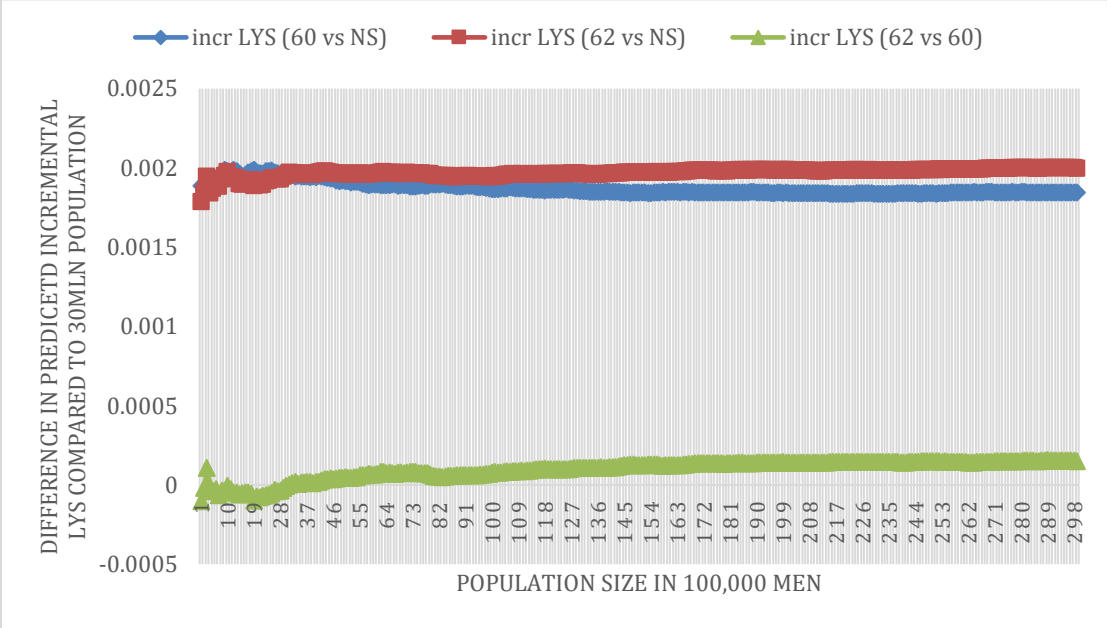
7.3. Determining the characteristics of the simulations

7.3.1. Addressing stochastic uncertainty

The number of simulated patients was determined using a graphical approach to assess the stabilisation of key outcomes, specifically incremental costs and incremental LYS. This approach was used to minimise stochastic uncertainty associated with random sampling. Differences in incremental outcomes across population samples were compared to the outcomes estimated for the entire England male population, assumed to be approximately 30 million. The figures

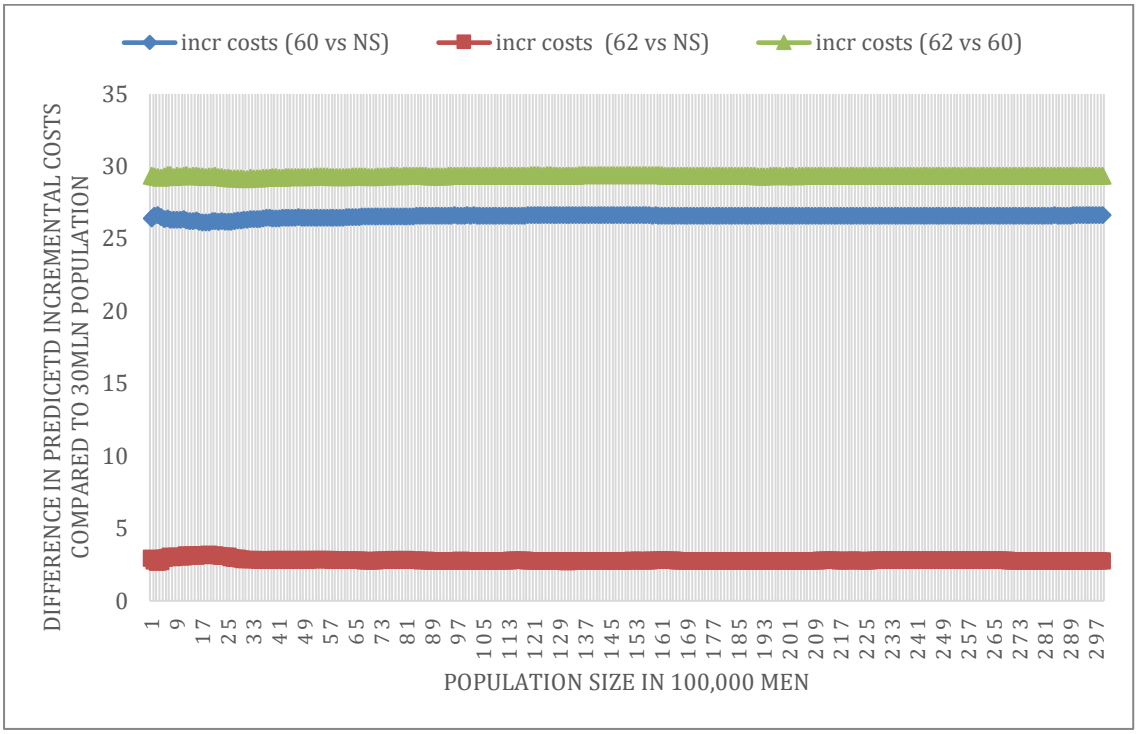
below present the plots used to evaluate outcome stability and guide the choice of an appropriate sample size. Based on this assessment, a minimum population size of 5 million was defined for the model to be used in the base case analysis (with 10 million preferred). However, due to the high computational burden of running such a large population, a smaller population of 1 million men was used in scenario analyses and probabilistic analyses.

Figure 11: Difference in incremental outcomes by population sample: life years saved



Legend: incr LYS –incremental life years saved; NS - no organised screening;

Figure 12: Difference in incremental outcomes by population sample: total costs



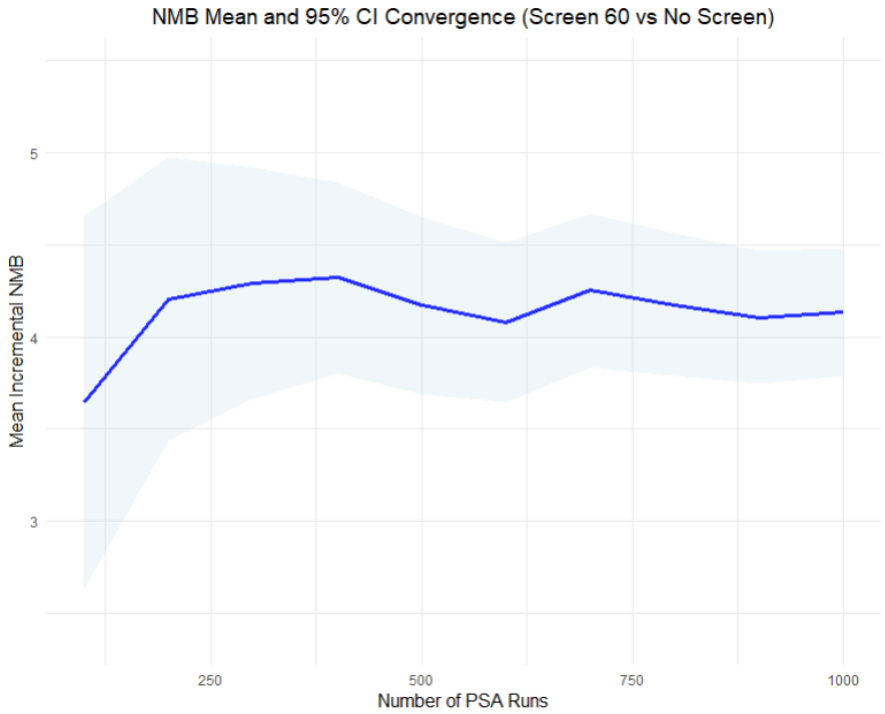
Legend: incr costs - incremental costs; NS - no organised screening;

7.3.2 Assessing convergence in probabilistic sensitivity analysis

Probabilistic sensitivity analysis quantifies parameter uncertainty in decision-making by sampling parameters from their respective probability distributions, rather than relying solely on mean or median values. Whilst literature and Health Technology Assessment (HTA) bodies recommend a "sufficient number" of simulations or running until "convergence," a precise definition of convergence is often lacking [67].

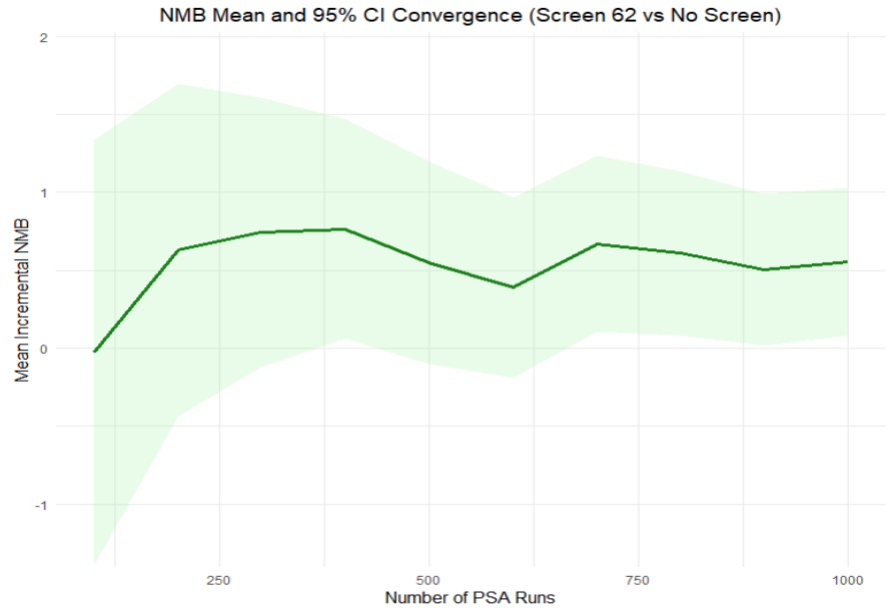
Given our need to model large populations and multiple subgroups, and to account for the stochastic nature of the model, we determined the minimum number of required probabilistic runs through visual inspection of convergence plots. These plots showed the point at which the results stabilised for general population screening in two age groups (60 and 62 years). Based on this visual assessment, 750 probabilistic runs were deemed appropriate for our analysis.

Figure 13: Convergence of the probabilistic runs for one-time screening of 60-year-olds compared to standard of care



Legend: NMB – net monetary benefit; PSA runs – probabilistic sensitivity analyses run.

Figure 14: Convergence of the probabilistic runs for one-time screening of 62-year-olds compared to standard of care



Legend: NMB – net monetary benefit; PSA runs – probabilistic sensitivity analyses run.

7.3.3 Approach to running the deterministic analysis and probabilistic sensitivity analysis

To evaluate risk-stratified screening in the general population as well as in subgroups with low representation (e.g. individuals of Black ethnicity comprising approximately 4% of the total population, and BRCA mutation carriers around 1 in 200), the model required a substantial level of complexity. This led to the following key limitations in model execution:

1. The model is stochastic, requiring simulation of a large number of individuals to adequately capture stochastic uncertainty. This demands significant computational capacity.
2. To ensure sufficient representation of high-risk subgroups, each subgroup needed to be simulated separately.
3. There are issues on applying discounting at the time of each intervention since this require each intervention to be evaluated separately and creates multiple "standard of care" comparators, each discounted at different times.

The consequences of these limitations are the following:

- Significant time was required to set up and run each analysis to properly address model uncertainty.
- A full incremental analysis could not be conducted, since each intervention was compared only to "standard of care" rather than to one another.

To address these limitations while meeting the project's requirement to simulate multiple analyses across different population groups, the following approach was adopted:

1. In the base case, using a cohort of 5 million individuals, each intervention was evaluated separately against a no-screening scenario, with discounting applied (a) from the cycle of the youngest age across all interventions, and (b) from the cycle in which each intervention began. This allowed to estimate an impact of the selected approach to discounting.

2. High-risk subgroups (men of Black ethnicity, BRCA1/2 carriers, and those with a family history) were simulated also using a cohort of 5 million individuals to ensure sufficient representation. For this, the men with relevant characteristics were sample with replacement from the original HSE population.
3. Scenario analyses and probabilistic sensitivity analyses were conducted using a cohort size of 1 million individuals to balance computational efficiency and statistical robustness. For probabilistic analysis all parameters were varied simultaneously unless they were set up as constants. The constant parameters included discounting and uptake rate. No structural uncertainty was included into the probabilistic analysis. The calibrated parameters were used in probabilistic analysis as correlated parameter sets.

7.3.4. Description of screening scenarios

All organised screening scenarios were compared against the standard of care. The standard of care was selected as the sole comparator because reconstructing a true "no screening" scenario was not feasible within the project timeframe and given the available data.

In this context, cancer cases detected in the standard care arm represent a mix of cases identified through symptoms, PSA testing performed during investigations for other conditions (i.e. incidental findings), and opportunistic screening. Modelling the comparator arm as a combined group of cancers detected under the standard of care is not expected to substantially affect the cost-effectiveness outcomes. This is because, in the model, average costs, utilities, and survival are assigned to each detected cancer case instead of pathway-specific follow-up strategies; consequently, the overall impact reflects the mean outcomes across these diagnostic pathways.

The decision to include only one comparator—the standard of care—was also driven by practical considerations. Modelling a "no screening" scenario would have required calibration to pre-PSA era data, which are outdated (see the calibration section). Moreover, the absence of reliable information on

opportunistic screening uptake makes it difficult to accurately construct either a "no screening" or a "current care" scenario. The standard of care was therefore deemed the most informative comparator, as it best represents the current real-world context.

To explore the potential net benefit of screening - that is, whether screening could be cost-effective under ideal conditions - we conducted both deterministic and probabilistic analyses assuming a perfect uptake rate. In this scenario, we assumed 100% participation across the entire screening and diagnostic pathway, not just full acceptance of the initial screening invitation.

It is important to note that the results from this "perfect uptake" scenario should not be interpreted as evidence that screening is cost-effective. Rather, they indicate that cost-effectiveness could be achieved in theory, under ideal but unrealistic conditions. In practice, such universal participation is unattainable; therefore, final conclusions about cost-effectiveness should be drawn from analyses using realistic uptake rates.

The description of the model set ups are described in the Supplementary sections E-G.

Single screening scenarios

To explore the potential for cost-effective screening, we first modelled single screening at different ages for each population subgroup: the general population, men of Black ethnicity, men with familial risk, and BRCA carriers. For the general population, single screening was evaluated at ages 50, 55, 58, 60, 62, 65, 68, 70, and 72 years. For men with familial risk, men of Black ethnicity, and BRCA carriers, additional single screening scenarios at ages 45 and 48 years were also assessed.

The simulated model outcomes are driven by underlying functions that capture trends and correlations. Because many model parameters - such as the probability of cancer onset and progression, PSA values, cancer survival, and other-cause mortality -are correlated with age, the resulting cost-effectiveness estimates also follow an age-related trend. As is typical in cancer screening models, cost-effectiveness initially improves with increasing screening age, since the prevalence of undiagnosed disease rises with age, leading to more true positives

and fewer false negatives. However, beyond a certain age, cost-effectiveness begins to decline. This decline occurs because competing mortality risks increase and the number of overdiagnosed cancers grows accordingly. Therefore, evaluating single screening strategies across different ages helps identify the turning points at which screening becomes, or ceases to be, cost-effective.

Repeat screening scenarios

Because screening is characterised by a combination of multiple characteristics, numerous repeat screening strategies could, in theory, be developed and evaluated in health economic analyses. However, because the mathematical screening model operates at the individual level, simulates the natural history of disease, and requires large population sizes to estimate clinical and economic outcomes, each screening scenario involves substantial simulation time. To ensure computational feasibility and focus on the most relevant options, the number of evaluated scenarios was limited to those with the highest likelihood of being cost-effective.

The selection of repeat screening scenarios was guided by the results of the single-screening analyses. Since the NMB from single screening follows a clear trend, we restricted the youngest and oldest ages for repeat screening invitations to those that produced the highest NMB in the single-screening results. This approach aligns with established modelling practices in other cancer screening models as repeat screening generally leads to higher costs and smaller incremental benefits compared with a one-off screening (e.g., the CAP prostate cancer model [42]).

Scenario analyses

To evaluate parameter and structural uncertainty, we conducted several scenario and sensitivity analyses:

1. Discount rate of **5%** applied to effects and **3.5%** to costs. This scenario explores methodological uncertainty.

2. Discount rate of **5%** applied to both effects and costs. This scenario explores methodological uncertainty.

3. Patients cannot die from cancer before reaching their symptomatic age (lead time scenario). This scenario explores structural uncertainty; feasibility of this scenario depends on how likely early diagnosis may result in earlier deaths due to treatment side effects.

4. Mortality extrapolated up to 70-year timeframe (instead of 15 years in the baseline). This scenario explores parameter and structural uncertainty, specifically in survival.

5. Sensitivity defined as a single threshold, assuming PSA values do not change with age. This scenario explores parameter and structural uncertainty, specifically in test sensitivity assumptions.

6. Sensitivity adjusted by age only for men without cancer. This scenario explores parameter and structural uncertainty, specifically in test sensitivity assumptions.

7. Assumes a lower health-related utility in the first-year post-diagnosis than in subsequent years. This scenario explores parameter uncertainty, specifically in utility assumptions.

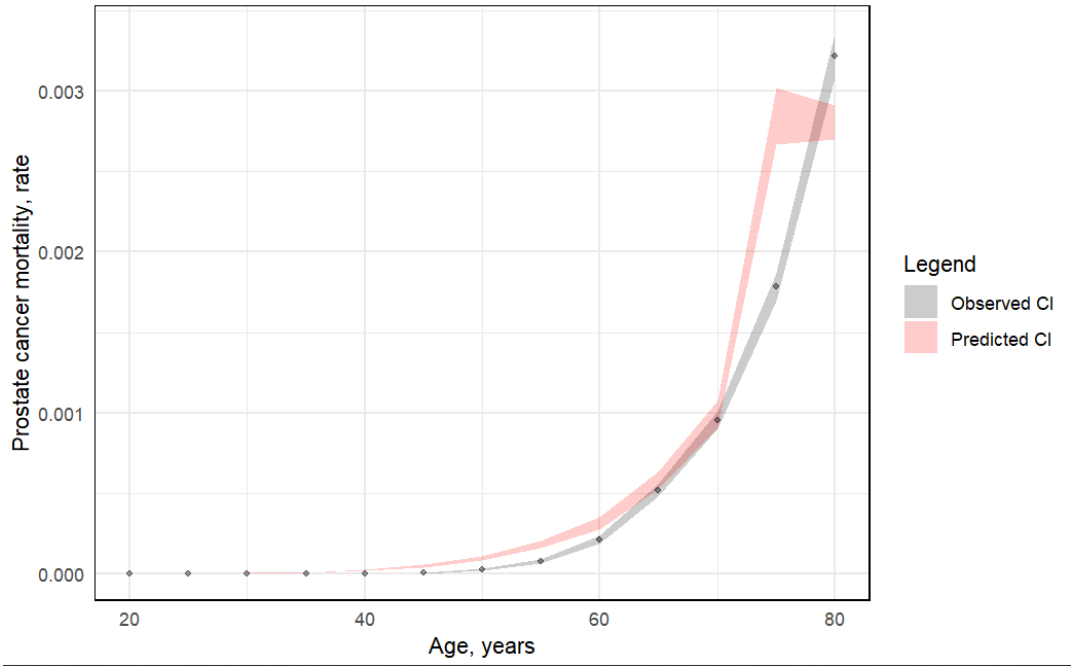
8. Scenario with improved fit to observed mortality by adjusting survival data. This was achieved by applying a multiplier of 0.8 to survival rates between ages 20–70 and 0.5 for all other ages across cancer stages (see Figure 15 for the resulting fit). This scenario is hypothetical and not data driven. It explores impact of better fit to data (mortality) in the model.

9. Discounting as per suggestions of the Green Book for both costs and effects and the decision threshold of £15,000 per QALY. This scenario explores methodological uncertainty (using different reference case to the NICE reference case).

10. Assuming a higher impact of harms related to cancer diagnosis by lowering the EQ-5D multipliers by 20% for each cancer stage. This scenario is hypothetical and not data driven. It explores impact of higher disease-associated harms in the model.

11. Exploring survival uncertainty by replacing extrapolated mortality in stages 3 and 4 for population aged 75 years old and older using mortality values in the year five. This scenario explores parameter and structural uncertainty, specifically in survival.

Figure 15: Fit to mortality after adjusting the survival in a scenario analysis 8



8. Modelling results: PSA screening with threshold 3ng/ml

8.1. General population

8.1.1. General population: single screening at different ages

As per the study protocol, the initial model runs were conducted for the general-risk population with a single screening round to identify the most cost-effective screening ages. In these scenarios, the standard care option (i.e. no organised screening) applied discounting from the age of the first screening intervention. However, because the timing of discounting may affect the NMB, each scenario was re-evaluated by applying discounting to both the intervention and the comparator (the standard of care) only starting from the intervention cycle, in scenarios where the NMB could be affected. In these cases, outcomes (LYs and QALYs), costs, and NMB were recalculated accordingly. Both results of the evaluation for the NMB are reported.

Screening led to an overall increase in lifetime incidence compared to standard care, with the magnitude of this increase being greater at older screening ages (Figure 16). However, screening also reduced prostate cancer mortality across all scenarios, with the largest reduction observed in the cohort screened at around age 65 (Figure 17).

Figure 16: Prediction of incidence with one-time screening in general population, per person

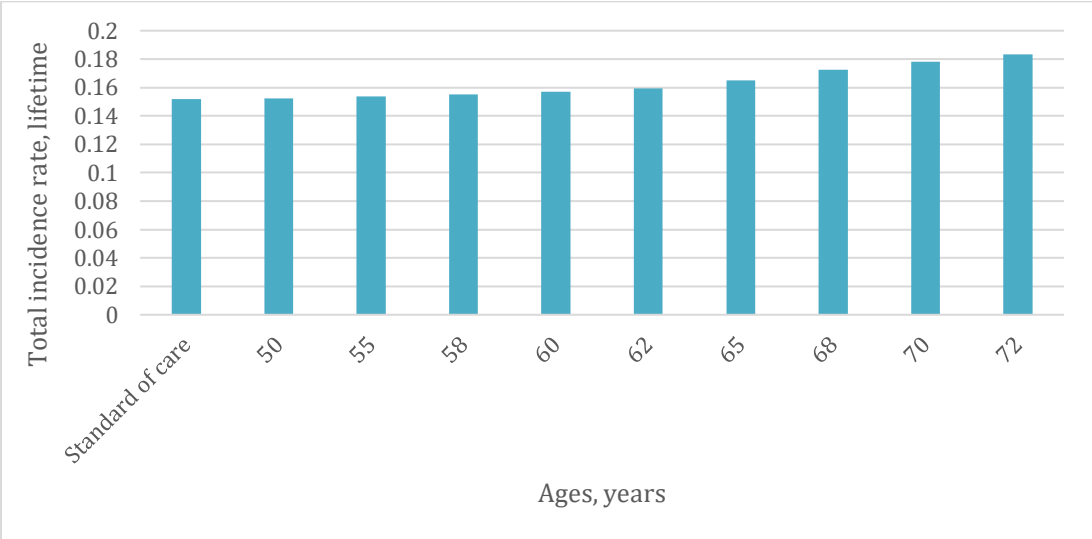
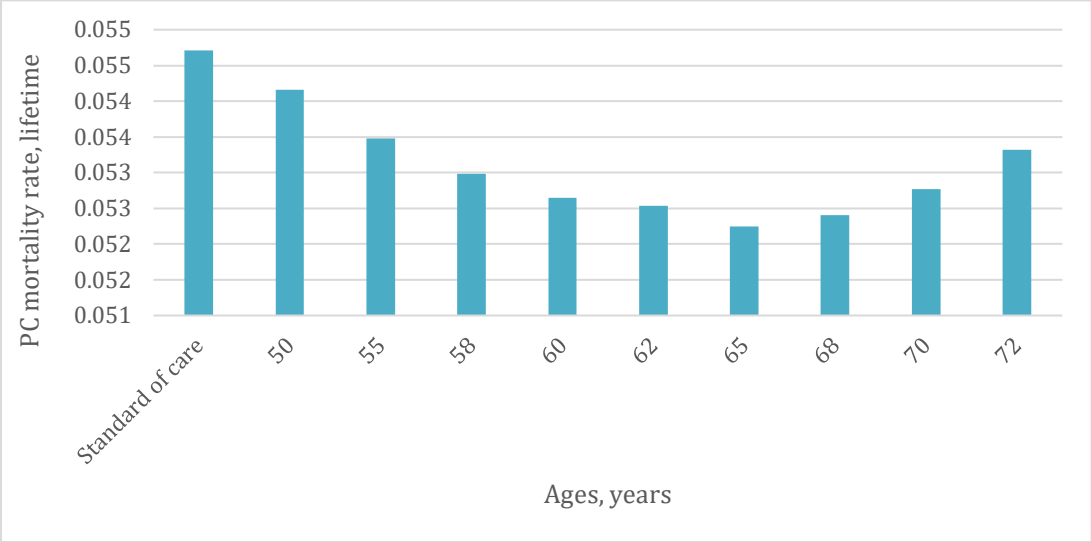


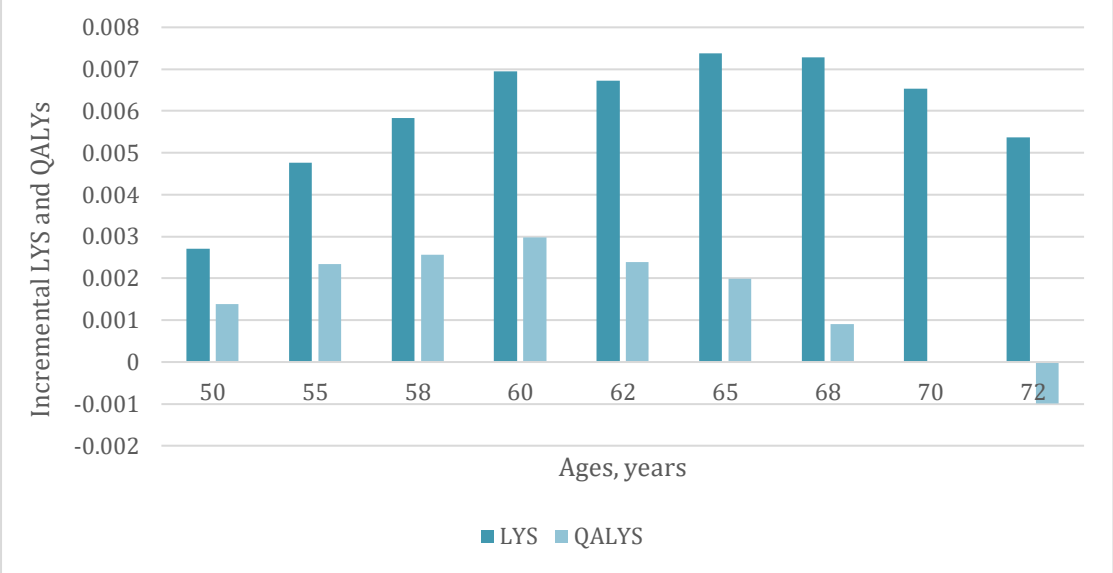
Figure 17: Prediction of lifetime prostate cancer mortality rate with one-time screening in general population



Legend: PC – prostate cancer

Screening led to gains in LYS across all age groups and increases in QALYs for those screened before age 70 (Figure 18).

Figure 18: Incremental LYS and QALYs per person with one-time screening in general population vs standard of care

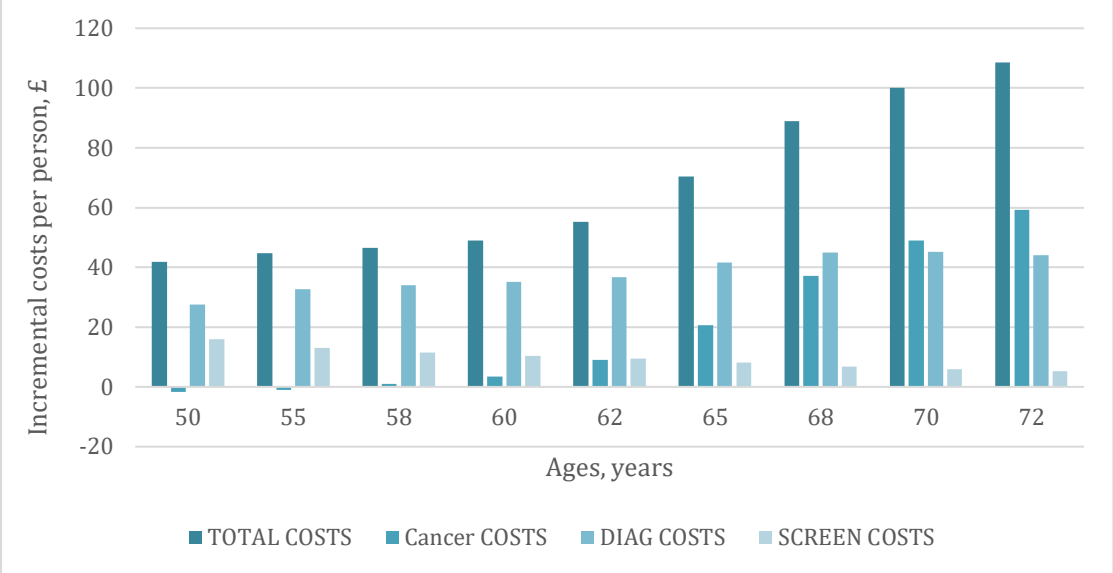


Legend: LYS – life years saved; QALYS – quality adjusted life years

Screening increased total costs (Figure 19). As expected, screening-related costs decreased with age due to fewer individuals being invited at older ages. In contrast, diagnostic follow-up costs for screen-positive cases and incremental cancer treatment costs generally increased with age, reflecting a higher

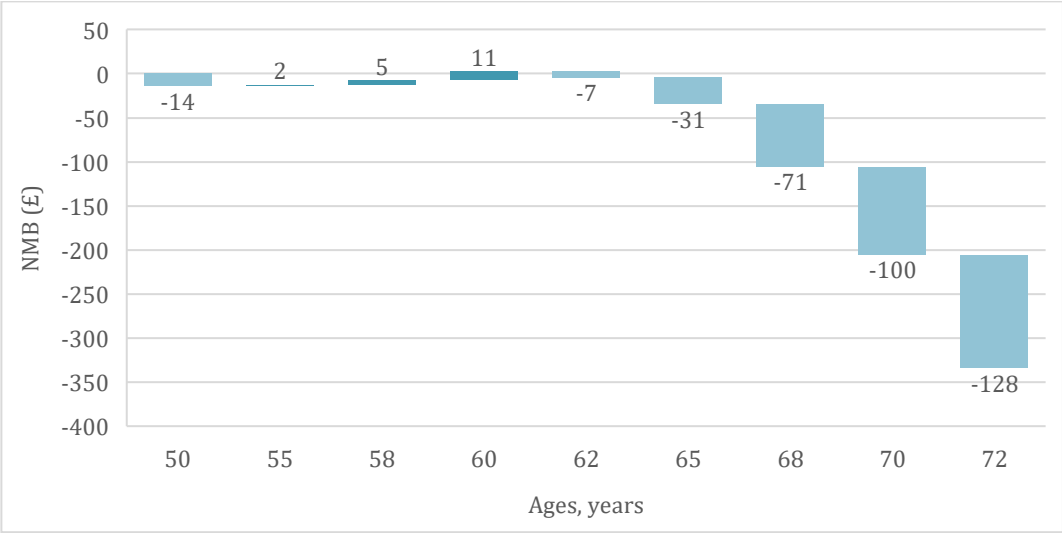
prevalence of undiagnosed cancers in older populations and increase in PSA levels by age.

Figure 19: Incremental costs per person with one-time screening in general population vs standard of care



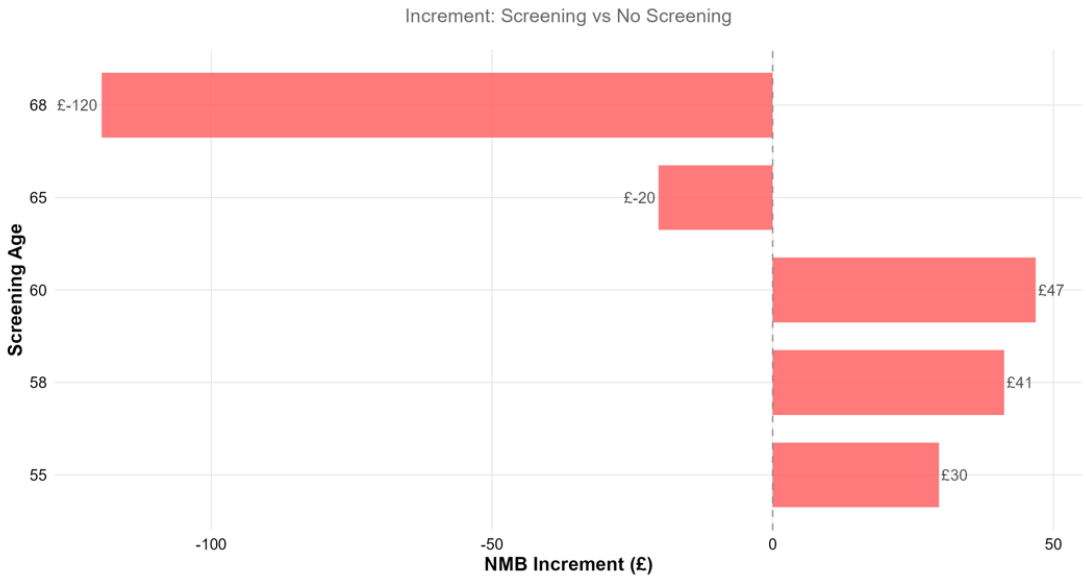
The incremental NMB of screening was slightly positive in the 55–60 age group, and the highest among those screened at ages 58 and 60 (Figure 20). The results did not change when discounting was applied at the time of the intervention for each intervention evaluated separately (Figure 21).

Figure 20: Incremental net monetary benefit (NMB) with one-time screening in general population vs standard of care



Legend: NMB - net monetary benefit

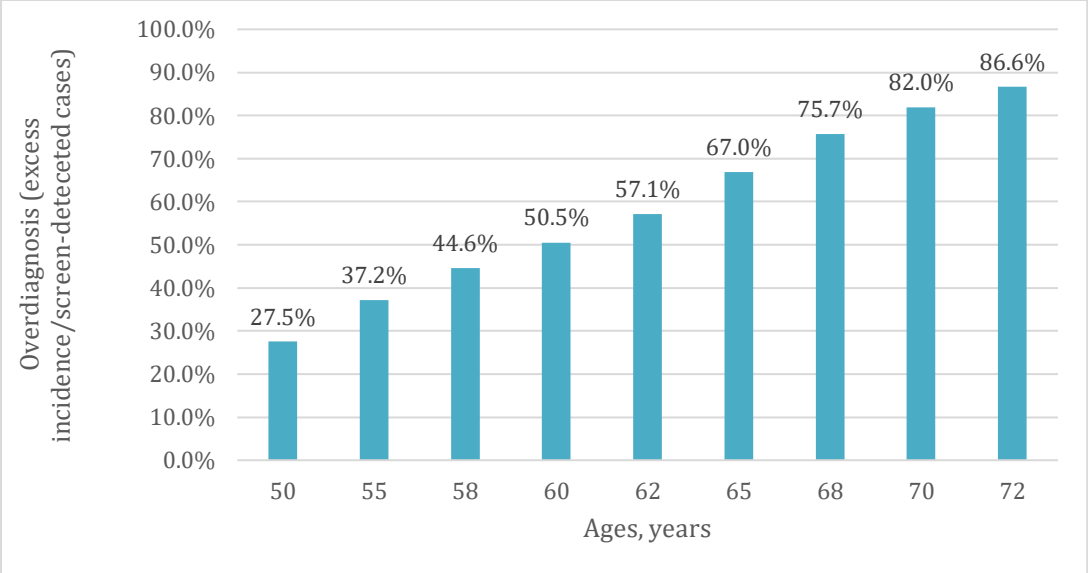
Figure 21: Incremental net monetary benefit with one-time screening in general population vs standard of care; scenario with adjusted discounting



Legend: NMB - net monetary benefit

However, screening also resulted in substantial overdiagnosis, with rates increasing at older ages due to competing mortality risks (Figure 22).

Figure 22: Overdiagnosis with one-time screening in general population (increased incidence divided by the number of screen-detected cases)

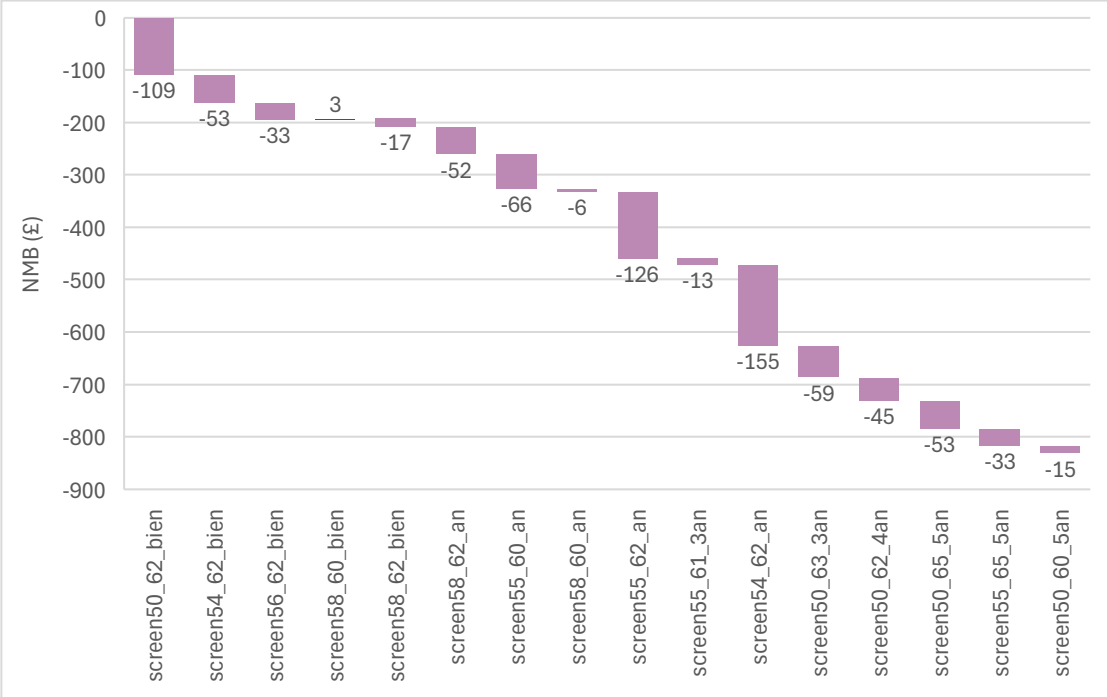


The results from the single screening analyses informed the design of the repeat screening strategies evaluated in the model.

8.1.2. General population: repeat screening scenarios

Repeat screening resulted in a negative NMB across all evaluated scenarios, except for biennial screening at ages 58 and 60 (i.e. two screening rounds per person), which yielded a small positive NMB (Figure 23). When this scenario was re-run with the discounting starting from the fist cycle of the intervention for both intervention and comparator arms, the NMB was positive (£17). However, a substantial proportion of cases in this scenario would be overdiagnosed (Figure 24); for biennial screening at ages 58 and 60, the overdiagnosis rate was estimated at 49%.

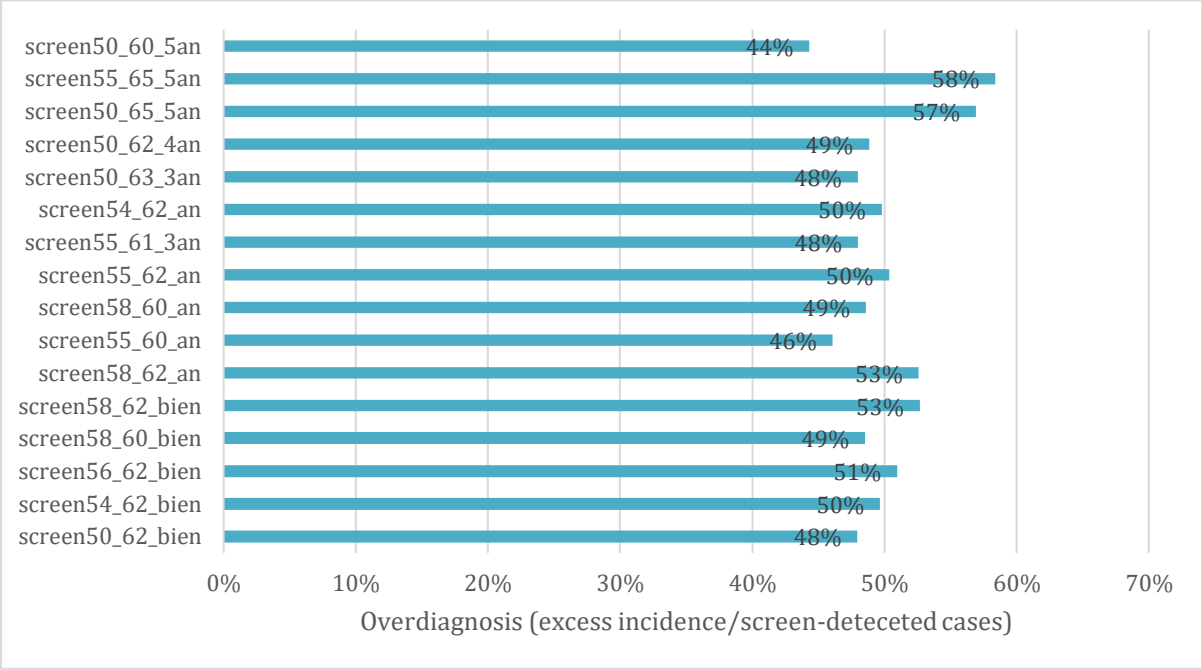
Figure 23: Incremental net monetary benefit (NMB) with repeat screening in general population vs standard of care



Legend: NMB - net monetary benefit

Annual screening at ages 58–62 (screen_58_62_an); Annual screening at ages 55–60 (screen_55_60_an); Annual screening at ages 58–60 (screen_58_60_an); Annual screening at ages 55–62 (screen_55_62_an); Annual screening at ages 54–62 (screen_54_62_an); Biennial screening at ages 50–62 (screen_50_62_bien); Biennial screening at ages 54–62 (screen_54_62_bien); Biennial screening at ages 58–60 (screen_58_60_bien); Biennial screening at ages 58–62 (screen_58_62_bien); Triennial screening at ages 55–61 (screen_55_61_3an); Triennial screening at ages 50–63 (screen_50_63_3an); Triennial screening at ages 50–59 (screen_50_59_3an); Screening every 4 years at ages 50–62 (screen_50_62_4an); Screening every 5 years at ages 50–65 (screen_50_65_5an); Screening every 5 years at ages 55–65 (screen_55_65_5an); Screening every 5 years at ages 50–60 (screen_50_60_5an).

Figure 24: Overdiagnosis rate with repeat screening in general population (increased incidence divided by the number of screen-detected cases)



Legend: Annual screening at ages 58–62 (screen_58_62_an); Annual screening at ages 55–60 (screen_55_60_an); Annual screening at ages 58–60 (screen_58_60_an); Annual screening at ages 55–62 (screen_55_62_an); Annual screening at ages 54–62 (screen_54_62_an); Biennial screening at ages 50–62 (screen_50_62_bien); Biennial screening at ages 54–62 (screen_54_62_bien); Biennial screening at ages 58–60 (screen_58_60_bien); Biennial screening at ages 58–62 (screen_58_62_bien); Triennial screening at ages 55–61 (screen_55_61_3an); Triennial screening at ages 50–63 (screen_50_63_3an); Triennial screening at ages 50–59 (screen_50_59_3an); Screening every 4 years at ages 50–62 (screen_50_62_4an); Screening every 5 years at ages 50–65 (screen_50_65_5an); Screening every 5 years at ages 55–65 (screen_55_65_5an); Screening every 5 years at ages 50–60 (screen_50_60_5an).

8.1.3. General population: probabilistic analysis

Probabilistic analysis was conducted for the following scenarios with positive NMB in deterministic runs: single screening at ages 58 and 60 and repeat screening at ages 58 and 60 (Figure 25) using the base case model inputs and assuming perfect screening uptake to evaluate the net impact of screening. The results indicated that uncertainty in the natural history parameters, specifically those correlated with age (see Supplementary C) was the most influential factor in determining cost-effectiveness, with the cost-effectiveness plane (Figure 26a) showing a split between three parameter spaces similar to the NHD parameters.

Single screening showed an approximately 80% probability of being cost-effective, though the credible intervals (CrI, the proportion of the probabilistic runs that included 95% of values) around the NMB were wide. For screening at age 58, the mean NMB was £87.2 with a 95% CrI of -£160 to £319; for screening

at age 60, the mean NMB was £69 with a 95% CrI of -£121 to £250. In contrast, repeat screening at ages 58 and 60 had a lower probability of cost-effectiveness (57%) and a mean NMB of £18.3 (95% CrI: -£318 to £304) with the decision thresholds of £20,000 per QALY (Figure 26b).

Figure 25: Incremental net monetary benefit (£) vs standard of care

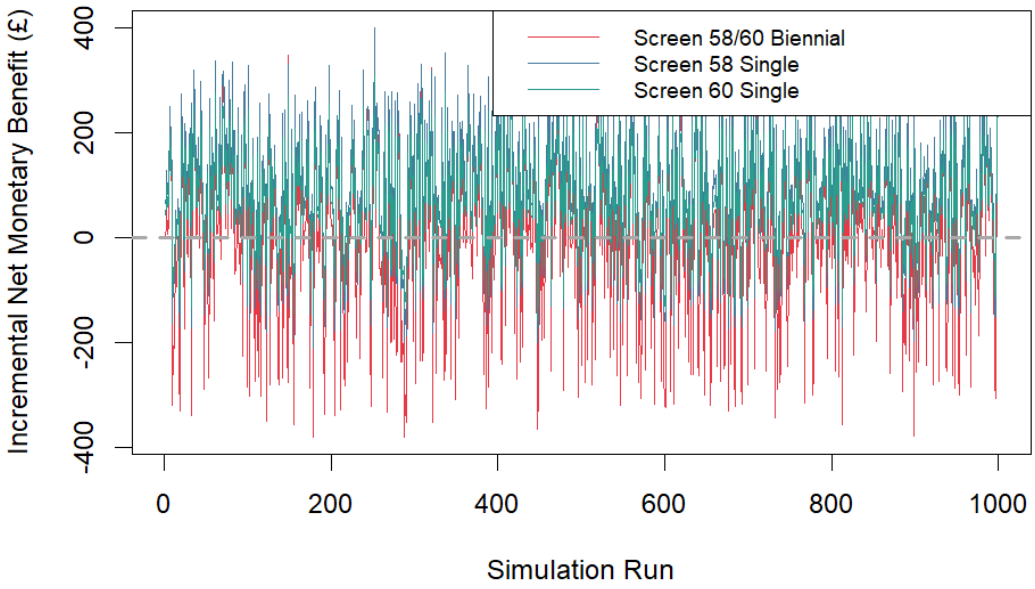
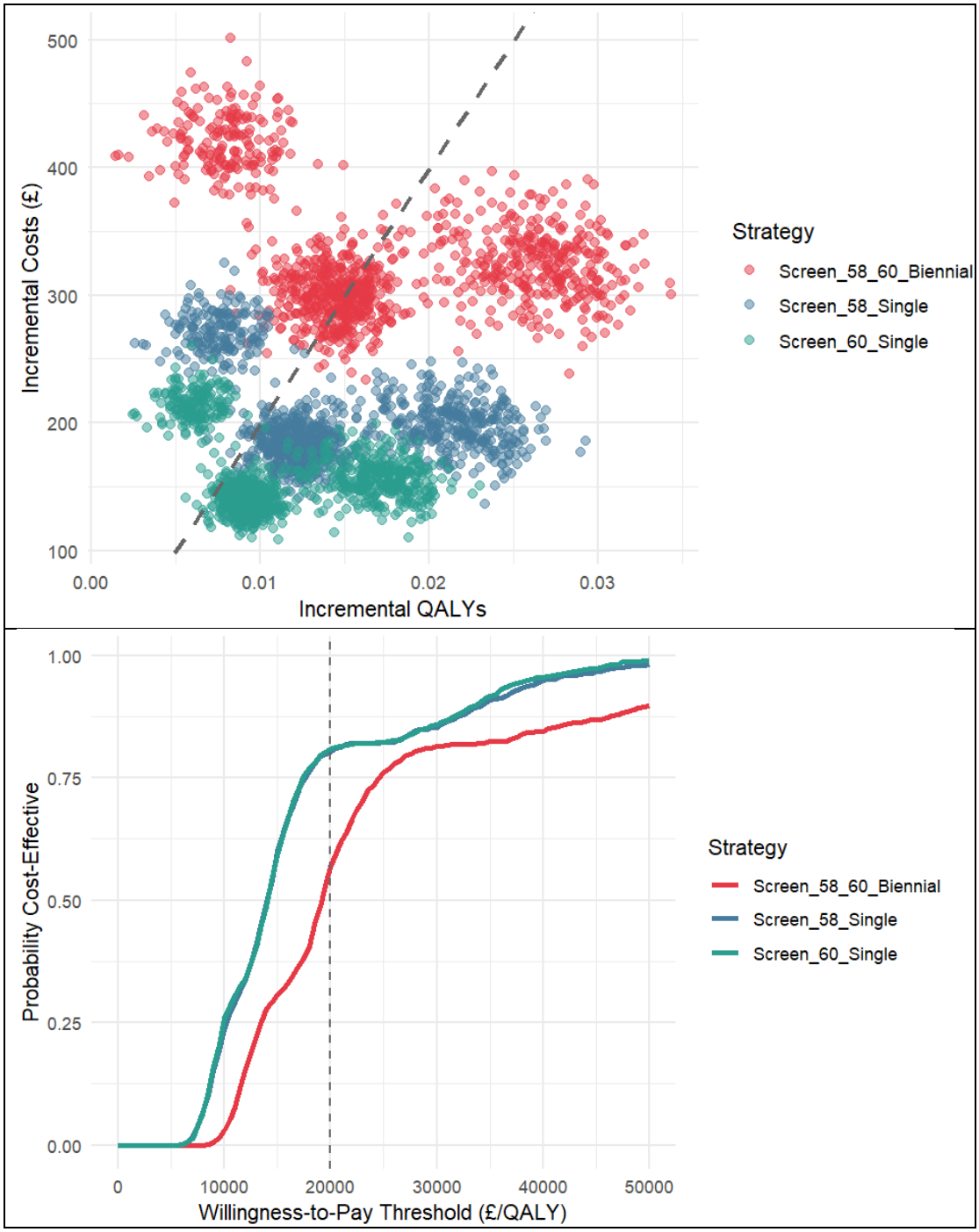


Figure 26: Cost-effectiveness plane (a) and cost-effectiveness acceptability curve (b)



Legend: QALY -quality adjusted life years

8.1.4. General population: scenario analyses

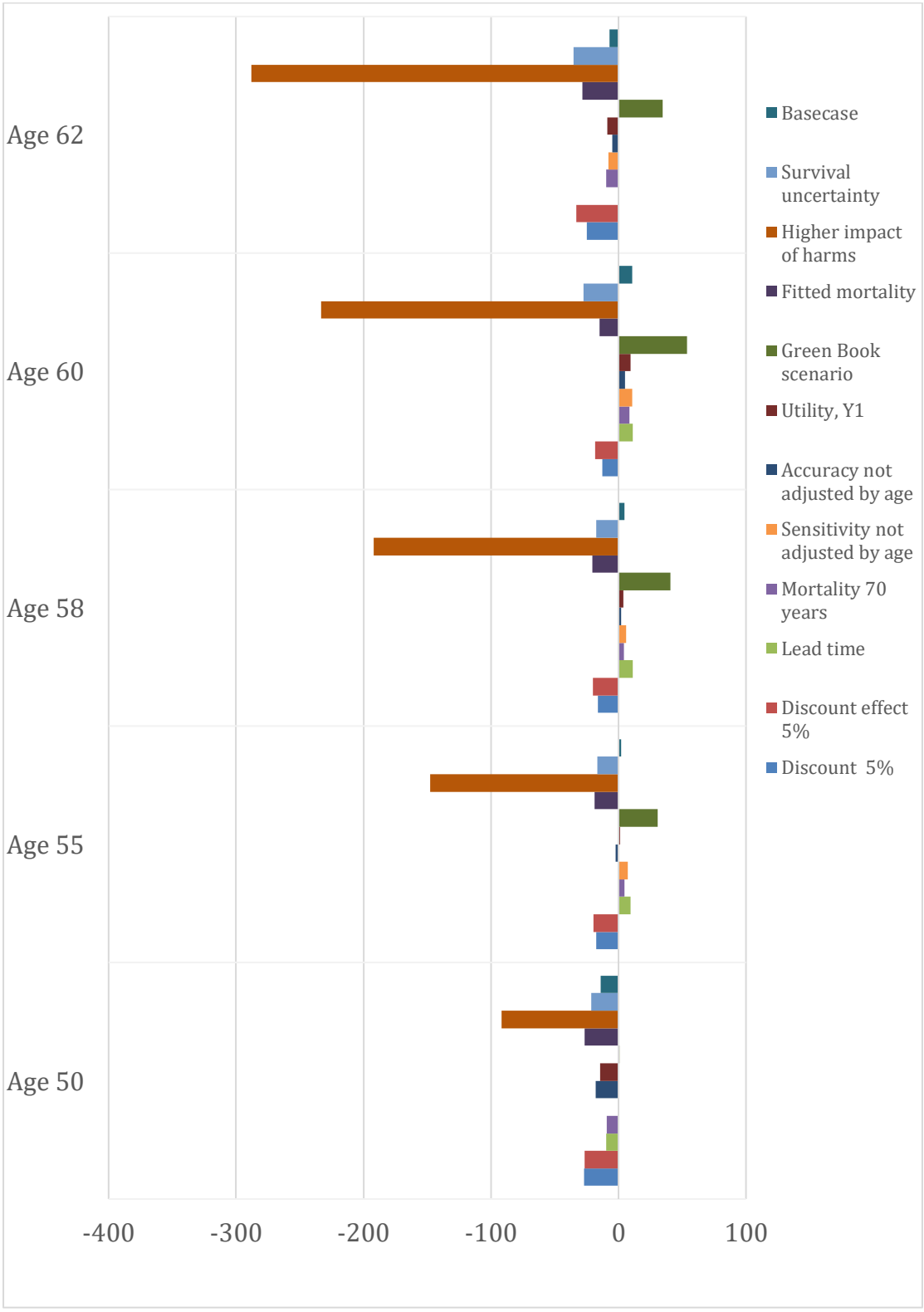
As described in previous sections of the report (section 7.3.4.), multiple deterministic scenario analyses were conducted for screening ages where the base case suggested screening might be potentially cost-effective. For single screening, these included ages 50, 55, 58, 60, and 62 years (Figure 27). For repeat screening, scenarios included biennial screening at ages 58 and 60, supplementary annual screening at ages 58–60, screening every 3 years at ages 55–61, and every 5 years at ages 50–60 (Figure 28).

Applying the Green Book scenario (i.e. dynamic by time with higher discounting for costs than effects and the £ 15,000 threshold) showed a positive NMB across all scenarios. Conversely, applying a higher discount rate of 5% for benefits had the opposite effect—screening was not cost-effective in any scenario. Assuming that cancer diagnosis resulted proportionally in higher harms (i.e. impact on quality of life) in all stages at diagnosis, resulted in highly negative NMB. Applying flat values for cancer-specific mortality in stage 3 and 4 cancers in men older than 70 years (the survival analysis with highest uncertainty) resulted in screening not be cost-effective in any age group. When the model did not overpredict mortality, screening also failed to be cost-effective across all scenarios. This suggests that achieving a better fit to the natural history of the disease will result in lower cost-effectiveness of screening.

