International Journal of Epidemiology, 2015, 264–277
doi: 10.1093/ije/dyu140
Advance Access Publication Date: 15 January 2015
Original article

Public Health Policy

Does screening for disease save lives in asymptomatic adults? Systematic review of meta-analyses and randomized trials

Nazmus Saquib,[1] Juliann Saquib[1] and John PA Ioannidis[1,2,3]*

[1]Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA [2]Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA, USA and [3]Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, CA, USA

*Corresponding author. Stanford Prevention Research Center, Medical School Office Building, Room X306, 1265 Welch Rd, Stanford, CA 94305, USA. E-mail: jioannid@stanford.edu

Accepted 16 June 2014

Abstract

Background: Several popular screening tests, such as mammography and prostate-specific antigen, have met with wide controversy and/or have lost their endorsement recently. We systematically evaluated evidence from randomized controlled trials (RCTs) as to whether screening decreases mortality from diseases where death is a common outcome.

Methods: We searched three sources: United States Preventive Services Task Force (USPSTF), Cochrane Database of Systematic Reviews, and PubMed. We extracted recommendation status, category of evidence and RCT availability on mortality for screening tests for diseases on asymptomatic adults (excluding pregnant women and children) from USPSTF. We identified meta-analyses and individual RCTs on screening and mortality from Cochrane and PubMed.

Results: We selected 19 diseases (39 tests) out of 50 diseases/disorders for which USPSTF provides screening evaluation. Screening is recommended for 6 diseases (12 tests) out of the 19. We assessed 9 non-overlapping meta-analyses and 48 individual trials for these 19 diseases. Among the results of the meta-analyses, reductions where the 95% confidence intervals (CIs) excluded the null occurred for four disease-specific mortality estimates (ultrasound for abdominal aortic aneurysm in men; mammography for breast cancer; fecal occult blood test and flexible sigmoidoscopy for colorectal cancer) and for none of the all-cause mortality estimates. Among individual RCTs, reductions in disease-specific and all-cause mortality where the 95% CIs excluded the null occurred in 30% and 11% of the estimates, respectively.

Conclusions: Among currently available screening tests for diseases where death is a common outcome, reductions in disease-specific mortality are uncommon and reductions in all-cause mortality are very rare or non-existent.

Key Messages

- We evaluated the evidence on 39 screening tests for 19 diseases where mortality is a common outcome.
- We found 48 randomized controlled trials and 9 meta-analyses that addressed either disease-specific or all-cause mortality.
- Reductions in disease-specific mortality were uncommon and reductions in all-cause mortality were very uncommon, or even non-existent with these screening tests.

Introduction

Screening for disease is a key component of modern health care. The rationale is simple and attractive—to detect diseases early in asymptomatic individuals and to treat them in order to reduce morbidity, mortality and the associated costs. However, the role of screening often comes into question. Some high-profile controversies have appeared lately in this regard. For example, for breast cancer, the United States Preventive Services Task Force (USPSTF) currently recommends against routine mammographic screening for women aged 40–49 years after retracting its previous recommendation in favour of mammography, as the data failed to show that benefit outweighed harm.[1] The decision against screening drew sharp criticism from various interest groups including patients who overestimate the benefit of screening.[2] Similarly, USPSTF now recommends against screening for prostate cancer in healthy men because harms from prostate specific antigen (PSA) screening exceed the benefit, trials do not show improvement in long-term survival[3] and screening carries a high risk of over-diagnosis with adverse consequences. Again, heated debates have been generated around this change of recommendation, both in the scientific and the popular press.

Some screening tests were entrenched in clinical and public health practice before randomized controlled trials (RCTs) became widely used. As the screening agenda encompasses a large number of tests, and new ones are continuously proposed, it is useful to reassess the evidence supporting their use. Our research question is whether recommended screening tests, among asymptomatic adults, have evidence from RCTs on mortality for diseases where death is a common outcome. In particular, is there evidence of mortality reduction, either disease-specific or all-cause, from screening? To this end, we have compiled and examined systematically the evidence from individual RCTs and meta-analyses thereof for screening tests that have been proposed for detecting major diseases in adults who have no symptoms.

Methods

Eligibility criteria

We assessed the diseases/disorders in adults, which USPSTF grouped in different clinical categories and made screening recommendations. We focused on the 'Cancer' and 'Heart and vascular diseases' categories, as well as type 2 diabetes mellitus and chronic obstructive pulmonary disease, because mortality is a common outcome for these diseases. We did not include diseases/disorders where mortality is not a common outcome, and that included the following clinical categories: infectious diseases; mental health conditions and substance abuse; metabolic, nutritional and endocrine disorders (except type 2 diabetes); musculoskeletal disorders; injury and violence; vision and hearing disorders; obstetric and gynaecological conditions; and miscellaneous (except chronic obstructive pulmonary diseases).

For the included diseases, we compiled a list of screening tests and assessed which of them are recommended by USPSTF, and whether they have randomized evidence on mortality outcomes. We defined screening as using a specific test on an otherwise asymptomatic, non-diseased population in order to detect a certain disease. We only considered evidence that compared mortality between screening and no-screening control groups. We did not consider screening/testing in already diseased individuals (e.g. patients who have diabetes mellitus or already have some cancer diagnosis).

Search strategies and documentation of evidence

We compiled information from USPSTF, Cochrane Database of Systematic Reviews and PubMed. We documented current recommendations and the corresponding level of evidence from USPSTF. We gathered meta-analytic evidence on screening from Cochrane and PubMed. In addition, we collected from PubMed information about individual RCTs on screening which had not been included in a published meta-analysis.

In the USPSTF website, we reviewed the documentation of RCT evidence for screening for each disease in adults (last update: January 2014).

We searched the Cochrane Database of Systematic Reviews using the search term 'screening' in title, abstract or keywords. We documented all systematic reviews on screening tests that had at least one eligible RCT or meta-analysis of several RCTs with mortality outcomes (last update: January 2014).

We searched PubMed using the search terms 'screen' or 'screening' or 'testing' in the title and 'death' or 'mortality' or 'survival' in title, abstract or keywords. We ran two searches for articles published in English; one was limited to RCTs and the other to meta-analyses (using 'type of publication' limit) (last update: January 2014).

Study selection: meta-analyses and individual trials

We screened, identified and organized the eligible meta-analyses by disease and the associated screening test (Cochrane, PubMed). When several meta-analyses were eligible on the same disease and screening test, we selected the most comprehensive ones (more trials, more long-term follow-up in included trials).

We screened and identified the eligible individual trials (PubMed). We organized the list first by screening test, then by trial name and then by year of publication. If there were more than one citation per trial, we selected the most recent publication. Simultaneously, we compiled a list of trials that were in the selected meta-analyses. We cross-checked the individual trials in PubMed with those in the selected meta-analyses to determine how many trials were in common. Finally, we compiled a list of individual trials—including those that were in common and those that were unique to either PubMed or a meta-analysis. Finally, if no meta-analysis was available, but multiple individual trials existed for a given screening test, we performed the meta-analysis ourselves using inverse variance synthesis with fixed effects.

Data extraction

From USPSTF, we documented the following for each disease: screening tests, recommendation statement, category of evidence, presence or absence of RCT evidence, and the specific population for whom the recommendation is applicable.

For each included meta-analysis of RCTs (Cochrane or PubMed) and single RCT (PubMed), we extracted the following: disease; screening intervention assessed; number of RCTs analysed; use of stratified analysis (yes, no) and, if so, types of strata; number of disease-specific deaths/total sample and disease-specific mortality risk estimates; and number of total deaths/total sample and all-cause mortality risk estimates. Data were extracted by two co-authors and any disagreement was resolved with discussion with the senior (third) author.

Presentation of mortality outcomes

Disease-specific mortality was defined as death attributed to the disease in question and all-cause mortality was defined as death from any cause; in both instances, the denominator was the total sample per randomized group and not those who were detected as diseased. We presented the treatment effect (risk estimates with 95% confidence interval) as they were reported in the original RCTs or meta-analyses.

Results

Evaluated screening tests in USPSTF

USPSTF provides evaluation of screening for 19 diseases where mortality is a common outcome (cancer $n = 12$, heart and vascular diseases $n = 5$, type 2 diabetes, chronic obstructive pulmonary disease) (Supplementary Table 1, available as Supplementary data at *IJE* online).

Overall, 39 different screening tests are addressed for these 19 diseases. Screening is recommended for 6 of the 19 diseases (for a total of 12 recommended tests out of 14 available tests for these 6 diseases). Randomized evidence with a mortality outcome is cited for only 5 diseases (breast, cervical and colorectal cancer, abdominal aortic aneurysm and type 2 diabetes) for 11 recommended tests among 13 assessed by USPSTF (Figure 1).

Randomized trials with data on mortality are not available for one disease (hypertension) where screening is recommended (one out of one test is recommended). Further, BRCA-gene mutation screening for breast cancer[4] and colonoscopy for colorectal cancer do not have randomized trials on their effectiveness, but they are both currently recommended for adults with a family history.

Screening is not recommended for the remaining 13 diseases where there are 25 available tests; of those, randomized trials with data on mortality are available only for 7 tests on 4 diseases: lung, oral, ovarian and prostate cancer. For breast cancer, screening for BRCA and mammography are recommended but clinical and self-examination of the breast are not recommended. Randomized evidence exists for mammography,[5] clinical and self-examination[6,7] but not for BRCA (there is a trial on genetic counselling but not for the screening test per se).

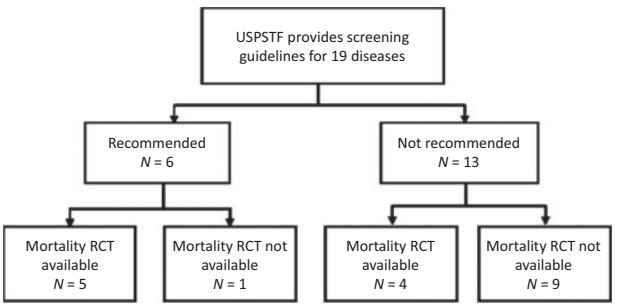

Figure 1. Flow diagram for randomized controlled trial (RCT) evidence from the United States Preventive Services Task Force (USPSTF).

