THE LANCET

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Sweeting MJ, Masconi KL, Jones E, et al. Analysis of clinical benefit, harms, and cost-effectiveness of screening women for abdominal aortic aneurysm. *Lancet* 2018; published online July 26. http://dx.doi.org/10.1016/S0140-6736(18)31222-4.

Supplement for

Should we screen women for abdominal aortic aneurysm? Analysis of clinical benefit, harms and cost-effectiveness

Michael J Sweeting, Katya L Masconi, Edmund Jones, Pinar Ulug, Matthew J Glover, Jonathan A Michaels, Matthew J Bown, Janet T Powell, Simon G Thompson

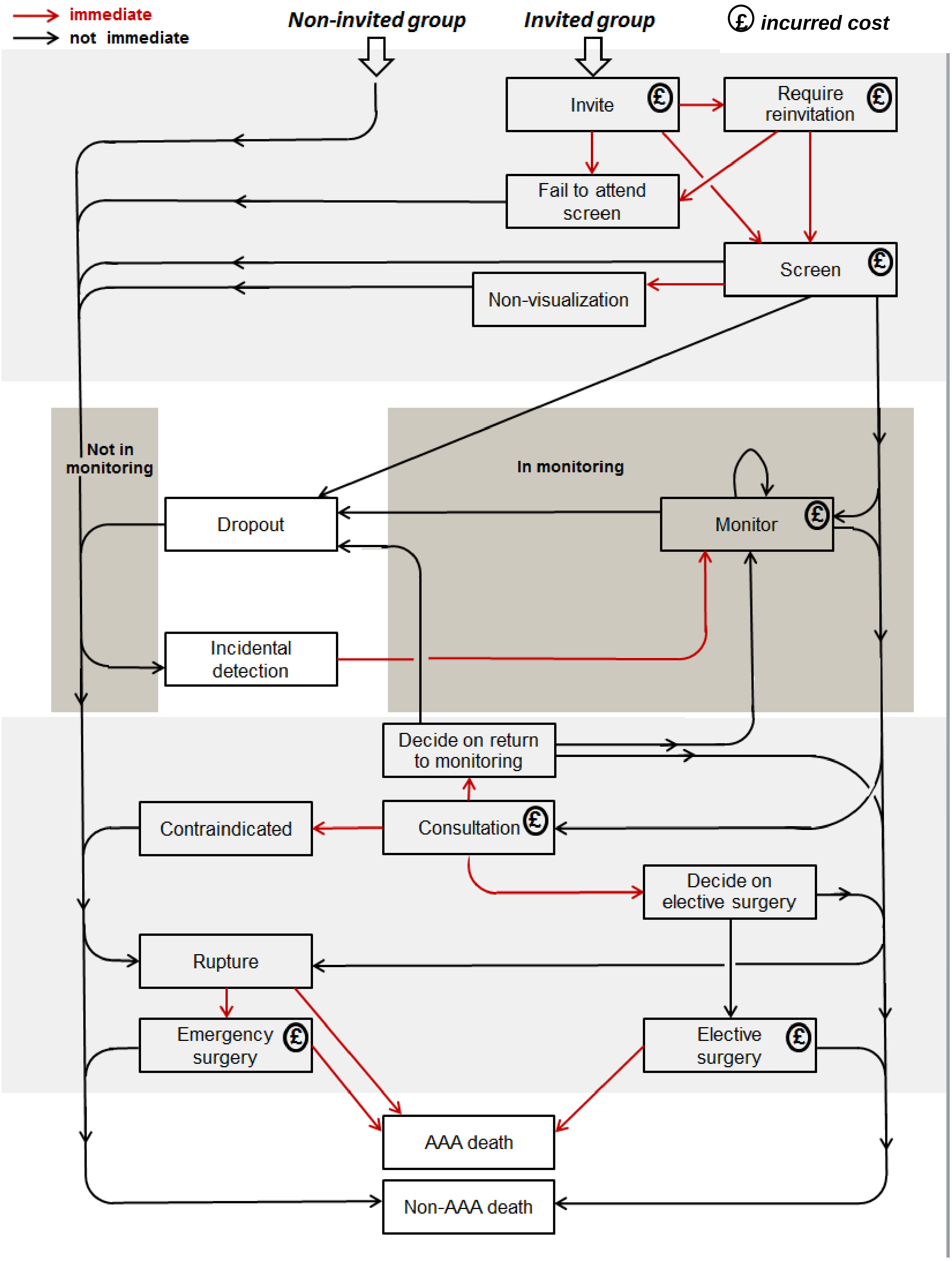
Containing:

Supplementary Methods

Supplementary Results

Supplementary Methods

Supplementary Figure 1. Structure of the discrete event simulation model



1. Baseline aortic diameter distribution

The distribution of diameters measured in the first 700,000 men screened in NAAASP[1] or from screening of 70 year old women in Sweden[2] were re-weighted to give the desired AAA prevalence in women, estimates of which were obtained from a systematic review[3]. A linear re-weighting approach was taken using the following algorithm:

1. Let p_{NAAASP} be the prevalence of AAA calculated in the NAAASP aortic diameter distribution for men and p_{WOMEN} the prevalence in women that we wish to re-calibrate the distribution to. Each aortic diameter size x (accurate to 1mm) in the NAAASP distribution has an associated probability weight $w(x)$ indicating the proportion of individuals in the distribution who were screened with that diameter. The weights sum up to 1. It follows that

$$p_{NAAASP} = \sum_{x \geq 3.0} w(x)$$

2. Given the desired prevalence, p_{WOMEN}, new weights $w^*(x)$ are calculated, as follows:

$$w^*(x) = f(x)w(x)$$

where $f(x) = a + bx$ is a linear function of x. The conditions that must be satisfied are
   i. $\sum_{x \geq 3.0} f(x)w(x) = p_{WOMEN}$
   ii. $\sum_x f(x)w(x) = 1$

The solution to this pair of simultaneous equations is

$$b = \frac{p_{NAAASP} - p_{WOMEN}}{p_{NAAASP} \sum_x x\, w(x) - \sum_{x \geq 3.0} x\, w(x)}$$

$$a = 1 - b \sum_x x\, w(x)$$

After re-weighting, some of the new weights may be negative. If this occurred, we set the weights to zero and then a further re-weighting step was performed to ensure the weights above the diagnosis threshold (e.g. 3.0cm) summed to the desired prevalence.