Other model adjustments—such as applying lower utility weights in the first year after diagnosis (to capture a higher immediate impact), not adjusting test sensitivity by age, extending mortality extrapolation beyond 70 years (compared to 15 years in the base case), and assuming that early-diagnosed cancers in the intervention arm would not result in earlier deaths compared to the comparator arm—had small impact on the NMB.

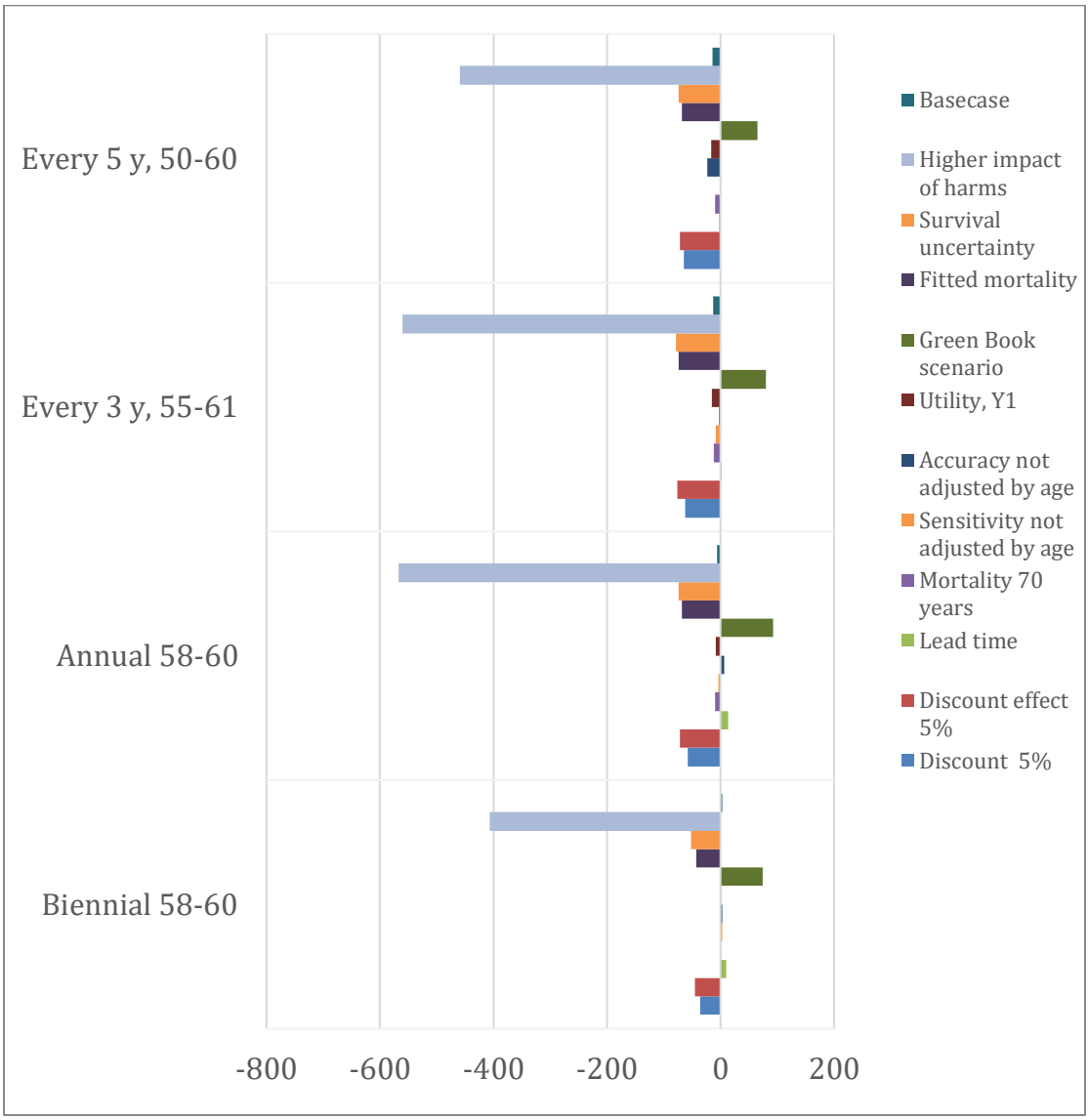
In all scenarios screening resulted in substantial level of overdiagnosis, which did not change substantially by scenarios (Supplementary H).

Figure 27: Net monetary benefit (£) in scenario analyses for one-time screening: general population



Legend: £15,000 threshold is used in the Green Book scenario and £20,000 in the other scenarios.

Figure 28: Net monetary benefit (£) in scenario analyses for repeat screening: general population



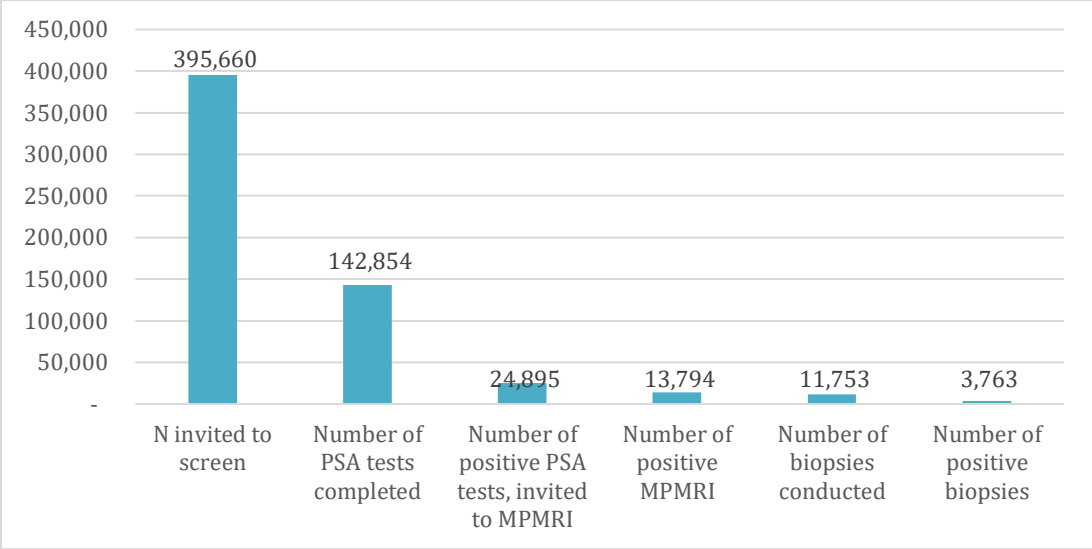
Legend: £15,000 threshold is used in the Green Book scenario and £20,000 in the other scenarios.

8.1.5. General population: impact of screening scaled to the population of England

Impact of screening on the population of England was assessed from perspective of resource requirements to implement screening and number of cases expected to be diagnosed through screening as well as expected difference in mortality between the standard of care and organised screening scenarios.

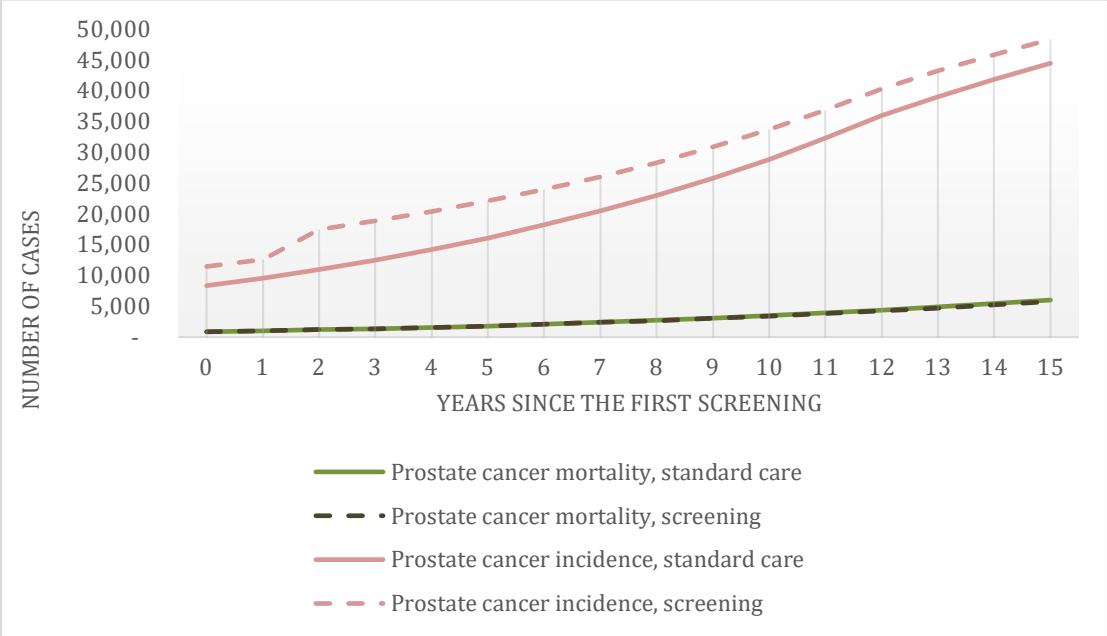
As expected, screening of men of average risk substantially increased resource use. In England, there are approximately 1.98 million men aged 55–59. Assuming 395,660 men aged 58 are invited for screening, such a programme would result annually in an estimated 142,854 completed PSA tests, 13,794 follow-up mpMRI procedures, and 11,753 biopsies (Figure 29).

Figure 29: Annual resource requirements in England for screening general population at ages 58 and 60



As stated above, it is estimated that 395,660 of men will be invited to start screening at age 58. It was estimated that over the 15 years following the implementation of an organised screening programme — started screening among all 58 y.o. men regardless of their risk (and repeating it at age 60)— an additional 4,618 prostate cancer cases are expected to be detected. The difference in prostate cancer deaths between the organised screening and standard care arms is projected to reach 153 cases in the 15 years post-implementation of screening in probabilistic analysis and 240 cases in deterministic analysis (Figure 30).

Figure 30. Difference in prostate cancer incidence and mortality between the organised screening and standard care arms among a cohort of 58-year-olds invited to screening.



8.2. Men of Black ethnicity

8.2.1. Men of Black ethnicity: single screening at different ages

Similar to the general population, the initial model runs were conducted for men of Black ethnicity with a single screening round to identify the most cost-effective screening ages.

Screening led to an overall increase in lifetime incidence compared to standard care, with the magnitude of this increase being greater at older screening ages (Figure 31). However, screening also reduced prostate cancer mortality across all scenarios, with the largest reduction observed in the cohort screened at ages 60-65 years (Figure 32).

Figure 31: Prediction of incidence with one-time screening in men of Black ethnicity, per person

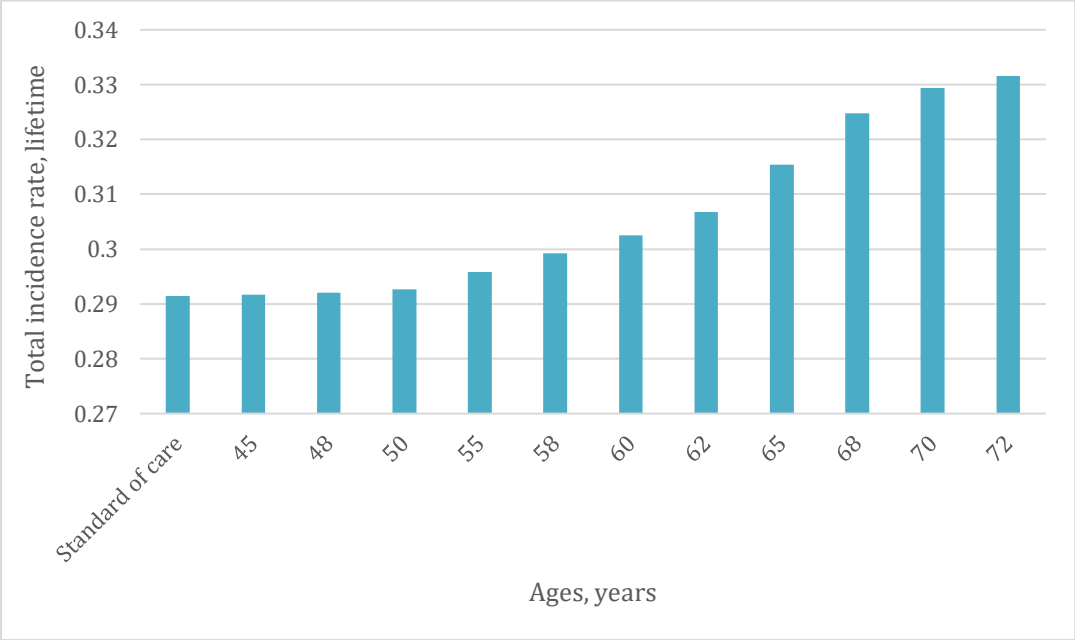
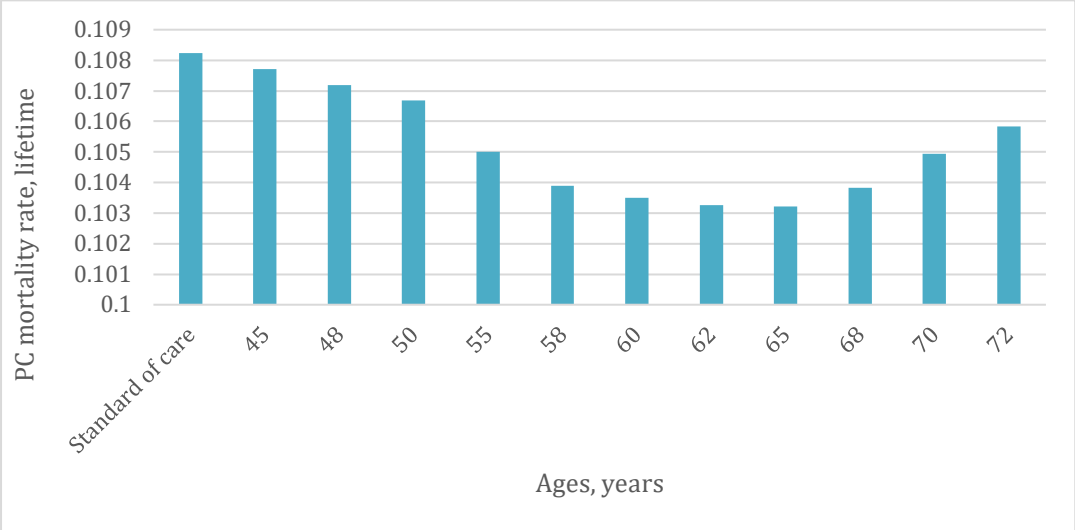


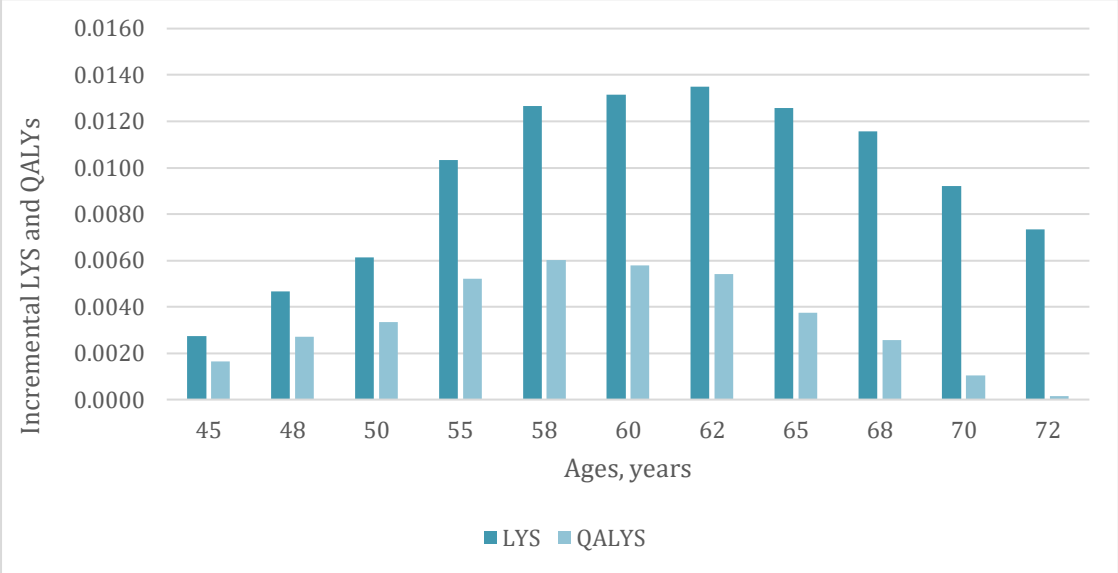
Figure 32: Prediction of lifetime prostate cancer mortality rate with one-time screening in men of Black ethnicity



Legend: PC -prostate cancer

Screening led to gains in LYS across all age groups (Figure 33) with the largest increment observed in ages 58-65 years.

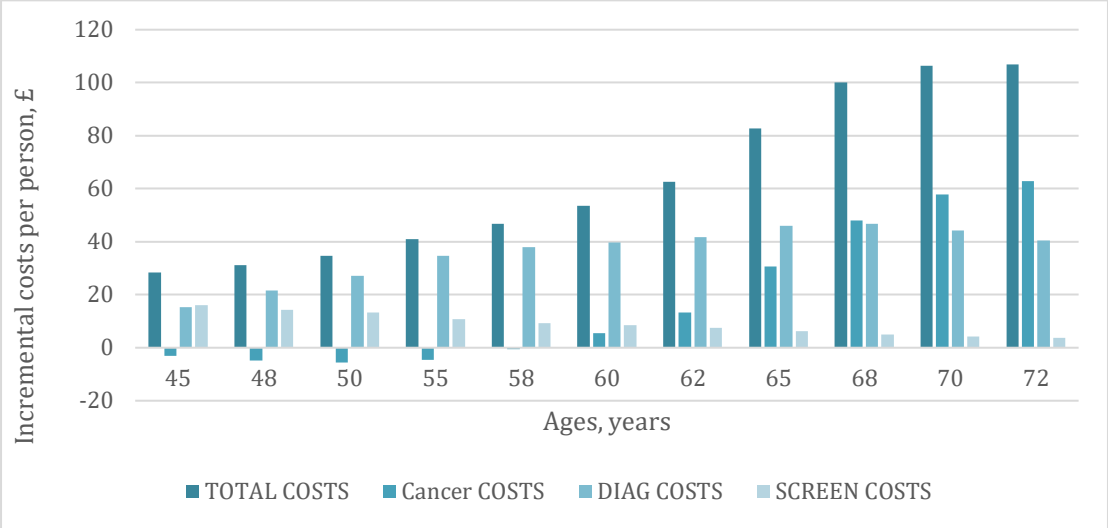
Figure 33: Incremental LYS and QALYs per person with one-time screening in men of Black ethnicity vs standard of care



Legend: LYS – life years saved; QALYS – quality adjusted life years

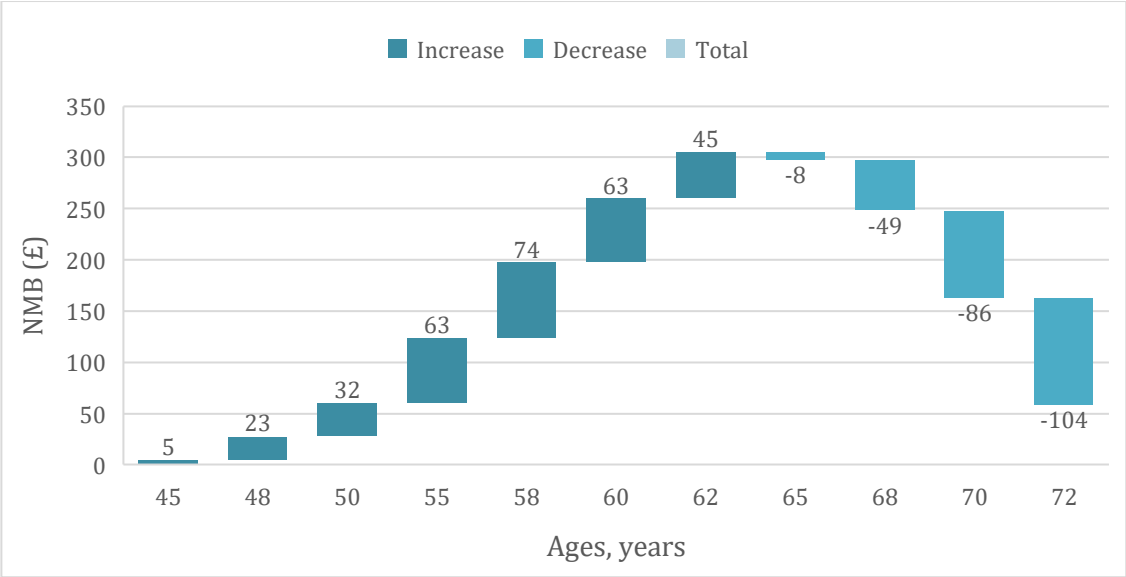
Similarly, screening led to increased total costs across all age groups. While cancer treatment costs were lower when screening occurred before age 60, the high diagnostic costs associated with follow-up of false positives and detection of indolent or slow-progressing cancers offset these savings (Figure 34).

Figure 34: Incremental costs per person with one-time screening in me of Black ethnicity vs standard of care



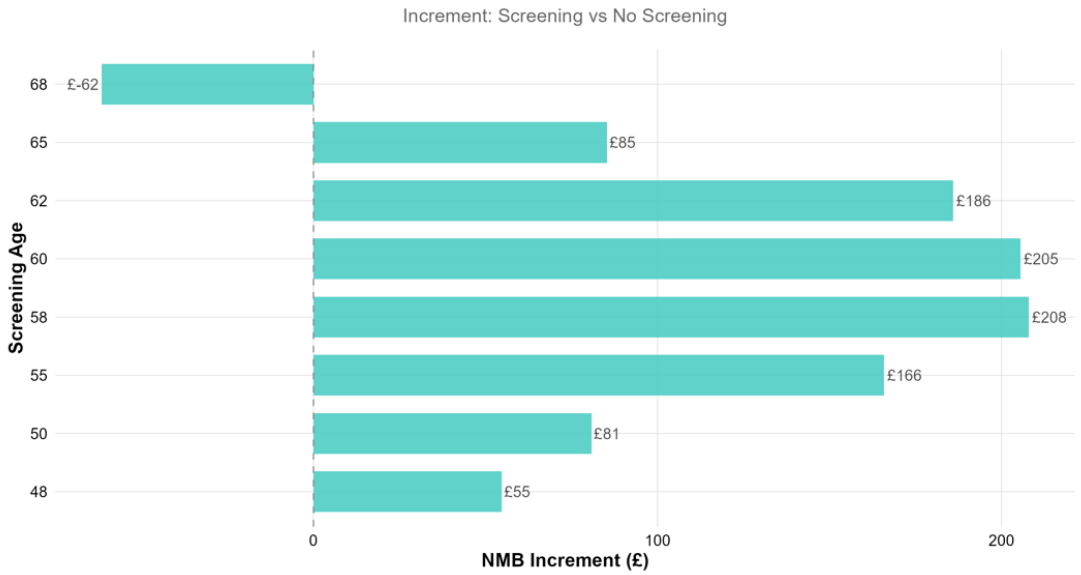
Incremental NMB was positive for screening at ages under 65, with the highest values observed in the 55–62 age group (Figure 35). These same scenarios remained cost-effective when discounting was applied from the cycle of intervention implementation for each scenario; additionally, screening men of Black ethnicity at age 65 was also found to be cost-effective under this approach (Figure 36), though the most cost-effective ages remained 55-62 and as expected the incremental NMBs at later age were higher.

Figure 35: Incremental net monetary benefit with one-time screening in men of Black ethnicity vs standard of care



Legend: NMB – net monetary benefit

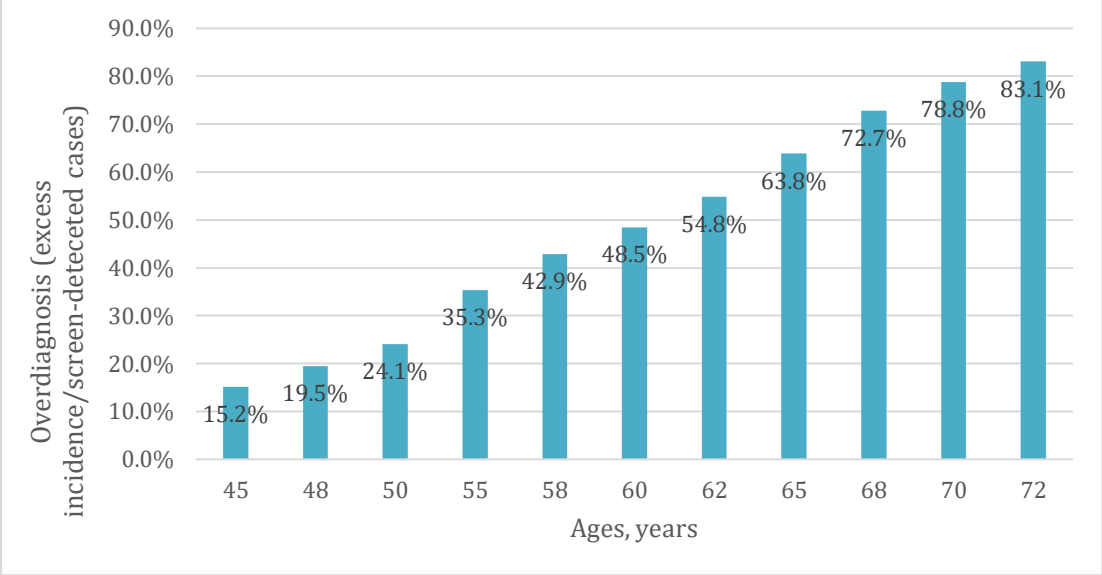
Figure 36: Incremental net monetary benefit with one-time screening in men of Black ethnicity vs standard of care; scenario with adjusted discounting



Legend: NMB – net monetary benefit

Screening resulted in overdiagnosis among men of Black ethnicity; as in the general population, the rate of overdiagnosis increased with age—from 15.2% of screen-detected cases at age 45 to 83% at age 72 (Figure 37).

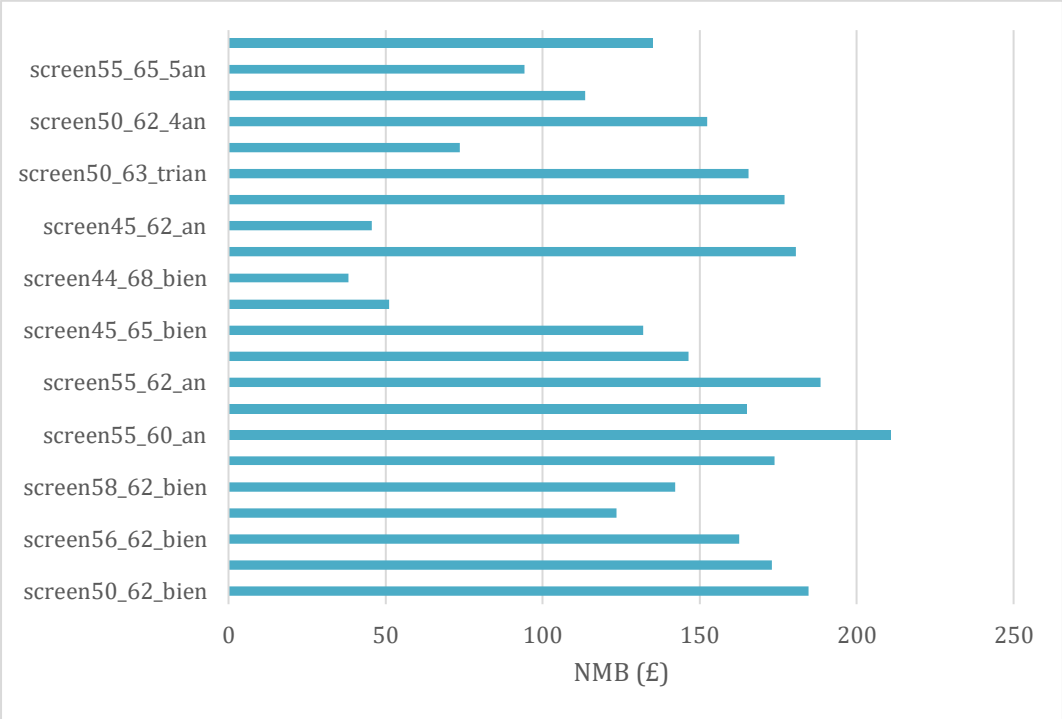
Figure 37: Overdiagnosis with one-time screening in men of Black ethnicity (increment in incidence divided by the number of screen-detected cases)



8.2.2. Men of Black ethnicity: repeat screening scenarios

Repeat screening was run for the single screening scenarios with largest NMB. The repeat screening resulted in positive incremental NMB across all evaluated scenarios (Figure 38). The largest incremental NMB was with annual screening at 55-60 years old. Screening in younger groups (<50 years) resulted in much smaller incremental NMB. Similar to the general population, a substantial proportion of cases would be overdiagnosed (Figure 39); with annual screening of 55–60-year-old men of Black ethnicity, 44% of prostate cancer cases diagnosed through screening would be overdiagnosed.

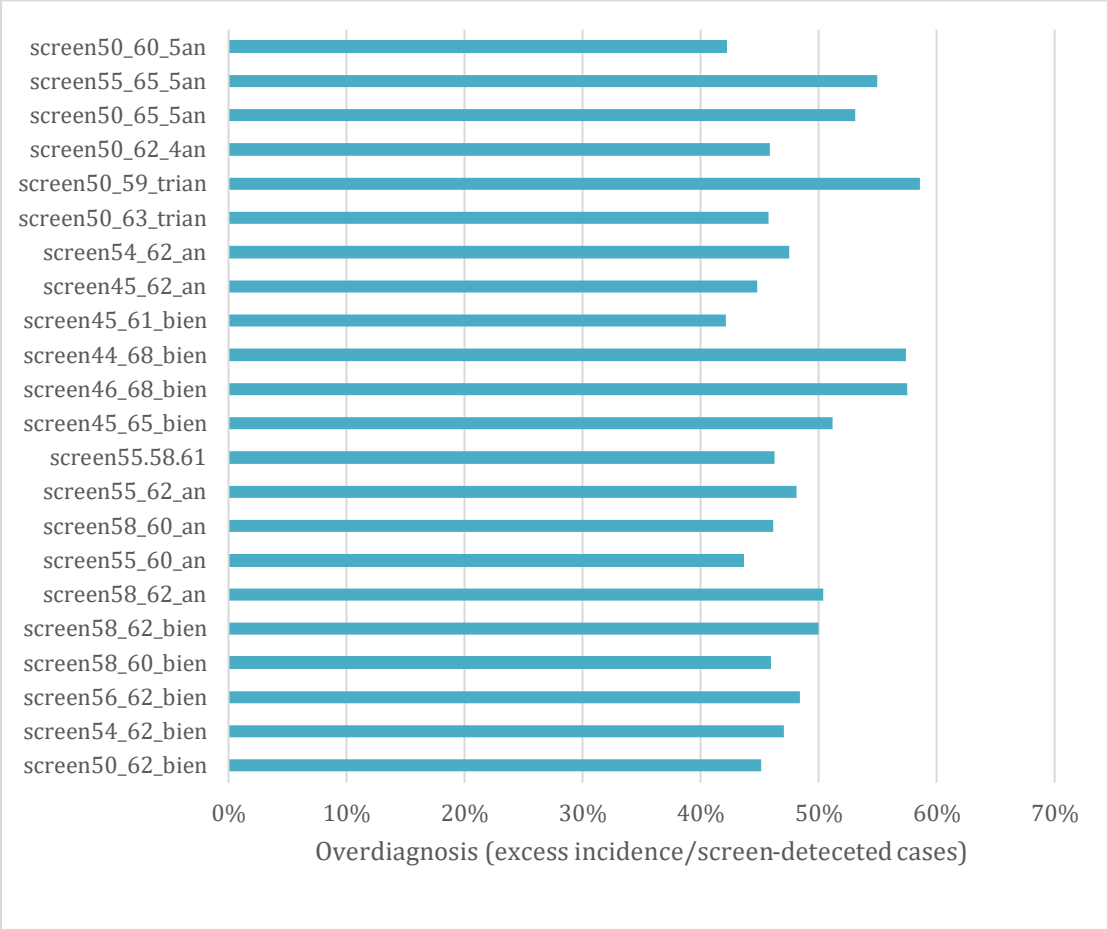
Figure 38: Incremental net monetary benefit with repeat screening of men of Black ethnicity vs standard of care



Legend: NMB - net monetary benefit

Screening every 5 years at ages 50–60 (screen_50_60_5an); Screening every 5 years at ages 55–65 (screen_55_65_5an); Screening every 5 years at ages 50–65 (screen_50_65_5an); Screening every 4 years at ages 50–62 (screen_50_62_4an); Triennial screening at ages 50–59 (screen_50_59_3an); Triennial screening at ages 50–63 (screen_50_63_trian); Annual screening at ages 54–62 (screen_54_62_an); Annual screening at ages 45–62 (screen_45_62_an); Biennial screening at ages 45–61 (screen_45_61_bien); Biennial screening at ages 44–61 (screen_45_68_bien); Biennial screening at ages 46–66 (screen_46_68_bien); Biennial screening at ages 45–65 (screen_45_65_bien); Triennial screening at ages 55–61 (screen_55.58.61); Annual screening at ages 55–62 (screen_55_62_an); Biennial screening at ages 58–60 (screen_58_60_bien); Annual screening at ages 55–60 (screen_55_60_an); Biennial screening at ages 58–62 (screen_58_62_bien); Annual screening at ages 58–60 (screen_58_60_an); Biennial screening at ages 56–62 (screen_56_62_bien); Biennial screening at ages 54–62 (screen_54_62_bien); Biennial screening at ages 50–62 (screen_50_62_bien).

Figure 39: Overdiagnosis rate with repeat screening of men of Black ethnicity (increment in incidence divided by the number of screen-detected cases)



Legend: Screening every 5 years at ages 50–60 (screen_50_60_5an); Screening every 5 years at ages 55–65 (screen_55_65_5an); Screening every 5 years at ages 50–65 (screen_50_65_5an); Screening every 4 years at ages 50–62 (screen_50_62_4an); Triennial screening at ages 50–59 (screen_50_59_3an); Triennial screening at ages 50–63 (screen_50_63_trian); Annual screening at ages 54–62 (screen_54_62_an); Annual screening at ages 45–62 (screen_45_62_an); Biennial screening at ages 45–61 (screen_45_61_bien); Biennial screening at ages 44–61 (screen_45_68_bien); Biennial screening at ages 46–66 (screen_46_68_bien); Biennial screening at ages 45–65 (screen_45_65_bien); Triennial screening at ages 55–61 (screen_55.58.61); Annual screening at ages 55–62 (screen_55_62_an); Biennial screening at ages 58–60 (screen_58_60_bien); Annual screening at ages 55–60 (screen_55_60_an); Biennial screening at ages 58–62 (screen_58_62_bien); Annual screening at ages 58–60 (screen_58_60_an); Biennial screening at ages 56–62 (screen_56_62_bien); Biennial screening at ages 54–62 (screen_54_62_bien); Biennial screening at ages 50–62 (screen_50_62_bien).

8.2.3. Men of Black ethnicity: probabilistic analyses

A probabilistic simulation of the Black male population was conducted for three scenarios (all run with perfect uptake): single screening at ages 40 and 45, and repeat screening from ages 50 to 62 at four-year intervals. The NMB was almost entirely positive for the first two scenarios and centred around zero for repeat screening (Figure 40). Similar to the men of general risk, the NMB was split into several clouds with screening in older ages, reflecting the uncertainty in calibrated NHD parameters correlated by age (see Supplementary C for this analysis).

Figure 40: Incremental net monetary benefit (£) vs standard of care (men of Black ethnicity scenarios, perfect uptake)

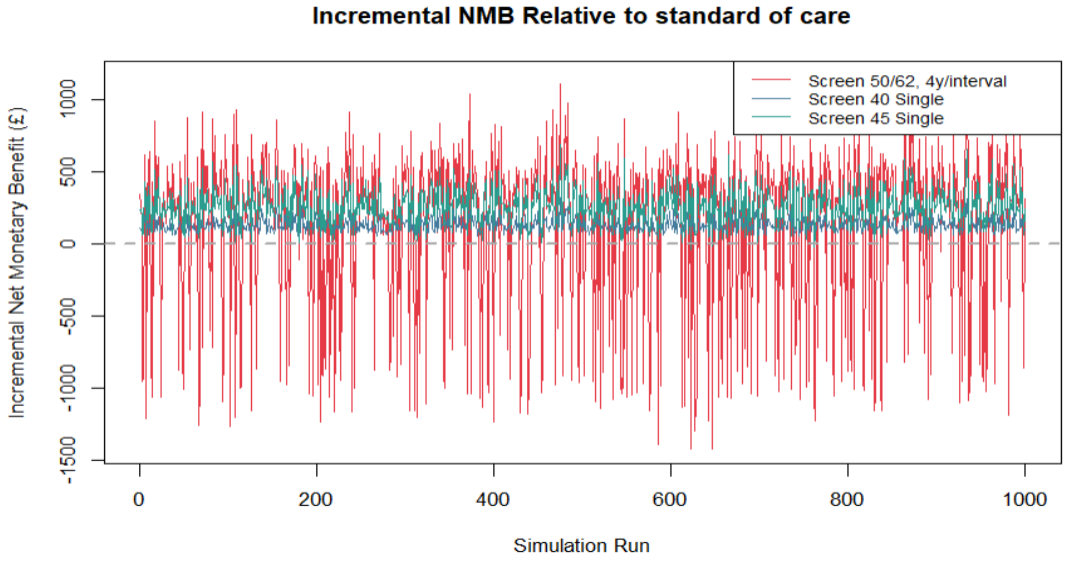
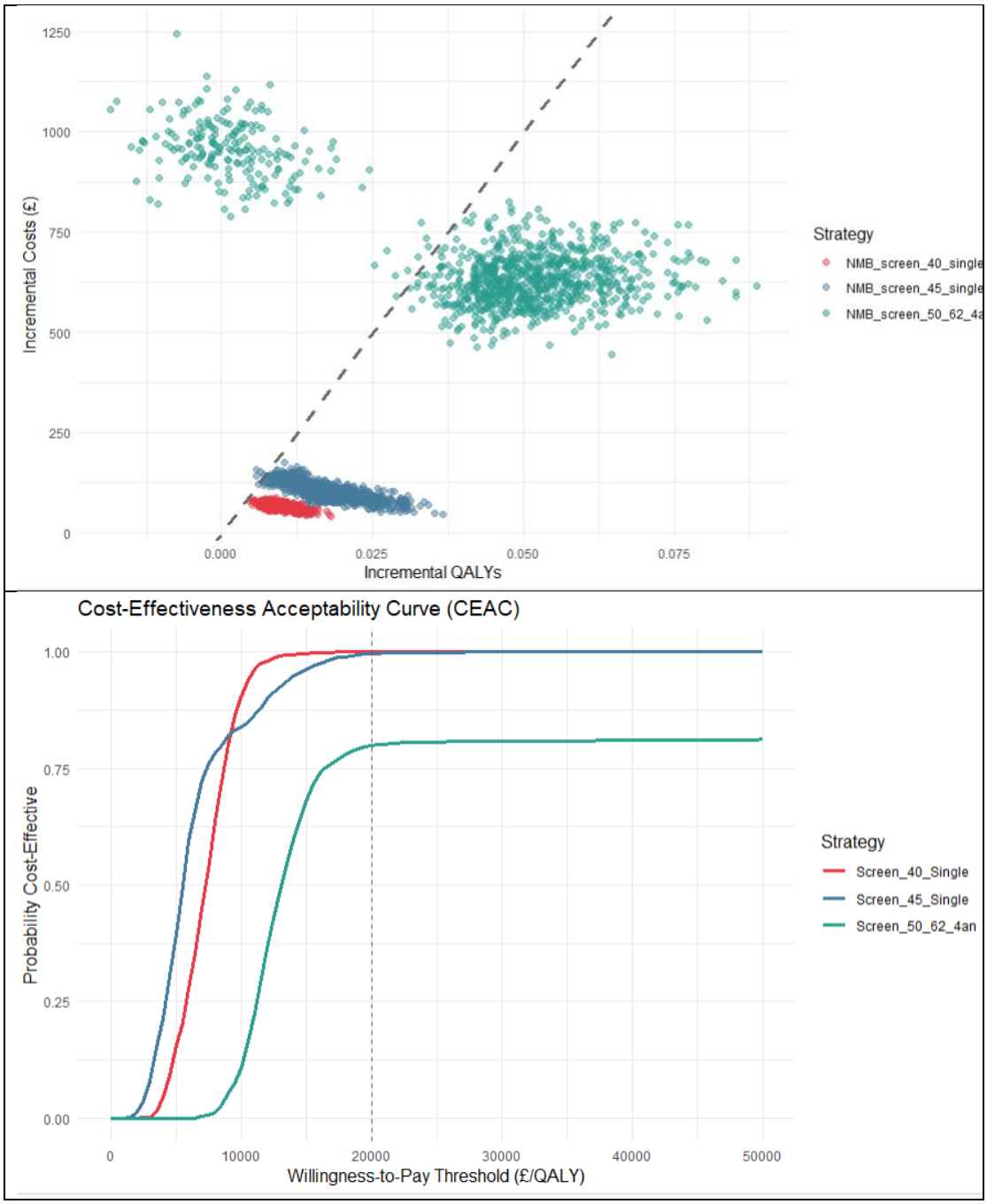


Figure 41: Cost-effectiveness plane (a) and cost-effectiveness acceptability curve (b) (men of Black ethnicity scenarios, perfect uptake)



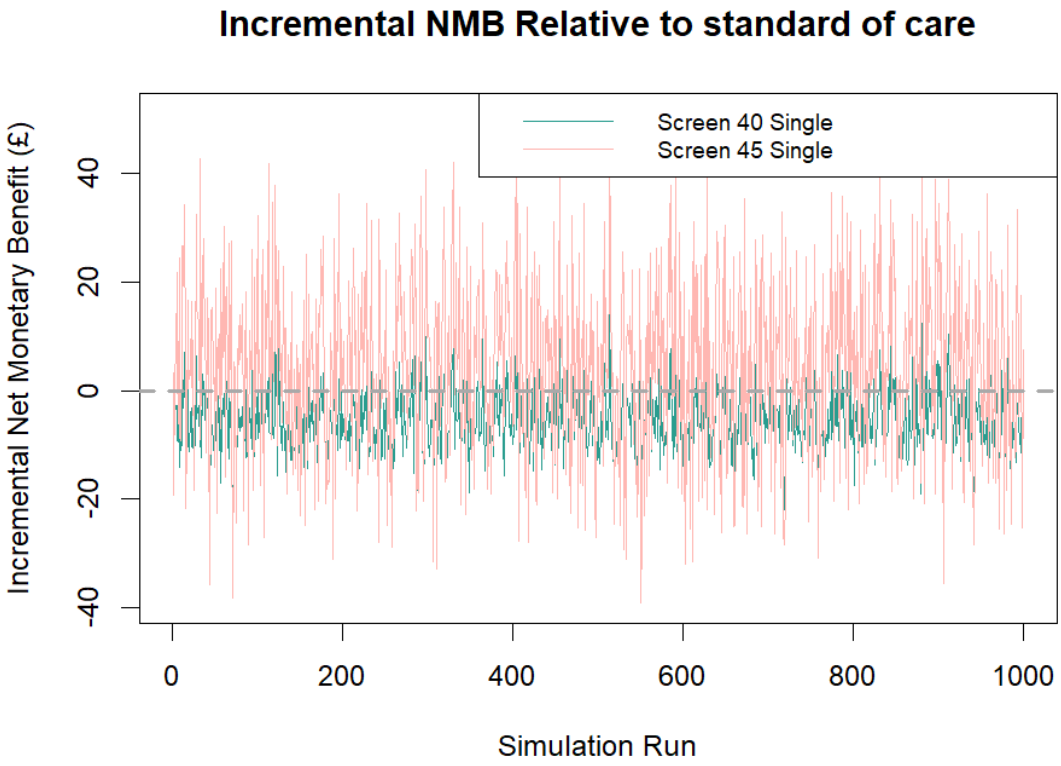
Uncertainty in the natural history parameters had a substantial influence on the cost-effectiveness of repeat screening but was less impactful for single screening (Figure 41a). Single screening demonstrated almost 100% probability of being cost-effective, with an incremental NMB of £124 (95% CrI: £55 to £235) for screening at age 40, and £259 (95% CrI: £31 to £518) for screening at age 45. The

repeat screening showed a mean incremental NMB of £148 (95% CrI: –£1,137 to £829).

Although the probabilistic sensitivity analysis indicated a potential for screening men of Black ethnicity to be cost-effective, this evaluation assumed perfect uptake—an idealised scenario of optimal screening implementation. Therefore, as with the general population, a further analysis was conducted using uptake rates observed in the CAP trial.

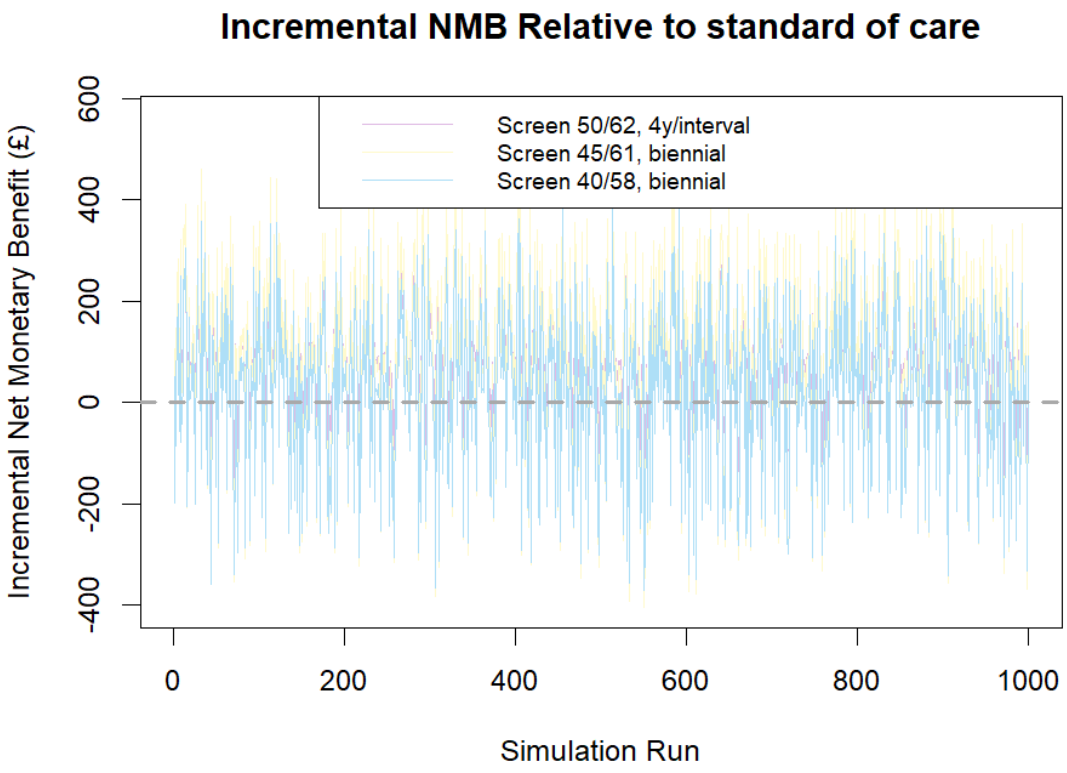
For single screening, running the model on 1 million men of Black ethnicity using uptake rates from the CAP trial showed that screening at age 40 resulted in negative incremental NMB in most simulation runs. Screening at age 45 produced results largely centred around zero, with a slight shift towards positive incremental NMB (Figure 42).

Figure 42: Incremental net monetary benefit (£) for single screening scenarios compared to the standard of care (men of Black ethnicity scenarios, CAP uptake)



For repeat screening, biennial screening from ages 40 to 58 showed high uncertainty, with incremental NMB centred around zero. In contrast, biennial screening from ages 45 to 61 and 4-yearly screening from ages 50 to 62 were both shifted towards positive incremental NMB (Figure 43).

Figure 43: Incremental net monetary benefit (£) for repeat screening scenarios compared to the standard of care (men of Black ethnicity scenarios, CAP uptake)

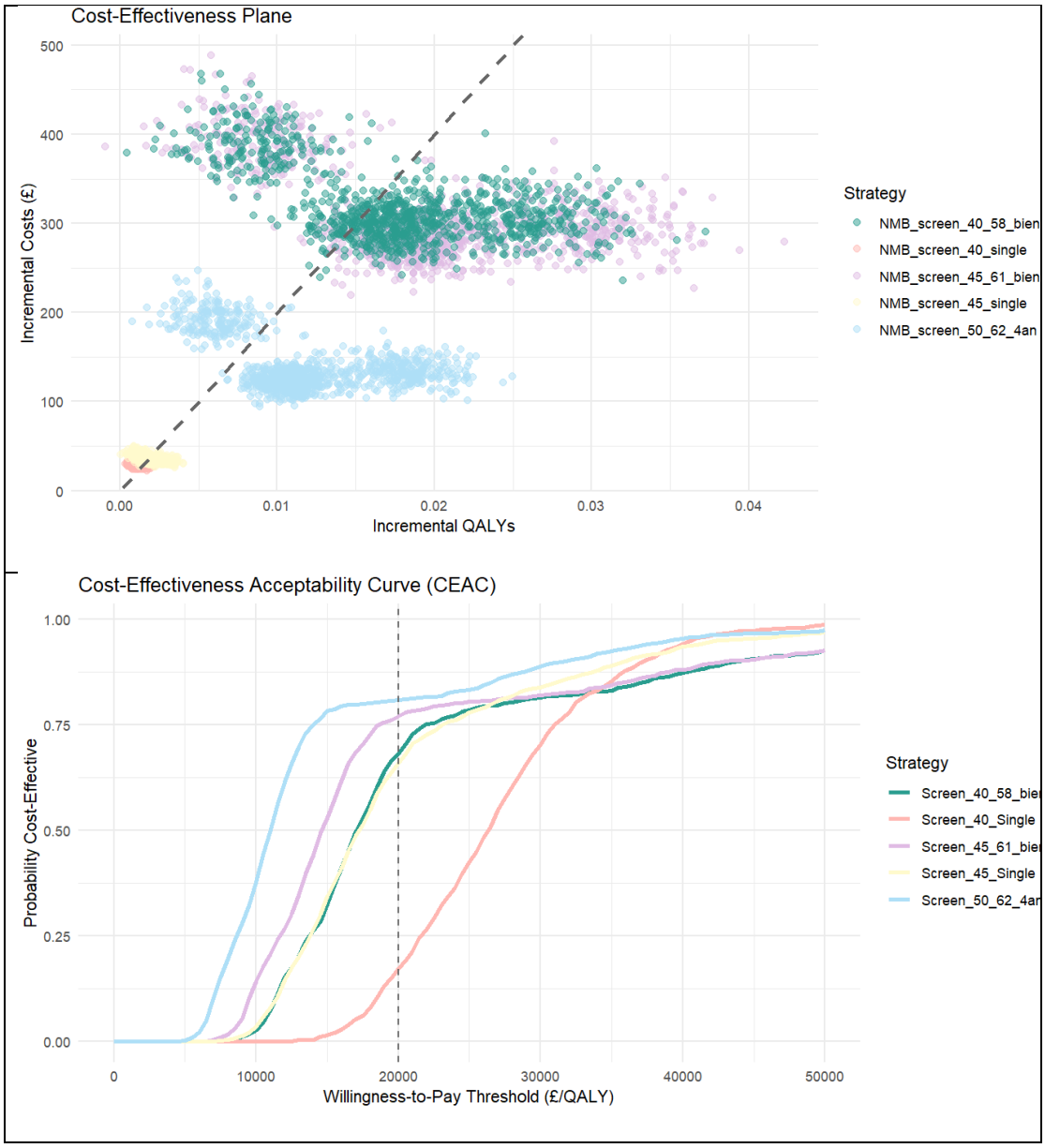


As seen on the cost-effectiveness plane, all scenarios showed relatively high uncertainty in cost-effectiveness results (Figure 44). This contrasts sharply with the probabilistic analysis assuming perfect uptake, where screening at ages 40 and 45 was cost-effective. The findings highlight the critical importance of achieving high uptake rates in the target population.

Single screening at age 40 had only 17% of probabilistic runs showing a positive incremental NMB (mean incremental NMB: -£6.1; 95% CrI: -£17 to £7). Confidence in cost-effectiveness was also low for screening at age 45, with 66% of runs positive (mean incremental NMB: £5; 95% CrI: -£26 to £34), Figure 44.

The highest mean incremental NMB and narrowest credible interval were observed for 4-yearly screening from ages 50 to 62. This strategy was cost-effective at the £20,000/QALY threshold in 81% of runs (mean incremental NMB: £99; 95% CrI: -£121 to £294). Biennial screening from ages 40 to 58 produced a mean incremental NMB of £36 (95% CrI: -£287 to £303) and was cost-effective in 68% of runs, while biennial screening from ages 45 to 61 had a mean incremental NMB of £87 (95% CrI: -£287 to £395) with 77% of runs positive.

Figure 44: Cost-effectiveness plane (a) and cost-effectiveness acceptability curve (b) (men of Black ethnicity scenarios, CAP trial uptake)



8.2.4. Men of Black ethnicity: scenario analyses

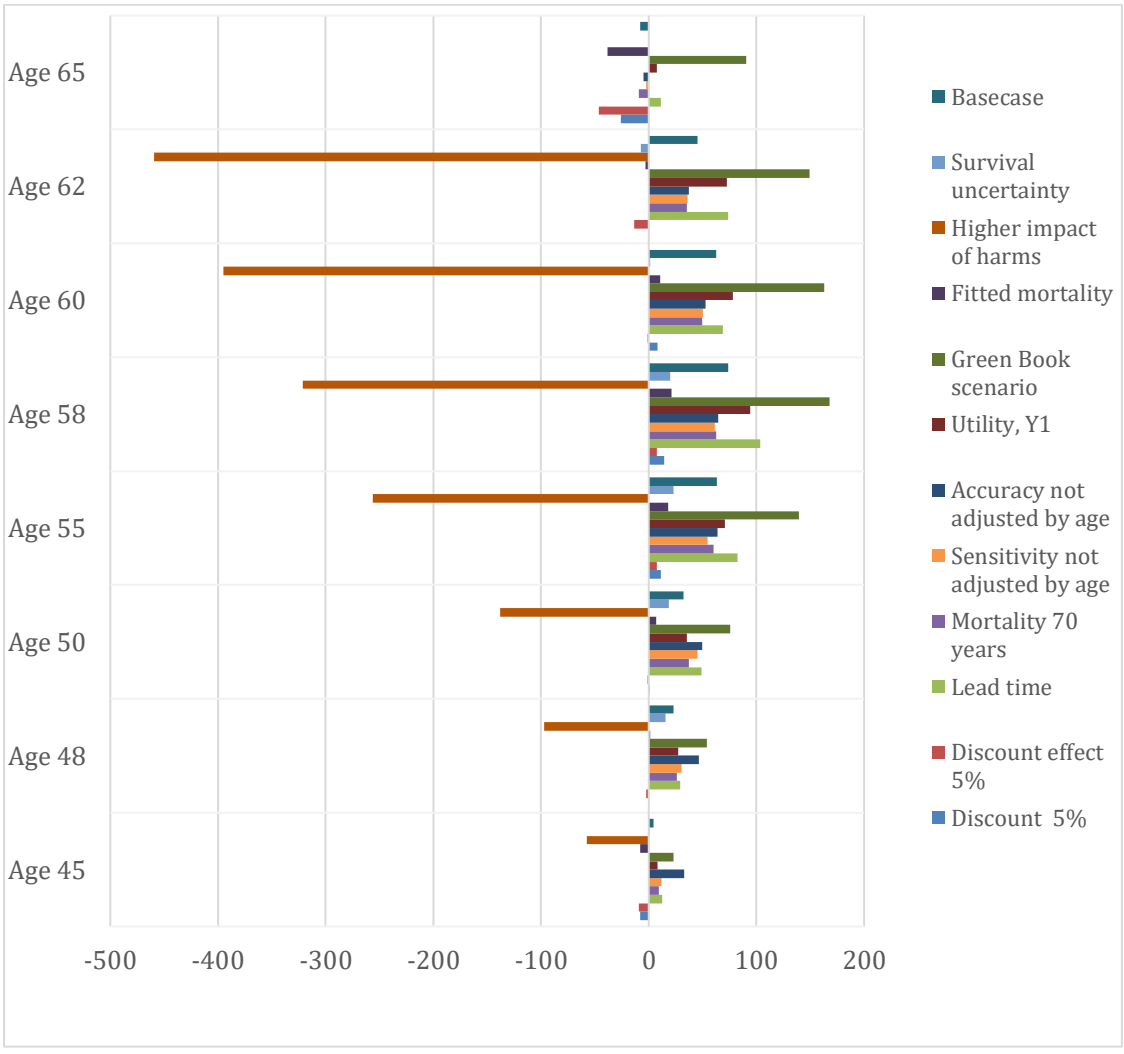
Multiple deterministic scenario analyses were conducted (Figure 45) to evaluate an impact of assumptions on incremental NMB.