Evaluated meta-analyses on screening tests in Cochrane and PubMed

The search produced 595 items in Cochrane and 125 items in PubMed; of those, 59 and 85, respectively, were assessed in full text. In Cochrane and PubMed, 12 and 44 meta-analyses, respectively, met the eligibility criteria; these included 8 Cochrane reviews that had also been presented as journal articles, thus there were 48 different eligible meta-analyses. These 48 meta-analyses were clustered by test and disease to identify the latest, non-overlapping meta-analysis on each topic. Eventually, eight meta-analyses were selected covering eight screening tests for six diseases;[3,5,6,8–12] additionally we performed ourselves a meta-analysis of the trials' data on screening with computer tomography (CT) for lung cancer, as there were several individual trials but no published meta-analyses (Figure 2).

Evaluated individual trials on screening tests in PubMed

The search produced 590 items; 83 records were evaluated further and 40 trials met the inclusion criteria. Of the 40, 28 trials had been included in at least one of the eight eligible meta-analyses mentioned above.[13–40] The other 12 trials[7,41–51] found in PubMed included mostly ($n = 9$) trials on topics for which there were no eligible previous meta-analyses; three trials[42,44,45] were excluded from the respective meta-analysis because the follow-up time was less than 5 years. Another eight trials[52–59] that had been included in the eight eligible meta-analyses were not captured by the PubMed search for trials; these were not picked by our PubMed search because one was in Russian language, two were not tagged as randomized controlled trials by PubMed and five did not have the search terms in their titles. Therefore, a total of 48 eligible RCTs were considered (Figure 3).

Meta-analytic and individual trial evidence by disease

Abdominal aortic aneurysm. Eight meta-analyses were found; we used the meta-analysis by Takagi *et al.*[12] that had included four trials (Chichester,[16] MASS,[60] Viborg.[13] Western Australia[35]) with the longest follow-up (≥10 years).

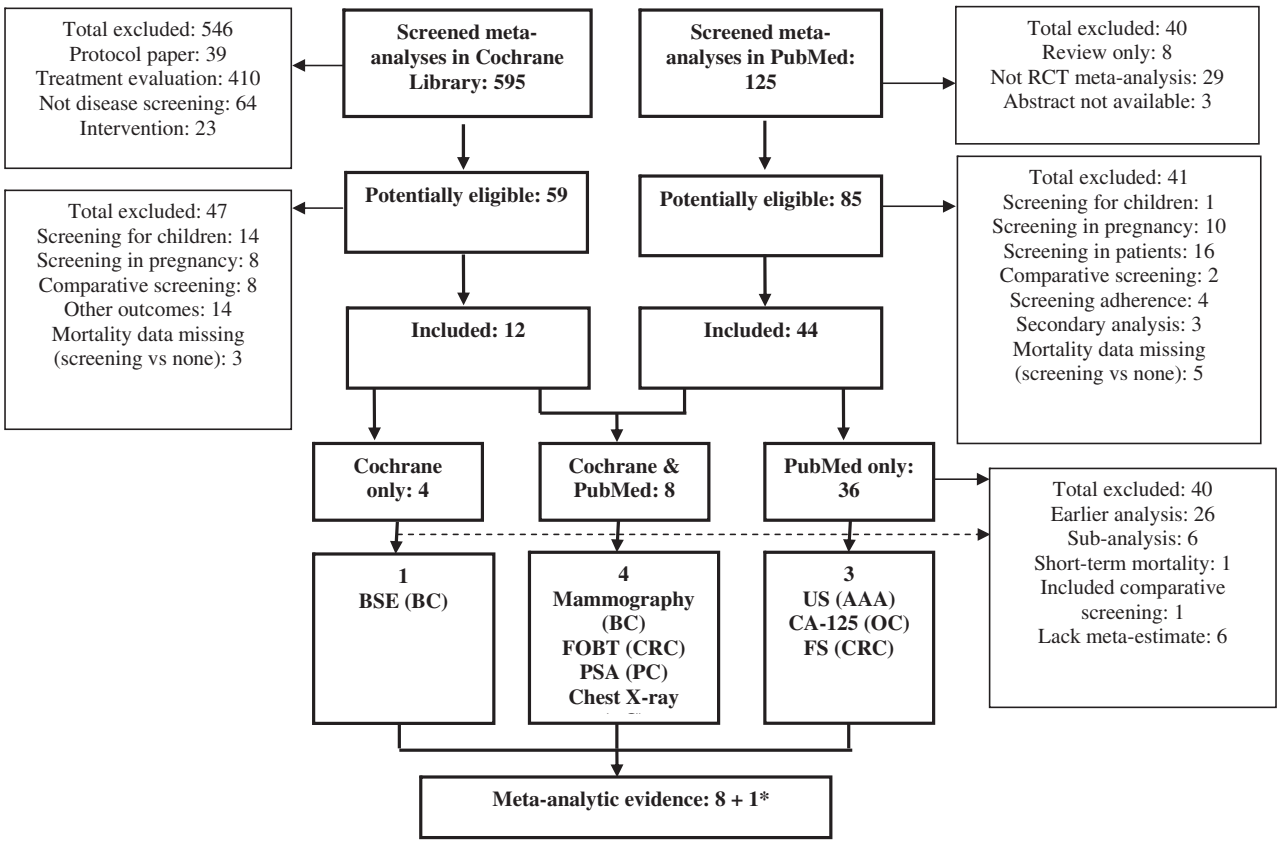

Figure 2. Flow diagram of meta-analytic search results from Cochrane Library and PubMed. RCT, randomized controlled trial; BSE, breast self-examination; BC, breast cancer; FOBT, fecal occult blood test; CRC, colorectal cancer; PSA, prostate-specific antigen; PC, prostate cancer; LC, lung cancer; US, ultrasound; AAA, abdominal aortic aneurysm; OC, ovarian cancer; FS, flexible sigmoidoscopy.

*Meta-analytic evidence: we conducted the meta-analyses for screening with computerized tomography (CT) scan using data from DANTE, DLCST and MILD trial (see Table 2).

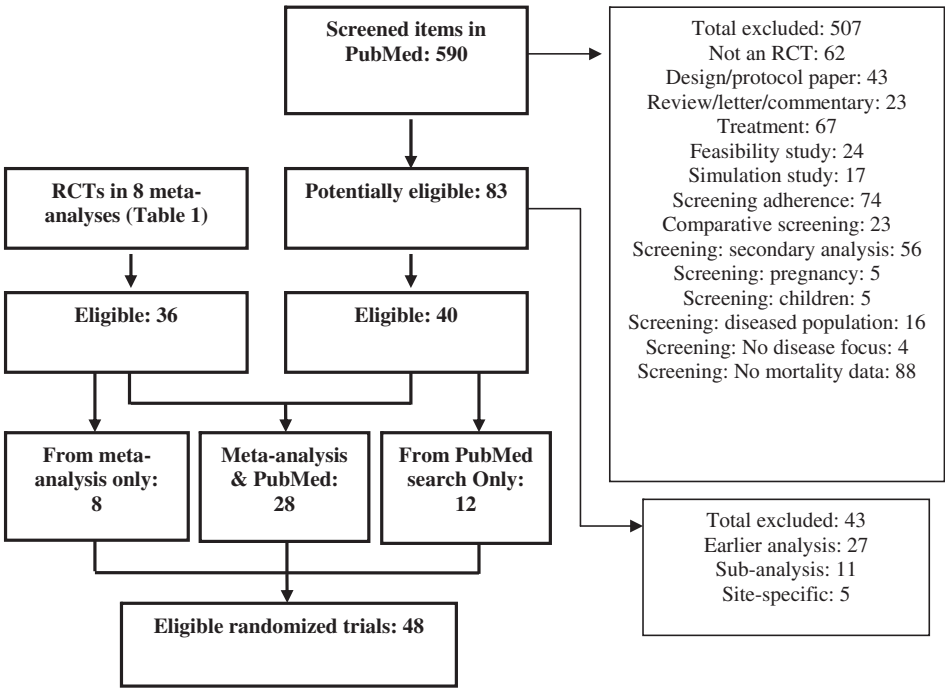

Figure 3. Flow diagram of individual trial evidence from PubMed search results and selected meta-analyses.

In addition, MASS[18] has published extended follow-up data (13 years) after the Takagi meta-analysis. The reason for excluding six meta-analyses was shorter follow-up.[61–66] The final one was excluded because it evaluated a 30-day mortality following elective surgery for aortic aneurysm.[67]

Breast cancer. Twelve meta-analyses of screening with mammography were found; Gotzsche *et al.*[5] had reviewed all eight trials (Canada 1980a,b,[17,68] Edinburgh,[14] Goteborg,[15] Malmo,[19] New York,[26] Stockholm,[20] Two-county[27] and UK age trial[24]) and reported the longest follow-up time (13 years). The other meta-analyses were excluded because they were earlier publications,[69–71] had fewer trials[72–76] or shorter follow-up[77] or selected a particular age group[78–82] or a sub-type of cancer.[83] For the Two-county study, Tabar *et al.* presented also disease-specific mortality estimates with longer follow-up (29 years);[27] the trial's estimates for all-cause mortality were extracted from Gotzsche *et al.*[5]

Only one meta-analysis[6] was found for screening with breast self-examination with two trials (Russia[58] and Shanghai[59]) and only a single trial (Mumbai[7]) for clinical breast examination.

Cervical cancer. Two single trials (Tamil Nadu[46] and Maharashtra[47] in India) were found on screening with visual inspection, human papilloma virus testing and cytological testing for cervical cancer.

Colorectal cancer. Four meta-analyses of screening with fecal occult blood (FOBT) test were found; Hewitson *et al.*[8] presented four trials (Funen,[21] Goteborg,[23] Minnesota [84] and Nottingham[25]) with the longest follow-up (11.7 to 18 years); the other three were excluded for including fewer trials[85] or shorter follow-up time.[86,87] After the Hewitson meta-analysis was published, the Minnesota study has published a 30-year follow-up.[49] Two meta-analyses[9,88] of screening with flexible sigmoidoscopy (single, multiple or in combination with FOBT) were also found; both included five trials: Telemark Polyp Study,[54] NCCPS,[31] UK trial,[29] SCORE,[40] and PLCO.[38] One meta-analysis was excluded because it did not provide all-cause mortality estimates.[88]

Hepatocellular cancer. Two reviews of screening with alpha-fetoprotein plus ultrasound were found; Wun *et al.*[89] included two trials (Toronto[90] and Shanghai,[91]) and Aghoram *et al.*[92] included three trials (Toronto,[90] Taiwan[93] and Shanghai[51,91]); neither review gave meta-analytical evidence. The Toronto and Taiwan trials evaluated comparative screening and therefore were not included in our evaluation of individual trials. The two reports from the Shanghai trial had discrepant results; the earlier report[91] did not show benefit with screening [odds ratio (OR) 0.81, 95% confidence interval (CI) = 0.54, 1.22] but the later one[51] showed benefit [relative risk (RR) = 0.63, 95% CI = 0.41, 0.98]. The later one has been included in the analysis as it had longer follow-up data.

