Supplementary Online Content

The RESCAN Collaborators; Bown MJ, Sweeting MJ, Brown LC, Powell JT, Thompson SG. Surveillance intervals for small abdominal aortic aneurysms: a meta-analysis. *JAMA*. 2013;309(8):806-813.

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix.

Methods: Applying the Newcastle-Ottawa Scale

To apply the Newcastle-Ottawa Scale (NOS)[1] to our context, studies for AAA growth and rupture have been scored separately. For eTable 3, we take AAA measurements as the 'exposure' and AAA growth as the 'outcome'. For eTable 5, we take AAA measurements as the 'exposure' and AAA rupture as the 'outcome'.

'Comparability' in the NOS refers to confounding within studies, not heterogeneity between studies. Confounding is not an issue in our studies because we are estimating overall summaries of growth and rupture rates applicable to the whole small-AAA population in the study. Thus whether age or smoking, for example, affects growth and rupture rates does not matter; the only issue is whether the study is representative of the intended population (and that is the separate first item on the NOS). Furthermore, all our analyses in the paper are presented separately for men and women, so the mixture / exclusion of sexes in particular studies does not matter.

Points are given to studies as follows. The higher the total score, the better the quality of the study.

Selection:

1. Representativeness of exposed cohort:
    - 2 points for screening only studies
    - 1 point for screening and hospital studies (mixed)
    - 0 points for hospital based studies

2. Selection of non-exposed cohort: 1 point given to all studies, since no-one is non-exposed.

3. Ascertainment of exposure: All studies based upon clinic records but 1 point given to all studies with clear, specified AAA measurement protocol.

4. Outcome of interest not present at start: 1 point given to all studies since rupture was not present at baseline.

Comparability:

1. Comparability: 1 point given to all studies, since confounding is not relevant to our analyses and interpretation [see comments above].

Outcome:

1. Assessment of outcome:1 point only given to studies which had clear diagnostic criteria for rupture. This is not applicable for AAA growth (eTable 3).

2. Follow-up long enough: 1 point given to all studies, since all contribute relevant follow-up information in the small AAA range.

3. Adequacy of follow-up: 1 point only for studies with 'yes' in final column of eTable 4.

Reference

1. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 3rd Symposium on Systematic Reviews: Beyond the Basics, 2000. www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed 22 January 2013].

eAppendix. *(continued)*

Statistical Methods

AAA Growth

For each study separately, the AAA diameter for patient i at time t, $y_i(t)$, is modelled using a linear random-effects growth model:

$$\begin{aligned} & y_i(t) = b_{0i} + b_{1i}t + \varepsilon_i(t) \\ & (b_{0i}, b_{1i})^T \sim N_2\left((\beta_0, \beta_1)^T, \Sigma\right) \\ & \varepsilon_i(t) \sim N(0, \sigma_w^2) \\ & \Sigma = \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix} \end{aligned} \quad , \qquad (0.1)$$

where b_{0i} and b_{1i} are the random intercept and slope, respectively, for patient i. The random-effects have mean β_0 and β_1 and variance-covariance matrix Σ, and $\varepsilon_i(t)$ is the residual error term. The time-origin ($t = 0$) is defined as the time of the baseline measurement for each individual (first measurement recorded between 3.0 and 5.4cm).

Given this model, the expected growth rate, b_1, for an individual with a single diameter measurement u taken at baseline, can be expressed as follows:

$$E[b_1 \mid y\ (0) = u] = \beta_1 + \frac{\rho\sigma_0\sigma_1}{\sigma_0^2 + \sigma_w^2}(u - \beta_0). \qquad (0.2)$$

Bayesian Markov chain Monte Carlo (MCMC) methods are used to fit the model to each dataset. Independent N(0,10^6) priors are placed on β_0 and β_1, with $\sigma_w^{-2} \sim Gamma(0.001, 0.001)$ and $\Sigma \sim Wishart(R, 3)$ where R is the prior guess at the variance-covariance matrix[1]. Inferences are based on two parallel chains, with between 10,500 and 17,500 iterations per chain depending on the study being fitted (the data from some studies requiring longer runs). The first 500 to 4000 iterations from each chain were discarded as burn-in. The R hat diagnostic[2] was calculated for each parameter to assess convergence with a value close to 1 indicating good convergence properties. The median, standard deviation, and 2.5th and 97.5th percentiles of the posterior distribution are then obtained for the expected growth rate given baseline diameter u, using Equation (0.2).

A random-effects meta-analysis is then conducted in a second stage using the growth rates obtained from each study and separately for each possible millimetre increase in baseline diameter from 3.0 to 5.4cm. Specifically, let $x_k(u)$ be the posterior median estimate of growth for study k given baseline diameter u and $s_k(u)$ the corresponding posterior standard deviation. Then the random-effects meta-analysis model is

$$\begin{aligned} & x_k(u) \sim N(\theta_k(u), s_k^2(u)) \\ & \theta_k(u) \sim N(\phi(u), \tau^2(u)) \end{aligned} \quad , \qquad (0.3)$$

where $\tau^2(u)$ is the between-study heterogeneity variance and $\phi(u)$ is the overall pooled effect for baseline diameter u.

The estimated time after baseline for which there is a 10% chance of crossing the threshold for surgery (5.5cm) is calculated as follows. Firstly, the expected AAA diameter at time t is $\mu_y(t) = \beta_0 + \beta_1 t$, the variance is $v_y(t) = \sigma_0^2 + t^2\sigma_1^2 + 2t\rho\sigma_0\sigma_1 + \sigma_w^2$ and the covariance between a measurement taken at time $t > 0$ and one taken at baseline (on the same individual) is $c_y(t) = \sigma_0^2 + t\rho\sigma_0\sigma_1$. Using these results, the expected AAA diameter at time t given a single diameter measurement u taken at baseline is

eAppendix. ***(continued)***

$$\mu_{y|y(0)=u}(t) = \mu_y(t) + \frac{c_y(t)}{v_y(0)}\left(u - \mu_y(0)\right)$$

using the properties of a bivariate normal distribution, and the variance of the measurement at t given baseline diameter u is

$$v_{y|y(0)=u}(t) = v_y(t) - \frac{c_y(t)^2}{v_y(0)}.$$

Hence, the probability of a measurement at time t being over the threshold for surgery (5.5cm) given baseline diameter u can be found from the tail area of a Gaussian distribution, as follows:

$$P_T(t \mid y(0) = u) = 1 - \Phi\left(\frac{5.5 - \mu_{y|y(0)=u}(t)}{\sqrt{v_{y|y(0)=u}(t)}}\right), \qquad (0.4)$$

where Φ is the standard normal cumulative distribution function. The probability in Equation (0.4) is calculated repeatedly over a fine grid of times for different baseline diameters.

For each grid time, the median, 2.5th and 97.5th percentiles of the posterior distribution for P_T are obtained. Since quantiles are invariant to one-to-one transformations, the posterior median time at which, for example, there is a 10% chance of being over the threshold is the grid time at which the posterior median probability is closest to 0.10. Similarly, the posterior 95% credibility range are the grid times where the 2.5th and 97.5th percentiles are closest to 0.10.

