FACULTY OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

PhD thesis

Lasse T Krogsbøll

Benefits and harms of general health checks and screening with urinary dipsticks

Principal supervisor: Peter C Gøtzsche

Co-supervisor: Karsten Juhl Jørgensen

Institute: Rigshospitalet

Department: The Nordic Cochrane Centre

Author: Lasse T. Krogsbøll

Titel: Gavnlige og skadelige virkninger af generelle helbredsundersøgelser og screening med urinstiks

Title: Benefits and harms of general health checks and screening with urinary dipsticks

Subject description: Systematic reviews of general health checks and screening with urinary dipsticks.

Submitted: 3 January 2015

This thesis has been submitted to the Graduate School of The Faculty of Health and Medical Sciences, University of Copenhagen.

Table of contents

Preface

The thesis is structured according to the guidelines of the Graduate School of Health and Medical Sciences at the University of Copenhagen.

The work on which this thesis is based was conducted at the Nordic Cochrane Centre, with a 3-month stay at the German Cochrane Centre.

The principal supervisor was Peter C Gøtzsche and Karsten Juhl Jørgensen was co-supervisor.

This thesis is built on the following four papers:

1. Krogsbøll LT, Jørgensen KJ, Grønhøj Larsen C, Gøtzsche PC. General health checks in adults for reducing morbidity and mortality from disease. Cochrane database of systematic reviews 2012;10:CD009009.
2. Krogsbøll LT. Guidelines for screening with urinary dipsticks differ substantially. Danish Medical Journal 2014;61(2):A4781.
3. Krogsbøll LT, Jørgensen KJ, Gøtzsche PC. Screening with urinary dipsticks for reducing morbidity and mortality. Cochrane database of systematic reviews **(accepted for publication).**
4. Krogsbøll LT, Jørgensen KJ, Gøtzsche PC. Downstream consequences of screening with urinary dipsticks - systematic review of observational studies. **(not yet submitted for publication).**

Acknowledgements

First, I wish to thank my two supervisors, Peter C. Gøtzsche and Karsten Juhl Jørgensen for their support in all phases of the research in this thesis. I also wish to thank my other colleagues at the Nordic Cochrane Centre for a stimulating environment that has made my time there a great learning experience as well as a genuinely good time. Most importantly, I wish to thank Helle Gerda Olsen for her enormous support.

Financial support

The research in this thesis was funded by the Nordic Cochrane Centre and Trygfonden.

English summary

General health checks consist of several screening tests, often accompanied by lifestyle counselling. They have been increasing in popularity during the 20th century. Their aim is to find and treat previously undetected disease and risk factors, and to prevent serious illness. Few of the tests commonly included in health checks have been studied in randomised trials, and the effects of health checks were generally unclear.

The aim of my PhD was to investigate the beneficial and harmful effects of general health checks, with a special focus on screening with urinary dipsticks, which is a common component in health checks. Four studies were conducted to this aim. The first was a Cochrane review of randomised trials of general health checks. The second was a study of the recommendations given by public health authorities and specialist societies regarding screening with urinary dipsticks. The third was a Cochrane review of randomised trials of screening with urinary dipsticks. The fourth was a review of observational studies in an attempt to quantify the harmful effects of screening with urinary dipsticks.

The results showed that health checks do not have documented beneficial effects on mortality, morbidity, and other important outcomes, while the harmful effects are poorly studied. However, there were signs that health checks lead to more diagnoses being made and to more drug therapy. Screening with urinary dipsticks has not been studied in randomised trials, and it is thus unknown whether it has beneficial effects. The harmful effects have not been adequately studied, but it is certain that it results in invasive diagnostic procedures and medicalisation of previously well persons, although the extent is unknown.

My conclusions are that general health checks do not seem to have a favourable balance between benefits and harms, while that balance is unknown for screening with urinary dipsticks. General health checks have been so thoroughly studied that further trials seem unnecessary unless significant improvements in risk assessment and preventive treatments occur. Screening with urinary dipsticks should be studied in randomised trials.

Danish summary

Generelle helbredstjek består af flere screeningstest, ofte ledsaget af livsstilsrådgivning, og har været tiltagende populære gennem det 20. århundrede. Formålet er at finde og behandle uerkendt sygdom og risikofaktorer, og at forebygge alvorlig sygdom. Få af de test der hyppigt indgår i helbredstjek har været undersøgt i randomiserede forsøg, og effekterne af helbredstjek var generelt uklare.

Formålet med denne ph.d. var at undersøge de gavnlige og skadelige virkninger af generelle helbredstjek, med et særligt fokus på screening med urinstiks, som er en hyppigt anvendt komponent i helbredstjek. Fire studier blev gennemført med henblik på dette. Det første studie var et Cochrane review af randomiserede forsøg med generelle helbredstjek. Det andet var en undersøgelse af hvilke anbefalinger der bliver givet af offentlige sundhedsmyndigheder og specialeselskaber om screening med urinstiks. Det tredje var et Cochrane review af randomiserede forsøg med screening med urinstiks. Det fjerde var et review af observationelle studier, i et forsøg på at kvantificere de skadelige virkninger af screening med urinstiks.

Resultaterne viste at generelle helbredstjek ikke har dokumenterede gavnlige virkninger på dødelighed, sygelighed og andre vigtige effektmål, mens de skadelige virkninger er sparsomt belyst. Der var dog tegn på at helbredstjek medfører flere diagnoser og mere medicinsk behandling. Screening med urinstiks har ikke været undersøgt i randomiserede forsøg, og det er således uvist om der er gavnlige virkninger. De skadelige virkninger er ikke blevet tilstrækkeligt undersøgt, men det er sikkert at urinstiks resulterer i invasive diagnostiske procedurer og sygeliggørelse af raske personer, omend i ukendt omfang.

Mine konklusioner er, at generelle helbredstjek ikke lader til at have en gunstig balance mellem gavnlige og skadelige virkninger, mens den balance er ukendt for screening med urinstiks. Generelle helbredstjek er blevet så grundigt undersøgt, at yderligere forsøg forekommer overflødige, medmindre der opstår betragtelige forbedringer i risikovurdering og forebyggende behandling. Screening med urinstiks bør studeres i randomiserede forsøg.

Introduction

The evolution of medicine in the 20th century, with its advances in diagnostic technologies and the advent of modifiable risk factors for disease, led to optimism about the possibilities of detecting disease early in persons without symptoms. The success in eradicating or preventing several infectious diseases contributed to this optimism.[1] Many disease processes pass through several phases before symptoms occur, and it could be beneficial if these processes were detected and treated before irreversible damage occurs. For example, cancers have better prognoses at early stages, and type 2 diabetes mellitus can cause complications after decades with few or no symptoms.

Screening is the application of a test to identify people who are at increased risk of suffering from a particular disease, or already have it, and general health checks combine several different screening tests and are often repeated regularly. They lack a consistent definition and vary in content, but their focus is usually on cardiovascular risk. They can be offered as part of an organised screening programme, or on the initiative of physician or patient, be offered by an employer, or be prompted by advertising from commercial providers. In addition to laboratory tests or imaging, a physical examination by a doctor is sometimes included, which is also in principle a screening test when directed at asymptomatic persons. Lifestyle counselling is also often included.

Most individual screening tests used in health checks have an inadequate evidence base.[2] Screening with urinary dipsticks is one of these, but its use is prevalent.[2–4] Urinary dipsticks can detect multiple substances in the urine that are associated with disease and can allow earlier identification of disease, and possibly a better prognosis through earlier treatment.

Screening

While some screenings lead to benefit, all screenings lead to harm.[5] For example, screening for colorectal cancer reduces mortality from colorectal cancer, but also leads to unnecessary colonoscopies and overdiagnosis and overtreatment of non-cancerous polyps, with a risk of serious complications such as colonic perforation.[6] In addition, some experience psychological harm from false positive results, or possibly from the invitation itself, as it highlights the risk of serious illness to an otherwise well citizen.

Benefit is not a given in screening, and convincing biological rationales have sometimes proved misleading. For example, the benefit of early identification of the childhood cancer neuroblastoma by a simple urine test appeared self-evident when it was introduced in Japan in 1985.[7] Initial observational studies suggested a substantial benefit, but much later controlled studies from Germany and Canada showed no effect on mortality but an increase in neuroblastoma incidence with screening, resulting in unnecessary treatment with surgery and chemotherapy.[8,9] The increased incidence with screening was caused by detection of cases that would have regressed spontaneously, i.e. cases that did not need therapy and in which detection was harmful.

Other screening tests have been evaluated in randomised trials during the past 4-5 decades, and in some cases the benefits have been doubtful,[10] in some cases benefit has been absent,[11,12] and in some cases the benefit seems real.[6,13] Although best described in cancer screening, the concepts of overdiagnosis and overtreatment apply to other areas as well, e.g. cardiovascular prevention.[14]

Components of health checks

Health checks differ from the screening tests mentioned above, in that they are collections of several tests bundled into one examination.

A central component in health checks has been the assessment of cardiovascular risk. A plethora of risk factors for cardiovascular disease has been identified and some are believed to be causal, e.g. high blood pressure, high cholesterol, smoking, male gender, and diabetes mellitus. Although screening for several of these risk factors have been the cornerstone of health checks for decades, no trials exist of screening for high blood pressure or high cholesterol in isolation, and only in 2012 was a randomised trial of screening for type 2 diabetes mellitus published.[15] Screening for cardiovascular risk factors is now viewed as a package, and the current approach is to assess and treat overall risk based on a combination of these factors, rather than treating them individually, unless severely elevated.[16] In contrast, knowledge about the benefit of treating risk factors comes mainly from trials of treating one risk factor at a time, e.g. high blood pressure. Also, there are gaps in the documentation of efficacy: Treatment of moderately and severely elevated blood pressure reduces morbidity and mortality, but this has not been demonstrated for treatment of mild hypertension.[17] Treatment of elevated blood glucose has been surprisingly poorly studied, as few trials have compared the effect of medical treatment with placebo on clinical outcomes. Intensive treatment of diabetes reduces microvascular complications, but not macrovascular complications or mortality, compared with usual care, and it increases the risk of harms.[18] Treatment with statins appears effective in both primary and secondary prevention,[19] but the effects are usually small in primary prevention. While harms of statins are reported to be few, most trials have been performed by or funded by producers of such drugs, raising concerns about bias.[20]

Examples of other common components in health checks are blood tests, electrocardiography, spirometry, and physical examination by a doctor. These tools have been developed to assess patients with medical complaints and are part of the standard armamentarium of clinicians. Their abilities as screening tools for people without relevant medical complaints have been poorly studied, if at all.

Considering these uncertainties, it is not a given that health checks lead to benefit, or that the harms caused in healthy individuals are acceptable. Previous reviews of the value of screening for cardiovascular risks have been inconclusive given the lack of trials with clinical outcomes.[21,22] A systematic review of the periodic health evaluation found beneficial effects on delivery of preventive health services and surrogate outcomes, but not on clinical outcomes.[23]

Choice of methods

Designs

The randomised trial is the preferred design for assessing medical interventions since it is more likely to give an unbiased answer to the research question. Health checks, and probably also dipstick screening, are susceptible to self-selection bias in observational studies, as the chance of having a health check is related to socio-economic status, education, and general health.[24–25] Since randomised trials are feasible for both interventions, we chose this design for systematic reviews of benefits and harms. Since there were no randomised trials on dipstick screening we chose to use observational studies for a review of harms, acknowledging the likely biases. For a study on the content of recommendations and guidelines for dipstick screening, a cross-sectional study was done using the internet and email contact. The specific methods used are described in detail in the papers.

Outcomes

A crucial point when assessing the evidence for an intervention is the choice of outcomes. In health checks, modification of cardiovascular risk factors is usually the cornerstone of the intervention, so some have viewed changes in risk factor levels as the most important outcomes. Observed reductions in cardiovascular risk factors have been used to calculate life-years gained,[26] and have formed the basis of modelling studies used for planning, economic analyses and political decision making.[27,28] In studies of renal disease, the degree of proteinuria and the estimated glomerular filtration rate are examples of surrogate outcomes.

Risk factors are by definition surrogate outcomes, and are thus named because they are substitutes for the real outcomes, i.e. those outcomes that directly matter to people. For example, high blood pressure is usually asymptomatic and usually only of interest because of the diseases it may cause, such as stroke. Since clinical outcomes, such as myocardial infarction, stroke, or death, are infrequent in prevention trials of healthy people, using them requires large sample sizes and long follow-up times. In contrast, a surrogate outcome such as blood pressure can be measured on all participants, and thus all participants may contribute information, instead of just those who experience events.

However, there are problems with this approach.[29] Extrapolating effects on clinical outcomes from effects on surrogate outcomes requires assumptions about the extent to which cardiovascular risk accumulated over a lifetime is reversible and about how fast the reduction in risk occurs. Since the main reason for using surrogate outcomes is that it allows smaller and shorter trials, extrapolating such measurements to effects on morbidity and mortality also requires assumptions about how the achieved changes in risk factors are maintained over longer periods. Furthermore, there are competing risks, which are particularly important in old people. Sometimes effects on clinical outcomes are modelled by using surrogate outcomes to estimate the number of persons that

will be offered treatment and applying treatment effects as estimated in trials. This approach also requires assumptions, such as the likely uptake of treatment, the adherence to suggested treatment, and the generalisability of treatment effects to the population in question. Industry trials often exclude persons with co-morbidity and with a high age and may thus not resemble the population that will actually be using the drugs.[30]

Using surrogate outcomes disregards possible harms from the drugs used to lower the risk factor levels, and from the health checks themselves and from follow-up tests of abnormal results. Another problem with risk factors is that they involve measurements on each participant at the end of the trial, and the results can therefore be biased when some participants drop out.

Paradoxically, proper validation of a surrogate outcome requires a large amount of information from trials with clinical outcomes,[31] exactly what they are meant to avoid, and even then uncertainty remains since new drugs may have new harmful effects. For example, torcetrapip lowered LDL-cholesterol and increased HDL cholesterol, which are well-know surrogate markers, but the treatment increased mortality.[32]

Given these uncertainties, we chose not to use risk factors as outcomes in our review. Instead, we chose total and disease-specific mortality as our primary outcomes, since they are least likely to be biased, and since they reflect both benefits and harms. For secondary outcomes, we chose morbidity and other parameters relevant to patients and society as a whole, e.g. admission to hospital.

Objectives

Since health checks are in widespread use we found it pertinent to investigate the available evidence. In addition to a review of randomised trials on health checks, we chose urinary dipsticks for closer study, as it is a common component in health checks and is widely used as a screening test at hospital admission. The objectives were:

- to estimate the beneficial and harmful effects of general health checks in a systematic review of randomised trials, focussing on clinically relevant outcomes such as morbidity and mortality (Paper 1)
- to assess the recommendations given by public authorities and specialist organisations regarding screening with urinary dipsticks (Paper 2)
- to estimate the beneficial and harmful effects of screening with urinary dipsticks in a systematic review of randomised trials, using clinically relevant outcomes (Paper 3)
- to estimate the harmful consequences of screening with urinary dipsticks in a systematic review, using observational studies in the absence of trials (Paper 4).

Discussion of papers

Paper 1

Our Cochrane review of general health checks showed that the evidence base was larger than expected, with several large trials having been conducted, and with no signs of beneficial effects on a range of important outcomes. The harms of health checks were inadequately studied or reported, but we did see some signs that health checks lead to more diagnoses and more treatment with drugs for hypertension.

The main conclusion of this review was that we could not show benefit on any outcomes. Screening should not be implemented unless there is clear evidence from randomised trials that the benefits outweigh the harms and our review clearly showed that such evidence does not exist for health checks.

A related conclusion was that it seemed unlikely that the intervention provided a net benefit. When considering the lack of effect in the mortality meta-analyses, combined with the lack of beneficial effects on the secondary outcomes, it is difficult to imagine an underlying true beneficial effect that was overlooked. This statement of probability was not arrived at using Bayesian statistical methods, but was the result of our judgement based on the data. Bayesian methods usually come to the same results as frequentist methods when non-informative priors are used, and using informative priors would have been a challenge as beliefs in health checks differ wildly, which the debate subsequent to publication of our review clearly demonstrated. However, we could have performed a Bayesian analysis using a wide range of priors. Trial sequential analysis[33] could also have been helpful, as it

allows a formalised assessment of futility in doing additional trials, i.e. whether enough information is currently available to reasonably allow a conclusion of no effect.

An important question is what can be concluded from the apparent lack of beneficial effects. In principle, it is never possible to prove with certainty that an intervention does not work. The general limitation to inductive reasoning applies here: repeated similar observations do not serve as a guarantee that future observations will be the same. Also, one could always imagine that a true effect exists but that the power was inadequate. Another way of phrasing this is that no matter how narrow the confidence interval may be in a meta-analysis, it is in principle always possible to imagine an even smaller real effect. Here it should be remembered that with medical interventions the direction of an effect where the confidence interval overlaps no effect is never guaranteed to be favourable, and medicine has always had an unfortunate tendency towards widespread use of harmful interventions.[34]

Another general objection to meta-analyses is that they combine effect estimates that were obtained using different interventions and participants. However, what constitutes unacceptable variation between study interventions, designs, and participants depends on the questions asked, as in clinical trials. When trying to answer a broad question, one must necessarily include a broad range of data, and we found the clinical and methodological heterogeneity to be acceptable in relation to answering our general question of whether health checks work. The statistical heterogeneity varied, but was generally acceptable. The most important tool to prevent unacceptable heterogeneity is pre-specification of eligibility criteria, defining a population of studies that can provide answers to the same general question. We pre-specified our eligibility criteria, review methods, and outcomes in the peer reviewed and published protocol before undertaking the review.