Table 1. Screening effect on mortality for major diseases: meta-analytic evidence from Cochrane library and PubMed

Disease	Screening	Disease-specific death			All-cause death			Disease-specific death	All-cause death
		No. RCTs	Screening death/sample	Control death/sample	No. RCTs	Screening death/sample	Control death/sample	Risk estimate (95% CI)	Risk estimate (95% CI)
Abdominal aortic aneurysm	Ultrasound[12,a]	3	221/43211	405/43238	4	19960/57181	20280/57195	0.55 (0.35, 0.86)[f]	0.98 (0.96, 1.00)[f]
Breast cancer	Mammography[5,b]	3	404/119504	572/172649	3	4644/119897	5671/173061	0.90 (0.79, 1.02)[d]	0.99 (0.95, 1.03)[e]
	Mammography[5,c]	4	633/170048	746/136889	5	14355/128749	11839/116119	0.75 (0.67, 0.83)[d]	0.99 (0.97, 1.01)[e]
	Breast self-exam[6]	2	292/190691	295/197844				1.05 (0.90, 1.24)[d]	
Colorectal cancer	Fecal occult blood test (FOBT)[8]	4	1476/172734	1592/156908	4	53666/172734	48202/156908	0.84 (0.78, 0.90)[d]	1.00 (0.99, 1.01)[d]
	Flexible sigmoidoscopy[9]	5	Not given	Not given	3	Not given	Not given	0.71 (0.61, 0.81)[d]	0.99 (0.91, 1.07)[e]
Lung cancer	Chest X-ray[10]	4	710/42668	629/38635	4	4757/54477	9639/115672	1.11 (1.00, 1.23)[d]	1.01 (0.94, 1.08)[e]
	Chest X-ray + cytology[10]	2	256/10194	293/10233	1	493/4968	491/5072	0.88 (0.74, 1.03)[d]	1.03 (0.91, 1.15)[e]
	CT scan[g]	3	53/5704	38/4971	3	158/5704	107/4971	1.23 (0.81, 1.87)[d]	1.34 (0.93, 1.93)[e]
Ovarian cancer	CA-125[11]	2	127/45211	118/45281	1	2924/34253	2914/34304	1.08 (0.84, 1.38)[d]	1.00 (0.96, 1.06)[e]
Prostate cancer	Prostate-specific-antigen[3]	5	698/156157	1318/185185	4	22833/125024	35790/169832	1.00 (0.86, 1.17)[d]	1.00 (0.96, 1.03)[e]

[a]Men.
[b]Optimal trial.
[c]Sub-optimal trial.
[d]Odds ratio.
[e]Relative risk or rate ratio.
[f]Hazard ratio.
[g]We conducted the meta-analyses for screening with CT scan using data from DANTE, DLCST, and MILD trial (see Table 2).

Table 2. Screening effect on mortality for major diseases: individual trial evidence from Cochrane library and PubMed

Disease	Screening test	Individual trial	Disease-specific death		All-cause death		Disease-specific death	All-cause death
			Screening death/sample	Control death/sample	Screening death/sample	Control death/sample	Risk estimate (95% CI)	Risk estimate (95% CI)
Abdominal aortic aneurysm	Ultrasound	Viborg[13]	14/6333	51/6306[f]	2184/6333	2234/6306[f]	0.27 (0.15, 0.49)[d]	0.97 (0.91, 1.03)[d]
		Western Australia[35,a]	18/19352	25/19352[f]	2232/19352	2571/19352[f]	0.72 (0.39, 1.32)[b]	0.98 (includes 1.00)[c]
		MASS[18]	224/33883	381/33887	13858/33883	14134/33887	0.58 (0.49, 0.69)[d]	0.97 (0.95, 0.99)[d]
		Chichester[16]	19/6333	55/6306	2931/6333	2964/6306	0.34 (0.20, 0.57)[d]	0.98 (0.93, 1.03)[d]
Breast cancer	Mammography	Canada 1980 [17,a,f]	105/25214	108/25216[f]	413/25214	413/25216[f]	0.97 (0.74, 1.27)[c]	1.00 (0.87, 1.14)c
		1980[b,68]	107/19711	105/19694[f]	734/19711	690/19694[f]	1.02 (0.78, 1.33)[c]	1.06 (0.96, 1.18)[c]
		Edinburgh[14]	156/28628	167/26026			0.79 (0.60, 1.02)[c]	
		Goteborg[15]	88/21650	162/29961	1430/21000	2241/29200	0.75 (0.58, 0.97)[c]	0.89 (0.83, 0.95)[c]
		Malmo[19]	87/20695	108/20783	2537/21088	2593/21195	0.81 (0.61, 1.07)[c]	0.98 (0.93, 1.04)[c]
		New York[26]	218/31000	262/31000	2062/30239	2116/30765	0.83 (0.66, 1.04)[c]	0.99 (0.94, 1.05)[c]
		Stockholm[20]	66/40318	45/19943	1768/39139	1036/20978	0.73 (0.50, 1.06)[c]	0.91 (0.85, 0.99)[c]
		Two-county[27,f]	351/77080	367/55985	6034/38568[f]		0.73 (0.59, 0.89)[d]	
		Kopparberg[71,f]			4829/38942	2796/18479[f]		1.03 (0.99, 1.08)[c,f]
		Ostergland[71]				4686/37675[f]		1.00 (0.96, 1.04)[c]
		UK Age trial[24]	105/53884	251/106956	960/53884	1975/106956	0.83 (0.66, 1.04)[c,f]	0.96 (0.89, 1.04)[c]
	Clinical breast exam	Mumbai[7]	22/75360	10/76178			$P = 0.304$[e]	
	Breast self-exam	Russia[6,58]	157/57712	164/64759			1.07 (0.86, 1.34)[c]	
		Shanghai[59]	135/132979	131/133085			1.03 (0.81, 1.31)[c]	
Cervical cancer	Visual inspection	Tamil Nadu, India[46]	83/31343	92/30958	1303/31343	977/30958	0.65 (0.47, 0.89)[d]	0.87 (0.78, 0.96)[d]
	Visual inspection	Maharashtra India[47,f]	56/34074[f]				0.86 (0.60, 1.25)[d,f]	
	Human papilloma virus		34/34126[f]	64/31488			0.52 (0.33, 0.83)[d,f]	
	Cytology		54/32058				0.89 (0.62, 1.27)[d]	
Cardiovascular disease	Echocardiography	Norwa [43]	250/3272	299/3589	880/3272	989/3589	0.91 (0.77, 1.08)[d]	0.97 (0.89, 1.06)[d]
Colorectal cancer	Fecal occult blood test (FOBT)	Funen[21]	362/30967	431/30966	12205/30967	12248/30966	0.84 (0.73, 0.96)[b]	0.99 (0.96, 1.03)[b]
		Goteborg[23]	252/34144	300/34164	10591/34144	10432/34164	0.84 (0.71, 0.99)[b]	1.02 (0.99, 1.06)[b]
		Minnesota – annual[49]	200/15570	295/15394	11072/15570	10944/15394	0.68 (0.56,0.82)[c]	1.00 (0.99, 1.01)[c]
		Minnesota – biennial[49]	237/15587	295/15394	11004/15587	10944/15394	0.78 (0.65, 0.93)[c]	0.99 (0.98, 1.01)[c]
		Nottingham[25]	593/76466	684/76384	20421/76466	20336/76384	0.87 (0.77, 0.97)[b]	1.00 (0.98, 1.03)[b]
	Flexible sigmoidoscopy with or without FOBT	Norwegian NCCPS[31]	24/13653	99/41092	Not given	Not given	0.73 (0.47, 1.13)[d,f]	1.02 (0.98, 1.07)[d]
	Single flexible sigmoidoscopy	Telemark[54]	1/400	4/399	62/400	40/399	0.25 (0.03, 2.23)[c]	1.55(1.04, 2.30)[c]
		UK trial[29]	189/57099	538/112939	6775/57099	13768/112939	0.69 (0.59, 0.82)[d]	0.97 (0.94, 1.00)[d]
		SCORE[40]	65/17136	83/17136	1202/17136	1233/17136	0.78 (0.56, 1.08)[c]	0.97 (0.90, 1.05)[c]
	Multiple flexible sigmoidoscopy	PLCO[38]	252/77445	341/77455			0.74 (0.63, 0.87)[c]	

(Continued)