2. An alternative diagnosis threshold for women

Data from aneurysm screening in 5140 women aged 70 in Uppsala and Dalarna, Sweden, were obtained to investigate an alternative threshold for AAA in women based on the definition of being 50% larger than a normal aortic diameter[4]. The mean (leading edge to leading edge) diameter in these women was 1.66cm and an aortic diameter of 2.5cm was 3.2 standard deviations (SDs) (or 51%) higher than the mean, whilst a diameter of 3.0cm was 5.2 SDs (or 81%) higher than the mean. In men screened in NAAASP an (inner to inner) diameter of 3.0cm is 3.4 SDs (or 68%) above the mean. This suggests that 2.5cm might be an appropriate alternative threshold for women.

3. A model for aortic growth

The evolution of an individual's aortic diameter over time affects many aspects of the health economic model, namely: 1) when an individual can be diagnosed, 2) planned surveillance intervals, 3) when an intervention can be considered, 4) the risk of rupture, 5) the probability of receiving EVAR rather than open repair, and 6) the operative mortality risk. Hence, the trajectory of the aortic diameter was modelled using a continuous-time linear mixed model. Letting y_{ij} be the aortic diameter, as measured using ultrasound, of woman i at time t_{ij}, $j = 1, \ldots, n_i$; so y_{i0} is the baseline diameter as measured at screening. A linear mixed model was specified as follows:

$$\log(y_{ij}) = b_{0i} + b_{1i}t_{ij} + \epsilon_{ij}$$
$$= m_{ij} + \epsilon_{ij}$$
$$(b_{0i}, b_{1i})^T \sim N_2(\beta, G)$$
$$\text{where } \epsilon_{ij} \sim N(0, \sigma_w^2)$$
$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$
$$G = \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix}$$

Each woman has two random effects: their intercept b_{0i} (true log aortic diameter at the time of screening), and their slope b_{1i} (rate of growth), measured on the log diameter scale. Correlation between an woman's underlying baseline log diameter and slope was incorporated through the correlation parameter ρ. The parameters σ_0^2 and σ_1^2 determine the between-person variability of the intercepts and slopes, respectively, whilst σ_w^2 determines the amount of variability due to measurement error.

The linear mixed model was fitted using data from repeated ultrasound measurements of the aortic diameter from 11 cohorts of women with AAA from the RESCAN collaboration[5], with a total of 1743 women providing 4800 person-years for analysis. Parameter estimates were obtained via restricted maximum likelihood estimation for each study separately, and in a second state, study-specific estimates were pooled via multivariate random-effects meta-analysis.

The 11 RESCAN cohorts were restricted to the diameter range of 3.0 to 5.5cm. As a result, external data sources and model extrapolation were used to sample true baseline diameters and growth rates for women outside of this range. The baseline diameter y_{i0} was sampled from a fixed distribution, which was specified using external data sources. The base case model used the distribution of diameters measured in the first 700,000 men screened in NAAASP, reweighted to give the desired AAA prevalence. An individual's random effects b_{0i} and b_{1i} were then generated conditional on their observed baseline diameter. A set of rules were developed to ensure that extrapolated growth rates below 3.0cm were sensible and approximated empirical data obtained from a group of men followed up over time with initial diameter 2.6-2.9cm[6]. The rules were as follows:

1. <u>If $y_{i0} \geq 3.0$ then random-effects were generated directly from the linear mixed model posterior distribution</u>
   Since estimated parameters from the linear mixed model are strictly relevant only to baseline diameters ≥3.0cm, then for individuals in this range, b_{0i} and b_{1i} are generated from their bivariate normal distribution conditional on the observed diameter, y_{i0}:
   $$(b_i | y_{i0}) \sim N_2(\mu_b, \Sigma_b)$$
   where
   $$\mu_b = \beta + \begin{pmatrix} \sigma_0^2 \\ \rho\sigma_0\sigma_1 \end{pmatrix} \frac{\log(y_{i0}) - \beta_0}{\sigma_0^2 + \sigma_w^2}$$
   $$\Sigma_b = \begin{pmatrix} \sigma_0^2 + \sigma_w^2 & \rho\sigma_0\sigma_1\sigma_w^2 \\ \rho\sigma_0\sigma_1\sigma_w^2 & \sigma_0^2\sigma_1^2(1-\rho^2) + \sigma_1^2\sigma_w^2 \end{pmatrix}$$

2. <u>If $y_{i0} < 3.0$ then an individual's true baseline diameter was set to their observed diameter</u>

This avoids shrinkage of the true baseline diameter upwards towards the mean in the RESCAN cohort used to fit the linear mixed model.

3. If $2.0 \leq y_{i0} < 3.0$ then an individual's rate of growth was generated from their posterior distribution conditional on b_{i0}:

$$(b_{1i}|b_{0i}) \sim N(\mu_{b1}, \sigma_{b1}^2)$$

where

$$\mu_{b1} = \beta_1 + \frac{\rho\sigma_1}{\sigma_0}(b_{0i} - \beta_0)$$
$$\sigma_{b1}^2 = (1 - \rho^2)\sigma_1^2$$

4. If $y_{i0} < 2.0$ then an individual's rate of growth to zero was set to zero
   This rule implies that no individuals measured below 2.0cm at baseline will grow during their lifetime.

The effect of the extrapolation rules set out above was investigated in validation studies conducted in men, with outputs from the model compared against data from the randomised Multicentre Aneurysm Screening Study; further details of which of given in Glover et al. 2018[7]. It should be noted that incremental effects (e.g. incremental QALYs, increments costs and the ICER) are robust to the choice of growth rates below the diagnosis threshold, since individuals below the diagnosis threshold at time of screening follow the same life course in both screened and non-screened populations.

The rate of AAA rupture was assumed to depend on the underlying AAA diameter and was modelled using a joint longitudinal and time-to-event model with the hazard of rupture for woman i at time t specified as

$$h_i(t) = \exp\left(\gamma + \alpha\left(b_{0i} + b_{1i}t_{ij}\right)\right)$$

where γ is the log baseline hazard and α is the log hazard ratio associated with a one unit increase in log aortic diameter (the expression in the inner brackets). The hazard function corresponds to a Gompertz distribution with shape parameter αb_{1i} and rate parameter $\exp(\gamma + \alpha b_{0i})$. The (primary) rupture risk was set to zero at the time a woman underwent a successful elective AAA operation.