The publication of our review spurred much debate. We obtained permission to co-publish the review in both the Cochrane Library and the BMJ[35], and subsequently JAMA[36] invited us to submit a 2-page summary. Debate occurred in these journals, in Ugeskrift for Læger, and elsewhere, and was covered in many news media. The Danish Minister of Health scrapped existing plans for implementing health checks, and in the UK the freshly introduced NHS Health Check programme was subjected to criticism for having been implemented without adequate supporting evidence. This made front page news of The Times[37] after we had sent a letter describing our experience of being denied the right to respond to a scientific attack that was published on the website of the NHS Health Check.[38]

In our review, and in the subsequent debate, we described several possible reasons for the observed lack of effect: opportunistic screening done by general practitioners in the control groups, inadequate uptake of health checks among high-risk groups, poorer performance of preventive drugs in real-life settings than in industry trial settings. The old age of many of the trials does not weaken the central conclusion about the lack of evidence for benefits of health checks, but does somewhat weaken the conclusions that can be drawn about the absence of effect in today's setting. However, subsequent to our review, the results on morbidity and mortality from the large Inter99 trial was published, and the main findings were similar to ours. This weakens the hopeful argument that although health checks did not work in the older trials included in our review, they might work today. A 2014 systematic review of health checks in general practice found benefits on surrogate outcomes such as blood pressure and cholesterol, but did not find benefits on total or cardiovascular mortality.[39] Although the authors, and others,[40] highlighted the improvements in surrogate

outcomes, the absence of benefit on mortality hints that surrogates are inappropriate in this context. A similar discrepancy was reported by the Inter99 study authors, who reported favourable effects on smoking, diet, physical activity, and binge drinking after 5 years, but a complete absence of benefit on clinical outcomes after 10 years.[41]

Paper 2

In the second study I investigated the guidance available to clinicians when deciding whether to offer screening with urinary dipsticks. In a thorough and systematic search I found no recommendations on screening with combined dipsticks, some on screening for bacteriuria with dipsticks, and little on screening for glucose, blood and protein. The conclusion was that clinicians are largely left to themselves when deciding whether to offer dipstick screening. It is no surprise that an easy and intuitively appealing test such the urinary dipstick is used for screening in the absence of guidance against it, as many people are intuitively attracted to the idea of early detection of disease. The study was limited by the selection of countries included and medical specialties searched, but it nonetheless represents a search effort that far surpasses that which any clinician would be likely to undertake to inform his or her clinical practice.

Paper 3

In our Cochrane review of screening with urinary dipsticks, we found no trials. This is startling as the intervention is prevalent in western countries, and as large-scale screening programmes have been running for years in several Asian countries. Thus, any beneficial effects are of unknown magnitude, and it is not a given that they exist at all. The fact that screening with dipsticks finds persons with asymptomatic disease does not prove that it is beneficial. Earlier detection may not

alter prognosis, and overdiagnosis and overtreatment plus the harms of invasive diagnostic procedures may outweigh any benefits.

As an example, asymptomatic IgA nephropathy detected by urine screening is likely less aggressive than symptomatic IgA nephropathy,[42,43] and may remain subclinical. If this is the case, treatment has less potential to cause benefit, but has the same potential to do harm, for example in the form of treatment with immunosuppressant drugs, and by lifelong nephrological follow-up and resulting medicalisation. The situation could be analogous to overdiagnosis of breast cancer in mammography screening, or overdiagnosis of prostate cancer by PSA screening. Combined dipsticks can detect many different diseases, and even if some of these get a better prognosis when caught early, this could be outweighed by harms from early detection of other diseases. This situation is analogous to health checks, which also combine tests with unknown benefits and harms. In the case of health checks, such combinations of common screening tests do not seem to carry any benefit, while in the case of urinary dipsticks the net effect is unknown.

Paper 4

In the fourth paper we tried to estimate the harms of screening with urinary dipsticks using data from observational studies. Although the dipstick test is harmless in itself, the subsequent diagnostic workup of positive results may not be. For example, sometimes kidney disease needs to be investigated with a kidney biopsy, and persistent microscopic haematuria is evaluated with a computed tomography scan (CT) with intravenous contrast and also often cystoscopy. Such procedures are harmful in themselves, and in the absence of proven benefit of dipstick screening, the general population should not be subjected to them. A quantification of the frequency of these and other harms was thus pertinent.

In 32 included studies we found that the frequency of positive tests varied greatly, as did the number of new diagnoses. We documented that imaging, invasive testing, medical and surgical treatment, and long-term follow-up sometimes results from dipstick screening, but were unable to give precise estimates due to the limited amount of available data combined with a large clinical and methodological heterogeneity. Our main conclusion was that the harms have been inadequately studied.

The harms of screening with urinary dipsticks would be best assessed in randomised trials, with the exception of long-term harms such as cancers resulting from radiation exposure from CT scans. Randomised trials would allow direct measurement of the frequencies of harmful events in screened and non-screened groups, and would allow a simple calculation of the effect. The second-best option would be non-randomised comparisons of populations invited to screening and similar populations not invited to screening. Given the absence of such studies, we chose the third best solution and assessed the frequency of harmful events in single cohorts of persons who had been screened with dipsticks. Since a non-screened population would also be expected to occasionally experience these events, this design leads to an overestimation of the effect of dipsticks, but as these events are probably rare in non-screened populations this bias may not be great. On the other hand, many studies followed up positive tests according to specific algorithms, which may have decreased the harm compared with the disorganised screening that is prevalent today. Also, few study authors had aimed to assess harms, but instead mainly focussed on the frequency of various positive findings, with the intent of showing a benefit. As a result, the harmful outcomes as well as the methods of data collection were usually poorly and haphazardly reported, and the numbers of invasive procedures were particularly uncommonly reported in the papers. These factors would lead

to underestimation of the harms. Finally, our approach does not provide us with an estimate of the population effect of screening with dipsticks, or the cost, but only the chance of experiencing the studied harms for persons who have actually been screened.

When reviewing trials the goal is always to be exhaustive, i.e. to find all existing eligible trials. In contrast, in reviews of observational studies, this is usually neither possible nor desirable due to the much larger number of observational studies, the lower standard of reporting, the more diverse terminology, and the greater risk of non-publication.[44] The types of studies we included in our review of harms are likely to be particularly vulnerable to these problems, as most of them were short-term and relatively inexpensive, and did not make comparisons between exposed and non-exposed participants. In addition, terminology is insufficiently defined and differentially used in this field, and these studies received many different design labels (e.g. case-series, cohort studies, cross-sectional studies), resulting in even greater identification problems. Thus, our review was not exhaustive, but tried to find a large and representative sample of the most important eligible studies.

In accordance with these limitations, we did not make strong or very specific conclusions, but rather presented plausible ranges of effects, with a call for more and better research.

Conclusions

Implications

Our review on health checks has already had implications, as the Danish Minister of Health scrapped the existing plans for introducing systematic health checks in the general population. It has also stimulated a debate in the UK about the NHS Health Check Programme, and about the evidence on which it is based, which is still ongoing. In the USA and Canada the results seem to have been received with less surprise, as a debate about health checks had been running for some time. However, health checks are still highly prevalent there, both because of their popularity and immediate face value appeal, but also because an annual health check is sometimes the only thing that insured Americans get from their health insurance company without co-payment, and for some people it is the only primary health care available. In this respect, it is important to highlight that the control groups in the trials we reviewed probably always had access to primary care when needed, and that the results cannot be used to justify removing people's only access to primary care.

The implications of our papers on screening with dipsticks are that randomised trials are needed to assess the beneficial and harmful effects, as the benefits are currently unknown and speculative, while harms do occur. Screening tests with undocumented benefits and poorly quantified harms should not be used, and recommendations from authorities and specialist societies should reflect this.

Future research

Further trials of bundles of screening tests as in health checks do not seem reasonable, especially as the findings of earlier trials have now been confirmed by the recent Inter99 trial. Future research

should be focused on potentially useful individual tests. It is imperative that future trials have a balanced approach to benefits and harms, with a rigorous evaluation of overdiagnosis, overtreatment, use of invasive testing, use of health care resources, and both short-term and long-term psychological effects.

With dipsticks, as with health checks, it seems advisable to investigate the individual components separately, and particularly screening for proteinuria is in need of trials, as chronic kidney disease is becoming increasingly common.

Authors conclusions

General health checks do not seem to be beneficial, and screening with urinary dipsticks have unknown benefits and harms.

Physicians and public health researchers and officials should remember the difference between patients seeking help and healthy citizens. In the words of Cochrane and Holland in 1971:

"We believe there is an ethical difference between everyday medical practice and screening. If a patient asks a medical practitioner for help, the doctor does the best he can. He is not responsible for defects in medical knowledge. If, however, the practitioner initiates screening procedures he is in a very different situation." [45]

Today, we should continue to demand clear and unambiguous evidence of a favourable balance between benefits and harms when contemplating screening, and politicians that are eager to benefit

public health with a new initiative should be advised to fund large scale trials instead of jumping to large scale implementation.

References

1. Charap MH. The periodic health examination: genesis of a myth. *Ann Intern Med* 1981;95(6):733–5.

2. Grønhøj Larsen C, Jørgensen KJ, Gøtzsche PC. Regular health checks: cross-sectional survey. *PLoS One* 2012;7(3):e33694.

3. Merenstein D, Daumit GL, Powe NR. Use and costs of nonrecommended tests during routine preventive health exams. *Am J Prev Med* 2006;30(6):521–7.

4. Prochazka A V, Lundahl K, Pearson W, Oboler SK, Anderson RJ. Support of evidence-based guidelines for the annual physical examination: a survey of primary care providers. *Arch Intern Med* 2005;165(12):1347–52.

5. Raffle A, Gray MBT-S evidence and practice. *Screening: evidence and practice*. Oxford University Press; 2007.

6. Holme Ø, Bretthauer M, Fretheim A, Odgaard-Jensen J, Hoff G. Flexible sigmoidoscopy versus faecal occult blood testing for colorectal cancer screening in asymptomatic individuals. *Cochrane database Syst Rev* 2013;9:CD009259.

7. Woods WG. Substitute "prostate cancer" for "neuroblastoma"? *J Clin Oncol* 2002;20(5):1154–5.

8. Woods WG, Gao R-N, Shuster JJ, Robison LL, Bernstein M, Weitzman S, et al. Screening of infants and mortality due to neuroblastoma. *N Engl J Med* 2002;346(14):1041–6.

9. Schilling FH, Spix C, Berthold F, Erttmann R, Fehse N, Hero B, et al. Neuroblastoma screening at one year of age. *N Engl J Med* 2002;346(14):1047–53.

10. Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. *Cochrane database Syst Rev* 2013;6:CD001877.

11. Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. *Cochrane database Syst Rev* 2013;1:CD004720.

12. Buys SS, Partridge E, Black A, Johnson CC, Lamerato L, Isaacs C, et al. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA* 2011;305(22):2295–303.

13. Hewitson P, Glasziou PP, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. *Cochrane Database Syst Rev* 2007;1:CD001216.

14. Welch HG, Schwartz L, Woloshin S. *Overdiagnosed - making people sick in the pursuit of health*. Boston: Beacon Press; 2011.

15. Simmons RK, Echouffo-Tcheugui JB, Sharp SJ, Sargeant LA, Williams KM, Prevost AT, et al. Screening for type 2 diabetes and population mortality over 10 years (ADDITION-Cambridge): a cluster-randomised controlled trial. *Lancet* 2012;380(9855):1741–8.

16. Graham I, Atar D, Borch-Johnsen K, Boysen G, Burell G, Cifkova R, et al. European guidelines on cardiovascular disease prevention in clinical practice: full text. Fourth Joint Task Force of the European Society of Cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by represent. *Eur J Cardiovasc Prev Rehabil* 2007;14 Suppl 2:S1–113.

17. Diao D, Wright JM, Cundiff DK, Gueyffier F. Pharmacotherapy for mild hypertension. *Cochrane Database Syst Rev* 2012;8:CD006742.

18. Hemmingsen B, Lund SS, Gluud C, Vaag A, Almdal TP, Hemmingsen C, et al. Targeting intensive glycaemic control versus targeting conventional glycaemic control for type 2 diabetes mellitus. *Cochrane database Syst Rev* 2013;11:CD008143.

19. Taylor F, Huffman MD, Macedo AF, Moore TH, Burke M, Davey Smith G, et al. Statins for the primary prevention of cardiovascular disease. *Cochrane database Syst Rev* 2013;1:CD004816.

20. Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. *Cochrane database Syst Rev* 2012;12:MR000033.

21. Sheridan SL, Viera AJ, Krantz MJ, Ice CL, Steinman LE, Peters KE, et al. The effect of giving global coronary risk information to adults: a systematic review. *Arch Intern Med* 2010;170(3):230–9.

22. Sheridan SL, Crespo E. Does the routine use of global coronary heart disease risk scores translate into clinical benefits or harms? A systematic review of the literature. *BMC Heal Serv Res* 2008;8(60):60.

23. Boulware LE, Marinopoulos S, Phillips KA, Hwang CW, Maynor K, Merenstein D, et al. Systematic review: the value of the periodic health evaluation. *Ann Intern Med* 2007;146(4):289-300.

24. Jørgensen T, Borch-Johnsen K, Thomsen TF, Ibsen H, Glümer C, Pisinger C. A randomized non-pharmacological intervention study for prevention of ischaemic heart disease: baseline results Inter99. *Eur J Cardiovasc Prev Rehabil* 2003;10(5):377–86.

25. Waller D, Agass M, Mant D, Coulter A, Fuller A, Jones L. Health checks in general practice: another example of inverse care? *BMJ* 1990;300(6732):1115–8.

26. Rasmussen S, Kilsmark J, Hvenegaard A, Thomsen JL, Engberg M, Lauritzen T, et al. Forebyggende helbredsundersøgelser og helbredssamtaler i almen praksis - en sundhedsøkonomisk analyse af "Sundhedsprojekt Ebeltoft". 2006;6(6). Available at: sundhedsstyrelsen.dk/publ/publ2006/cemtv/ebeltoft/ebeltoft_ce_rap.pdf (accessed 2 December 2014).

27. Pedersen KM, Bech M, Vrangbæk K. *The Danish Health Care System : An Analysis of Strengths , Weaknesses , Opportunities and Threats*. Copenhagen Consensus Center; 2011. Available at: www.copenhagenconsensus.com/sites/default/files/ConsensusReportDanishHealth_final.pdf (accessed 2 December 2014).

28. Department of Health. *Economic modelling for vascular checks*; 2008. Available at: www.healthcheck.nhs.uk/document.php?o=225. (accessed 2 December 2014).

29. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;125(7):605–13.

30. Uijen AA, Bakx JC, Mokkink HGA, van Weel C. Hypertension patients participating in trials differ in many aspects from patients treated in general practices. *J Clin Epidemiol* 2007;60(4):330–5.

31. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. *Stat Med* 2012;31(25):2973-84.

32. Barter PJ, Caulfield M, Eriksson M, Grundy SM, Kastelein JJ, Komajda M, et al. Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med* 2007;357(21):2109–22.

33. Thorlund K, Engstrøm J, Wetterslev J, Brok J, Imberger G, Gluud C. *User manual for Trial Sequential Analysis*; 2011. Available at: www.ctu.dk/tsa/files/tsa_manual.pdf (accessed 2 December 2014).

34. Porter R. *Ve og vel.* Rosinante;2001:834.

35. Krogsbøll LT, Jørgensen KJ, Grønhøj Larsen C, Gøtzsche PC. General health checks in adults for reducing morbidity and mortality from disease: Cochrane systematic review and meta-analysis. *BMJ* 2012;345:e7191.

36. Krogsbøll LT, Jørgensen KJ, Gøtzsche PC. General health checks in adults for reducing morbidity and mortality from disease. *JAMA* 2013;309(23):2489-90.

37. Smyth C. Health checks on over-40s condemned as "useless." *The Times;* August 20, 2013:1–2.

38. Krogsbøll L, Jørgensen K, Gøtzsche P. Health check check. *The Times;* August 20, 2013:25.

39. Si S, Moss JR, Sullivan TR, Newton SS, Stocks NP. Effectiveness of general practice-based health checks: a systematic review and meta-analysis. *Br J Gen Pract* 2014;64(618):e47–53.

40. Lauritzen T, Borch-Johnsen K, Sandbæk A. [General practice plays a role in health check screening.]. *Ugeskr Laeger* 2014;176(12).

41. 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.

42. Takebayashi S, Yanase K. Asymptomatic urinary abnormalities found via the Japanese school screening program: a clinical, morphological and prognostic analysis. *Nephron* 1992;61(1):82–8.

43. Ito K, Kawaguchi H, Hattori M. Screening for proteinuria and hematuria in school children--is it possible to reduce the incidence of chronic renal failure in children and adolescents? *Acta Paediatr Jpn* 1990;32(6):710–5.

44. Higgins J, Green S, (editors). *Cochrane handbook for systematic reviews of interventions version 5.1 [updated March 2011]*. The Cochrane Collaboration; 2011. Available at: www.cochrane-handbook.org.