To obtain pooled estimates from a random-effects meta-analysis, the log of the posterior median time is used as the effect estimate, whose standard error on the log scale is approximated as

$$s \approx \frac{(\log(t_{97.5}) - \log(t_{50}))}{1.96}, \qquad (0.5)$$

where t_{50} and $t_{97.5}$ are the posterior median and 97.5th percentile of time. A random-effects meta-analysis is then conducted using the log estimate of time and its standard error separately for the baseline diameters 3.0, 3.5, 4.0, 4.5, and 5.0cm as specified in Equation (0.3). This is also done repeatedly for various P_T, ranging from 1% to 15%.

eAppendix. *(continued)*

AAA Rupture

A joint model for the longitudinal (growth) and time-to-event (rupture) processes is used to estimate the rupture rate conditional on baseline diameter. The model is as follows:

$$\begin{aligned}
&\text{Longitudinal process}\\
&y_i(t) = b_{0i} + b_{1i}t + \varepsilon_i(t) = m_i(t) + \varepsilon_i(t)\\
&(b_{0i}, b_{1i})^T \sim N_2\left((\beta_0, \beta_1)^T, \Sigma\right)\\
&\varepsilon_i(t) \sim N(0, \sigma_w^2)\\
&\Sigma = \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix} \qquad , \qquad (0.6)\\
&\text{Time-to-event process}\\
&h_i(t) = \exp(\gamma + \alpha m_i(t))
\end{aligned}$$

where the longitudinal process is as specified previously in (0.1), and the time-to-event process is defined by the hazard of rupture $h_i(t)$ for patient i at time t. The association parameter, α, is the log hazard ratio for the effect of an underlying diameter on risk of rupture, whilst γ is the baseline log-hazard. A time-to-event process that used a non-constant baseline hazard was also investigated using a Weibull proportional hazards model. However, there was found to be no evidence from the data that the baseline hazard changed over time.

The hazard of rupture at baseline given baseline diameter *u* is therefore calculated as

$$h(0 \mid y(0) = u) = \exp(\gamma + \alpha u). \qquad (0.7)$$

Bayesian MCMC methods are used to fit the model . In addition to the parameters specified in the growth model we used the priors $\gamma \sim N(0,10^4)$ and $\alpha \sim N(0,10^4)$. Inferences are based on two parallel chains, with between 31,000 and 110,000 iterations per chain depending on the study being fitted (the data from some studies requiring longer convergence and runs). After excluding a burn-in and thinning the chain, inferences were based on 5,000 iterations per chain. The R hat diagnostic[2] was calculated for each parameter to assess convergence with a value close to 1 indicating good convergence properties. The median, standard deviation, 2.5th and 97.5th percentiles are obtained from the posterior distribution of the hazard (0.7) and log-hazard for each millimetre diameter from 3.0cm to 5.4cm. The posterior median log-hazard and its standard deviation are extracted from each study and combined in a random-effects meta-analysis as described in (0.3).

eAppendix. *(continued)*

Given baseline diameter, u, the predicted hazard, cumulative hazard and survival functions at time $t > 0$ are

$$h(t \mid y(0) = u) = \exp\left(\gamma + \alpha\ \mu_{y|y(0)=u}(t)\right)$$

$$H(t \mid y(0) = u) = \left[\frac{\partial \mu_{y|y(0)=u}(t)}{\partial t}\right]^{-1} \left(h(t \mid y(0) = u) - h(0 \mid y(0) = u)\right),$$

$$S(t \mid y(0) = u) = \exp\left(-H(t \mid y(0) = u)\right)$$

where the integrated hazard, H, can be written in a closed form since $\mu_{y|y(0)=u}(t)$ is a linear function of t. The probability of rupture by time *t* given baseline diameter *u* can therefore be written as

$$P_R(t \mid y(0) = u) = 1 - S(t \mid y(0) = u). \qquad (0.8)$$

The probability of rupture is evaluated over a fine grid of possible times. The posterior median, 2.5th and 97.5th percentiles of the probability P_R are obtained at each time. The posterior median time at which there is, for example, a 1% chance of rupturing is then the grid time at which the posterior median probability is closest to 0.01. Similarly, the posterior 95% credibility range are the grid times where the 2.5th percentile and 97.5th percentile of P_R are closest to 0.01.

Pooled estimates from a random-effects meta-analysis are obtained as before by taking the log of the posterior median time as the effect estimate, whose standard error on the log scale is approximated as in Equation (0.5). The meta-analysis is conducted separately for the baseline diameters 3.0, 3.5, 4.0, 4.5, and 5.0cm as specified in Equation (0.3). This is also done repeatedly for various P_R, ranging from 0.25% to 5%.

References for Statistical Methods

1. Sweeting MJ, Thompson SG. Making predictions from complex longitudinal data, with application to planning monitoring intervals in a national screening programme. *J Roy Stat Soc Ser A* 2012; 175: 569-586.
2. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Computnl Graph Statist* 1998: 7, 434-455.

eAppendix. *(continued)*

Results: Sensitivity analysis to length of follow-up

Restricting follow-up to a maximum of two years following the last AAA scan reduced the numbers of ruptures considerably, to 113 in 11,262 men and 39 in 1314 women. Compared to the unrestricted analysis, the pooled rupture rates in men were slightly lower for diameters 3.0 to 4.0cm (eFigure 7). There was little change in the pooled rupture rate for a 4.5cm AAA (increased from 3.2 to 3.3 per 1000 person-years) whilst the pooled rupture rate for a 5.0cm AAA increased from 6.4 to 8.3 per 1000 person-years. Between-study heterogeneity reduced considerably in this restricted analysis and this is reflected in narrower prediction intervals. For women, pooled rupture rates decreased slightly for 3.0cm and 3.5cm diameters, but remained relatively unchanged for 4.0-5.0cm diameters (eFigure 8).

eTable 1. Studies From Which Individual Patient Data Were Not Available

Author	Number of patients	Country	Study type
Brown[1]	476	Canada	observational
Bjorck (*unpublished*)	179	Sweden	Prospective observational
Cao[2]	178	Italy	RCT
Lederle[3]	567	USA	RCT
Santilli[4]	790	USA	Screening
Schlosser[5]	147	Netherlands	Prospective observational
Schouten[6]	150	Netherlands	Retrospective observational
Total	2487		

References for eTable 1

(1) Brown PM, Zelt DT, Sobolev B. The risk of rupture in untreated aneurysms: the impact of size, gender, and expansion rate. *J Vasc Surg* 2003; 37(2):280-284.

(2) Cao P, De Rango P, Verzini F, Parlani G, Romano L, Cieri E et al. Comparison of surveillance versus aortic endografting for small aneurysm repair (CAESAR): Results from a randomised trial. *Eur J Vasc Endovasc Surg* 2011 A.D.; 41(1):13-25.

(3) Lederle FA, Wilson SE, Johnson GR, Reinke DB, Littooy FN, Acher CW et al. Immediate repair compared with surveillance of small abdominal aortic aneurysms. *N Engl J Med* 2002; 346(19):1437-1444.