Table . Continued

Disease	Screening test	Individual trial	Disease-specific death		All-cause death		Disease-specific death	All-cause death
			Screening death/sample	Control death/sample	Screening death/sample	Control death/sample	Risk estimate (95% CI)	Risk estimate (95% CI)
Type 2 diabetes	Fasting blood glucose + HbA1c	ADDITION-Cambridge[50]	Not given	Not given	1532/16047	377/4137	1.26(0.75, 2.10)[d]	1.06 (0.90,1.25)[d]
Hepatocellular carcinoma	Alphafetoprotein + ultrasound	Shanghai[51]	32/9373	54/9443			0.63 (0.41, 0.98)[c]	
	Alphafetoprotein	Qidong, China[41]	218/3712	109/1869	353/3712	175/1869	1.01 (0.81, 1.26)[c]	1.01 (0.85, 1.21)[c]
Lung cancer	Chest X-ray	Czech[22]	247/3171	216/3174	341/3171	292/3174	1.14 (0.96, 1.36)[c]	1.16 (1.00, 1.35)[c]
		Kaiser[53]	44/5156	42/5557	585/5156	643/5557	1.13 (0.74, 1.72)[c]	0.98 (0.88, 1.09)[c]
		Mayo Lung[57]	337/4618	303/4593	688/4618	665/4593	1.11 (0.95, 1.28)[c]	1.03 (0.93, 1.14)[c]
		North London[52]	82/29723	68/25311			1.03 (0.74, 1.42)[c]	
		John Hopkins[56,f]	141/5226	173/5161			0.80 (0.65, 1.00)[c]	
		Memorial Sloan-Kettering[34,f]	115/4968	120/5072	493/4968	491/5072	0.98 (0.76, 1.26)[c]	1.03 (0.91, 1.15)[c]
		PLCO[36,f]	1213/77445	1230/77456			0.99 (0.87, 1.22)[c]	
	CT scan	DANTE[42,f]	20/1276	20/1196	46/1276	45/1196	0.94 (0.50, 1.73)[c]	0.96 (0.64, 1.43)[c]
		DLCST[45]	15/2052	11/2052	61/2052	42/2052	1.36 (0.63, 2.96)[c]	1.45 (0.98, 2.14)[c]
		MILD[44]	18/2376	7/1723	51/2376	20/1723	1.64 (0.67, 4.01)[d]	1.40 (0.82, 2.38)[d]
Oral cancer	Visual exam	Kerala, India -overall[48]	138/96517	154/95356			0.88 (0.69, 1.12)[c]	
		Kerala, India –high risk[48,f]	129/45791	147/39151			0.76 (0.60, 0.97)[c]	
		Kerala, India –low risk[48,g]	9/50726	7/56205			1.36 (0.57, 3.26)[c]	
Ovarian cancer	CA – 125	PLCO[30]	118/34253	100/34304	2924/34253	2914/34304	1.18 (0.82, 1.71)[c]	1.01 (0.96, 1.06)[c]
		UK[32]	9/10958	18/10977			0.50 (0.22, 1.11)[c]	
Prostate cancer	Prostate-specific-antigen	ERSPC[103]	261/82816	363/99184	13082/82816	15717/99184	0.86 (0.73, 1.01)[c]	1.00 (0.98, 1.02)[c]
		Norkopping[37]	30/1494	130/7532			1.16 (0.78, 1.73)[c]	
		PLCO[28]	158/38340	145/38345			1.09 (0.87, 1.36)[c]	
		Quebec[33]	153/31133	75/15353			1.01 (0.76, 1.33)[c]	
		Stockholm[55]	53/2374	506/24772	986/2374	10328/24772	1.09 (0.83, 1.45)[c]	1.00 (0.95, 1.05)[c]

[a]The estimate for all-cause mortality in the Western Australia study is based on the ratio of the age standardized mortality data provided in Table 5 of Norman *et al.*[35] Although Lindholt *et al.*[62] show a decrease in all-cause mortality with longer follow-up (odds ratio = 0.91 (0.88, 0.95), this seems to be an artefact of age imbalance in the two arms; and Takagi *et al.*[12] show no difference between groups in the 65 to 74 age stratum, which is well age-balanced (hazard ratio = 0.99 (0.96, 1.02).

[b]Odds ratio.

[c]Relative risk or rate ratio.

[d]Hazard ratio.

[e]Not specified.

[f]High-risk participants (smokers and alcohol users).

[g]Low-risk participants.

One additional single trial (Qidong, China[41]) of screening with only alpha-fetoprotein was also found.

Lung cancer. Three meta-analyses of screening with chest X-ray were found; Manser *et al.*[10] presented seven trials of screening with chest X-ray (Czech,[22] Erfurt County,[94] Kaiser,[53] Mayo lung,[57] North London,[52] Johns Hopkins[56] and Memorial Sloan-Kettering[34]) and has been selected for analysis. However, the Erfurt County study was excluded from the individual trial evidence (Table 2) because of its non-randomized design. The other two meta-analyses on chest X-ray were excluded because they were earlier publications and contained non-randomized data.[95,96] Data from PLCO[36] (chest X-ray) was not included in any of the meta-analyses but we presented its estimates in the individual trial evidence. There were four trials (DLCST,[45] MILD,[44] DANTE[42] and NLST[97]) on computer tomography (CT) scan, but no available meta-analyses. We excluded NLST as it evaluated comparative screening (CT scan vs chest X-ray). We recorded the estimates from the other three trials and conducted our own meta-analysis.

Oral cancer. One review[98] of screening with visual examination was found. It contained only one trial from Kerala, India.[48,99] The estimate with longer follow-up data from that trial was presented in the individual trial evidence.[48]

Ovarian cancer. One meta-analysis on screening with CA-125 was found;[11] it contained two individual trials (PLCO,[30] UK[32]).

Prostate cancer. Six meta-analyses of screening with prostate specific antigen (PSA) were found. Four were by Ilic *et al.*[3,100–102] and contained the same data from five trials (ERSPC,[103] Norrkoping,[104] PLCO,[105] Quebec[33] and Stockholm[55]); we used the estimates from the most recent publication.[3] The other two were not used because they included site-specific data (e.g. French ERSPC and Gothenburg are part of original ERSPC) or used non-randomized data.[106,107] Of the individual trials, we also included updated estimates for Norrkoping[37] and PLCO.[28]

Cardiovascular disease. One individual trial of screening with echocardiography was found.[43]

Type 2 diabetes mellitus. One individual trial of screening with fasting blood glucose and haemoglobin A1c (HbA1c) was found.[50]

Synopsis of RCT evidence (meta-analytic and individual) for mortality

Meta-analytic evidence

As shown in Table 1, meta-analyses of randomized trials were available for nine screening tests on six diseases. The 95% CIs excluded the null in 4 out of 11 available estimates (36%) of disease-specific mortality, but in none out of 10 available estimates for all-cause mortality. Disease-specific mortality was reduced with ultrasound for abdominal aortic aneurysm in men;[12] mammography for breast cancer;[5] and fecal occult blood test[8] and flexible sigmoidoscopy[9] for colorectal cancer. The range of relative risk reduction in these four cases was between 16% and 45%. Relative risk estimates for all-cause mortality were all very close to 1.00 (range 0.98–1.03).

Individual trial evidence

As shown in Table 2, we compiled evidence from 48 randomized trials on 19 screening tests for 11 diseases. The 95% CIs excluded the null in 16 out of 54 reported estimates (30%) (some trials reported more than one estimate, e.g. in different subgroups) for disease-specific mortality and for 4 out of 36 reported estimates (11%) for all-cause mortality. The range of relative risk reduction in the 16 cases with improved disease-specific mortality was between 13% and 73% (median 29%) and in the four cases of improved all-cause mortality it was between 3% and 13% (median 10%).

Disease-specific mortality was reduced with ultrasound for abdominal aortic aneurysm in the Viborg,[13] MASS[60] and Chichester[16] trials; with mammography for breast cancer in the Goteborg[15] and Two-county[27] trials; with visual inspection for cervical cancer in the Tamil Nadu[46] and Maharashtra[47] trials; with FOBT for colorectal cancer in the Funen,[21] Goteborg,[23] Minnesota[49] and Nottingham[25] trials; with flexible sigmoidoscopy for colorectal cancer in the UK trial[29] and PLCO;[38] with alpha-fetoprotein and ultrasound for hepatocellular cancer in the Shanghai[51] trial; and with visual examination for oral cancer in the Kerala[99] trial. Overall, seven tests for six diseases had at least one RCT with a disease-specific mortality benefit: of those, three diseases had also been documented in meta-analyses.

All-cause mortality was reduced with ultrasound in abdominal aortic aneurysm in MASS;[18] with mammography in breast cancer in Goteborg[15] and Stockholm;[20] and with visual examination for cervical cancer in Tamil Nadu.[46] Mammography and ultrasound for aortic aneurysm had no all-cause mortality benefits in the respective meta-analyses including all the relevant trials. Visual examination for cervical cancer had also been assessed in

another trial that did not report results on all-cause mortality.[47]

Discussion

Our comprehensive overview shows that there are currently at least 48 RCTs and 9 non-overlapping meta-analyses that have evaluated the impact of any screening test vs no screening on mortality in asymptomatic adults for diseases where mortality is a common outcome. Documented reductions in disease-specific mortality in randomized trials of screening are uncommon. Reduction in all-cause mortality is even more uncommon in single trials and has not been documented in the latest available meta-analysis of multiple trials for any of the examined topics. This overview offers to researchers, policy makers and healthcare providers a synthesis of RCT evidence on the potential benefits of screening on mortality, and is timely in the wake of recent controversies around breast and prostate cancer screening.

Of the handful of trials that have reported survival benefits from screening, it is likely that in a few of them the benefit is substantially overestimated. For example, visual inspection of the cervix with acetic acid (cervical cancer screening) offered a 13% estimated relative risk reduction for all-cause death in one trial[46] conducted in rural India. Women in the screened group received other interventions apart from screening, such as correction of anaemia and measurement of blood pressure. Hence this large difference in total mortality, if true, was likely the result of multiple interventions and not the screening alone (cervical cancer does not account for 13% of all deaths even in rural India). Similarly, a mortality reduction shown in the Shanghai trial for screening in hepatocellular cancer[51] is in question. The earlier paper[91] from that trial did not report a risk estimate but only reported percent survival; a subsequent Cochrane review[89] used the survival data to calculate a risk estimate with 95% CIs that did not exclude the null. In the same way, the original publication[35] from the Western Australia study for screening in abdominal aortic aneurysm did not report a relative risk estimate for all-cause mortality; a subsequent meta-analysis[62] calculated a mortality reduction with 95% CIs that excluded the null, but this did not take into account the substantial age imbalance that existed between the study groups; and another more recent meta-analysis[12] that realized this caveat had 95% CIs that did not exclude the null.