45. Cochrane AL, Holland WW. Validation of screening procedures. *Br Med Bull* 1971;27(1):3-8.

General health checks in adults for reducing morbidity and mortality from disease (Review)

Krogsbøll LT, Jørgensen KJ, Grønhøj Larsen C, Gøtzsche PC

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2012, Issue 10

http://www.thecochranelibrary.com

TABLE OF CONTENTS

[Intervention Review]

General health checks in adults for reducing morbidity and mortality from disease

Lasse T Krogsbøll[1], Karsten Juhl Jørgensen[1], Christian Grønhøj Larsen[1], Peter C Gøtzsche[1]

[1]The Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark

Contact address: Lasse T Krogsbøll, The Nordic Cochrane Centre, Rigshospitalet, Blegdamsvej 9, 7811, Copenhagen, 2100, Denmark. ltk@cochrane.dk. l.t.krogsboll@gmail.com.

Editorial group: Cochrane Effective Practice and Organisation of Care Group.
Publication status and date: New, published in Issue 10, 2012.
Review content assessed as up-to-date: 4 July 2012.

Citation: Krogsbøll LT, Jørgensen KJ, Grønhøj Larsen C, Gøtzsche PC. General health checks in adults for reducing morbidity and mortality from disease. *Cochrane Database of Systematic Reviews* 2012, Issue 10. Art. No.: CD009009. DOI: 10.1002/14651858.CD009009.pub2.

ABSTRACT

Background

General health checks are common elements of health care in some countries. These aim to detect disease and risk factors for disease with the purpose of reducing morbidity and mortality. Most of the commonly used screening tests offered in general health checks have been incompletely studied. Also, screening leads to increased use of diagnostic and therapeutic interventions, which can be harmful as well as beneficial. It is, therefore, important to assess whether general health checks do more good than harm.

Objectives

We aimed to quantify the benefits and harms of general health checks with an emphasis on patient-relevant outcomes such as morbidity and mortality rather than on surrogate outcomes such as blood pressure and serum cholesterol levels.

Search methods

We searched *The Cochrane Library*, the Cochrane Central Register of Controlled Trials (CENTRAL), the Cochrane Effective Practice and Organisation of Care (EPOC) Trials Register, MEDLINE, EMBASE, Healthstar, CINAHL, ClinicalTrials.gov and WHO International Clinical Trials Registry Platform (ICTRP) to July 2012. Two authors screened titles and abstracts, assessed papers for eligibility and read reference lists. One author used citation tracking (Web of Knowledge) and asked trialists about additional studies.

Selection criteria

We included randomised trials comparing health checks with no health checks in adults unselected for disease or risk factors. We did not include geriatric trials. We defined health checks as screening general populations for more than one disease or risk factor in more than one organ system.

Data collection and analysis

Two authors independently extracted data and assessed the risk of bias in the trials. We contacted authors for additional outcomes or trial details when necessary. For mortality outcomes we analysed the results with random-effects model meta-analysis, and for other outcomes we did a qualitative synthesis as meta-analysis was not feasible.

Main results

We included 16 trials, 14 of which had available outcome data (182,880 participants). Nine trials provided data on total mortality (155,899 participants, 11,940 deaths), median follow-up time nine years, giving a risk ratio of 0.99 (95% confidence interval (CI) 0.95 to 1.03). Eight trials provided data on cardiovascular mortality (152,435 participants, 4567 deaths), risk ratio 1.03 (95% CI 0.91 to 1.17) and eight trials on cancer mortality (139,290 participants, 3663 deaths), risk ratio 1.01 (95% CI 0.92 to 1.12). Subgroup and sensitivity analyses did not alter these findings.

We did not find an effect on clinical events or other measures of morbidity but one trial found an increased occurrence of hypertension and hypercholesterolaemia with screening and one trial found an increased occurence of self-reported chronic disease. One trial found a 20% increase in the total number of new diagnoses per participant over six years compared to the control group. No trials compared the total number of prescriptions, but two out of four trials found an increased number of people using antihypertensive drugs. Two out of four trials found small beneficial effects on self-reported health, but this could be due to reporting bias as the trials were not blinded. We did not find an effect on admission to hospital, disability, worry, additional visits to the physician, or absence from work, but most of these outcomes were poorly studied. We did not find useful results on the number of referrals to specialists, the number of follow-up tests after positive screening results, or the amount of surgery.

Authors' conclusions

General health checks did not reduce morbidity or mortality, neither overall nor for cardiovascular or cancer causes, although the number of new diagnoses was increased. Important harmful outcomes, such as the number of follow-up diagnostic procedures or short term psychological effects, were often not studied or reported and many trials had methodological problems. With the large number of participants and deaths included, the long follow-up periods used, and considering that cardiovascular and cancer mortality were not reduced, general health checks are unlikely to be beneficial.

PLAIN LANGUAGE SUMMARY

General health checks for reducing illness and mortality

General health checks involve multiple tests in a person who does not feel ill with the purpose of finding disease early, preventing disease from developing, or providing reassurance. Health checks are a common element of health care in some countries. To many people health checks intuitively make sense, but experience from screening programmes for individual diseases have shown that the benefits may be smaller than expected and the harms greater. One possible harm from health checks is the diagnosis and treatment of conditions that were not destined to cause symptoms or death. Their diagnosis will, therefore, be superfluous and carry the risk of unnecessary treatment.

We identified 16 randomised trials which had compared a group of adults offered general health checks to a group not offered health checks. Results were available from 14 trials, including 182,880 participants. Nine trials studied the risk of death and included 155,899 participants and 11,940 deaths. There was no effect on the risk of death, or on the risk of death due to cardiovascular diseases or cancer. We did not find an effect on the risk of illness but one trial found an increased number of people identified with high blood pressure and high cholesterol, and one trial found an increased number with chronic diseases. One trial reported the total number of new diagnoses per participant and found a 20% increase over six years compared to the control group. No trials compared the total number of new prescriptions but two out of four trials found an increased number of people using drugs for high blood pressure. Two out of four trials found that health checks made people feel somewhat healthier, but this result is not reliable. We did not find that health checks had an effect on the number of admissions to hospital, disability, worry, the number of referrals to specialists, additional visits to the physician, or absence from work, but most of these outcomes were poorly studied. None of the trials reported on the number of follow-up tests after positive screening results, or the amount of surgery used.

One reason for the apparent lack of effect may be that primary care physicians already identify and intervene when they suspect a patient to be at high risk of developing disease when they see them for other reasons. Also, those at high risk of developing disease may not attend general health checks when invited. Most of the trials were old, which makes the results less applicable to today's settings because the treatments used for conditions and risk factors have changed.

With the large number of participants and deaths included, the long follow-up periods used in the trials, and considering that death from cardiovascular diseases and cancer were not reduced, general health checks are unlikely to be beneficial.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON *[Explanation]*

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk with intervention				
Total mortality Deaths Follow-up: 4-22 years	**75 per 1000**	**74 per 1000** (71 to 77)	**RR 0.99** (0.95 to 1.03)	155,899 (9 studies)	⊕⊕⊕⊕ **high**	
Cardiovascular mortality Deaths from cardiovascular causes Follow-up: 4-22 years	**37 per 1000**	**38 per 1000** (34 to 43)	**RR 1.03** (0.91 to 1.17)	152,435 (8 studies)	⊕⊕⊕◯ **moderate**	There was substantial heterogeneity which may reflect the different outcome definitions used in the trials
Cancer mortality Cancer deaths Follow-up: 4-22 years	**21 per 1000**	**21 per 1000** (19 to 24)	**RR 1.01** (0.92 to 1.12)	139,290 (8 studies)	⊕⊕⊕⊕ **high**	

*The **assumed risk** is the median control group risk across studies. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **RR:** Risk ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

BACKGROUND

Description of the condition

General health checks are common elements of health care in some countries (Han 1997; Holland 2009). Historically, general health checks of the healthy public is a recent phenomenon. The evolution of medicine in the latter half of the 20th century has yielded a great increase in diagnostic methods and increased expectations that many diseases can be prevented or discovered before there is irreversible damage.

Description of the intervention

General health checks involve a contact between a health professional and a person that is not motivated by symptoms and where several screening tests are performed to assess general health. The purpose is to prevent future illness through earlier detection of disease or risk factors, or to provide reassurance. The terminology is confusing. Multiphasic screening, periodic health examination and preventive health checks are examples of terms used to describe the intervention. Some studies have investigated the effect of a single health check and some have examined the effect of consecutive checks, and the diagnostic tests included vary considerably. We use the broad term 'general health check', which is frequently used by lay people and in advertising.

Few of the screening tests commonly included in general health checks have been evaluated according to accepted criteria, that is in high-quality randomised trials (UK National Screening Committee 2010). Whilst the benefits and harms of treatments for conditions such as hypertension and diabetes have been extensively studied in randomised trials, screening asymptomatic people for these conditions has not (Norris 2008; Sheridan 2003). When screening for individual conditions has been studied in randomised trials, the outcome has varied. For example, screening for prostate cancer does not appear to substantially reduce disease-specific mortality but has important harms (Djulbegovic 2010), whereas testing for faecal occult blood prevents one in six colorectal cancer deaths though at the cost of a large number of invasive examinations in healthy people (Hewitson 2007).

Health checks may be offered to the general population as part of a national policy or private health insurance, or employers may offer them to their employees. They may also be purchased by the individual from commercial providers or provided by general practitioners. Health checks may be quite comprehensive and use advanced technologies, such as computed tomography or magnetic resonance imaging, although these interventions are not recommended for health checks because of unproven benefit and risk of harms (FDA 2011).

Some general health checks include a conversation with a health professional, possibly a questionnaire, and sometimes also a physical examination by a doctor. In essence these manoeuvres are screening tests, although a conversation may not be perceived as such. Lifestyle interventions are also frequently administered during a health check, for example advice on diet and smoking. This is not screening but behavioural intervention, and appears to be of varying value. For example, systematic reviews have not shown a value for multiple risk factor interventions in general populations (Ebrahim 2011). There may be a small effect of modification of dietary fat, but the ideal type of modification is not clear (Hooper 2011). However, simple advice on quitting smoking has been shown to have an effect (Stead 2008).

Importantly, primary care physicians sometimes advise health checks or selected screening tests for patients that they think might benefit from them when they see the patients for other reasons. Such clinically motivated testing is often considered an integral part of primary care practice and it is against this background that the effect of systematic health checks are measured.

How the intervention might work

General health checks are expected to reduce morbidity and mortality through earlier detection and treatment of diseases and risk factors for diseases. For example, early detection of hypertension or hypercholesterolaemia may lead to reductions in morbidity and mortality through treatment. Screening may detect precursors to disease, for example colorectal adenomas or cervical dysplasia the treatment of which may prevent cancer from developing. Also, identification of signs or symptoms of manifest disease that the person had not deemed important may be beneficial. Counselling on diet, weight and smoking may also be of value. Healthy people may feel reassured, which could decrease worry. The preventive nature of general health checks implies that most effects would be expected to have a latency of several years.

Screening healthy people can also be harmful. While we cannot be certain that screening leads to benefit, all medical interventions can lead to harm. A well-known example is overdiagnosis of latent cancers or carcinoma in situ, which might not have progressed to become symptomatic or might have regressed spontaneously (Welch 2004). Furthermore, false positive test results can lead to unnecessary invasive diagnostic tests that may cause harm; and drug treatment of people with risk factors such as high cholesterol and elevated blood glucose can have adverse effects, also in people who would not have developed manifest disease. False positive test results may cause unnecessary worry (Brewer 2007), and false negative results may lead to a false sense of security and delay medical attention when needed. Further, being labelled as having a disease, or even just as being at increased risk of getting a disease, may negatively impact healthy peoples' views of themselves (Barger 2006; Haynes 1978). It may also make it more difficult to obtain life and health insurance in some countries. Last but not least, there is a financial cost for patients and society in identifying and treating risk factors and diseases that might never have manifested themselves as illness or shortened life.

Why it is important to do this review

General health checks are mixtures of screening tests few of which have been adequately studied, and it is not clear whether they do more good than harm. A systematic review of the periodic health evaluation, which included both trials and observational studies, found mixed results on clinical outcomes, except for patient worry where a beneficial effect was seen in one trial (Boulware 2006; Boulware 2007). The definition of the intervention was narrow and relatively few trials were included. Two other reviews focused on using global coronary risk scores, which is a common component of health checks (Sheridan 2008; Sheridan 2010). One included studies in which the effect of calculating the risk score could be isolated and it did not find any studies reporting on long-term clinical events. Two out of four studies found that the intervention increased prescription of cardiovascular drugs (Sheridan 2008). Another review focused on the effect of giving global coronary risk information to adults (Sheridan 2010). The authors found that the intervention improved the participants' perception of risk and that it may increase the intent to initiate prevention, but they found no studies reporting on actual event rates. We saw a need for a broad and comprehensive review of the randomised trials, with a focus on clinically important outcomes rather than surrogate outcomes. We chose not to review observational studies because the risk of bias is too great in relation to the expected effect sizes.

OBJECTIVES

To quantify the benefits and harms of general health checks.

METHODS

Criteria for considering studies for this review

Types of studies

Randomised trials of general health checks compared with no health checks. We had no language restrictions. We included trials regardless of funding source.

Types of participants

Inclusion criteria

Adults, regardless of gender and ethnicity. The setting had to be primary care or the community. We included trials regardless of whether they were directed at the general population or a more narrow group, for example employees of a company.

Exclusion criteria

We did not include studies described as specifically targeting older people, or which only included people aged 65 years or more (see Differences between protocol and review). Studies in populations of patients or people with specific known risk factors or diseases were excluded, for example studies in people with hypertension or ischaemic heart disease.

Types of interventions

Screening for more than one disease or risk factor and in more than one organ system, whether performed only once or repeatedly. This definition excludes trials of screening for single diseases, for example prostate cancer, and trials of single screening tests which may detect more than one disease, for example spirometry.

We accepted trials which included a lifestyle intervention (for example advice on diet, smoking and exercise) in addition to screening since this is a fairly well-defined intervention that is often incorporated into health checks.

We included trials regardless of the type of healthcare provider, for example a doctor, nurse, or other health professional.

Types of outcome measures

Some trials and observational studies have investigated the effects of health checks on surrogate outcomes, for example cardiovascular risk factors, health behaviours, or cancer screening rates, and some have found positive effects, albeit generally small. However, there can be serious problems with using surrogate outcomes (Fleming 1996).

First, assessing the effect of changes in a surrogate outcome on morbidity and mortality is difficult and unreliable and requires modelling with assumptions that are difficult to test. There may be latency of effects (Ebrahim 2011; Hooper 2011) and uncertainty regarding the degree of reversibility of the risk. For example, quitting smoking reduces the risk of coronary heart disease and mortality, but slowly and probably not completely (Ben-Schlomo 1994; Cook 1986). Also, it is difficult to know to what degree changes in risk factors and behaviours are maintained in the long term. Second, the use of surrogate outcomes disregards the harmful effects of follow-up diagnostic procedures and treatments. A recent example is the drug rosiglitazone for diabetes, which reduced the surrogate outcome blood glucose but caused serious heart disease (Lehman 2010; Nissen 2010). This was not recognised in trials using surrogate outcomes only. Third, in order to measure changes in risk factors and health behaviours the participants need to attend a follow-up session or answer questionnaires. Since it is impossible to blind the intervention group, and since the intervention is often partly behavioural, biased loss to follow-up is to be expected. For example, people with adverse health behaviours might not feel inclined to confront the researchers again, which could lead to spurious improvements in surrogate outcomes in an

available case analysis or a last observation carried forward analysis. Also, the lack of blinding may cause biased reporting of health behaviours.

For these reasons, we focused on outcomes that directly reflect the beneficial and harmful effects of health checks on the health of the participants and which can be reliably ascertained with long follow-up. We chose total and disease-specific mortality as our primary outcomes because these are less likely to be biased than other outcomes, are of direct relevance to participants, and capture both beneficial and harmful effects. However, we included some outcomes that are susceptible to attrition bias and reporting bias because they are important and cannot be assessed in other ways, for example self-reported health and worry.

Primary outcomes

- All-cause mortality
- Disease-specific mortality

Secondary outcomes

- Morbidity (e.g. myocardial infarction)
- New diagnoses (total and condition-specific)
- Admission to hospital
- Disability (preferably patient-reported)
- Patient worry
- Self-reported health
- Number of referrals to specialists
- Number of non-scheduled visits to general practitioners
- Number of additional diagnostic procedures due to positive screening tests
- New medications prescribed and frequency and type of surgery
- Absence from work

Harms

The harmful effects of health checks are reflected in the above outcomes. The major harms are overdiagnosis, adverse psychological and behavioural effects, complications related to follow-up investigations, and unnecessary treatments instigated as a result of overdiagnosis. While diagnostic, preventive and therapeutic activity can lead to improved health, they are also often harmful and should be balanced by reductions in morbidity and mortality to be justified. Estimating overdiagnosis will not be possible for all diseases due to the broad scope of the review and because increased incidence is a goal for some conditions, for example diabetes, but a problem for others, for example prostate cancer. These questions are more appropriately addressed in reviews of screening for individual diseases. However, a quantification of the change in the incidence of individual conditions is still valuable even though it may represent both beneficial and harmful effects. Another possible harm is a negative effect on health behaviours, for example failure to quit smoking due to reassurance of good health. Such effects would also be captured by the chosen outcomes.

Search methods for identification of studies

Related systematic reviews were identified by searching the Database of Abstracts of Reviews of Effectiveness (DARE) and the databases listed below. Studies were identified using the following bibliographic databases, sources, and approaches.

The Cochrane Central Register of Controlled Trials (CENTRAL) (2010, Issue 11), part of the *The Cochrane Library* at www.thecochranelibrary.com.

MEDLINE on Ovid (1948 to current), MEDLINE In-Process.