(4) Santilli SM, Littooy FN, Cambria RA, Rapp JH, Tretinyak AS, d'Audiffret AC et al. Expansion rates and outcomes for the 3.0-cm to the 3.9-cm infrarenal abdominal aortic aneurysm. *Journal of Vascular Surgery* 2002; 35(4):666-671.

(5) Schlosser FJ, Tangelder MJ, Verhagen HJ, van der Heijden GJ, Muhs BE, van der Graaf Y et al. Growth predictors and prognosis of small abdominal aortic aneurysms. *J Vasc Surg* 2008; 47(6):1127-1133.

(6) Schouten O, van Laanen JH, Boersma E, Vidakovic R, Feringa HH, Dunkelgrun M et al. Statins are associated with a reduced infrarenal abdominal aortic aneurysm growth. *Eur J Vasc Endovasc Surg* 2006; 32(1):21-26.

eTable 2. Summary of Individual Patient Datasets[$]

Study (reference)	Mean calendar year at baseline	Threshold for intervention (cm)	Measurement modalities used	Internal / External diameter measured	Number men / women	Mean follow-up (years) men/women	Number of small AAA ruptures men/women	Crude rupture rate (per 1000 person-years) men / women
Western Australia[1]	1997	5.0 or 5.5cm by centre	US only	External	685 / 0	8.24 / NA	4 / NA	0.71 / NA
Bournemouth, UK[2]	2002	5.5	US only	External	677 / 0	3.44 / NA	NA / NA	NA / NA
Chichester, UK[3]	1999	6.0 later 5.5	US only	Internal	1405 / 99	4.45 / 4.42	43 / 8	6.88 / 18.26
Edinburgh, UK[4]	NA*	5.5	US only	External	670 / 382	2.89 / 2.42	NA / NA	NA / NA
Gloucestershire, UK[5]	2000	5.5	US only	Internal	1981 / 0	4.70 / NA	34 / NA	3.65 / NA
Huntingdon, UK[6]	1995	4.5	US only	External	629 / 0	3.87 / NA	2 / NA	0.82 / NA
Leeds, UK[7]	2004	5.5	US & CT	External	220 / 47	3.27 / 3.14	NA / NA	NA / NA
Leicester, UK[8]	2002	5.5	US only	External	899 / 0	3.28 / NA	NA / NA	NA / NA
Manchester, UK[9]	2005	5.5	US only	External	837 / 258	2.41 / 2.41	6 / 5	2.97 / 8.03
MASS, UK[10]	1998	5.5	US only	Internal	1122 / 0	5.42 / NA	33 / NA	5.42 / NA
Tromso, Norway[11]	1995	5.5	US only	External	179 / 45	8.59 / 8.16	2 / 2	1.30 / 5.45
PIVOTAL, USA[12]	2007	5.0	US & CT	External	619 / 96	0.92 / 0.96	0 / 1	0.00 / 10.84
Propranolol, Canada[13]	1996	5.0 or 5.5 by centre	US only	External	460 / 88	2.47 / 2.39	3 / 0	2.64 / 0.00
Galdakao, Spain[14]	2001	5.0	US & CT	External	859 / 64	3.93 / 2.55	5 / 1	1.47 / 6.14
Stirling, UK[15]	2003	5.5	US & CT	No set protocol	331 / 125	3.08 / 3.34	4 / 5	3.92 / 11.98
Gävle, Sweden[16]	2003	5.0 or 5.5 by centre	US only	External	184 / 59	2.46 / 2.52	1 / 0	2.21 / 0.00
UKSAT, UK[17]	1993	5.5	US only	External	1747 / 480	2.38 / 2.65	32 / 28	7.68 / 22.00
Viborg, Denmark[18]	1996	5.0	US only	External	224 / 0	6.09 / NA	9 / NA	6.59 / NA

[$] Adapted from Sweeting et al[19]. NA - data not collected as part of original study. * Edinburgh study provided data at 6-month intervals only with no exact dates.

eTable 3. Quality Scores of the Studies of AAA Growth According to the Components of the Newcastle-Ottawa Scale, and Total Score (Higher Scores Indicate Better Quality)

	Selection[a]				Comparability[a]	Outcome[a]			**Total score**
Question	1	2	3	4	1	1	2	3	
Study:									
Western Australia[1]	2	1	1	1	1	n/a	1	0	7
Bournemouth, UK[2]	0	1	1	1	1	n/a	1	1	6
Chichester, UK[3]	1	1	0	1	1	n/a	1	1	6
Edinburgh, UK[4]	0	1	1	1	1	n/a	1	0	5
Gloucestershire, UK[5]	2	1	1	1	1	n/a	1	1	8
Huntingdon, UK[6]	2	1	0	1	1	n/a	1	1	7
Leeds, UK[7]	0	1	0	1	1	n/a	1	0	4
Leicester, UK[8]	2	1	0	1	1	n/a	1	0	6
Manchester, UK[9]	1	1	1	1	1	n/a	1	1	7
MASS, UK[10]	2	1	1	1	1	n/a	1	1	8
Tromso, Norway[11]	1	1	0	1	1	n/a	1	1	6
PIVOTAL, USA[12]	0	1	1	1	1	n/a	1	1	6
Propranolol, Canada[13]	0	1	1	1	1	n/a	1	0	5
Galdakao, Spain[14]	0	1	0	1	1	n/a	1	0	4
Stirling, UK[15]	0	1	1	1	1	n/a	1	1	6
Gavle, Sweden[16]	1	1	1	1	1	n/a	1	1	7
UKSAT, UK[17]	0	1	1	1	1	n/a	1	1	6
Viborg, Denmark[18]	1	1	1	1	1	n/a	1	1	7

[a]For definition of scores, see "**eAppendix: Applying the Newcastle-Ottawa Scale"** Possible total score range: 0 = worst quality, 8 = best quality; n/a = not applicable

eTable 4. Quality Indicators for Individual Patient Datasets

Study (reference)	Imaging protocol specified clearly	Professional observers	Measurement variability assessed	Study type	Explicit size intervention policy at start (intervention diameter cm)	Patient censorship defined systematically
Western Australia[1]	yes	yes	no	RCT	no	no
Bournemouth, UK[2]	yes	yes	no	obs	yes (5.5)	yes
Chichester, UK[3]	no	for the majority	later, not for RCT	RCT* + later obs	yes, initially 6.0 later 5.5	yes
Edinburgh, UK[4]	yes	yes	no	obs*	no	no
Gloucestershire, UK[5]	yes but changed during follow- up	yes	no	pop	yes (5.5)	yes
Huntingdon, UK[6]	no	yes	no	pop	yes (4.5)	yes
Leeds, UK[7]	no	yes	yes	obs	yes (5.5)	no
Leicester, UK[8]	no	yes	no	pop	variable, either 5.0 or 5.5	no
Manchester, UK[9]	yes	yes	yes	obs	yes men (5.5), women (5.0)	yes
MASS, UK[10]	yes	yes	yes	RCT	yes (5.5)	yes
Tromso, Norway[11]	no	yes	no	pop	no	yes
PIVOTAL, USA[12]	yes	yes	yes	RCT	yes	yes
Propranolol, Canada[13]	yes	yes	yes	RCT	variable, either 5.0 or 5.5	no
Galdakao, Spain[14]	no	yes	no	obs	yes (5.0)	no
Stirling, UK[15]	yes	yes	no	obs	yes (5.5)	yes
Gavle, Sweden[16]	yes	yes	no	RCT	no (usually 5.0 or 5.5)	yes
UKSAT, UK[17]	yes	yes	yes	RCT + obs**	yes (5.5)	yes
Viborg, Denmark[18]	yes	yes	yes	RCT + pop	yes (5.0)	yes