Six RESCAN studies provided data on both AAA growth and rupture. The model was fitted separately within each study before pooling estimates using multivariate random-effects meta-analysis. Since ruptures were rare, we used data from both 1071 women and 5358 men, contributing 49 and 92 AAA rupture events, respectively, and a total of 21,658 person-years of follow-up. We allowed for sex differences in AAA diameter and rate of rupture by including sex as a covariate in both the longitudinal (growth) and time-to-event (rupture) sub-models. A linear relationship between log(diameter) and time was assumed to model the growth of an aneurysm.

4. Operative mortality and non-intervention rates

Data on operative mortality rates for both endovascular and open aneurysm repairs, and elective and emergency operations were extracted from the UK National Vascular Registry (NVR)[8] and Hospital Episode Statistics (HES) data[9], which contains details of all admissions, outpatient appointments and A&E attendances at NHS hospitals in England. NVR contains data on in-hospital mortality and HES contains data on both 30-day and in-hospital mortality. NVR was the principal source used for surgical parameters for women since data from this registry were used to create age and AAA diameter-specific estimates using logistic regression models. The NVR in-hospital mortality was then adjusted to reflect the (greater) 30-day mortality with a log odds ratio corresponding to the 30-day mortality vs. in-hospital mortality in HES. EVAR was used in ~60% of elective repairs recorded in NVR, but in <50% for women aged less than 75[10]. The overall estimated 30-day mortality rates were 2.4% for elective endovascular repair, 8.1% for elective open repair, 35.9% for emergency endovascular repair, and 44.2% for emergency open repair. Non-intervention rates were obtained from a systematic review[11].

5. Incidental detection rate

In the discrete event simulation model all incidental detections were assumed to thereafter follow the same surveillance protocol as a screen-detected AAA (i.e. surveillance for those detected below the intervention threshold, and referral for consideration of surgery for those detected above the intervention threshold).

Data on the incidental detection rate were obtained from a study conducted in Canterbury, New Zealand in which 165 new incidental AAAs were detected in men and women from CT scans over the period of 4.25 years[12]. About a quarter of all detected AAAs (incidental and known) were in women. Assuming this proportion also applies to the incidental AAAs and that 97% of AAAs were in individuals aged 65 and over, then there would be approximately 40 AAAs detected in women aged ≥65 years. From census data, the 2006 population of women ≥65 years for the catchment area (Canterbury, West Coast, and Timaru regions of South Island, New Zealand) is approximately 43,500. Based on a prevalence of 0.74% for women ≥65 years[3], 321 of these women have an aneurysm. This would indicate an incidental detection rate of approximately 40/(321 x 4.25) = 2.93 per 100 person-years for women ≥65 years with an AAA. This is similar to the rate of 4.6 per 100 person-years used in the most recent health-economic model for men[13].

The rate is also similar to data from electronic hospital records of women aged 65 years and over undergoing CT scanning obtained from the University Hospital of South Manchester in 2014; 2494 women underwent an abdominal CT during this period, and 65 AAAs were identified. Of these, 53 were newly identified AAAs, but only 7 were referred on to vascular surgeons to be followed up with surveillance or elective surgery. The population (women ≥65 years) of the referral catchment area for the university hospital is approximately 24,500. Assuming that 181 (0.74%) of these women have an aneurysm this would indicate an incidental detection rate to a surveillance programme of approximately 7/181 = 3.9 per 100 person-years for women ≥65 years with an AAA.

6. Cost discounting

The cost discounting rate of 3.5% was as recommended by the UK Treasury (Finance Ministry)[14].

Supplementary Table 1. Input parameters for the reference case, probability distributions used in probabilistic sensitivity analyses (PSA) and deterministic sensitivity analyses (DSA) inputs

Parameter	Source	Reference model	PSA	DSA
Screening				
Re-invitation proportion	NAAASP (unpublished)	142,127 / 594,376 ≈ 0.239	None	None
Attendance proportion	Scott et al. 2002 [15]	218 / 300 ≈ 0.727	Beta(218,82)	None
Non-visualisation proportion	NAAASP (unpublished)	1652 / 470,531 ≈ 0.0035	None	None
AAA size distribution at screening	NAAASP (unpublished)	NAAASP distribution, reweighted to give 0.0043 prevalence	NAAASP distribution based on uncertain prevalence (see below)	Uppsala distribution, reweighted to give 0.0043 prevalence
Prevalence proportion	Ulug et al. 2016 [3]	0.0042756	Based on Normal (–5.45054, 0.32321^2) distribution for logit(p)	a) 0.0021378 b) 0.0085512
AAA growth & rupture				
AAA growth	Thompson et al. 2013 [5]	Mixed linear model for log AAA diameter *	Using variance – covariance matrix for the 6 parameters **	None
AAA rupture	Thompson et al. 2013 [5]	Joint model for log rupture rates and log underlying AAA diameter †	Using variance – covariance matrix for the 2 parameters ‡	None
Surveillance				
Dropout rate from surveillance	NAAASP (unpublished)	1072 / 19,650 ≈ 0.0546 per year	Gamma(1072, 19650)	a) 0.0273 per year b) 0.1092 per year
Incidental detection rate	Khashram et al. 2015 [12]	40 / 1364.25 ≈ 0.0293 per year	Gamma(40, 1364.25)	a) 0.0147 per year b) 0.0586 per year
Delay from 5.5+cm scan to consultation	NAAASP (unpublished)	10.6 days	None	None
Consultation scan	Thompson et al. 2013 [5], Singh et al. 2003 [16]	CT is on average 0.244cm greater than US; measurement error SD 0.19cm for CT	None	None
Decision at consultation: proportion returned to surveillance	N/A	Modelled directly from AAA measurements by CT	N/A	N/A
Decision at consultation: non-intervention proportion	Meta-analysis from four hospitals (Ulug et al. 2017[11])	0.34226 of those not returned to surveillance	Based on Normal (–0.65324, 0.13502^2) distribution for logit(p)	0.233 at age 80 of those not returned to surveillance. Odds ratio 1.20 per year increase in age
Decision at consultation: proportion elective surgery	N/A	1 – 0.34226 = 0.65774 of those not returned to surveillance	Obtained by subtraction	Obtained by subtraction
Delay from consultation scan to elective surgery	NAAASP (unpublished)	70.8 days	None	None