EMBASE on Ovid (1947 to current).

Cumulative Index to Nursing and Allied Health Literature CINAHL on EBSCOhost (1980 to current).

Healthstar on Ovid (1966 to 2010).

Cochrane Effective Practice and Organisation of Care Review Group (EPOC) Specialised Register, Reference Manager.

ClinicalTrials.gov.

World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP).

Search strategies were developed by the EPOC Trials Search Coordinator (TSC), Michelle Fiander, in consultation with the authors. Strategies reflect an iterative development process whereby the TSC developed a series of test strategies the results of which were screened by the authors for relevance. Based on this feedback, the TSC added or deleted terms and search strategies were finalized. Two MEDLINE strategies were run: MEDLINE Strategy A (Appendix 1), run in August 2010; MEDLINE Strategy B (Appendix 2), run in November 2010. Strategy B served as the basis for translations to other databases. Neither date nor language restrictions were applied. Duplicates were removed both in the Ovid interface and in Reference Manager software. Searches were conducted in November to December 2010; all databases were searched from the database start date forward. Two methodological search filters were used to limit retrieval to the appropriate study design and interventions of interest: the Cochrane randomised controlled trial (RCT) sensitivity and precision maximizing filter (Higgins 2011); and the EPOC filter to identify non-RCT study designs. Strategies for searches in *The Cochrane Library*, EMBASE, CINAHL, and the EPOC Register are in Appendix 3. An updated search was run in July 2012 (Appendix 4).

Searching other resources

We searched the reference lists of included studies and used citation tracking (Web of Knowledge) for all articles describing eligible trials. We asked authors of the included studies if they were aware of any other published, unpublished, or ongoing studies that could meet our inclusion criteria.

Data collection and analysis

Selection of studies

Two authors (LTK and CGL or KJJ) independently assessed the potential relevance of all titles and abstracts identified through the searches and full-text copies of potentially eligible articles were assessed. Disagreements were resolved through discussion, involving the other authors (KJJ and PCG) when necessary. Two authors independently searched reference lists (LTK and KJJ) and one author used citation tracking (Web of Knowledge) on included articles.

Data extraction and management

Two authors (LTK and KJJ) independently extracted data from the included trials and entered them into a piloted data extraction form. When relevant information was missing from the reports we contacted the authors.

The following data were extracted from all included trials: study design, diagnostic tests used, total study duration, the number of participants allocated to each arm, number lost to follow-up for each outcome, baseline comparability, setting, age, country, and date of study. We extracted the number of events or rates for mortality, hospitalisation (one or more), surgery, new medications, referrals to specialists and diagnostic procedures required because of positive screening tests, and for the number of physician visits. For ordinal scale outcomes we extracted the mean value; standard deviation; and name, range, and direction of the scale. When these data formats were not available we extracted what was possible to extract, including narrative accounts if the actual numbers were missing.

Assessment of risk of bias in included studies

We used the Cochrane risk of bias tool. The domains formally assessed were: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting, and other biases. We assessed the risk of contamination of the control group under 'Other bias'. We also assessed the randomised groups for baseline comparability.

Measures of treatment effect

We preferred data from intention-to-treat analyses (ITT). When these were not available, we assessed the possible bias resulting from missing data. For mortality, we used the risk ratio. Ranking scales were treated as continuous data when possible. For all measures we used 95% confidence intervals.

Unit of analysis issues

For cluster randomised trials we preferably used effect estimates and standard errors from analyses which took the clustering into account. When such estimates were not available we disregarded the effect of clustering and investigated the impact of this in a sensitivity analysis.

Assessment of heterogeneity

Clinical and methodological differences between trials were assessed before any meta-analyses were done, and we judged whether trials could be pooled. Heterogeneity was investigated with the I^2 statistic, which describes the variation between trials in relation to the total variation.

Assessment of reporting biases

Outcome reporting bias is difficult to assess but we noted whether the outcomes that we considered important were reported. When the study design implied that data on other outcomes than the ones reported might have been investigated, we asked the authors for further data.

Data synthesis

As specified in our protocol, we used random-effects model meta-analyses. Due to the need to use published effect estimates in two trials, we used the generic inverse variance method available in RevMan. For outcomes other than mortality we summarised the results in tables and did a qualitative synthesis.

Subgroup analysis and investigation of heterogeneity

We planned the following subgroup analyses:

- only one health check versus several;
- physical examination by physician;
- interventions that included advice on lifestyle;
- age of trial;
- geographical location of trial;
- high versus low risk of bias;
- long versus short follow-up.

Sensitivity analysis

We decided to include cluster randomised trials despite anticipating that we had to ignore the clustering in some cases, and despite the greater risk of unsuccessful randomisation. To investigate the robustness of our results, we planned a sensitivity analysis excluding cluster randomised trials.

RESULTS

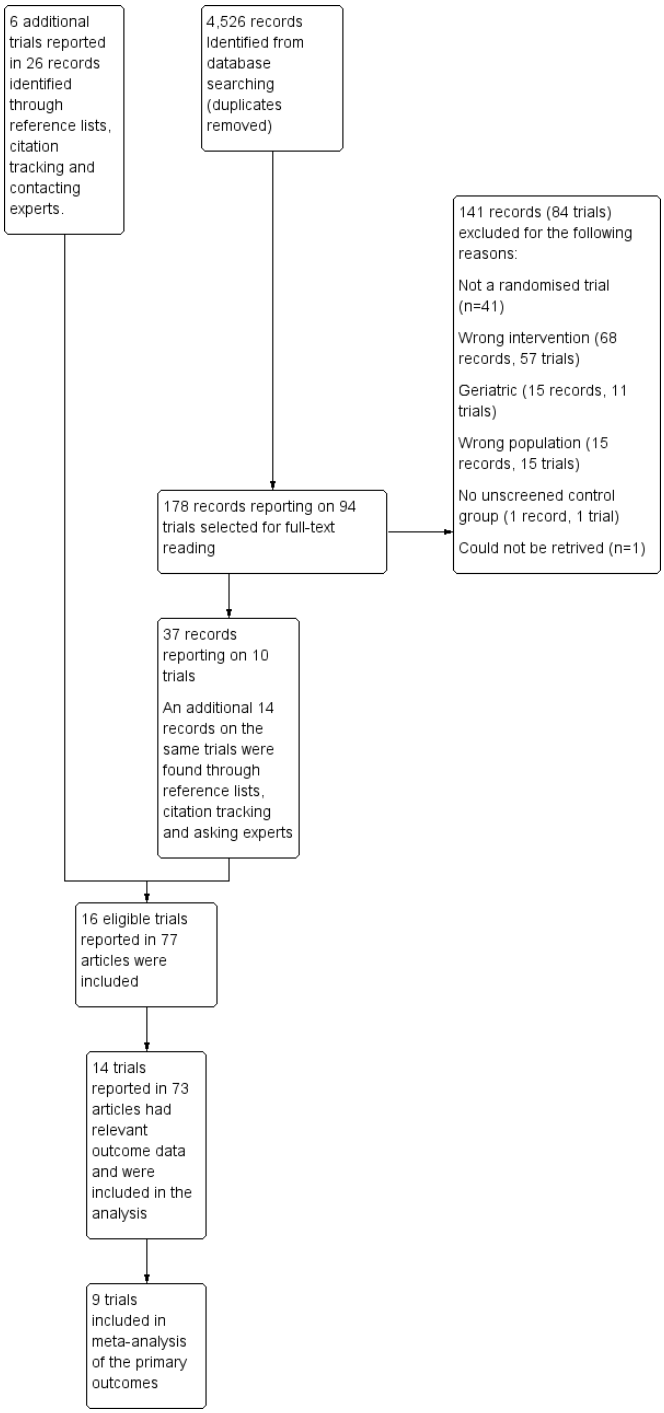
Description of studies

See: Characteristics of included studies; Characteristics of excluded studies.

Results of the search

The search yielded 4526 records after removal of duplicates. From these we selected 178 articles for full-text assessment, of which we excluded 141. Forty-one of the excluded articles did not report on a randomised trial, 68 articles (57 trials) studied a non-relevant intervention (for example reminder systems for physicians or lifestyle interventions), 15 articles (11 trials) were geriatric, 15 articles (15 trials) studied people who were selected for diseases or risk factors thus not representing a general population, one did not have an unscreened control group, and one could not be retrieved. This left 37 articles reporting on 10 trials that were eligible for inclusion. An additional six trials reported in 26 articles were identified through searching reference lists and citation tracking. We identified a further 14 articles on the included trials through searching reference lists and citation tracking. Thus 16 trials reported in 77 articles were included; but since two trials never published their results (New York 1971; Titograd 1971), 14 trials reported in 73 articles were analysed (see Figure 1).

Figure 1. Study flow diagram.



Included studies

The 14 trials included in the analyses varied in size from 533 randomised persons in the Northumberland trial (Northumberland 1969) to 57,460 in the WHO trial (WHO 1971). The total number of participants was 182,880 with 76,403 allocated to health checks and 106,477 to control. Nine trials with 155,899 participants reported a total of 11,940 deaths (Ebeltoft 1992; Göteborg 1963; Göteborg 1970; Kaiser Permanente 1965; Malmö 1969; OXCHECK 1989; South-East London 1967; Stockholm 1969; WHO 1971). The length of follow-up for mortality varied from 4 to 22 years, and it also varied within trials for different outcomes. The trials that did not report on mortality were often small (Mankato 1982; Northumberland 1969; Salt Lake City 1972), with the exception of the British Family Heart study (Family Heart 1990) which included 12,924 persons. The Inter99 trial (Inter99 1999) has not yet published results for mortality.

The setting was general practice in five trials (Family Heart 1990; Ebeltoft 1992; Northumberland 1969; OXCHECK 1989; South-East London 1967), the community in eight trials (Göteborg 1963; Göteborg 1970; Inter99 1999; Kaiser Permanente 1965; Malmö 1969; Mankato 1982; Salt Lake City 1972; Stockholm 1969), and the workplace in one trial (WHO 1971). As per our inclusion criteria, they included people that were not selected for diseases or risk factors. Four trials randomised households or couples (Family Heart 1990; OXCHECK 1989; Salt Lake City 1972; South-East London 1967), one randomised factories (WHO 1971), and nine randomised persons.

The interventions can be broadly classified into two categories: screening focused on cardiovascular risk factors with a strong lifestyle intervention component, and broad screenings using many tests (often called multiphasic screening in older publications) but often without an important lifestyle intervention component. The broad type of screening was mainly seen in trials that started in the 1960s and 1970s (Göteborg 1963; Kaiser Permanente 1965; Malmö 1969; Northumberland 1969; Salt Lake City 1972; South-East London 1967) and in the Ebeltoft trial (Ebeltoft 1992) in the 1990s. Five trials included screening for cancer. The tests used were chest radiographs (Göteborg 1963; Malmö 1969); chest radiographs and faecal occult blood testing (South-East London 1967); chest radiographs, mammography and cervical smears (Salt Lake City 1972); and chest radiographs, sigmoidoscopy, mammography and pelvic examinations (Kaiser Permanente 1965). See Table 1 for an overview of the interventions used.

The uptake in the first screening round ranged between 50% (Mankato 1982) and 90% (Ebeltoft 1992) with a median of 82%. The Kaiser Permanente trial did not have screening rounds but used continuous urging of the intervention group by written invitations and phone calls to utilise a pre-paid health check.

Here we present a description of the included trials. The references to trials are labelled with the year of the start of the trial. For additional details the reader is referred to the Characteristics of included studies section.

The Göteborg 1963 trial

(Göteborg 1963)

This included all men born in 1913 and living in Göteborg, Sweden in 1962. Randomisation was done using date of birth, in a 1:2 ratio, resulting in groups sizes of 1010 (intervention) and 1956 (control). The allocation sequence was predictable but since all eligible persons were included and allocated before any contact was made the risk of selection bias was low. The intervention group was invited for three rounds of screening (1963, 1967, and 1973) and the control group was not contacted. All participants were followed for mortality over 15 years.

The first screening was performed by staff at a local hospital and included an interview about cardiovascular symptoms and chronic bronchitis, a questionnaire on social data, smoking, personal and family history, a questionnaire on cardiovascular symptoms, weight, height, skinfold thickness, blood pressure, electrocardiography, urinalysis (protein, glucose, osmolality), blood samples (cholesterol, triglycerides, fasting blood sugar, haematocrit, sedimentation rate, creatinine, serum protein electrophoresis, sodium, potassium, chlorides, blood group), chest x-ray, measurement of heart volume, general physical examination, and an examination by an ophthalmologist. Half of the intervention group had a psychiatric interview and the other half were given a psychiatric questionnaire. At the second screening, in 1967, the examination also included a work test at maximum load. The 1973 examination was unclearly described but at least included height, weight, skinfold thickness, and questions about morbidity, well-being, and utilisation of medical care.

The Kaiser-Permanente trial

(Kaiser Permanente 1965)

This trial investigated the effects of broad (multiphasic) screening with 16 years of follow-up. In 1964, a sample of members of the Kaiser-Permanente Health Plan in San Francisco and Oakland aged 35 to 54 years were divided into an intervention group (n = 5156) and a control group (n = 5557) using an allocation algorithm based on membership numbers, which was likely to yield comparable groups. This was done before any contact was made with the trial participants and the risk of selection bias was low. The intervention began on 1 January 1965 and participants alive at that date were included in the analyses, giving analysed group

sizes of 5138 (intervention) and 5536 (control). The control group was larger than the intervention group due to identity mix-ups and exclusion from the intervention group of people who had moved too far away from the study centre. The excluded participants were included in an analysis of mortality after 11 years, without changes to the results (Dales 1979).

Participants in the intervention group were urged annually, by telephone and letter, to have a multiphasic screening examination that was available to members of the Kaiser health plan. The intervention continued for 16 years. The control group participants received questionnaires about their health but were not urged to be screened and were not informed about the experiment. However, as part of their health plan they were able to request the same multiphasic screening examination as the intervention group and did so to a large extent. After 16 years of intervention the mean number of health checks was 6.8 in the intervention group and 2.8 in the control group. In the intervention group 15.7% of the participants had never had a health check, compared to 36.2% in the control group. Thus, the contrast between the groups was not substantial.

The screening intervention was broad and included a medical history, clinical examination, chest x-ray, laboratory tests, mammography, and recommendations for gynaecologic examinations and sigmoidoscopy for people over 40 years, but no explicit lifestyle component (see full list at Characteristics of included studies). Additional testing was done according to computerised advice rules and the judgement of the clinicians in charge of the screening. There was a follow-up visit with a physician or nurse for interpretation of the results.

The outcomes relevant to this review were total mortality, cause-specific mortality, morbidity, hospitalisation, physician visits, prescriptions, disability, and number of new diagnoses. A weakness of this trial is that participants leaving the health plan were considered lost to follow-up for all outcomes except mortality, resulting in more than 35% having been lost after 16 years. Only people leaving California were lost to follow-up for mortality and the authors assessed this to be 8% to 18% of deaths (Friedman 1986).

The South-East London Screening Study

(South-East London 1967)

This trial began in 1967 and was set in two general practices in London, England. All registered people aged 40 to 64 years were included and they were randomised all at once before any contact was made with them. The randomisation was unclearly described but involved alternate allocation of couples from an alphabetically arranged list. There was also some form of matching, but this was not described in detail.

The trial is reported in several papers by different sets of authors and the sizes of the compared groups after randomisation differ between publications. An early paper stated that the group sizes were 3460 (intervention) and 3337 (control) (Trevelyan 1973), but in the main paper they were reported as 3876 (intervention) and 3353 (control) (South-East London Study Group 1977). Furthermore, only 3292 (intervention) and 3132 (control) were included in the mortality analyses. Another paper explained that 579 spouses of eligible participants who were outside the defined age range were originally included in the study and invited for screening (D'Souza 1976) but they appear to have been excluded at the time of analysis, possibly to avoid bias from expanding the intervention group with people at ages less likely to benefit from screening. However, this does not fully explain the discrepancies.

The intervention group was invited for two rounds of multiphasic screening, done independently from the participants' own general practitioners. The screening included a physical examination, medical history, a questionnaire on symptoms, height and weight measurements, vision and hearing tests, chest x-ray, spirometry, electrocardiogram (ECG), blood pressure, blood chemistry and faecal occult blood testing. The control group was not invited, and the authors wrote that the control group did not show any interest in screening and that none were screened (Trevelyan 1973).

After five years, both the intervention group and the control group were screened using the same tests, except for the questionnaire and faecal occult blood testing. Follow-up for mortality and usage of health services continued for a further four years. One later report described the five-year survey as being "non-prescriptive (in the sense that no therapeutic activity was expected to result from it)" but did not describe how this was ensured (Stone 1981). Screening the control group after five years biased the nine-year results towards no effect.

The Northumberland trial

(Northumberland 1969)

In 1969, all men aged 50 to 59 in seven general practices in the UK were included and randomised by date of birth into three groups. People with serious illnesses were excluded. One group was screened with a full physical examination (n = 242), although the contents of this were not described. A control group was not invited for screening (n = 291). A third group was sent a questionnaire about health issues and were invited for examination if certain symptoms were present, for example persistent cough or haematuria (n = 275). Follow-up was done after 18 months and was based on patient records. The outcome included in this review is physician visits. Other relevant outcomes were reported but in a way we could not use.