Study type: RCT, randomised controlled trial; pop, consecutive population screening observational; obs, consecutive hospital referrals observational (except **). All studies were prospective.
* Studies with pre-1990 quality imaging included
** Recruitment catchment for RCT, not consecutive series

eTable 5. Quality Scores of the Studies of AAA Rupture According to the Components of the Newcastle-Ottawa Scale, and Total Score (Higher Scores Indicate Better Quality)

	Selection[a]				Comparability[A]	Outcome[a]			**Total score**
Question	1	2	3	4	1	1	2	3	
Study:									
Western Australia[1]	2	1	1	1	1	0	1	0	7
Chichester, UK[3]	1	1	0	1	1	0	1	1	6
Gloucestershire, UK[5]	2	1	1	1	1	1	1	1	9
Huntingdon, UK[6]	2	1	0	1	1	0	1	1	7
Manchester, UK[9]	1	1	1	1	1	1	1	1	8
MASS, UK[10]	2	1	1	1	1	1	1	1	9
Tromso, Norway[11]	1	1	0	1	1	0	1	1	6
PIVOTAL, USA[12]	0	1	1	1	1	0	1	1	6
Propranolol, Canada[13]	0	1	1	1	1	0	1	0	5
Galdakao, Spain[14]	0	1	0	1	1	0	1	0	4
Stirling, UK[15]	0	1	1	1	1	0	1	1	6
Gavle, Sweden[16]	1	1	1	1	1	0	1	1	7
UKSAT, UK[17]	0	1	1	1	1	1	1	1	7
Viborg, Denmark[18]	1	1	1	1	1	1	1	1	8

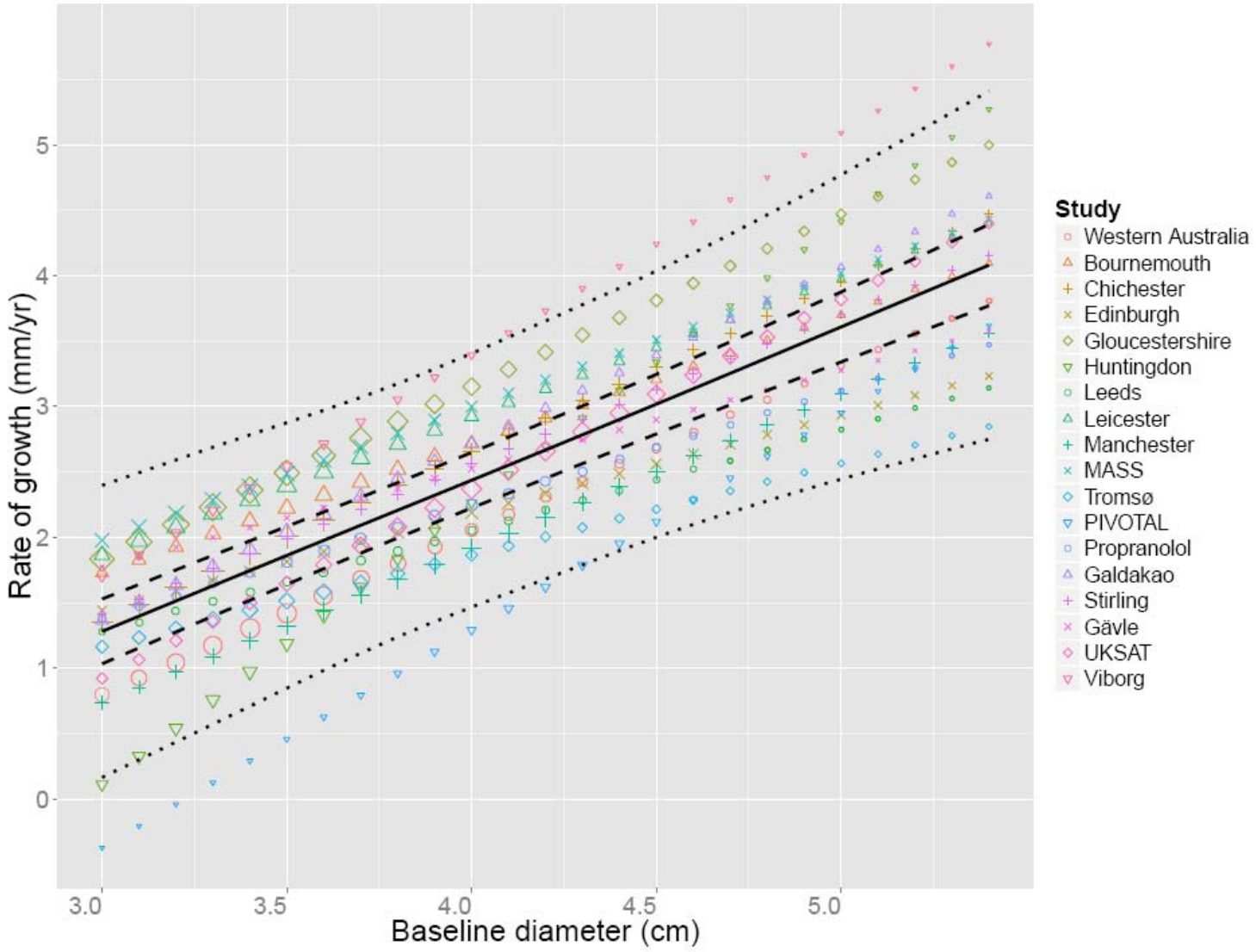
[a]For definition of scores, see "**eMethods: Applying the Newcastle-Ottawa Scale"** Possible total score range: 0 = worst quality, 9 = best quality

References for eTables 2 – 5

These references describe the study methodology. In several instances, studies have been extended since these publications so that the number of patients reported in the tables is greater than in the publication cited.