There are many potential underlying reasons for the overall poor performance of screening in reducing mortality: the screening test may lack sufficient sensitivity and specificity to capture the disease early in its process; there are no markedly effective treatment options for the disease; treatments are available but the risk-benefit ratio of the whole screening and treatment process is unfavourable; or competing causes of death do not allow us to see a net benefit. Often, these reasons may coexist. Whether screening saves lives can only be reliably proven with RCTs.[108] However, even for newly proposed tests, we suspect that their adoption in practice may evade RCT testing. A very large number of tests continuously become available due to technological advancement.[109] One may be tempted to claim a survival benefit of screening based on observational cohorts showing improved survival rates,[110] but these are prone to lead-time and other types of bias. Even RCTs can be biased sometimes, as has been discussed and hotly debated in the controversy over mammography.[71]

Some limitations should be acknowledged in our overview. First, we synthesized randomized evidence, but did not include data from other research designs, such as cohort and case-control studies. However, as we stated above, non-randomized studies have serious limitations. Non-randomized studies may provide useful suggestions and insights, but typically these would be less definitive, unless the effect is very robust and large, and most screening tests do not seem to have large effects on mortality. Second, one should acknowledge that given the many competing causes of death, it is very difficult to document reductions in all-cause mortality, unless the disease of interest is a leading cause of death and extremely large RCTs are performed. This is the reason why we also addressed comprehensively all the available data on disease-specific mortality. Third, we used broad search terms in PubMed and in Cochrane to maximize the capture of relevant trials and meta-analyses. It is possible that a few trials may have been missed, but it is unlikely that we have missed major trials that had found mortality benefits. As a quality check, we also matched our search results with the USPSTF documents. We found that we had not missed any trials that USPSTF has cited, whereas we have detected several additional recent trials that USPSTF did not cite (not unexpected since the USPSTF updates the evidence periodically). Finally, we did not include evidence on the effectiveness of one screening test against another (i.e. comparative screening). Nevertheless, it is difficult to interpret a trial that shows that a screening test is better than an older comparator, when it is unknown whether the older comparator does more good than harm.

To avoid uncertainty and a continuing conundrum in the world of screening for disease, we need to choose the appropriate study design and outcome, depending on the disease, to evaluate the effectiveness of screening tests. We argue that for diseases where short- and medium-term mortality are a relatively common outcomes, RCT should be the default evaluation tool and disease-specific and

all-cause mortality should be routinely considered as main outcomes. Our overview suggests that even then, all-cause mortality may hardly ever be improved. One may argue that a reduction in disease-specific mortality may sometimes be beneficial even in the absence of a reduction in all-cause mortality. Such an inference would have to consider the relative perception of different types of death by patients (e.g. death by cancer vs death by other cause), and it may entail also some subjectivity. For diseases where mortality outcomes are potentially important but only in the very long term, one has to consider whether the use of other, intermediate outcomes and/or other quasi-experimental designs that may be performed relatively quickly with very large sample sizes (e.g. before and after the introduction of a test) are meaningful alternatives to very long-term RCTs or may add more bias and confusion in a field that has already seen many hot debates. Screening may still be highly effective (and thus justifiable) for a variety of other clinical outcomes, besides mortality. However, our overview suggests that expectations of major benefits in mortality from screening need to be cautiously tempered.

Supplementary Data

Supplementary data are available at *IJE* online.

Conflict of interest: None declared.

References

1. Force USPST. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Int Med* 2009;151:716–26, W–236.
2. Hudson B, Zarifeh A, Young L, Wells JE. Patients' expectations of screening and preventive treatments. *Ann Fam Med* 2012;10:495–502.
3. Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. *Cochrane Database Syst Rev* 2013;1:CD004720.
4. Nelson HD, Huffman LH, Fu R, Harris EL, Force USPST. Genetic risk assessment and BRCA mutation testing for breast and ovarian cancer susceptibility: systematic evidence review for the U.S. Preventive Services Task Force. *Ann Int Med* 2005;143:362–79.
5. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2013;6: CD001877.
6. Kosters JP, Gotzsche PC. Regular self-examination or clinical examination for early detection of breast cancer. *Cochrane dDtabase Syst Rev* 2003;2:CD003373.
7. Mittra I, Mishra GA, Singh S *et al*. A cluster randomized, controlled trial of breast and cervix cancer screening in Mumbai, India: methodology and interim results after three rounds of screening. *Int J Cancer* 2010;126:976–84.
8. Hewitson P, Glasziou P, Watson E, Towler B, Irwig L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. *Am J Gastroenterol* 2008;103:1541–49.
9. Littlejohn C, Hilton S, Macfarlane GJ, Phull P. Systematic review and meta-analysis of the evidence for flexible sigmoidoscopy as a screening method for the prevention of colorectal cancer. *Br J Surg* 2012;99:1488–500.
10. Manser R, Lethaby A, Irving LB *et al*. Screening for lung cancer. *Cochrane Database Syst Rev* 2013;6:CD001991.
11. Reade CJ, Riva JJ, Busse JW, Goldsmith CH, Elit L. Risks and benefits of screening asymptomatic women for ovarian cancer: a systematic review and meta-analysis. *Gynecol Oncol* 2013;130:674–81.
12. Takagi H, Goto SN, Matsui M, Manabe H, Umemoto T. A further meta-analysis of population-based screening for abdominal aortic aneurysm. *J Vasc Surg* 2010;52:1103–08.
13. Lindholt JS, Juul S, Fasting H, Henneberg EW. Preliminary ten year results from a randomised single centre mass screening trial for abdominal aortic aneurysm. *Eur J Vasc Endovasc Surg* 2006;32:608–14.
14. Alexander FE, Anderson TJ, Brown HK *et al*. 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet* 1999;353:1903–08.
15. Bjurstam N, Bjorneld L, Warwick J *et al*. The Gothenburg Breast Screening Trial. *Cancer* 2003;97:2387–96.
16. Lindholt JS, Sorensen J, Sogaard R, Henneberg EW. Long-term benefit and cost-effectiveness analysis of screening for abdominal aortic aneurysms from a randomized controlled trial. *Br J Surg* 2010;97:826–34.
17. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Int Med* 2002;137 (**5 Part 1**):305–12.
18. Thompson SG, Ashton HA, Gao L, Buxton MJ, Scott RA, Multicentre Aneurysm Screening Study G. Final follow-up of the Multicentre Aneurysm Screening Study (MASS) randomized trial of abdominal aortic aneurysm screening. *Br J Surg* 2012;99:1649–56.
19. Andersson I, Janzon L. Reduced breast cancer mortality in women under age 50: updated results from the Malmo Mammographic Screening Program. *J Natl Cancer Inst Monogr.* 1997;22:63–67.
20. Frisell J, Lidbrink E, Hellstrom L, Rutqvist LE. Followup after 11 years – update of mortality results in the Stockholm mammographic screening trial. *Breast Cancer Res Treat* 1997;45: 263–70.
21. Kronborg O, Jorgensen OD, Fenger C, Rasmussen M. Randomized study of biennial screening with a faecal occult blood test: results after nine screening rounds. *Scand J Gastroenterol* 2004;39:846–51.
22. Kubik AK, Parkin DM, Zatloukal P. Czech Study on Lung Cancer Screening: post-trial follow-up of lung cancer deaths up to year 15 since enrollment. *Cancer* 2000;89(**Suppl 11**):2363–68.
23. Lindholm E, Brevinge H, Haglind E. Survival benefit in a randomized clinical trial of faecal occult blood screening for colorectal cancer. *Br J Surg* 2008;95:1029–36.
24. Moss SM, Cuckle H, Evans A *et al*. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *Lancet* 2006;368:2053–60.

25. Scholefield JH, Moss S, Sufi F, Mangham CM, Hardcastle JD. Effect of faecal occult blood screening on mortality from colorectal cancer: results from a randomised controlled trial. *Gut* 2002;50:840–44.
26. Shapiro S, Venet W, Strax P, Venet L, Roeser R. Ten- to fourteen-year effect of screening on breast cancer mortality. *J Natl Cancer Inst* 1982;69:349–55.
27. Tabar L, Vitak B, Chen TH *et al*. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology* 2011;260:658–63.
28. Andriole GL, Crawford ED, Grubb RL III *et al*. Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up. *J Natl Cancer Inst* 2012;104:125–32.
29. Atkin WS, Edwards R, Kralj-Hans I *et al*. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet* 2010;375: 1624–33.
30. Buys SS, Partridge E, Black A *et al*. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA* 2011;305:2295–303.
31. Hoff G, Grotmol T, Skovlund E, Bretthauer M, Norwegian Colorectal Cancer Prevention Study G. Risk of colorectal cancer seven years after flexible sigmoidoscopy screening: randomised controlled trial. *BMJ* 2009;338:b1846.
32. Jacobs IJ, Skates SJ, MacDonald N *et al*. Screening for ovarian cancer: a pilot randomised controlled trial. *Lancet* 1999;353: 1207–10.
33. Labrie F, Candas B, Cusan L *et al*. Screening decreases prostate cancer mortality: 11-year follow-up of the 1988 Quebec prospective randomized controlled trial. *Prostate* 2004;59: 311–18.
34. Melamed MR, Flehinger BJ, Zaman MB, Heelan RT, Perchick WA, Martini N. Screening for early lung cancer. Results of the Memorial Sloan-Kettering study in New York. *Chest* 1984;86: 44–53.
35. Norman PE, Jamrozik K, Lawrence-Brown MM *et al*. Population based randomised controlled trial on impact of screening on mortality from abdominal aortic aneurysm. *BMJ* 2004;329:1259.
36. Oken MM, Hocking WG, Kvale PA *et al*. Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. *JAMA* 2011;306:1865–73.
37. Sandblom G, Varenhorst E, Rosell J, Lofman O, Carlsson P. Randomised prostate cancer screening trial: 20 year follow-up. *BMJ* 2011;342:d1539.
38. Schoen RE, Pinsky PF, Weissfeld JL *et al*. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. *N Engl J Med* 2012;366:2345–57.
39. Schroder FH, Hugosson J, Roobol MJ *et al*. Prostate-cancer mortality at 11 years of follow-up. *N Engl J med* 2012;366: 981–90.
40. Segnan N, Armaroli P, Bonelli L *et al*. Once-only sigmoidoscopy in colorectal cancer screening: follow-up findings of the Italian Randomized Controlled Trial – SCORE. *J Natl Cancer Inst* 2011;103:1310–22.
41. Chen JG, Parkin DM, Chen QG *et al*. Screening for liver cancer: results of a randomised controlled trial in Qidong, China. *J Med Screen*. 2003;10:204–09.
42. Infante M, Cavuto S, Lutman FR *et al*. A randomized study of lung cancer screening with spiral computed tomography: three-year results from the DANTE trial. *Am J Resp Crit Care Med* 2009;180:445–53.
43. Lindekleiv H, Lochen ML, Mathiesen EB, Njolstad I, Wilsgaard T, Schirmer H. Echocardiographic screening of the general population and long-term survival: a randomized clinical study. *JAMA Int Med* 2013;173:1592–98.
44. Pastorino U, Rossi M, Rosato V *et al*. Annual or biennial CT screening versus observation in heavy smokers: 5-year results of the MILD trial. *Eur J Cancer Prev* 2012;21:308–15.
45. Saghir Z, Dirksen A, Ashraf H *et al*. CT screening for lung cancer brings forward early disease. The randomised Danish Lung Cancer Screening Trial: status after five annual screening rounds with low-dose CT. *Thorax* 2012;67:296–301.
46. Sankaranarayanan R, Esmy PO, Rajkumar R *et al*. Effect of visual screening on cervical cancer incidence and mortality in Tamil Nadu, India: a cluster-randomised trial. *Lancet* 2007; 370:398–406.
47. Sankaranarayanan R, Nene BM, Shastri SS *et al*. HPV screening for cervical cancer in rural India. *N Engl J Med* 2009;360: 1385–94.
48. Sankaranarayanan R, Ramadas K, Thara S *et al*. Long term effect of visual screening on oral cancer incidence and mortality in a randomized trial in Kerala, India. *Oral Oncol* 2013;49:314–21.
49. Shaukat A, Mongin SJ, Geisser MS *et al*. Long-term mortality after screening for colorectal cancer. *N Engl J Med* 2013;369: 1106–14.
50. Simmons RK, Echouffo-Tcheugui JB, Sharp SJ *et al*. Screening for type 2 diabetes and population mortality over 10 years (ADDITION-Cambridge): a cluster-randomised controlled trial. *Lancet* 2012;380:1741–48.
51. Zhang BH, Yang BH, Tang ZY. Randomized controlled trial of screening for hepatocellular carcinoma. *J Cancer Res Clin Oncol* 2004;130:417–22.
52. Brett GZ. Earlier diagnosis and survival in lung cancer. *Br Med J* 1969;4:260–62.
53. Friedman GD, Collen MF, Fireman BH. Multiphasic Health Checkup Evaluation: a 16-year follow-up. *J Chron Dis* 1986; 39:453–63.
54. Hoff G, Thiis-Evensen E, Grotmol T, Sauar J, Vatn MH, Moen IE. Do undesirable effects of screening affect all-cause mortality in flexible sigmoidoscopy programmes? Experience from the Telemark Polyp Study 1983-1996. *Eur J Cancer Prev* 2001; 10:131–37.
55. Kjellman A, Akre O, Norming U, Tornblom M, Gustafsson O. 15-year followup of a population based prostate cancer screening study. *J Urol* 2009;181:1615–21; discussion 1621.
56. Levin ML, Tockman MS, Frost JK, Ball WC Jr. Lung cancer mortality in males screened by chest X-ray and cytologic sputum examination: a preliminary report. *Recent results Cancer Res* 1982;82:138–46.