The Malmö trial

(Malmö 1969)

The study population was defined as all men born in 1914 and living in Malmö, Sweden in early 1969. All men born in even-numbered months were invited to screening (n = 809) and all men born in uneven-numbered months were not (n = 804). This

method of allocation sequence generation is obsolete, but since all eligible participants were included and randomised before any contact was made the risk of selection bias was low. The screening intervention was broad and included blood pressure, blood tests (cholesterol, triglycerides, haematocrit), urinalysis (glucose, albumin), height, weight, electrocardiography, spirometry, nitrogen washout for measuring pulmonary dead space, sputum cytology, chest x-ray, venous occlusion plethysmography (arterial blood flow), an interview, a questionnaire, and a physical examination. Of the 178 participants classified as heavy smokers in the intervention group, a random sample of 51 were offered a group counselling intervention to quit. Participants with hypertension or impaired lung function were followed up and treated at a hospital rather than by their general practitioner. This may have biased the results in favour of the intervention group.

The participants' primary care physicians were not involved with the study and the control group was not contacted. Information on mortality and hospitalisations was gathered from public registers after five years, with 1% loss to follow-up. Cause of death was ascertained blinded to randomised group by one person using autopsy reports and hospital records.

The Stockholm trial

(Stockholm 1969)

This trial aimed to assess the effect of one general health examination on long-term mortality. The participants were men and women aged 18 to 65 years living around Stockholm. A complex stratified randomisation scheme was used which purposely introduced baseline imbalances (see Characteristics of included studies for description). The authors used Cox regression, in which they controlled for the baseline imbalances introduced by the randomisation scheme as well as sex and age. We obtained mortality data from the authors and supplemented this analysis with a fixed-effect model meta-analysis combining the effects in each of the 12 strata, and got results nearly identical to those originally reported. The numbers randomised were 3064 (screening) and 29,122 (control). Participants in the intervention group were invited to one screening while the control group was not. Both groups were sent a questionnaire before randomisation. The screening consisted of blood pressure; blood tests (not specified); ECG; exercise tests; a physical examination; social, psychiatric and medical interviews; eye and dental examinations. Participants with an identified need for specialist services were directly referred, whereas participants were instructed to contact their primary care physician for other identified issues. Simple services like reassurance and prescription of simple medications (not specified) were provided by the researchers.

Participants were followed up for mortality in registers over a period of 22 years. The outcomes studied were total mortality, cardiovascular mortality, cancer mortality, and mortality from accidents and intoxications. Data on hospitalisation were collected but not published.

The Göteborg 1970 trial

(Göteborg 1970)

The aims of this trial were to reduce cardiovascular risk factors and to measure the effect on morbidity and mortality. The trial started in 1970 and included all men in Göteborg, Sweden, who were born in 1915 to 1922 and 1924 t0 1925. These were randomised to an intervention group (n = 10,004) and two control groups (n = 10,011 and 10,007). The intervention group was invited to screening at baseline and after four years. The screening was focused on cardiovascular risk factors and included blood pressure, total serum cholesterol, height, weight, ECG, a questionnaire on family history of cardiovascular disease and risk factors, and an interview. Elevated risk factors were treated with lifestyle advice and drugs according to simple decision rules based on cut-off values for individual risk factors (see Characteristics of included studies). Thus, the standard of follow-up and care was likely to be different compared to the control groups.

In one of the control groups, a random 2% were invited to screening at baseline and an 11% sample after four years. The purpose of this was to compare changes in risk factors. The other control group was never contacted. We chose to pool both control groups for our meta-analysis.

The participants were followed in registers for mortality and morbidity until the end of 1983, with a mean follow-up time of 11.8 years. For our analysis of cardiovascular mortality we combined fatal coronary heart disease and fatal stroke.

The WHO trial

(WHO 1971)

Conducted in five countries (UK, Belgium, Poland, Spain, Italy) this trial had aims similar to the multifactor primary prevention trial in Göteborg (Göteborg 1970), but used a different design and was set in the workplace. It started as one trial in the UK but was soon expanded to include other countries using similar methods. Results from Spain were never included in the analysis of events. This decision was made before results were available to the investigators and was due to the fact the Spanish part of the trial was started later than the others. Factories were recruited for participation, matched in pairs, and these pairs were then randomised to either intervention or control. The method of randomisation was not described but allocation was concealed and demographic and prognostic variables were balanced at baseline. The number of factories were 80, providing 40 pairs. Only the male employees were included. The sizes of the groups as randomised were 30,489 (intervention) and 30,392 (control). To assess baseline balance and study the effect on risk factors, a 10% random sample of the control group was invited to screening. These participants were not included in the analysis of events and the numbers analysed were thus 30,489 (intervention) and 26,971 (control).

The screening included blood pressure, total serum cholesterol, weight and a questionnaire on smoking, physical activity and symptoms of coronary heart disease. The men at highest risk (10% to 20%, which varied between centres) were called for an interview with a physician and given lifestyle advice and medical treatment of risk factors. In addition, the intervention factories had a campaign of health education aimed at reducing risk factors.
Annually, a random 5% of the intervention group were invited to screening in order to assess changes in risk factors. At the end of trial, all in both the intervention and control factories were invited to screening. Follow-up was at between five and six years (differed between centres). Mortality was assessed for all, but morbidity was only assessed for people still employed to avoid detection bias.
No results for cardiovascular mortality were reported (including stroke and other causes) so instead we used the reported results for coronary heart disease mortality in our meta-analysis. For total and coronary heart disease mortality, we used reported effect estimates from an analysis which took clustering into account. For cancer mortality no such estimate was reported so we disregarded the clustering.

The automated multiphasic health testing (AMHT) study

(New York 1971)
This trial was set up in the Health Insurance Plan of Greater New York (HIP) with the aim of investigating whether health checks could reduce the gap in health status and health behaviour between poor and non-poor persons. The study included families with at least one person aged 12 to 74 years. The exact size of the sample was unclear, but about 7,000 non-poor persons and somewhat fewer poor persons were mentioned as being the intervention group. The control group was said to be 20% of this size. The intervention included blood pressure, height, weight, skinfold thickness, ECG, pulse rate, chest x-ray, audiometry, dental survey, visual acuity, tonometry, spirometry, glucose challenge, blood tests (cholesterol, total protein, albumin, calcium, total bilirubin, urea nitrogen, uric acid, haemoglobin, white blood cell count, syphilis test), urine tests (pH, protein, glucose, blood, acetone), sickle cell trait, urine culture (women only), instruction in breast self-examination, mammography (women aged 40+ years), and Pap smear. The trial was designed to investigate disability and absence from work. Mortality data were also to be gathered. The AMHT programme was discontinued after the first screening round but follow-up was planned to continue. We have not found reports of the results.

Titograd 1971

(Titograd 1971)
This study was set up in Titograd, former Yugoslavia, in collaboration between Yugoslavian and American researchers. A random sample was drawn from the population aged 30 to 49 years, and randomly divided into an intervention (n = 6577) and a control group (n = 6573). A 20% random subsample of both groups were interviewed at baseline. The intervention group was invited for screening at baseline and at two-year intervals. Follow-up of positive test results and treatment of identified conditions was done according to specified regimens. The intervention included blood pressure, cholesterol, height and weight, ECG, spirometry, glucose tolerance, chest x-ray, red and white blood cell counts, blood sedimentation rate, blood urea nitrogen, cervical smear, visual acuity and fundus examination, Wassermann reaction (syphilis), urinalysis (not specified), and a latex fixation test (unclear which antibodies were tested for). The control group was not invited for screening. Analysis of morbidity, disability, mortality, and medical care utilisation was planned after six years, and if no effect was observed the trial would be continued for a further four years. We have not found reports of the results of this trial.

Salt Lake City 1972

(Salt Lake City 1972)
This trial was conducted in 1972 to 1973 and studied the effects of one multiphasic screening examination on disability and utilisation of health care. The study sample consisted of random samples from three groups in Salt Lake City, USA: 200 low-income families with a pre-paid healthcare programme, 200 low-income families with no pre-paid healthcare programme, and 166 middle-income families who had volunteered for a study of health care. The participants were randomised by family to the intervention or control in a 3:2 ratio. The number of families in each group were not reported but the number of participants in the intervention group was 642 and in the control group it was 454. All were interviewed at baseline for information about health status, number of disability days caused by illness, patterns of healthcare utilisation, health knowledge, attitudes toward the healthcare system, and hypochondriasis. The intervention group was offered one multiphasic screening consisting of a very broad array of tests including five different x-ray studies, mammography, cervical cytology, spirometry, ECG, blood pressure, tonometry, audiometry, visual acuity, venereal disease survey, 12 blood tests and six urine tests. The control group was not offered screening. All outcomes were ascertained through a second interview one year later. Those who changed economic status, did not attend for screening, did not consult their physician about screening results, or who did not participate in the one-year follow-up were excluded. This resulted in 49% of the intervention group and 82% of the control group participants being included in analyses. The relevant outcomes studied were hospitalisation, physician visits, and disability.

The Minnesota Heart Health Program

(Mankato 1982)
This trial randomised addresses representing the entire community to intervention (n = 1156) or control (n = 1167). In the inter-

vention group, the whole household was invited for screening but only one person from each household aged 25 to 74 years, selected randomly, was followed up and included in the analyses. After one year, the participants in the intervention group who attended the initial screening were re-invited for a follow-up screening and the control group was invited for their first time. The screening included blood pressure, cholesterol, height, weight, expired air carbon monoxide, and leisure time physical activity. Participants received health education at each measurement station. Each family spent 20 minutes with a health educator to review the results and receive further advice. Participants were referred to their regular physician for treatment when necessary. Only persons who participated in the screening were included in analyses, which resulted in missing outcome data for more than 50%. The trial was conducted during a population-based programme to educate about risk factors for coronary heart disease. The relevant outcome reported was use of antihypertensive medication.

OXCHECK

(OXCHECK 1989)

Starting in 1989, this trial included 11,090 persons aged 35 to 64 years who were registered with one of five general practices in the UK and who returned an initial questionnaire. Participants were randomised by household into four groups before contact was made. The first group had health checks at year one and year four, the second group at years two and four, the third group at years three and four, and the last group only at year four. Participants in the first two groups were further randomised to annual re-checks or no re-checks. The first three groups constituted the intervention groups with differing lengths of follow-up and 'dose' of the intervention, and the last group was a control group.

The health checks were performed by specially trained nurses and included measurement of blood pressure, total cholesterol, height and weight; and questionnaires on personal and family medical history, lifestyle, diet, exercise rates, and alcohol consumption. Participants were given individualised counselling on reduction of risk factors and offered follow-up visits with the nurse, as needed. The groups were compared for changes in risk factors and health behaviours. The trial was designed for studying changes in risk factors and not mortality, but we obtained mortality data from the authors.

The British Family Heart Study

(Family Heart 1990)

Thirteen matched pairs of general practices were randomised to either intervention or control (external control group). In the intervention practices, men aged 40 to 59 years were randomised to either intervention or control (internal control group) and their partners were included. The number of people randomised was not clear but the numbers analysed were 3436 (intervention), 3576 (internal control), and 5912 (external control). The intervention group was invited for screening and lifestyle intervention at baseline. The screening included blood pressure, cholesterol, blood glucose, body mass index (BMI), waist/hip ratio, smoking status, and medical history. A coronary risk score (Dundee) was communicated to each participant and the frequency of follow-up examinations was determined by this score together with other individual risk factors. Lifestyle advice was given and personally negotiated lifestyle changes were recorded. After one year both the intervention and control groups were invited for follow-up screening. Only those participants who attended their first health check were included in the analyses, that is at baseline for the intervention group and after one year for the control group. Relevant outcomes were self-reported prevalence of hypertension, hypercholesterolaemia, diabetes, and coronary heart disease; self-reported health; and use of selected medications.

The Ebeltoft trial

(Ebeltoft 1992)

This trial began in 1992 and studied the effects of broad health checks and lifestyle interventions in general practice. The initial population was all 3464 residents aged 30 to 49 years living in the Ebeltoft municipality, Denmark, in 1991. A random sample of 2000 participants (invitation failed for administrative reasons in an additional 30 persons) were mailed an invitation and a questionnaire. Persons who returned the questionnaire (n = 1507) were individually randomised into two intervention groups and a control group. The first intervention group (n = 502) was offered a health check at baseline and after two years, with a written response about the results and recommendations for follow-up. The second intervention group (n = 504) was offered the same plus annual 45-minute lifestyle discussions with the general practitioner. The third group (n = 501) had usual care.

The health checks included an assessment of cardiovascular risk (blood pressure, cholesterol, smoking, family history, sex, age, body mass index), ECG, liver enzymes, creatinine, blood glucose, HIV status (optional), spirometry, urinary dipstick for albumin and blood, BMI, CO concentration in expired air, physical endurance, and vision and hearing tests.

All three groups were invited for screening after five years with 25% to 31% loss to follow-up. The main outcomes were cardiovascular risk factors but self-reported health and worry were also measured. Data on mortality, physician visits, referrals, and hospitalisation were collected through registers, and two comparisons were made: 1) between the three intervention groups, and 2) between the 2000 randomly invited to participate in the trial and the 1434 not invited. The first comparison may have had diminished external validity due to self-selection in returning the questionnaire, and the questionnaire itself may have contaminated the control group. Furthermore, hospitalisations and referrals were compared after eight years of follow-up even though the control group

was screened after five years. The second comparison did not have these problems but had low contrast since only about half of the participants invited to participate in the trial were eventually invited to health checks. We chose the eight-year mortality results from the second comparison for our meta-analysis, and for the qualitative analyses we present results from both comparisons.

Inter99

(Inter99 1999)

This recently concluded trial investigated the effects of health checks and two kinds of lifestyle interventions. All 61,301 persons aged 30, 35, 40, 45, 50, 55 and 60 years and living in 11 municipalities in the south-western part of Copenhagen County on 2 December 1998 were included. A random sample of 13,016 persons were invited to screening and the remaining 48,285 constituted the control group. The intervention groups and a random sample of 5264 persons in the control group had questionnaires at baseline and after one, three, and five years of follow-up. All participants were followed up through central registers.

The screening included blood pressure, height and weight, waist and hip circumference and ratio, fasting blood samples (high density lipoprotein (HDL), triglyceride, total cholesterol, very low density lipoprotein (VLDL), low density lipoprotein (LDL)), glucose tolerance test, spirometry, and ECG. Absolute 10-year risk of ischaemic heart disease was assessed using the PRECARD computer program and individual counselling on risk factors and adverse health behaviours was given.

High-risk participants were offered four health checks (at baseline and years one, three, and five), low-risk participants were offered two (at baseline and year five). The intervention group was further randomised into high or low intensity treatment of risk factors. The high intensity group participants, who had a high risk of ischaemic heart disease, were offered six sessions of group counselling during a four to six month period and were re-invited for a similar intervention after one and three years. Participants in the low intensity group were not offered group counselling but were referred to their general practitioner. The control group was not contacted.

Mortality data are not published yet. The results on self-reported health were based on a comparison between the intervention group and the 11% subsample of the control group that had questionnaires. Those who returned the baseline questionnaire were included in an analysis of repeated measurements of self-reported health, giving sample sizes of 6784 (intervention) and 3321 (control group).

Risk of bias in included studies

Risk of bias varied considerably between trials, but in general there were problems in most trials. The two major issues were lack of blinding and missing outcome data, whereas selection bias was unlikely in most trials.

For mortality, seven out of nine trials reporting on this outcome had low risk of selection bias, and eight of nine were at low risk of attrition bias for that particular outcome. Kaiser Permanente (Kaiser Permanente 1965), the South-East London Screening Study (South-East London 1967), and the Ebeltoft Health Promotion Study (Ebeltoft 1992) were biased towards no effect because of contamination and low contrast between groups, and in the OXCHECK (OXCHECK 1989) we prioritised power over contrast in the merging of groups. Four trials were biased by design in favour of the screening group (Göteborg 1963; Göteborg 1970; Malmö 1969; WHO 1971). One of the most reliable trials (Inter99 1999) has not yet published mortality results.

For other outcomes, detection bias, biased reporting of subjective outcomes, and biased drop-out were major concerns in many of the trials. In particular, the patient-reported outcomes should be viewed with caution due to the lack of blinding. Readers are referred to the risk of bias figures for an overview (Figure 2; Figure 3).

Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

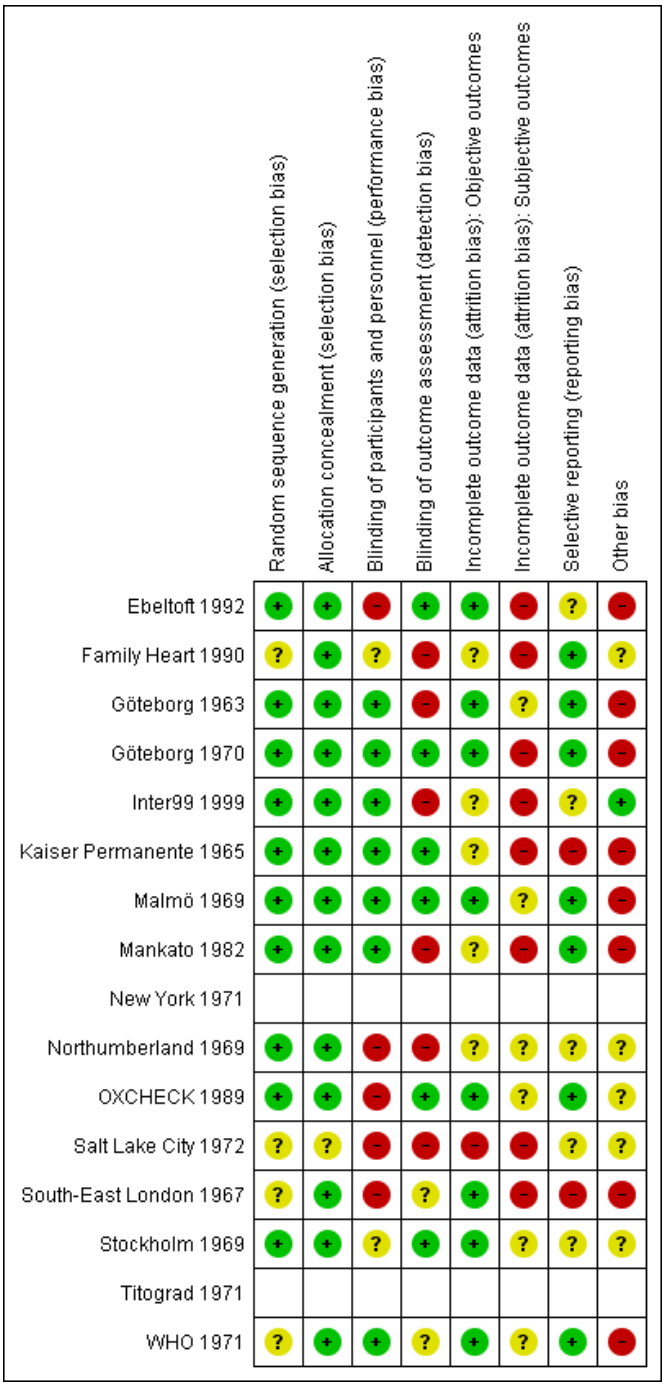
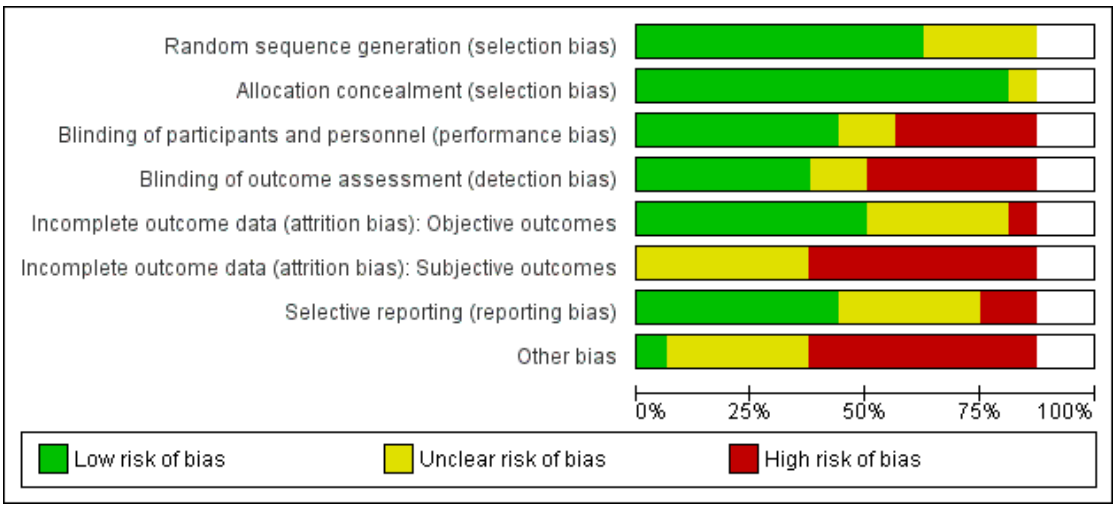


Figure 3.



Allocation

Six trials used a genuinely random method for generating the randomisation sequence (Ebeltoft 1992; Göteborg 1970; Inter99 1999; Mankato 1982; OXCHECK 1989; Stockholm 1969). In four trials, we could not determine how the sequence was generated (Family Heart 1990; Salt Lake City 1972; South-East London 1967; WHO 1971). In four trials, the sequence was predictable (for example date of birth) (Göteborg 1963; Kaiser Permanente 1965; Malmö 1969; Northumberland 1969) but these trials used designs where participants were included through lists or registers and allocated before any contact was made and we judged the risk of selection bias to be low.

We judged allocation to be adequately concealed in 13 trials (Ebeltoft 1992; Family Heart 1990; Göteborg 1963; Göteborg 1970; Inter99 1999; Kaiser Permanente 1965; Malmö 1969; Mankato 1982; Northumberland 1969; OXCHECK 1989; South-East London 1967; Stockholm 1969; WHO 1971), reflecting the use of a pre-randomised design. It was unclear in one trial (Salt Lake City 1972).

We thus judged 10 trials as likely to be free from selection bias (Ebeltoft 1992; Göteborg 1963; Göteborg 1970; Inter99 1999; Kaiser Permanente 1965; Malmö 1969; Mankato 1982; Northumberland 1969; OXCHECK 1989; Stockholm 1969). In four trials, we could not rule out selection bias. In the WHO trial (WHO 1971), the Salt Lake City trial (Salt Lake City 1972), and the British Family Heart Study (Family Heart 1990) there was no description of the sequence generation. In the South-East London Screening Study (South-East London 1967) the randomisation included use of a matching procedure which was unclearly described, and the sizes of the groups varied between publications.

Blinding

True blinding was not possible for the intervention group but could be achieved for the control group and the participants' primary care physicians by not informing them about the trial, and by gathering outcome data through registers. This may not be unethical because the participants were not patients in need of treatment, and the control group suffered no harm by being studied in this way. One trial attempted to create some degree of blinding by simply urging people to have a health check, which they were already entitled to by their health plan membership (Kaiser Permanente 1965).

Performance bias

Performance bias in this context meant differences in medical attention and preventive and screening activities resulting from knowledge of allocation.

In seven trials, the risk of performance bias was low (Göteborg 1963; Göteborg 1970; Inter99 1999; Kaiser Permanente 1965; Malmö 1969; Mankato 1982; WHO 1971), in two trials it was unclear (Family Heart 1990; Stockholm 1969), and in five trials the

risk was high (Ebeltoft 1992; Northumberland 1969; OXCHECK 1989; Salt Lake City 1972; South-East London 1967) because the primary care physicians clearly had knowledge of the status of their patients. For example, in one trial primary care physicians had lifestyle conversations with a subset of their own patients (Ebeltoft 1992), and in one trial there was a sticker on the medical records indicating the allocation (OXCHECK 1989). We expect the effects of these biases to be small due to the fact that these were predominantly healthy people with relatively few health issues requiring care.

Detection bias

We present a single assessment of the risk of detection bias for each trial, although there were exceptions for some outcomes in some trials. The reader is referred to the Characteristics of included studies section for detailed assessments.

Six trials had a low risk for most outcomes (Ebeltoft 1992; Göteborg 1970; Kaiser Permanente 1965; Malmö 1969; OXCHECK 1989; Stockholm 1969), two trials had unclear risk (South-East London 1967; WHO 1971), and six trials had a high risk (Family Heart 1990; Göteborg 1963; Inter99 1999; Mankato 1982; Northumberland 1969; Salt Lake City 1972).

Of the three trials that adjudicated the cause of death given on death certificates, one did this blinded (Malmö 1969), one unblinded (Göteborg 1963), and in one it was unclear (WHO 1971). The other six trials reporting on mortality used public registers or death certificates without re-classification (Ebeltoft 1992; Göteborg 1970; Kaiser Permanente 1965; OXCHECK 1989; South-East London 1967; Stockholm 1969). The Inter99 trial (Inter99 1999) has not yet published mortality results or details about cause of death ascertainment.

We considered answers to questionnaires to be at high risk of bias due to the lack of blinding.

Incomplete outcome data

Objective outcomes

For objective outcomes (for example mortality, physician visits) we judged the risk of attrition bias to be low in eight trials (Ebeltoft 1992; Göteborg 1963; Göteborg 1970; Malmö 1969; OXCHECK 1989; South-East London 1967; Stockholm 1969; WHO 1971), unclear in five trials (Family Heart 1990; Inter99 1999; Kaiser Permanente 1965; Mankato 1982; Northumberland 1969), and high in one trial. The Salt Lake City trial (Salt Lake City 1972) excluded participants who changed economic status, did not attend for screening, did not consult their physician about screening results, or did not participate in the one-year follow-up. This resulted in only 49% of the intervention group and 82% of the control group participants being included in the analyses. In the Kaiser Permanente trial, the authors considered participants as lost to follow-up when they left the Kaiser health plan. This resulted in the loss of more than one third of participants for most outcomes. For mortality, only people leaving California were lost. Registers were used and the authors estimated the loss to be 8% to 18% over the 16-year study period (Friedman 1986). Other trialists had access to mortality registers with much fewer losses (Ebeltoft 1992; Göteborg 1963; Göteborg 1970; Malmö 1969; OXCHECK 1989; South-East London 1967; Stockholm 1969; WHO 1971). In the WHO trial (WHO 1971), cancer mortality was not reported from the Belgian part of the trial. The reason given for this was that all non-coronary deaths were only categorised as such, without detailing the cause of death, as per the trial's protocol. The risk of bias due to this was unclear.

Subjective outcomes

In unblinded trials, attrition bias (bias due to incomplete outcome data in those lost to follow-up) is a threat to any outcome which is dependent on the active participation of participants for follow-up, for example answering a questionnaire, even when numbers lost to follow-up are similar in the groups. None of the trials were at low risk of attrition bias, six trials did not report subjective outcomes (Göteborg 1963; Malmö 1969; Northumberland 1969; OXCHECK 1989; Stockholm 1969; WHO 1971) and the risk was high in all other trials (Ebeltoft 1992; Family Heart 1990; Göteborg 1970; Inter99 1999; Kaiser Permanente 1965; Mankato 1982; Salt Lake City 1972; South-East London 1967).

Five trials investigated the possible effects of the missing data. In the Inter99 trial, the authors investigated the effects of non-response with logistic regression on serial measurements of self-reported health. They found that extreme values of self-reported health were associated with non-response but judged it unlikely to have seriously biased the results (Pisinger 2009). The British Family Heart Study (Family Heart 1990) used imputation with the last observation carried forward in the analysis of self-reported health and found no important differences. In another analysis they found twice as many smokers among non-attenders as among attenders. The Minnesota Heart Health Program trial (Mankato 1982) and the OXCHECK (OXCHECK 1989) trial found similar evidence of bias in relation to smoking but no large differences for other variables. In the Ebeltoft trial (Ebeltoft 1992), the authors reported in a letter that there were no differences in sex, age, baseline smoking, and baseline BMI between non-attenders in the intervention and control groups, but did not present the data (Engberg 2002c). Important differences might not be statistically significant when the numbers are small.

None of the trials used optimal imputation techniques (for example multiple imputation). Last observation carried forward may give biased results, and the direction of the bias is unpredictable. Also, there might be differences in unmeasured factors, such as motivation and ability to change lifestyle, and we advice caution in interpreting these outcomes.

Selective reporting

We found seven trials to be at low risk of reporting bias (Family Heart 1990; Göteborg 1963; Göteborg 1970; Malmö 1969; Mankato 1982; OXCHECK 1989; WHO 1971), in five trials the risk was unclear (Ebeltoft 1992; Inter99 1999; Northumberland 1969; Salt Lake City 1972; Stockholm 1969) and in two trials the risk of reporting bias was high. In the Kaiser Permanente trial (Kaiser Permanente 1965), data on surgery, prescriptions, and reasons for hospitalisation were collected but not published. Also, results on new diagnoses were collected and reported in early publications but not for the planned study period. In the South-East London Screening Study (South-East London 1967), data on referrals, prescriptions, and investigations carried out were collected but not reported.

Other potential sources of bias

Four trials had a design that could favour the screening group (Göteborg 1963; Göteborg 1970; Malmö 1969; WHO 1971). In these trials, conditions identified at screening were treated and followed at a special clinic or by the researchers whereas the control group used their regular physicians.

Screening of the control group (contamination) would dilute both the beneficial and the harmful effects of the intervention. The number of participants in the control group having health checks was only assessed in two trials. In the Kaiser Permanente trial (Kaiser Permanente 1965), after 16 years, the mean number of health checks in the control group was 2.8 compared with 6.8 in the screening group. Only 36.2% of the control group had not had a health check compared to 15.7% of the screening group. However this result cannot be generalised to the other trials, or other populations, mainly because the participants were all members of the same health plan with access to the same high-profiled multiphasic health screening. Also, screening has long been more popular in the US than in, for example, Europe. In the South-East London Screening Study (South-East London 1967) there was very little interest in screening among the participants in the control group, and none were screened for the first five years (Trevelyan 1973). However, the control group was offered screening after five years, which biased the nine-year results towards no effect.

The British Family Heart Study (Family Heart 1990) used both an internal and an external control group in order to investigate contamination. They found similar results when comparing with either control group indicating that contamination was not a big problem. In the Ebeltoft Health Promotion Study (Ebeltoft 1992), which was set in a small town, the authors noted that the trial appeared to have a large positive influence on the health behaviours of the control group (Lauritzen 2012). Also, the control group was offered screening after five years while some data were collected for eight years. The Mankato trial (Mankato 1982) was conducted during a health promotion campaign, which may have diminished the effect of the intervention.

In summary, we found six trials with a low risk of contamination (Göteborg 1963; Göteborg 1970; Inter99 1999; Malmö 1969; Stockholm 1969; WHO 1971), four trials in which it was unclear (Family Heart 1990; Northumberland 1969; OXCHECK 1989; Salt Lake City 1972), and four trials with a high risk of contamination (Ebeltoft 1992; Kaiser Permanente 1965; Mankato 1982; South-East London 1967). For the OXCHECK trial, we chose to combine all three intervention groups to achieve more power, accepting a loss of contrast. However, the results were similar when analysing the results for maximum contrast, that is only comparing those screened in year one with those in year four.

Two trials randomised only people who had returned an initial questionnaire on health and lifestyle (Ebeltoft 1992; OXCHECK 1989). This limited the external validity because of self-selection of people with an interest in health and lifestyle (Pill 1988; Waller 1990).

Effects of interventions

See: **Summary of findings for the main comparison** General health checks for preventing morbidity and mortality from disease

Total mortality

Nine trials reported on total mortality. Seven had a low risk of selection bias (Ebeltoft 1992; Göteborg 1963; Göteborg 1970; Kaiser Permanente 1965; Malmö 1969; OXCHECK 1989; Stockholm 1969) and two had an unclear risk (South-East London 1967; WHO 1971). The length of follow-up was four (OXCHECK 1989), five (Malmö 1969), between five and six (WHO 1971), eight (Ebeltoft 1992), nine (South-East London 1967), 11.8 (Göteborg 1970), 15 (Göteborg 1963), 16 (Kaiser Permanente 1965), and 22 years (Stockholm 1969). In total, the meta-analysis included 155,899 persons and 11,940 deaths. The median event rate in the control group was 7% and the range was 2% to 16%. We did not find an effect of general health checks on total mortality in the pooled analysis, risk ratio (RR) of 0.99 (95% CI 0.95 to 1.03). There was no heterogeneity (I^2 = 0%). Subgroup and sensitivity analyses did not alter the results.

Disease-specific mortality

For cardiovascular mortality (152,435 persons, 4567 deaths), the pooled point estimate was 1.03 (95% CI 0.91 to 1.17) but with large heterogeneity (I^2 = 64%). One possible explanation for the heterogeneity was the different definitions of the outcome among trials. For example, the WHO trial only reported mortality from coronary heart disease (WHO 1971) and the South-East London Screening Study grouped mortality from stroke with mortality from diseases in the central nerveous system, which meant that we could not include it (South-East London 1967). Another possible reason was unrecognised bias in the outcome assessment. One trial found a large reduction in cardiovascular mortality (Malmö 1969),

RR of 0.42 (95% CI 0.23 to 0.77), while another found a large increase (South-East London 1967), RR of 1.54 (95% CI 1.09 to 2.17). The Kaiser Permanente trial (Kaiser Permanente 1965) found a reduction in a pre-specified composite of potentially postponable causes of death, which included colorectal cancer and hypertension related disorders. Ischaemic heart disease was not a part of the composite. Subgroup and sensitivity analyses did not alter the results, nor explain heterogeneity. The two trials at high risk of performance bias showed a harmful effect of the intervention, but we consider this a chance finding.

For cancer mortality (139,290 persons, 3663 deaths) the pooled point estimate was 1.01 (95% CI 0.92 to 1.12) with moderate heterogeneity (I^2 = 33%). Subgroup and sensitivity analyses did not alter the results. The Göteborg 1970 trial (Göteborg 1970) found a reduction in cancer mortality, RR of 0.87 (95% CI 0.76 to 0.99). This was surprising since that trial only screened for cardiovascular risk. Furthermore, the intervention was not successful in reducing smoking. We believe that the result may be due to chance.

Morbidity

Few trials reported on well-defined clinical events. The Göteborg 1970 trial (Göteborg 1970) did not find effects on non-fatal coronary heart disease (CHD), RR of 1.03 (95% CI 0.92 to 1.14), non-fatal stroke (RR 1.12, 95% CI 0.93 to 1.35), combined fatal and non-fatal CHD (RR 0.99, 95% CI 0.91 to 1.07), or combined fatal and non-fatal stroke (RR 1.01, 95% CI 0.86 to 1.20). The results from the WHO trial (WHO 1971) were suggestive of an effect on non-fatal myocardial infarction (RR 0.85, 95% CI 0.72 to 1.01) and combined fatal and non-fatal coronary heart disease (RR 0.90, 95% CI 0.80 to 1.01). The OXCHECK (OXCHECK 1989) authors supplied us with data on incident cancers. When pooling the three intervention groups and comparing with the control group the risk ratio was 1.12 (95% CI 0.85 to 1.48). When using only the group screened at year one, for maximum contrast, the risk ratio was 1.17 (95% CI 0.85 to 1.63).