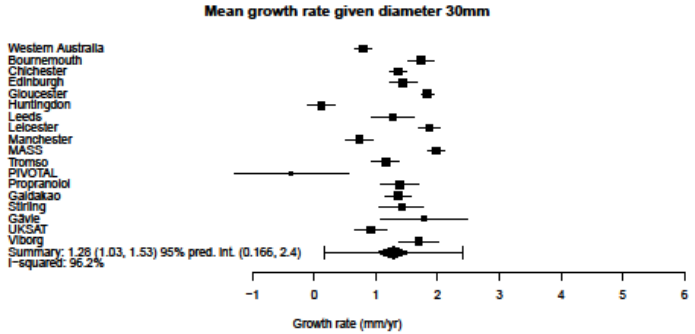
1. Norman PE, Jamrozik K, Lawrence-Brown MM, Le MT, Spencer CA, Tuohy RJ *et al.* Population based randomised controlled trial on impact of screening on mortality from abdominal aortic aneurysm. *BMJ* 2004; 329: 1259.
2. Parvin S (personal communication).
3. Ashton HA, Gao L, Kim LG, Druce PS, Thompson SG, Scott RA. Fifteen-year follow-up of a randomized clinical trial of ultrasonographic screening for abdominal aortic aneurysms. *Br J Surg* 2007; 94: 696-701.
4. Mofidi R, Goldie VJ, Kelman J, Dawson AR, Murie JA, Chalmers RT. Influence of sex on expansion rate of abdominal aortic aneurysms. *Br J Surg* 2007; 94: 310-314.
5. McCarthy RJ, Shaw E,Whyman MR, Earnshaw JJ, Poskitt KR, Heather BP. Recommendations for screening intervals for small aortic aneurysms. *Br J Surg* 2003; 90: 821-826.
6. Vardulaki KA, Prevost TC,Walker NM, Day NE, Wilmink AB, Quick CR *et al*. Growth rates and risk of rupture of abdominal aortic aneurysms. *Br J Surg* 1998; 85: 1674-1680.
7. Parry DJ, Al-Barjas HS, Chappell L, Rashid T, Ariëns RA, Scott DJ. Haemostatic and fibrinolytic factors in men with a small abdominal aortic aneurysm. *Br J Surg* 2009; 96: 870-877.
8. Salem M, Rayt HS, Hussey G, Raffelt S, Nelson CP, Sayers RD, et al. Should Asian men be included in abdominal aortic aneurysm screening programmes? *Eur J Vasc Endovasc Surg* 2009;38:748-9.
9. McCollum CN (personal communication) and see http://aaa.screening.nhs.uk/greater-manchester.
10. Ashton HA, Buxton MJ, Day NE, Kim LG, Marteau TM, Scott RA *et al*.; Multicentre Aneurysm Screening Study Group. The Multicentre Aneurysm Screening Study (MASS) into the effect of abdominal aortic aneurysm screening on mortality in men: a randomised controlled trial. *Lancet* 2002; 360: 1531-1539.
11. Solberg S, Singh K, Wilsgaard T, Jacobsen BK. Increased growth rate of abdominal aortic aneurysms in women. The Tromsø study. *Eur J Vasc Endovasc Surg* 2005; 29: 145-149.
12. Ouriel K, Clair DG, Kent KC, Zarins CK; Positive Impact of Endovascular Options for Treating Aneurysms Early (PIVOTAL) Investigators. Endovascular repair compared with surveillance for patients with small abdominal aortic aneurysms. *J Vasc Surg* 2010; 51: 1081-1087.
13. Propranolol Aneurysm Trial Investigators. Propranolol for small abdominal aortic aneurysms: results of a randomized trial. *J Vasc Surg* 2002; 35: 72-79.
14. Vega de Céniga M, Gómez R, Estallo L, Rodríguez L, Baquer M, Barba A. Growth rate and associated factors in small abdominal aortic aneurysms. *Eur J Vasc Endovasc Surg* 2006; 31: 231-236.
15. Holdsworth RJ (personal communication).
16. Karlsson L, Gnarpe J, Bergqvist D, Lindback J, Pärsson H. The effect of azithromycin and *Chlamydophilia pneumonia* infection on expansion of small abdominal aortic aneurysms – a prospective randomized double-blind trial. *J Vasc Surg* 2009; 50: 23-29.
17. Brady AR, Thompson SG, Fowkes FG, Greenhalgh RM, Powell JT; UK Small Aneurysm Trial Participants. Abdominal aortic aneurysm expansion: risk factors and time intervals for surveillance. *Circulation* 2004; 110: 16-21.
18. Lindholt JS, Juul S, Fasting H, Henneberg EW. Screening for abdominal aortic aneurysms: single centre randomised controlled trial. *BMJ* 2005; 330: 750.
19. RESCAN Collaborators (Sweeting MJ, Thompson SG, Brown LC, Powell JT). Meta-analysis of individual patient data to examine factors affecting growth and rupture of small abdominal aortic aneurysms. *Br J Surg* 2012; 99: 655-665.

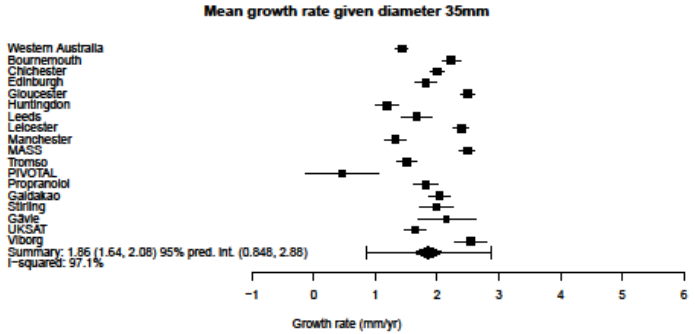
eFigure 1. Estimated Mean Growth Rates in Men

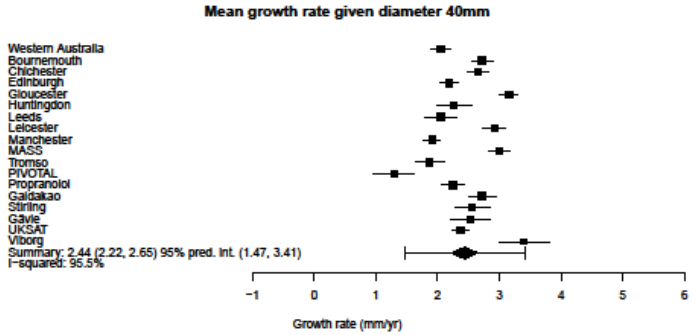


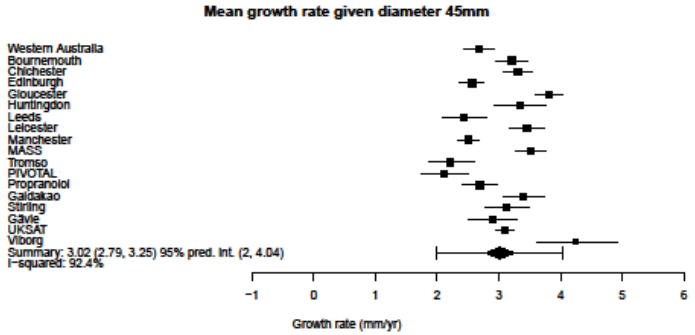
Given baseline AAA diameter, in men, by study and overall (with 95% confidence and prediction intervals). The area of each symbol is inversely proportional to its standard error.