57. Marcus PM, Bergstralh EJ, Fagerstrom RM *et al*. Lung cancer mortality in the Mayo Lung Project: impact of extended follow-up. *J Natl Cancer Inst* 2000;92:1308–16.
58. Semiglazov VF, Moiseenko VM, Manikhas AG *et al*. [Interim results of a prospective randomized study of self-examination for early detection of breast cancer (Russia/St Petersburg/WHO). *Voprosy onkologii* 1999;45:265–71.
59. Thomas DB, Gao DL, Ray RM *et al*. Randomized trial of breast self-examination in Shanghai: final results. *J Natl Cancer Inst* 2002;94:1445–57.
60. Thompson SG, Ashton HA, Gao L, Scott RA, Multicentre Aneurysm Screening Study G. Screening men for abdominal aortic aneurysm: 10 year mortality and cost effectiveness results from the randomised Multicentre Aneurysm Screening Study. *BMJ* 2009;338:b2307.
61. Eckstein HH, Bockler D, Flessenkamper I, Schmitz-Rixen T, Debus S, Lang W. [Ultrasonographic screening for the detection of abdominal aortic aneurysms.] *Deutsches Arzteblatt international* 106:657–63.
62. Lindholt JS, Norman P. Screening for abdominal aortic aneurysm reduces overall mortality in men. A meta-analysis of the mid- and long-term effects of screening for abdominal aortic aneurysms. *Eur J vasc Endovasc Surg* 2008;36:167–71.
63. Takagi H, Kawai N, Umemoto T. Abdominal aortic aneurysm: Screening reduces all cause mortality in men. *BMJ* 2007; 335:899.
64. Cosford PA, Leng GC. Screening for abdominal aortic aneurysm. *Cochrane Database Syst Rev* 2007;2:CD002945.
65. Takagi H, Tanabashi T, Kawai N, Kato T, Umemoto T. Abdominal aortic aneurysm screening reduces mortality: meta-analyses of randomized, controlled trials. *Eur J Vasc Endovasc Surg* 2007;33:132–33.
66. Fleming C, Whitlock EP, Beil TL, Lederle FA. Screening for abdominal aortic aneurysm: a best-evidence systematic review for the U.S. Preventive Services Task Force. *Ann Int Med* 2005; 142:203–11.
67. Lindholt JS, Norman PE. Meta-analysis of postoperative mortality after elective repair of abdominal aortic aneurysms detected by screening. *Br J Surg* 2011;98:619–22.
68. Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years. *J Natl Cancer Inst* 2000;92:1490–99.
69. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2009;4:CD001877.
70. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2006;4:CD001877.
71. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev.* 2011;1: CD001877.
72. Shen Y, Yang Y, Inoue LY, Munsell MF, Miller AB, Berry DA. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. *J Natl Cancer Inst* 2005;97:1195–203.
73. Hackshaw A. EUSOMA review of mammography screening. *Ann Oncol* 2003;14:1193–95.
74. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann INt Med* 2002;137(**5 Part 1**):347–60.
75. Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;355:129–34.
76. Glasziou PP. Meta-analysis adjusting for compliance: the example of screening for breast cancer. *J Clin Epidemiol* 1992;45:1251–56.
77. Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of screening mammography. A meta-analysis. *JAMA* 1995;273:149–54.
78. Magnus MC, Ping M, Shen MM, Bourgeois J, Magnus JH. Effectiveness of mammography screening in reducing breast cancer mortality in women aged 39-49 years: a meta-analysis. *J Women's Health* 2011;20:845–52.
79. Bastardis-Zakas K, Iatrakis G, Navrozoglou I *et al*. Maximizing the benefits of screening mammography for women 40-49 years old. *Clin Exp Obstet Gynecol* 2010;37:278–82.
80. Hendrick RE, Smith RA, Rutledge JH III, Smart CR. Benefit of screening mammography in women aged 40-49: a new meta-analysis of randomized controlled trials. *J Natl Cancer Inst Monogr* 1997;22:87–92.
81. Smart CR, Hendrick RE, Rutledge JH III, Smith RA. Benefit of mammography screening in women ages 40 to 49 years. Current evidence from randomized controlled trials. *Cancer* 1995;75:1619–26.
82. Elwood JM, Cox B, Richardson AK. The effectiveness of breast cancer screening by mammography in younger women. *Online J Curr Clin Trials* 1993;Doc No 32.
83. Autier P, Hery C, Haukka J, Boniol M, Byrnes G. Advanced breast cancer and breast cancer mortality in randomized controlled trials on mammography screening. *J Clin Oncol* 2009;27:5919–23.
84. Mandel JS, Church TR, Ederer F, Bond JH. Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. *J Natl Cancer Inst* 1999;91:43437.
85. Towler B, Irwig L, Glasziou P, Haas M, Plunkett A, Salkeld G. The potential benefits and harms of screening for colorectal cancer. *Aust J Public health* 1995;19:24–28.
86. Towler B, Irwig L, Glasziou P, Kewenter J, Weller D, Silagy C. A systematic review of the effects of screening for colorectal cancer using the faecal occult blood test, hemoccult. *BMJ* 1998;317:559–65.
87. Heresbach D, Manfredi S, D'Halluin PN, Bretagne JF, Branger B. Review in depth and meta-analysis of controlled trials on colorectal cancer screening by faecal occult blood test. *Eur J Gastroenterol Hepatol* 2006;18:427–33.
88. Elmunzer BJ, Hayward RA, Schoenfeld PS, Saini SD, Deshpande A, Waljee AK. Effect of flexible sigmoidoscopy-based screening on incidence and mortality of colorectal cancer: a systematic review and meta-analysis of randomized controlled trials. *PLoS Med* 012;**9**:e1001352.
89. Wun YT, Dickinson JA. Alpha-fetoprotein and/or liver ultrasonography for liver cancer screening in patients with chronic hepatitis B. *Cochrane Database Syst Rev* 2003;2:CD002799.
90. Sherman M, Peltekian KM, Lee C. Screening for hepatocellular carcinoma in chronic carriers of hepatitis B virus: incidence and prevalence of hepatocellular carcinoma in a North American urban population. *Hepatology* 1995;22:432–38.
91. Yang B, Zhang B, Xu Y *et al*. Prospective study of early detection for primary liver cancer. *J Cancer Res Clin Oncol* 1997;123:357–60.