For single screening, the NMB was positive in all except one (higher impact of cancer diagnosis on utilities, i.e. higher harms related to screening) tested scenarios at ages 55 and 58 years. Higher screening-related harms (negative impact on cost-effectiveness) and using the Green Book scenario (positive impact on cost-effectiveness) were the most impactful scenarios on NMB values.

For repeat screening, none of the programme designs resulted in a positive incremental NMB overall. However, screening men of Black ethnicity aged 50–62 once every four years showed positive incremental NMBs in all but two scenarios (applying a 5% discount rate for effects and higher harms for cancer diagnosis) (Figure 46).

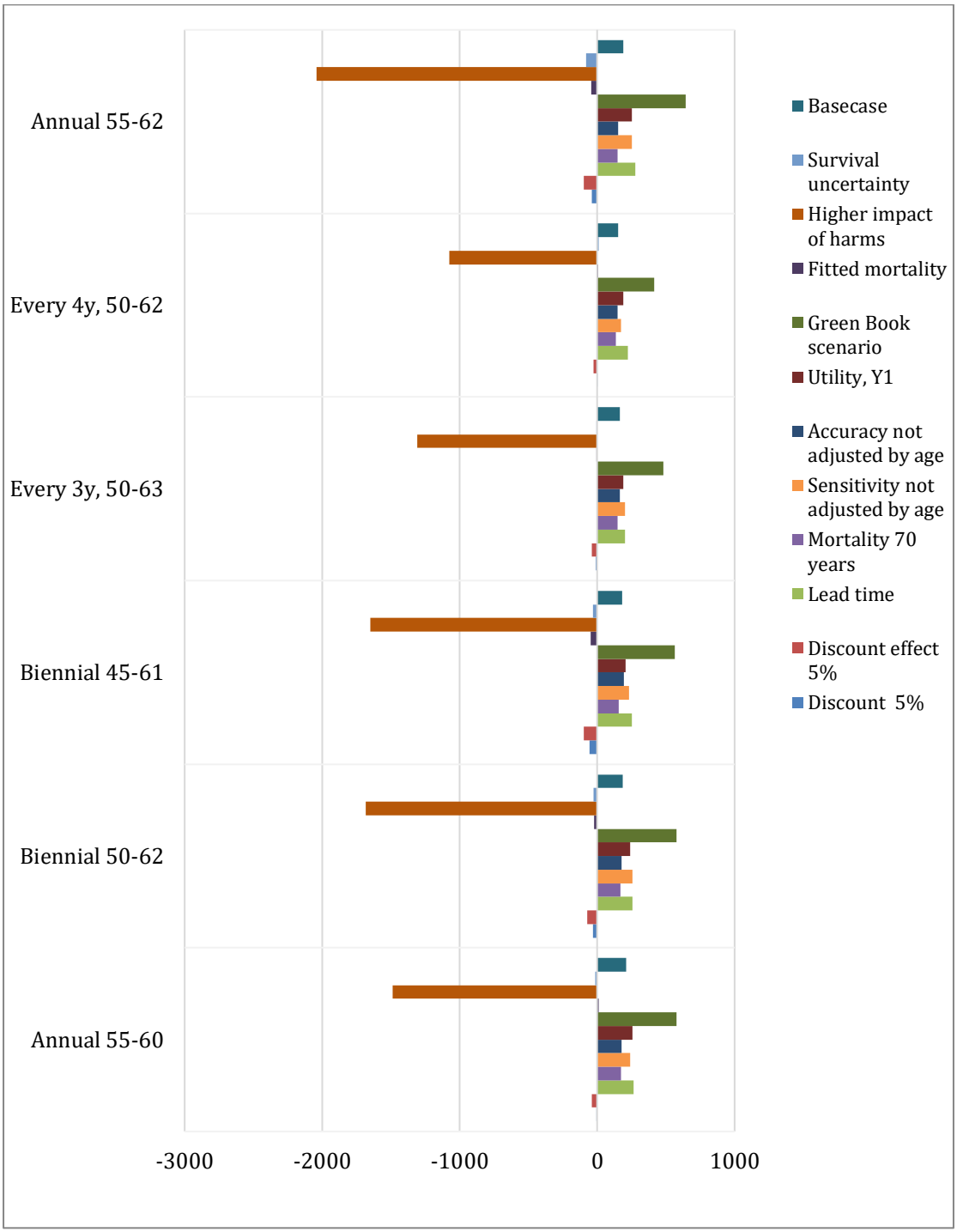
Similar to the general population, applying discounting by the approach of the Green Book and higher screening-related harms were the most impactful on cost-effectiveness results.

Figure 45: Net monetary benefit (£) in scenario analyses for one-time screening: men of Black ethnicity



Legend: £15,000 threshold is used in the Green Book scenario and £20,000 in the other scenarios.

Figure 46: Net monetary benefit (£) in scenario analyses for repeat screening: men of Black ethnicity

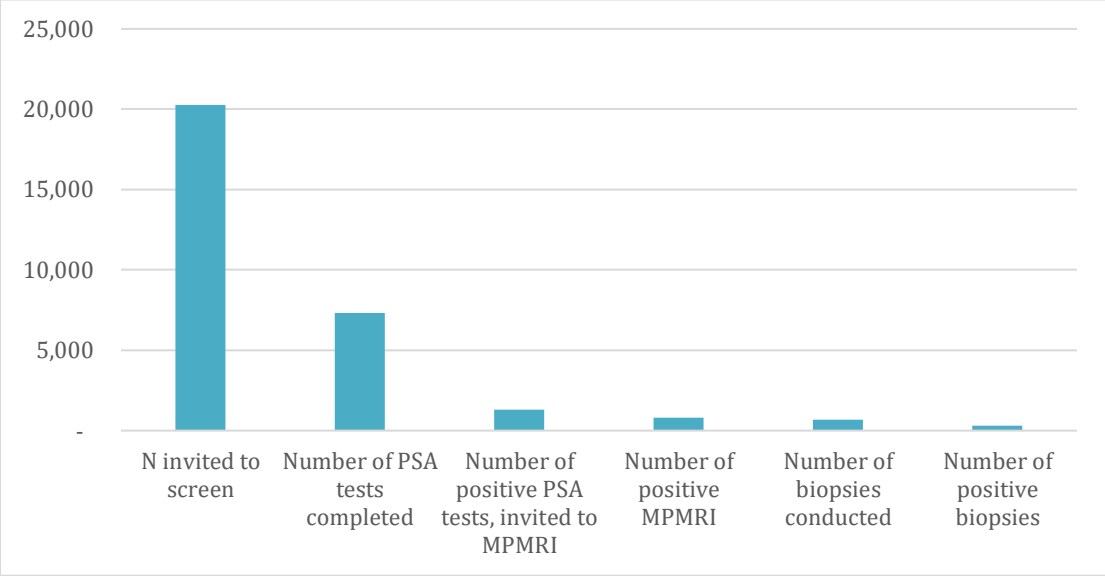


Legend: £15,000 threshold is used in the Green Book scenario and £20,000 in the other scenarios.

8.2.5. Men of Black ethnicity: impact of screening scaled to the population of England

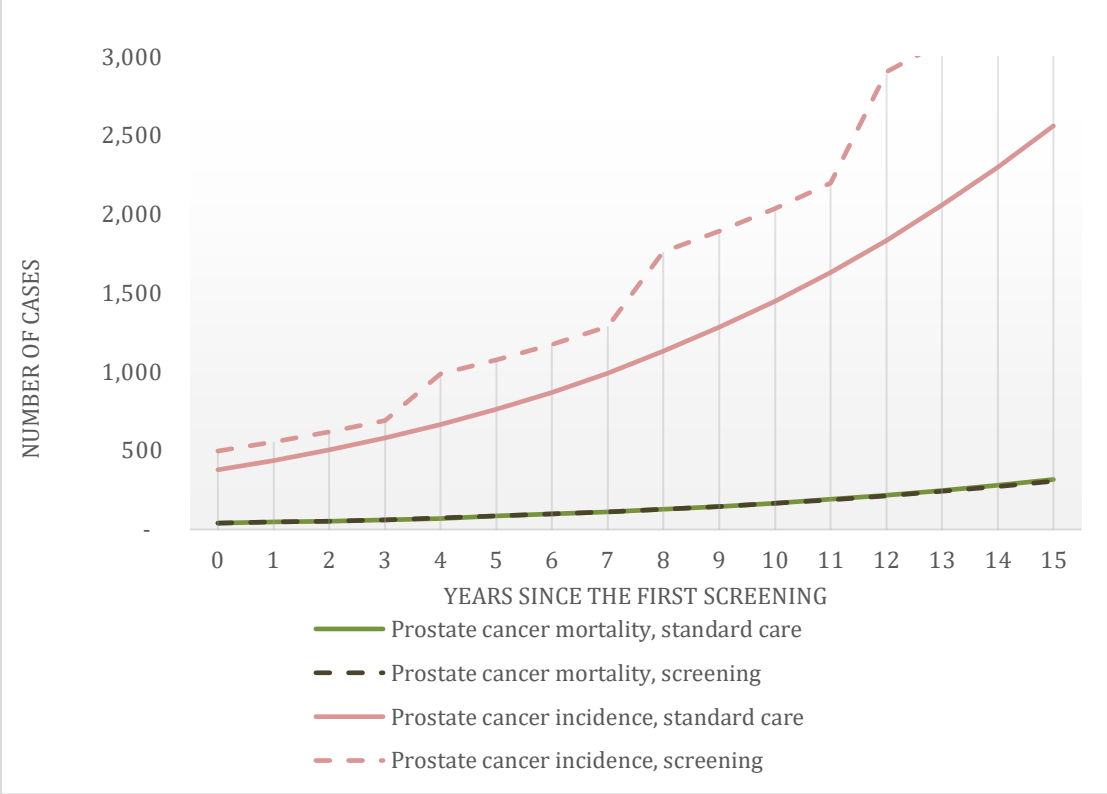
As only 5% of the population in England are Black, and with 2,027,500 men aged 50–55 years reported by the ONS in 2022, it is estimated that inviting to screening 50-year-old men of Black ethnicity (from ages 50 to 62 with a 4-year interval, which was selected as one of the most cost-effective strategies with lower resource use) would result in approximately 7,300 PSA screening completed annually. This would lead to around 790 follow-up mpMRI scans and 670 additional biopsies each year when the programme is fully enrolled (Figure 47).

Figure 47: Annual resource requirements in England for screening men of Black ethnicity aged 50 to 62 every four years



Over the 15 years following the implementation of an organised screening programme — screening men of Black ethnicity aged 50 to 62 every four years resulted in an additional 571 prostate cancer cases are expected to be detected when screening every 4 years and 1,142 addition prostate cancer cases are expected to be detected when screening every 2 years. The difference in prostate cancer deaths between the organised screening for men of Black ethnicity and standard care arms is projected to reach 1 case per cohort (or 7 cases in a fully rolled out programme) in probabilistic analyses and 11 cases in deterministic analysis after 15 years of follow up (Figure 48).

Figure 48. Difference in prostate cancer incidence and mortality between the organised screening and standard care arms among a cohort of 50-year-old Black men invited to be screened every four years from age 50 to 62.



8.3. Men with familial risk

8.3.1. Familial risk: single screening at different ages

Similar to other population groups, the initial model runs were conducted for men with familial risk with a single screening round to identify the most cost-effective screening ages.

The same trend of increased incidence of prostate cancer by age was observed in this subgroup, with a substantial increase after the age 62 (Figure 49). Similarly, there was a reduction in mortality, with the largest reduction observed in the cohort screened at ages 60-68 years (Figure 50).

Figure 49: Prediction of incidence with one-time screening in men with familial risk, per person

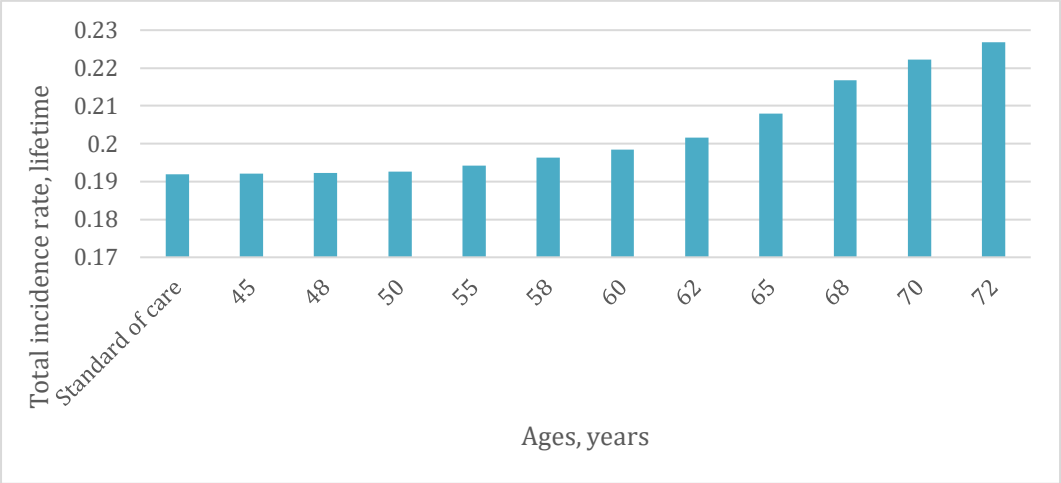
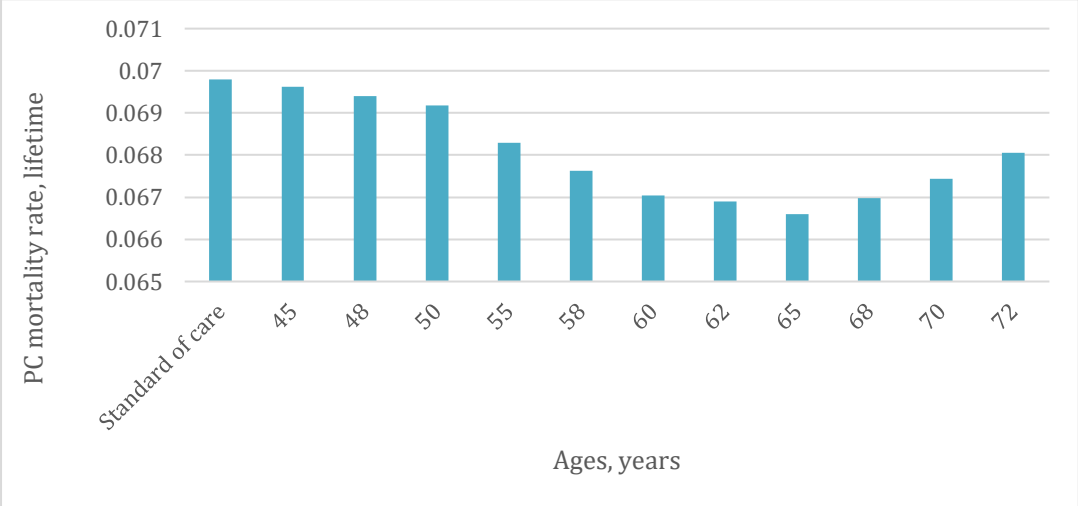


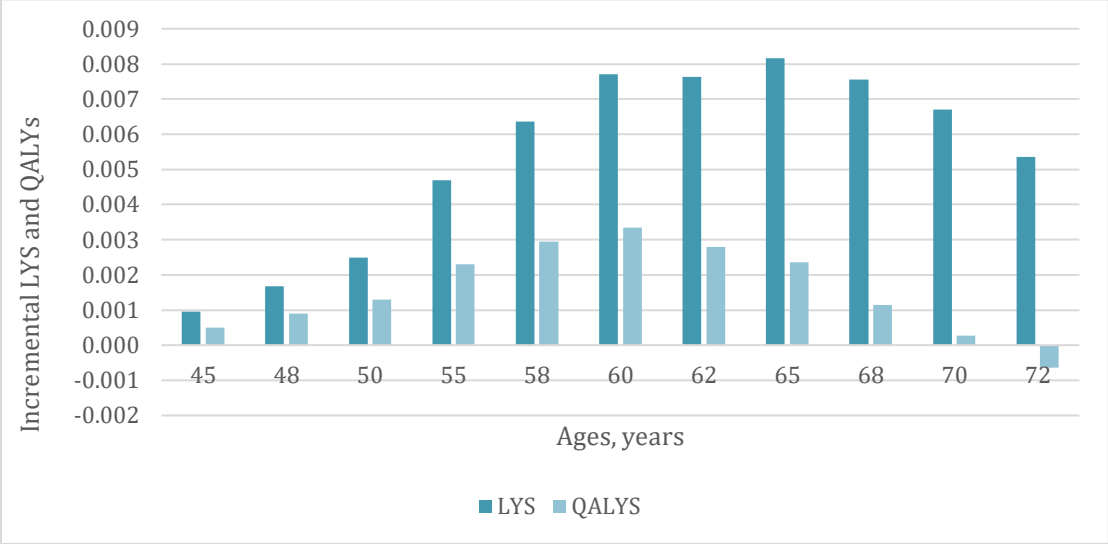
Figure 50: Prediction of lifetime prostate cancer mortality rate with one-time screening in men with familial risk



Legend: PC -prostate cancer

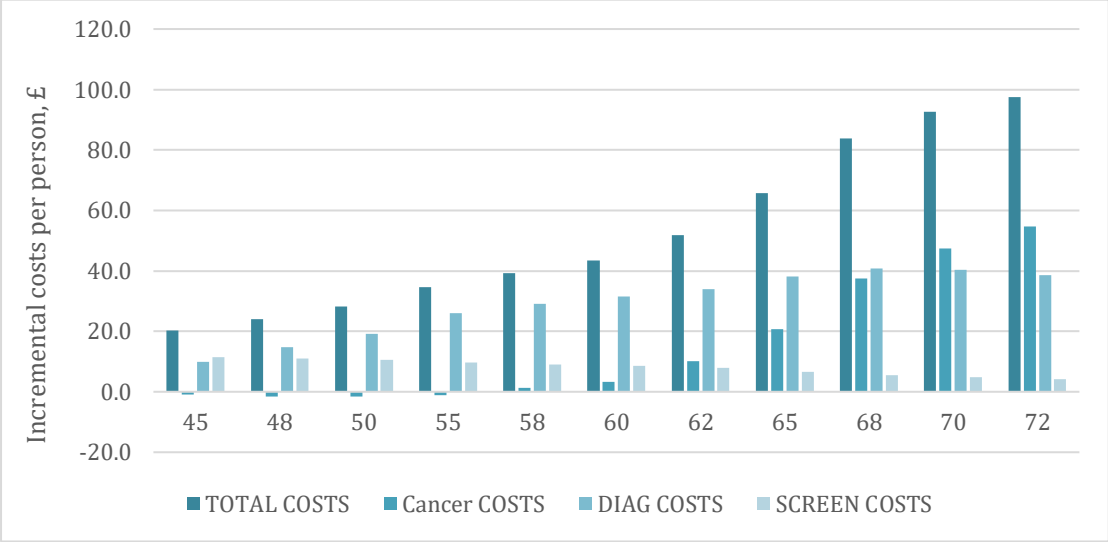
Screening led to gains in LYS across all age groups and increases in QALYs for those screened before age 72 (Figure 51). Screening also increased total costs (Figure 52). As expected, screening-related costs decreased with age due to fewer individuals being invited at older ages. Cancer treatment costs were lower in those younger than 58. The diagnostic follow-up costs for screen-positive cases and incremental cancer treatment costs generally increased with age up to the age 70, reflecting a higher prevalence of undiagnosed cancers in older populations.

Figure 51: Incremental LYS and QALYs per person with one-time screening of men with familial risk vs standard of care



Legend: LYS – life years saved; QALYS -quality adjusted life years

Figure 52: Incremental costs per person with one-time screening of men with familial risk vs standard of care



The incremental NMB of screening was positive in the 55–62 age group with discounting applied starting from the same model cycle (Figure 53) and in 48-65 group when each intervention was evaluated separately (Figure 54). However, this did not change the conclusion on the most cost-effective age of screening, which in both cases was age 60.

Figure 53: Incremental net monetary benefit with one-time screening in men with familial risk vs standard of care

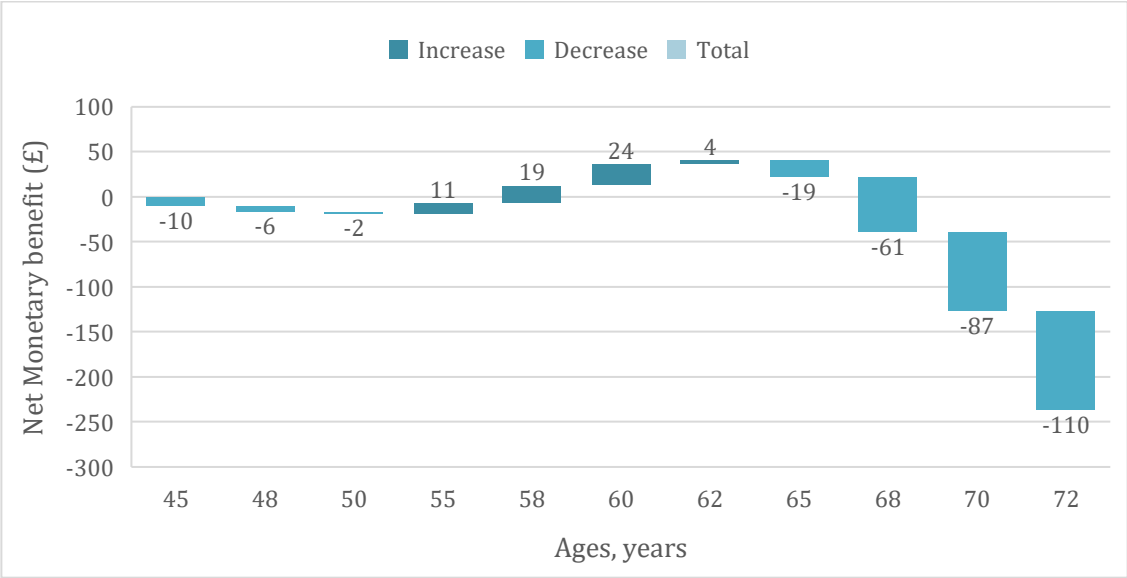
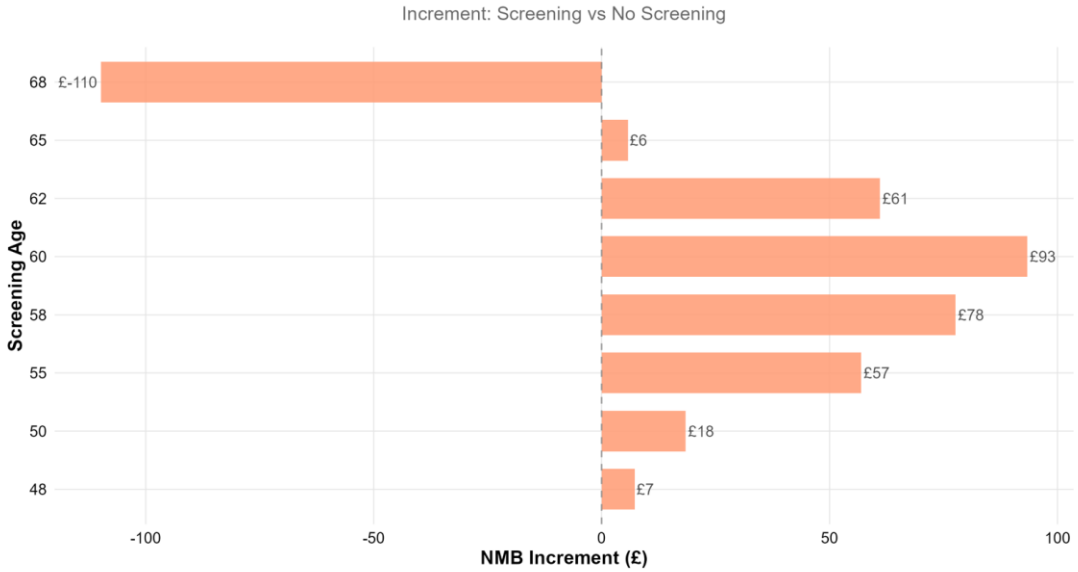


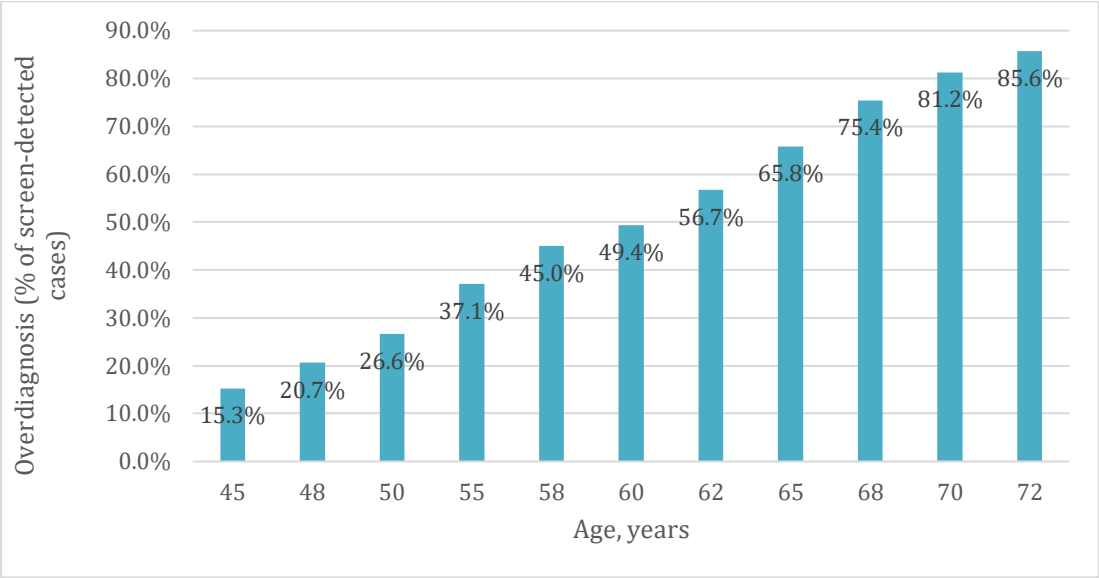
Figure 54: Incremental net monetary benefit with one-time screening in men with familial risk vs standard of care; scenario with adjusted discounting



Legend: NMB – net monetary benefit

Screening also resulted in substantial overdiagnosis rate with 15% of prostate cancer cases overdiagnosed if men with familial risk are screened at age 45 to 86% of cases if men are screened at age 72 (Figure 55).

Figure 55: Overdiagnosis with one-time screening of men with familial risk (increased incidence divided by the number of screen-detected cases)

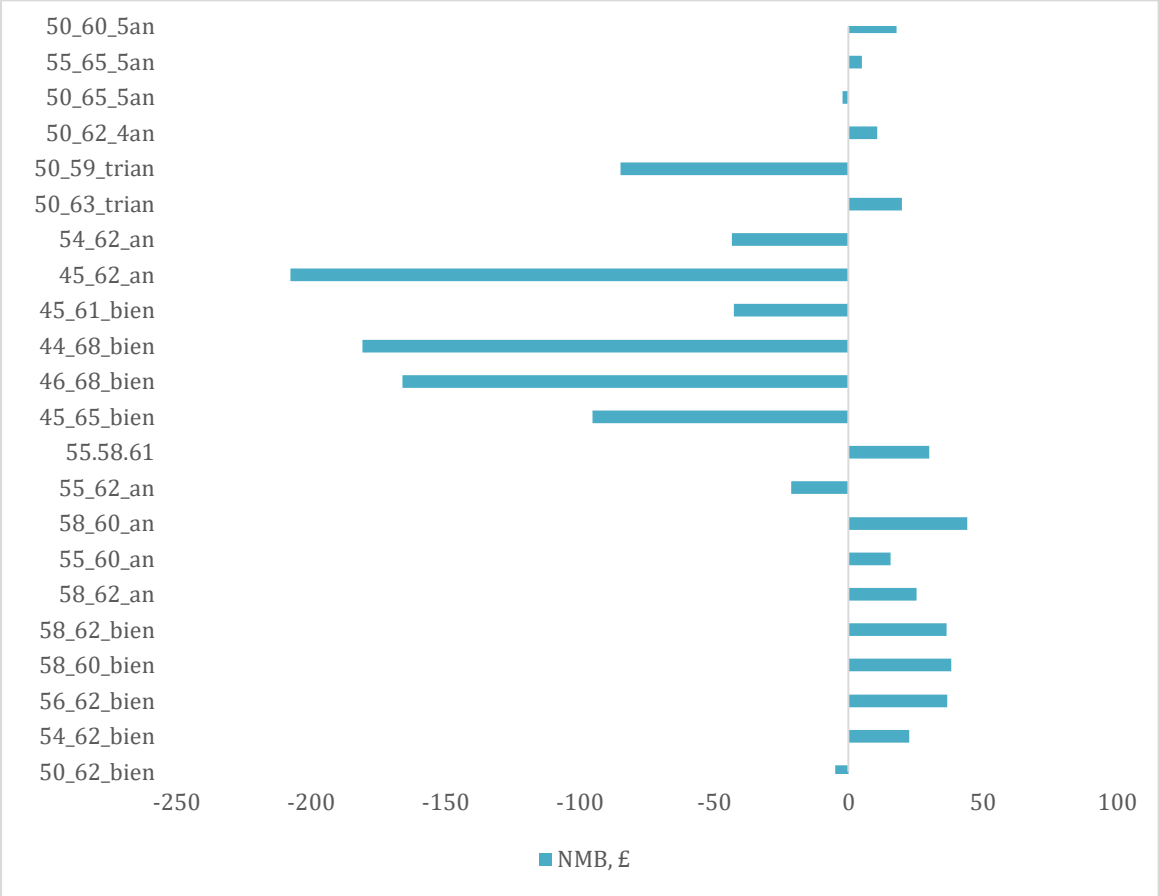


8.3.2. Familial risk: repeat screening scenarios

The results from the single screening analyses informed the design of the repeat screening strategies evaluated in the model.

Screening was cost-effective in some scenarios but not in others. It was cost-effective with annual screening between ages 55-60, 58-60 and 58–62; biennial screening at ages 50–62, 58–60, 56–62, 54–62, and 50–62; triennial screening at ages 55–61 and 50–63; screening every 4 years at ages 50–62; and screening every 5 years at ages 50-60 and 55–65 (Figure 56). In all scenarios, screening led to substantial overdiagnosis, with rates ranging from 46% to 63% (Figure 57).

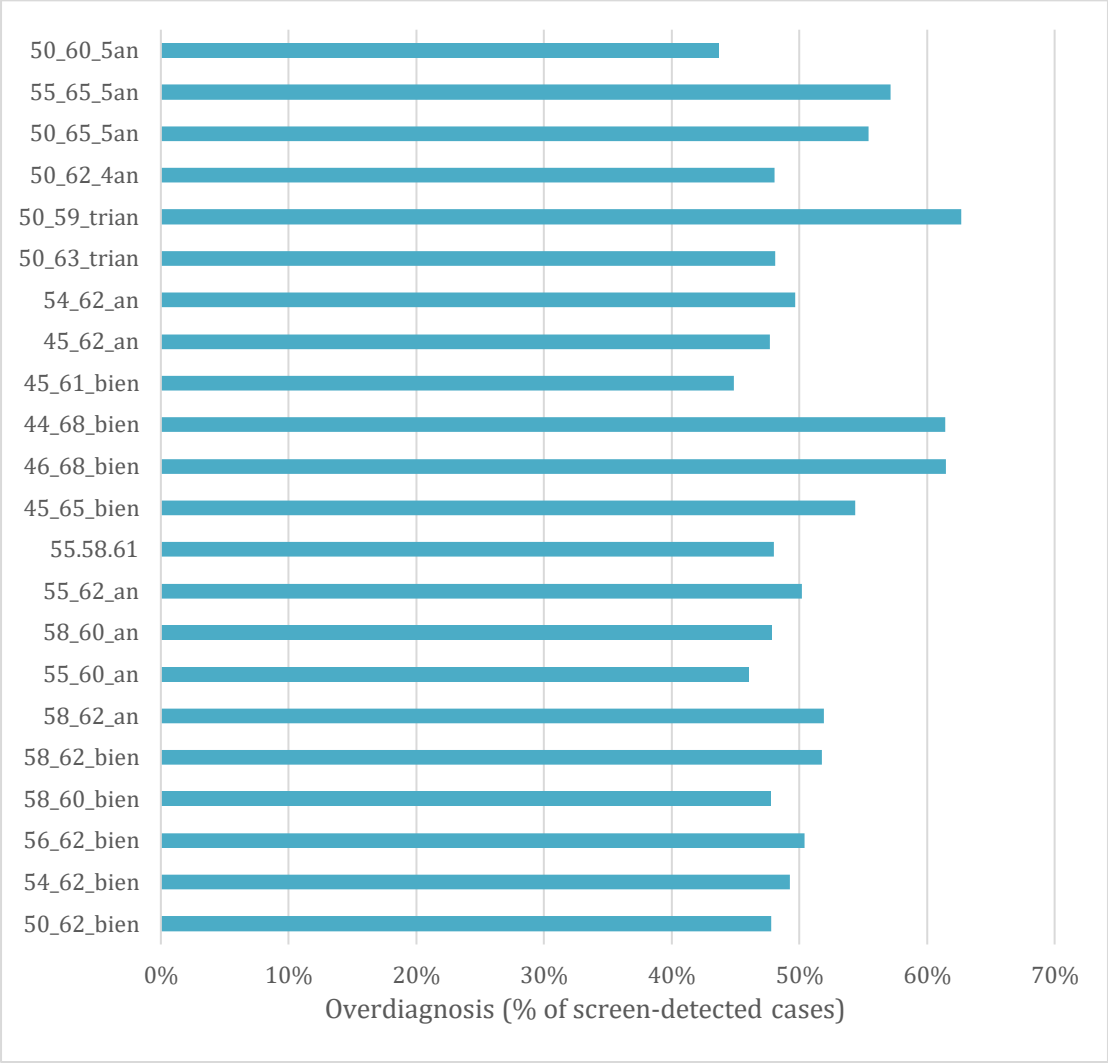
Figure 56: Incremental net monetary benefit with repeat screening in men with familial risk vs standard of care



Legend: NMB – net monetary benefit;

Screening every 5 years at ages 50–60 (screen_50_60_5an); Screening every 5 years at ages 55–65 (screen_55_65_5an); Screening every 5 years at ages 50–65 (screen_50_65_5an); Screening every 4 years at ages 50–62 (screen_50_62_4an); Triennial screening at ages 50–59 (screen_50_59_3an); Triennial screening at ages 50–63 (screen_50_63_trian); Annual screening at ages 54–62 (screen_54_62_an); Annual screening at ages 45–62 (screen_45_62_an); Biennial screening at ages 45–61 (screen_45_61_bien); Biennial screening at ages 44–61 (screen_45_68_bien); Biennial screening at ages 46–66 (screen_46_68_bien); Biennial screening at ages 45–65 (screen_45_65_bien); Triennial screening at ages 55–61 (screen_55.58.61); Annual screening at ages 55–62 (screen_55_62_an); Biennial screening at ages 58–60 (screen_58_60_bien); Annual screening at ages 55–60 (screen_55_60_an); Biennial screening at ages 58–62 (screen_58_62_bien); Annual screening at ages 58–60 (screen_58_60_an); Biennial screening at ages 56–62 (screen_56_62_bien); Biennial screening at ages 54–62 (screen_54_62_bien); Biennial screening at ages 50–62 (screen_50_62_bien).

Figure 57: Overdiagnosis rate with repeat screening in men with familial risk (increment in incidence divided by the number of screen-detected cases)



Legend: Screening every 5 years at ages 50–60 (screen_50_60_5an); Screening every 5 years at ages 55–65 (screen_55_65_5an); Screening every 5 years at ages 50–65 (screen_50_65_5an); Screening every 4 years at ages 50–62 (screen_50_62_4an); Triennial screening at ages 50–59 (screen_50_59_3an); Triennial screening at ages 50–63 (screen_50_63_trian); Annual screening at ages 54–62 (screen_54_62_an); Annual screening at ages 45–62 (screen_45_62_an); Biennial screening at ages 45–61 (screen_45_61_bien); Biennial screening at ages 44–61 (screen_45_68_bien); Biennial screening at ages 46–66 (screen_46_68_bien); Biennial screening at ages 45–65 (screen_45_65_bien); Triennial screening at ages 55–61 (screen_55.58.61); Annual screening at ages 55–62 (screen_55_62_an); Biennial screening at ages 58–60 (screen_58_60_bien); Annual screening at ages 55–60 (screen_55_60_an); Biennial screening at ages 58–62 (screen_58_62_bien); Annual screening at ages 58–60 (screen_58_60_an); Biennial screening at ages 56–62 (screen_56_62_bien); Biennial screening at ages 54–62 (screen_54_62_bien); Biennial screening at ages 50–62 (screen_50_62_bien).

8.3.3. Familial risk: probabilistic analyses

Probabilistic analysis for screening men who have first degree relatives with prostate, breast or ovarian cancers was first conducted assuming perfect uptake to assess the net impact of screening, demonstrating consistently that single screening had higher NMB than repeat screening (Figure 58 and Figure 59).

Screening men with familial risk at age 40 was not cost-effective (probability of cost-effectiveness: 0.3%; mean incremental NMB –£11, 95% CrI –£19 to –£2). In contrast, single screening at older ages (50, 58, and 60) showed substantial uncertainty around cost-effectiveness, with small positive mean incremental NMBs and credible intervals centred around zero, despite 76–81% of probabilistic runs yielding positive incremental NMBs. Specifically, the mean incremental NMBs (95% CrIs) were £5.6 (–£22 to £27), £23 (–£73 to £83), and £15 (–£62 to £60) for screening at ages 50, 58, and 60, respectively (Figure 60).

These findings were broadly similar to those for screening the general population but with slightly reduced uncertainty. Repeat screening strategies were not cost-effective: only 3% of probabilistic runs were positive for triennial screening from ages 50 to 63 (mean incremental NMB –£96, 95% CrI –£302 to £4), and 18% were positive for screening once every four years from ages 50 to 62 (mean incremental NMB –£55, 95% CrI –£241 to £37).

Figure 58: Incremental net monetary benefit (£) vs standard of care (men with familial risk, perfect uptake, single screening)

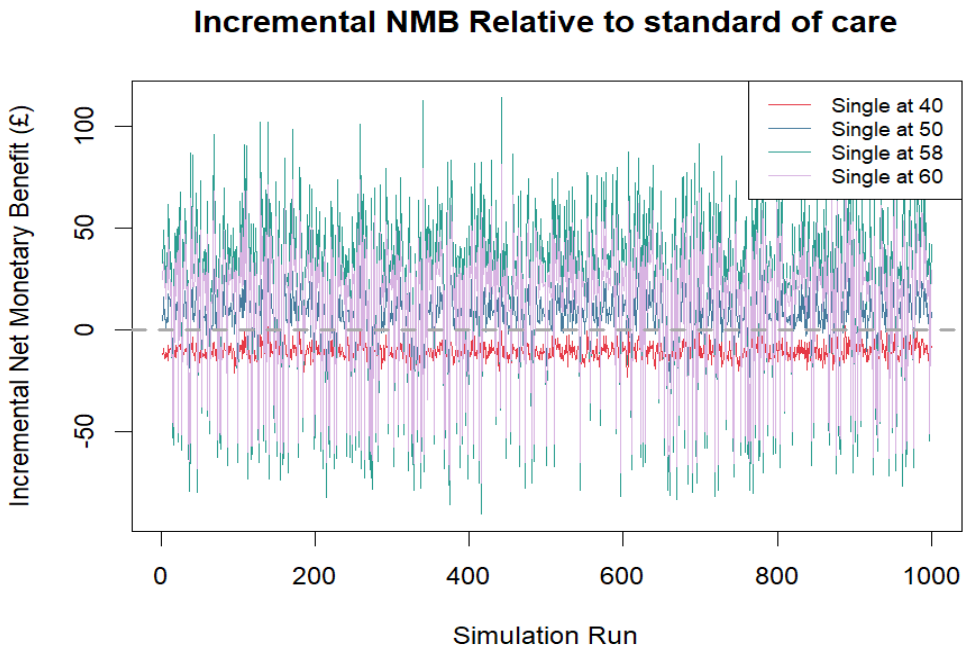


Figure 59: Incremental net monetary benefit (£) vs standard of care (men with familial risk, perfect uptake, repeat screening)

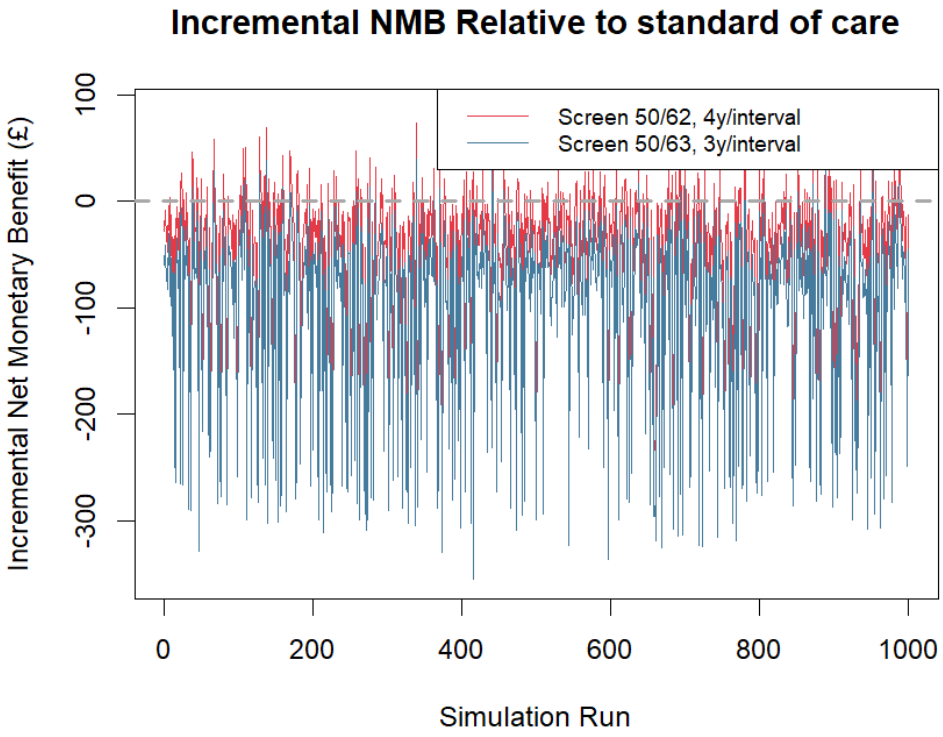
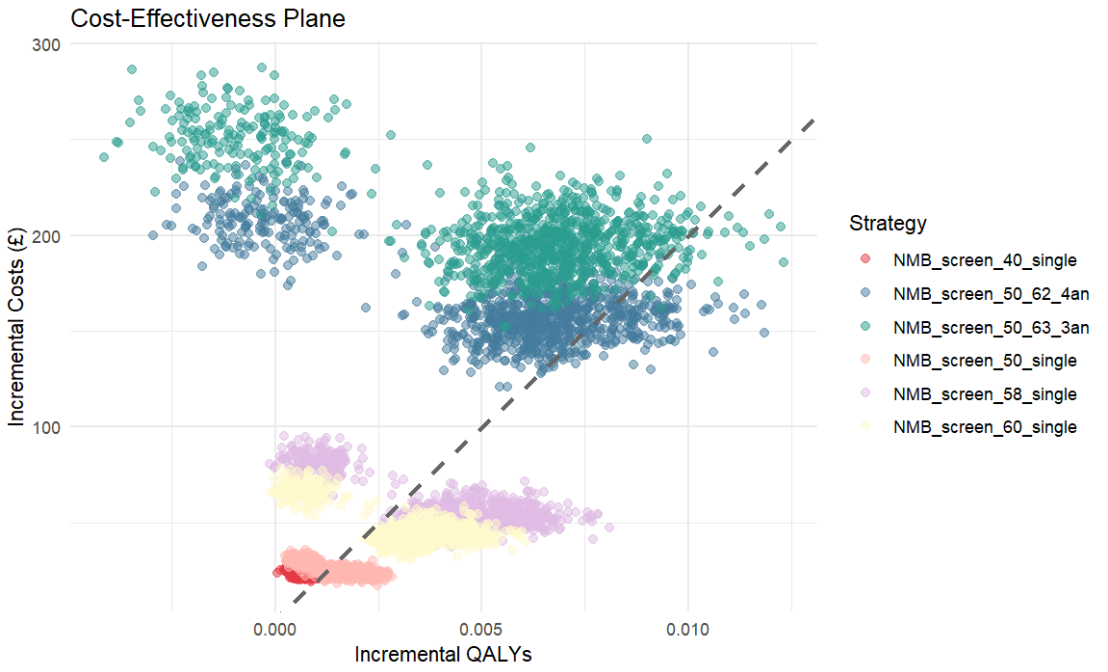
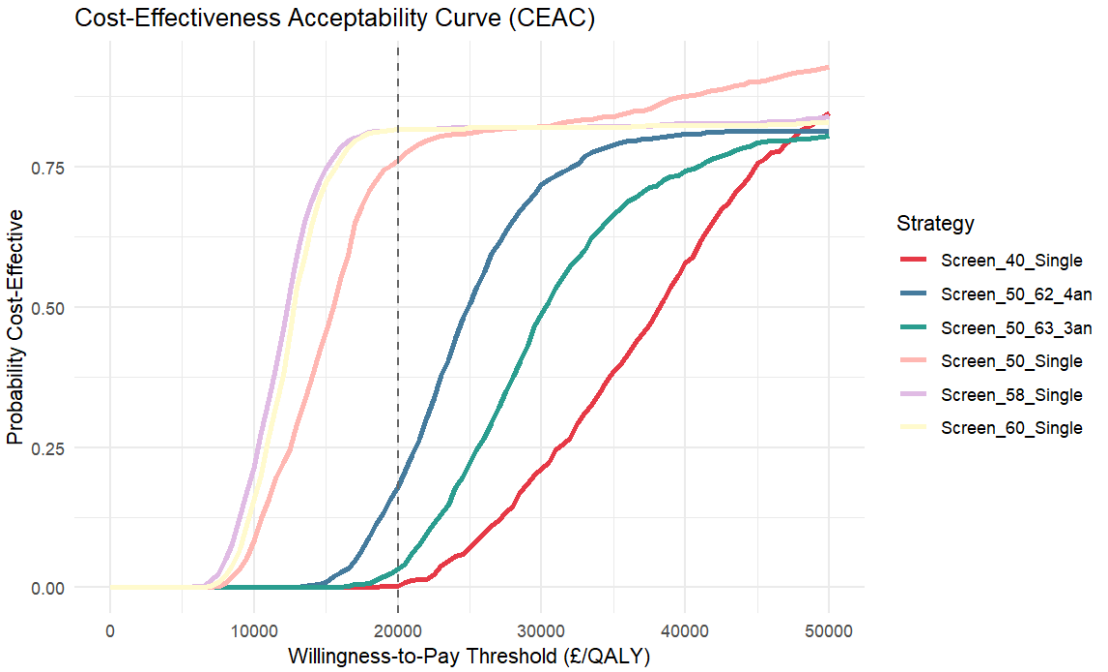


Figure 60: Cost-effectiveness plane (a) and cost-effectiveness acceptability curve (b) (men with familial risk, perfect uptake)





Probabilistic analysis for screening men with familial risk was also performed in the 58–60 age group, using the uptake rates from the CAP trial. This aimed to assess the impact of screening under more realistic conditions and to compare it with repeat screening of men in the general population.

The probabilistic analysis revealed substantial uncertainty in the cost-effectiveness of PSA screening for men with familial risk, even in this most cost-effective deterministic scenario (Figure 61).

While screening had a 78% probability of being cost-effective, the 95% CrI was wide, similar to the all-men population, with a mean incremental NMB of £59 (95% CrI: –£120 to £230; Figure 62). This indicates that, for men with familial risk, considerable uncertainty remains regarding the cost-effectiveness of PSA screening.

Figure 61: Incremental net monetary benefit (£) vs standard of care (men with familial risk, CAP uptake, annual screening of 58–60-year-old men)

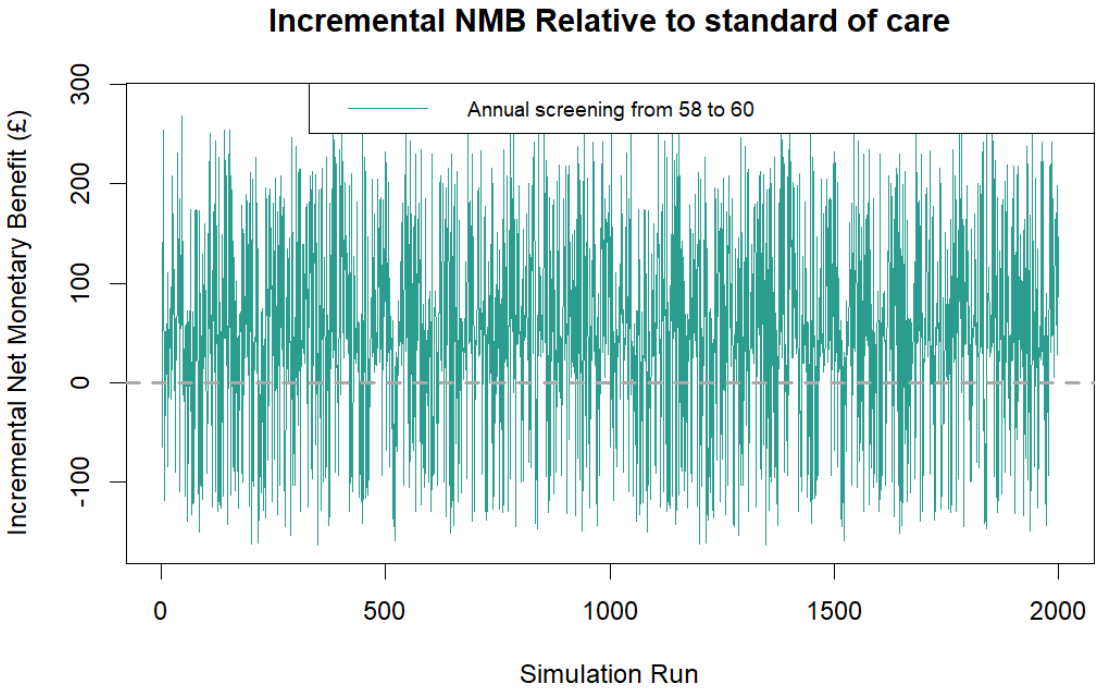
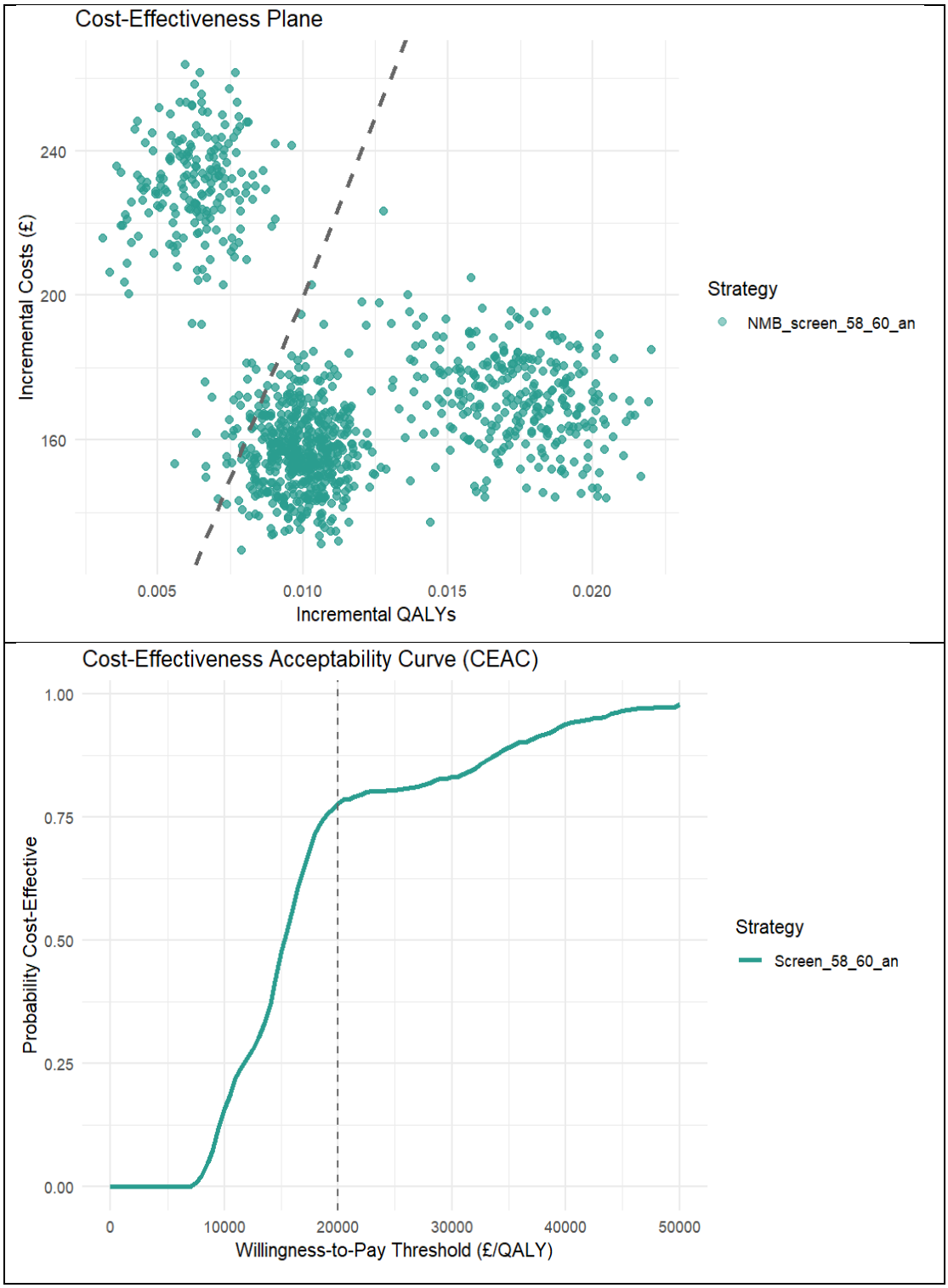


Figure 62: Cost-effectiveness plane (a) and cost-effectiveness acceptability curve (b) (men with familial risk, CAP uptake, annual screening of 58–60-year-old men)

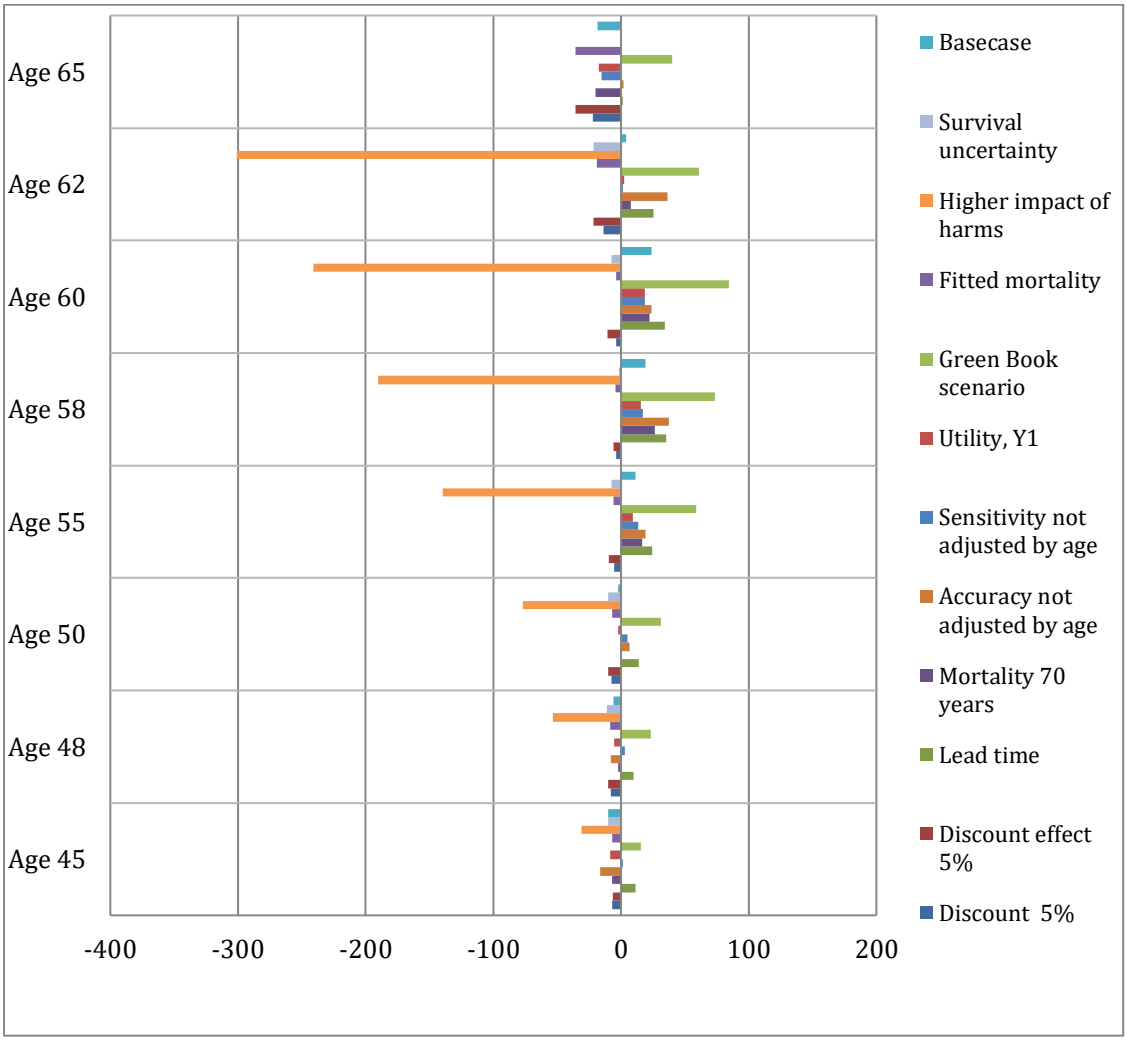


8.3.4. Familial risk: scenario analyses

As in other populations, applying higher impact of harms related to cancer diagnosis (i.e. lower HRQoL values for each cancer stage) and the Green Book approach to discounting costs and effects and using £15,000 decision threshold had the greatest impact on incremental NMB estimates.