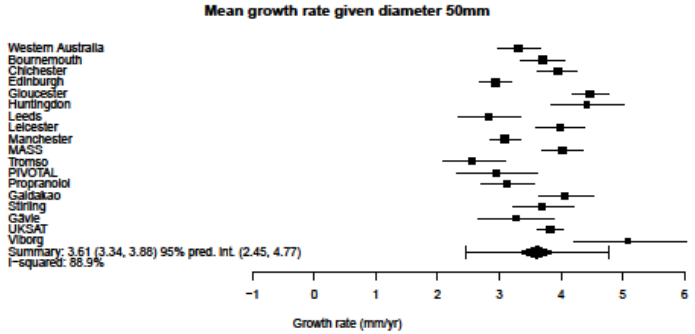
eFigure 2. Forest Plots of Mean AAA Growth Rate (Estimate and 95% CI) in Men According to Baseline AAA Diameter



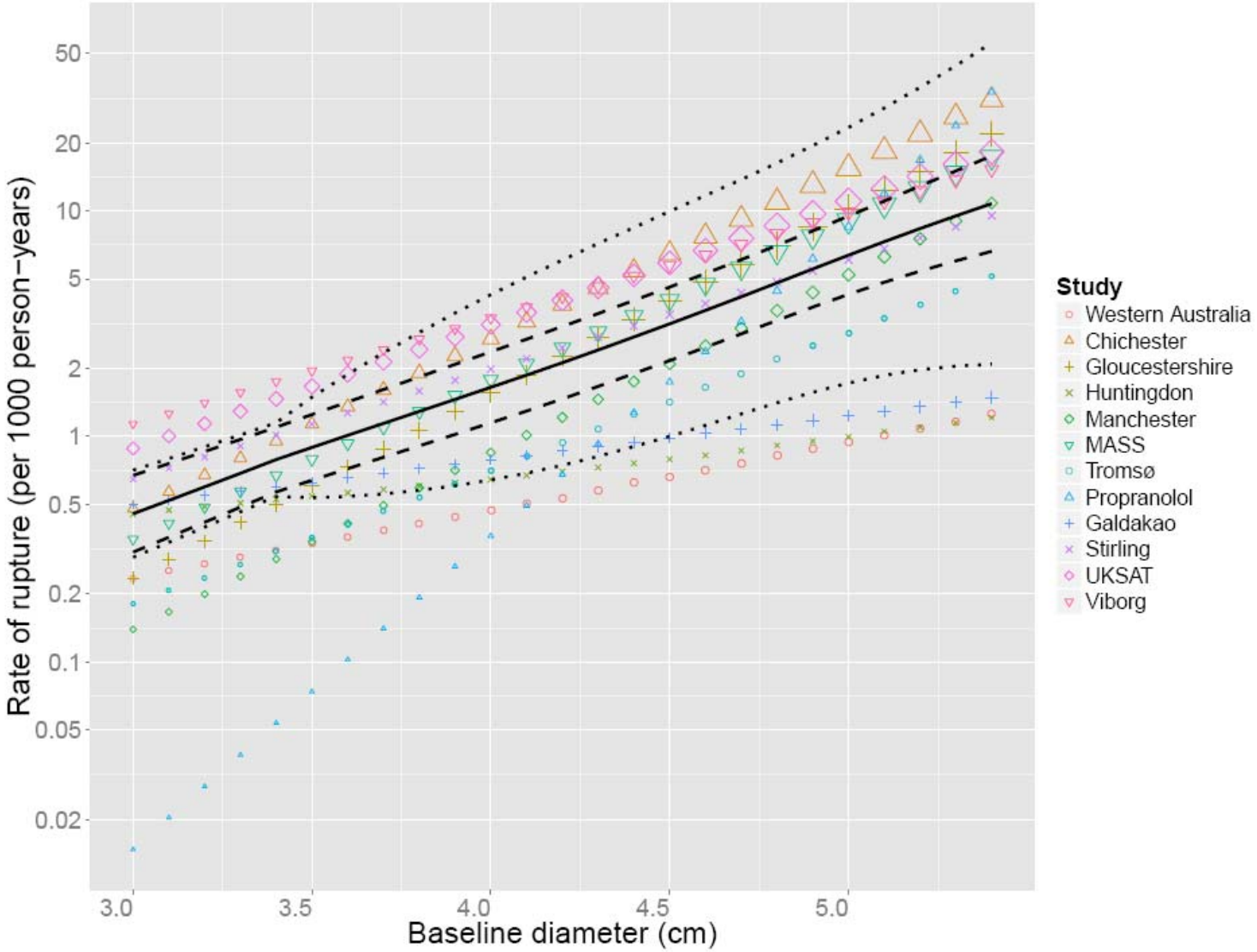






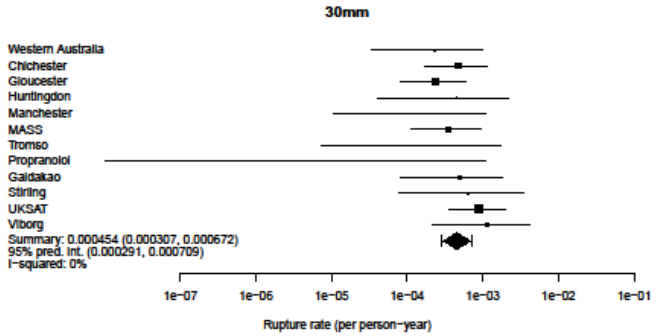


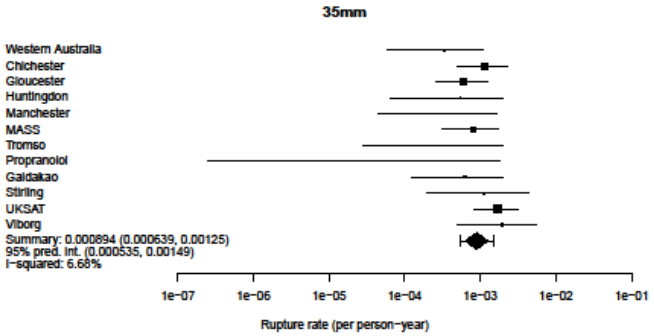
eFigure 3. Rupture Risk According to Underlying AAA Diameter in Men

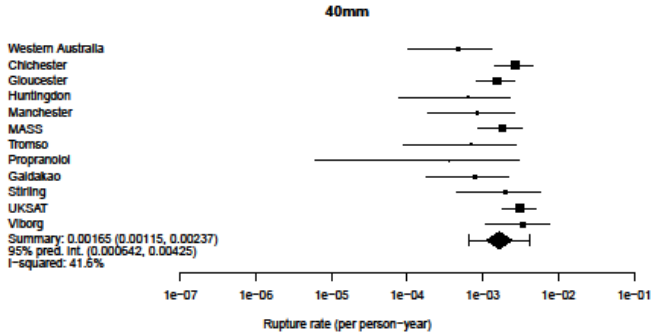


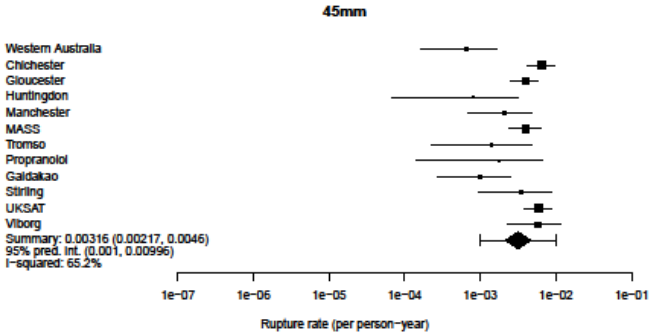
Data are by study and overall (log scale, with 95% confidence and prediction intervals). The area of each symbol is inversely proportional to its standard error.