92. Aghoram R, Cai P, Dickinson JA. Alpha-foetoprotein and/or liver ultrasonography for screening of hepatocellular carcinoma in patients with chronic hepatitis B. *Cochrane Database Syst Rev* 2012;9:CD002799.
93. Wang JH, Chang KC, Hung CH, Chen CH, Lu SN. Hepatocellular carcinoma surveillance with 4 versus 12 months interval for patients with chronic viral hepatitis – a randomized community study. *Hepatology* 2011;54(**Suppl 1**):1378A.
94. Wilde J. A 10 year follow-up of semi-annual screening for early detection of lung cancer in the Erfurt County, GDR. *Eur Resp J* 1989;2:656–62.
95. Manser RL, Irving LB, Byrnes G, Abramson MJ, Stone CA, Campbell DA. Screening for lung cancer: a systematic review and meta-analysis of controlled trials. *Thorax* 2003;58:784–89.
96. Manser RL, Irving LB, Stone C, Byrnes G, Abramson M, Campbell D. Screening for lung cancer. *Cochrane Database Syst Rev* 2004;1:CD001991.
97. National Lung Screening Trial Research T, Aberle DR, Adams AM *et al*. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;365:395–409.
98. Brocklehurst P, Kujan O, Glenny AM *et al*. Screening programmes for the early detection and prevention of oral cancer. *Cochrane Database Syst Rev* 2010;11:CD004150.
99. Sankaranarayanan R, Ramadas K, Thomas G *et al*. Effect of screening on oral cancer mortality in Kerala, India: a cluster-randomised controlled trial. *Lancet* 2005;365:1927–33.
100. Ilic D, O'Connor D, Green S, Wilt TJ. Screening for prostate cancer: an updated Cochrane systematic review. *BJU Int* 2011;107:882–91.
101. Ilic D, O'Connor D, Green S, Wilt T. Screening for prostate cancer: a Cochrane systematic review. *Cancer Causes Control* 2007;18:279–85.
102. Ilic D, O'Connor D, Green S, Wilt T. Screening for prostate cancer. *Cochrane Database Syst Rev* 2006;3:CD004720.
103. Schroder FH, Hugosson J, Roobol MJ *et al*. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* 2009;360:1320–28.
104. Sandblom G, Varenhorst E, Lofman O, Rosell J, Carlsson P. Clinical consequences of screening for prostate cancer: 15 years follow-up of a randomised controlled trial in Sweden. *Eur Urol* 2004;46:717–23; discussion 724.
105. Andriole GL, Crawford ED, Grubb RL III *et al*. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med* 2009;360:1310–19.
106. Lumen N, Fonteyne V, De Meerleert G *et al*. Population screening for prostate cancer: an overview of available studies and meta-analysis. *Int J Urol* 2012;19:100–08.
107. Djulbegovic M, Beyth RJ, Neuberger MM *et al*. Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2010;341:c4543.
108. Measurement of progress against cancer. Extramural Committee to Assess Measures of Progress Against Cancer. *J Natl Cancer Inst* 1990;82:825–35.
109. Tarini BA. The current revolution in newborn screening: new technology, old controversies. *Arch Ped Adolesc Med* 2007;161:767–72.
110. Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? *JAMA* 2000;283:2975–78.

International Journal of Epidemiology, 2015, 277–278
doi: 10.1093/ije/dyu268
Advance Access Publication Date: 15 January 2015

Commentary on N Saquib et al. Does screening for disease save lives in asymptomatic adults? Systematic review of 5 meta-analyses and randomized trials

Paul G Shekelle

RAND Corporation, 1776 Main Street, Santa Monica, CA 90407, USA. E-mail: shekelle@rand.org

In this issue, Drs Saquib, Saquib and Ioannidis perform a valuable service by reviewing the evidence that screening for various diseases save lives. The authors examined the randomized controlled trials (RCTs) and meta-analyses of trials of the various screening strategies, and then assessed the outcomes of disease-specific mortality and all-cause mortality. They found that evidence of an effect on disease-specific mortality was relatively uncommon, and that evidence of an effect on all-cause mortality was essentially non-existent. The authors conclude that the effects of screening on mortality are likely to be modest at best, and that future evaluations of screening tests for diseases where short- and medium-term mortality are common, RCTs should be the default evaluation tool and disease-specific and all-cause mortality should be the main outcomes.

This raises the larger question of what should be the evidence upon which to base decisions about the appropriateness of screening tests, which by definition are

administered to large masses of the general population who do not exhibit any signs or symptoms of the disease. The purpose of screening is to detect in a preclinical phase the presence of a disease or precursor to a disease whose subsequent clinical course can then be ameliorated or even eliminated with treatment at that stage, in comparison with beginning treatment when the patient develops signs or symptoms of the disease (I am intentionally ignoring the issue of screening for genetic diseases, where a better understanding of prognosis in the absence of effective treatment can be considered an important outcome, as well as the prevention of a genetic disease in any offspring).

What needs to be considered then, are the outcomes of the disease for which screening is being contemplated. Although death is of course an important outcome, it is not the only outcome, and for some diseases may not even be the most important outcome. Many chronic diseases, such as heart failure, diabetes and chronic obstructive pulmonary disease (all three included in the set of diseases studied in this analysis) have numerous symptoms and outcomes other than mortality, such as dyspnoea, blindness, kidney failure and amputation. Even in the absence of any effect on mortality it is easy for me, as a primary care clinician, to imagine that patients would highly value any screening test and intervention that decreased the risk or severity of these outcomes. So I do not agree with the authors that, for these diseases, any screening test should be assessed with mortality as the main outcome. Whether these values are common among a broad community of patients deserves further study. Then there is the issue of patient preferences for different outcomes. Even if there is no effect on all-cause mortality, my clinical experience is that most patients would prefer some other cause of death to a death from cancer. Therefore for most diseases, I do not think that all-cause mortality should be considered the main outcome. Where the authors' data are most compelling is the evidence or lack thereof for a disease-specific effect on mortalities of cancers. Here, my clinical experience is that what patients are most concerned about is death from that cancer, be it lung, prostate, breast etc., and reducing the risk of that outcome is their paramount concern. It is hard for me to imagine having any enthusiasm for a screening test for cancer without convincing evidence that it would reduce disease-specific mortality.

The second issue I wish to comment on is what constitutes convincing evidence. The authors claim that this must come from randomized controlled trials with one group being offered screening and the other group not getting screened. For the most part, I agree with them. But there are exceptions. Cervical cancer screening has not been subjected to the kind of randomized controlled trial advocated by the authors, yet the observational evidence that mass screening programmes have had a beneficial effect is sufficiently strong to conclude that there is a cause-and-effect relationship. However, this is a historical issue, and I can agree with the authors that newly proposed tests should be subject to randomized trials assessing their benefits and harms.

In sum, the evidence synthesized by Drs Saquib, Saquib and Ioannides should be considered by anyone contemplating clinical practice guidelines about screening or proposing new screening tests. We have let too much get into routine practice without an adequate evaluation, and once widely disseminated, it can be very difficult to re-orient patient expectations and clinical behaviours to an understanding that a randomized trial comparing screening with no screening is ethically justified.

Conflict of interest: The author has received royalties for writing two chapters for *UpToDate*. In addition he receives an honorarium from ECRI for serving on the National Guidelines Clearing House and National Quality Clearing House (NGC/NQMC) Committee.

Commentary: Screening: a seductive paradigm that has generally failed us

International Journal of Epidemiology, 2015, 278–280
doi: 10.1093/ije/dyu267
Advance Access Publication Date: 15 January 2015

PC Gøtzsche

Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark. E-mail: pcg@cochrane.dk

Screening healthy people has face value and great public and political appeal. It looks so simple, and yet screening is fraught with difficulties. These start already with the terminology, and common slogans like, 'Catch the disease early, before it has produced any symptoms!' are misleading on two counts.

First, disease means lack of ease, which is not what we understand by being healthy; but people who work

with screening tend to forget that they deal with healthy people. For example, women being invited to mammography screening are often called patients in scientific articles.

The second error is the assumption that the disease is caught early. That is rarely the case, and breast cancer is again a good example. If we assume that the growth rate for a particular cancer is constant, then the women have harboured the cancer for 21 years on average before it is large enough to be detected by mammography screening.[1] Finding precursors to cancer is of course an entirely different matter. Screening with flexible sigmoidoscopy identifies polyps and vaginal smear finds carcinoma *in situ*.

A third problem with screening is that it always causes harm. Sometimes it also leads to benefits, and sometimes the benefits are sufficiently large to outweigh the harms. The main focus in screening trials should therefore be to quantify the harms, but this has rarely been the case, if ever. Screening trials focus on disease-specific mortality, which may seem natural, but it is the wrong outcome. Screening leads to overdiagnosis, and interventions that are beneficial for real patients can be lethal for healthy overdiagnosed people. Radiotherapy of overdiagnosed women may kill at least as many as those who are spared dying from breast cancer by attending breast screening.[2]

Total mortality should therefore be the primary outcome in screening trials of mortality, and Saquib *et al.* report a systematic review in this issue of the journal that aimed at clarifying whether screening lowers total mortality for diseases that carry a high disease-specific mortality.[3] They focused on cancer, cardiovascular diseases, type 2 diabetes and chronic obstructive pulmonary disease. They did not find any screening trials for hypertension or chronic obstructive pulmonary disease. Disease-specific mortality was reduced with ultrasound for abdominal aortic aneurysm in men, mammography for breast cancer and faecal occult blood test and flexible sigmoidoscopy for colorectal cancer, but the risk ratio point estimates for all-cause mortality were all very close to 1.00 (range 0.98–1.03).

Screening proponents often say that disease-specific mortality is the right outcome, arguing that in order to show an effect on total mortality, trials would become unrealistically large. I believe this argument is invalid, for both scientific and ethical reasons. We do randomized trials in order to avoid bias, and our primary outcome should therefore not be a biased one. Drug interventions are usually more common in a screened group, and they tend to increase mortality for a variety of non-disease related reasons.[4]

From an ethical perspective, it is problematic to screen the whole population in a certain age group without knowing whether this makes people live longer, while knowing almost certainly that it makes people less happy. It took 50 years after the first randomized trial of mammography started before we knew what the psychological consequences are of the many false-positive findings.[5] A specially designed questionnaire was developed using focus groups and women who had attended screening were followed up for 3 years. Even after so long a time, those who had experienced a false-positive diagnosis had an anxiety level (and other psychological problems) that fell between that for women with breast cancer and women who had always been told they did not have cancer. This study showed for the first time that the psychological harms of breast screening are substantial and long-lasting, and they affect a huge number of healthy women, as the cumulative risk of a false-positive result after 10 mammograms ranges from about 20% to 60%.[6] Added to this comes the psychological harm inflicted on all the overdiagnosed women who do not know that they are overdiagnosed but think that they suffer from a fatal disease. It is therefore pretty clear that any utility analysis that takes the psychological harms of breast screening into account will come out negative, as was recently reported by the Swiss Medical Board.[7]

Saquib *et al.* found no screening trials for hypertension and only one for diabetes, ADDITION-Cambridge, for which the risk ratio for all-cause mortality was 1.06. In our systematic review of general health checks,[8] 7 of the 16 trials screened for diabetes, and likely all of them screened for hypertension (in one, the screening tests were not specified). Although we had 11 940 deaths, we did not find an effect on total mortality (risk ratio 0.99, 95% confidence interval 0.95 to 1.03). We could not include the most recent trial, as it was published in 2014.[9] It investigated the effect of systematic screening for risk factors for ischaemic heart disease and lifestyle counselling. This trial was large but it also failed to find an effect on total mortality: 3163deaths occurred, and the hazard ratio was 1.00 (0.91 to 1.09).