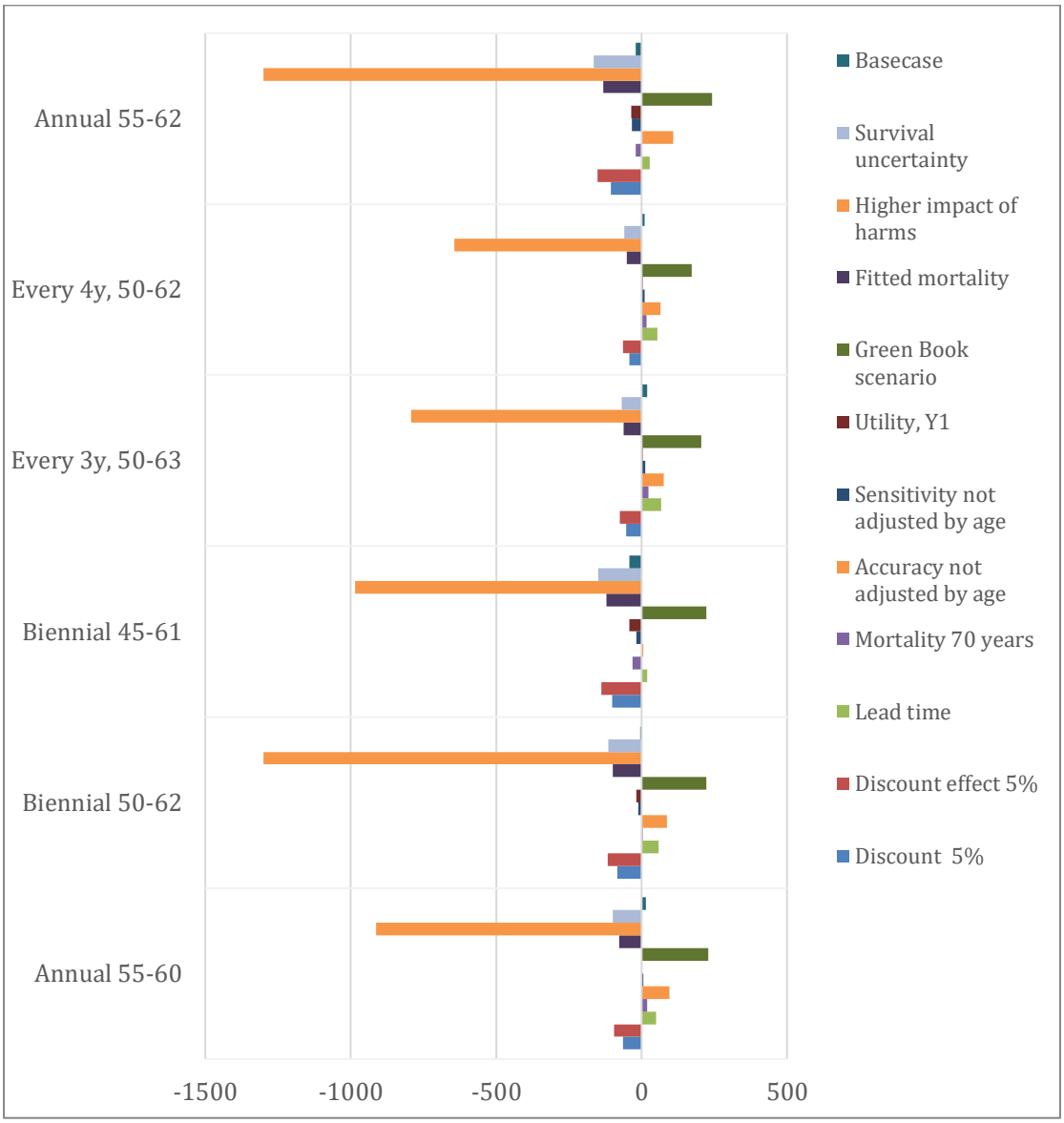
Sensitivity analyses also showed that single screening scenarios were more likely to result in a positive incremental NMB compared to repeat screening (Figure 63, Figure 64). Single screening at ages 58 and 60 produced the highest NMB estimates. Among repeat screening options, screening every four years between ages 50–62 or every three years between ages 50–63 showed the fewest scenarios with negative incremental NMB.

Figure 63: Net monetary benefit (£) in scenario analyses for one-time screening: men with familial risk



Legend: £15,000 threshold is used in the Green Book scenario and £20,000 in the other scenarios.

Figure 64: Net monetary benefit (£) in scenario analyses for repeat screening: men with familial risk



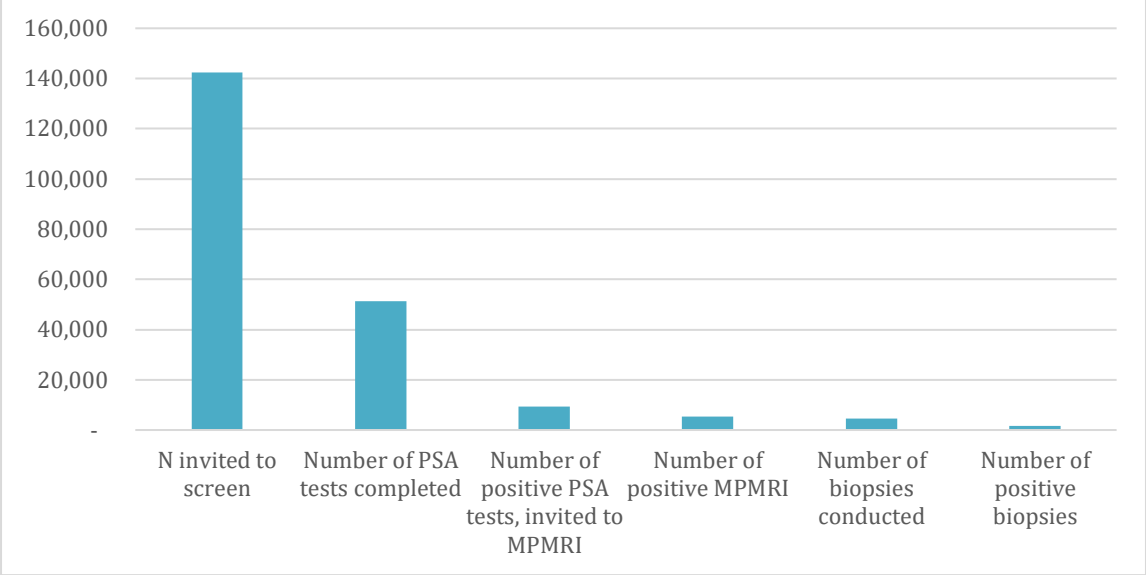
Legend: £15,000 threshold is used in the Green Book scenario and £20,000 in the other scenarios.

8.3.5. Men with familial risk: impact of screening scaled to the population of England

In the model, 46.9% of men had at least one first-degree relative with breast, ovarian, or prostate cancer by the end of their lifetime. By age 58, the probability of having such a family history was 0.7776, corresponding to 36% of men aged 58. This translated to an estimated 142,438 men with known familial risk being invited to screening at age 58. Annually, this would require inviting approximately

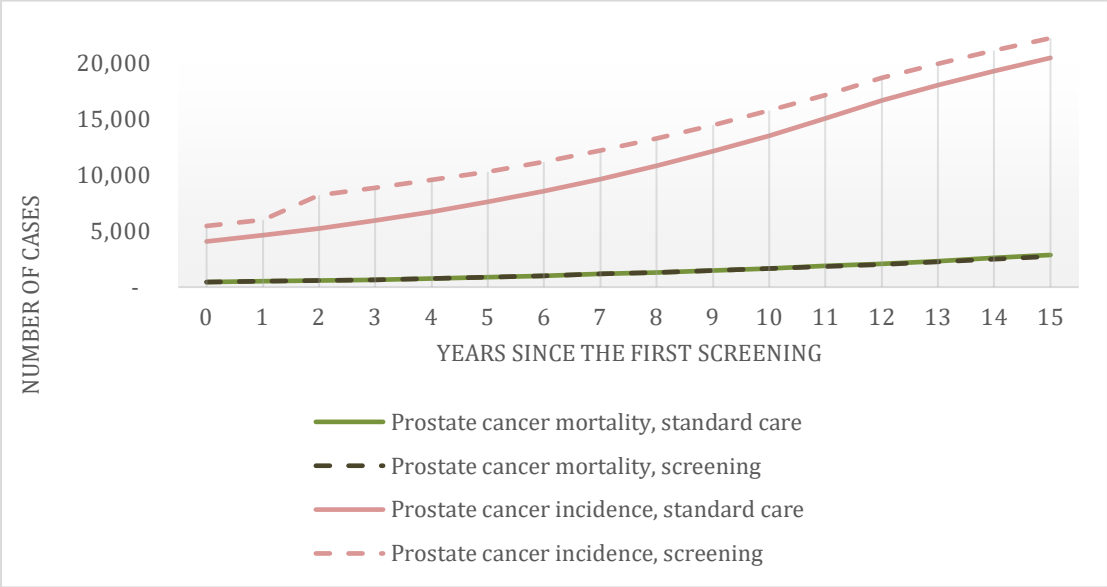
9,000 men for follow-up mpMRI and performing around 4,500 biopsies (Figure 65).

Figure 65: Resource requirement in England to screen men with familial risk at ages 58 to 60 annually



Over the 15 years following the implementation of an organised screening programme — screening men with familial risk at ages 58 and 60 resulted in an additional 2,170 prostate cancer cases expected to be detected in probabilistic analysis. The difference in prostate cancer deaths between the organised screening for men with familial risk and standard care arms is projected to reach 58 cases after 15 years of follow up in probabilistic analysis and 132 cases in deterministic analysis (Figure 66).

Figure 66. Difference in prostate cancer incidence and mortality between the organised screening and standard care arms among a cohort of 58-year-old men with familial risk invited to be screened at ages 58 and 60.



8.4. BRCA carriers

As in the other population groups, the initial model runs for BRCA carriers were conducted with a single screening round to identify the most cost-effective screening ages. An increase in incidence rates was observed up to age 72 (Figure 67), with the greatest impact on mortality achieved by a single screen at ages 60 - 62 (Figure 68).

8.4.1. BRCA carriers: single screening at different ages

Figure 67: Prediction of incidence with one-time screening in BRCA carriers, per person

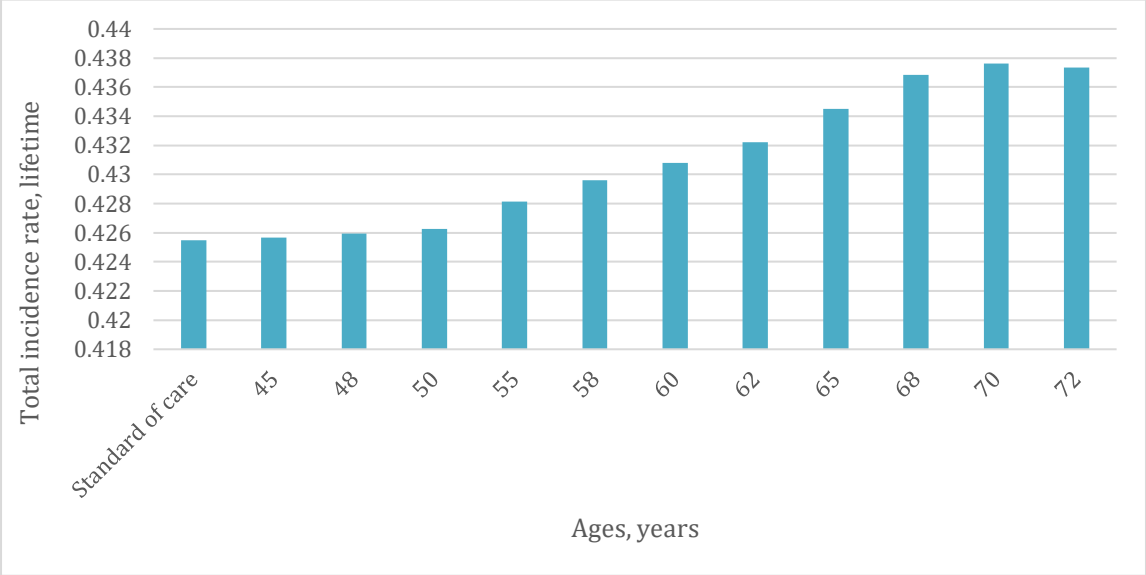
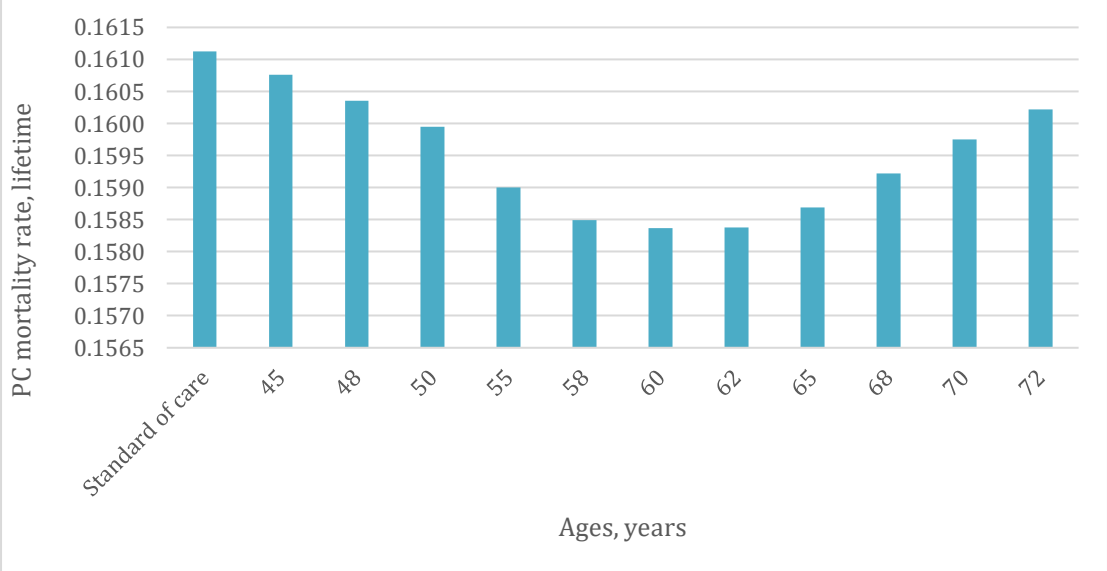


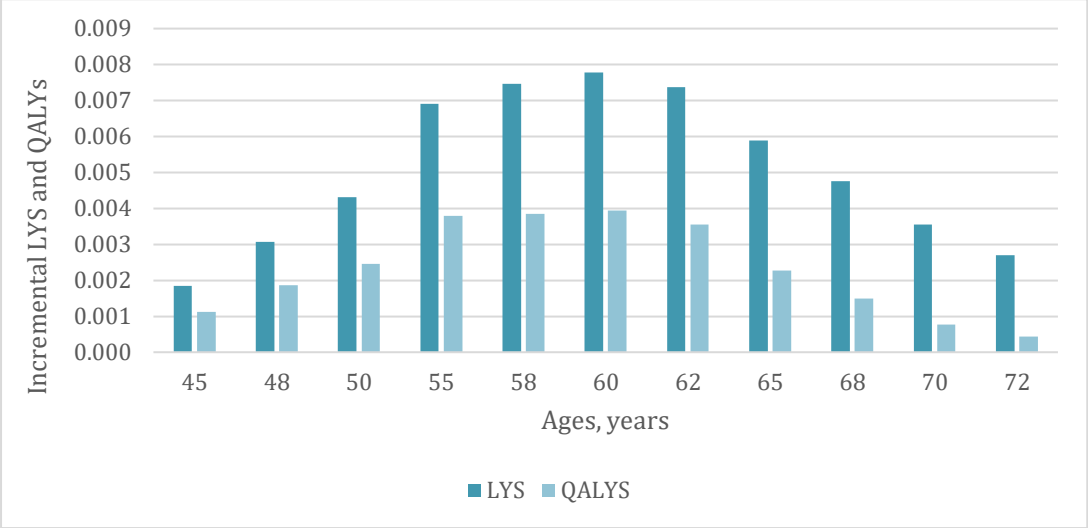
Figure 68: Prediction of lifetime prostate cancer mortality rate with one-time screening in BRCA carriers



Legend: PC -prostate cancer

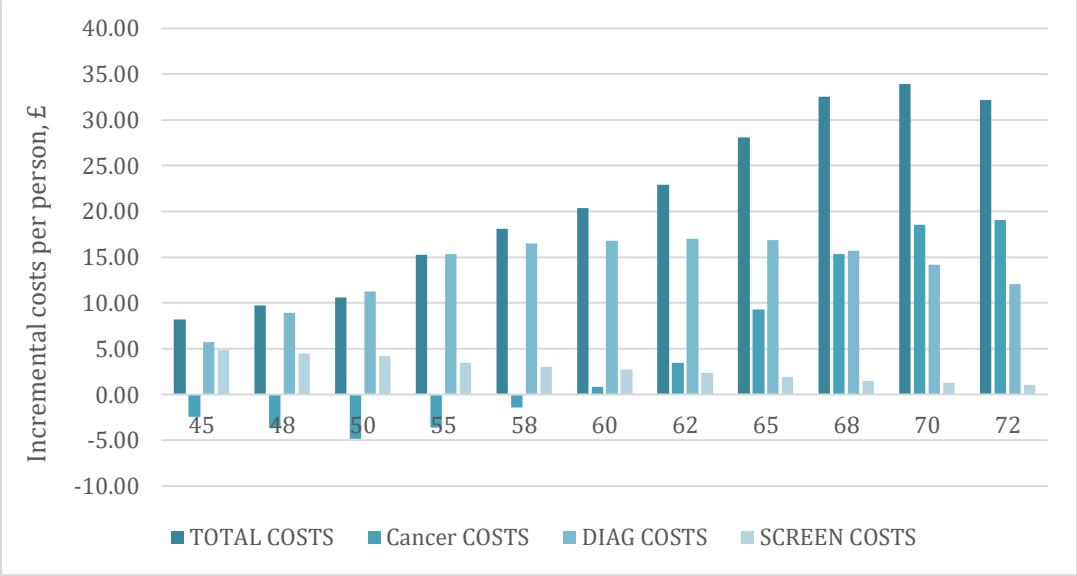
Screening resulted in increments in both LYS and QALYs in all ages (Figure 69), and in lower treatment costs with screening before ages 60 but higher total costs in all age groups (Figure 70).

Figure 69: Incremental LYS and QALYs per person with one-time screening in BRCA carriers vs standard of care



Legend: LYS – life years saved; QALYS -quality adjusted life years

Figure 70: Incremental costs per person with one-time screening in BRCA carriers vs standard of care



NMB was positive for single screening conducted before age 68, with the highest values observed for screening at ages 55–60 (Figure 71). Modelling each intervention separately—by applying discounting from the cycle of the first

intervention for both the screening and no-screening arms—did not change the conclusions regarding the most cost-effective age. However, under this approach, screening at age 68 was also found to be cost-effective (Figure 72).

Figure 71: Incremental net monetary benefit with one-time screening in BRCA carriers vs standard of care

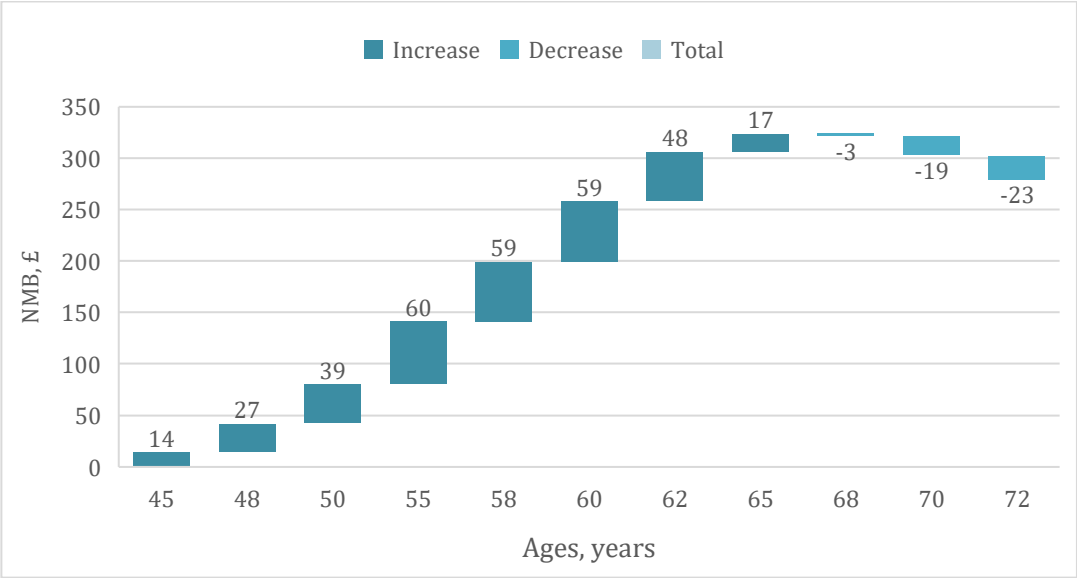
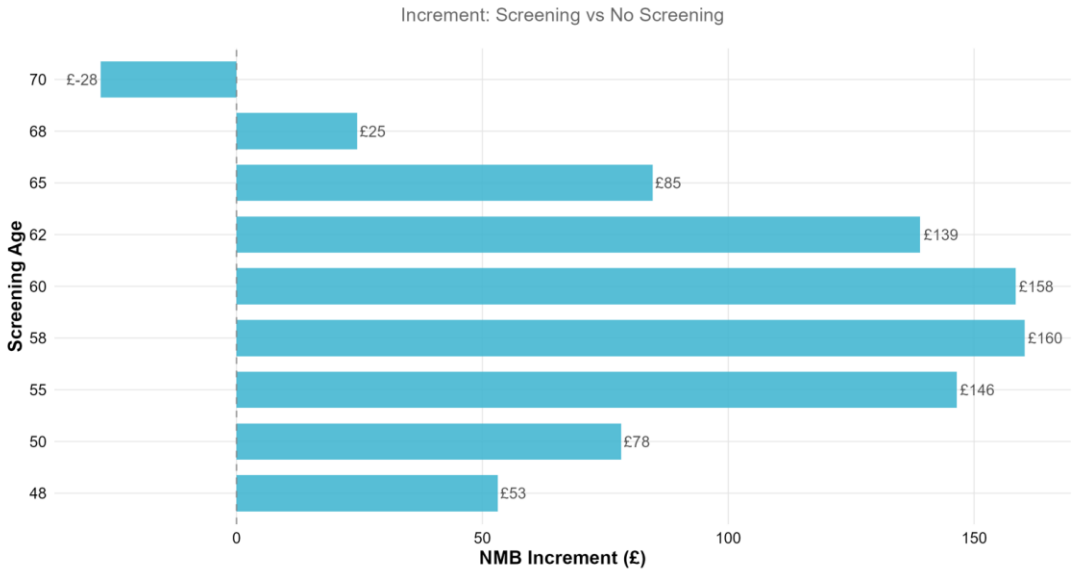
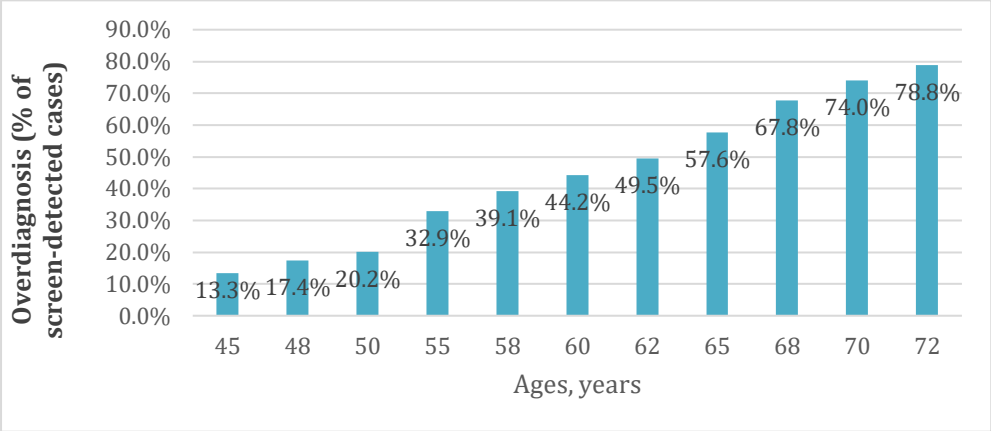


Figure 72: Incremental net monetary benefit with one-time screening in BRCA carriers vs standard of care; scenario with adjusted discounting



Similar to other population groups, screening in BRCA carriers led to substantial overdiagnosis, ranging from 13% with a single screen at age 45 to 79% at age 72 (Figure 73).

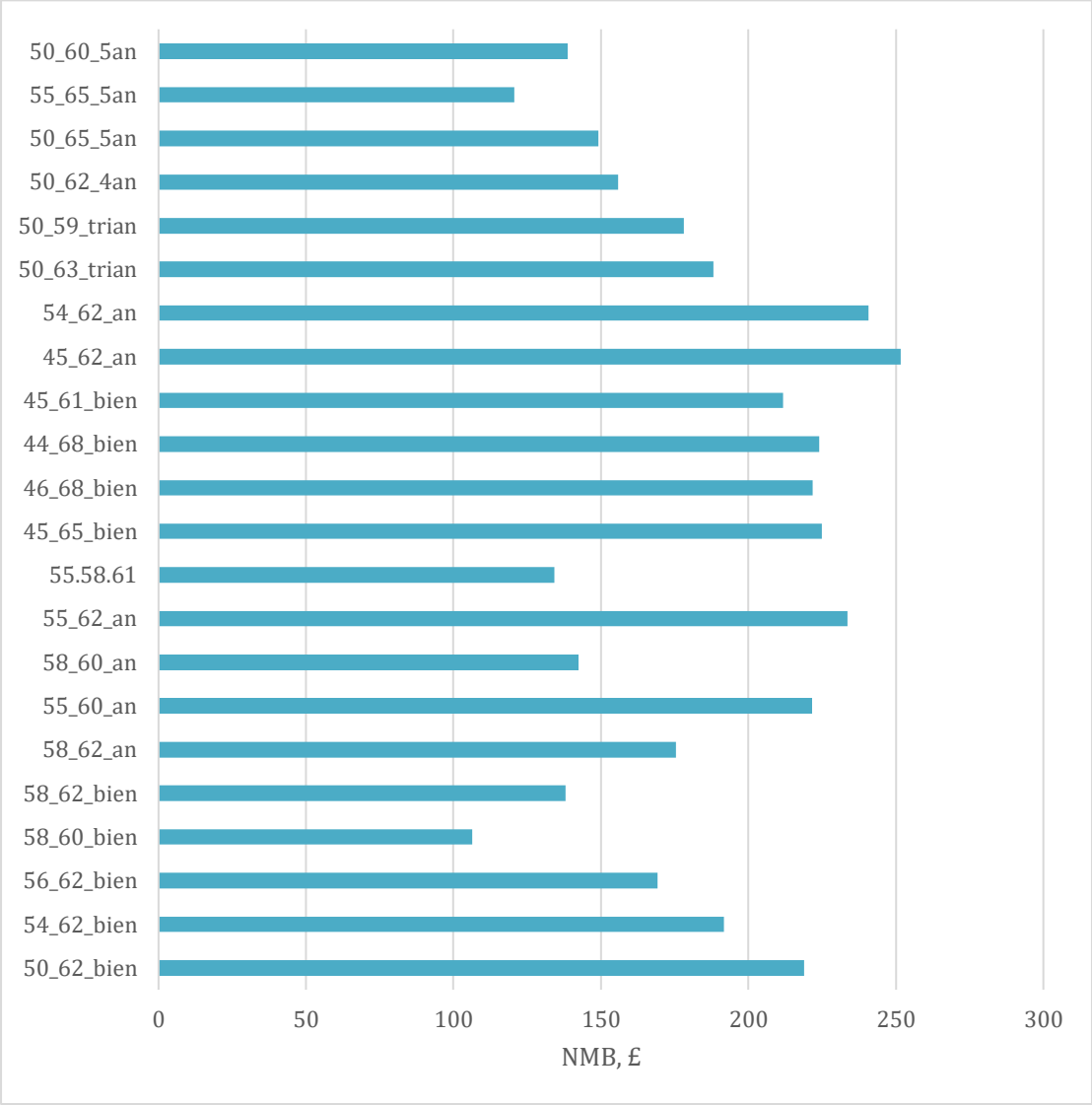
Figure 73: Overdiagnosis with one-time screening in BRCA carriers (increased incidence divided by the number of screen-detected cases)



8.4.2. BRCA carriers: repeat screening scenarios

Repeat screening was evaluated by modelling various interventions with age range of 45 to 68 years informed by the single-screening analyses. Annual screening of BRCA carriers from ages 45 to 62 emerged as the strategy with highest net monetary benefit (Figure 74). All repeat screening scenarios were also associated with high estimated rates of overdiagnosis (Figure 75).

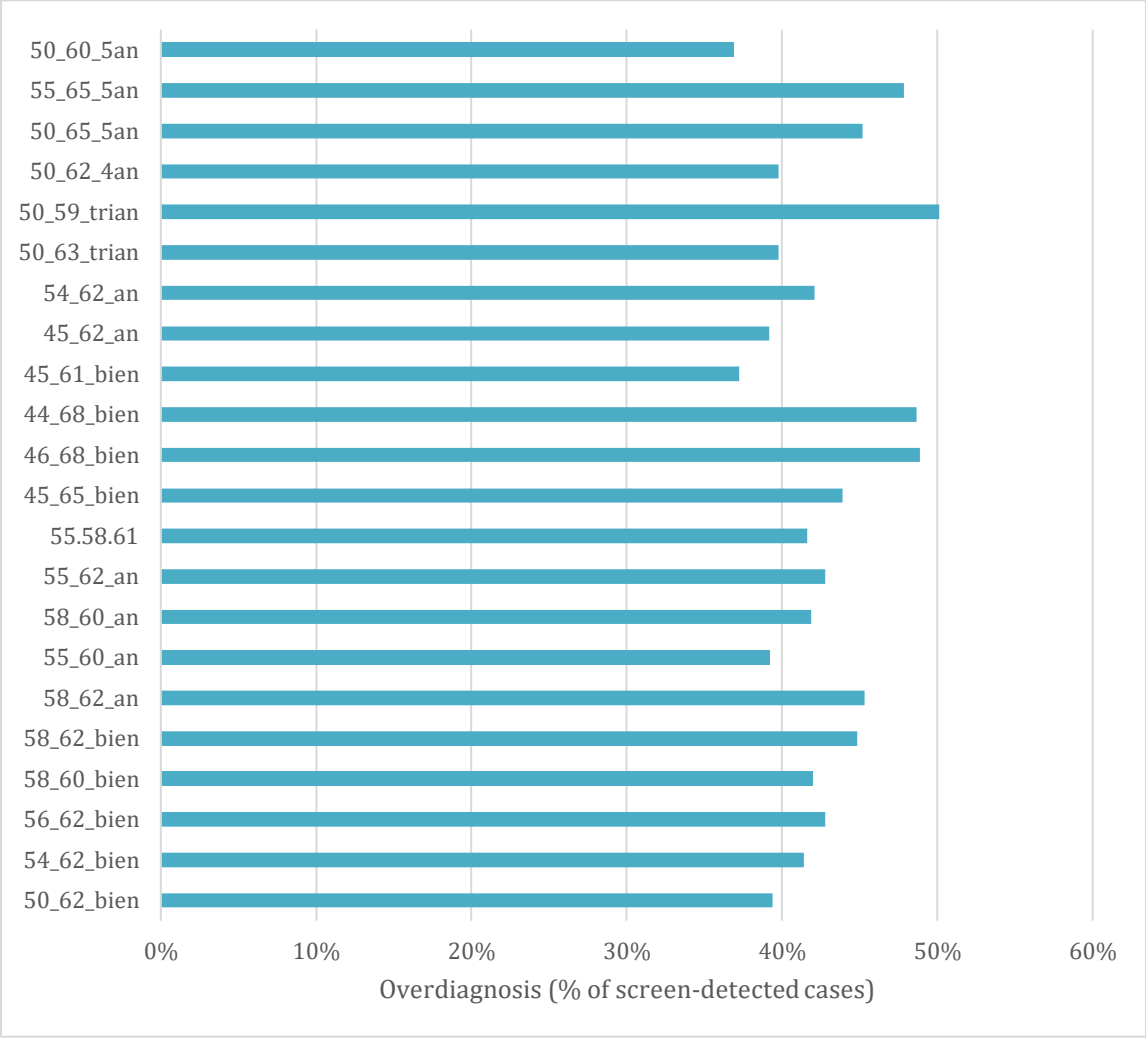
Figure 74: Incremental net monetary benefit with repeat screening in BRCA carriers vs standard of care



Legend: NMB – net monetary benefit

Screening every 5 years at ages 50–60 (screen_50_60_5an); Screening every 5 years at ages 55–65 (screen_55_65_5an); Screening every 5 years at ages 50–65 (screen_50_65_5an); Screening every 4 years at ages 50–62 (screen_50_62_4an); Triennial screening at ages 50–59 (screen_50_59_3an); Triennial screening at ages 50–63 (screen_50_63_trian); Annual screening at ages 54–62 (screen_54_62_an); Annual screening at ages 45–62 (screen_45_62_an); Biennial screening at ages 45–61 (screen_45_61_bien); Biennial screening at ages 44–61 (screen_45_68_bien); Biennial screening at ages 46–66 (screen_46_68_bien); Biennial screening at ages 45–65 (screen_45_65_bien); Triennial screening at ages 55–61 (screen_55.58.61); Annual screening at ages 55–62 (screen_55_62_an); Biennial screening at ages 58–60 (screen_58_60_bien); Annual screening at ages 55–60 (screen_55_60_an); Biennial screening at ages 58–62 (screen_58_62_bien); Annual screening at ages 58–60 (screen_58_60_an); Biennial screening at ages 56–62 (screen_56_62_bien); Biennial screening at ages 54–62 (screen_54_62_bien); Biennial screening at ages 50–62 (screen_50_62_bien).

Figure 75: Overdiagnosis rate with repeat screening in BRCA carriers (increased incidence divided by the number of screen-detected cases)



Legend: Screening every 5 years at ages 50–60 (screen_50_60_5an); Screening every 5 years at ages 55–65 (screen_55_65_5an); Screening every 5 years at ages 50–65 (screen_50_65_5an); Screening every 4 years at ages 50–62 (screen_50_62_4an); Triennial screening at ages 50–59 (screen_50_59_3an); Triennial screening at ages 50–63 (screen_50_63_trian); Annual screening at ages 54–62 (screen_54_62_an); Annual screening at ages 45–62 (screen_45_62_an); Biennial screening at ages 45–61 (screen_45_61_bien); Biennial screening at ages 44–61 (screen_45_68_bien); Biennial screening at ages 46–66 (screen_46_68_bien); Biennial screening at ages 45–65 (screen_45_65_bien); Triennial screening at ages 55–61 (screen_55.58.61); Annual screening at ages 55–62 (screen_55_62_an); Biennial screening at ages 58–60 (screen_58_60_bien); Annual screening at ages 55–60 (screen_55_60_an); Biennial screening at ages 58–62 (screen_58_62_bien); Annual screening at ages 58–60 (screen_58_60_an); Biennial screening at ages 56–62 (screen_56_62_bien); Biennial screening at ages 54–62 (screen_54_62_bien); Biennial screening at ages 50–62 (screen_50_62_bien).

8.4.3. BRCA carriers: probabilistic analyses

Probabilistic analyses were initially conducted assuming perfect uptake to assess the net impact of screening while accounting for parameter uncertainty.

The scenarios analysed included single screening at ages 45, 48, 50, and 55, as well as repeat screening at ages 46–62 and 50–62 with two-year intervals.

All single-screening scenarios produced positive incremental NMB values (Figure 76). Similarly, all repeat-screening scenarios showed incremental NMB values above zero (Figure 77).

As shown on the cost-effectiveness plane, both single and repeat screening scenarios with perfect uptake had a 100% probability of being cost effective for BRCA carriers (with previously detected genetic mutations) at a threshold of £20,000 per QALY (Figure 78). The mean incremental NMBs (95% CrI) for single screening at ages 45, 48, 50, and 55 were £125 (63; 200), £176 (87;282), £113 (59; 181), and £209 (102; 346), respectively. For repeat screening every four years, the mean incremental NMBs (95% CrI) were £324 (56–584) for ages 46–62 and £297 (65–532) for ages 50–62.

Figure 76: Incremental net monetary benefit (£) vs standard of care (BRCA carriers, perfect uptake, single screening)

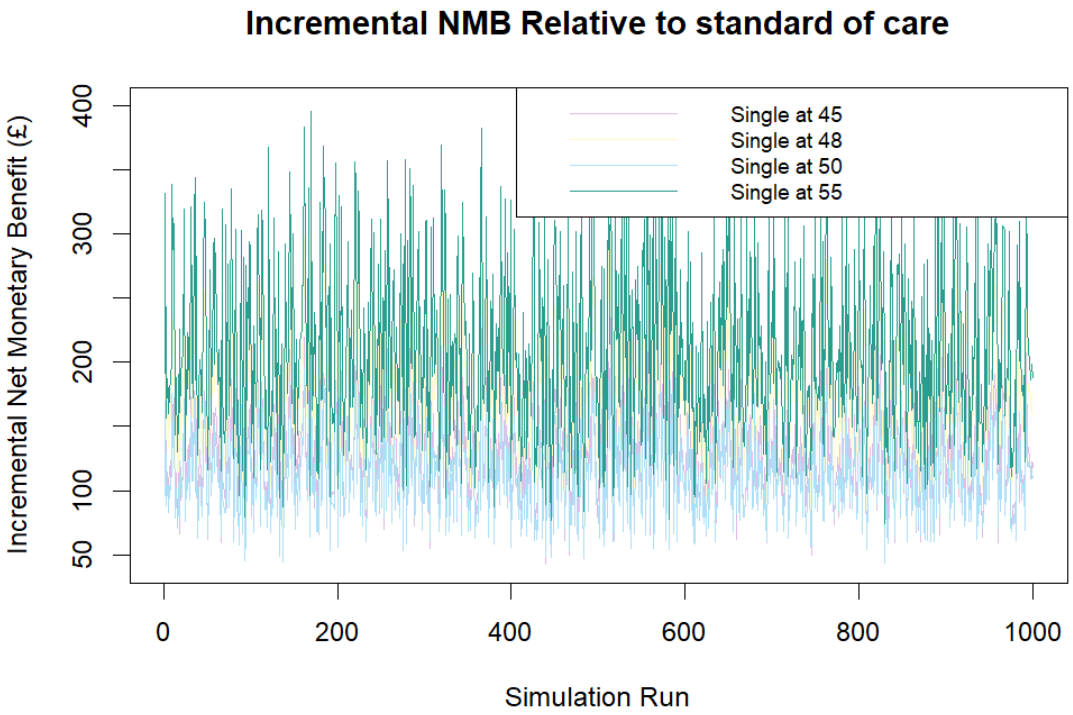


Figure 77: Incremental net monetary benefit (£) vs standard of care (BRCA carriers, perfect uptake, repeat screening)

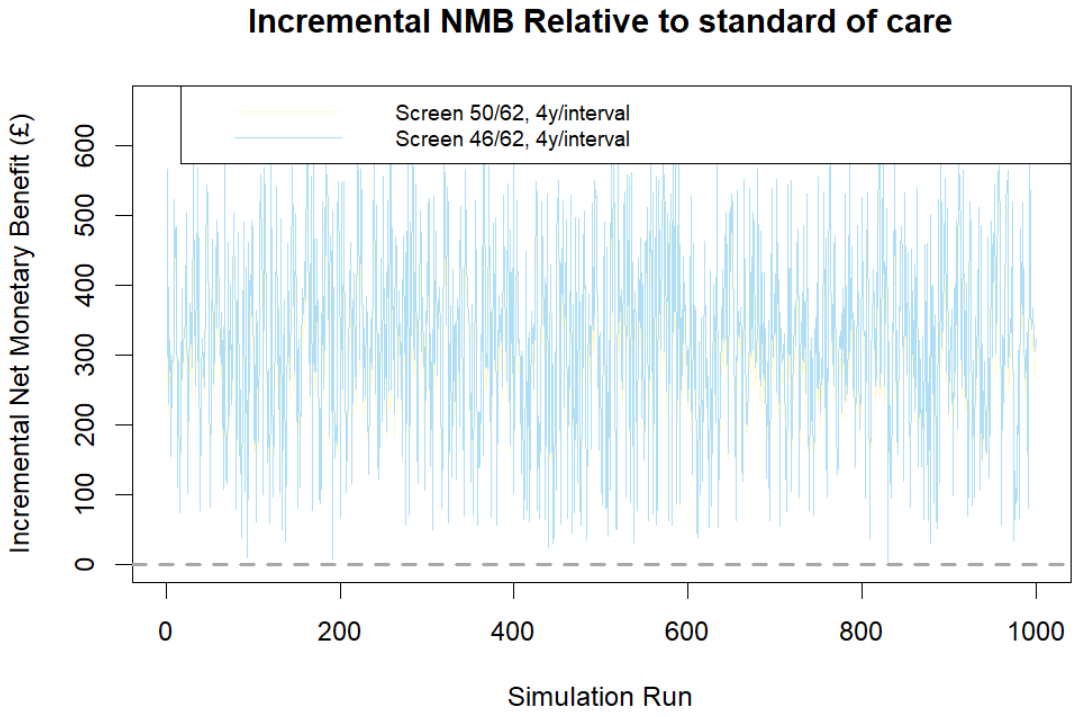
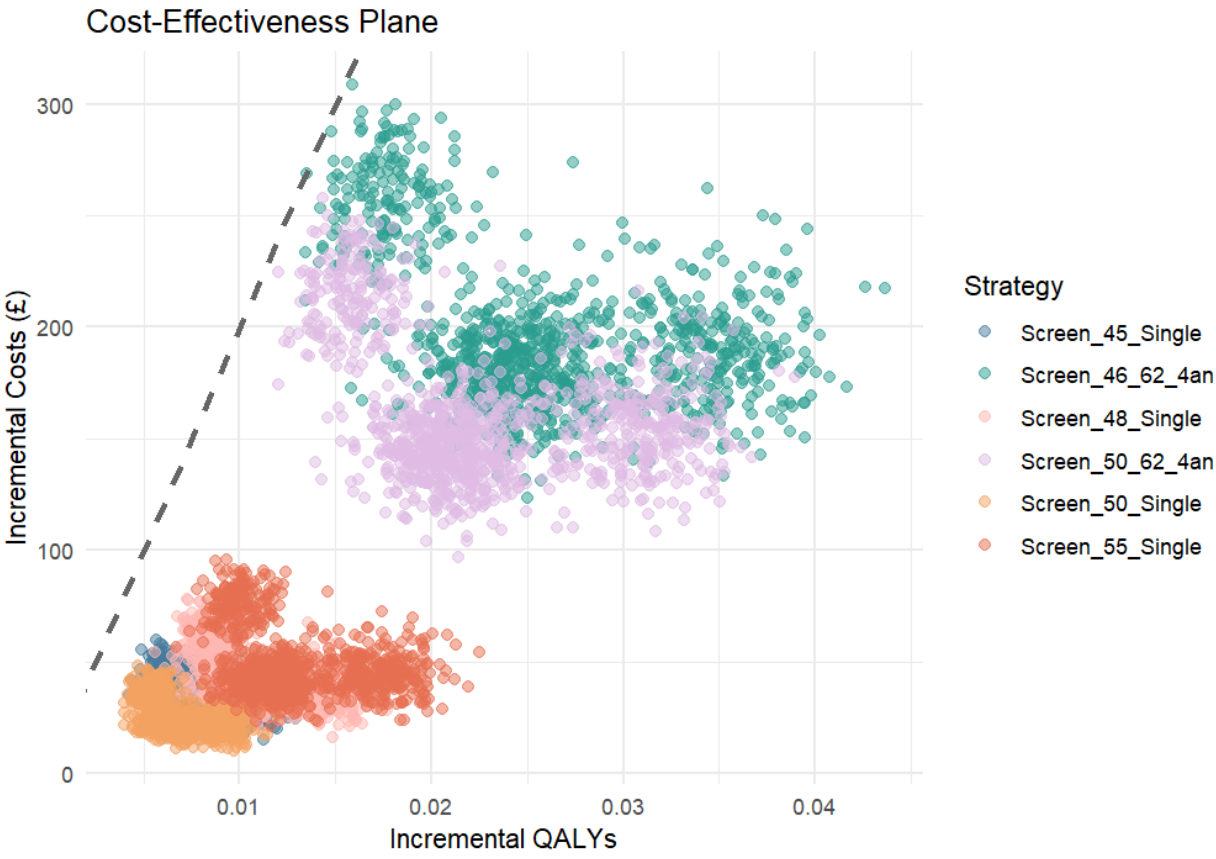
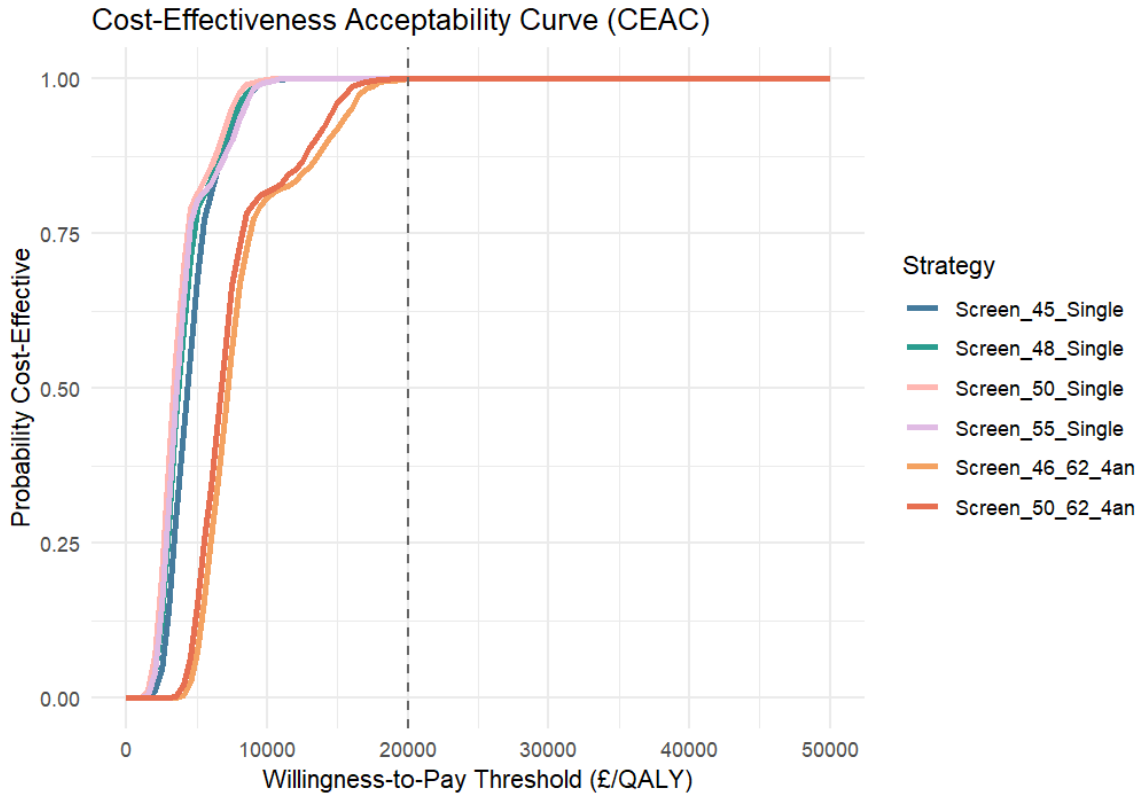


Figure 78: Cost-effectiveness plane (a) and cost-effectiveness acceptability curve (b) (BRCA carriers, perfect uptake)

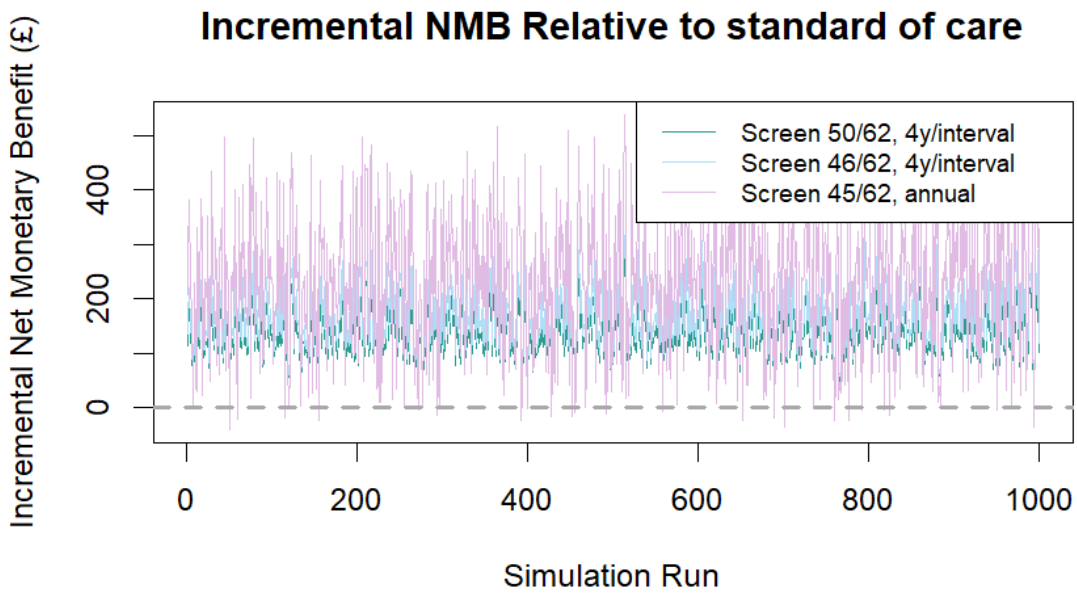




As with PSA screening in the Black population, the cost effectiveness of screening with perfect uptake for BRCA carriers does not provide sufficient confidence regarding its potential under realistic uptake levels. Therefore, we also conducted probabilistic evaluations of repeat screening for prostate cancer among BRCA carriers using uptake inputs from the CAP trial. The scenarios assessed included screening every four years from ages 46–62 and 50–62, as well as the most cost-effective scenario identified in deterministic runs—annual screening from ages 45 to 62.

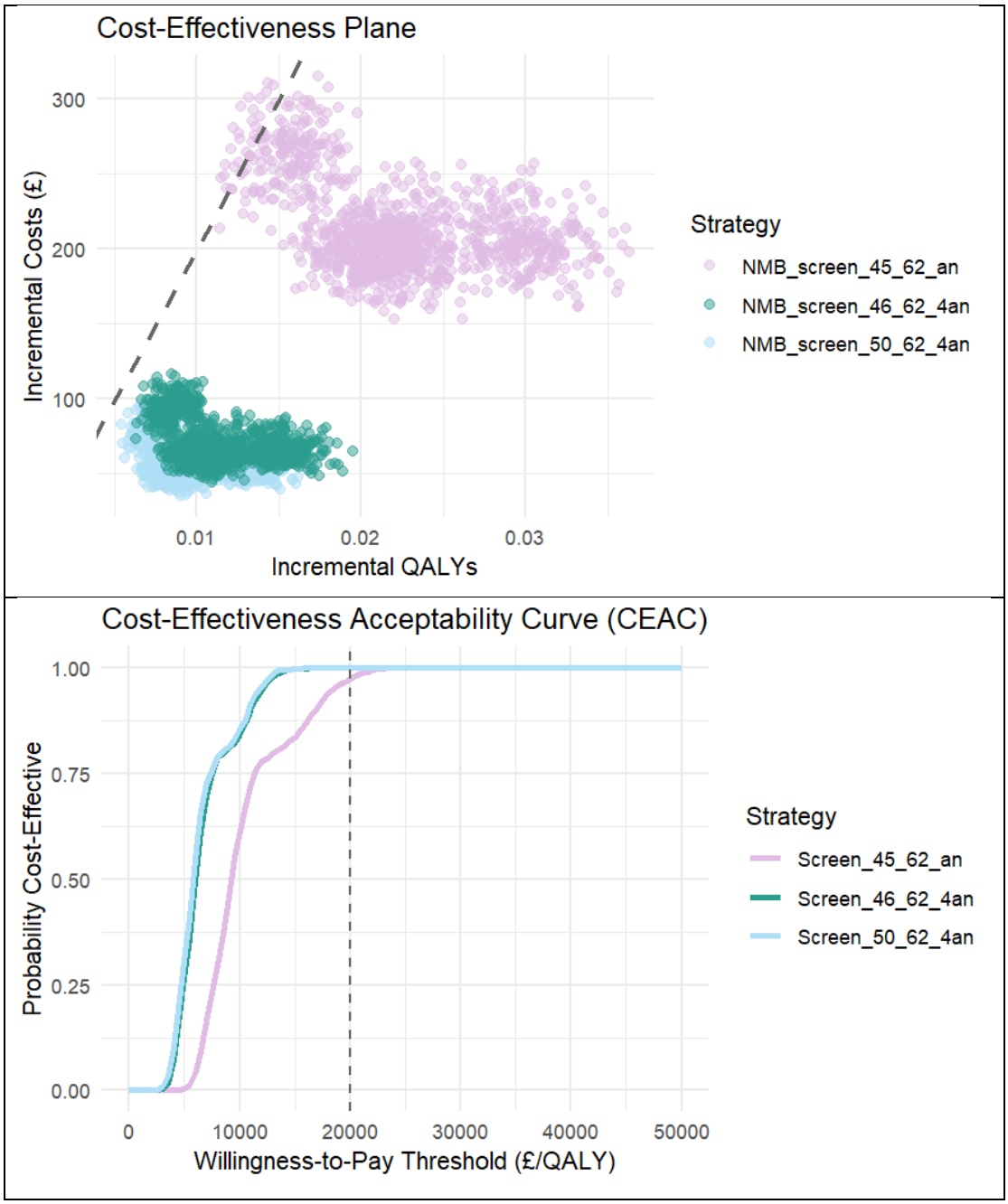
Applying uptake from the CAP trial in the probabilistic analysis did not change the conclusion on cost-effectiveness of PSA screening in men with established BRCA status (Figure 79).