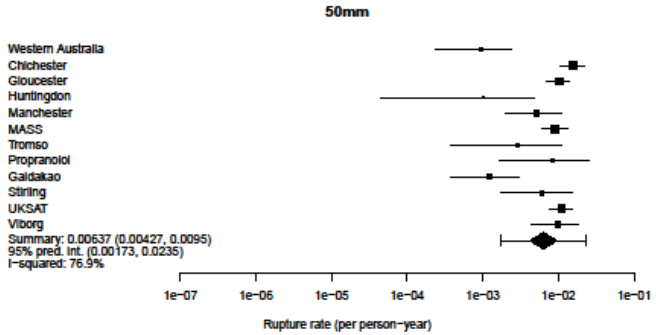
eFigure 4. Forest Plots of Rupture Risk (estimate and 95% CI, log scale) in Men According to Underlying AAA Diameter



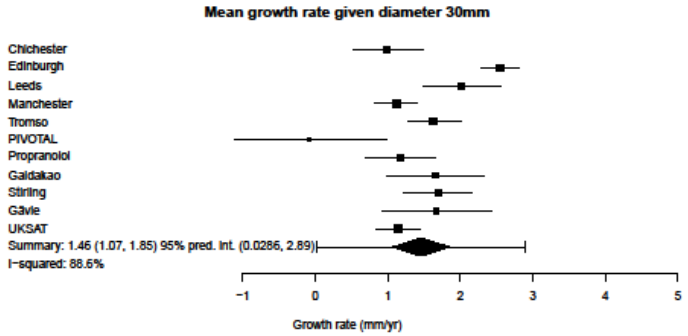


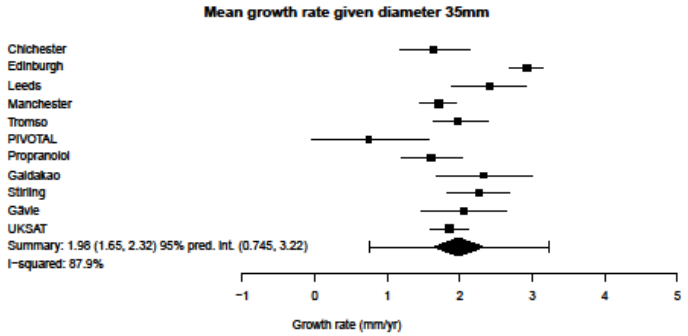


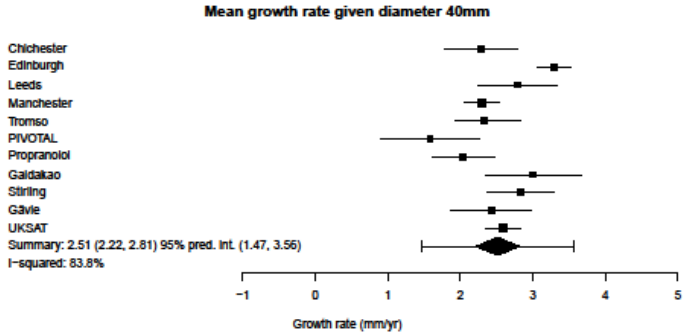


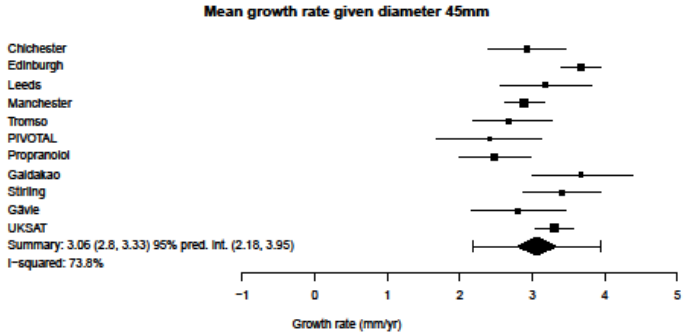


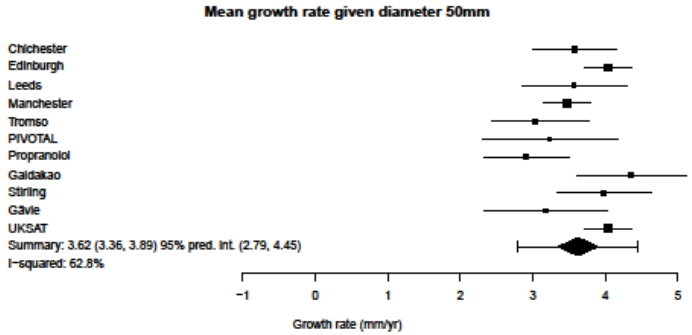
eFigure 5. Forest Plots of Mean AAA Growth Rate (estimate and 95% CI) in Women According to Baseline AAA Diameter