Four other trials reported some measure of morbidity.

The Kaiser Permanente trial (Kaiser Permanente 1965) found that after seven years 61% of the intervention group reported having a chronic condition compared to 54% in the control group, and that this difference was statistically significant. The conditions were not defined and were likely to have included elevated risk factors like blood pressure or blood glucose.

The South-East London Screening Study (South-East London 1967) did not find effects on the prevalence of angina, ischaemic changes on electrocardiogram, or bronchitic symptoms after five years. For angina the prevalence was 21.9% (screening) and 22.4% (control group), for ischaemic changes 17.9% (screening) and 16.6% (control), and for bronchitic symptoms 29.0% (screening) and 30.6% (control). They also specified the reasons for hospitalisation, using broad categories such as cardiovascular causes, central nervous system causes, and neoplasms, but did not find differences.

The Malmö trial (Malmö 1969) reported reasons for hospitalisations in categories, for example ischaemic heart disease, cerebrovascular disease, and neoplasms, and did not find differences between groups. There was low power due to the stratification in to disease categories. See the results on total hospitalisation below.

The British Family Heart Study (Family Heart 1990) investigated the effect on the prevalence of four conditions. They found substantially more persons with self-reported high blood pressure and high cholesterol in the screening group, slightly more men with self-reported diabetes in the screening group, and no effect on self-reported coronary heart disease. After one year, 6.9% of the control group men had high blood cholesterol compared to 14% of the screening group. For women the results were 3.8% (control) and 9.7% (screening). For high blood pressure, the results for the men were: 14.8% (control) and 17.1% (screening); and for the women: 13.0% (control) and 16.2% (screening). For diabetes, the results for the men were: 1.7% (control) and 3.3% (screening); and for the women: 1.1% (control) and 1.2% (screening). For coronary heart disease, the results for the men were: 5.5% (control) and 5.9% (screening); and for the women: 1.1% (control) and 1.9% (screening). The results were similar when the authors calculated the results within each practice and pooled results. The results were at risk of detection bias and attrition bias.

In summary, we did not find an effect of health checks on morbidity in terms of actual illness, but they may increase the number of people diagnosed with elevated risk factors, as expected.

New diagnoses

In addition to conditions identified through the screening itself, screening might increase diagnostic activity between scheduled screenings due to increased physician contact in relation to follow-up visits or due to a lowered threshold for consulting a physician. Cumulative rates of new diagnoses over time in the screened and unscreened groups would allow an assessment of the full effect of screening on diagnostic activity. However, only one trial reported such results (Kaiser Permanente 1965), but only for the first six years. In a 40% sample, that trial found a sharp divergence in the mean annual number of new diagnoses per participant immediately after the intervention started, with the differences being statistically significant each year. By adding the results for each year we found a mean number of new diagnoses per participant of 4.3 in the screening group and 3.6 in the control group. This corresponded to a 20% increase. The trial lasted for 16 years but follow-up for new diagnoses was not continued.

Four trials reported on the findings at the first screening of the intervention group but without comparisons with the control group over time. The South-East London Screening Study (South-East London 1967) found an average of 2,3 diseases per person at the first screening. Of these 53% were not previously known. The

Ebeltoft trial (Ebeltoft 1992) reported the percentage of participants with abnormal findings prompting health advice at the initial screening to be 76%. The most common reasons were raised CO concentration in expiratory air in smokers (37%), low physical endurance (30%), poor hearing (19%), poor sight (12%), and being overweight (16%). Increased cardiovascular risk was found in 11%, hypercholesterolaemia in 10%, hypertension in 10%, and elevated liver enzymes in 13%. The Salt Lake City Trial (Salt Lake City 1972) found a total of 2031 abnormalities in 384 people screened. This trial used very broad biochemical screening.
In summary, health checks were likely to increase the number of new diagnoses, but the outcome was poorly reported in most trials.

Admission to hospital

Five trials reported on hospitalisation using different measures, for example admission rates, number of people admitted once or more, or number of days in hospital.
The Kaiser Permanente trial (Kaiser Permanente 1965) reported the mean number of days in hospital over 18 years of follow-up. The results were 10.00 days in the intervention group and 10.38 days in the control group (P = 0.13, Wilcoxon rank sum test reported in article). Roughly one third of participants had missing data for this outcome. The South-East London Screening Study (South-East London 1967) reported the number of participants admitted to hospital once or more during nine years of follow-up, risk ratio of 1.04 (95% CI 0.96 to 1.13). The amount of missing data was unclear but was probably low for this outcome. The Malmö trial (Malmö 1969) also studied the number admitted once or more and found similar results, risk ratio of 1.05 (95% CI 0.92 to 1.20). There were 3% to 5% missing data. The Salt Lake City trial (Salt Lake City 1972) compared hospitalisation rates before and after the intervention and did not find an effect, but they did find an effect on the number of nights in hospital in one of three subgroups. The result was unreliable due to biased exclusions after randomisation. The Ebeltoft trial (Ebeltoft 1992) compared admission rates in the two intervention groups with the control group and did not find an effect after eight years, rate ratio of 0.91 (95% CI 0.63 to 1.32). They also compared the random sample invited to participate in the trial with all not invited and found similar results, rate ratio of 0.97 (95% CI 0.80 to 1.18). There were 5% missing data.
In summary, we did not find an effect on admission rates, number of people admitted once or more, or number of days in hospital.

Disability

Three trials investigated the effect on disability. The Kaiser Permanente trial (Kaiser Permanente 1965) found that after 16 years 31% of the screening group and 30% of the control group reported total or partial disability on a questionnaire. Attrition was roughly one third and response rates around 75%, which left only half of the people randomised in this analysis. The South-East London Screening Study (South-East London 1967) found that 2.5% in the screening group and 1.8% in the control group reported major disability after five years. There were between 40% and 50% missing data in this analysis. The Salt Lake City trial (Salt Lake City 1972) compared the number of disability days before and after the intervention and did not find an effect.
In summary, we did not find an effect on disability but the results were unreliable due to a high risk of attrition bias and reporting bias.

Worry

Only two trials reported relevant results, using scales measuring psychological distress.
The Ebeltoft trial (Ebeltoft 1992) used the General Health Questionnaire (GHQ-12) at baseline and after one and five years. A decrease in score indicates a beneficial effect of the intervention. After one year, the change from baseline in the screening groups was an increase of 0.05 and in the control group a decrease of 0.16, P = 0.6. After five years, the screening group had a decrease of 0.23 and the control group had a decrease of 0.39, P = 0.73. They also investigated subgroups of smokers, overweight participants, people who were informed of an elevated risk and people informed of no elevated risk, and did not find effects. Participation was 79.2% after five years.
The South-East London Screening Study (South-East London 1967) used the Middlesex Hospital Questionnaire on a subset of participants after five years. In the anxiety domain of the scale, the authors found significantly lower scores in the intervention group among men (lower scores are better). When pooling men and women, we found a mean score of 4.14 (SD = 3.38, n = 602) in the intervention group and 4.48 (SD = 3.63, n = 572) in the control group, P = 0.097 (t-test, equal variances). In the other domains assessed with this scale ('phobic', 'obsessional', 'somatic', 'depression', 'hysteria') there were no effects. Follow-up was roughly 90%.
In summary, we did not find that screening caused or reduced worry, but only long-term effects were investigated in the trials.

Self-reported health

Four trials reported on self-reported health.
The South-East London Screening Study (South-East London 1967) found that after five years 53.6% of the screening group and 56.5% of the control group reported good or excellent health in the preceeding two weeks (X^2 = 3.274, P = 0.07).
The Ebeltoft trial (Ebeltoft 1992) used a five-point scale at baseline and after five years. After five years, 69.9% and 71.6% of the two intervention groups reported good or excellent health compared to 71% of the control group. Data on change from baseline were only available in a graph. This showed that approximately 12%

in the intervention groups had an improvement in self-reported health compared to approximately 20% in the control group. Approximately 60% in the intervention groups had no change compared to approximately 52% in the control group. In all groups approximately 28% had worsened self-reported health.
In the British Family Heart Study (Family Heart 1990) 79.5% of the screening group and 75.7% of the internal control group reported good or excellent health after one year. This analysis used last observation carried forward for missing data. The pooled difference, taking into account the 13 different practices, was 3.8% in favour of screening, P = 0.004.
The Inter99 trial (Inter99 1999) used SF-12 and found significantly slower deterioration of both physical and mental health components in the intervention group. For mental health, the difference after five years was approximately 2 on a 100-point scale, where 50 is the mean of a reference population and the standard deviation is set to 10. The effect was smaller for physical health but was difficult to assess because of baseline imbalances in scores. The authors found indications of biased non-response.
In summary, two out of four trials found small beneficial effects on self-reported health but they may be due to bias.

Referrals to specialists

Only one trial (Ebeltoft 1992) reported on this outcome, but the results could not be used in our analysis. This was because the authors only had data from 1995 to 1999 but the screening took place in 1992 to 1993 (intervention groups screened) and 1997 (intervention groups and control group screened). This means that the expected increase in referrals following the intervention was not included in the analysis, and that any contrast between groups would be diminished by the 1997 screening. The authors made two comparisons and did not find effects in either analysis. When comparing the screening and control groups, the rate ratio was 1.04 (95% CI 0.85 to 1.26). When comparing the random sample invited to participate in the trial versus all eligible people not invited, the rate ratio was 0.94 (95% CI 0.84 to 1.06).

Number of non-scheduled visits to general practitioners

Five trials reported on physician visits. The length of follow-up was between one and nine years, with missing outcome data ranging between 5% (Ebeltoft 1992) and 51% (Salt Lake City 1972).
The Kaiser Permanente trial (Kaiser Permanente 1965) found a mean number of physician visits of 16.0 in both groups after five years, not including the screenings themselves. The results were reported without measures of uncertainty and data on this outcome were collected from a 20% subsample, which reduces power.
The South-East London Screening Study (South-East London 1967) did not find an effect on the mean annual number of physician visits. It was not clear whether the screening visits were included in this, and we cannot tell whether the results were from the five-year or nine-year follow-up. Participants who left the study before one year were excluded from the analyses (14% from the screening group and 13% from the control group).
The Northumberland trial (Northumberland 1969) found an average number of consultations per participant of 5.4 in the screening group and 5.0 in the control group over 1½ years. This did not include the screenings themselves. When adding the screenings the results were 6.3 in the screening group and 5.0 in the control group. The type of health check was not specified, and there was a high risk of detection bias.
The Salt Lake City trial (Salt Lake City 1972) did not find effects after one year, but this result was unreliable. The screening visits were not included in the analysis.
The Ebeltoft trial (Ebeltoft 1992) found an increased rate of physician visits after five years in the screening plus health discussion group compared to the control group, rate ratio of 1.15 (95% CI 1.02 to 1.31) but not in the screening only group compared to controls, rate ratio of 1.01 (95% CI 0.89 to 1.15). When comparing all those invited to participate in the trial with all not invited, the rate ratio was 1.01 (95% CI 0.93 to 1.10). However, this comparison included data from 1992 to 1999 and thus included the screening of the control group in 1997, diluting any differences between groups. The authors found a significant downwards trend in the rate ratio over time favouring the intervention, but in the absence of an overall effect this is not a relevant observation. It likely reflects the initial increase in visits generated by the screenings themselves, which gave a high starting point for the trend analysis. Similarly, the 1997 screening of the control group would be expected to cause an increase in physician visits in the control group, further contributing to the downward trend.
In summary, we did not find an effect on physician visits. Most trials did not include the screening visits in the analysis.

Number of additional diagnostic procedures required because of positive screening tests

None of the trials reported on this outcome.
The Kaiser Permanente trial (Kaiser Permanente 1965) reported the mean number of laboratory tests per participant after five and 10 years, based on a 20% sample. After five years it was 23.8 in the screening group and 23.3 in the control group. The data after 10 years were not reported but it was stated in a narrative that there was no difference. The number of laboratory tests did not include the tests used at screening.

Prescriptions and surgery

None of the trials reported the total number of prescriptions, new drugs prescribed, or the number of operations performed. This is unfortunate since these are important factors for balancing the benefits and harms of health checks, and for estimating the costs. Five trials provided some results of relevance.

The Göteborg 1970 trial (Göteborg 1970) examined random samples of the intervention group and control group 1 and found that after 10 years of follow-up 26.0% of the intervention group used antihypertensive medications compared to 19.6% in the control group (Chi^2 = 16.41, P < 0.0001, our calculation). The Kaiser Permanente trial (Kaiser Permanente 1965) reported in a narrative that prescription rates gathered from pharmacies showed a non-significant trend towards increased prescription in the screening group, but only data from years six and seven were analysed. The Ebeltoft trial (Ebeltoft 1992) presented data on self-reported use of selected types of drugs after five years. In the screening groups, 4.8% reported using blood pressure medication compared to 6.8% in the control group (X^2 = 1.42, P = 0.23, our calculation). For diuretics, the figures were 3.7% (screening) and 3.9% (control group), and for heart medication they were 0.9% (screening) and 1.0% (control). The British Family Heart Study (Family Heart 1990) reported in a narrative that there was no difference between the intervention and control groups regarding use of drugs to lower blood pressure or cholesterol, or for diabetes. The Mankato trial (Mankato 1982) reported that the proportion of participants on blood pressure medication after one year was 13.8% in the intervention group and 9.8% in the control group (P < 0.05).

In summary, we cannot make firm conclusions on total drug use. Two out of four trials found increased use of antihypertensive medication, but there was a high risk of bias in all the results. None of the trials studied the amount of surgery used.

Absence from work

Two trials reported on absence from work (Kaiser Permanente 1965; South-East London 1967). Neither trial found an effect, and neither trial reported the exact results but only mentioned their findings in a narrative.

Subgroup and sensitivity analyses

We planned and performed several subgroup and sensitivity analyses. Some of the resulting subgroups were based on very few trials but are presented for completeness. They should be interpreted with caution. We found no convincing patterns in any subgroup or sensitivity analysis.

For outcomes not included in the meta-analyses we considered the same factors. We were not able to discern any patterns except that the more recent trials often had a strong focus on lifestyle interventions, often had changes in risk factors as their primary outcomes, and were designed accordingly (shorter follow-up) (Ebeltoft 1992; Family Heart 1990; Mankato 1982; OXCHECK 1989).

DISCUSSION

Summary of main results

We did not find an effect of general health checks on total or cause-specific mortality. For total mortality our confidence interval includes a 5% reduction and a 3% increase, both of which would be clinically relevant. However, for the causes of death most likely to be influenced by health checks, cardiovascular and cancer-specific mortality, there were no reductions either. A substantial latency of effects on mortality would be expected but we included several trials with very long follow-up. Our results suggest that the lack of an effect on total mortality is not a chance finding, nor due to low power, but that there may in fact be no or only a minimal effect of the intervention on mortality in general non-geriatric populations.

We did not find an effect on morbidity, hospitalisations, disability, visits to the physician, number of referrals, or absence from work. We found indications of an increase in the number of new diagnoses as well as descriptions of large numbers of abnormal findings at the initial screenings. We also found indications of increased use of antihypertensive drugs, but this outcome was poorly studied. We did not find an effect on measures of psychological distress but this was also sporadically reported and only for long-term effects. Two out of four trials found a possible small improvement in self-reported health but this may have been due to bias. None of the trials studied the number of follow-up tests after positive screening results, nor the amount of surgery resulting from the intervention.

In general, the outcomes expected to reflect beneficial effects of the intervention were better studied and reported than the harmful effects. We expected the number of new diagnoses and initiated treatments to be reported since these are important elements of screening, but this was rarely the case. Only one trial reported the number of new diagnoses in the two groups, and only for the first six years although the intervention was continued for 16 years (Kaiser Permanente 1965). Drug use was only assessed for selected drugs and was mainly self-reported, with a risk of attrition bias and detection bias because the screening groups could not be blinded. We also expected the number of follow-up tests and referrals to specialists to be reported since they also reflect the burden placed by screening upon the participants and the health-care system. However, these outcomes were rarely reported. Without knowing the amount of 'downstream' investigations following screening, it is not possible to evaluate the harms or costs. This has been long recognised for screening interventions in general (Raffle & Gray 2007).

Increased diagnostic and therapeutic activity would be expected if general health checks led to improved health, at least in the short term, as this is the main mechanism of the intervention. However, more diagnoses and more treatment in the absence of health improvements would indicate overdiagnosis and overtreatment. Overdiagnosis is the diagnosis of conditions which were not destined to cause symptoms or affect the longevity of the patient if they had not been detected at screening, and is an inherent risk in any screening programme. Overdiagnosis leads to overtreatment

and perhaps increased anxiety and undesirable effects on peoples' image of their own health. These harms have been documented in cancer screening and are also obvious harms in screening for cardiovascular risk factors, as reflected in the large numbers needed to treat in primary prevention of cardiovascular disease (Welch, Schwartz and Woloshin 2011).