Supplementary Table 1 continued

Elective operations				
Proportion receiving EVAR vs. open repair	NVR (unpublished)	0.67 at age 80, AAA diameter 6.0cm. Odds ratio 1.10 per year increase in age, 0.74 per cm increase in diameter	Based on multivariate normal from logistic regression parameters	0.3396 based on systematic review of EVAR suitability
EVAR 30-day operative mortality	NVR [10], HES (unpublished)	0.027 at age 80, AAA diameter 6.0cm. Odds ratio 1.002 per year increase in age, 0.97 per cm increase in diameter	Based on multivariate normal from logistic regression parameters	0.0223 based on systematic review
Open repair 30-day operative mortality	NVR [10], HES (unpublished)	0.103 at age 80, AAA diameter 6.0cm. Odds ratio 1.07 per year increase in age, 1.08 per cm increase in diameter.	Based on multivariate normal from logistic regression parameters	a) 0.0537 based on systematic review b) 0.05
Re-intervention rate after successful EVAR	EVAR1 RCT [17]	20.3 and 6.4 per 100 women-years during 31-120 and >120 days respectively	Based on Gamma(3, 15) and Gamma(27, 421) respectively	None
Re-intervention rate after successful open repair	EVAR1 RCT [17]	0.0	None	a) Based on DREAM/OVER RCT rates in men, since these trials include incisional hernias. Overall rate across two trials combined: 4.4 and 2.9 per 100 women-years during 31-120 and >120 days respectively
Long-term AAA mortality rate after successful EVAR	EVAR1 RCT [17]	1.799 per 100 women-years	Based on Gamma(8, 444.7)	None
Long-term AAA mortality rate after successful open repair	EVAR1 RCT [17]	0.499 per 100 women-years	Based on Gamma(2, 400.8)	None
Emergency operations				
% operated after rupture	Literature review (unpublished), IMPROVE RCT [18]	0.25	Based on Normal(0.25, 0.05^2), with truncation to within [0,1]	None
Proportion receiving EVAR vs. open repair	NVR (unpublished)	0.18 at age 80. Odds ratio 1.04 per year increase in age	Based on multivariate normal from logistic regression parameters	None
EVAR 30-day operative mortality	NVR [10], HES (unpublished)	0.35 at age 80. Odds ratio 1.06 per year increase in age	Based on multivariate normal from logistic regression parameters	0.32 based on systematic review
Open repair 30-day operative mortality	NVR [10], HES (unpublished)	0.46 at age 80. Odds ratio 1.03 per year increase in age	Based on multivariate normal from logistic regression parameters	0.51 based on systematic review
Re-intervention rate after successful EVAR	IMPROVE RCT [18]	15.8 per 100 women-years	Based on Gamma(9, 57)	None
Re-intervention rate after successful open repair	IMPROVE RCT [18]	2.3 per 100 women-years	Based on Gamma(2, 85)	None
Long-term AAA mortality rate after successful EVAR	IMPROVE RCT [18]	0.0	None	0.985 per 100 women-years based on men
Long-term AAA mortality rate after successful open repair	IMPROVE RCT [18]	1.613 per 100 women-years	Based on Gamma(2, 124)	1.437 per 100 women-years based on men

Supplementary Table 1 continued

Costs				
Invitation, re-invitation	NAAASP (unpublished)	£1.80	In all cases: Based on Normal(log(base-case estimate), 0.114^2) for log costs	In all cases: a) Base-case estimate * 0.80 b) Base-case estimate * 1.25
Screening scan	NAAASP (unpublished)	£34.11		
Surveillance scan	NAAASP (unpublished)	£72.03		
Consultation for elective surgery	MASS [19], NHS Reference costs 2014/15	£328.64		
Elective EVAR repair	EVAR1 [17], HES (unpublished), NHS Reference costs 2014/15	£13,844		
Elective open repair	EVAR1 [17], HES (unpublished), NHS Reference costs 2014/15	£13,060		
Emergency EVAR repair	IMPROVE [18], HES (unpublished), NHS Reference costs 2014/15	£16,154		
Emergency open repair	IMPROVE [18], HES (unpublished), NHS Reference costs 2014/15	£17,613		
Surveillance after operations	Expert opinion, NHS Reference costs 2014/15	£258.16 annually after EVAR, £196.79 at 6 weeks after open repair		
Re-intervention after EVAR	EVAR1 [17]	£7,546		
Re-intervention after open repair	EVAR1 [17]	£8,986		
Miscellaneous				
Non-AAA mortality rate	ONS	ONS 2012-14 data by single year of age, ages 65-94	None	None
Overall QoL / utilities	EuroQol-5D	0.81 for age 55-64; 0.78 for age 65-74, 0.71 for age ≥75	None	None
QoL harms of screening	Ashton et al. 2002 [20]	No effect	None	Utility decrements of -0.01 for AAA diagnosis during surveillance,
QoL harms of surgery	EVAR1 [17], IMPROVE [18]	No effect	None	Utility decrements of -0.02 EVAR elective and -0.07 Open elective (3 months), -0.04 EVAR emergency and -0.10 Open emergency (3 years), -0.10 contraindicated (remaining lifetime)
Discounting rates	N/A	a) Undiscounted b) 3.5% per year for costs, 3.5% per year for life-years	None	None

NAAASP – National Abdominal Aortic Aneurysm Screening Programme
NVR – National Vascular Registry
HES – Hospital Episodes Statistics
EVAR1 RCT – EVAR-1 Randomised Controlled Trial
IMPROVE – IMPROVE Randomised Controlled Trial

* Slope ($\beta_1 = 0.052$), Intercept ($\beta_0 = 1.33$), Slope log SD ($\log(\sigma_1) = -3.28$), Intercept log SD ($\log(\sigma_0) = -1.99$), Arctanh correlation ($\text{atanh}(\rho) = 0.41$), Residual log SD ($\log(\sigma_w) = -2.96$)

** $N(\mu, \Sigma)$ where $\mu = (0.052 \quad 1.33 \quad -3.28 \quad -1.99 \quad 0.41 \quad -2.96)$, and

$$\Sigma = \begin{pmatrix} 0.000015 & & & & & \\ 6.5 \times 10^{-6} & 0.000568 & & & & \\ 0.000028 & -0.000752 & 0.009516 & & & \\ 0.000186 & -0.001364 & 0.005153 & 0.011569 & & \\ -0.000125 & -0.000418 & -0.000047 & 0.000843 & 0.011419 & \\ -0.000087 & -0.001800 & 0.002401 & 0.005566 & 0.005260 & 0.013688 \end{pmatrix}$$

† Association with diameter ($\gamma_1 = 5.47$), Intercept ($\gamma_0 = -12.40$)

‡ $N(\mu, \Sigma)$ where $\mu = (5.47, -12.40)$, and $\Sigma = \begin{pmatrix} 1.5892 & -2.2178 \\ -2.2178 & 3.1406 \end{pmatrix}$

7. Patient and public involvement

Public interest groups were set up to support this research by author MJB. No formal qualitative research was conducted.