It is worth noting that when screening does not work, it might be because beneficial effects are outweighed by harmful ones. Diabetes drugs, for example, are approved on the basis of their glucose-lowering effect without knowing what they do to patients. And the only large trial of tolbutamide ever performed was stopped prematurely because the drug increased cardiovascular mortality.[4] Rosiglitazone was once the most-sold diabetes drug in the world, but it was taken off the market in Europe in 2010 as it causes myocardial infarction and cardiovascular death; and pioglitazone has been linked to heart failure and bladder cancer.[4]

Screening is popular, but we need to be much more careful in the future when we contemplate approaching healthy people with our screening tests, and should demand much stronger evidence than when we treat patients.

Conflict of interest: None declared.

References

1. Gøtzsche PC, Jørgensen KJ, Zahl PH, Maehlen J. Why mammography screening has not lived up to expectations from the randomised trials. *Cancer Causes Control* 2012;23:15–21.
2. Baum M. Harms from breast cancer screening outweigh benefits if death caused by treatment is included. *BMJ* 2013;346:f385.
3. Saquib N, Saquib J, Ioannidis J. Does screening for disease save lives in asymptomatic adults? Systematic review of meta-analyses and randomized trials. *Int J Epidemiol* 2015;44:264–77.
4. Gøtzsche PC. *Deadly Medicines and Organised Crime: How Big Pharma Has Corrupted Health Care*. London: Radcliffe Publishing, 2013.
5. Brodersen J, Siersma VD. Long-term psychosocial consequences of false-positive screening mammography. *Ann Fam Med* 2013;11:106–15.
6. Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2013;6:CD001877.
7. Swiss Medical Board. *Systematisches Mammographie-Screening*. 2013. www.medicalboard.ch/fileadmin/docs/public/mb/Fachberichte/2013-12-15_Bericht_Mammographie_Final_Kurzfassung_e.pdf (date last accessed, 1 October 2014).
8. Krogsbøll LT, Jørgensen KJ, Grønhøj Larsen C, Gøtzsche PC. General health checks in adults for reducing morbidity and mortality from disease. *Cochrane Database Syst Rev* 2012;10:CD009009.
9. Jørgensen T, Jacobsen RK, Toft U, Aadahl M, Glümer C, Pisinger C. Effect of screening and lifestyle counselling on incidence of ischaemic heart disease in general population: Inter99 randomised trial. *BMJ* 2014;348:g3617.

International Journal of Epidemiology, 2015, 280–282
doi: 10.1093/ije/dyu269
Advance Access Publication Date: 15 January 2015

Commentary: Tempering expectations of screening: what is the most authoritative advice we can give, given the data that we have?

Paul Taylor

Institute of Health Informatics, University College London, 222 Euston Road, London NW1 2DA, UK.
E-mail: p.taylor@ucl.ac.uk

The most authoritative basis for supporting a medical intervention is a meta-analysis of all sufficiently rigorous relevant randomized controlled trials. In this issue Saquib, Saquib and Ioannidis present an unprecedentedly thorough survey of 9 meta-analyses and 48 trials representing the best available evidence for the effectiveness of a range of screening interventions.[1] Some of the evidence reviewed has been argued over before. In the case of breast cancer, probably

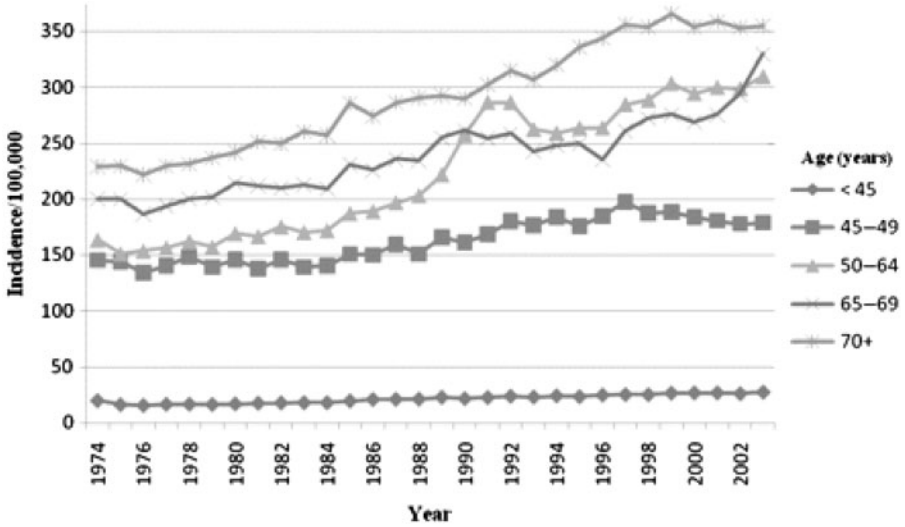

Figure 1. Incidence of breast cancer by age group in the UK from 1974 to 2004. (Reproduced from Duffy et al.[5]).

the most debated screening intervention, a series of large trials in the 1970s and 1980s provided what seemed to be clear evidence that screening saved lives, and countries across the developed world introduced programmes. Yet, in 2000, a meta-analysis concluded that there was no reliable evidence for breast screening.[2] The authors, Olsen and Gøtzsche, had identified eight trials but argued that the results of the six more or less favourable trials could not be trusted and that only the two more equivocal trials were sound.

The inclusion of one of the larger trials, the Swedish two-counties trial, should have been enough to reverse Olsen and Gøtzsche's conclusion.[3] They had argued that this trial had to be excluded in part because the average ages of the two groups were slightly different. The point is not that this difference—which was only 5 months—affected the results, but rather that any difference between two such large samples casts suspicion on the claim that they were randomly allocated. A further concern was that the investigators reviewing deaths among participants knew which arm of the trial a woman was in when they decided whether to count her death as caused by breast cancer death or not.

Whether or not these concerns are sufficient to warrant the exclusion of the trial is a matter of judgement and judgements, in this case, differed. Olsen and Gøtzsche were not the first or the last to attempt a meta-analysis of breast screening: there have been more reviews than there are trials to review. The arguments have been bitter, but have led towards consensus. Gøtzsche updated his analysis in 2011,[3] including more trials and finding overall support for the conclusion that screening reduces breast cancer deaths. An independent panel of UK experts, commissioned to look at the evidence, published a report in 2012 that drew on Gøtzsche's revised review to conclude that screening does reduce breast cancer deaths.[4] The United States Preventative Services Task Force has made a similar assessment.[6] Saquib, Saquib and Ioannidis, following Gøtzsche's updated analysis, give breast cancer as a case where screening reduces disease-specific mortality.[1]

But there's the rub. If breast cancer deaths are reduced, but all-cause mortality is unaffected, is this because detecting the latter requires that more statistical power be deployed? Or is it, as Gøtzsche has suggested, because the harms of screening increase deaths from other causes? The most serious cause of harm is overdiagnosis. The independent UK panel took the view that the best estimate of overdiagnosis could be provided by comparing the rates of cancer detection in the screened and the unscreened groups of randomized controlled trials. The problem is that when most trials ended, screening was offered to the women in the control groups, creating overdiagnosis in the follow-up period. The panel therefore restricted their attention to three trials in which no screening was offered to the control group during follow-up. This is a very limited set of data. Saquib, Saquib and Ioannidis ignore the question of harms presumably because there simply are not enough RCT data to review.

It is striking that almost all the patients screened in the reviewed trials that show a benefit due to screening, had their ultrasound, mammogram, sigmoidoscopy or faecal occult blood test in the past century, many of them in the 1970s and 1980s. For many cancers the benefits of early detection have been attenuated since then as a consequence of improvements in the treatment of late-stage disease. Trials of screening are expensive. Tens, sometimes hundreds, of thousands of participants are required and follow-up periods of 10 and 20 years are needed. Saquib, Saquib and Ioannidis's review lists only 48 trials. Restricting ourselves to this subset of the available data may be the best defence against methodological error, but in a changing world it clearly limits our capacity to base policy on relevant evidence.

Data other than those from trials could be used to provide evidence about the benefits and harms of screening.

The above graph, for example, shows a spike in the incidence of cancer in women of 50 to 64 years of age following the start of screening programme. We can use this to calculate how much overdiagnosis there is if we can estimate (i) the gradual increase in incidence observed before 1988—which presumably would have continued along the same trajectory had screening not been introduced—and (ii) the compensatory drop in incidence in older women who have been through screening. Unfortunately the aggregation of uncertainties in the calculation of these two figures means that wildly different estimates of overdiagnosis rates can be derived and indeed are derived.[6] We need a process similar to that which has allowed a degree of consensus to emerge on the validity of evidence from moderately flawed clinical trials, before we can use the data collected in the course of routine screening.

The abstract of this review[1] concludes: 'Among currently available screening tests for diseases where death is a common outcome, reductions in disease-specific mortality are uncommon and reductions in all-cause mortality are very rare or non-existent'. As I read it, 'uncommon' equates to 30% and 'very rare or non-existent' to 11%. The 30% figure is presented as disappointing. Perhaps it is, but remember that even an advocate of screening would expect a good proportion of trials to fail. One issue that is not discussed is the impact of our increasing capacity to stratify populations on the basis of risk. This should allow us to optimize screening programmes and improve outcomes. The cautious tempering of expectations advised by Saquib, Saquib and Ionnidis is prudent but should not be overdone.

Conflict of interest: None declared.

References

1. Saquib N, Saquib J, Ioannidis J. Does screening for disease save lives in asymptomatic adults? Systematic review of meta-analyses and randomized trials. *Int J Epidemiol* 2015;44:264–77.
2. Gøtzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;355:129–34.
3. Tabar L, Fagerberg G, Duffy SW, Day NE. The Swedish two county trial of mammographic screening for breast cancer: recent results and calculation of benefit. *J Epidemiol Community Health* 1989;43:107–14.
4. Gøtzsche, Peter C, and Margrethe Nielsen. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2011;CD001877:4.
5. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012;380:1778–86.
6. Nelson H, Tyne K, Naik A, Bougatsos C, Chan B, Humphrey L. Screening for Breast Cancer: An Update for the U.S. Preventive Services Task Force. *Ann Intern Med* 2009;151:727–37.
7. Duffy SW, McCann J, Godward S *et al*. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *J Med Screen* 2010;17:25–30.
8. Duffy SW, Chen TH-H, Smith RA, Amy Ming-Fang Yen AM-F, Tabar L. Real and artificial controversies in breast cancer screening. *Breast Cancer Manag* 2013;2:519–28.