Figure 79: Incremental net monetary benefit (£) compared to the standard of care (BRCA carriers, CAP uptake)



All repeat screening scenarios were cost-effective; however, screening at a 4-year interval demonstrated a higher probability of cost-effectiveness (100%) compared to annual screening (97.5%). The mean NMB and 95% CrI were £136 (£48; £240) for screening every four years between ages 50–62, £157 (£53; £274) for screening every four years between ages 46–62, and £234 (£2; £466) for annual screening between ages 45–62 (Figure 80).

Figure 80: Cost-effectiveness plane (a) and cost-effectiveness acceptability curve (b) (BRCA carriers, perfect uptake)



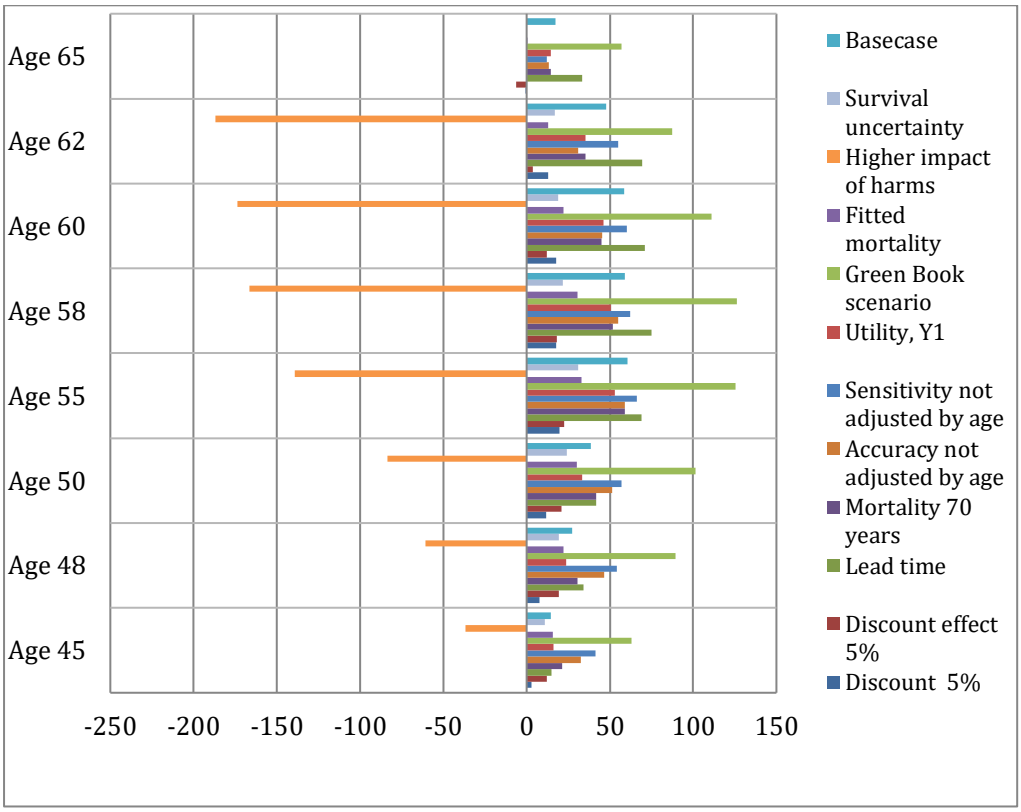
8.4.4. BRCA carriers: scenario analyses

As in other populations, applying higher impact of harms (this is however a hypothetical scenario to demonstrate an impact of harms on the model and this scenario is not supported by data) and the Green Book approach to discounting costs and effects had the greatest impact on incremental NMB estimates. However, in contrast to other population subgroups the scenario that prevents the

simulated population from dying from cancer diagnosed through organised screening before reaching their age of diagnosis in the comparator arm (standard of care), i.e. called here "lead time scenario", also had substantial impact on predicted NMB, especially in single-time screening runs (Figure 81, Figure 82). This happens presumably because cancer developed in BRCA carriers was assumed to be more aggressive in the model.

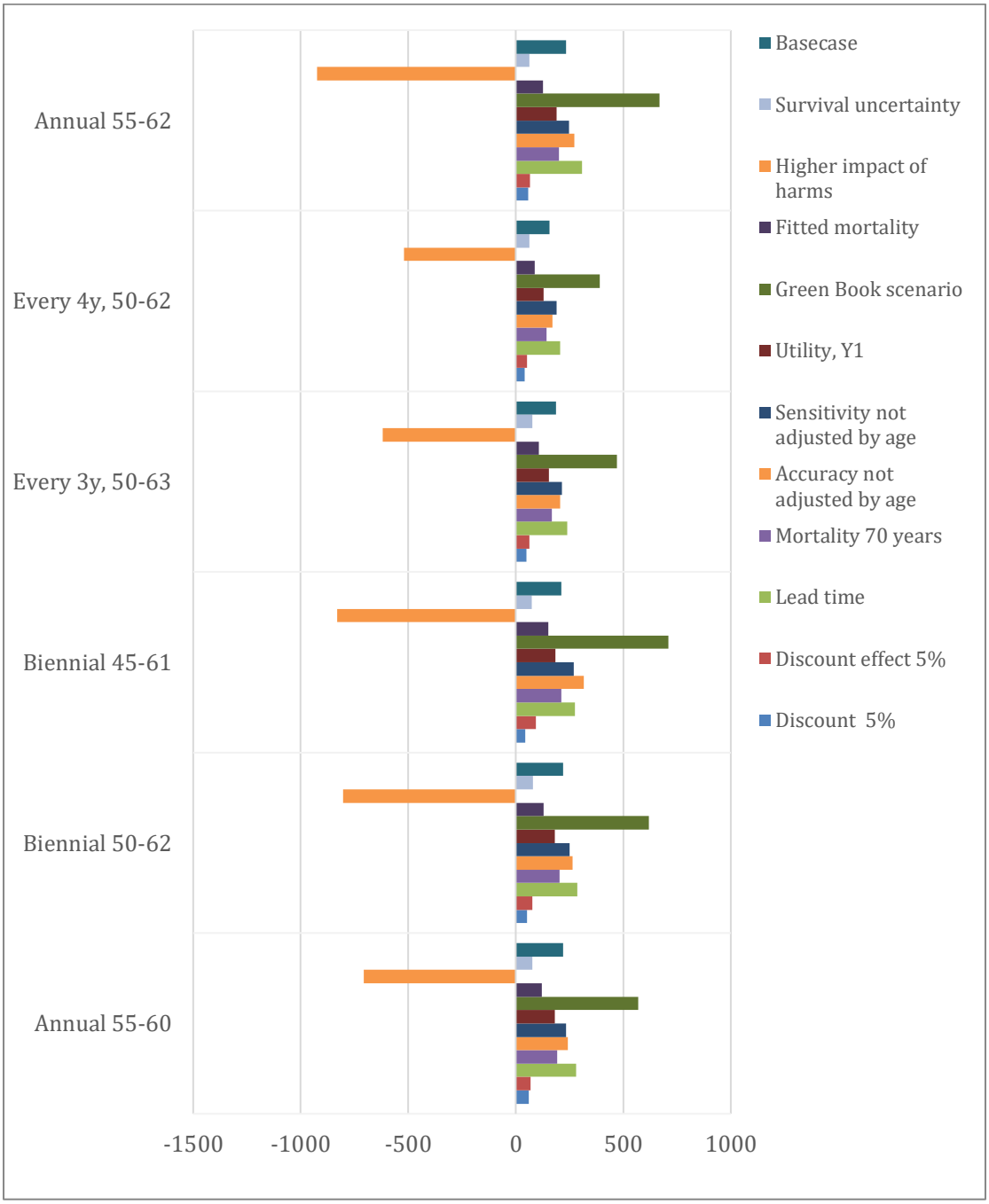
In single-screening scenarios for those who were screened at age 62 years and younger and in all repeat screenings, all except one run screening scenarios resulted in positive incremental NMB.

Figure 81: Net monetary benefit (£) in scenario analyses for one-time screening: BRCA carriers



Legend: £15,000 threshold is used in the Green Book scenario and £20,000 in the other scenarios.

Figure 82: Net monetary benefit (£) in scenario analyses for repeat screening: BRCA carriers



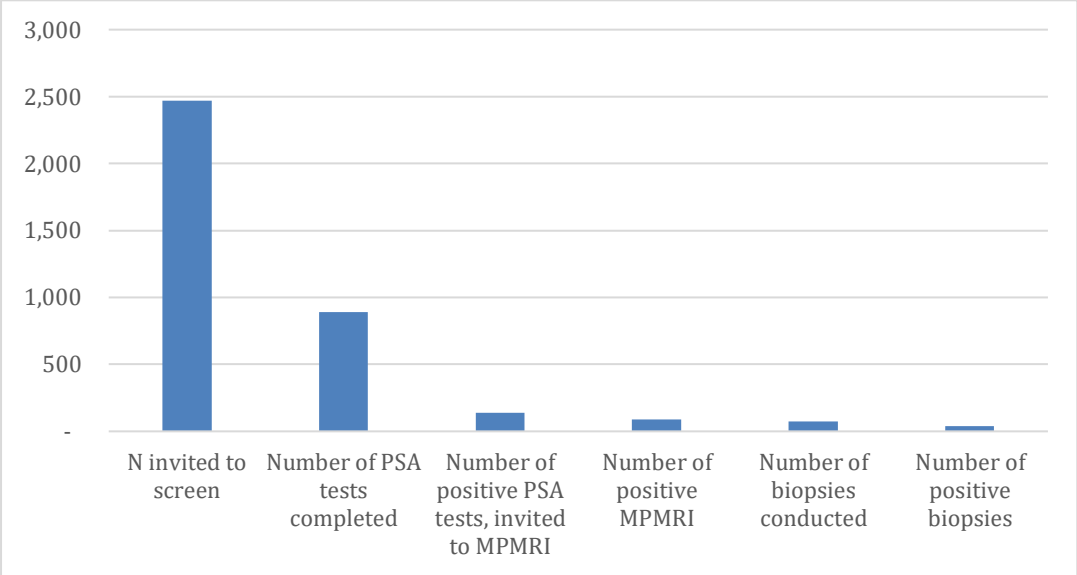
Legend: £15,000 threshold is used in the Green Book scenario and £20,000 in the other scenarios.

8.4.5. BRCA carriers: impact of screening scaled to the population of England

Biennial screening of BRCA carriers from 45 to 61 years was cost-effective in both deterministic and probabilistic scenarios. While annual screening was very cost-effective in deterministic analysis it would likely incur substantial administrative burden and could negatively impact uptake—factors not accounted for in the current model due to a lack of available data.

The resource use per cohort invited to be screened, was not substantial considering the small target population. The probability of a 45-year-old man being a BRCA carrier is only 0.00661. With a reported population of 1.9 million men aged 45–50 years in England, even if all were tested for BRCA status, only around 2,500 men would be eligible for screening. This would result in approximately 140 men invited for follow-up after a positive PSA test, 87 positive mpMRIs, and 64 biopsies performed each year, of which an estimated 41 would be cancer-positive per cohort screened (Figure 83).

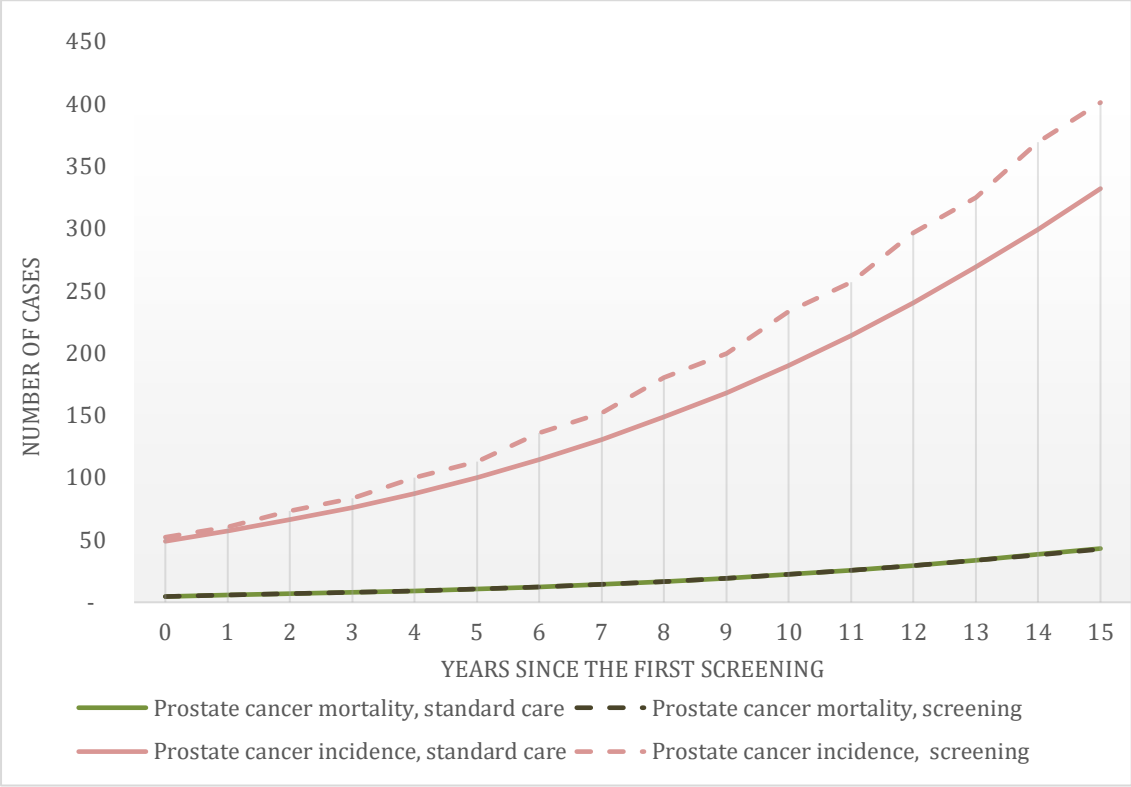
Figure 83: Resource requirement in England to screen biennially BRCA carriers from age 45 to 61



Over the 15 years following the implementation of an organised screening programme — screening BRCA carriers biennially from ages 45 to 61 resulted in an additional 69 prostate cancer cases expected to be detected in deterministic analyses. The difference in prostate cancer deaths in 15 years of follow up

between the organised screening for BRCA carriers and standard care arms is projected to result in 0.5 deaths per cohort screened or 4 deaths in a fully rolled out screening program (Figure 84).

Figure 84. Difference in prostate cancer incidence and mortality between the organised screening and standard care arms among a cohort of 45-year-old BRCA carriers invited to be screened biennially from ages 45 years to 50 years



9. Modelling results: risk-stratified screening

The following scenarios were evaluated for the risk-stratified screening:

Scenario 1: run for men of general risk and familial risk

Men are invited for screening at age 58. If diagnosed with cancer, they proceed with treatment and surveillance. If their first PSA test is >=3.5ng/ml (The NICE suggested threshold) but they test negative in the follow-up tests (MRI and biopsy) men are invited back at age 60 years. If their PSA is < 3.5ng/ml, they are not re-invited. The same scenario is conducted with the PSA test threshold of 3 ng/ml

Scenario 2: run for Black and BRCA carriers

Men are invited for screening at age 40. If diagnosed with cancer, they proceed with treatment and surveillance. If their first PSA test is >=2.5ng/ml but they test negative in the follow-up tests (MRI and biopsy) men are invited back biennially till age 58. If their PSA at the first test is < 2.5ng/ml, they are not re-invited. The same scenario is conducted with the PSA test threshold of 3 ng/ml.

Scenario 3: run for Black and BRCA carriers

Men are invited for screening at age 45. If diagnosed with cancer, they proceed with treatment and surveillance. If their first PSA test is >=2.5ng/ml but they test negative in the follow-up tests (MRI and biopsy) men are invited back biennially till age 61. If their PSA at the first test is < 2.5ng/ml, they are not re-invited. The same scenario is conducted with the PSA test threshold of 3 ng/ml.

Scenario 4: run for Black and BRCA carriers

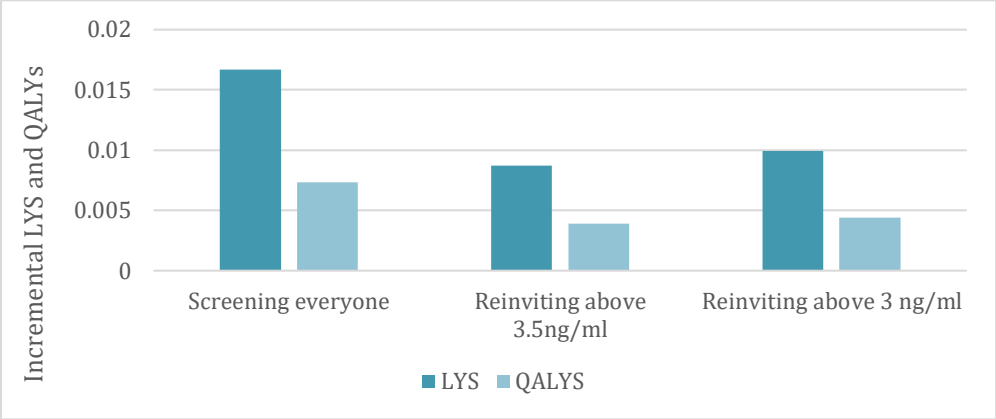
Men are invited for screening at age 45. If diagnosed with cancer, they proceed with treatment and surveillance. If their first PSA test is >=2.5ng/ml but they test negative in the follow-up tests (MRI and biopsy) men are invited back each 4 years up to age 61. If their PSA is < 2.5ng/ml, they are not re-invited. The same scenario is conducted with the PSA test threshold of 3 ng/ml.

To address the impact of discounting, all risk stratified strategies were modelled with the discounting applied in the first cycle of the intervention.

9.1. Risk stratified screening for men of average risk

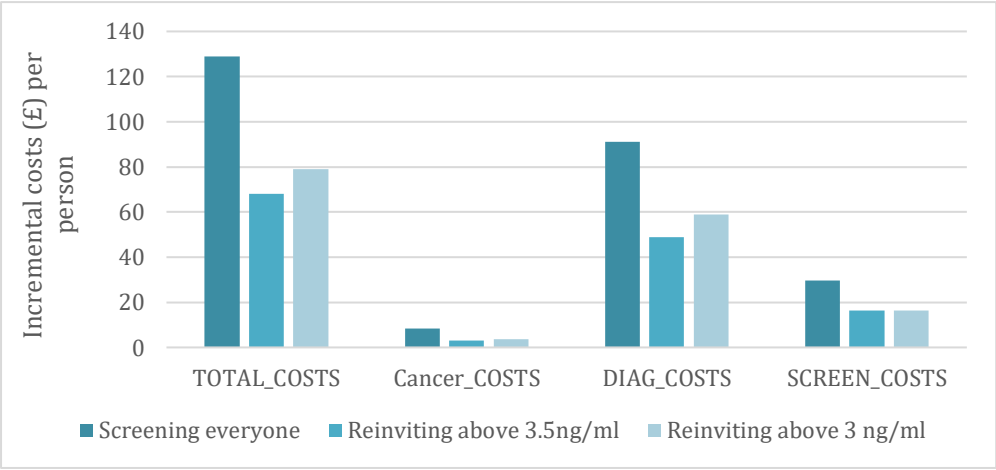
In the deterministic analysis, screening all men at ages 58 and 60 resulted in greater LYS and QALYs than reinviting at age 60 only those with PSA levels at age 58 above specific thresholds (either the NICE-recommended threshold of 3.5 ng/ml for this age group or 3 ng/ml as used in non-risk-stratified screening). However, the non-risk-stratified strategy also led to substantially higher screening, diagnostic, and treatment costs (Figures 85 and Figure 86). The rate of overdiagnosis predicted with these strategies was 49% if everyone is invited to be screened and 47% if only men with elevated PSA levels are invited the second time.

Figure 85: Incremental LYS and QALYs per person with screening general population at ages 58 and 60 vs standard of care



Legend: LYS – life years saved; QALYS -quality adjusted life years

Figure 86: Incremental costs per person with screening general population at ages 58 and 60 vs standard of care



Screening everyone at both ages (58 and 60) had the highest incremental NMB when compared pairwise to standard of care, with an incremental NMB of £17 at a willingness-to-pay threshold of £20,000 per QALY. This was followed by reinviting men with PSA >3.5 ng/ml (£10) and PSA >3 ng/ml (£9). However, the full incremental deterministic analysis showed that reinviting with a threshold of 3ng/ml was dominated with a slight negative incremental NMB (–£1) relative to reinviting those with PSA >3.5 ng/ml. Reinviting men with PSA >3.5 ng/ml was identified as the most cost-effective option on the efficiency frontier, though the difference between the interventions was small.

In the probabilistic analysis, screening all men using a risk-stratified approach (i.e. reinviting at age 60 only those with elevated PSA levels at the initial screen but no prostate cancer diagnosis) was not more cost-effective compared to inviting all men at ages 58 and 60(Figure 87). The mean incremental NMB compared to the standard of care in probabilistic analysis was £14 (95% CrI: –£127 to £129) for screening all men, £9 (–£65 to £81) for re-inviting those with PSA ≥3.5 ng/ml (as recommended by NICE), and £8 (–£78 to £91) for re-inviting those with PSA ≥3 ng/ml. The uncertainty in cost-effectiveness was similar across all three strategies, with probabilities of being cost-effective ranging from 61% to 64% (Table 31). The results of the deterministic scenario analysis, using the discounting approach recommended in the Green Book, were broadly consistent with those observed in the base case (Table 32).

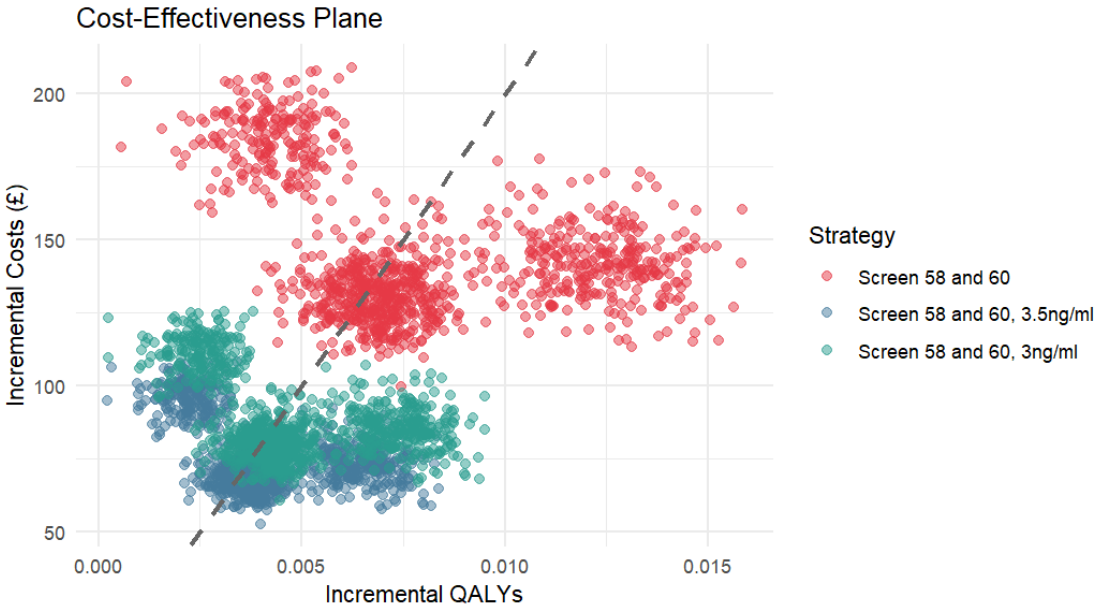
Table 31: Incremental net monetary benefit vs standard of care and other interventions, men of unknown risk

Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental deterministic analysis), £	NMB (Full incremental probabilistic analysis), mean £, p	NMB (standard of care), deterministic, £	NMB vs standard of care, probabilistic, mean £ (CrI), p
Standard of care	£ 8,157	44.871					
Reinviting above 3.5ng/ml	£ 8,225	44.875	£ 17,449	£ 10	£ 8.9, p= 64%	£ 10	£ 9 (-65;81), p=64%
Reinviting above 3 ng/ml	£ 8,236	44.876	£ 17,864	- £ 1	- £0.5, p= 49%	£ 9	£ 8 (-78;91), p=61%
Screening everyone	£ 8,286	44.879	£ 17,648	£ 7	£ 5.3, p= 60%	£ 17	£14 (-127;149), p=61%

Table 32: Incremental net monetary benefit vs standard of care and other interventions (deterministic using the Green Book scenario), men of unknown risk

Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental analysis), £	NMB (standard of care), £
Standard of care	£ 8,167	47.513			
Reinviting above 3.5ng/ml	£ 8,235	47.520	£ 9,328	£ 41	£ 41
Reinviting above 3 ng/ml	£ 8,246	47.521	£ 9,500	£ 4	£ 46
Screening everyone	£ 8,296	47.526	£ 9,338	£ 33	£ 78

Figure 87: Cost effectiveness plane for screening general population at ages 58 and 60 vs standard of care



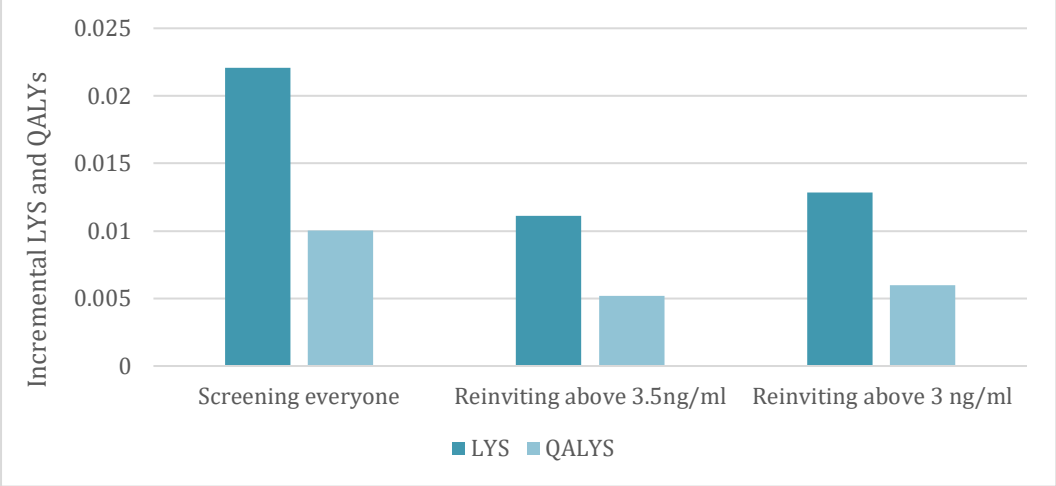
***Legend**: Screen 58 and 60 – All individuals are invited for screening at ages 50 and 58; Screen 58 and 60, 3ng/ml– Only those with PSA levels above 3 ng/ml at the first screen are reinvited at age 60; Screen 58 and 60, 3.5ng/ml – Only those with PSA levels above 3.5 ng/ml at the first screen, as recommended by NICE, are reinvited at age 60.*

9.2. Risk stratified screening for men with familial risk

Screening men with familial risk twice at ages 58 and 60 resulted in greater LYS and QALYs but also higher costs than reinviting at age 60 only those with PSA levels at age 58 above specific thresholds (either the NICE-recommended threshold of 3.5 ng/ml for this age group or 3 ng/ml as used in non-risk-stratified screening), see Figures 88 and 89. The rate of overdiagnosis predicted with these

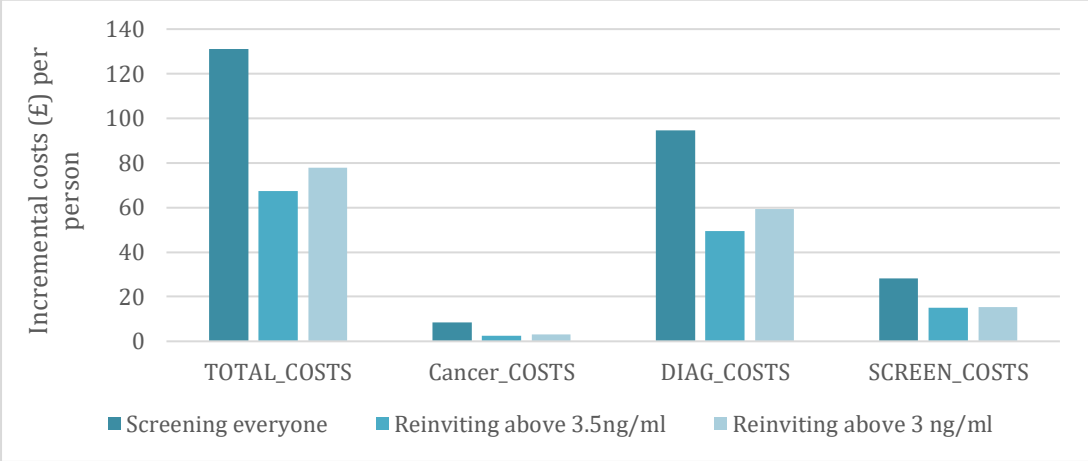
strategies was 48.5% if everyone is invited to be screened and 46% if only men with elevated PSA levels are invited the second time.

Figure 88: Incremental LYS and QALYs per person with screening men with familial risk at ages 58 and 60 vs standard of care



Legend: LYS – life years saved; QALYS -quality adjusted life years

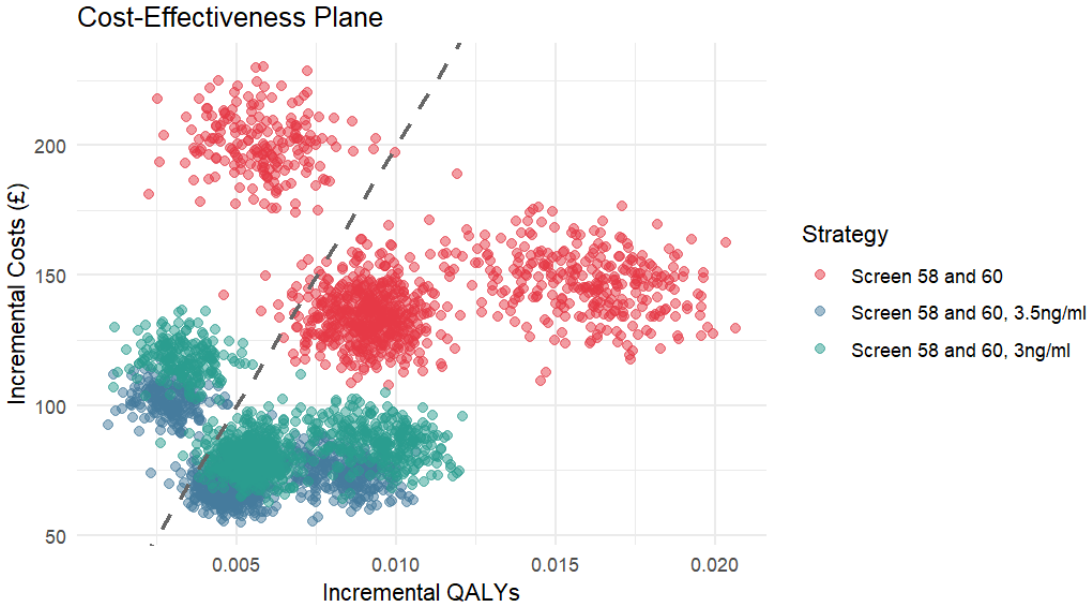
Figure 89: Incremental costs with screening men with familial risk at ages 58 and 60 (compared to the standard of care), per person



Screening men with familial risk at both ages had the highest incremental NMB when compared pairwise to standard of care in both deterministic and probabilistic analyses, with an incremental NMB of £70 (deterministic) and £59 (probabilistic, 95% CrI -£119; £230) at a willingness-to-pay threshold of £20,000 per QALY. This was followed by reinviting men with PSA >3 ng/ml (£41 in deterministic and £35, 95% CrI -£61; £120 in probabilistic) and PSA >3.5 ng/ml (£37 in deterministic and £32 95% CrI -£72; £137 in probabilistic).

In the full incremental analysis, screening all men with familial risk at ages 58 and 60 also emerged as the most cost-effective strategy in both deterministic and probabilistic analyses in the base case and Green Book scenarios (Table 33 and Table 34). The probability of organised screening interventions being cost-effective was higher for men with familial risk compared to those at general risk. However, the credible intervals were wide and included negative values, indicating substantial uncertainty (Figure 90).

Figure 90: Cost effectiveness plane for screening men with familial risk at ages 58 and 60 vs standard of care



***Legend**: Screen 58 and 60 – All individuals are invited for screening at ages 50 and 58; Screen 58 and 60, 3ng/ml– Only those with PSA levels above 3 ng/ml at the first screen are reinvited at age 60; Screen 58 and 60, 3.5ng/ml – Only those with PSA levels above 3.5 ng/ml at the first screen, as recommended by NICE, are reinvited at age 60.*

Table 33: Incremental net monetary benefit vs standard of care and other interventions, men with familial risk

Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental deterministic analysis), £	NMB (Full incremental probabilistic analysis), mean £, p	NMB (standard of care), deterministic, £	NMB vs standard of care, probabilistic, mean £ (CrI), p
Standard of care	£ 8,636	44.777					
Reinviting above 3.5ng/ml	£ 8,704	44.782	£ 12,944	£ 37	£ 32, p=78%	£37	£ 32 (-61; 120), p =79%
Reinviting above 3 ng/ml	£ 8,714	44.783	£ 13,080	£ 5	£ 3.5, p=71%	£41	£ 35 (-71;137), p=79%
Screening everyone	£ 8,767	44.787	£ 13,051	£ 29	£ 24, p=78%	£70	£ 59 (-120; 230), p=79%

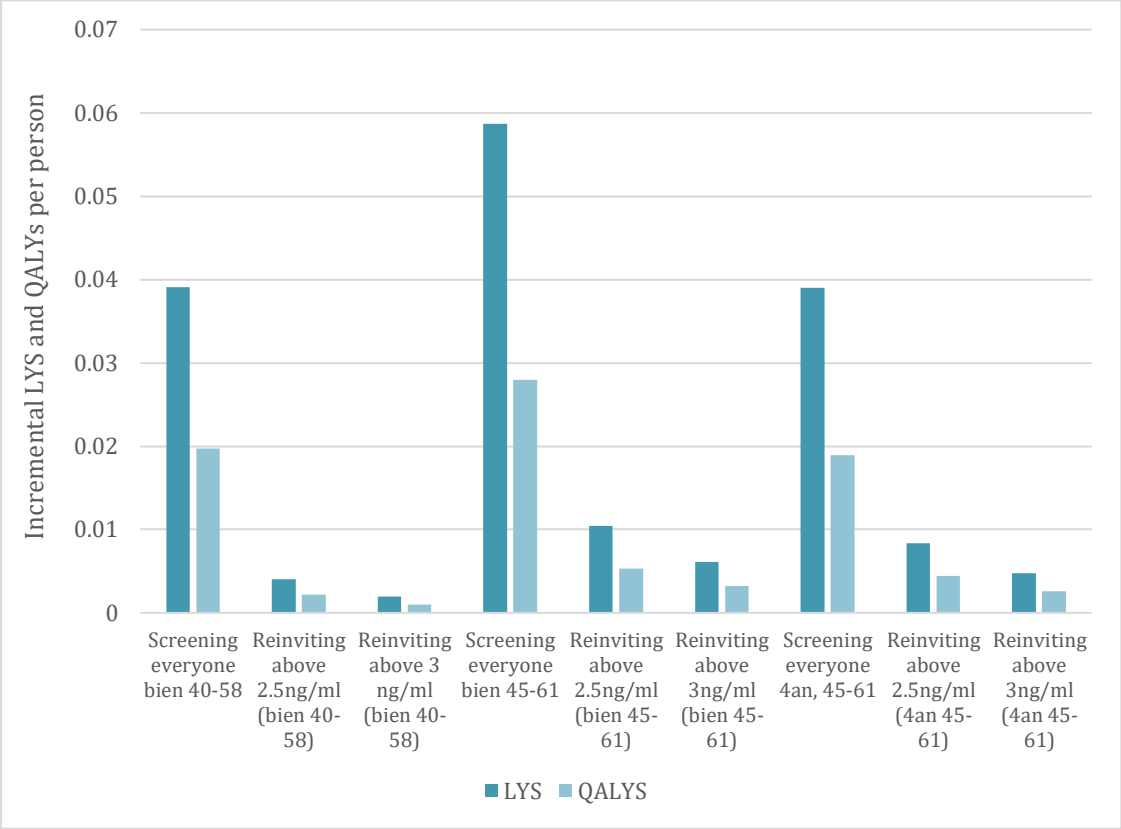
Table 34: Incremental net monetary benefit vs standard of care and other interventions (deterministic using the Green Book scenario), men with familial risk

Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental analysis), £	NMB (standard of care) , £
Standard of care	£ 8,646	47.377			
Reinviting above 3.5ng/ml	£ 8,713	47.387	£ 7,020	£ 77	£ 77
Reinviting above 3 ng/ml	£ 8,724	47.388	£ 7,082	£ 11	£ 87
Screening everyone	£ 8,777	47.396	£ 7,019	£ 62	£ 149

9.3. Risk stratified screening for men of Black ethnicity

In the deterministic analysis, screening men of Black ethnicity led to greater LYS and QALYs across all scenarios, with and without risk stratification. Expectedly, the benefit was more pronounced in strategies involving multiple rounds of screening per individual—i.e. when all men were re-invited regardless of PSA level—compared to strategies that only re-invited those with elevated PSA results from the initial round (Figure 91).

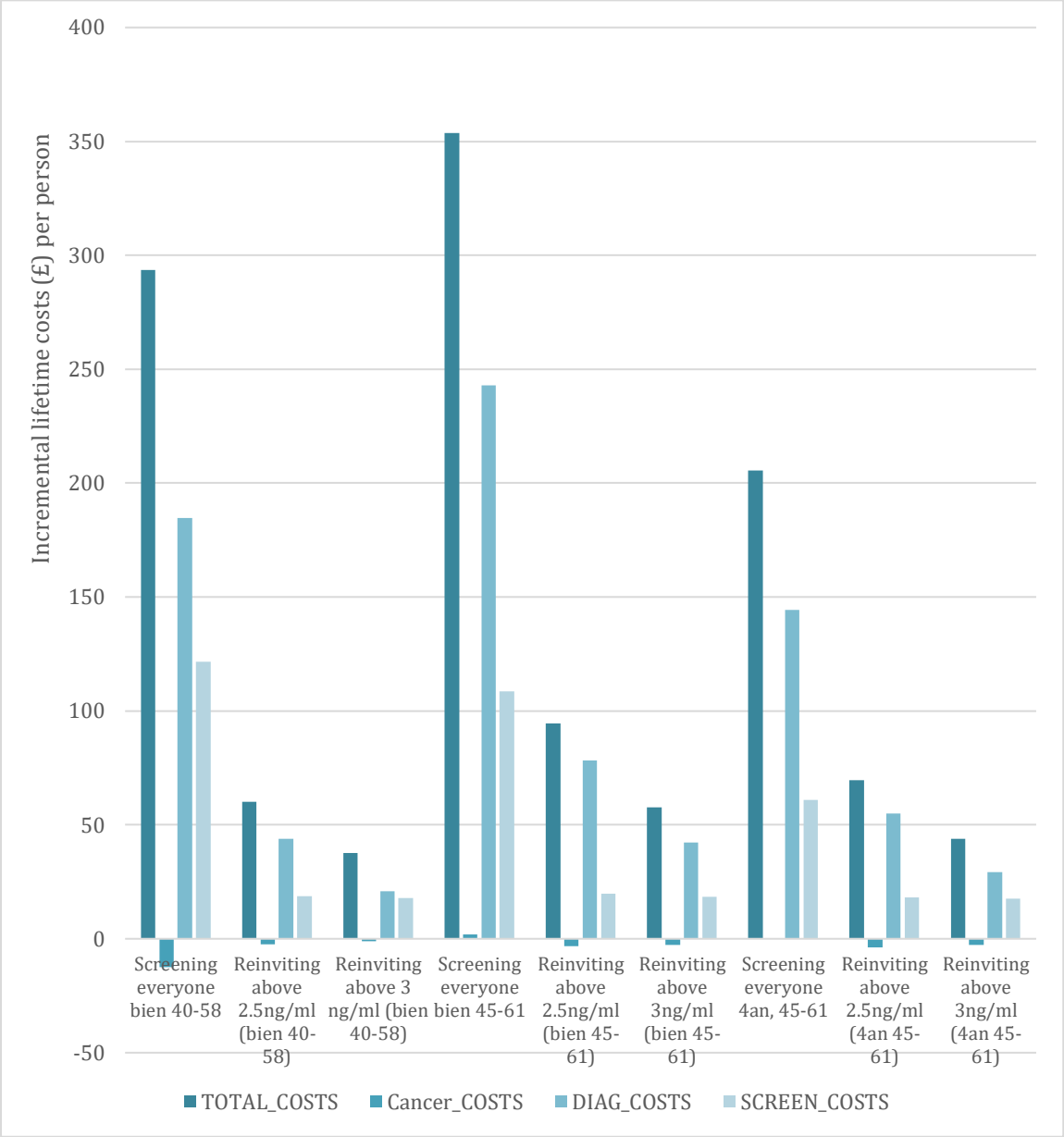
Figure 91: Incremental LYS and QALYs per person with screening men of Black ethnicity vs standard of care



Legend: LYS – life years saved; QALYS -quality adjusted life years

Screening also resulted in higher overall costs. Although earlier screening (i.e. starting at age 40) reduced treatment costs in some scenarios, total costs increased due to the added costs of screening and diagnostic procedures (Figure 92).

Figure 92: Incremental costs per person with screening men of Black ethnicity vs standard of care



Among scenarios where screening started at age 40, screening everyone was the most cost-effective strategy in both deterministic and probabilistic pairwise and full incremental analyses in the base case and Green book analyses (Table 35 and Table 36). A single screen at age 40 followed by selective re-invitation based on PSA levels was not cost-effective in deterministic analysis. In the probabilistic full incremental analysis, all comparators had a 75–80% probability of being cost-effective at the NICE threshold (Figure 93). Using a lower PSA threshold of 2.5 ng/ml led to slightly lower costs compared to 3.5 ng/ml, but this did not alter

the conclusion that screening everyone remained the most cost-effective option. As in other population subgroups, the differences between strategies were small.

For deterministic scenarios initiating screening at age 45, all strategies were cost-effective compared to standard care (Table 37). However, biennial screening of all men from ages 45 to 61 was the most cost-effective strategy on the efficiency frontier in both base case and Green Book (Table 38) scenarios.

In probabilistic analyses for screening strategies starting from the age of 45, strategies where only men with high PSA levels were invited to the repeat screening, after their first test at age 45 was positive, had negative or low incremental NMB (Figure 94). Screening once in two or four years from age 45 to 61 had positive incremental NMB, though with wide CrI with around 20% of probabilistic runs resulting in negative NMB.

Biennial screening from age 45 to 61 emerged as the most cost-effective strategy for Black men in probabilistic analysis. However, the confidence that it is more cost-effective than quadrennial screening (every four years) for the same age group is low.

Figure 93: Cost effectiveness plane for screening men of Black ethnicity starting at age 40 vs standard of care

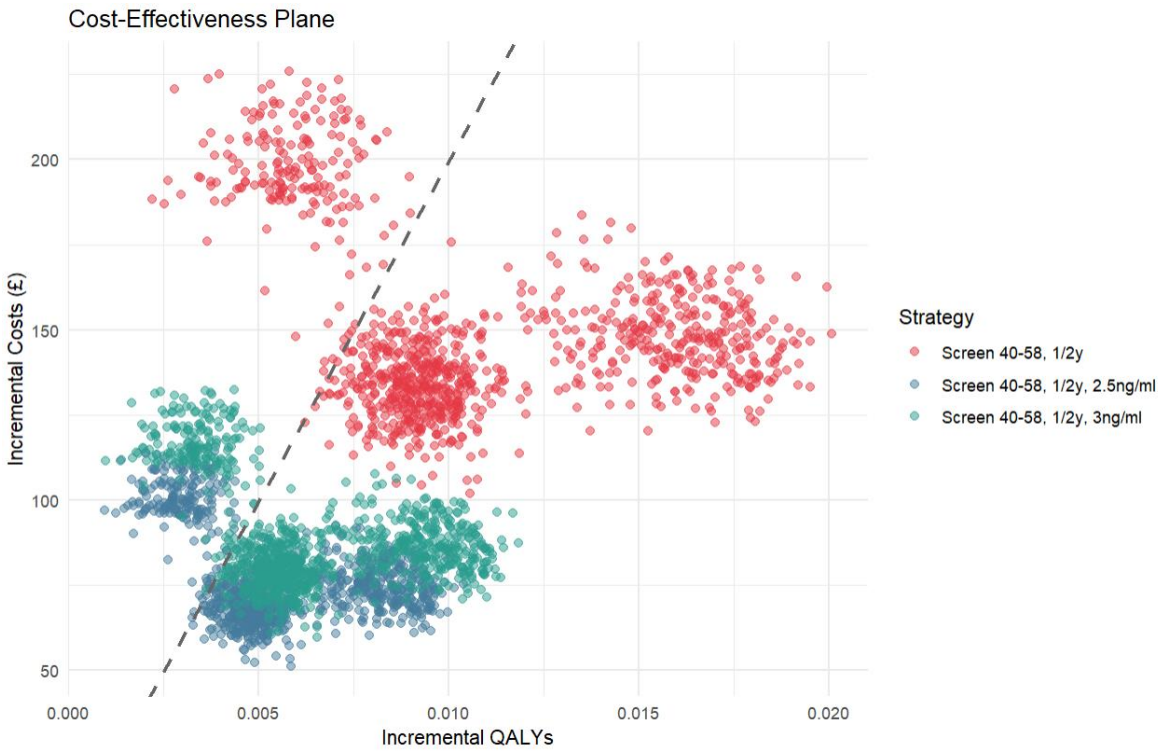


Figure 94: Cost effectiveness plane for screening men of Black ethnicity starting at age 45 vs standard of care

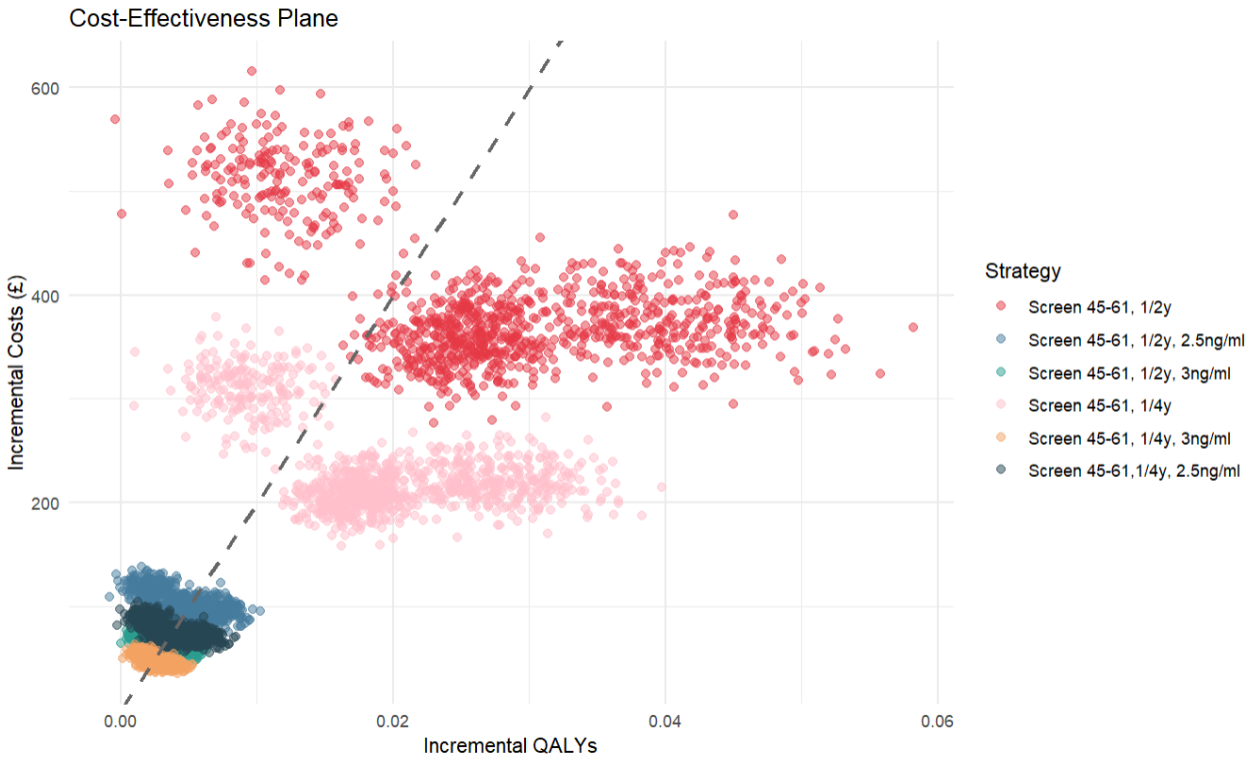


Table 35: Incremental net monetary benefit vs standard of care and other interventions with age for risk stratification 40, men of Black ethnicity

Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental deterministic analysis), £	NMB (vs standard of care), deterministic, £	NMB mean and 95%CrI (vs standard of care), probabilistic, £
Standard of care	£ 5,683	36.176				
Reinviting above 3 ng/ml	£ 5,721	36.177	£ 36,636	-£ 17.09	-£17.09	40 (-70; 136), p=81%
Reinviting above 2.5ng/ml	£ 5,743	36.178	£ 27,589	-£ 17	-£17	36 (-59; 118), p=81%
Screening everyone	£ 5,976	36.196	£ 14,906	£ 100	£ 100	67 (-113;229), p=81%

Table 36: Incremental net monetary benefit vs standard of care and other interventions with age for risk stratification 40 (deterministic with Green Book scenario), men of Black ethnicity

Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental analysis), £	NMB (vs standard of care), £
Standard of care	£ 5,806	42.544			
Reinviting above 3 ng/ml	£ 5,844	42.547	£ 13,932	£ 3	£ 3
Reinviting above 2.5ng/ml	£ 5,866	42.549	£ 11,288	£ 20	£ 20
Screening everyone	£ 6,097	42.594	£ 5,778	£ 465	£ 465

Table 37: Incremental net monetary benefit vs standard of care and other interventions with age for risk stratification 45, men of Black ethnicity

Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental probabilistic analysis), £	NMB (Full incremental deterministic analysis), mean £, p	NMB (standard of care), deterministic, £	NMB vs standard of care, probabilistic, mean £ (CrI), p
Standard of care	£ 6,678	38.987					
Reinviting above 3ng/ml (4an 45-61)	£ 6,719	39.005	£ 2,283	£ 7, p =67%	£ 315	£315	£ 7 (-53;45), p=67%
Reinviting above 3ng/ml (bien 45-61)	£ 6,736	38.990	£17,911	- £ 6, p =15%	-£309	£7	£ 0.7 (-53;47), p=54%
Reinviting above 2.5ng/ml (4an 45-61)	£ 6,748	38.992	£15,716	£ 5, p =72%	-£297	£19	£ 5 (-67;70), p=62%
Reinviting above 2.5ng/ml (bien 45-61)	£ 6,773	38.993	£17,894	- £ 12, p =10%	-£304	£11	£ -7 (-101;72), p=47%
Screening everyone 4an, 45-61	£ 6,884	39.006	£ 10,849	£ 158, p =83%	-£142	£173	£ 151 (-197;441), p=81%
Screening everyone bien 45-61	£ 7,032	39.015	£12,633	£ 4, p =62%	- £109	£206	£ 155 (-377; 583), p=79%

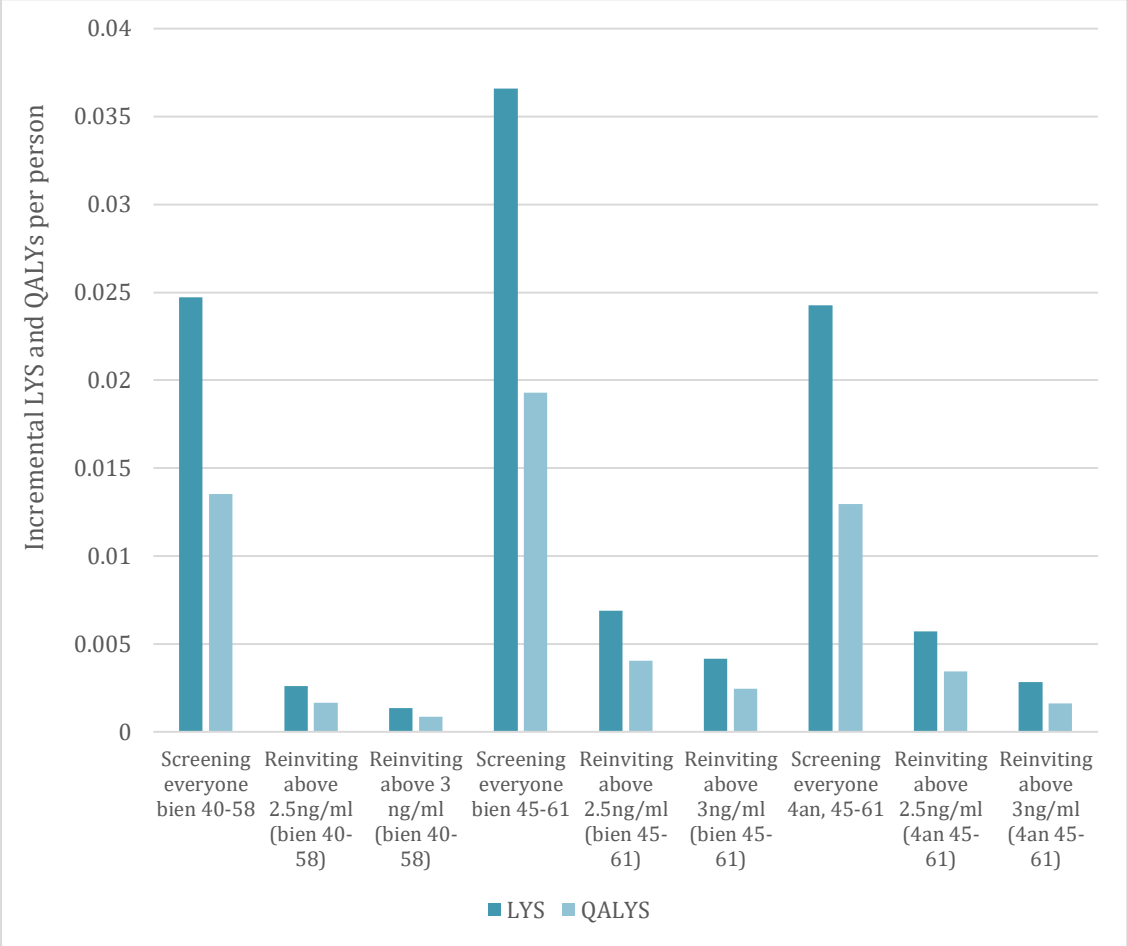
Table 38: Incremental net monetary benefit vs standard of care and other interventions with age for risk stratification 45 (deterministic with Green Book discounting), men of Black ethnicity

Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental analysis), £	NMB (vs standard of care), £
Standard of care	£ 6,759	44.141			
Reinviting above 3ng/ml (4an 45-61)	£ 6,800	44.162	£ 1,930	£ 275	£ 275
Reinviting above 3ng/ml (bien 45-61)	£ 6,817	44.148	£ 8,349	-£ 229	£ 46
Reinviting above 2.5ng/ml (4an 45-61)	£ 6,829	44.150	£ 7,325	-£ 202	£ 73
Reinviting above 2.5ng/ml (bien 45-61)	£ 6,854	44.152	£ 8,145	-£ 195	£ 80
Screening everyone 4an, 45-61	£ 6,965	44.185	£ 4,662	£ 183	£ 458
Screening everyone bien 45-61	£ 7,114	44.207	£ 5,387	£ 176	£ 634

9.4. Risk stratified screening for BRCA carriers

Screening of BRCA carriers led to greater LYS and QALYs across all scenarios compared to the standard care. Similar to the men of Black ethnicity, the benefit was more pronounced in strategies involving multiple rounds of screening per individual—i.e. when all men were re-invited regardless of PSA level—compared to strategies that only re-invited those with elevated PSA results from the initial round (Figure 95).