During the development phase of this research men and women attending a public information event about the management of AAA at the (UK) University Hospitals of Leicester NHS Trust were invited to join a focus group and assist with the design of this research for the purpose of developing the funding application. Four men and two women attended an initial meeting in July 2015. All the men had screen-detected small AAA and one of the women was the partner of one of the men. The aim of this initial meeting was to establish if screening women for AAA was a public research priority and explore patient and public priorities to be examined in the research. This contributed to the overall concept of the research by confirming the general acceptability of screening programmes but highlighted that one of the key areas of importance to potential patients is the acceptability/risks of treatment for screen-detected diseases. This confirmed that the proposed aims of the research were valid and the design was appropriate to meet public research priorities.

The initial focus group convened in the design phase of the project had significant knowledge of AAA and AAA screening. To address this another project specific group was established that was representative of the target population. Through television and radio broadcast interviews in Leicestershire women were invited to participate in this second focus group. 11 women responded and attended three meetings over the duration of the project (January 2016, August 2016 and March 2017). One women had a strong family history of AAA (2 first-degree relatives) and one woman's husband had previously undergone an AAA repair. The majority (9 women) had family members who had been affected by AAA. The aim of these meetings was to confirm the findings from the initial focus group, obtain feedback regarding the aims of the project, to ensure that outputs were representative of the information relevant to the public and to provide a public perspective on the overall study results.

At the initial project specific focus group meeting (January 2016) the concept of screening was discussed. Evidence for and against screening women for AAA was presented verbally as a means to start an overall discussion about screening. The overall theme arising from this initial meeting was that the reassurance of a negative screen would be the main benefit for most women. All members of the focus group thought that AAA screening should be offered to women. A specific discussion was held with the focus group regarding the acceptability of treatment (surgery) for AAA. With the knowledge that AAA repair was a higher risk procedure for women the focus group thought that most women would want to undergo AAA repair if feasible. The group were asked about whether they would want to undergo AAA repair if this were indicated, particularly with the knowledge that women have higher perioperative risk than men. The women thought that providing the overall risks were considered that most women would want to undergo and AAA repair. The effect of age on perioperative risk was raised by members of the group who also suggested that older women may not want screening as they would not want to know or undergo surgery if diagnosed with an AAA.

A second meeting in August 2016 was used to explore the specific themes of targeting AAA screening for women at high-risk groups such as smokers. Having previously identified that the main benefit of screening for most women was the reassurance provided by a negative screening, the group thought that targeted screening would not be desirable since the main positive effect of screening would be denied to a large proportion of women.

A final focus group meeting was held in March 2017. At this meeting the results of the SWAN project were available. This meeting was first used to re-discuss and clarify the themes identified in the previous meetings. The focus group confirmed that AAA screening was highly acceptable to women and that they would all attend if invited. They thought that most women would attend if invited. The group confirmed that screening should be offered to all women rather than being targeted at high risk groups.

Following this initial discussion the group were provided with the following written plain English summary of the results of the SWAN project, written for the National Institute for Health Research official project report:

"Abdominal aortic aneurysms (AAAs) are bulges in the main blood vessel in the abdomen. If an AAA gets too large it can burst (rupture) and this is usually fatal. While an AAA does not usually have any symptoms and is unlikely to cause problems until it bursts, AAAs can be easily diagnosed by a simple ultrasound screening scan. In the UK, men aged 65 are offered an ultrasound scan to look for an AAA and just over 1 in 100 men who are screened have an AAA. Men found to have an AAA are offered an operation to prevent the aneurysm bursting if it is large, or offered regular scans to monitor their AAA if it is small.

Women are not currently screened for AAA, mainly because they are less likely to have AAAs than men. Currently there is no information on whether screening for AAA would save lives from AAA rupture in women, or whether this would be cost-effective for the NHS. In this research we have gathered together a wide range of available information about AAAs in women to find out if screening women for AAA might be effective. We have developed a computer program to analyse all of this information and simulate what would happen if women were screened for AAA.

