Michael C. Dalsing, MD, SECTION EDITOR

A systematic review of short-term vs long-term effectiveness of one-time abdominal aortic aneurysm screening in men with ultrasound

Muhammad Usman Ali, MD,[a,b] Donna Fitzpatrick-Lewis, MSW,[a,c] Meghan Kenny, MA,[a,b] John Miller, MD,[d] Parminder Raina, PhD,[a,b] *and* Diana Sherifali, RN, PhD,[a,c] *Hamilton, Ontario, Canada*

ABSTRACT

Background: An up-to-date systematic review on the long-term benefits of one-time abdominal aortic aneurysm (AAA) screening in men with ultrasound is required as new evidence is available. This report was produced for the Canadian Task Force on Preventive Health Care to provide evidence on screening for AAA with ultrasound. The aim of this systematic review was to examine the short-term (3-5 years of follow-up) vs long-term (13-15 years of follow-up) effectiveness of one-time screening for AAA in men.

Methods: This systematic review considered studies from the most recent U.S. Preventive Services Task Force review on AAA screening and passed through the screening process with citations identified in our search up to April 2017 (PROSPERO registration #CRD42015019047).

Results: Based on pooled estimates from four population-based randomized controlled trials with moderate-quality evidence, one-time AAA screening in men showed significant reductions in AAA-related mortality and AAA rupture rate, with a reduction of 43% for AAA-related mortality (risk ratio [RR], 0.57; 95% confidence interval [CI], 0.44-0.72; number needed to screen [NNS], 796) and 48% for AAA rupture rate (RR, 0.52; 95% CI, 0.35-0.79; NNS, 606) in short-term follow-up and a reduction of 34% for AAA-related mortality (RR, 0.66; 95% CI, 0.47-0.93; NNS, 311) and 35% for AAA rupture rate (RR, 0.65; 95% CI, 0.51-0.82; NNS, 264) in long-term follow-up. The effect on all-cause mortality was nonsignificant ($P = .14$) for short-term follow-up but marginally significant for long-term follow-up (RR, 0.99; 95% CI, 0.98-1.00; $P = .03$; NNS, 164). One-time AAA screening in men was also associated with a significant increase in the number of elective AAA-related procedures and a subsequent decrease in the number of emergency AAA procedures and 30-day postoperative mortality at both short-term and long-term follow-ups. We found no differences for one-time AAA screening in 30-day postoperative mortality due to elective and emergency operations compared with control groups.

Conclusions: Population-based one-time screening for AAA with ultrasound in asymptomatic men aged 65 years and older remains beneficial during the longer term after screening has ceased, with significant reductions in AAA mortality and AAA rupture rate, and hence avoids unnecessary AAA-related deaths. The sensitivity analyses also showed that the benefits of AAA screening were more pronounced in men at a mean age of <70 years with a relatively lower prevalence of AAA than in men at a mean age of >70 years with a relatively higher prevalence of AAA. Future research should explore the long-term benefits of a targeted AAA screening approach based on risk factors such as age, sex, smoking status, family history, aortic diameter, and baseline risk of rupture. (J Vasc Surg 2018;68:612-23.)

Keywords: Abdominal aortic aneurysm; Number needed to screen; Screening; Ultrasound; Mortality; Systematic review

Abdominal aortic aneurysm (AAA) continues to be a population health concern as it is estimated that >20,000 Canadians are diagnosed with AAA. It is an asymptomatic condition, and a ruptured AAA has an 80% mortality rate.[1] A recent systematic review[2] and report completed for the Canadian Task Force on Preventive Health Care demonstrated a significant reduction of 43% for AAA-related mortality and 48% for rupture at a short-term follow-up of 3 to 5 years using data from four population-based randomized controlled trials (RCTs; Multicentre Aneurysm Screening Study [MASS],[3-7] Chichester,[8-11] Viborg,[12-17] and Western Australia[18-20]). However, the long-term follow-up analysis is considered to be incomplete as data were not available from one of the larger trials for one-time AAA screening (ie, the Western Australia trial[20]).

Research examining the short-term vs long-term benefits of ultrasound screening for AAA in men older than 50 years was completed by Wilmink et al[7] for the Huntingdon Aneurysm Screening Programme (HASP). Their study concluded that after 5 years of screening,

From the McMaster Evidence Review and Synthesis Centre,[a] Department of Health Research Methods, Evidence and Impact, Faculty of Health Sciences,[b] School of Nursing, Faculty of Health Sciences,[c] and Department of Surgery, Faculty of Health Sciences,[d] McMaster University.

Author conflict of interest: none.

Correspondence: Diana Sherifali, RN, PhD, School of Nursing, Faculty of Health Sciences, McMaster University, 1280 Main St W, Health Sciences Centre 3N28F, Hamilton, Ontario, Canada L8S 4K1 (e-mail: dsherif@mcmaster.ca).