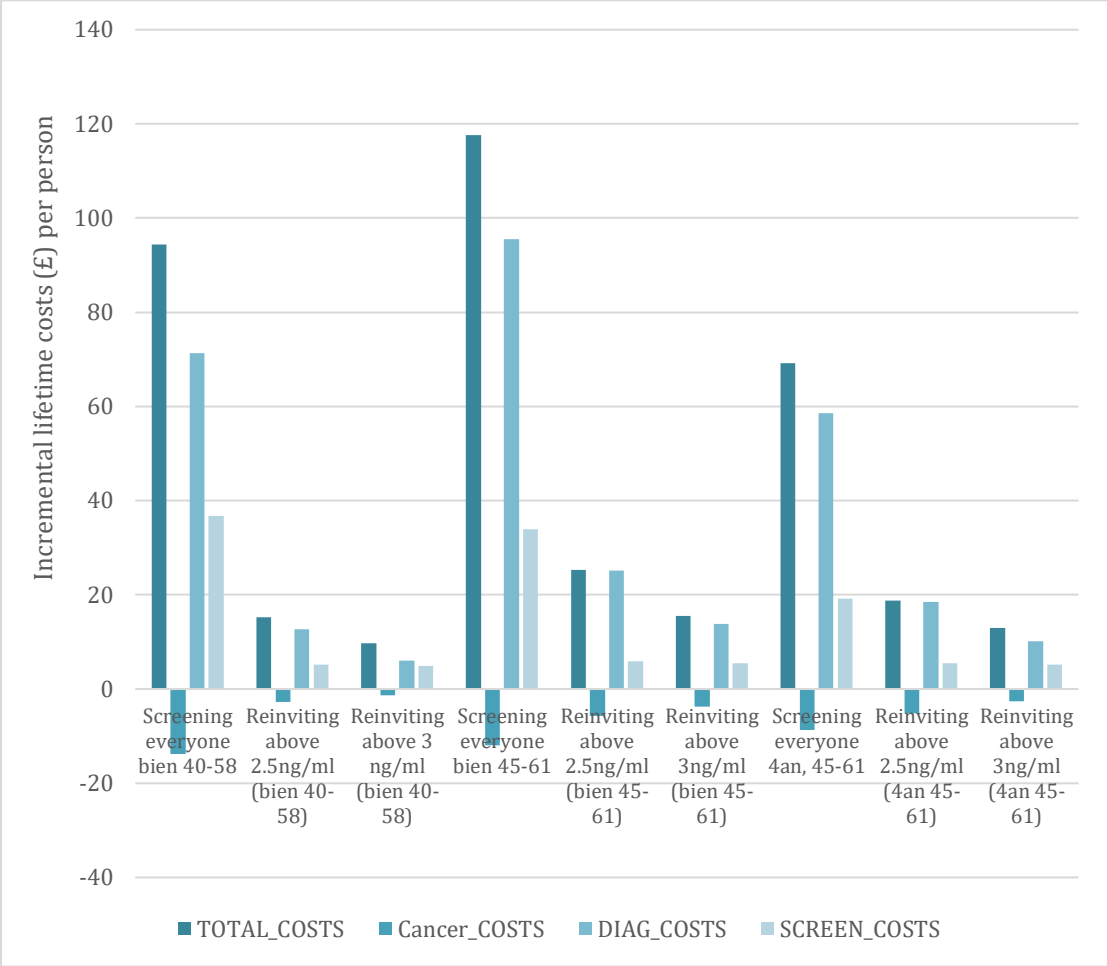
Figure 95: Incremental LYS and QALYs per person with screening general population at ages 58 and 60 vs standard of care



Legend: LYS – life years saved; QALYS -quality adjusted life years

Screening also resulted in higher costs in all scenarios, even though the treatment costs were lower in all evaluated screening programmes (Figure 96).

Figure 96: Incremental costs per person with screening general population at ages 58 and 60 vs standard of care



For scenarios that initiated screening at age 40, biennial screening from age 40 to 58 was the most cost-effective strategy in both deterministic and probabilistic analyses in the base case (Table 39) and Green Book (Table 40) scenarios. A less cost-effective approach was to screen at age 40 and then reinvite only those above a pre-defined PSA threshold (Figure 97).

In scenarios which started screening at age 45, all strategies were found to be cost-effective compared to standard care (Table 41, Figure 98). Similar to the findings for men of Black ethnicity, biennial screening for all men from ages 45 to 61 was the most cost-effective strategy on the efficiency frontier in both base case (Table 41) and Green Book scenario (Table 42).

Overall, for men with an established BRCA status, biennial screening from age 45 to 61 emerged as the most cost-effective strategy across all scenarios examined.

Table 39: Incremental net monetary benefit vs standard of care and other interventions with age for risk stratification 40, BRCA carriers

Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental deterministic analysis), £	NMB (Full incremental probabilistic analysis), mean £, p	NMB (standard of care), deterministic, £	NMB vs standard of care, probabilistic, mean £ (CrI), p
Standard of care	£ 6,892	35.728					
Screening at 40 and reinviting biennially till 58 if PSA above 3 ng/ml	£ 6,901	35.729	£11,046	£ 8	£ 4, p=78%	£8	£4 (-5; 15), p=78%
Screening at 40 and reinviting till 58 if PSA above 2.5ng/ml	£6,907	35.729	£9,156	£ 10	£ 11, p=96%	£18	15 (-2; 34), p=96%
Screening everyone biennially from 40-58	£6,986	35.741	£6,976	£ 158	£ 163, p=100%	£176	£ 178 (45; 317), p=100%

Table 40: Incremental net monetary benefit vs standard of care and other interventions with age for risk stratification 40 (deterministic with Green Book scenario), BRCA carriers

Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental analysis), £	NMB (vs standard of care), £
Standard of care	£7,016	41.854			
Screening at 40 and reinviting till 58 if PSA above 3 ng/ml	£7,025	41.856	£5,202	£18	£18
Screening at 40 and reinviting till 58 if PSA above 2.5ng/ml	£7,031	41.857	£4,275	£20	£38
Screening everyone biennially from 40-58	£7,111	41.887	£2,863	£364	£402

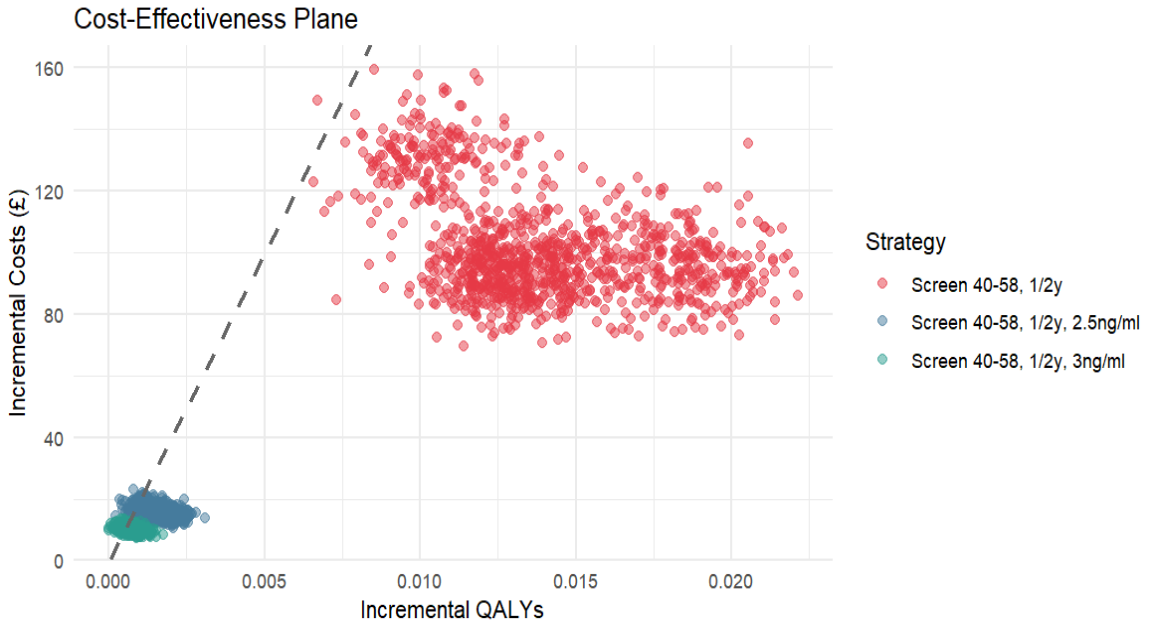
Table 41: Incremental net monetary benefit vs standard of care and other interventions with age for risk stratification 45, BRCA carriers

Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental deterministic analysis), £	NMB (Full incremental probabilistic analysis), mean £, p	NMB (standard of care), deterministic, £	NMB vs standard of care, probabilistic, mean £ (CrI), p
Standard of care	£8,121	38.353					
Reinviting above PSA 3ng/ml (1/2y, 45-61)	£8,137	38.356	£6,313	£34	£ 28, p=100%	£34	£ 28 (7; 51), p=100%
Reinviting above PSA 2.5ng/ml (1/4y, 45-61)	£8,140	38.357	£5,507	£16	£ 13, p=100%	£50	£ 41 (13; 73), p=100%
Reinviting above PSA 2.5ng/ml (1/2y, 45-61)	£8,147	38.357	£6,291	£6	£ 5, p=89%	£55	£ 46 (12; 82), p=100%
Screening everyone 1/4y, 45-61	£8,191	38.366	£5,341	£135	£ 137, p=100%	£190	£ 183 (61;319), p=100%
Screening everyone 1/2y, 45-61	£8,239	38.372	£6,095	£78	£ 70, p=98%	£268	£ 252 (63; 448), p=100%

Table 42: Incremental net monetary benefit vs standard of care and other interventions with age for risk stratification 45 (deterministic with Green Book scenario), BRCA carriers

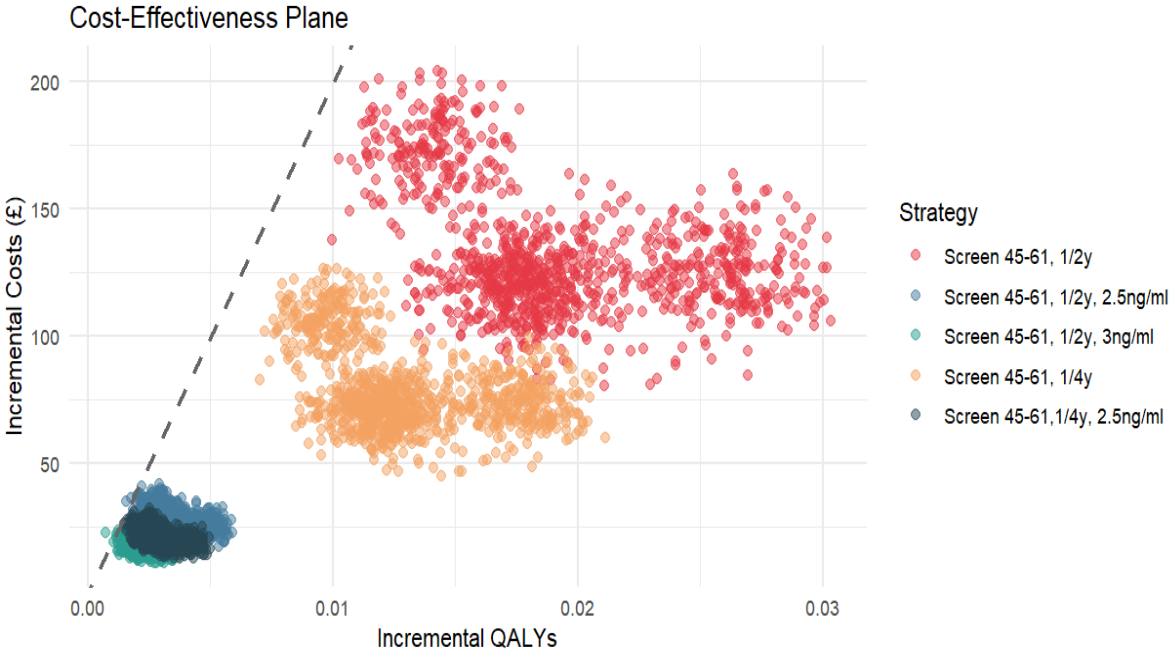
Scenario	Costs, £	Effects	ICER (to standard of care), £	NMB (Full incremental analysis), £	NMB (vs standard of care), £
Standard of care	£8,176	43.428			
Reinviting above PSA 3ng/ml (1/2y, 45-61)	£ 8,192	43.432	£ 3,897	£ 47	£ 47
Reinviting above PSA 2.5ng/ml (1/4y, 45-61)	£ 8,196	43.434	£ 3,252	£ 25	£ 71
Reinviting above PSA 2.5ng/ml (1/2y, 45-61)	£ 8,202	43.445	£ 3,506	£ 14	£ 85
Screening everyone 1/4y, 45-61	£ 8,248	43.455	£ 2,623	£ 255	£ 341
Screening everyone 1/2y, 45-61	£ 8,298	43.469	£ 2,964	£ 154	£ 494

Figure 97: Cost effectiveness plane for screening BRCA carriers starting at age 40 vs standard of care*



**These scenarios were run with the smaller population of 300,000 men and so are subject to stochastic uncertainty*

Figure 98: Cost effectiveness plane for screening BRCA carriers starting at age 45 vs standard of care



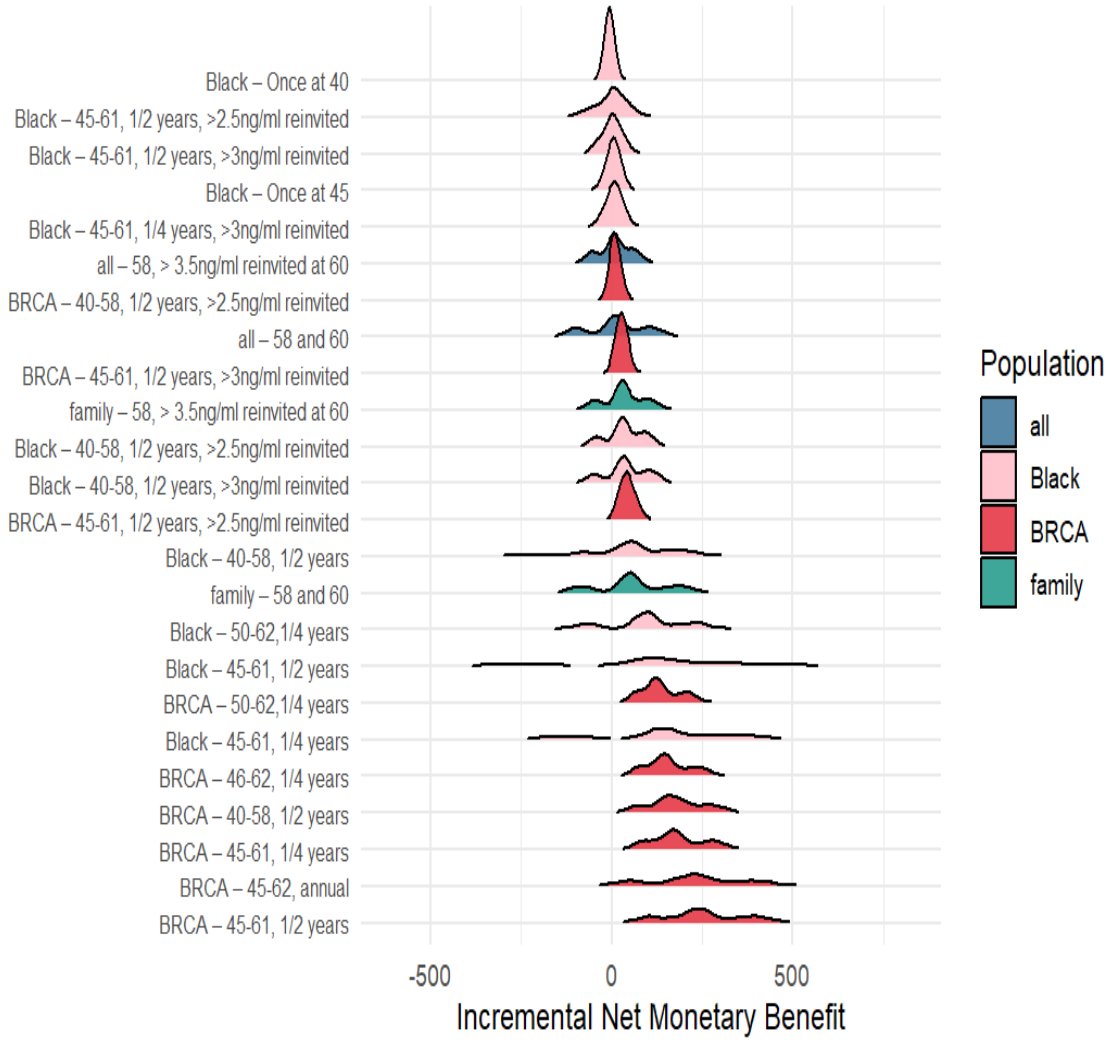
10. Discussion and Conclusions

10.1. Summary of the Modelling Results

As expected, the cost-effectiveness of screening increased with the population's risk and cancer's aggressiveness. Among the four subgroups analysed—all-risk men, Black men, men with familial risk, and BRCA carriers—screening BRCA carriers was the most cost-effective. In contrast, screening the all-risk population was associated with the most uncertainty (Figure 99).

The risk stratification based on initial PSA values was not a cost-effective approach for any of the subgroups. Reinviting all individuals within a given risk subgroup for repeat testing proved to be more effective than basing repeat tests on the results of their initial test.

Figure 99: Incremental net monetary benefit for various screening strategies in four population subgroups: probabilistic analyses



In all four population groups, applying the Green Book scenario—where costs are discounted more heavily than effects, discounting rates vary over time, and the willingness-to-pay threshold is set at £15,000—made screening appear more favourable compared with the NICE reference case. Both the choice of discounting method and the discount rate substantially influenced the cost-effectiveness estimates.

Across all four subgroups, assuming greater harms associated with a cancer diagnosis (modelled as lower utility values for each cancer stage) made screening appear substantially less cost-effective compared with the base case. It is important to note that this scenario is hypothetical and does not reflect real-world evidence on the harms of prostate cancer diagnosis.

A series of additional scenarios were explored to assess uncertainty related to model structure, data inputs, and methodological choices. If the natural history model was calibrated more closely to key targets (e.g., mortality), screening became less cost-effective relative to the base case. Likewise, long-term survival estimates for stages 3 and 4 were extrapolated from only five years of data; replacing these extrapolations with a flat annual mortality rate also reduced the cost-effectiveness of screening.

There is ongoing debate about whether the current NICE threshold of £20,000–£30,000 per QALY is too high, with empirical evidence suggesting a value closer to £15,000 per QALY. Furthermore, screening programmes offered to asymptomatic individuals are often expected to be evaluated against lower thresholds than treatments for symptomatic patients. For this analysis, the NICE reference case with a £20,000 threshold was applied. However, consensus is still needed on the appropriate discounting approach and cost-effectiveness threshold for evaluating screening programmes.

General population: Conclusions

Screening men in England may lead to some benefits in mortality reduction but will also result in substantial overdiagnosis (i.e. detection of cancers that would not have caused symptoms or death in the absence of organised screening).

There is considerable uncertainty about whether screening at ages 58–60 - the ages with the highest positive NMB in deterministic analyses -is cost-effective, with a probability below 60%. Offering an initial screen at age 58 and re-inviting only those above the PSA threshold at age 60 did not improve cost-effectiveness.

The model assumes a perfect correlation between age and PSA levels; under a more realistic, imperfect correlation, risk-stratified repeat screening would be even less cost-effective. Likewise, improving model fit to calibration data or avoiding extrapolation to mortality would reduce NMB and further decrease the probability that screening is cost-effective. Screening all men, irrespective of their risk profile, would also impose substantial demands on healthcare resources.

Familial risk: Conclusions

The cost-effectiveness of screening men with familial risk is highly uncertain, mirroring findings in the general population. In this subgroup, screening at younger ages (<58 years) produces lower NMB than screening at ages 58-60. Using the first PSA result to guide re-invitation was not cost-effective.

Uncertainty is driven largely by limited evidence on the natural history of prostate cancer in men with familial risk - including age-specific onset and progression, the relationship between GGG and progression, correlations between risk factors, and age- and state-dependent changes in PSA. Screening this group would require substantial additional resources, and nearly half of those resources would not benefit men diagnosed through screening because of overdiagnosis.

Results in the familial-risk subgroup were broadly similar to those in all men (albeit somewhat more favourable) partly because a large share of the general male population - over one-third of 60-year-olds - was estimated to have at least one first-degree relative with breast, ovarian, or prostate cancer. As a result, the subgroup is not markedly different from the general population.

The PICO for this project defined familial risk as having any first-degree relative with prostate cancer. Future research should explore whether screening men with a family history of aggressive disease (e.g., early-onset cancers) would be cost-effective; however, suitable data to support such analyses are limited and may require new experimental or observational evidence.

Men of Black ethnicity: Conclusions

Screening men of Black ethnicity leads to reductions in mortality and increases in LYS and QALYs, but also increases overdiagnosis. Screening of Black men has higher probability of cost-effectiveness in probabilistic analyses than screening men of general risk. Nonetheless, substantial uncertainty remains, driven by assumptions regarding the natural history of prostate cancer, test sensitivity, screening uptake, and discounting of costs and effects.

The strategy with the most favourable and least uncertain benefit–harm profile was screening every four years between ages 50 and 62. Screening success, however, would depend heavily on achieving high uptake and the resources needed to support this. Feasibility of achieving higher uptake than observed in experimental studies - 36% in the CAP trial and 27% uptake in a GP-based PSA screening study by Langley at all (2025)[68] - should be explored.

Additional uncertainty arises from ethnicity-related assumptions (e.g., how ethnicity affects disease progression and time to diagnosis). The model also did not include mixed-race men due to data limitations; for instance, observational data often categorise "mixed race" without distinguishing Black–White mixed backgrounds. Furthermore, if calibration were improved to better fit mortality or if alternative survival assumptions were applied, screening would be less cost-effective, as scenario analyses showed.

Considering data limitation, it is uncertain whether screening men of Black ethnicity is cost effective, and it also carries the risk of overdiagnosis. This group would benefit from having better data on how cancer develops, progresses, and what is expected participation rate in screening would be for this population group. Given the small proportion of Black men in the English population, the national-level impact of an organised programme on resource use or mortality would be modest.

BRCA carriers: Conclusions

Screening BRCA carriers is cost-effective, even when extended to older ages such as 62 years. Despite the more aggressive nature of prostate cancer in this group, screening still leads to overdiagnosis, with rates increasing at older screening ages. Probabilistic and sensitivity analyses show high certainty that biennial screening from age 45 to 61 is cost-effective. However, given the small population

eligible for screening, the overall resource use and impact on national mortality would be minimal.

This study did not evaluate the cost-effectiveness of BRCA testing itself, which should be examined separately. In addition, because prostate cancer–specific data for BRCA carriers are limited, the model assumed similar costs, utilities, and stage-specific survival to those of other groups, adding uncertainty to projections for this population.

In the model, each individual was assigned BRCA1 and BRCA2 status based on their familial risk profile and the population prevalence of these mutations; however, the analysis was conducted for BRCA1 and BRCA2 carriers combined. Evidence on the impact of BRCA1 on prostate cancer onset and progression is inconclusive. If the analysis were restricted to BRCA2 carriers only, the NMB would be higher than for the combined BRCA1/2 group, reflecting model assumptions that BRCA2 carriers have a higher risk of cancer onset and more aggressive disease progression.

10.2. Comparative predictions of overdiagnosis

Overdiagnosis represents one of the principal harms of prostate cancer screening [69]. However, estimation of overdiagnosis depends not only on assumptions of natural history disease of prostate cancer, survival, and competing mortality risk, but also on definition of overdiagnosis. Besides the SCHARR prostate cancer model, other well-established natural history model frameworks include the CISNET prostate cancer models (PSAPC, MISCAN-PRO, and SCANS), the CAP model, and the Karlsson/Stockholm Prostata model [26, 28, 42, 70-73].

10.2.1. Description of the models used for comparison

PSA-Prostate Cancer Model (FHCRC)

PSA-Prostate Cancer (PSAPC) model is a microsimulation model developed by the Fred Hutchinson Cancer Research Centre (FHCRC) in the US [27, 72, 74]. It is one of the earliest CISNET prostate cancer natural history models and primarily used to explain the PSA screening in U.S. prostate cancer incidence and mortality trends. PSAPC model assumed that Gleasson Grade does not change over time, and cancer progression is linked to PSA biomarker growth. Model calibration used U.S. SEER incidence trends before and after the introduction of PSA testing, with the

ERSPC trial mortality relative risk used for external validation. Model parameters, including PSA growth rates, screening dissemination, and biopsy compliance are informed by data from the PCPT and PLCO trials.

MISCAN-PRO (Erasmus MC)

The microsimulation screening analysis prostate cancer model (MISCAN-PRO) is a microsimulation model developed by the Erasmus Medical Centre (Rotterdam, the Netherlands) research team, extended the earlier MISCAN framework (originally developed for colorectal and breast cancer) to prostate cancer[25, 69, 75]. The model evaluates impact of PSA screening on prostate cancer incidence and mortality and was used to reconstruct to the results of Rotterdam section of the ERSPC trial and trends in the US population. The latest version of MISCAN-PRO allows tumour grade to progress over time. Calibration was based on data from the ERSPC-Rotterdam trial and Dutch cancer registries, targeting age-specific incidence, stage, and detection rates. In the U.S. version, calibration was performed to match baseline incidence and PSA testing patterns from SEER data [73].

SCANS (Michigan)

The Self-Consistency Analysis of Surveillance (SCANS) model developed by the University of Michigan, is an analytic mathematical model of prostate cancer natural history. SCANS conceptualize prostate cancer as a stochastic process encompassing cancer onset, preclinical detectability, diagnosis, and mortality. The model allows tumour grade to progress over time but does not explicitly link progression to PSA dynamics. Parameters such as disease progression and sojourn time are estimated through Bayesian framework and parametric fitting using real- world population level data from SEER, ERSPC, and PLCO trials, enabling projections of incidence, lead time, and overdiagnosis at the population level.

Karlsson/Stockholm Prostata model

The Stockholm Prostata Model is an individual microsimulation model which was branched from the FHCRC prostate cancer model by the research team at Karolinska Institute[71]. The model links PSA growth with prostate cancer progression and adopts the FHCRC model to the Swedish context using linked

national registries data (SPBR and PCBaSe). Model calibration was based on Swedish age-specific stage distributions, survival data, and the ERSPC incidence rate ratio, while validation reproduced observed Swedish prostate cancer incidence and mortality trends. The model represents the general male population in Sweden, simulating individual life histories from age 35 onward. It is primarily used to evaluate screening strategies for men aged 55–69 years.

CAP-based UK model

The CAP-based UK model is an individual microsimulation model based on cluster randomised trial of PSA testing for prostate cancer (i.e. the abbreviation - CAP)[42]. The model was primarily developed by a team from the University of Bristol to evaluate the cost-effectiveness of PSA screening in the UK using data from CAP and the ProtecT trials. The model extends the Stockholm Prostata Model [71] described above. The model to represent the natural history of prostate cancer through age- and PSA-dependent transitions from healthy, screen-detectable, clinically diagnosed, to death. Model calibration was performed to UK Office for National Statistics (ONS) and CAP trial incidence data by age and Gleason score, with the validation referenced ERSPC and CAP mortality relative risk.

10.2.2. Comparison of overdiagnosis estimation in the models

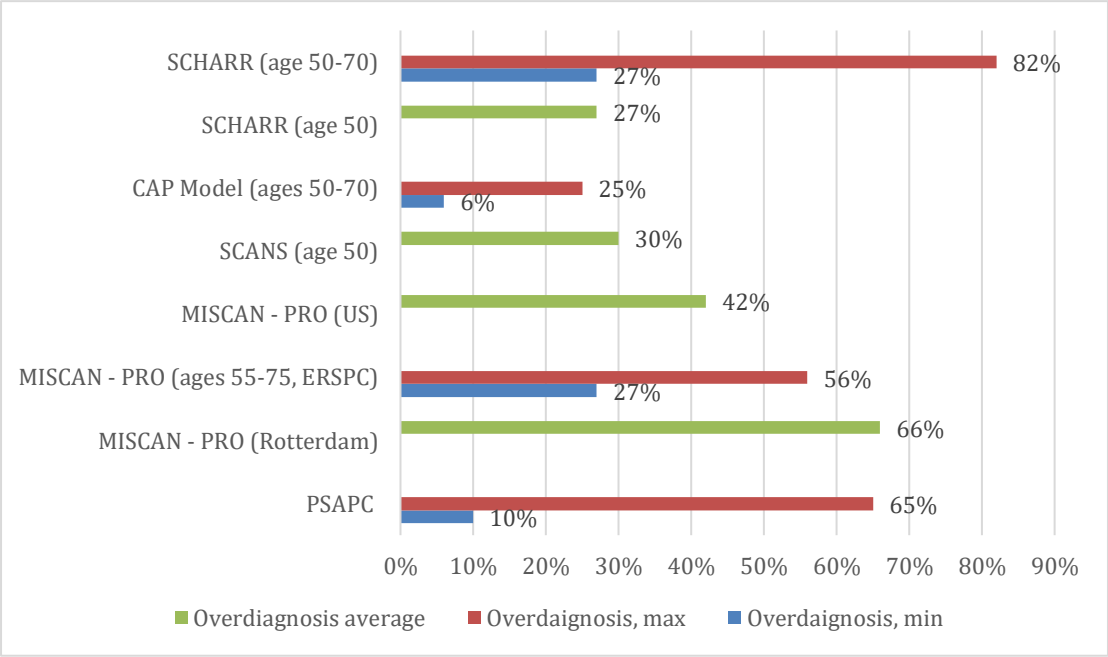
In the reviewed models, overdiagnosis was generally defined as the detection of cancers that would not have been diagnosed in the absence of screening. This was typically reported either as a proportion of all cancer incidence or as a proportion of screen-detected incidence. In the SCHARR model, overdiagnosis was defined as the proportion of screen-detected cases, based on consultations with clinical experts (see Supplementary I, Phase 2). Because the choice of denominator has a substantial impact on predicted levels of overdiagnosis, Figure 100 presents results only from models that applied the same definition.

Estimates of overdiagnosis varied widely (6–82%), both across and within models. These differences were driven by the epidemiological data used for calibration and parameterisation, as well as by the assumed screening ages.

The lowest range of overdiagnosis was predicted by the UK CAP trial when evaluated over a 10-year follow-up period (6–25%). The MISCAN-PRO model showed substantial variation depending on the epidemiological inputs and screening ages. Its minimum estimates were similar to those from the SCHARR model (e.g., 27% at age 50), although SCHARR predicted higher overdiagnosis at older ages. On average, MISCAN-PRO predicted higher overdiagnosis than SCHARR, although this may reflect differences in screening schedules, which were not always clearly reported.

In the SCANS analytical model, overdiagnosis following a single screen at age 50 was broadly similar to SCHARR's estimates for the same age (30% vs 27%). The PSAPC model, like SCHARR, reported a wide range of estimates, with higher overdiagnosis at older ages and with more intensive screening. However, the overall magnitude remained lower than that predicted by SCHARR across all ages.

Figure 100: Overdiagnosis (proportion of screen-detected cancers) under one-time screening of the general population in prostate cancer screening models



Overdiagnosis is a model-dependent construct rather than a directly observable outcome. Differences in model structure, underlying assumptions, definitions, calibration targets, and assumed screening strategies make results challenging to interpret and difficult to compare directly across models. Compared with other models, the SCHARR model predicted similar levels of overdiagnosis to the

CISNET models for screening at younger ages, but higher levels at older ages. These differences may reflect variations in modelling assumptions - such as the age-related increase in PSA values in the SCHARR model - or differences in key parameters, including age-specific prostate cancer mortality.

10.3. Model limitations

The model predictions are subject to substantial uncertainty, primarily due to limitations in the knowledge of the NHD, as well as test accuracy estimates. The model did not incorporate a correlation between ethnicity and familial history because of no data describing three-factors correlations: BRCA carrier, familial history and ethnicity. However, the impact of this limitation was considered to be minimal in the model, since the RR of Black ethnicity was not adjusted for familial history in the used literature.

The model calibration could not achieve a simultaneously good fit to all calibration targets. The scenarios demonstrate that if better fit to mortality is achieved, screening in all scenarios would be less cost effective. Uncertainty in the natural history parameters informed through calibration - specifically, the concentration of parameter estimates into several distinct regions of the parameter space - led to the probabilistic NMB results forming multiple clusters. This clustering became more pronounced at older screening ages, reflecting using age as a correlated factor across several model functions.

As it is described above, the model was calibrated to the standard of care scenario, in the absence of data on opportunistic screening uptake by age and individual risk. Calibration results showed that the NHD parameters were split across several distinct regions of parameter space, what impacted probabilistic results. No informative priors were imposed during calibration, except for two assumptions: that the probability of clinical diagnosis under standard care increases with cancer stage, and that higher-grade cancers (based on GGG) are more aggressive than lower-grade ones. Considerable uncertainty remains regarding the true NHD of prostate cancer.

There is also notable uncertainty in how PSA values change with age across different groups—men without cancer, men with low-grade cancer, and men with

high-grade cancer. This includes both general trends in age-related PSA increases and the specific trajectories of PSA progression by cancer status.

The model assumes that cancer progresses faster in older age. However, much of the model input data are underrepresented in younger (<55 years) and older (>75 years) age groups. This affects both trial-based and population-level data, where these age groups are less prevalent, resulting in increased uncertainty in model predictions for these subpopulations.

The model applies flat uptake rate; future model development should incorporate variable uptake based on results of the previous screening decisions.

There was inconsistency in expert opinion on whether prostate cancer is more aggressive in men of Black ethnicity than in other ethnic groups. Phase 1 stakeholder engagement supported the view that prostate cancer is more aggressive in this population, and this assumption was incorporated into the model. However, Phase 2 stakeholders disagreed with this assumption. If the latter view is correct, the model may overestimate the cost-effectiveness of screening in men of Black ethnicity. In such a case, the predictions for this group would more closely resemble those for men with familial risk, as both groups share similar hazard ratios for cancer onset.

The model assumes that the prognosis of prostate cancer in BRCA carriers after diagnosis depends solely on the stage at which the cancer is detected. Consequently, costs, utility values, and survival outcomes by stage for BRCA carriers were assumed to be comparable to those of other subgroups. No evidence was identified to support alternative assumptions.

10.4. Data limitations and requirements

The model development revealed several important gaps in knowledge about the natural history of prostate cancer and PSA screening. Addressing these limitations through future clinical studies would help to reduce modelling uncertainty and support more informed decision-making around prostate cancer screening.

Prevalence of BRCA Mutations in the Population

The model assumed prevalence rates of 1 in 381 for BRCA1 and 1 in 277 for BRCA2, based on Maxwell et al. (2016) and expert opinion within the reference group which supported the development of the model. However, in discussion on the draft model, these estimates were considered to be too high by some and not based on the UK population, suggesting prevalence rates of 1 in 1,428 for BRCA1 and 1 in 416 for BRCA2, based on UK Biobank data.

The research team acknowledges potential difference in data but also notes that UK Biobank is not representative of the general UK population, for instance it underrepresents individuals of Black ethnicity. Since both ethnicity and BRCA1/2 mutations are genetically inherited, reliance on Biobank data may underestimate BRCA mutation prevalence.

Improved estimates of BRCA1/2 prevalence in a representative UK population are therefore essential for future research on high-risk groups.

Impact of BRCA1 Mutation on Cancer Onset Risk

The model assumed that BRCA1 carriers have an increased risk of prostate cancer onset. However, this assumption was questioned in discussion on the draft model report. Other cited sources did not find which found significant association between BRCA1 mutation and prostate cancer incidence.

Data on Black Ethnicity

The model assumed that Black men have a higher probability of presenting with advanced GGG at diagnosis, and - since progression speed in the model is correlated with GGG - a faster disease progression overall. This assumption was informed by US-based data, as no equivalent UK data were identified (see Section 5.5).

However, the higher GGG at diagnosis observed in US studies may reflect health inequalities rather than biological differences. Therefore, these data may not be generalisable to the UK context.

There is a clear need for UK-specific data on the Black population, particularly regarding GGG distribution by age at diagnosis, stage at diagnosis by age,

incidence-to-mortality ratios, and uptake of and response to opportunistic and organised screening programmes.

Distribution of GGG by Age in Younger and Older Populations

Clinical experts generally believe that older men have higher GGG scores. In discussion on the draft model report, it was suggested that national statistics and CAP trial data support that older men in no-screening arm indeed have a higher proportion of GGG3-5, but not of GGG1-2. Instead, they have a greater proportion of unclassified grades.

Data on GGG distribution among younger men remain highly uncertain due to small sample sizes and restrictive trial inclusion criteria.

Future studies, particularly those classifying missing grades and reflecting post-COVID distributions, would help reduce uncertainty in the natural history of prostate cancer.

Prostate Cancer Incidence by Age and Stage (TNM or Stage I–IV Classification)

Converting between staging systems introduces additional uncertainty. As recent UK data on age- and stage-specific incidence were unavailable, several assumptions were required (see Section 5.7). A significant proportion of cancers were unstaged based on the NHS data. Similar to the CISNET prostate cancer models and following expert consultation, it was assumed that unstaged cancers were evenly distributed across the four stages. However, not all experts agreed with this assumption.

No data specific for prostate cancer were identified to confirm or refute this approach, highlighting the need for studies examining how missing stage data could be better classified.

Prevalence of Undiagnosed Prostate Cancer

There are no prostate cancer screening studies using highly sensitive tests to estimate the prevalence of undiagnosed disease, nor are there large-scale pathology studies. Consequently, the prevalence of undiagnosed cancer, speed of progression, and rate of clinical diagnosis in the absence of screening remain

uncertain. Although prostate cancer is generally considered a slow-growing disease, there is disagreement in the modelling studies about the exact rate of progression and whether it varies by age.

Clinical Presentation Rate and Opportunistic Screening Coverage

In current practice, prostate cancer may be detected incidentally (during investigation for other suspected conditions), through opportunistic screening, or following symptomatic presentation. In all cases, men may receive a PSA test.

According to Martins et al. (2018), around 40% of men have had at least one PSA test, but most of these were performed for reasons other than opportunistic screening. This implies the difficulty of distinguishing the opportunistic screening - PSA testing initiated by the health service when someone presents for something else and does not have any symptoms – from incidental findings when someone presents with symptoms but have another disease suspected or symptomatic patients when someone has a suspected prostate cancer and get the PSA test within the symptomatic pathway. The UK clinical experts also emphasised that this estimate is imprecise and likely to vary by time and region.

Due to this uncertainty, we were unable to construct a group were tested while asymptomatic as a consequence of self-selection through currently available routes. This was a limitation of the modelled standard care arm against which organised screening was compared.

While constructing such a scenario could be informative, it would either rely on outdated pre-PSA data or require detailed information on the proportion of men undergoing truly opportunistic screening (i.e. without clinical suspicion) by age, ethnicity, familial risk, or BRCA status. Such data would improve the reliability of comparator arms in future modelling projects.

Sensitivity of mpMRI

Sensitivity and specificity values for mpMRI and biopsy were based on Ahmed et al. (2017). As no data were identified describing variation in mpMRI sensitivity by stage or grade, uniform sensitivity was assumed across all cancers.

Future clinical data should explore how mpMRI sensitivity varies by stage and grade to improve model accuracy.

Changes in PSA Levels by Age

No data were found describing changes in PSA levels with age among men without prostate cancer. Clinical opinions diverged: some experts suggested PSA increases with age in all men, while others believed this occurs only in men with cancer.

Some studies also indicated higher PSA levels among Black men, though these did not simultaneously account for age, stage, and ethnicity.

Further research on PSA trajectories across health states and risk profiles is needed to support more informed decision-making.

Correlations Between Diagnostic Test Accuracies

The model did not incorporate correlations between different diagnostic tests, as no suitable data were identified. Future studies examining correlations in test positivity across multiple diagnostic tests would help to improve the realism of future prostate cancer models.

Screening uptake and screening invitation costs

It is recognised that real-world screening uptake is likely to be lower than that observed in clinical trials, and that uptake for repeat screening may differ from uptake for a single screening round. Uptake by age, ethnicity, and screening round should therefore be estimated in future trials.

Future studies should also collect data on the resources required to invite each individual, including indirect costs, reminder systems, and information-support infrastructure.

10.5. Future modelling work based on additional data

Future model updates should incorporate better clinical data to reduce uncertainty in both the NHD and test accuracy. In particular, the following data are essential to improve model validity and calibration:

1. **PSA testing uptake**: Data on opportunistic PSA test uptake by calendar year, age, and individual risk factors (e.g., ethnicity, BRCA carrier status, familial risk). Where possible, access to individual-level GP records should be pursued.

2. **Stage and grade at diagnosis**: Information on stage at diagnosis and GGG by year, age, and ideally by risk group.

3. **PSA trajectories:** Longitudinal data on changes in PSA levels over time in men without prostate cancer, and those with low-grade and high-grade diagnosed cancers.

4. **Prognostic data by age**: GGG distribution at diagnosis, progression patterns, and survival outcomes among younger men (<55 years) and older men (>75 years), who are currently underrepresented in existing datasets.

10.6. Future research questions

The current project evaluated three screening strategies: (1) single screening using a PSA threshold of 3 ng/ml, (2) repeat screening with the same threshold, and (3) risk-adapted screening, where invitation to repeat testing was based on the initial PSA result using either the NICE-recommended age-specific threshold or a flat threshold of 3 ng/ml regardless of age.

Although the original proposal included modelling the impact of using the Stockholm3 algorithm among men with a Charlson Comorbidity Score below 3, this component was not implemented. The required data were not publicly available, and the Stockholm3 research team only offered data access under conditions that included a full embargo on research outputs. Future modelling questions may focus on evaluating cost effectiveness of biomarkers in hugh-risk groups.

While some risk-stratified screening approaches were modelled, the interventions assessed were not exhaustive. More different risk stratification strategies can be modelled after consultations with clinical experts.

The literature suggests a number of additional strategies that aim to improve the diagnostic accuracy of screening. These include consideration of PSA density alongside PSA level, the use of machine learning algorithms following two annual

PSA tests, and enhanced triage (e.g. targeted biopsy) based on combined risk indicators.

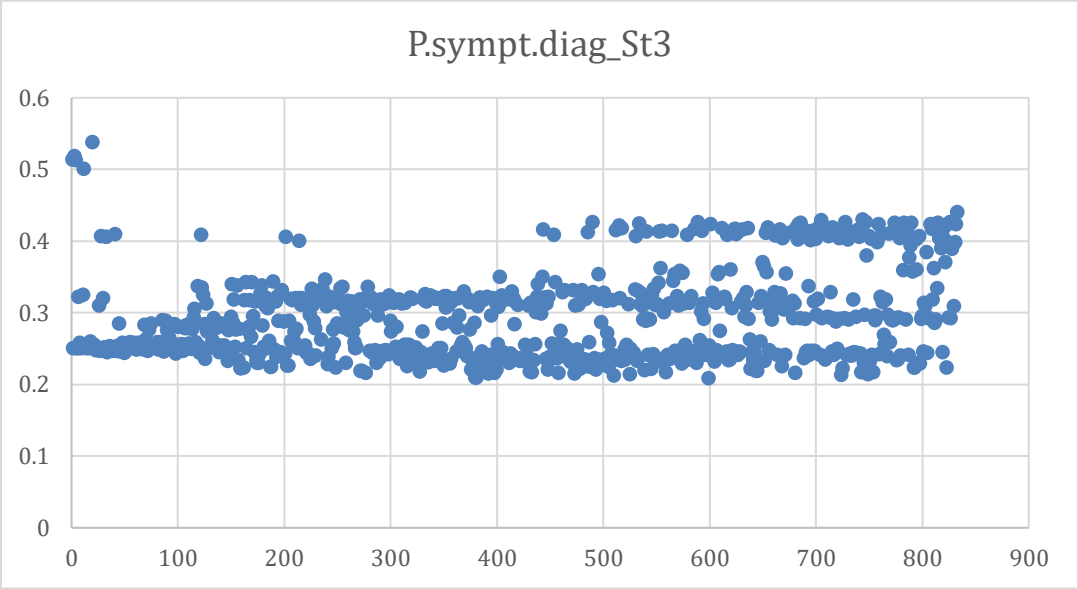
Further approaches worth exploring include targeted screening of men with multiple risk factors—for example, those with both a family history and Black ethnicity, men with more than one first-degree relative affected, or those who have relatives diagnosed at a younger age.

Finally, the cost-effectiveness of BRCA testing in men was not evaluated in this study. This represents an important opportunity for higher-risk groups such as those with a family history. Future work in this direction should account for the broader health and economic implications of genetic testing for BRCA status, including cascade testing in relatives.

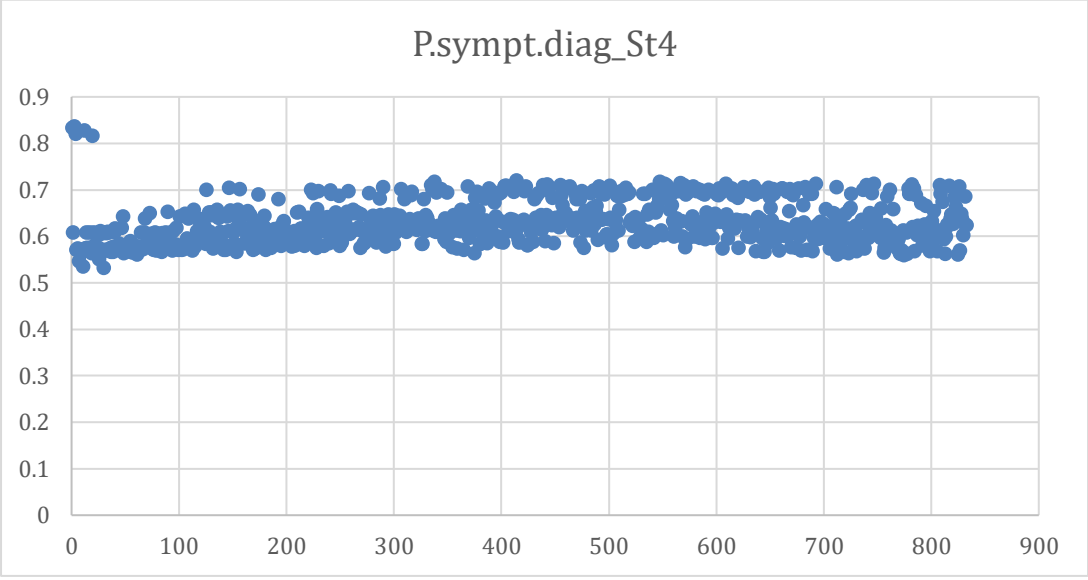
Supplementary A: Comparison of parameter space for probability of onset parameter

Some of the parameters from the posterior samples with high likelihood and so high probability to be sampled in probabilistic analysis (the parameters closer to 0 on x axes have higher likelihood), were sampled from very distinct parameter spaces. Such parameters are presented on the Figures below.

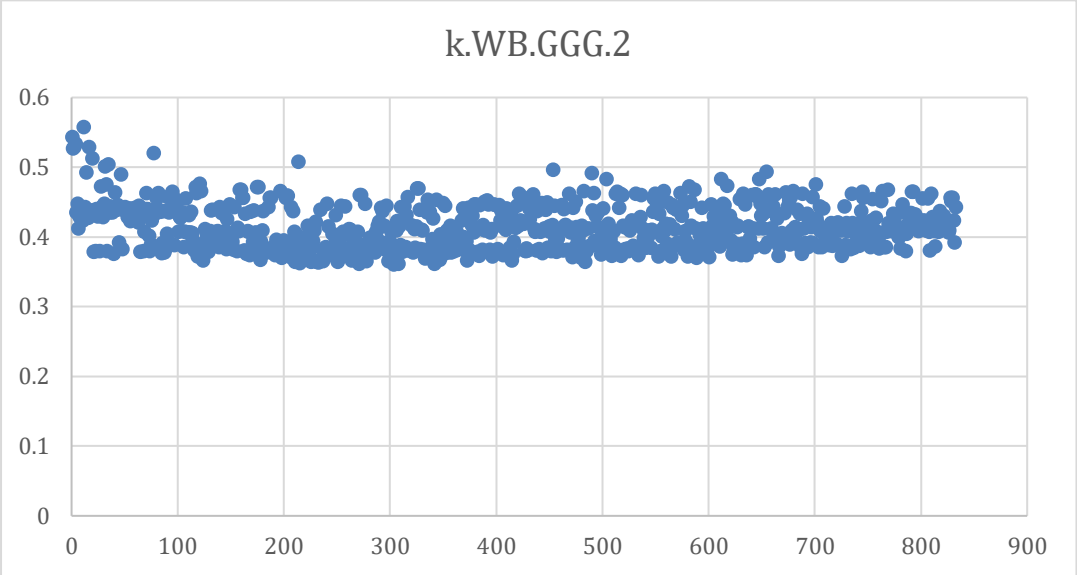
Supplementary Figure S1. Posterior distribution for the probability of clinical diagnosis in stage 3 (annual)



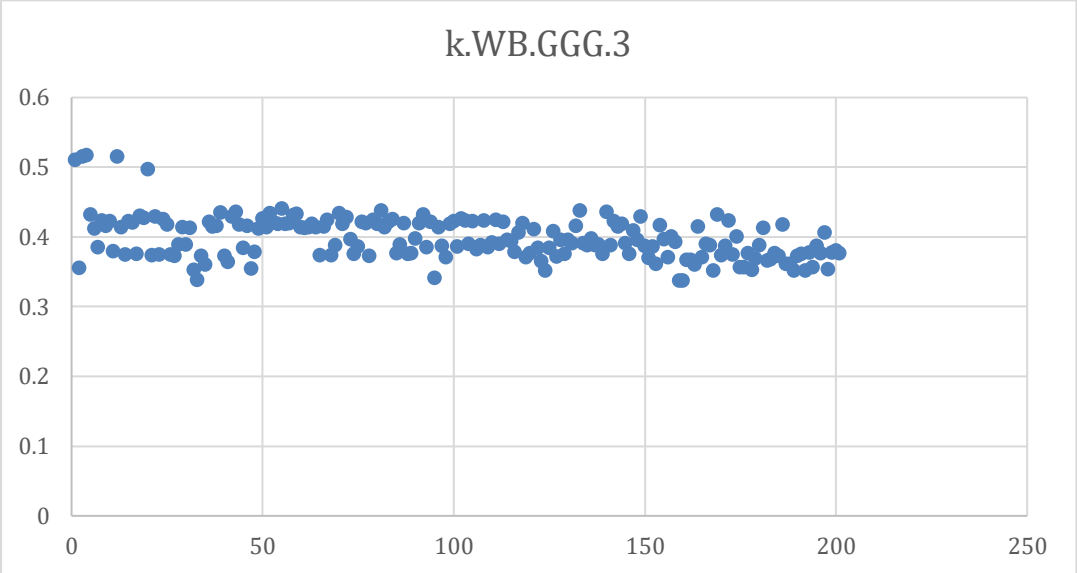
Supplementary Figure S2. Posterior distribution for the probability of clinical diagnosis in stage 4 (annual)



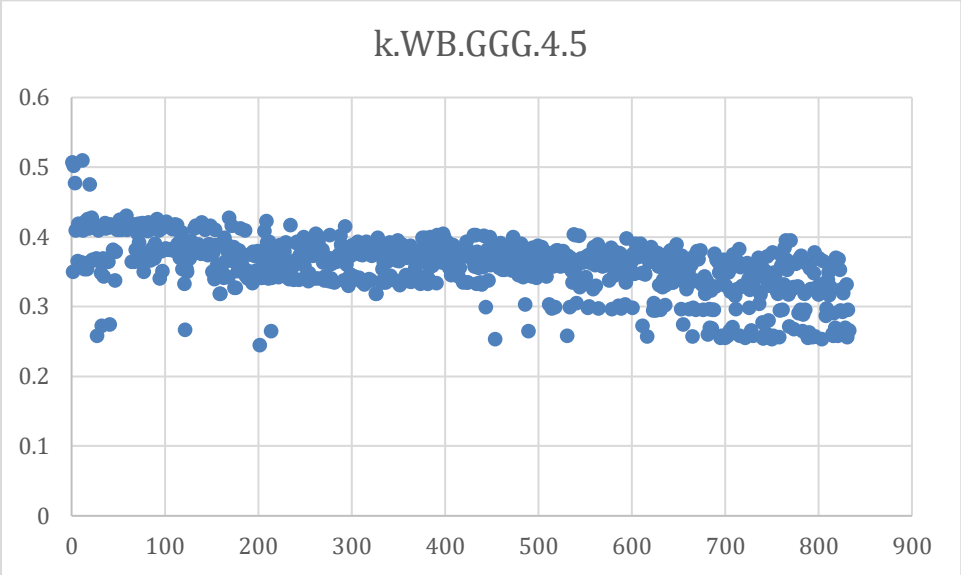
Supplementary Figure S3. Posterior distribution for the coefficient defining the speed of cancer progression in GGG2



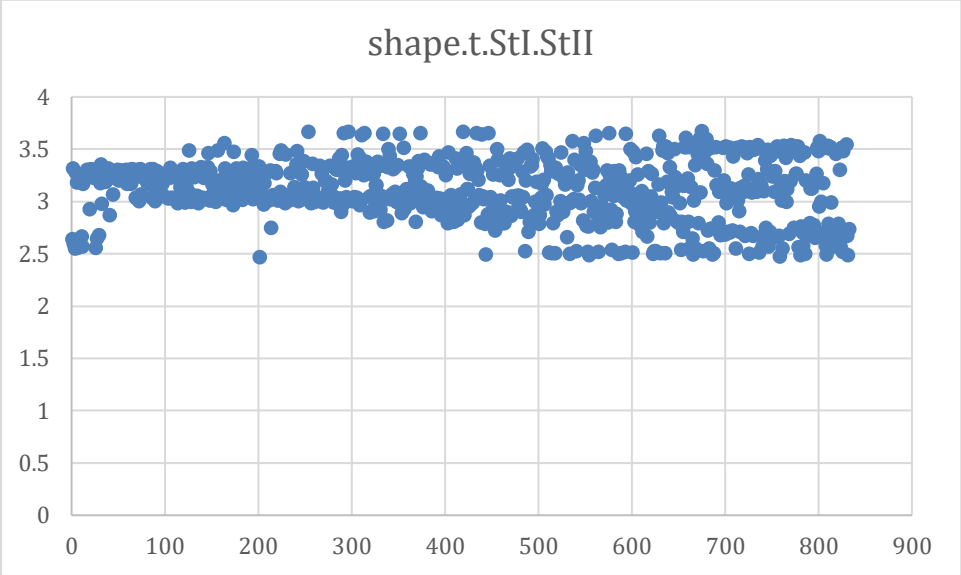
Supplementary Figure S4. Posterior distribution for the coefficient defining the speed of cancer progression in GGG3



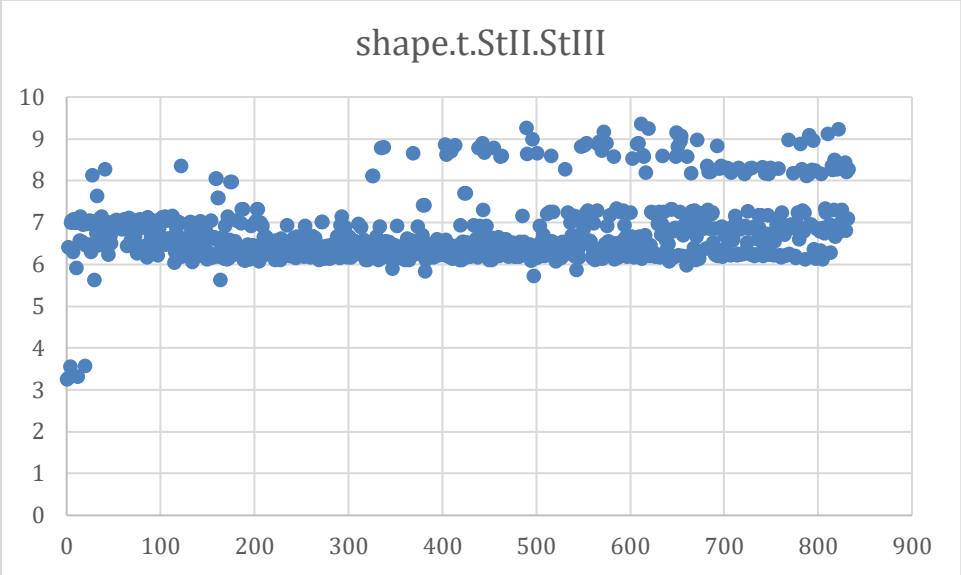
Supplementary Figure S5. Posterior distribution for the coefficient defining the speed of cancer progression in GGG4-5



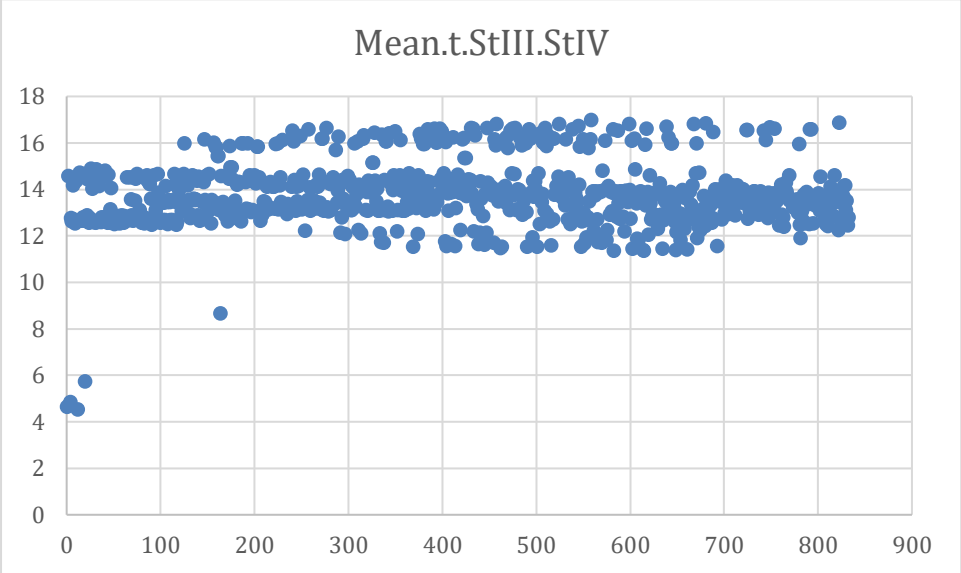
Supplementary Figure S6. Posterior distribution for the shape of the Weibull distribution defining time of progression from stage 1 to stage 2



Supplementary Figure S7. Posterior distribution for the shape of the Weibull distribution defining time of progression from stage 2 to stage 3



Supplementary Figure S8. Posterior distribution for the mean of the Weibull distribution defining time of progression from stage 3 to stage 4 in GGG1



Supplementary B. prostate cancer model data verification checks

Comparison of inputs used in the model and reported in the model report and technical documentation:

- **HSE file size and age limit**: Identical (correctly excludes individuals under age 20 and identical dataset size)
- **BRCA uptake age**: Identical (correctly implements age-specific probabilities from Forde et al. 2020)
- **Family history allocation**: Identical (correctly implements non-random allocation based on BRCA status)
- **Palliative care costs by age and mortality cause**: Identical (correctly implements Table 24 values with appropriate inflation)
- **GGG allocation:** Fixed (now includes proper values for age 30, same as age 35 in the report)
- **Cancer risk based on familial history HRs**: Identical (correctly implements HRs for PC: 1.94, BC: 1.23, OC: 3.45)
- **Familial history probabilities based on BRCA status**: Identical (Table 3 values correctly implemented)
- **Accuracy of PSA test**: Identical (sensitivities of 0.214 for GGG1, 0.449 for GGG2-5, 0.973 for Stage 4, specificity 0.851)
- **Accuracy of diagnostic MP-MRI**: Identical (sensitivity 0.88, specificity 0.45)
- **Accuracy of LATP biopsy**: Identical (sensitivity 0.52 for GGG1, 0.85 for GGG2+, specificity 0.98)
- **Utility decrements and multipliers**: Identical (correct implementation of Table 18 values and biopsy decrement)
- **Sojourn time**: Identical (correctly calibrated to 13.4 years for 50-69 year-olds)
- **Calibration parameters**: Identical
- **Incidence of prostate cancer by age**: Minor discrepancy (713.5 vs 714 per 100,000 for ages 80-90+, negligible impact). **Corrected**.
- **Incidence by stage and age**: Identical
- **Mortality by age**: Identical
- **GGG by age and stage**: Identical
- **Inflation and cost calculations**: Mostly identical with one exception:
    - **PSA Test cost inflation**: Identical (£5.91 → £7.81)
    - **Stage costs inflation**: Identical
    - **MP-MRI cost inflation**: Discrepancy (report shows £270.25, model uses £316.01, ***17% higher***). **Corrected.**

Supplementary C. Internal validation of the model

We validated the model with by initially getting someone experienced in developing economic models in R software (DP), to test the sensitivity of model outputs to extreme parameter inputs. If the tests looked like they produced results that could not explained, they were followed up in detail by a modeller experienced in this model (LM), as this an efficient use of staff time, given the model complexity. The face validity of initial results were assessed by LM and DP and were potential issues were identified, we went back through the model to try to see if these results were logical and if not, we then identified the errors in the model code. More comprehensive validation checks, e.g. rebuilding the base case model, were ruled out for this project as given the complexity of the model they could not be feasibly conducted within the timelines of the project.