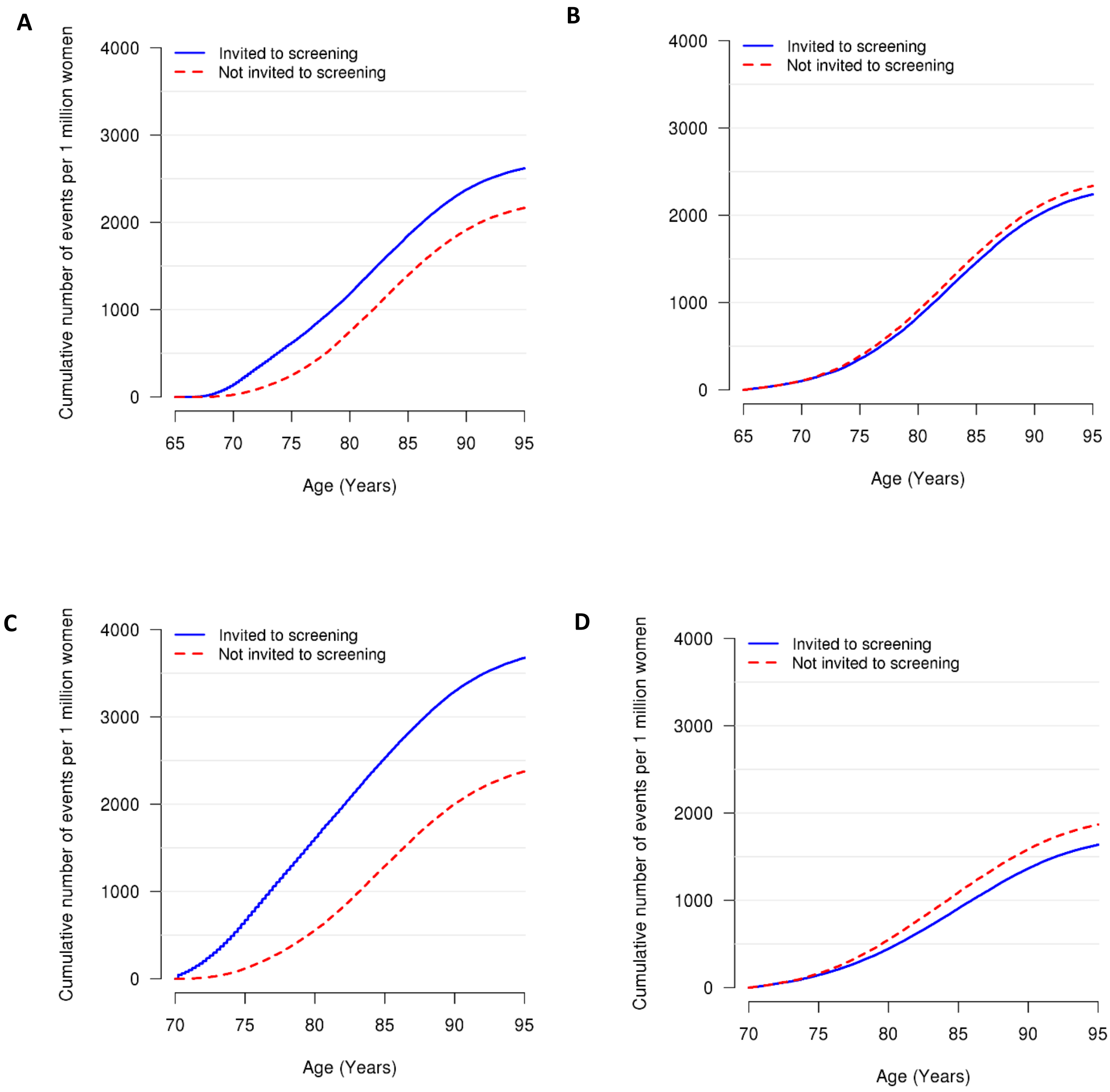
Our research has shown that if women were offered the same screening as men this would have a very minor effect on the overall life-expectancy of women, gaining on average just over one day of life per woman invited to screening. Although there is considerable uncertainty, we estimate that around 4100 women would need to be invited to screening to prevent one death from AAA, and that screening would cost £150,000 per death from AAA prevented.

Based on our findings, a national AAA screening programme for women would not be cost-effective for the NHS."

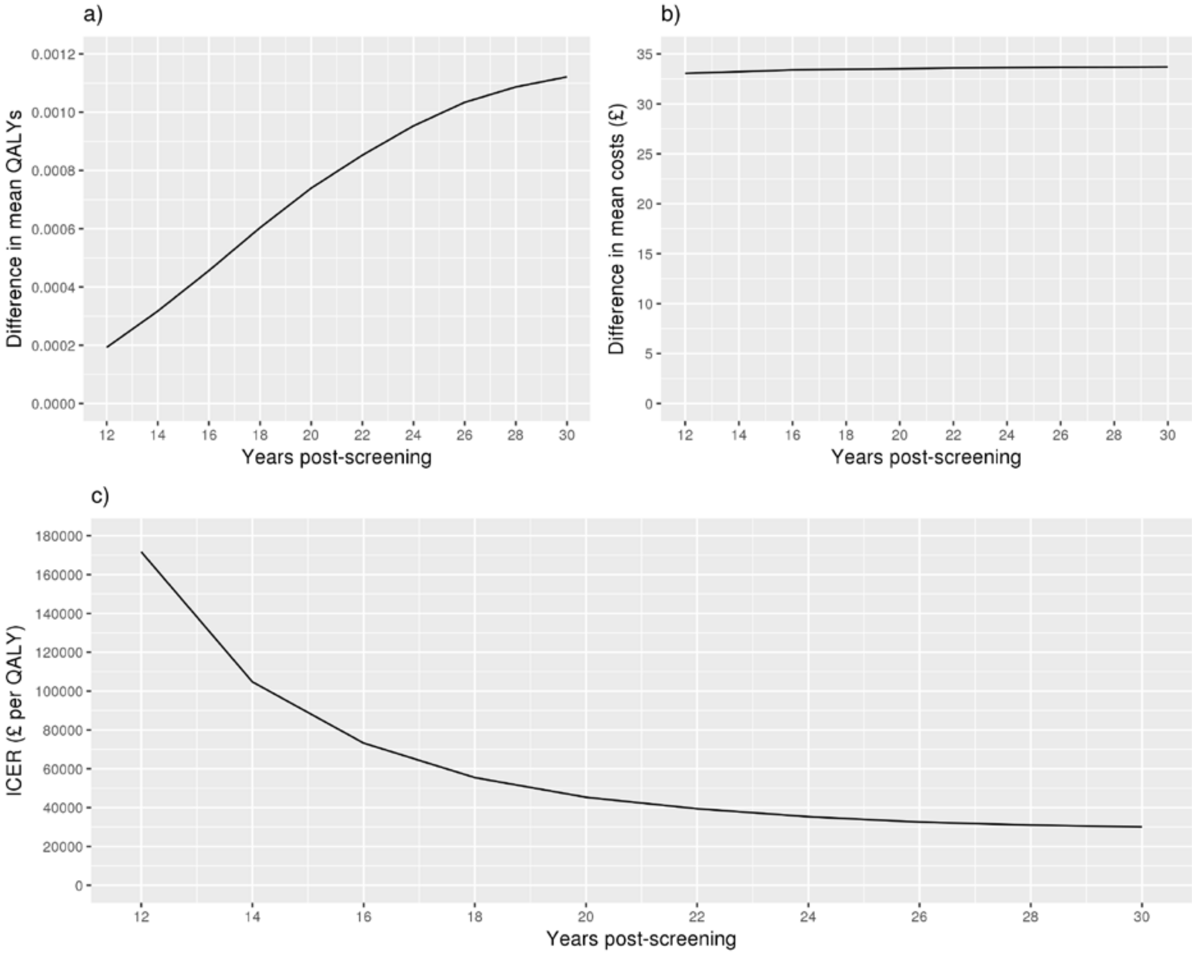
Following the presentation of this plain English summary the themes previously identified were re-discussed. Based on the results presented, the women present thought that targeted screening may be better than no screening at all for women. Despite the negative cost-effectiveness results the members of the focus group thought that AAA screening would still have significant positive benefits for most women. The group thought that the positive effects of a normal screening scan should be investigated as a research priority going forward and that this should be combined with a more detailed assessment of quality of life in screen-negative women.