The editors and reviewers of this article have no relevant financial relationships to disclose per the JVS policy that requires reviewers to decline review of any manuscript for which they may have a conflict of interest.

0741-5214

https://doi.org/10.1016/j.jvs.2018.03.411

the reduction in the incidence of ruptured AAA was 40% (95% confidence interval [CI], 3%-74%; number needed to screen [NNS], 1380). However, after 13 years of screening, the reduction in incidence of ruptured AAA was 73% (95% CI, 58%-82%), and the NNS decreased to 505 men.[7] This warranted the need to update and to review the evidence after inclusion of the recently published long-term follow-up data from the Western Australia trial.[20]

The aim of this updated systematic review was to compare the short-term and long-term effectiveness of one-time AAA screening with ultrasound, which we were unable to perform in our earlier publication on screening for AAA in asymptomatic adults[2] as the long-term data from the Western Australia trial were not available. We also aimed to explore the possible effect of various population- and study-level factors on outcomes of interest, specifically AAA-related mortality, all-cause mortality, AAA rupture rates, AAA-related procedures, and 30-day postoperative mortality. These data will further support the Canadian Task Force on Preventive Health Care in updating its previous (1991) recommendation on AAA screening.[21]

METHODS

For full details, see the original PROSPERO CRD42015019047. The same methods were used to update this review, and similar methods have been used by and are reported in other publications authored by our review team.[22-24]

Search strategy. The original literature search updated the search done for the 2014 U.S. Preventive Services Task Force review on screening of AAA using the same search strategy; for the purposes of this update, an additional update was done to include all possible long-term data.[2] We searched MEDLINE, Embase, and Cochrane Central Register of Controlled Trials. We also searched PubMed for any relevant publisher-supplied nonindexed citations. The searches covered the time period since the last update of the U.S. Preventive Services Task Force search (April 2015-April 2017). English and French studies as well as reference lists of on-topic systematic reviews were reviewed. Inclusion and exclusion criteria have been published elsewhere and were consistently applied in this update.[2]

Study selection, data abstraction, and quality assessment. Two reviewers independently selected studies for possible inclusion. At the title and abstract level, any citation that was selected for inclusion by either reviewer moved to full-text review. At that level, any disagreement was discussed between reviewers, and a third party was involved to help reach consensus as necessary. Full data extraction, including characteristics of included studies and risk of bias (assessed using the Cochrane risk of bias framework),[25] was completed by one reviewer and verified by a second reviewer. Disagreements were resolved through consensus between the two reviewers. The Grading of Recommendations Assessment, Development, and Evaluation (GRADE)[26] system was used to assess the strength and the quality of evidence using GRADEPro software.[27] The quality of outcome-based bodies of evidence was assessed for risk of bias due to limitations in design, indirectness, inconsistency of findings, imprecision, and reporting bias (such as publication bias). Meta-analyses were conducted where appropriate.

Data synthesis. For the primary outcomes of effectiveness of one-time AAA screening (ie, AAA-related mortality, all-cause mortality, AAA rupture rates, AAA-related procedures, and 30-day postoperative mortality), we used number of events to generate the summary measures of effect in the form of risk ratio (RR) by DerSimonian and Laird random-effects models with Mantel-Haenszel method.[28] The primary subgrouping in each meta-analysis was based on length of follow-up, that is, short-term (3-5 years) and long-term (13-15 years) follow-up. The estimates of absolute risk reduction (ARR), absolute risk increase (ARI), NNS, and number needed to operate (NNO) were added. The NNS and NNO were estimated using the control group event rate and pooled RR with the 95% CI obtained from the meta-analysis.[29]

To evaluate statistical stability and robustness of results and to account for any potential bias (such as CIs being inappropriately wide) in pooled estimates, we performed sensitivity analysis based on type of pooling method, that is, DerSimonian-Laird random-effects model, fixed-effects model, Peto one-step odds ratio, and time to event (hazard ratios) model.[30,31] To explore heterogeneity across included studies, we further performed metaregression and subgroup analyses for various population- and study-level factors that may influence the pooled effect estimates, such as mean age at screening, country, year of recruitment, screening adherence rate, study risk of bias, study sample size, baseline AAA prevalence, and baseline prevalence of AAA ≥5.5 cm.[32] The Cochran Q ($\alpha = .05$) was employed to detect statistical heterogeneity, and the I^2 statistic quantified the magnitude of statistical heterogeneity between studies, where I^2 of 30% to 60% represents moderate and I^2 of 60% to 90% represents substantial heterogeneity across studies.[33] All analyses were performed using Review Manager (RevMan version 5.3; The Nordic Cochrane Center, Copenhagen, Denmark), Stata (version 14; StataCorp LP, College Station, Tex), and GRADEpro Guideline Development Tool (McMaster University, Hamilton, Ontario, Canada) software packages.

RESULTS

Search results

After removal of duplicates, 677 citations were identified. At title and abstract screening, we excluded 489