Extreme parameter input validation tests

We broadly followed the principles of black box checks in Tappenden et al[76]. We conducted the follow extreme parameter inputs to the test the model as detailed below.

Table 1: A summary of the black box checks conducted on the MIMIC-prostate model

Model test	Error identified?	Details
Changed the uptake of PSA screening to 0, 0.25, 0.36 (base case) 0.5, 0.75,1	No	Model results behaved as expected. As more people where screened, number of identified cancers rose
Changed the sensitivity and specificity of MP-MRI to 0 and 1	No	Model results behaved as expected. Fewer cancers were identified when MP-MRI sensitivity and specificity were 0 Many more cancers were identified when MP-MRI sensitivity and specificity were 1
Changed the sensitivity and specificity of biopsy testing to 0 and 1	No	Model results behaved as expected. Fewer cancers were identified when biopsy sensitivity and specificity were 0 More cancers were identified when biopsy sensitivity and specificity were 1
Changed the utility estimation so every patient's utility was 1 and all decrements were 0	Yes	Life years did not equal QALYs. Error was identified as a small typo in the outcomes estimation recording QALYs in the wrong column. When fixed life years equalled QALYs.
Set all screening costs to £0 and then to £100,000	No	Costs behaved as expected, when all screening costs were 0 the only costs remaining were associated with the treatment of cancers. When all costs associated with screening

		where £100,000 total costs were substantially higher than when costs were £0.
We set everyone in the model to have: **1) Neither BCRA1 or BCRA 2 genes** **2) BCRA1 genes** **3) BCRA2 genes**	No	The model behaved as expected. Compared to the base case (where a proportion of people have BCRA1 or BCRA 2), fewer cancer cases were found when no one had BCRA genes More cancer cases were identified in the scenarios where everyone either had BCRA1 or BCRA2 genes.
We set everyone in the model to have familial history of: **Breast Cancer, Ovarian Cancer & Prostate cancer without BCRA mutations** **Breast Cancer, Ovarian Cancer & Prostate cancer with BCRA mutations**	No	The model behaved as expected. The total incidence of prostate cancer at 80 years was higher in both of these scenarios than in the base case model. The incidence was higher when the familial history was with BCRA mutations compared to familial history without BCRA mutations.
Changed the relative risks of developing cancer associated with ethnic status, BCRA status and familial history on developing prostate cancer and GGG score to be **Very harmful, 10** **Very protective, 0.1**	No	We compared the results to the base case model. Setting all risk factors separately (i.e. one by one) to be very harmful made the model behave as predicted with increases in the incidence of cancer. Setting all risk factors separately to be protective made the model behave as predicted with deceases in the incidence of cancer.
Set the mortality from prostate cancer to be 0	No	The model behaved as expected, life expectancy increased when death from prostate cancer was not possible
PSA, prostate specific antigen; MP MRI, multiparametric magnetic resonance imaging; QALYs, quality adjusted life years		

Deterministic results checking

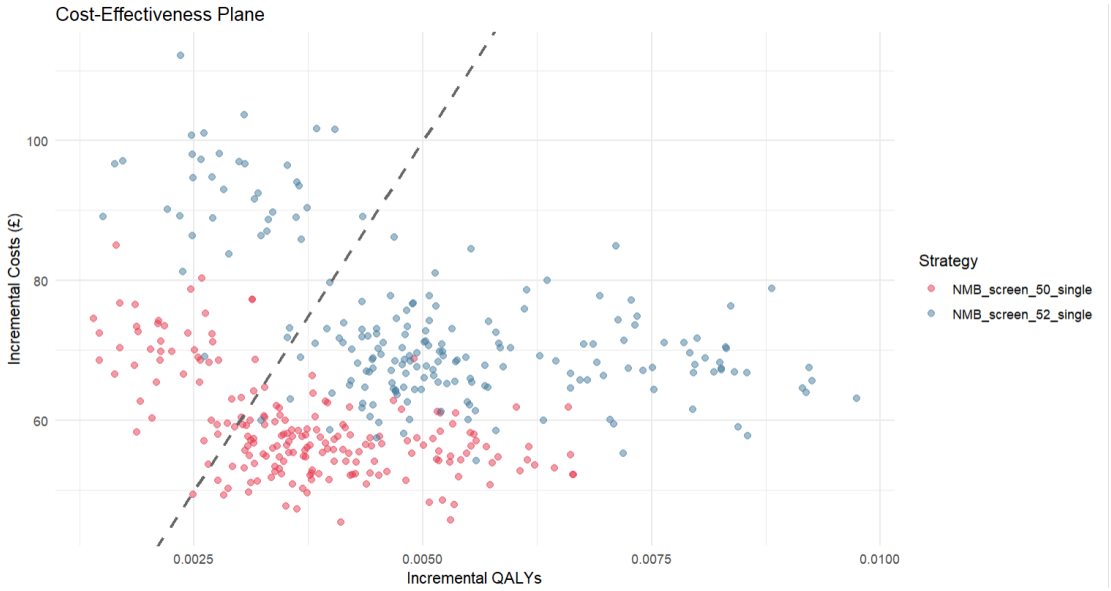
The model was run with a single year of screening starting in two-year age bands. We identified that the incidence of cancer changed unexpectedly when screening changed over 10-year age bands (e.g. between 58 and 60, 68 and 70). Extensive checks of the model code were conducted. An error was identified that NICE PSA

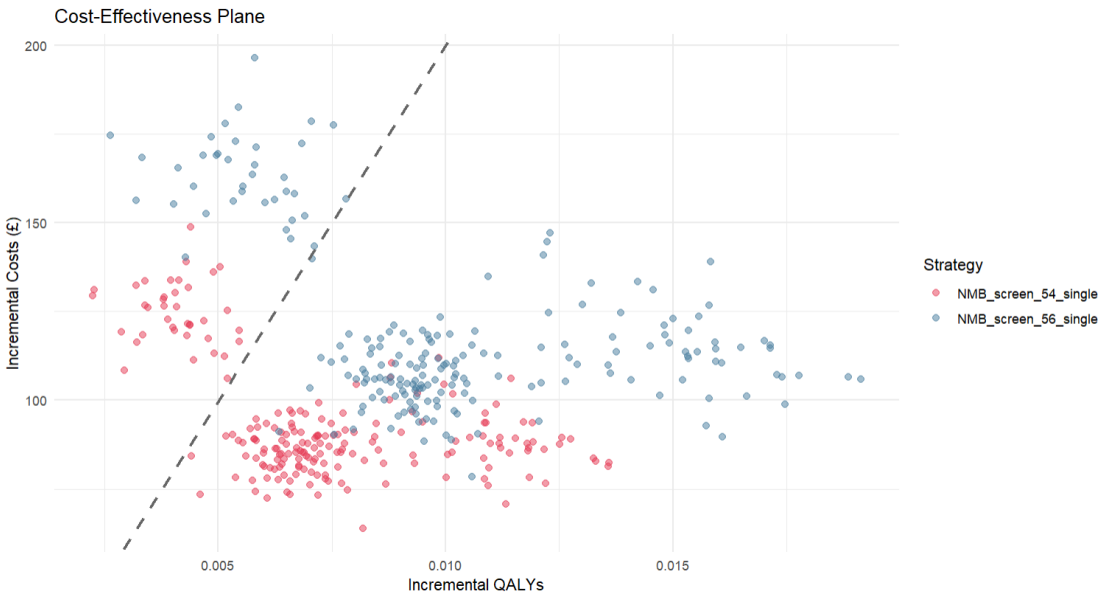
thresholds were being used to determine PSA screening results in the base case model. This error was addressed in the final version of the model.

Probabilistic results checking

The results of the probabilistic analysis were further examined to assess whether the observed distribution of predicted NMB on the cost-effectiveness planes could be explained by the NHD parameters. To investigate this, the model was run for family members using 100 probabilistic simulations for single screening conducted at ages 46, 48, 50, 52, 54, and 56.

The resulting plots indicate that the influence of NHD parameters is smaller at younger ages (rather than that these predictions change in a specific age). This is consistent with expectations, given the exponential relationship between age and cancer onset. Additionally, cancer aggressiveness is associated with the GGG, and the probability of a person having a higher GGG increases with age. Moreover, the parameters defining cancer aggressiveness appear to cluster into two distinct regions, further contributing to the observed patterns.


Cost-Effectiveness Plane
Incremental Costs (£)
100
80
60
0.0025
0.0050
0.0075
0.0100
Incremental QALYs
Strategy
NMB_screen_50_single
NMB_screen_52_single


Cost-Effectiveness Plane
Incremental Costs (£)
200
150
100
0.005
0.010
0.015
Incremental QALYs
Strategy
NMB_screen_54_single
NMB_screen_56_single

Supplementary D. Model checks through scenario analyses

1. Discount rate of 5% applied to effects and 3.5% to costs.

Fig 1a. General population

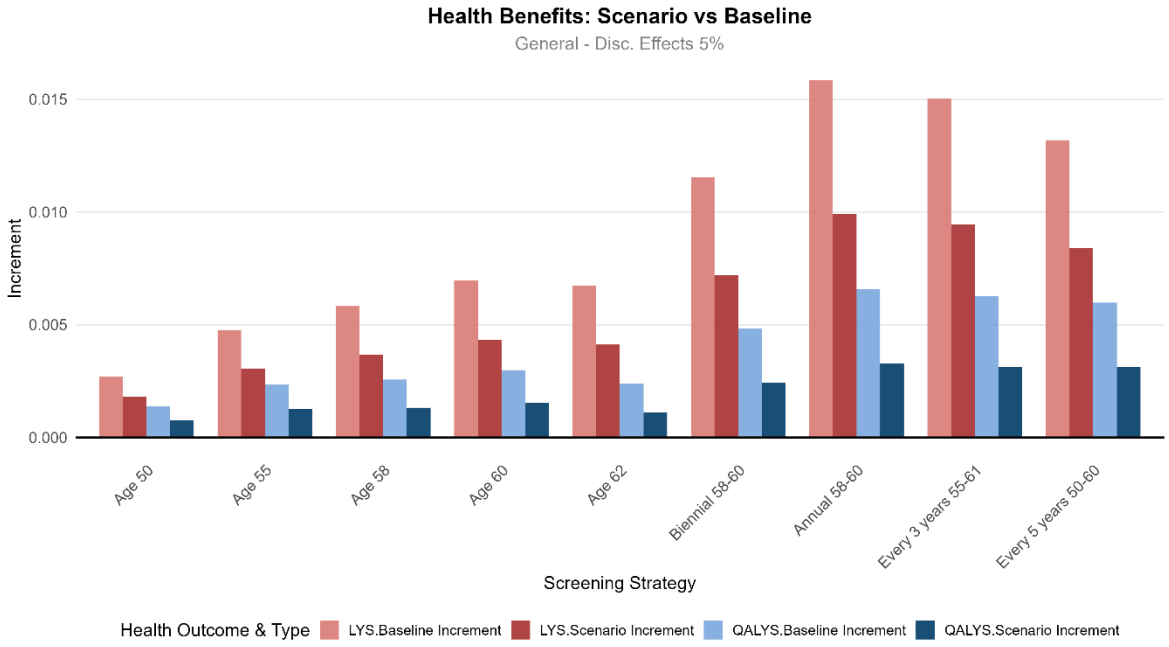
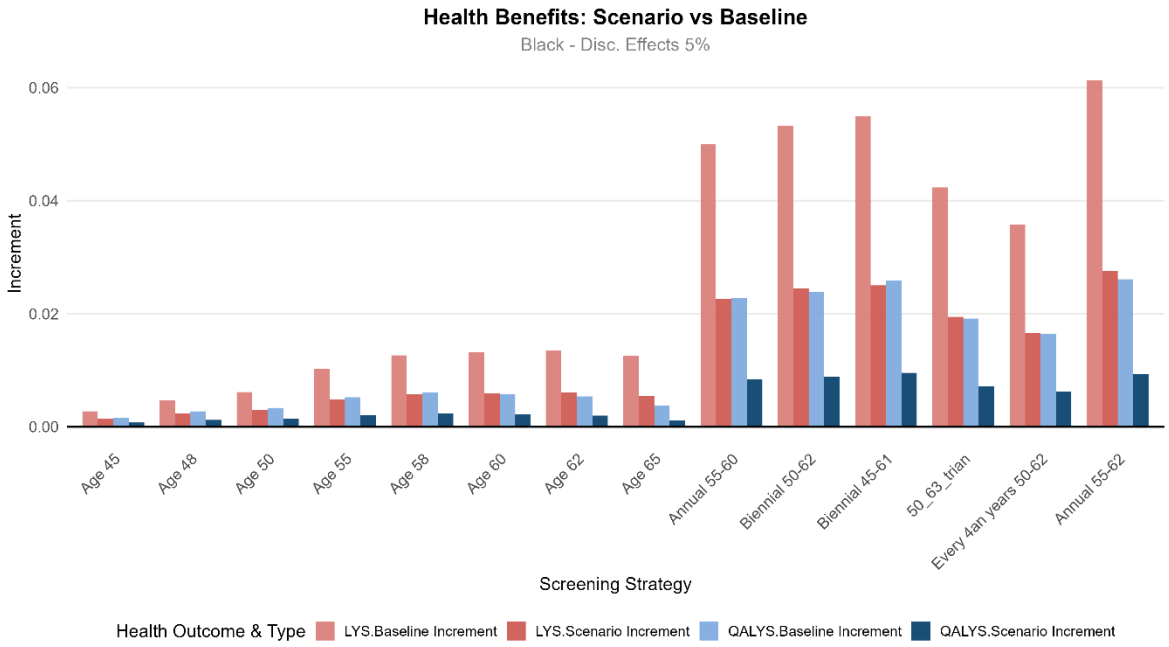
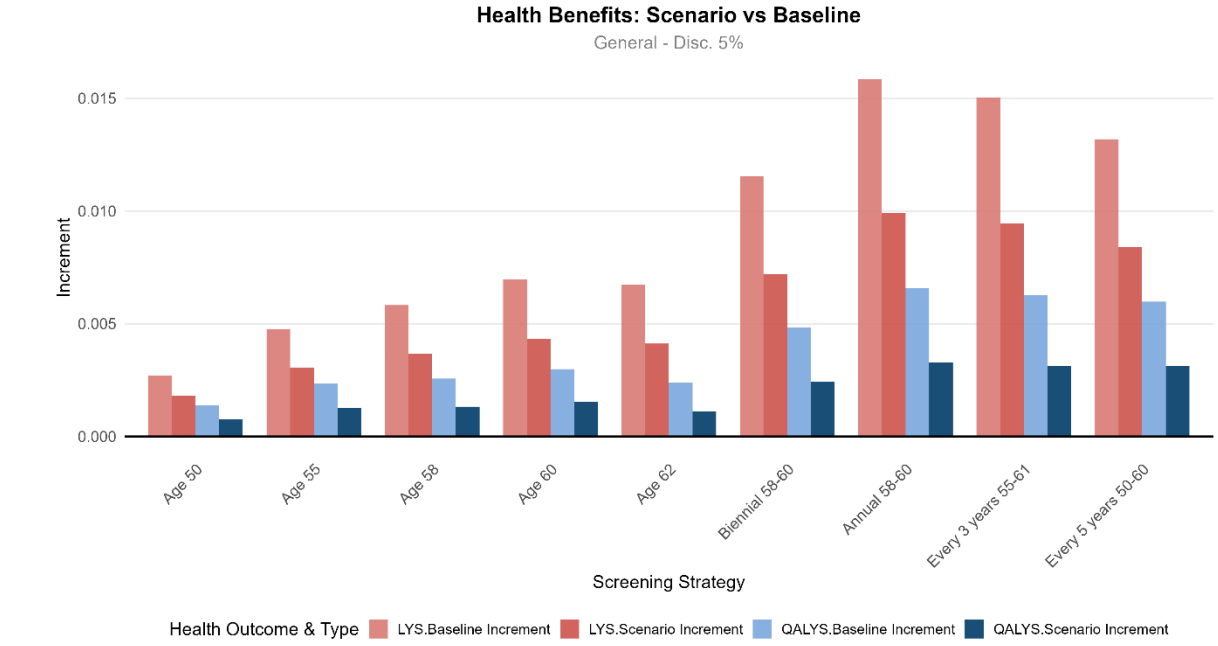


Fig 1b. Black population



2. Discount rate of 3.5% applied to both effects and costs.

Fig 2a. General population



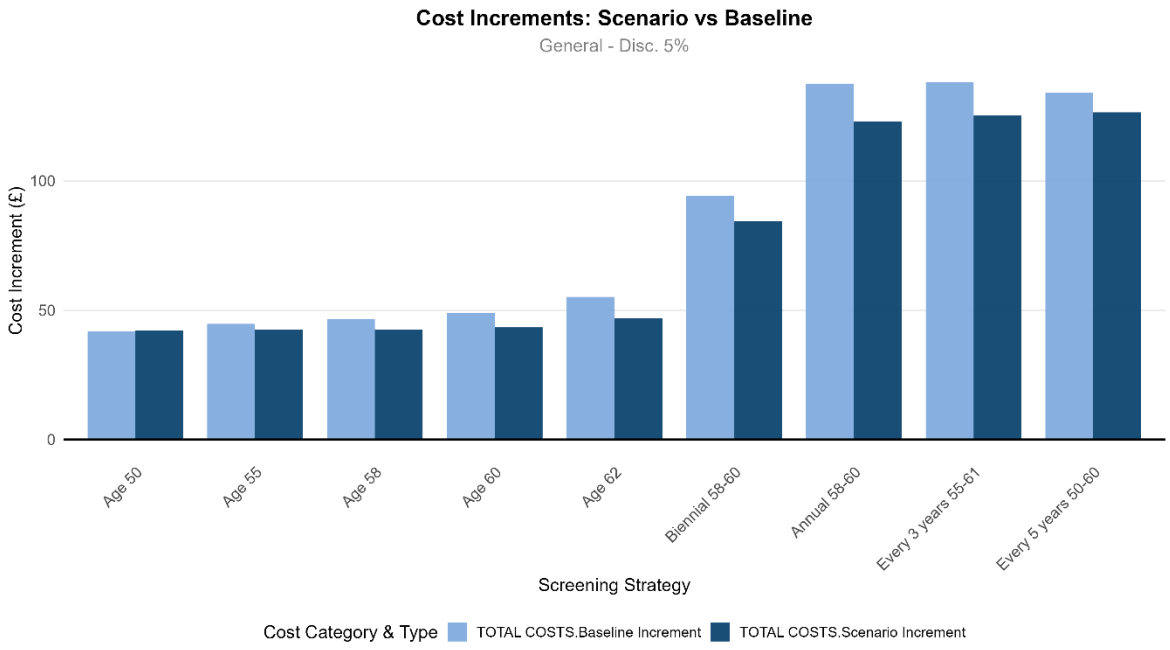
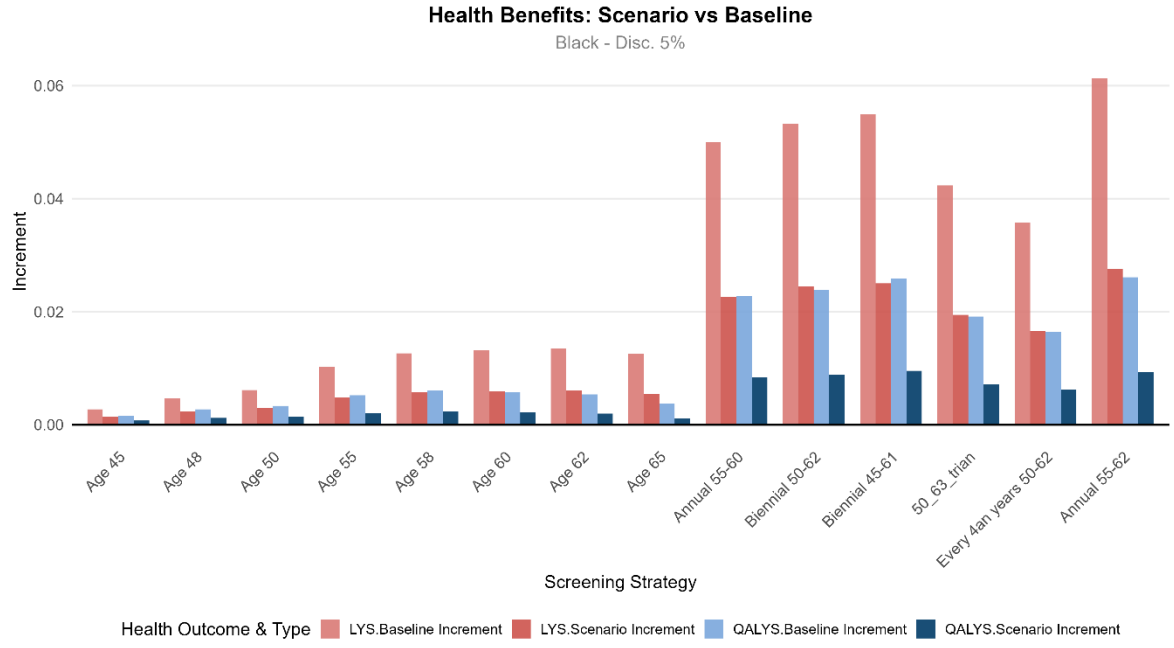
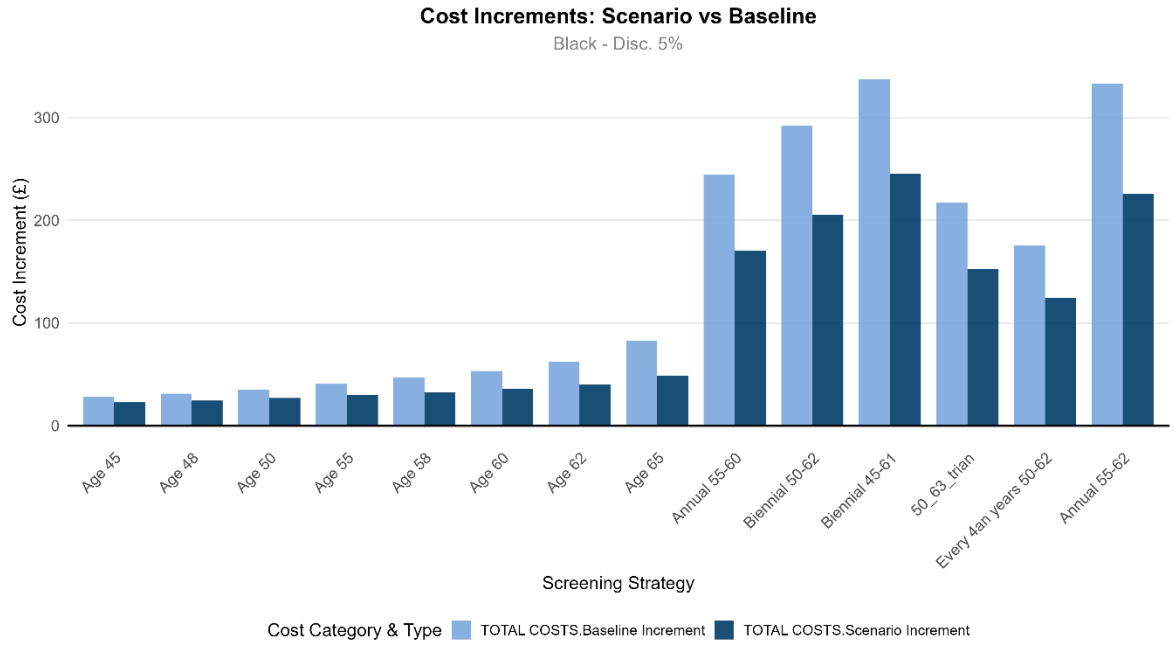


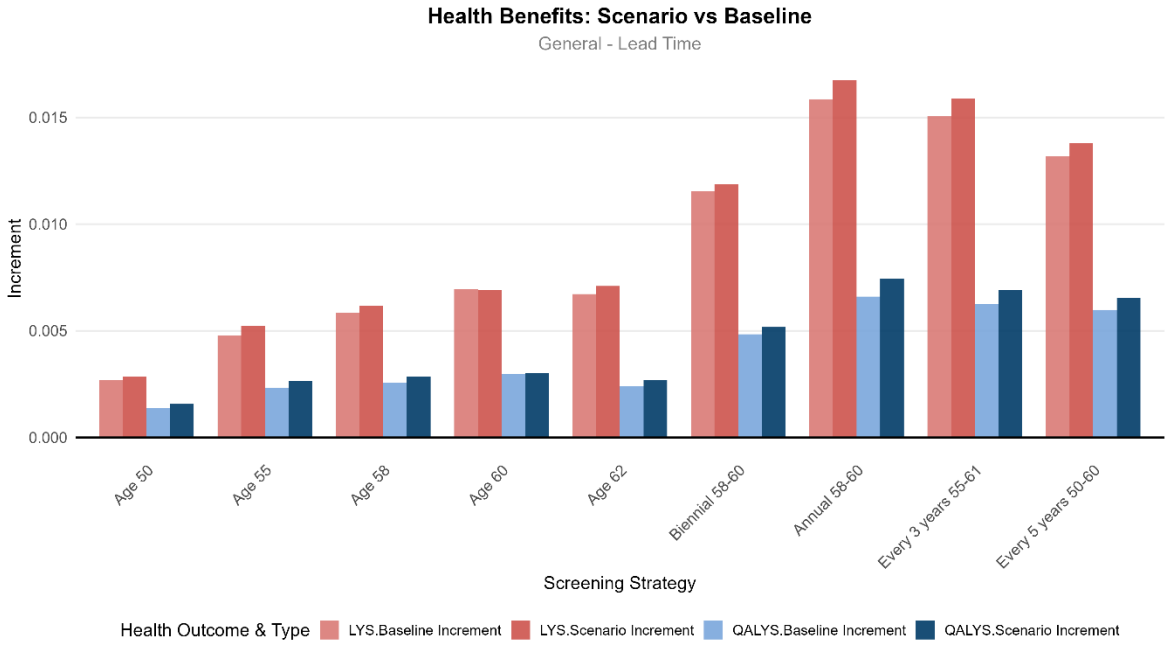
Fig 2b. Black population





3. Patients cannot die from cancer before reaching their symptomatic age (lead time scenario).

Fig 3a General population



4. Mortality extrapolated up to 70-year timeframe (instead of 15 years in the baseline).

Fig 4a General population

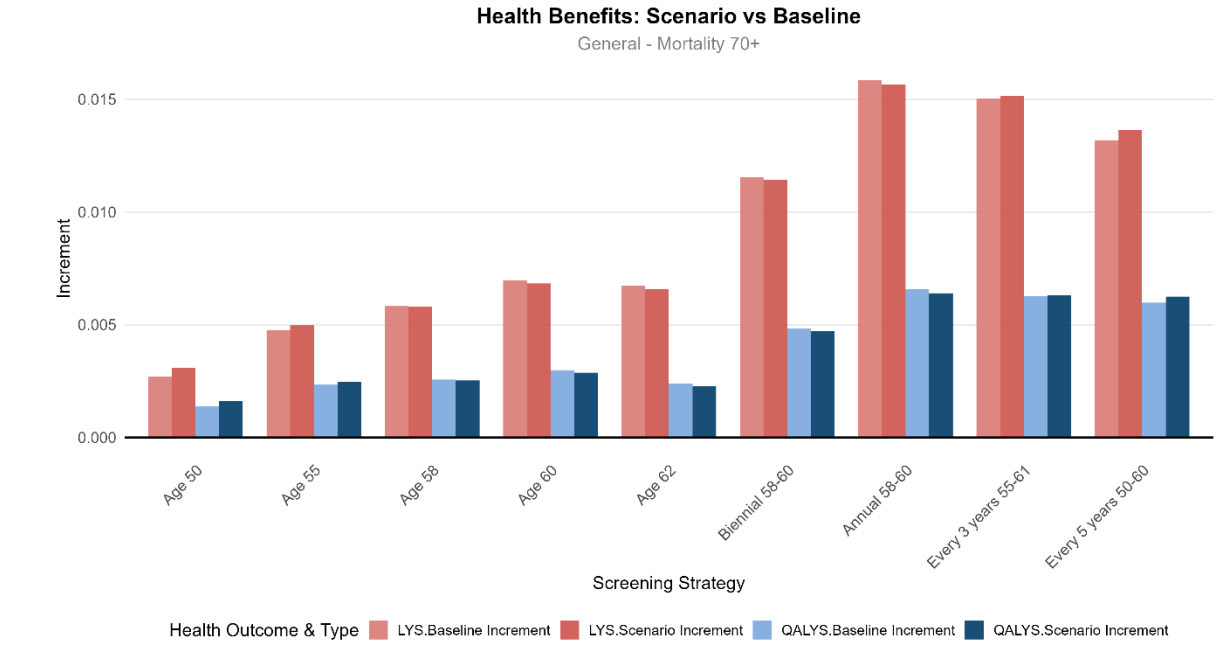
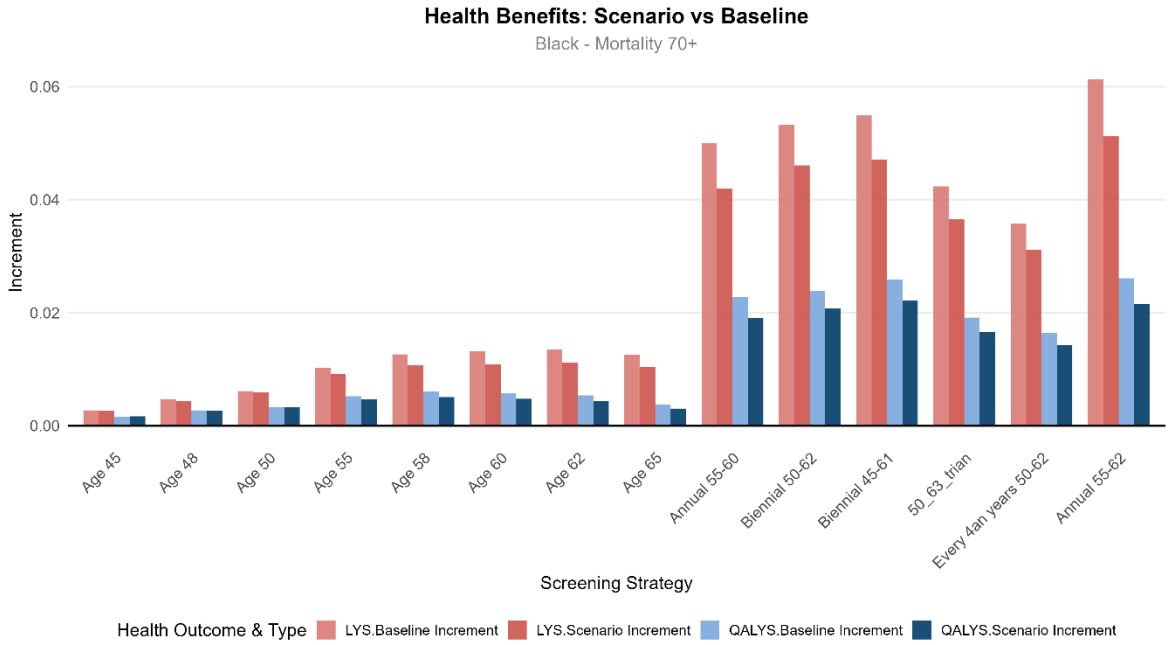


Fig 4b. Black population



5. Perfect uptake of screening and diagnostic testing.

Fig 5a General population

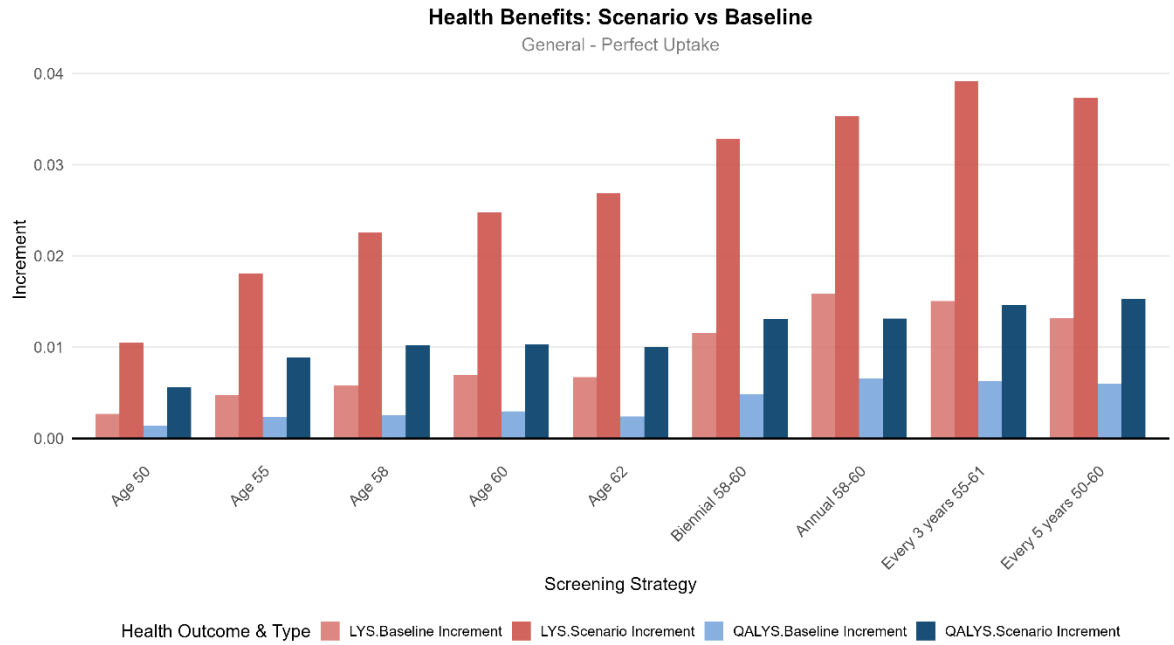
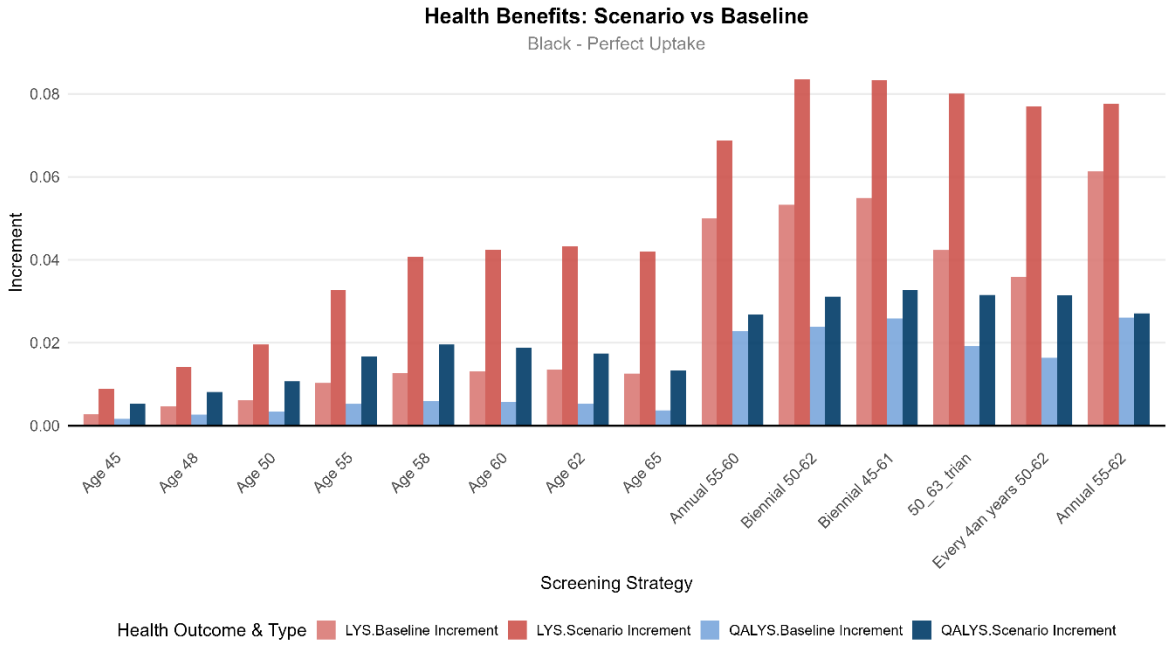


Fig 5b. Black population



6. Sensitivity defined as a single threshold, assuming PSA values do not change with age.

Fig 6a General population

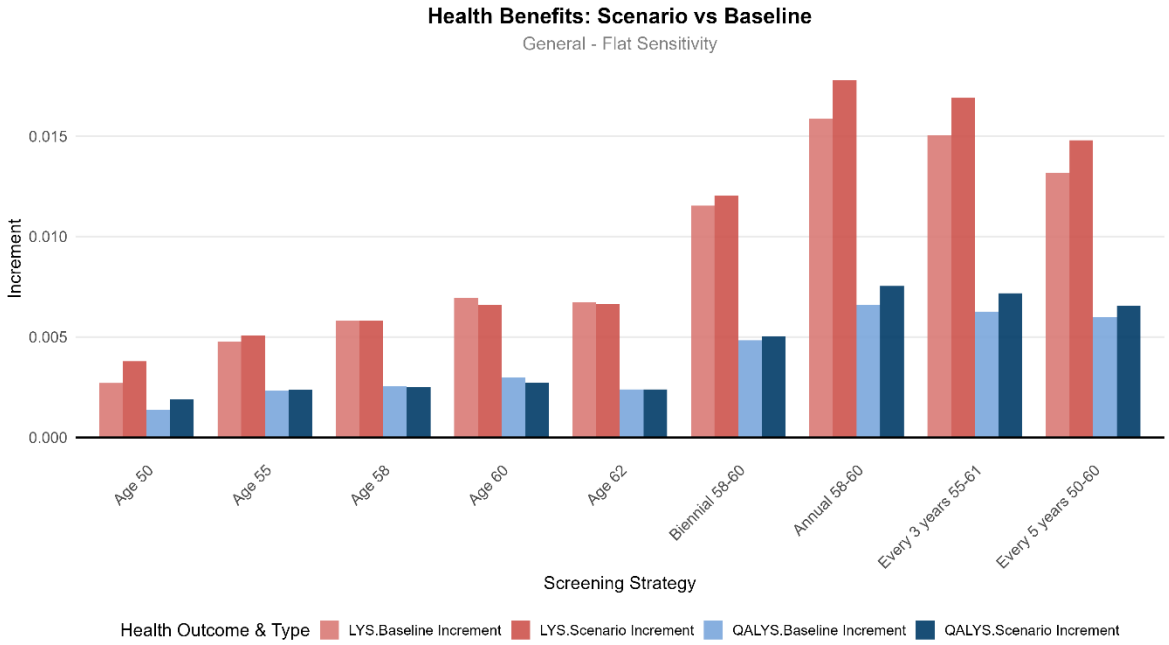
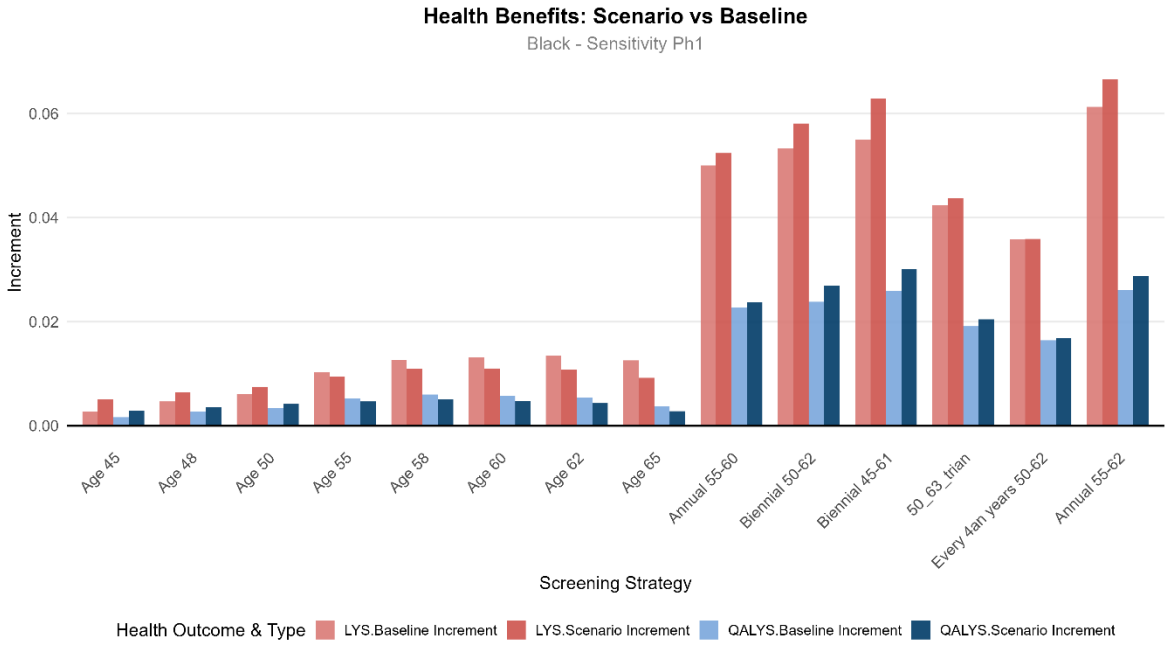


Fig 6b. Black population



7. Sensitivity adjusted by age only for men without cancer.

Fig 7a General population

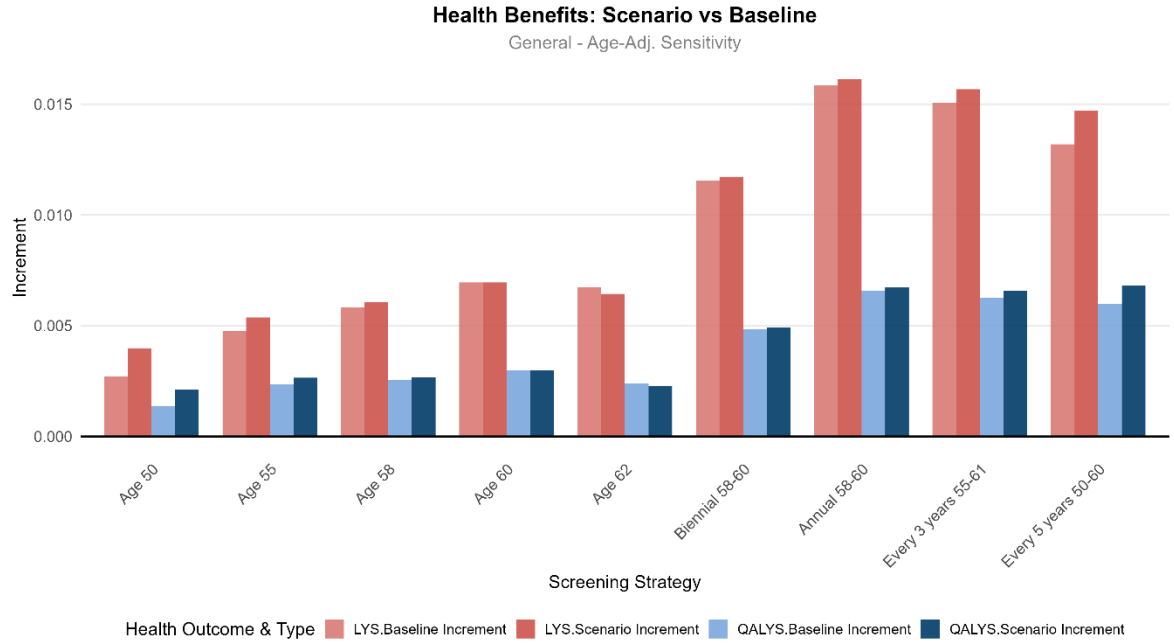
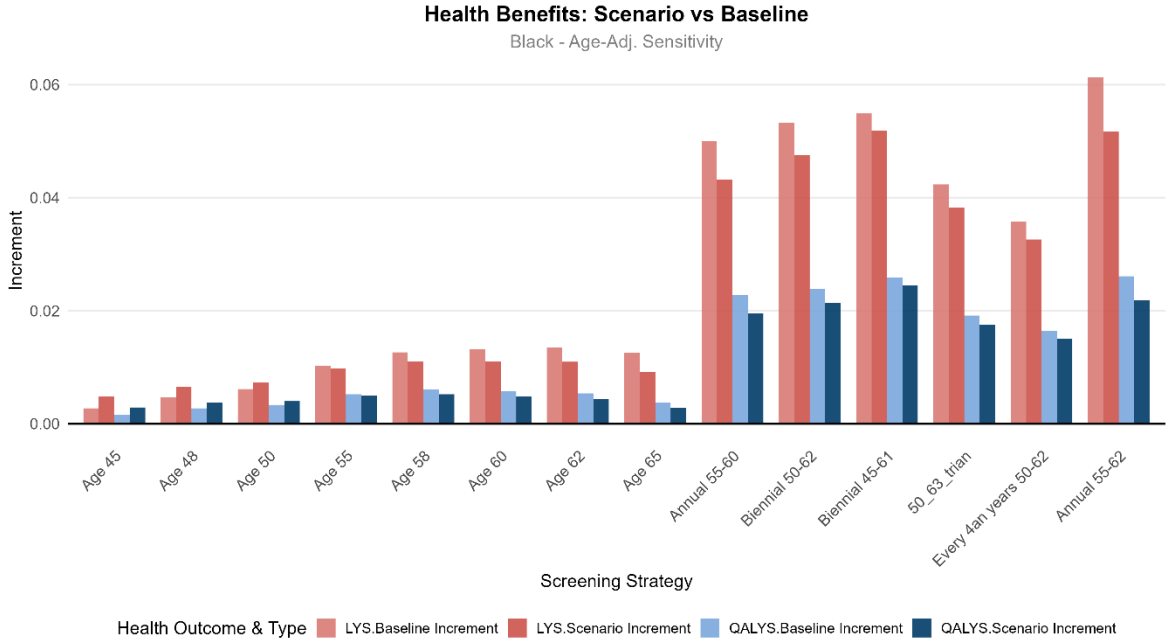


Fig 7b. Black population



8. Assumes a lower health-related utility in the first-year post-diagnosis than in subsequent years.

Fig 8a General population

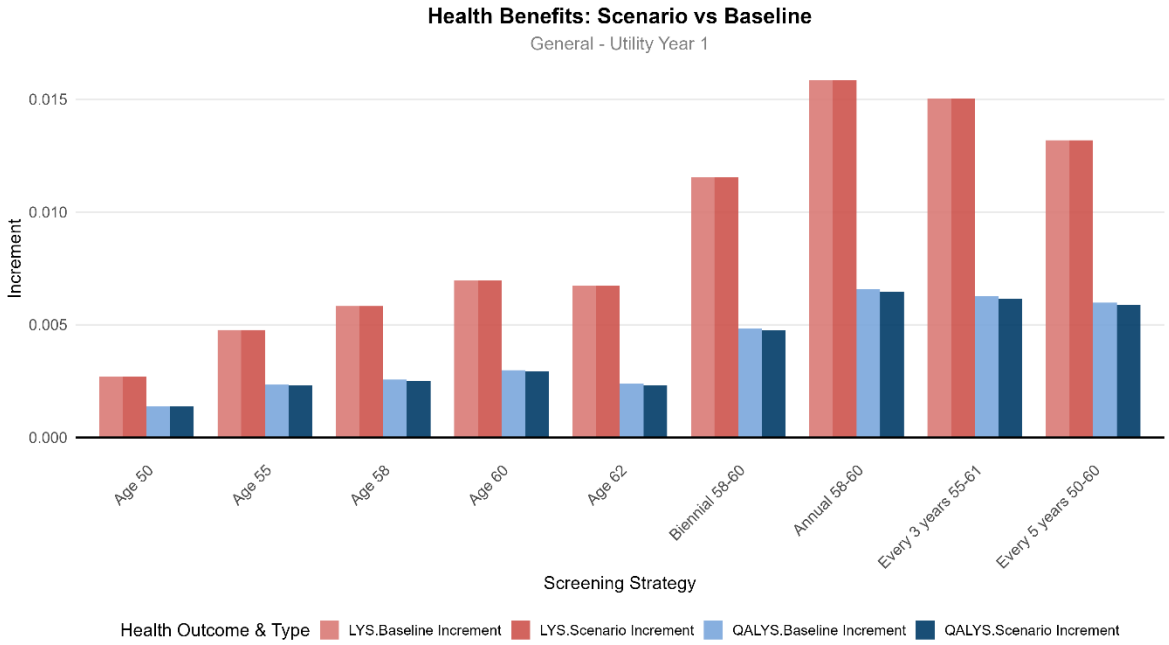
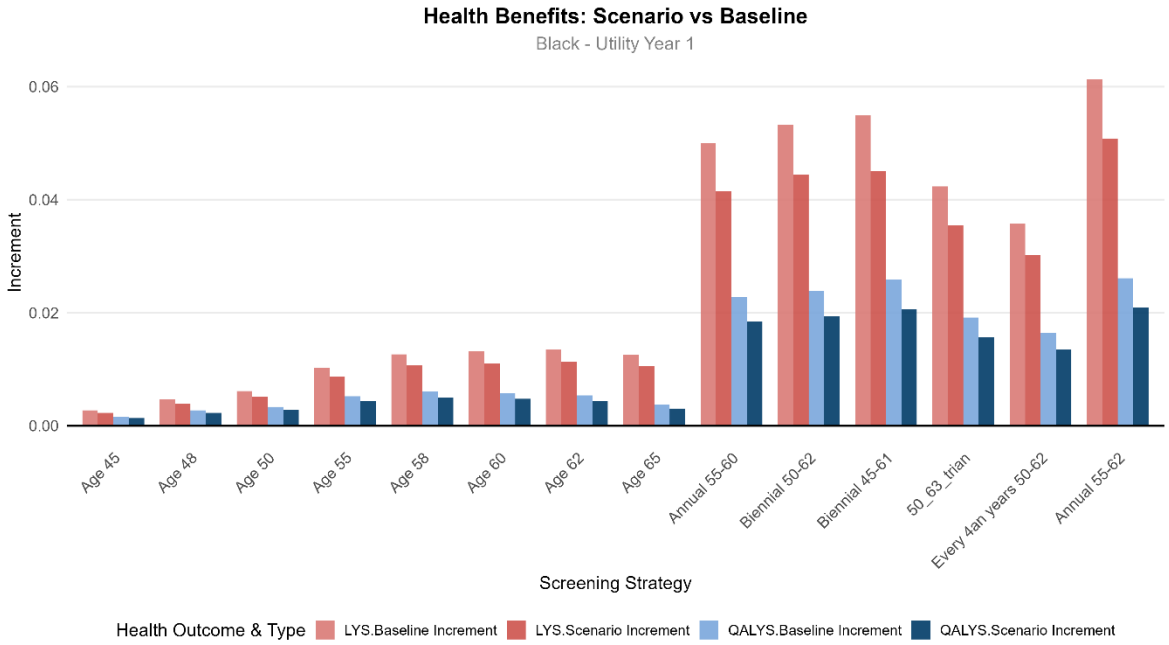


Fig 8b. Black population



9. Scenario with improved fit to observed mortality by adjusting survival data. This was achieved by applying a multiplier of 0.8 to survival rates between ages 20–70 and 0.5 for all other ages across cancer stages (see Figure 12 for the resulting fit).