Supplementary Results

Supplementary Figure 2. A) Cumulative elective operations and B) cumulative emergency operations in the invited to screening vs. not invited to screening groups in the reference case per 1 million women. C) Cumulative elective operations and D) cumulative emergency operations in the invited to screening vs. not invited to screening groups in the best alternative strategy per 1 million women.

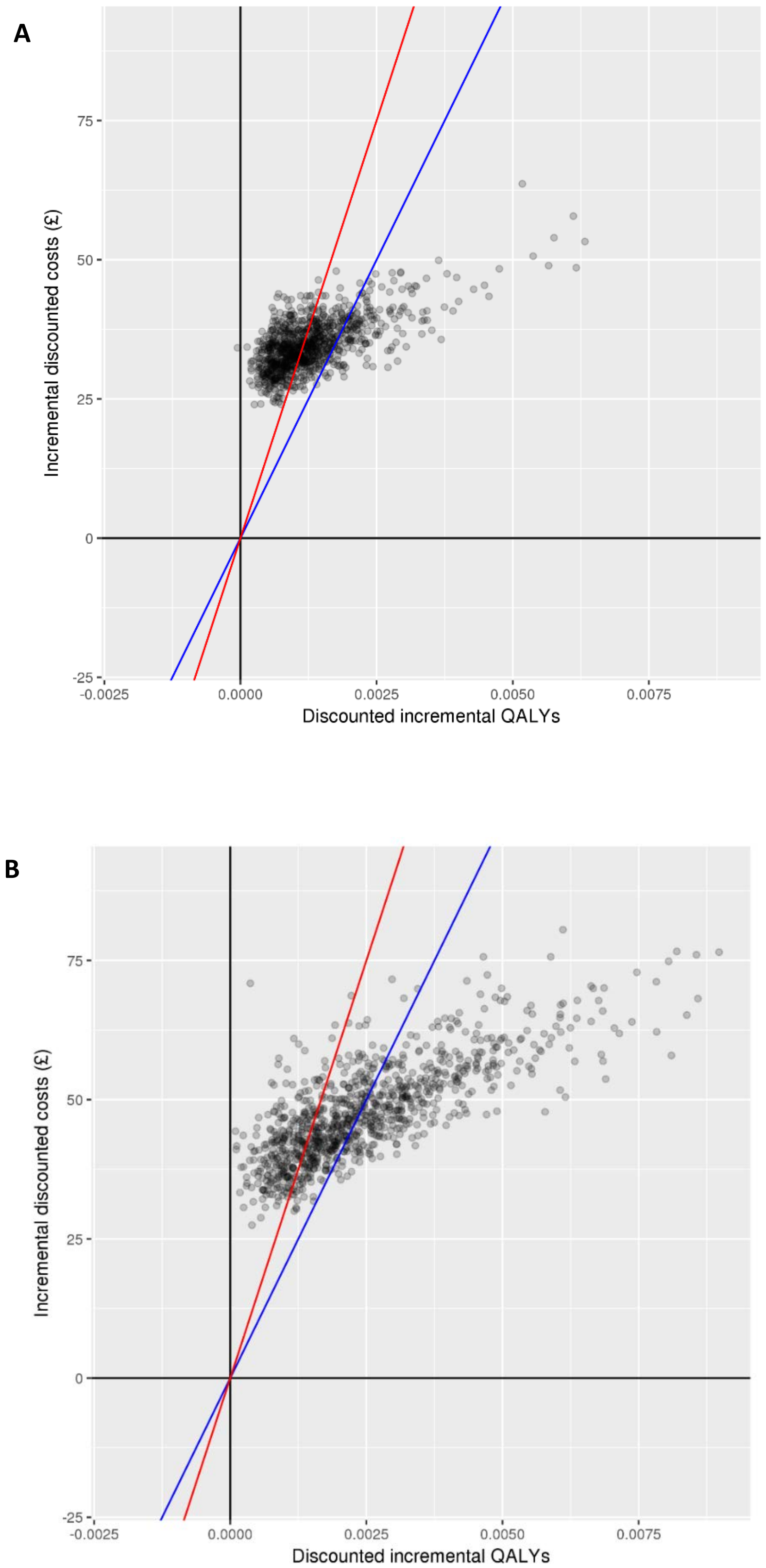


Supplementary Figure 3. Estimates of a) incremental QALYs, b) costs and c) the cost-effectiveness ratio over time in the reference case, up to 30 years after invitation to screening.



Supplementary Figure 4. Cost-effectiveness of invitation to AAA screening with 1,000 probabilistic sensitivity analysis iterations for A) the reference case, and B) the best alternative screening strategy. The blue and red lines indicate willingness-to-pay thresholds of £20,000 and £30,000 per QALY.

QALY – Quality adjusted life-year.



Supplementary Table 2. Numbers of AAA ruptures in the reference case and best alternative strategy, for 1 million women

	Reference case		Best alternative	
Number of AAA ruptures	Not invited to screening N=9,235 (100%)	Invited to screening N=8,839 (100%)	Not invited to screening N=7,465 (100%)	Invited to screening N=6,555 (100%)
Screened normal, no further contact	-	4,273 (48%)	-	1,761 (27%)
Failed to attend (not invited in no screening arm) or non-visualised aorta	7,465 (81%)	2,048 (23%)	6,101 (82%)	1,991 (30%)
Under surveillance	515 (6%)	689 (8%)	358 (5%)	646 (10%)
After dropping out of surveillance	514 (6%)	891 (10%)	371 (5%)	1,027 (16%)
After undergoing vascular consultation, but before surgery	44 (0.5%)	48 (0.5%)	32 (0.4%)	41 (0.6%)
After being turned down for surgery	697 (8%)	890 (10%)	603 (8%)	1,089 (17%)

Supplementary Table 3. Effect of health related quality of life decrements on mean QALYs and the incremental cost-effectiveness ratio

Quality adjustment	QoL weights	Length of change	Reference case			Alternative scenario		
			Mean QALYs		ICER	Mean QALYs		ICER
			Not invited	Invited		Not invited	Invited	
Age only	0.78 (Age<75) 0.71 (Age $\geq$ 75)	-	10.4484	10.4495	30,000	8.7257	8.7277	23,000
AAA diagnosis†	-0.01	Under surveillance	10.4478	10.4486	43,000	8.7247	8.7253	76,000
Elective surgery‡	-0.02 [EVAR] -0.07 [Open]	3 months	10.4483	10.4495	30,000	8.7257	8.7276	23,000
Emergency surgery¥	-0.04 [EVAR] -0.10 [Open]	3 years	10.4481	10.4492	30,000	8.7255	8.7275	23,000
Elective surgery contraindicated*	-0.10	Lifetime	10.4479	10.4488	35,000	8.7251	8.7266	30,000
AAA diagnosis, elective & emergency surgery and contraindication	All of the above	As above	10.4469	10.4476	52,000	8.7239	8.7241	278,000

QoL – Quality of life, QALY – Quality adjusted life-year, ICER – Incremental cost-effectiveness ratio (£ per QALY)

† Investigating reduction in EQ-5D of 0.01 from diagnosis to end of surveillance.