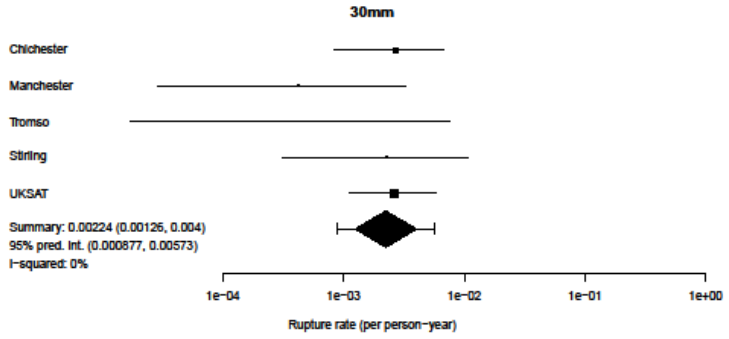


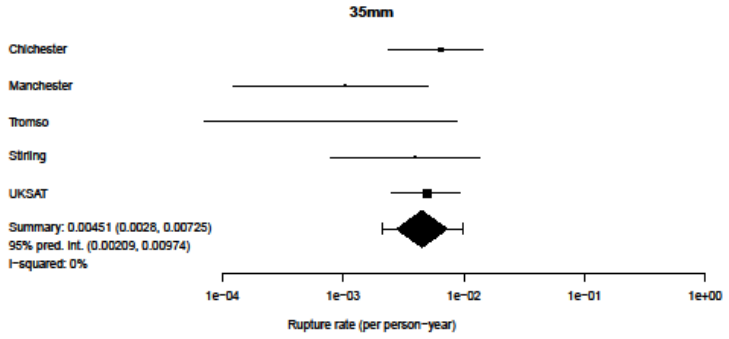


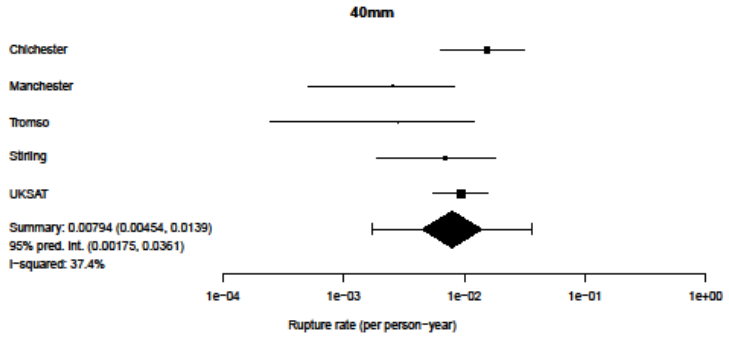


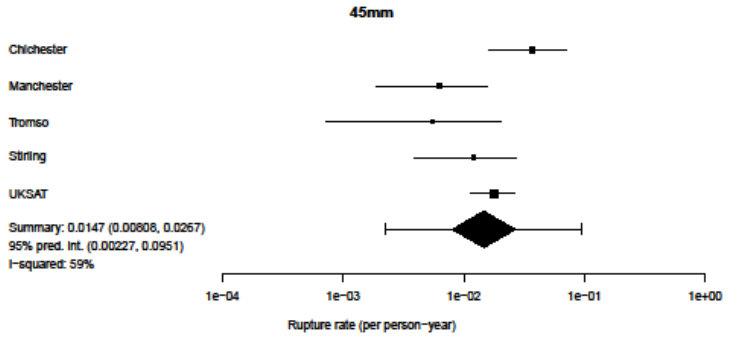


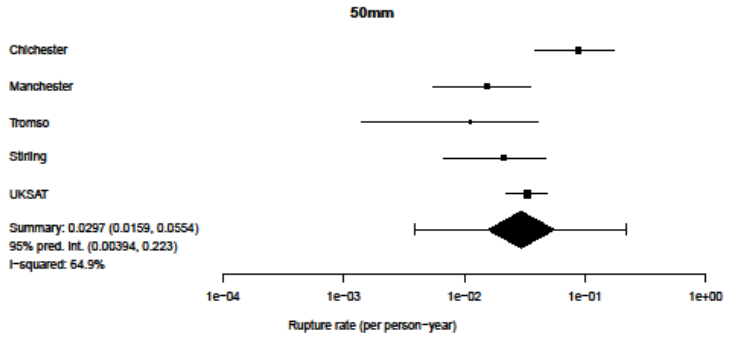
eFigure 6. Forest Plots of Rupture Risk (Estimate and 95% CI, Log Scale) in Women According to Underlying AAA Diameter



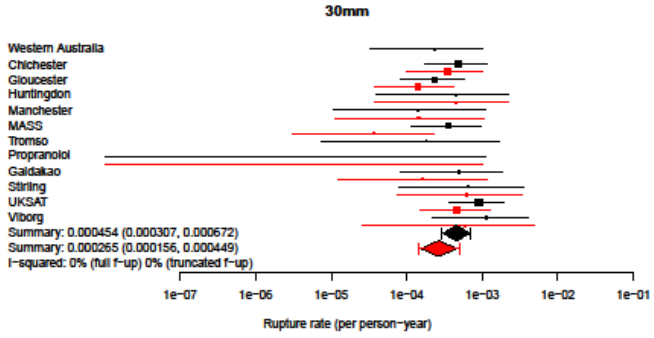


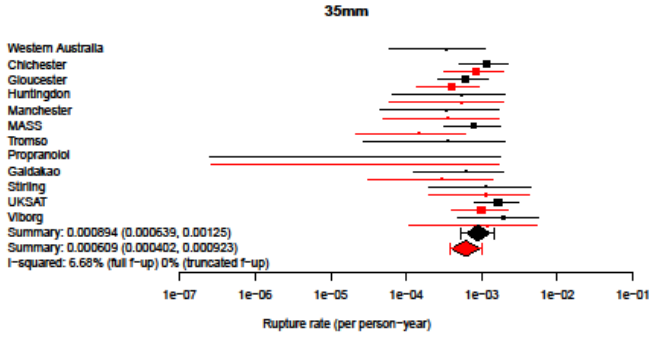


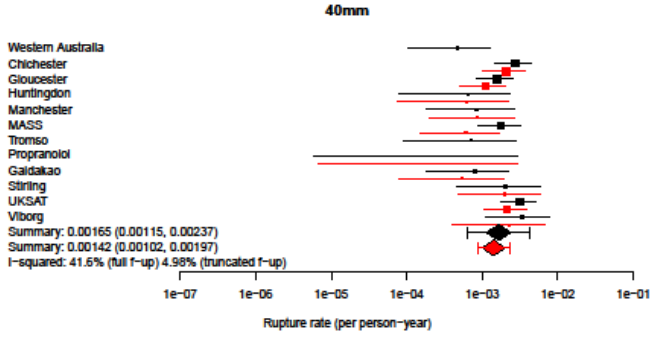


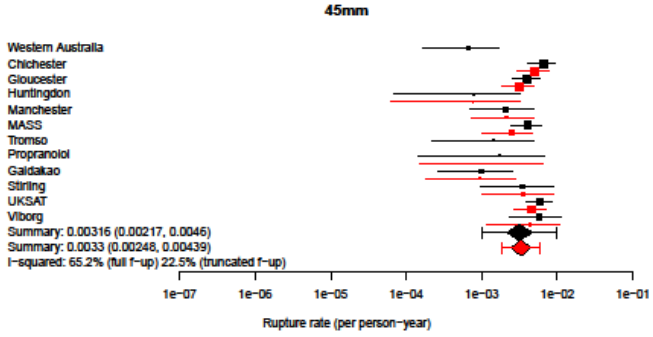


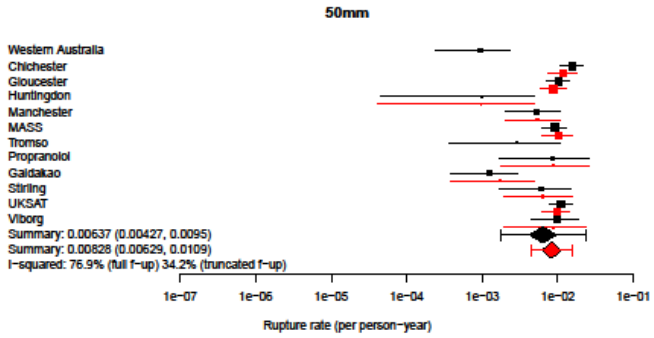
eFigure 7. Forest Plots of Rupture Risk (Estimate and 95% CI, Log Scale) in Men Comparing Main Analysis (Shown in Black) to Sensitivity Analysis Where Follow-up Is Restricted to Two Years After Last AAA Scan (shown in red)











eFigure 8. Forest Plots of Rupture Risk (Estimate and 95% CI, Log Scale) in Women Comparing Main Analysis (Shown in Black) to Sensitivity Analysis Where Follow-up Is Restricted to Two Years After Last AAA Scan (Shown in Red)