Fig 9a General population

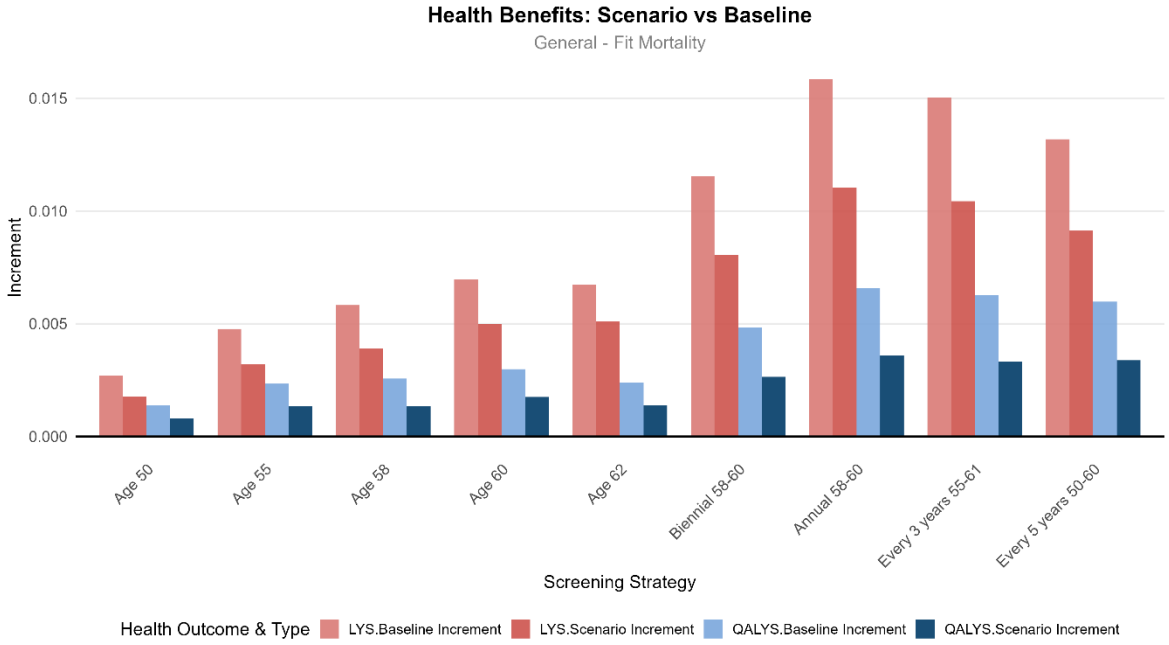
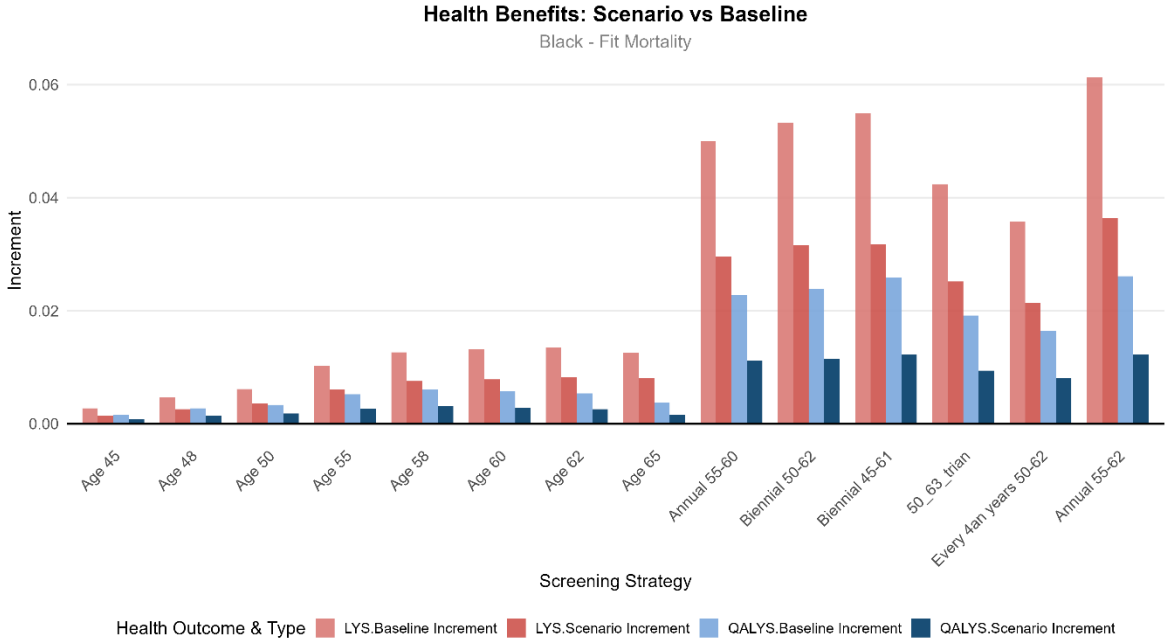


Fig 9b. Black population



10. Discounting as it is in the Green Book

Fig 10a General population

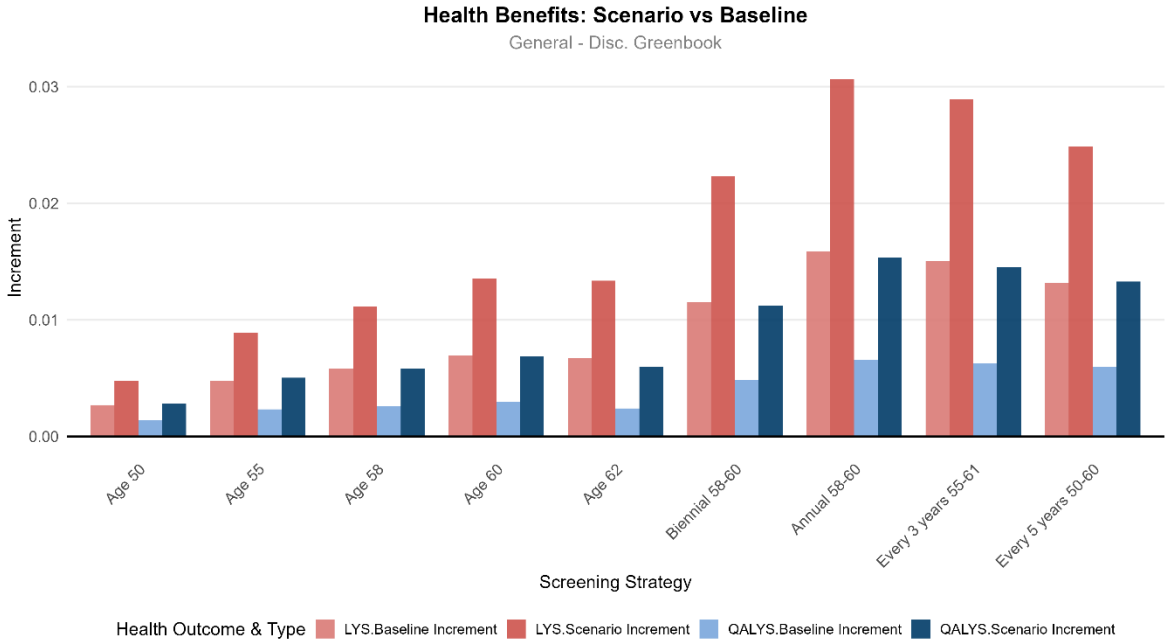
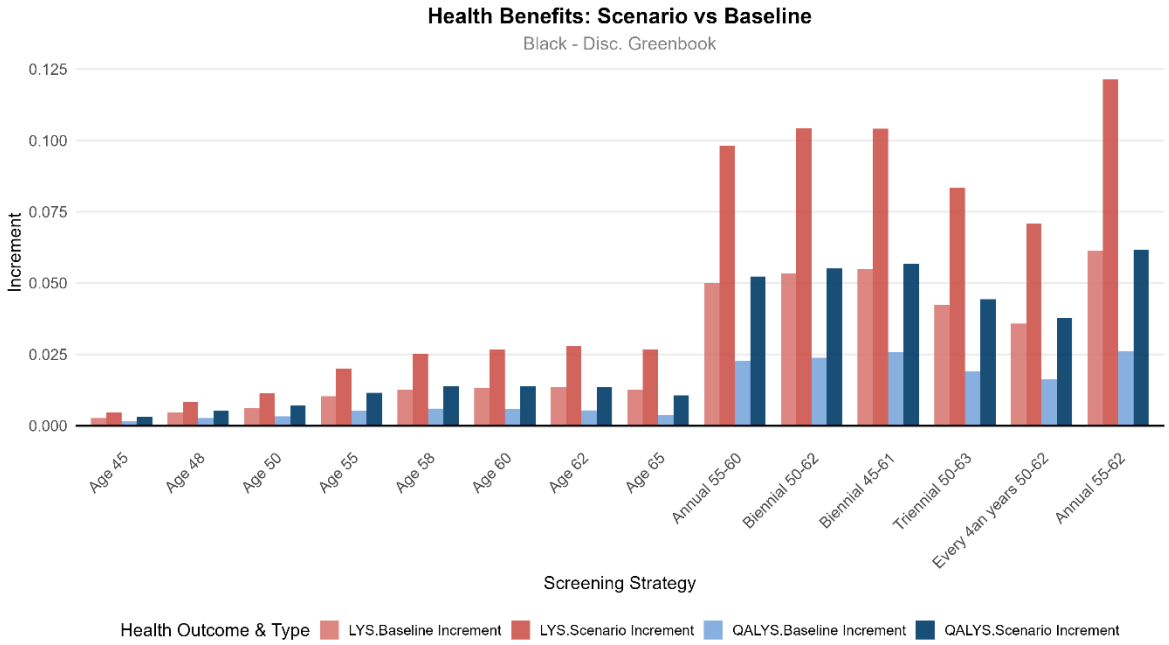


Fig 10b. Black population



Supplementary E. Setting up the model to run base case and deterministic scenarios

E1. Base case scenario

E1.1. Single screening at different ages among general risk population

Step 1: Set Up Parameters in run_model.R

Open the script run_model.R and configure the following parameters for the base-case deterministic analysis. These control the model behaviour, population details, discounting, and simulation settings.

```
run_mode      <- "Deterministic"  # Run mode: "Deterministic" or "PSA"
cl            <- 1            # Cycle length (in years)
d.c           <- 0.035        # Discount rate for costs
d.e           <- 0.035        # Discount rate for effects
n.t           <- 80           # Number of annual cycles
t.dw          <- Age # Start of discounting. Start at the cycle when the first screening occurs
Lead_time     <- 0            # No lead time in base case
mort.limit    <- 15           # Mortality extrapolated over 15 years (longest time possible)
Utility_adjust <- 0       # 0 - no adjustment by year since diagnosis.
pop_run       <- "all"        # Run model for the full population
pop_screen    <- "all"        # Screen the full population
wtp           <- 20000        # Willingness-to-pay threshold (£)
nsample       <- 100000       # Sample size of the population
iter_DA       <- 10           # Number of deterministic iterations
cohort_age    <- 20           # Starting age of the cohort
Fixed_threshold <- 1          # Use a fixed cost-effectiveness threshold
```

To run the model with perfect uptake to see the net (maximum) cost-effectiveness of screening.

These settings define the assumptions for a base-case deterministic run and should be modified as needed for scenario or sensitivity analyses.

Step 2: Set Up and Run Deterministic Scenarios in run_DA.R

In the script run_DA.R, follow the steps below to prepare, run, and store results for multiple deterministic scenarios (e.g., different screening start ages).

a) Clear Existing Results

At the top of the script (around line 6), remove any previously stored results to avoid conflicts or data carryover.

b) Reinitialise Matrices

c) Set Screening Scenario Parameters

Within the loop or in individual runs, set parameters for each scenario:

```
PSA_age      <- age          # Start age of screening
scr_frequency <- 1           # Screening frequency (e.g., annual)
PSA_age_end   <- age         # End age (same as start for one-time screen)
```

These variables control the timing and frequency of the screening intervention.

d) Run the Model for Each Scenario

Make sure model_output matches the structure expected by the results matrix.

e) Save or Export Results

E1.2. Repeat screening at different ages among general risk population

Step 1: Set Up Parameters in run_model.R

Identical to E1.1.

Step 2: Set Up and Run Deterministic Scenarios in run_DA.R

Include the most cost-effective ages in analysis.

To set up repeat screening, set scr_frequency (screening frequency) between 1 (annual) and 5 (once in 5 years).

PSA_age_end should correspond to the age of the last screening.

E1.3. Single and repeat screening at different ages among high-risk population

The set-up is exactly the same as for the general population. Each subgroup needs to be simulated in turns. To do that on the script *run_model.R* set up population to run and population to screen to be the same group. For example, for Black ethnicity, set up:

```
pop_run      <- "Black"      # Run model for the full population
pop_screen   <- "Black"      # Screen the full population
```

All the other script set-up should be similar to E1.1. and E1.2.

E1.4. Scenarios

For general and high-risk sub-groups run the following scenarios:

N	Scenario	Set up
1	The model considers that population in screening arm will not die from cancer before reaching their age of symptomatic (or opportunistic screening) diagnosis in no organised screening scenario	Lead_time==1
2	Mortality from prostate cancer is extrapolated over 70 years (15 years in the base case)	mort.limit <-70
	PSA values change by age only in men without cancer	
	Survival in the model is readjusted for stage 3,4 cancers to fit mortality data	Apply a relative risk to the survival curves to fit mortality data
3	Utilities are assumed to be lower in the first year since diagnosis	Utility_adjust==1

4	Higher (5% and 10%) discounting for costs and effects	d.c <- 0.05 # Set discounting for costs
		d.e <- 0.05 # Set discounting for effects

5	Higher (5% and 10%) discounting for effects only	d.c <- 0.035 # Set discounting for costs
		d.e <- 0.05 # Set discounting for effects

6	Impact of the start for discounting	t.dw set to 1 and run for all comparators; t.dw set to the cycle of screening implementation for the screening age when screening changes from cost-effective to not cost effective
	The Green Book (2022) discounting	Declining long term discounting rate for costs and health as in green book https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government/the-green-book-2020#a6-discounting
7	Impact of uptake	Replace uptake parameters with the perfect uptake (param_space$Uptake[ ] ==1)
8		

Supplementary F. Description of the population parameters

The population used in this analysis is derived from the Health Survey for England (HSE) 2018 and 2019 data, as described in the main report.

The initial population includes the following variables:

Seriala: Unique identifier for each individual.

age_0: Age at the start of the model (20 years).

age: Current age of the individual.

imd: Index of Multiple Deprivation, indicating socioeconomic status (1 to 5 quintiles).

ethnic: Ethnic group of the individual, where Black == 2, Asian == 3

EQ5D: A standardised measure of health-related quality of life.

weighting: Survey weights to ensure the sample is representative of the general population.

Additional Variables and Risk Calculations complementing HSE data:

The initial HSE data are augmented with several derived variables and risk calculations through R scripts, focusing on genetic and familial cancer risks.

Lifetime risks and probabilities

Genetic Risk Factors:

i.BRCA1.status: Binary variable (0 or 1) indicating the presence of a BRCA1 mutation, sampled based on a predefined probability (p.BRCA1).

i.BRCA2.status: Binary variable (0 or 1) indicating the presence of a BRCA2 mutation, sampled based on a predefined probability (p.BRCA2).

Familial Cancer History:

Probabilities for familial history of specific cancers are calculated for each individual, taking into account their BRCA status:

p.i.Famil.BC: Probability of familial history of breast cancer.

p.i.Famil.OC: Probability of familial history of ovarian cancer.

p.i.Famil.PC: Probability of familial history of PC.

Based on these probabilities, binary indicators for familial history are then sampled:

i.Famil.BC: Binary variable (0 or 1) indicating familial history of breast cancer.

i.Famil.OC: Binary variable (0 or 1) indicating familial history of ovarian cancer.

i.Famil.PC: Binary variable (0 or 1) indicating familial history of PC.

Annual prostate cancer Risk:

p.i.PC: Individual annual probability of prostate cancer onset. This risk is calculated using relative risks associated with various attributes, including BRCA1 status, BRCA2 status, familial history of breast, ovarian, and PCs, and ethnic background (specifically Black and Asian ethnicity). The calculation adjusts for the population mean of these attributes.

Population Sampling and Subgroup Selection

The R scripts also include functionalities to sample and subset the population based on specific criteria:

f.pop.set: This function samples individuals from the population to achieve a defined sample size (nsample) and assigns a starting age (cohort_age) to all individuals in the sampled population. It also reassigns a unique PID (Personal Identifier) to each person.

f.resample.pop.risk: This function allows for resampling of specific risk groups from the population, enabling analyses focused on particular cohorts. The available risk groups are:

"Family": Individuals with a familial history of breast, ovarian, or prostate cancer (i.Famil.BC=1 or i.Famil.OC=1 or i.Famil.PC=1).

"BRCA1,2": Individuals with a BRCA1 or BRCA2 mutation (i.BRCA1.status=1 or i.BRCA2.status=1).

"Black": Individuals identified as ethnically Black (ethnic=2).

"high_risk": A broader high-risk group including individuals with familial cancer history, BRCA mutations, or Black ethnicity.

"all": The entire population without specific risk-group filtering.

Supplementary G. Description of the modelling outcomes

All the outcomes are calculated per weighted individual at the model start. This means that the outcomes recorded during each cycle are divided by the population weight. For deterministic runs, the cumulative results are calculated over the lifetime and averaged across multiple stochastic deterministic cycles. For probabilistic analysis, the lifetime results are saved for each PSA loop.

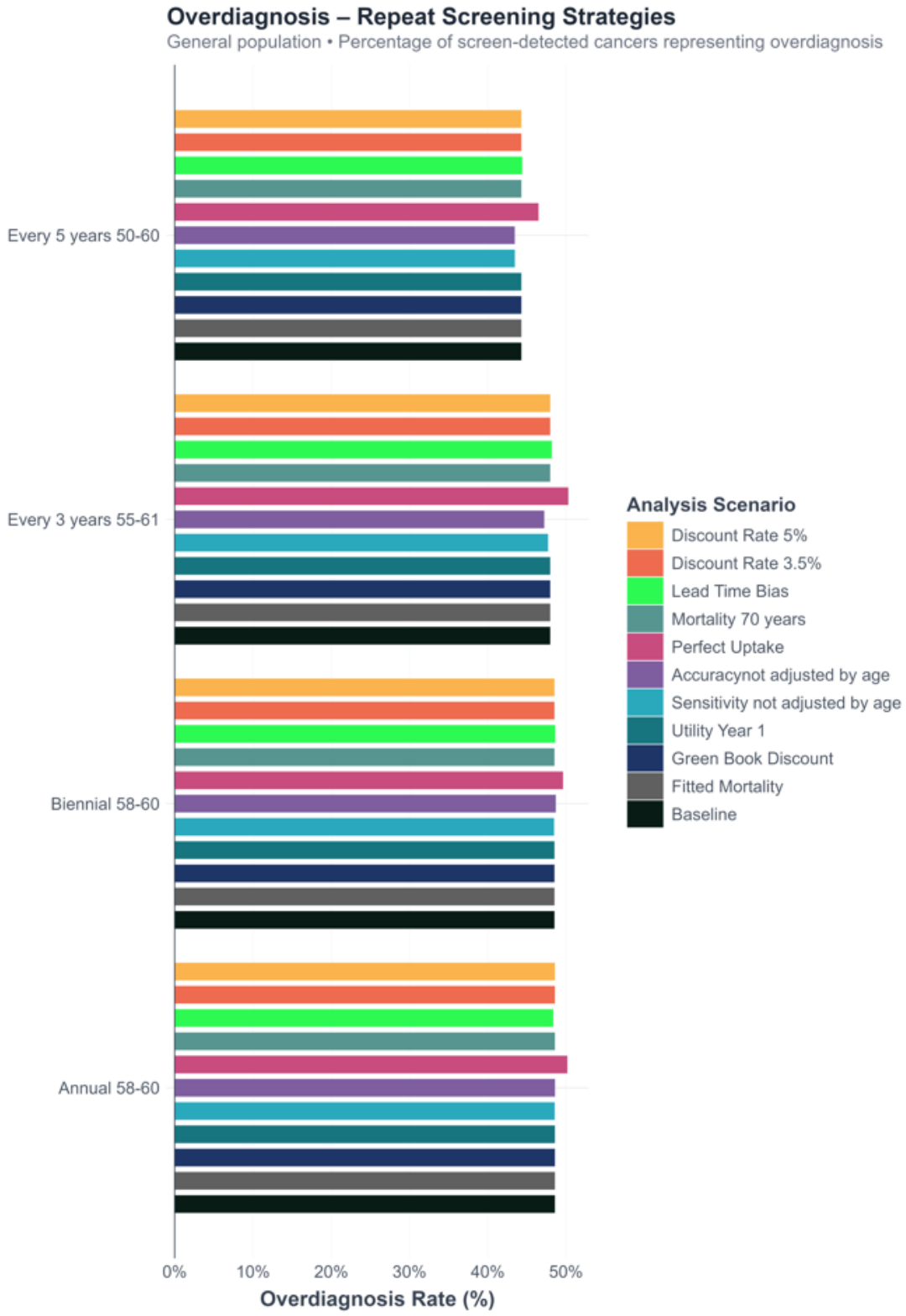
The description of the outcomes is reported in the table below.

N	Outcome in the model	Description
1	TOTAL_COSTS	Total costs (sum of treatment, diagnostic, and screening costs).
2	Cancer_COSTS	Total weighted treatment costs. The palliative care costs are included in treatment costs.
3	DIAG_COSTS	Total weighted diagnostic costs for following up screen positive cases. The diagnostic costs for symptomatic patients are included into the Cancer_COSTS
4	SCREEN_COSTS	Total weighted screening costs, includes the costs for PSA invitations and PSA tests
5	LYS	Total weighted LYS (discounted)
6	QALYS	Total weighted QALYs (discounted)
7	LYS_n.d	Total weighted LYS (undiscounted)
8	QALYS_n.d	Total weighted QALYs (undiscounted)
9	St1_SYMPT	Weighed cumulative proportion of population diagnosed with Stage 1 symptomatic cancer
10	St2_SYMPT	Weighed cumulative proportion of population diagnosed with Stage 2symptomatic cancer
11	St3_SYMPT	Weighed cumulative proportion of population diagnosed with Stage 3 symptomatic cancer
12	St4_SYMPT	Weighed cumulative proportion of population diagnosed with Stage 4 symptomatic cancer
13	ALL_incidence	Weighed cumulative proportion of population diagnosed with cancer

14	St1_SCRN	Weighed cumulative proportion of population diagnosed with Stage 1 screen-detected cancer
15	St2_SCRN	Weighed cumulative proportion of population diagnosed with Stage screen-detected cancer
16	St3_SCRN	Weighed cumulative proportion of population diagnosed with Stage 3 screen-detected cancer
17	St4_SCRN	Weighed cumulative proportion of population diagnosed with Stage 4 screen-detected cancer
18	St1_MORT	Weighed cumulative proportion of population diagnosed with Stage 1 cancer and died from cancer
19	St2_MORT	Weighed cumulative proportion of population diagnosed with Stage 2 cancer and died from cancer
20	St3_MORT	Weighed cumulative proportion of population diagnosed with Stage 3 cancer and died from cancer
21	St4_MORT	Weighed cumulative proportion of population diagnosed with Stage 4 cancer and died from cancer
22	MORT	Weighed cumulative proportion of population died from cancer
23	GGG1_St1_2	Weighed cumulative proportion of population diagnosed with Stage 1 or 2 cancer having GGG1
24	GGG2_St1_2	Weighed cumulative proportion of population diagnosed with Stage 1 or 2 cancer having GGG2
25	GGG3_St1_2	Weighed cumulative proportion of population diagnosed with Stage 1 or 2 cancer having GGG3
26	GGG45_St1_2	Weighed cumulative proportion of population diagnosed with Stage 1 or 2 cancer having GGG4,5
27	GGG1_St3_4	Weighed cumulative proportion of population diagnosed with Stage 3 or 4cancer having GGG1
28	GGG2_St3_4	Weighed cumulative proportion of population diagnosed with Stage 3 or 4cancer having GGG2
29	GGG3_St3_4	Weighed cumulative proportion of population diagnosed with Stage 3 or 4cancer having GGG3,4
30	GGG45_St3_4	Weighed cumulative proportion of population diagnosed with Stage 3 or 4cancer having GGG4,5

31	Invite_PSA	Weighed cumulative proportion of population invited to PSA screening
32	PSA	Weighed cumulative proportion of population who attended PSA screening
33	Positive_PSA	Weighed cumulative proportion of population who were screen-positive
34	Invite_MP-MRI	Weighed cumulative proportion of population invited to PSA screening
35	MP-MRI	Weighed cumulative proportion of population who attended MP-MRI to follow up screen positive test
36	Positive_MP-MRI	Weighed cumulative proportion of population who were MP-MRI-positive
37	Invite_Biopsy	Weighed cumulative proportion of population invited to biopsy follow up
38	Biopsy	Weighed cumulative proportion of population who attended biopsy
39	Positive_Biopsy	Weighed cumulative proportion of population who were biopsy-positive
40	FP	False positive cases (all with positive biopsy and no cancer)
41	FN	False negative cases (all who were tested negatively at least in one of the diagnostic pathway stages biopsies and have cancer)
42	PC	True positive prostate cancer cases

Supplementary H. Predictions of overdiagnosis in scenarios



Supplementary I. Summary of stakeholder meetings

Phase 1 Stakeholder meeting

28/05/2024 Stakeholder Attendees: Dr Helen Hanson (geneticist), Dr Jim Catto (clinician), Natalia Norori (Prostate Cancer UK), Amy Rylance (Prostate Cancer UK)

31/06/2024 Stakeholder Attendees: Dr Bill Cross (clinician), Dr Hashim Ahmed (clinician)

1. Glossary of names:

Helen Hanson: HH

Jim Catto: JC

Natalia Norori: NN

Amy Rylance: AR

Bill (William) Cross: BC

Hashim Ahmed: HA

Olena Mandrik: OM

This document presents stakeholder comments. Please refer to the shared PowerPoint slides for more context on each topic. Any major changes made since the meetings have also been highlighted here.

2. Natural history model structure

All stakeholders agreed with the general model structure

3. Modelling prostate cancer risk

HH: most genetic testing is done on the basis of cancer history, so the prevalence of BRCA1, 2 source we are using might be higher than in reality.

NN: concerned about the quality of ethnicity data from HSE due to mixed black and black patients sometimes being misreported.

SCHARR changes to the modelling plan: SCHARR made no changes to the modelling plan, as the HSE includes population weights. Therefore, even if certain cohorts are underrepresented, the population weights should address any issues.

4. Disease onset assumptions

HH: having both BRCA1,2 mutations do not increase your probability of cancer. If you have both mutations, you should be assigned the highest risk of the two.

SCHARR changes to the modelling plan: having both mutations does not increase ones probability of cancer with the highest RR from BRCA1 and BRCA2 risks assigned.

5. Prostate cancer incidence rate

OM: suggestion is to use the age trend until age 79, then the average incidence for 79 to 100.

There was no objection to this.

AR, NN, JC: increase in cancer rate in older ages but decrease in clinical diagnosis due to the differences in testing by age.

BC & HA: both recommend using 2018 data instead of 2020 data as this is not contaminated by Covid.

SCHARR changes to the modelling plan:

1) Incidence data for 2018 is going to be used in the model calibration.
2) The incidence by age is going to be used till age 79 years old.
3) The average incidence is going to be used for those who are 79-100 years old.
4) The model will assume an increase in cancer onset by age and decrease in clinical diagnosis from age 75 years.

6. Stage distribution

BC & HA: concerned about the age of this data as it is from 2014. Suggested alternative sources from the National Prostate Cancer Audit and Caroline Moore at UCL.

NN: would expect that unknown cancers are distributed equally across stages, based on data from Wales (to be shared), where they saw a decrease in the number of unknown cases, and no one stage saw an increase as a result.

SCHARR changes to the modelling plan: Unknown cancers will be equally re-allocated across stages 1-4.

7. Gleason grade group distribution by stage and age

AR, JC & NN: were all concerned about the quality of the 2020 data for Gleason Grade Group (GGG) distribution due to Covid.

Note: The originally presented data on the slide represented GGG distribution with age adjustments (i.e. subtracting sojourn time reported in CAP from the age of incidence). This resulted in high proportion of patients with unclassified GGG in younger patients.

HA & BC: GGG 1 percentages will be much lower for locally advanced. GGG1 in locally advanced cancer would be lower than 16%.

BC: suggested using only the data for 60-79 year olds (average), with 1-3% of GGG1 in stage 3 PC, and for metastatic cancer, GGG 1 should be 0.

BC & HA: in the absence of data, it should be ok to assume that the distribution of the GGG is the same among those aged 80+ as in those aged 75-79.

SCHARR changes to the modelling plan:

1) The ages used in the GGG onset will be the same as the GGG at diagnosis, considering that the model will assume no changes in the GGG during undiagnosed period. However, 2020 data will still be used as no earlier sources with such detailing of GGG distribution by age was found.
2) As initially suggested, two targets for GGG by stage and age will be used informed by the Get Data Out (2023):
    a. Localised stage: Stages I and II
    b. Advanced stages: Stages III and IV

This data for the calibration target will be modified before using the following way:

a. Stages 1,2: Average proportion of patients in GGG at time of diagnosis in localised stage for ages 30-59 and 60 and above will be used.

   Stages 3,4: Average proportion of patients in GGG at time of diagnosis in locally advanced stage for ages 60-79 will be used for all ages. For stage 4, it will be assumed that no patients have GGG 1.

8. Progression of undiagnosed cancer

JC: considers using the general sojourn time to be a suitable substitute.

AR: more confident in the Gleason Grade Group being a predictor of progression that age.

HA & BC: do not feel like they can comment whether prostate cancer progresses slower in older age groups. They suggested to check with Caroline Moore for potential data on this.

HA: we might be challenged on not modelling grade change over time.

BC: suggested looking at the active surveillance data to inform the change in progression for the Gleason Grade Group (GGG).

Note:

Patients on active surveillance have change in the GGG, with Richard et al (2020) for example reporting upgrading of GGG in 16% of patients over one year (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7654679/). However, changing the assumption on constant GGG will likely not decrease uncertainty due to the following reasons:

(a) Patients on active surveillance are mainly symptomatically diagnosed patients. This means that the change in GGG for these patients will likely be much quicker than for asymptomatic patients. For example, Bloom at all report that negative fusion biopsy is associated with a reduced risk of Grade Group progression (HR 0.41, 95% CI 0.22-0.77, p <0.01) (https://www.auajournals.org/doi/full/10.1016/j.juro.2018.07.051).

(b) Modification of the GGG based on symptomatic patient data will also assume re-calculating the GGG at time of onset considering the sojourn

time. This could create additional uncertainty due to small number of patients and a lot of missing values in older age groups.

SCHARR changes to the modelling plan:

1) We decided not to change the assumptions on constant GGG from time of onset to diagnosis in the GGG.
2) We will use sojourn time for 50–69-year-old reported in the CAP trial as the calibration target to inform the model.

9. Disease progression assumptions

No comments from stakeholders

10. Survival

No comments from stakeholders

11. Prostate cancer diagnostic pathway

No comments from stakeholder

12. Accuracy of PSA test

All stakeholders concerned about the use of 4 ng/ml as the threshold rather than 3 ng/ml for the PSA test. A number of alternative sources were suggested: NICE published guidelines, CAP data, IMPACT study.

AR: mentioned that age-specific thresholds might be used in symptomatic cases. Whereas the 3ng threshold would be applied for non-symptomatic cases. Suggested that for screening we can use one threshold of 3ng/ml rather than age-specific thresholds.

HA & BC: agree with using the threshold of 3ng/ml.

HA: suggests using Prostagram study (UK population) as there is some sensitivity data for threshold 3, but the study has a small population size of 410 patients.

SCHARR changes to the modelling plan:

Before the second meeting, we adjusted the PSA threshold values to 3ng/ml, which we will continue to use. Considering no data, we will use the sensitivity of 4ng/ml for Stage 4 patients, retrieved from the symptomatic cohort assuming that

stage 4 patients would become symptomatic within one year and that there would not be many patients with the PSA level between 3-4ng/ml at this stage.

13. Accuracy of Multiparametric magnetic resonance imaging (MP-MRI) & Transrectal Ultrasound Scan (TRUS) biopsy

AR: there has been a shift towards an increased use of transperineal biopsies, instead of TRUS biopsy. However, currently more TRUS biopsies are still being used. Suggests incorporating transperineal biopsies as well to better reflect future diagnostic pathways.

HA & BC: a lot more people using transperineal biopsy now, so also agree that we should change to use this.

HA: NICE diagnostic review of transperineal biopsy brought together all the sources on the sensitivities of this. They suggest that transperineal biopsies have a higher accuracy but not clear how much higher.

SCHARR changes to the modelling plan:

Base case will include using the transperineal biopsy instead of the TRUS biopsy.

14. Utility multipliers

JC: agreed that generally stage 1 & 2 have little change in the HRQoL.

AR: would have expected a larger fall in HRQoL for those in stage 3. Source of assumption is not known.

OM: suggested 2 scenarios with utilities: prostate cancer specific utility multipliers in the base case and general cancer utilities in the scenario.

BC: 5-7 days for biopsy harms seems reasonable, especially transperineal has fewer harms.

HA: RCTs in the US for TRUS and transperineal biopsies can provide some data on utilities related to the biopsy.

SCHARR changes to the modelling plan: No changes. prostate cancer specific utility multipliers will be used in the base case.

15. Diagnostic costs

AR: more detail is needed on the uptake of these diagnostic procedures.

HA: there is cost data in the NICE diagnostic assessment for transperineal biopsy published 2-3 years ago which should have recent costs.

16. Treatment costs

AR, NN & JC: more up-to-date sources are needed for costs due to changes in possible treatments. National Prostate Cancer Audit for treatment allocation was suggested as a source.

AR: there are new novel hormonal treatments for stage 4 and even stage 3 that people will continue to receive beyond the first year.

JC: ProtecT trial, in particular, is out of date because it does not include multi-model therapies or hormonals.

HA & BC: Extrapolation of standard treatment (surgery, radiotherapy) costs is reasonable. Additional systemic therapy is very rarely prescribed to Stage 1 or 2 patients. The systemic therapy is assigned to newly diagnosed patients. Up to 60% of late-stage patients can get target treatment. National prostate cancer audit data could be used.

17. Surveillance & palliative care costs

JC: people tend to receive active surveillance for their entire lifetime after diagnosis (50%), or they are treated within 5 years (50%).

BC: surveillance should be distinguished between active surveillance and the surveillance after the radical treatment. The latter one will be less costly. After the radical treatment patients are followed mainly remotely – 6 months PSA for 3 years and then annual PSA for the lifetime.

SCHARR changes to the modelling plan on costs:

1. We will use 5 types of costs:

- Standard treatment costs (surgery, SACT, radiotherapy)

- Systemic Therapy (stage 3, 4, according to the National Prostate Cancer Audit)
- Active surveillance costs (those who have no treatment)
- Surveillance costs (all alive patients until they relapse)
- Palliative care costs.

2. We will recalculate the average annual costs assuming that the patients who didn't receive standard treatment are on active surveillance (Stage 1-37.8%, Stage 2-68.6%, Stage 3- 76.6%, stage 4 – 50.4%). The active surveillance will be costed up to patient's death from other causes or the year of cancer death (and so the assumed cancer progression or relapse). The standard treatment costs will be extrapolated over the time using ProtecT data. The surveillance costs will be added from year 2, assuming the proportion of patients who are not on active surveillance. The systemic therapy costs will be added to a proportion of patients who are in Stages 3,4 according to the Prostate Cancer Audit.
3. Palliative care costs will be added to each patient dying from cancer to reflect the last -year/ relapse costs.

Phase 2 Stakeholder meeting with Derek Rosario (1)

Meeting summary 20/05/2025

Present: Lena Mandrik, Dan Pollard, Annabel Rayner, Jessica E Forsyth, Maria Hanini; ROSARIO, Derek (SHEFFIELD TEACHING HOSPITALS NHS FOUNDATION TRUST)

Presented:

- General modelling assumptions
- Model structure (NHD)
- Modelling prostate cancer risk for onset and progression
- Risk factors
- Correlations between risk factors
- Modelling process (order of events in the model)
- Summary of model changes (from Phase 1 to Phase 2)

Suggestions Received:

- No objections to the model structure or general assumptions.
- Agreement that incorporation of correlations for prevalence of BRCA1/2 status, family history, and ethnicity is limited by data availability.
- Agreed with updated data on prostate cancer distribution by age and stage, based on NHS Digital (2021) and CMA Stage (2013–2021).
- Agreed with revised survival data for stages 1 and 2.
- It was noted that mortality in ProtecT may be underestimated due to underrepresentation of higher-grade group (GGG) cancers. It was suggested to consider PCPT for mortality estimates.
    - **Note**: JF reviewed the PCPT publication, which only reports overall all-cause mortality and therefore cannot be used to inform net prostate cancer survival by stage.
- **Survival extrapolation**: It was suggested to conduct a scenario analysis (or consider as a base case) using 15-year cancer-specific mortality (i.e. assuming individuals who survive 15 years post-diagnosis are assumed to die from other causes).
- Agreed with GGG distribution data at onset.

- Agreed with lifetime cost extrapolation and inclusion of palliative care costs for both prostate cancer-related and other-cause mortality.
- Agreed with updates to the referent population for EQ-5D based on the EEPRU report (Alava, 2022). It was suggested to apply a flat multiplier for year 1 and subsequent years.
- Agreed that PSA sensitivity can be based on PSA allocation, and that PSA values vary with age. It was suggested that the base case should account for PSA variation in both cancer and non-cancer states.
- Agreed with BRCA knowledge status estimates. It was acknowledged that these are likely underestimated due to not accounting for ongoing and future uptake of genetic testing. It was noted that due to the small size of the BRCA population, the overall impact is likely negligible.
- Agreed with the updated simulated population approach (using HSE 2018 and 2019 data, and starting simulations from age 20).
- **Calibration**: It was suggested that calibration results indicate a good model fit. It was proposed that the survival scenario limited to 15 years post-diagnosis could be considered for the base case.
- **Screening scenarios**:
    - It was suggested that the current NICE thresholds may lack strong scientific justification.
    - It was proposed that the BRCA population does not need to be evaluated separately.
    - Recommended screening scenarios to evaluate:
        - Fixed threshold (3 ng/ml) screening for both average- and high-risk group
        - Lower threshold (0.75 or 1.5) screening for high-risk populations (Black ethnicity, BRCA carriers, family history
        - Risk-stratified screening using NICE thresholds

Phase 2 Stakeholder meeting with Derek Rosario (2)

Meeting Summary - Clinical Detection Testing Results Review

Date: 04/07/2025

Present: Lena Mandrik, Jessica E Forsyth, Maria Hanini; ROSARIO, Derek (SHEFFIELD TEACHING HOSPITALS NHS FOUNDATION TRUST)

Presented:

- Model validation against CAP study data
- Initial model results

Suggestions Received:

- **Model validation:** Agreed that CAP study methodology presents significant limitations for validation purposes.
- Noted that clinical consensus indicates fundamental flaws in CAP methodology, with strong recommendation that alternative validation targets should be prioritised.
- **Natural history modelling:** Agreed on need for better research to inform progression rate and other parts of the natural history. Strong emphasis on data limitations, particularly absence of screening data for older men (70-80 years), which limits the modelling.
- **Screening age recommendations:** No objections to age-based screening logic.
- **Overdiagnosis calculations:** Informed that the clinical definition includes cases that would not result in mortality. Suggested using screen-detected cases as denominator rather than all diagnosed cases, referencing PROTECT study data showing 9 out of 10 diagnosed but untreated cases die from other causes.
- **PSA kinetics integration:** Strong support for incorporating PSA trajectory analysis. Suggested early baseline testing (ages 40, 45, 50) could inform risk-stratified screening intervals based on PSA kinetics rather than age alone. PSA at age 40 predicts trajectory with three distinct patterns: flat curves, slow age-related increases, and exponential rises.
- Agreement that higher baseline PSA levels justify more frequent screening intervals for cost-effectiveness. It was noted that while PSA density show

moderate increases in diagnosis, they don't provide clear information about outcomes or aggressive cancer identification.

- **Ethnic disparities:** Suggested that prostate cancer is more prevalent in Men of Black ethnicity but not more aggressive; evidence suggests twice the incidence and mortality rates but identical mortality to incidence ratio for Black and non-Men of Black ethnicity.
- **Uptake:** Suggested testing 100% MRI uptake scenarios (realistic uptake could reach ~98%).
- **Research gaps**: Comprehensive discussion of limitations including lack of screening data for men aged 70-80, insufficient natural history data, need for better PSA kinetics models, and limited understanding of aggressive cancer identification.

References

1. Squires, H., et al., *A Framework for Developing the Structure of Public Health Economic Models.* Value Health, 2016. **19**(5): p. 588–601.
2. *Health Survey for England 2018*, N. Digital, Editor. 2018.
3. Perez-Cornago, A., et al., *Prospective investigation of risk factors for prostate cancer in the UK Biobank cohort study.* British Journal of Cancer, 2017. **117**(10): p. 1562–1571.
4. Maxwell, K.N., et al., *Population Frequency of Germline BRCA1/2 Mutations.* Journal of Clinical Oncology, 2016. **34**(34): p. 4183–4185.
5. Nyberg, T., M. Tischkowitz, and A.C. Antoniou, *BRCA1 and BRCA2 pathogenic variants and prostate cancer risk: systematic review and meta-analysis.* British Journal of Cancer, 2022. **126**(7): p. 1067–1081.
6. Forde, C., et al., *Uptake of pre-symptomatic testing for BRCA1 and BRCA2 is age, gender, offspring and time-dependent.* J Med Genet, 2020.
7. Petrucelli, N., M.B. Daly, and T. Pal, *BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer*, in *GeneReviews(®)*, M.P. Adam, et al., Editors. 1993, University of Washington, Seattle

Copyright © 1993-2024, University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved.: Seattle (WA).

8. Tehranifar, P., et al., *Validation of family cancer history data in high-risk families: the influence of cancer site, ethnicity, kinship degree, and multiple family reporters.* Am J Epidemiol, 2015. **181**(3): p. 204–12.
9. Kiciński, M., J. Vangronsveld, and T.S. Nawrot, *An epidemiological reappraisal of the familial aggregation of prostate cancer: a meta-analysis.* PLoS One, 2011. **6**(10): p. e27130.
10. Bruner, D.W., et al., *Relative risk of prostate cancer for men with affected relatives: Systematic review and meta-analysis.* International Journal of Cancer, 2003. **107**(5): p. 797–803.
11. Perez-Cornago, A., et al., *Prospective investigation of risk factors for prostate cancer in the UK Biobank cohort study.* Br J Cancer, 2017. **117**(10): p. 1562–1571.
12. Ren, Z.J., et al., *First-degree family history of breast cancer is associated with prostate cancer risk: A systematic review and meta-analysis.* BMC Cancer, 2019. **19**(1): p. 871.
13. Beebe-Dimmer, J.L., et al., *Risk of Prostate Cancer Associated With Familial and Hereditary Cancer Syndromes.* Journal of Clinical Oncology, 2020. **38**(16): p. 1807–1813.
14. Sudlow, C., et al., *UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age.* PLOS Medicine, 2015. **12**(3): p. e1001779.
15. Digital, N., *Prostate cancer incidence (ICD-10 C61) by Gleason grade group*. 2023, NHS Digital.
16. Sheridan, T.B., et al., *Change in prostate cancer grade over time in men followed expectantly for stage T1c disease.* J Urol, 2008. **179**(3): p. 901–4; discussion 904–5.
17. Navarro, S., et al., *Racial/ethnic disparities in patient experiences with care and Gleason score at diagnosis of prostate cancer: a SEER-CAHPS study.* Cancer Causes & Control, 2022. **33**(4): p. 601–612.
18. Statistics, O.f.N., *Cancer Survival in England: adults diagnosed between 2013 and 2017 and followed up to 2018*, O.f.N. Statistics, Editor. 2019.
19. England, P.H., *Geographical Patterns of cancer survival in England: Adults diagnosed 2014 to 2018 and followed up to 2019*. 2020.

20. Service, N.C.R.a.A., *Get Data Out: Prostate*. 2023.
21. Digital, N., *Cancer Survival in England, cancers diagnosed 2016 to 2020, followed up to 2021*. 2023, NHS Digital: England.
22. Wills, L.P., C., *10-year cancer survival by stage for patients diagnosed in the East of England, 2007 to 2017*. 2021, Public Health England & Cancer Research UK.
23. Hamdy, F.C., et al., *Fifteen-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer.* N Engl J Med, 2023. **388**(17): p. 1547–1558.
24. Digital, N., *Cancer Registration Statistics, England 2018*. 2020, NHS Digital: England.
25. E, B.R.D.G.H.E.d.K.H.W., *CISNET Model Profile PROSTATE – ERASMUS MC*. 2009, Cancer Intervention and Surveillance Modeling Network (CISNET), Erasmus University Medical Center.
26. G, T.A.S.A.W.J.M.R.Z.M.G., *SCANS (Michigan) Prostate Cancer Model Profile*. 2008, Cancer Intervention and Surveillance Modeling Network (CISNET), University of Michigan: Ann Arbor (MI), USA.
27. W, E.R.G.R.I.L.K.J.H., *FHCRC (PSAPC) Prostate Cancer Model Profile*. 2009, Cancer Intervention and Surveillance Modeling Network (CISNET), Fred Hutchinson Cancer Research Center: Seattle, WA.
28. Martin, R.M., et al., *Prostate-Specific Antigen Screening and 15-Year Prostate Cancer Mortality: A Secondary Analysis of the CAP Randomized Clinical Trial.* JAMA, 2024. **331**(17): p. 1460–1470.
29. E.W., S., *Clinical Prediction Models. A Practical Approach to Development, Validation, and Updating*. 2019.
30. Jalal, H., et al., *An Overview of R in Health Decision Sciences.* Med Decis Making, 2017. **37**(7): p. 735–746.
31. Vehtari, A., et al., *Pareto smoothed importance sampling.* Journal of Machine Learning Research, 2024. **25**(72): p. 1–58.
32. Elvira, V. and L. Martino, *Advances in Importance Sampling*, in *Wiley StatsRef: Statistics Reference Online*. p. 1–14.
33. Fan, Y. and S. Sisson, *ABC samplers.* arXiv preprint arXiv:1802.09650, 2018.
34. Ankerst, D.P., et al., *Prostate Cancer Prevention Trial risk calculator 2.0 for the prediction of low- vs high-grade prostate cancer.* Urology, 2014. **83**(6): p. 1362–7.
35. Merriel, S.W.D., et al., *Systematic review and meta-analysis of the diagnostic accuracy of prostate-specific antigen (PSA) for the detection of prostate cancer in symptomatic patients.* BMC Med, 2022. **20**(1): p. 54.
36. Reza, H.S., et al., *Age-specific reference ranges of prostate-specific antigen in the elderly of Amirkola: A population-based study.* Asian J Urol, 2021. **8**(2): p. 183–188.
37. Zheng, Z., et al., *Tumor characteristics, treatments, and survival outcomes in prostate cancer patients with a PSA level < 4 ng/ml: a population-based study.* BMC Cancer, 2020. **20**(1): p. 340.
38. Ahmed, H.U., et al., *Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study.* Lancet, 2017. **389**(10071): p. 815–822.
39. England, L.R.C.o.S.o., *National Prostate Cancer Audit: State of the Nation Report*. 2024.
40. Goldberg, H., et al., *Comparison of Magnetic Resonance Imaging and Transrectal Ultrasound Informed Prostate Biopsy for Prostate Cancer Diagnosis in Biopsy Naïve Men: A Systematic Review and Meta-Analysis.* J Urol, 2020. **203**(6): p. 1085–1093.
41. Kanagarajah, A., et al., *A systematic review on the outcomes of local anaesthetic transperineal prostate biopsy.* BJU international, 2023. **131**(4): p. 408–423.

42. Keeney, E., et al., *Cost-Effectiveness Analysis of Prostate Cancer Screening in the UK: A Decision Model Analysis Based on the CAP Trial.* Pharmacoeconomics, 2022. **40**(12): p. 1207–1220.
43. NICE. *Prostate cancer: diagnosis and management*. 2021 [cited 2024 07/08]; Available from: https://www.nice.org.uk/guidance/ng131/chapter/Recommendations#localised-and-locally-advanced-prostate-cancer.
44. England, N., *Cancer Quality of Life Survey*. 2024: National Disease Registration Service.
45. Noble, S.M., et al., *The ProtecT randomised trial cost-effectiveness analysis comparing active monitoring, surgery, or radiotherapy for prostate cancer.* Br J Cancer, 2020. **123**(7): p. 1063–1070.
46. Alava M, P.S., Wailoo A, *ESTIMATING EQ-5D BY AGE AND SEX FOR THE UK.*, in *REPORT BY THE DECISION SUPPORT UNIT*. 2022, School of Health and Related Research, University of Sheffield, UK.
47. Kasivisvanathan, V., et al., *MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis.* New England Journal of Medicine, 2018. **378**(19): p. 1767–1777.
48. Li, L., J.L.H. Severens, and O. Mandrik, *Disutility associated with cancer screening programs: A systematic review.* PLoS One, 2019. **14**(7): p. e0220148.
49. Mowatt, G., et al., *The diagnostic accuracy and cost-effectiveness of magnetic resonance spectroscopy and enhanced magnetic resonance imaging techniques in aiding the localisation of prostate abnormalities for biopsy: a systematic review and economic evaluation.* Health Technol Assess, 2013. **17**(20): p. vii–xix, 1–281.
50. (NICE), N.I.f.H.a.C.E., *Health Economic Model Report*. 2019.
51. Nicholson, A., et al., *The clinical effectiveness and cost-effectiveness of the PROGENSA® prostate cancer antigen 3 assay and the Prostate Health Index in the diagnosis of prostate cancer: a systematic review and economic evaluation.* Health Technol Assess, 2015. **19**(87): p. i–xxxi, 1–191.
52. *Prostate cancer diagnosis and treatment. Clinical guideline CG175*. 2014, National Collaborating Centre for Cancer. p. 511.
53. Merriel, S.W.D., et al., *Early economic evaluation of magnetic resonance imaging for prostate cancer detection in primary care.* BJUI Compass, 2024. **n/a**(n/a).
54. England, L.T.R.C.o.S.o., *National Prostate Cancer Audit: Annual Report 2020*. 2021.
55. Thomas, C., et al., *The Costs and Benefits of Risk Stratification for Colorectal Cancer Screening Based On Phenotypic and Genetic Risk: A Health Economic Analysis.* Cancer Prev Res (Phila), 2021. **14**(8): p. 811–822.
56. Wills, L., et al., *Estimating surgery, radiotherapy and systemic anti-cancer therapy treatment costs for cancer patients by stage at diagnosis.* Eur J Health Econ, 2024. **25**(5): p. 763–774.
57. NICE, *Resource impact template: Enzalutamide for treating hormone-sensitive metastatic prostate cancer (TA712)*, NICE, Editor. 2021.
58. Committee, J.F. *Docetaxal*. 2006; Available from: https://bnf.nice.org.uk/drugs/docetaxel-specialist-drug/.
59. Committee, J.F. *Enzalutamide*. 2021 [cited 2024; Available from: https://bnf.nice.org.uk/drugs/enzalutamide/.
60. Committee, J.F. *Apalutamide*. 2021 [cited 2024; Available from: https://bnf.nice.org.uk/drugs/apalutamide/.
61. Agarwal, N., et al., *Health-related quality of life after apalutamide treatment in patients with metastatic castration-sensitive prostate cancer (TITAN): a randomised, placebo-controlled, phase 3 study.* Lancet Oncol, 2019. **20**(11): p. 1518–1530.

62. Round, J., L. Jones, and S. Morris, *Estimating the cost of caring for people with cancer at the end of life: A modelling study.* Palliat Med, 2015. **29**(10): p. 899–907.
63. Diernberger, K., et al., *Incompatible: end-of-life care and health economics.* BMJ Supportive & Palliative Care, 2021. **11**(3): p. 296.
64. Martin, R.M., et al., *Effect of a Low-Intensity PSA-Based Screening Intervention on Prostate Cancer Mortality: The CAP Randomized Clinical Trial.* JAMA, 2018. **319**(9): p. 883–895.
65. Hugosson, J., et al., *A 16-yr Follow-up of the European Randomized study of Screening for Prostate Cancer.* Eur Urol, 2019. **76**(1): p. 43–51.
66. Roobol, M.J., et al., *European Study of Prostate Cancer Screening — 23-Year Follow-up.* New England Journal of Medicine, 2025. **393**(17): p. 1669–1680.
67. Hatswell, A.J., et al., *Probabilistic Sensitivity Analysis in Cost-Effectiveness Models: Determining Model Convergence in Cohort Models.* PharmacoEconomics, 2018. **36**(12): p. 1421–1426.
68. Langley S, U.L.S., Uribe J et al. *A novel system to identify those prostate cancer patients who failed to present during Covid19 pandemic -results from over 18,000 PSA tests*. in *BAUS*. 2025. Manchester.
69. Draisma, G., et al., *Gleason score, age and screening: Modeling dedifferentiation in prostate cancer.* International Journal of Cancer, 2006. **119**(10): p. 2366–2371.
70. De Vulder, N., et al., *Safety and efficacy of software-assisted MRI-TRUS fusion-guided transperineal prostate biopsy in an outpatient setting using local anaesthesia.* Abdominal Radiology, 2023. **48**(2): p. 694–703.
71. Karlsson, A., et al., *A natural history model for planning prostate cancer testing: Calibration and validation using Swedish registry data.* PLoS One, 2019. **14**(2): p. e0211918.
72. Etzioni, R., et al., *Asymptomatic incidence and duration of prostate cancer.* Am J Epidemiol, 1998. **148**(8): p. 775–85.
73. Mariotto, A.B., et al., *Reconstructing PSA testing patterns between black and white men in the US from Medicare claims and the National Health Interview Survey.* Cancer, 2007. **109**(9): p. 1877–86.
74. Gulati, R., et al., *Calibrating disease progression models using population data: a critical precursor to policy development in cancer control.* Biostatistics, 2010. **11**(4): p. 707–19.
75. J, D.G.E.R.T.A.M.A.W.E.G.R.F.E.d.K.H., *Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context.* Journal of the National Cancer Institute, 2009. **101**(6): p. 374–383.
76. Tappenden, P. and J.B. Chilcott, *Avoiding and identifying errors and other threats to the credibility of health economic models.* Pharmacoeconomics, 2014. **32**(10): p. 967–79.