‡ Evidence from EVAR-1 randomised controlled trial showed a 3% reduction in QoL for EVAR and a 9% reduction for open repair from 0-3 months post-surgery[17]. Hence, we investigate a reduction of EQ-5D of 0.02 in those undergoing EVAR and 0.07 in those undergoing open repair.

¥ Evidence from IMPROVE trial showed EQ-5D of 0.76 (EVAR) and 0.66 (open repair) at 3 months, 0.78 (EVAR) and 0.71 (open repair) at 12 months and 0.74 (EVAR) and 0.73 (open repair) at 36 months post-surgery[18]. Assuming EQ-5D of zero at operation, a return to usual quality of life by 12 months for EVAR and 36 months for open repair, we investigate an average reduction in utility of 0.04 and 0.10 for EVAR and open repair, respectively over 3 years.

* Investigating reduction in EQ-5D of 0.10 for remaining life from non-intervention for surgery. Reduced life-years in those contraindicated not accounted for in the model, likely resulting in too severe a reduction in mean QALYs.

References

1. Jacomelli J, Summers L, Stevenson A, Lees T, Earnshaw JJ. Impact of the first 5 years of a national abdominal aortic aneurysm screening programme. *Br J Surg* 2016; **103**(9): 1125-31.
2. Svensjö S, Björck M, Wanhainen A. Current prevalence of abdominal aortic aneurysm in 70-year-old women. *British Journal of Surgery* 2013; **100**(3): 367-72.
3. Ulug P, Powell J, Sweeting M, Bown M, Thompson S. Meta-analysis of the current prevalence of screen-detected abdominal aortic aneurysm in women. *British Journal of Surgery* 2016; **103**(9): 1097-104.
4. Kent KC. Clinical practice. Abdominal aortic aneurysms. *N Engl J Med* 2014; **371**(22): 2101-8.
5. Thompson SG, Brown LC, Sweeting MJ, et al. Systematic review and meta-analysis of the growth and rupture rates of small abdominal aortic aneurysms: implications for surveillance intervals and their cost-effectiveness. *Health technology assessment (Winchester, England)* 2013; **17**(41): 1-118.
6. Darwood R, Earnshaw JJ, Turton G, et al. Twenty-year review of abdominal aortic aneurysm screening in men in the county of Gloucestershire, United Kingdom. *J Vasc Surg* 2012; **56**(1): 8-13.
7. Glover MJ, Jones E, Masconi KL, Sweeting MJ, Thompson SG, collaborators S. Discrete event simulation for decision making in health care: lessons from abdominal aortic aneurysm screening. *Med Decis Making* 2018; **38**(4): 439-51.
8. Vascular Services Quality Improvement Programme. NVR annual report, 2015.
9. NHS Digital. Hospital Episode Statistics. 2016.
10. Sidloff DA, Saratzis A, Sweeting MJ, et al. Sex differences in mortality after abdominal aortic aneurysm repair in the UK. *Br J Surg* 2017; **104**(12): 1656-64.
11. Ulug P, Sweeting MJ, von Allmen RS, Thompson SG, Powell JT, SWAN Collaborators. Morphological suitability for endovascular repair, non-intervention rates, and operative mortality in women and men assessed for intact abdominal aortic aneurysm repair: systematic reviews with meta-analysis. *The Lancet* 2017; **389**(10088): 2482-91.
12. Khashram M, Jones G, Roake J. Prevalence of abdominal aortic aneurysm (AAA) in a population undergoing computed tomography colonography in Canterbury, New Zealand. *European Journal of Vascular and Endovascular Surgery* 2015; **50**(2): 199-205.
13. Glover M, Kim L, Sweeting M, Thompson S, Buxton M. Cost-effectiveness of the National Health Service abdominal aortic aneurysm screening programme in England. *British Journal of Surgery* 2014; **101**(8): 976-82.
14. Her Majesty's Treasury. The green book: appraisal and evaluation in central government. London TSO; 2003.
15. Scott R, Bridgewater S, Ashton H. Randomized clinical trial of screening for abdominal aortic aneurysm in women. *British Journal of Surgery* 2002; **89**(3): 283-5.
16. Singh K, Jacobsen BK, Solberg S, et al. Intra- and interobserver variability in the measurements of abdominal aortic and common iliac artery diameter with computed tomography. The Tromso study. *Eur J Vasc Endovasc Surg* 2003; **25**(5): 399-407.
17. Patel R, Sweeting MJ, Powell JT, Greenhalgh RM, investigators Et. Endovascular versus open repair of abdominal aortic aneurysm in 15-years' follow-up of the UK endovascular aneurysm repair trial 1 (EVAR trial 1): a randomised controlled trial. *Lancet* 2016; **388**(10058): 2366-74.
18. IMPROVE Trial Investigators. Comparative clinical effectiveness and cost effectiveness of endovascular strategy v open repair for ruptured abdominal aortic aneurysm: three year results of the IMPROVE randomised trial. *BMJ* 2017; **359**: j4859.
19. Thompson SG, Ashton HA, Gao L, Buxton MJ, Scott RA, Multicentre Aneurysm Screening Study Group. Final follow-up of the Multicentre Aneurysm Screening Study (MASS) randomized trial of abdominal aortic aneurysm screening. *Br J Surg* 2012; **99**(12): 1649-56.
20. Ashton H, Buxton M, Day N, Kim L, Marteau T, Scott R. The Multicentre Aneurysm Screening Study (MASS) into the effect of abdominal aortic aneurysm screening on mortality in men: a randomised controlled trial. *The Lancet* 2002; **360**(9345): 1531-9.