The psychological consequences of general health checks were investigated to a somewhat greater extent, although only in a minority of trials. An interesting result is that we did not find harmful effects on measures of psychological distress, self-reported health, or absence from work. Two trials found beneficial effects on self-reported health, but the effects were small and could be due to bias. One systematic review (Boulware 2007) found beneficial effects of periodic health evaluations on worry in one trial of elderly people (Patrick 1999), and a systematic review of coronary heart disease risk scores found no harmful effects in two fair quality studies (Sheridan 2008). Regarding hypertension, cross-sectional studies have found that people diagnosed with hypertension had poorer self-reported health regardless of whether they were correctly diagnosed or not (Barger 2006; Bloom 1981). However, a review of cohort studies found mixed effects on absenteeism and fair quality evidence that screening for hypertension does not cause adverse psychological effects (Sheridan 2003). One review found short-term adverse psychological effects from predicting a person's risk of illness, but no long-term effects (Shaw 1999). Similarly, a review of trials of any kind of screening found no long-term effect on anxiety, depression, or quality of life, but the reviewers were not able to make conclusions about short-term effects (Collins 2011). None of the trials we reviewed reported on short-term adverse psychological effects.

The lack of measurable effects indicates that general health checks did not work as intended in the included trials. Below, we explore possible reasons for the apparent lack of effect as well as challenges in generalising the results to the present day.

Bias

Three trials in our mortality meta-analyses were biased towards no effect (Ebeltoft 1992; Kaiser Permanente 1965; South-East London 1967), and in one trial we prioritised power over contrast in the merging of intervention groups (OXCHECK 1989). In a post hoc sensitivity analysis, removing these trials from the analyses did not change the results and only marginally expanded the confidence intervals.

Type of health check

Many of the older trials investigated very broad screening regimens, with a large potential for detecting abnormalitites. Healthy people frequently harbour pathology that can be discovered by examination, imaging (Furtado 2005; Xiong 2005), or biopsy (Welch 2004), but this is not necessarily beneficial and it may be harmful (Welch, Schwartz and Woloshin 2011). The results from the Kaiser Permanente trial (Kaiser Permanente 1965) suggested that it was as they found increases in mortality due to lymphohaematopoietic cancers and suicide. This may be a random finding but the pattern appeared after seven years and continued throughout the full 16 years of the trial. The increase in available diagnostic tests might lead to more invasive follow-up procedures today and more drug treatment and surgery, for example for prostate and thyroid cancer, with resulting harms. Today, no authorities recommend health checks as broad as studied in some of the older trials but they are still common, particularly among commercial providers (Grønhøj Larsen 2012). In contrast, the recently introduced National Health Service (NHS) Health Check in the UK is focused on cardiovascular risk and diabetes, with fewer tests.

Most of the trials that reported on mortality did not have an explicit lifestyle intervention component, but we do not expect this element to be particularly important. Multiple risk factor interventions directed at general populations for the primary prevention of coronary heart disease have been extensively studied and found to be without effect on total and coronary heart disease-specific mortality, or the number of cardiovascular events (Ebrahim 2011). One of the trials in our review included a randomised comparison between screening with and without scheduled face-to-face lifestyle conversations, but found no effect (Ebeltoft 1992).

Developments in therapy

Developments in preventive drug therapy might produce a different effect on cardiovascular outcomes today compared to when the identified trials were performed. For example, use of statins and angiotensin converting enzyme inhibitors instead of harmful drugs such as clofibrate (WHO 1984) and reserpine (Healy 2004) is likely to provide a considerable improvement. However, we cannot be certain that developments in drug treatments are always beneficial to patients because some modern drugs may have serious side effects that are not known at present. For example, the diabetes drug rosiglitazone was on the market for 10 years before being withdrawn because it causes serious heart disease (Lehman 2010; Nissen 2010), and tiotropium mist inhalers for chronic obstructive pulmonary disease have recently been shown to increase mortality (Singh 2011). Also, poor trial reporting of harms from commonly used preventive drugs, such as statins (Taylor 2011), may mean that adverse effects are more common and more serious than we think (Golomb 2012).

Thresholds for treating cardiovascular risk factors and diabetes are lower today than at the time most of the included trials were conducted. This has lead to increased prescription of preventive drugs with demonstrated efficacy, for example statins (Taylor 2011) and antihypertensives (Wright 2009). However, the balance between benefits and harms may be unfavourable when the absolute risks are low, such as in a screened population, or when used in more heterogeneous populations with more co-morbidities. For exam-

ple, the populations used for testing antihypertensive drugs were usually younger and had less co-morbidity than the typical patient in general practice (Uijen 2007). Thus, we cannot know whether results would be better today. Morbidity and mortality results from the Inter99 trial (Inter99 1999) will inform about the effect of health checks in a modern setting.

Therapy for identified disease has improved in many areas and this might lead to better effects of health checks over time. However, in the meta-analyses arranged by year of trial start there are no visible time trends (Analysis 1.1; Analysis 1.14; Analysis 1.27), and the idea of increasing benefits over time therefore remains hypothetical.

Self-selection

People who accept an invitation to a health check are often different from those who do not. They tend to have higher socioeconomic status (Pill 1988), lower cardiovascular risk (Waller 1990), less cardiovascular morbidity (Jørgensen 2003), and lower mortality (Göteborg 1970). Thus, systematic health checks may not reach those who need prevention the most, and they have been called 'another example of inverse care' (Waller 1990).

Clinically motivated testing

Another possible reason for the lack of beneficial effects is that many physicians already carry out screening for cardiovascular risk factors or diseases in patients that they judge to be at high risk when they see them for other reasons. This is often considered an integral part of primary care practice. Clinically motivated testing may already have resulted in the identification of many people at high risk thus eroding the potential for a benefit from systematic screening.

Potential biases in the review process

In the meta-analyses, we ignored clustering by family in two trials (OXCHECK 1989; South-East London 1967) and by factory in the analysis of cancer mortality from the WHO trial (WHO 1971). In a pre-specified sensitivity analysis, excluding cluster randomised trials resulted in very little change to the results.

We attempted to contact authors and succeeded in 10 cases (Ebeltoft 1992; Göteborg 1963; Göteborg 1970; Inter99 1999; Malmö 1969; Mankato 1982; OXCHECK 1989; South-East London 1967; Stockholm 1969; WHO 1971). We often had questions about trial methods but, since most trials were quite old, there is a risk that some answers may have been inaccurate.

Agreements and disagreements with other studies or reviews

The existing systematic review of health checks included observational studies and geriatric studies but used a different definition of the intervention and included fewer trials (Boulware 2006; Boulware 2007). The trials reviewed by us are largely different but the results are broadly in line for the overlapping outcomes of total mortality, hospitalisation, disability, and the number of new diagnoses (disease detection). For worry, the previous review found one trial which showed a beneficial effect whereas we found two trials without an effect on this outcome.

We did not include geriatric trials because they included interventions other than screening for disease and risk factors, and lifestyle interventions. A systematic review of 89 trials of complex interventions to improve physical function and maintain independent living in elderly people found beneficial effects on the risk of not living at home, nursing home admission, falls, hospital admissions, and physical function, but not mortality (Beswick 2008). In the subgroup of 28 trials of geriatric assessments for elderly people representing the general population, the results were similar except no effect on hospitalisation was found. Thus, the results were similar to ours except on outcomes of special relevance to older people, where important benefits were found.

AUTHORS' CONCLUSIONS

Implications for practice

Our results do not support the use of general health checks aimed at a general population outside the context of randomised trials. Our results do not imply that physicians should stop clinically motivated testing and preventive activities as such activities may be an important reason why an effect of general health checks has not been shown. Public healthcare initiatives to systematically offer general health checks should be resisted, and private suppliers of the intervention do so without support from the best available evidence.

Implications for research

We suggest that future research be directed at the individual components of health checks, for example screening for cardiovascular risk factors, chronic obstructive pulmonary disease, diabetes, or kidney disease. We also suggest that surrogate outcomes such as changes in risk factors are not used for assessing the benefits of health checks since they do not capture harmful effects and may lead to misleading conclusions. The required large randomised trials with long follow-up are expensive but not nearly as expensive as the implementation of ineffective or harmful screening programmes. The results on total and cause-specific mortality from the Inter99 study will be published soon and will reflect the effect of the intervention in a modern setting. If these results are also negative, there would seem to be little reason to embark on further

randomised trials of general health checks until new treatments for risk factors and early disease could substantially alter our expectations for an effect.

ACKNOWLEDGEMENTS

We wish to thank the trialists who helped us: Guy De Backer, Walter W Holland, Sven-Olof Isacsson, Torben Jørgensen, Olof Lannerstad, Torsten Lauritzen, David Murray, Charlotta Pisinger, Lennart Welin, and Lars Wilhelmsen for additional information on their trials; and David Mant, Alice Fuller, Holger Theobald, and Janus L Thomsen for providing unpublished outcome data.

We also wish to thank the EPOC Trials Search Co-ordinator, Michelle Fiander, for designing, conducting and updating the searches, and the peer reviewers for their valuable comments.

REFERENCES

References to studies included in this review

Ebeltoft 1992 ***{published and unpublished data}***

Christensen B, Engberg M, Lauritzen T. No long-term psychological reaction to information about incrased risk of coronary heart disease in general practice. *European Journal of Cardiovascular Prevention and Rehabilitation* 2004;**11**(3): 239–43.

Engberg M, Christensen B, Karlsmose B, Lous J, Lauritzen T. Can systematic general health screening and patient-physician health discussions improve the cardiovascular profile of the population? A randomized controlled trial in general practice with a 5-year follow-up [Kan systematiske generelle helbredsundersøgelser og helbredssamtaler forbedre den kardiovaskulære risikoprofil i befolkningen? En randomiseret og kontrolleret undersøgelse i almen praksis med fem års opfølgning]. *Ugeskrift for Laeger* 2002; **164**(25):3354–60.

* Engberg M, Christensen B, Karlsmose B, Lous J, Lauritzen T. General health screenings to improve cardiovascular risk profiles: A randomized controlled trial in general practice with 5-year follow-up. *The Journal of Family Practice* 2002; **51**(6):546–52.

Geest TA, Engberg M, Lauritzen T. Discordance between self-evaluated health and doctor-evaluated health in relation to general health promotion. *Scandinavian Journal of Primary Health Care* 2004;**22**(3):146–51.

Kanstrup H, Refsgaard J, Engberg M, Lassen JF, Larsen ML, Lauritzen T. Cholesterol reduction following health screening in general practice. *Scandinavian Journal of Primary Health Care* 2002;**20**(4):219–23.

Karlsmose B, Lauritzen T, Engberg M, Parving A. A randomised controlled trial of screening for adult hearing loss during preventive health checks. *The British Journal of General Practice* 2001;**51**(466):351–5.

Larsen EL, Thomsen JL, Lauritzen T, Engberg M. [Forebyggende helbredsundersøgelser og helbredssamtaler i almen praksis – en analyse af patientperspektivet. En analyse med særligt fokus på Sundhedsprojekt Ebeltoft og andre randomiserede studier]. Medicinsk Teknologivurdering - puljeprojekter 6(7). Copenhagen (Denmark): Sundhedsstyrelsen, Center for Evaluering og Medicinsk Teknologivurdering 2006.

Lauritzen T. [Helbredsundersøgelser og samtaler reducerer deltagernes risikoscore for hjerte–kar–sygdom]. *Månedsskrift for Almen Praksis* 2012;**90**(3):245–50.

Lauritzen T, Jensen MS, Thomsen JL, Christensen B, Engberg M. Health tests and health consultations reduced cardiovascular risk without psychological strain, increased healthcare utilization or increased costs. An overview of the results from a 5-year randomized trial in primary care. The Ebeltoft Health Promotion Project (EHPP). *Scandinavian Journal of Public Health* 2008;**36**(6):650–61.

Lauritzen T, Leboeuf-Yde C, Lunde IM, Nielsen KD. Ebeltoft project: baseline data from a five-year randomized, controlled, prospective health promotion study in a Danish population. *The British Journal of General Practice* 1995;**45** (399):542–7.

Lauritzen T, Nielsen KD, Leboeuf-Yde C, Lunde IM. [The health project Ebeltoft: health check ups and discussions in general practice. Basic data from a 5-year, prospective, randomized, controlled population study]. *Ugeskrift for Laeger* 1997;**159**(25):3940–6.

Rasmussen SR, Kilsmark J, Hvenegaard A, Thomsen JL, Engberg M, Lauritzen T, et al.[Forebyggende helbredsundersøgelser og helbredssamtaler i almen praksisEn sundhedsøkonomisk analyse af »Sundhedsprojekt Ebeltoft]. Medicinsk Teknologivurdering - puljeprojekter; 6 (6). Copenhagen (Denmark): Sundhedsstyrelsen, Center for Evaluering og Medicinsk Teknologivurdering 2006.

Rasmussen SR, Thomsen JL, Kilsmark J, Hvenegaard A, Engberg M, Lauritzen T, et al.Preventive health screenings and health consultations in primary care increase life expectancy without increasing costs. *Scandinavian Journal of Public Health* 2007;**35**(4):365–72.

Thomsen JL, Karlsmose B, Parner ET, Thulstrup AM, Lauritzen T, Engberg M. Secondary healthcare contacts after multiphasic preventive health screening: a randomized trial. *Scandinavian Journal of Public Health* 2006;**34**(3):254–61.

Thomsen JL, Parner ET. Methods for analysing recurrent events in health care data. Examples from admissions in Ebeltoft Health Promotion Project. *Family Practice* 2006; **23**(4):407–13.

Thomsen JL, Parner ET, Karlsmose B, Thulstrup AM, Lauritzen T, Engberg M. Effect of preventive health screening on long-term primary health care utilization. A randomized controlled trial. *Family Practice* 2005;**22**(3):

242–8.

Family Heart 1990 *{published data only}*

Davies G, Pyke S, Kinmonth AL, Family Heart Study Group. Effect of non-attenders on the potential of a primary care programme to reduce cardiovascular risk in the population. *BMJ* 1994;**309**(6968):1553–6.

Family Heart Study Group. British family heart study: its design and method, and prevalence of cardiovascular risk factors. *The British Journal of General Practice* 1994;**44** (397):62–7.

* Family Heart Study Group. Randomised controlled trial evaluating cardiovascular screening and intervention in general practice: principal results of British Family Heart Study. *BMJ* 1994;**308**(6924):313–20.

Marteau TM, Kinmonth AL, Pyke S, Thompson SG, Family Heart Study Group. Readiness for lifestyle advice: self-assessments of coronary risk prior to screening in the British Family Heart Study. *The British Journal of General Practice* 1995;**45**(390):5–8.

Marteau TM, Kinmonth AL, Thompson S, Pyke S, British Family Heart Study Group. The psychological impact of cardiovascular screening and intervention in primary care: a problem of false reassurance?. *The British Journal of General Practice* 1996;**46**(411):577–82.

Pyke SD, Wood DA, Kinmonth AL, Thompson SG, Family Heart Study Group. Change in coronary risk and coronary risk factor levels in couples following lifestyle intervention. *Archives of Family Medicine* 1997;**6**(4):354–60.

Göteborg 1963 *{published data only}*

Grimby G, Bjure J, Aurell M, Ekström-Jodal B, Tibblin G, Wilhelmsen L. Work capacity and physiologic responses to work. Men born in 1913. *The American Journal of Cardiology* 1972;**30**(1):37–42.

Larsson B, Björntorp P, Tibblin G. The health consequences of moderate obesity. *International Journal of Obesity* 1981;**5** (2):97–116.

Tibblin G. A population study of 50-year-old men. An analysis of the non-participation group. *Acta Medica Scandinavica* 1965;**178**(4):453–9.

Tibblin G. High blood pressure in men aged 50 -- a population study of men born in 1913. *Acta Medica Scandinavica* 1967;**470**:1–84.

Tibblin G, Aurell E, Hjortzberg-Nordlund H, Paulin S, Risholm L, Sanne H, et al.A general health examination of a random sample of 50-year-old men in Göteborg. *Acta Medica Scandinavica* 1965;**177**:739–49.

* Tibblin G, Welin L, Larsson B, Ljungberg IL, Svärdsudd K. The influence of repeated health examinations on mortality in a prospective cohort study, with a comment on the autopsy frequency. The study of men born in 1913. *Scandinavian Journal of Social Medicine* 1982;**10**(1):27–32.

Göteborg 1970 *{published data only}*

* Wilhelmsen L, Berglund G, Elmfeldt D, Tibblin G, Wedel H, Pennert K, et al.The multifactor primary prevention trial in Göteborg, Sweden. *European Heart Journal* 1986;**7**(4): 279–88.

Wilhelmsen L, Ljungberg S, Wedel H, Werkö L. A comparison between participants and non-participants in a primary preventive trial. *Journal of Chronic Diseases* 1976; **29**(5):331–9.

Wilhelmsen L, Tibblin G, Werkö L. A primary preventive study of Gothenburg, Sweden. *Preventive Medicine* 1972;**1** (1):153–60.

Inter99 1999 *{published data only}*

Jørgensen T, Borch-Johnsen K, Thomsen TF, Ibsen H, Glümer C, Pisinger C. A randomized non-pharmacological intervention study for prevention of ischaemic heart disease: baseline results Inter99. *European Journal of Cardiovascular Prevention and Rehabilitation* 2003;**10**(5):377–86.

Jørgensen T, Ladelund S, Borch-Johnsen K, Pisinger C, Schrader AM, Thomsen T, et al.Screening for risk of cardiovascular disease is not associated with mental distress: the Inter99 study. *Preventive Medicine* 2009;**48**(3):242–6.

* Pisinger C, Ladelund S, Glümer C, Toft U, Aadahl M, Jørgensen T. Five years of lifestyle intervention improved self-reported mental and physical health in a general population: the Inter99 study. *Preventive Medicine* 2009;**49** (5):424–8.

Kaiser Permanente 1965 *{published data only}*

Collen M. Cost analysis of the Kaiser Foundations systemized health evaluation. *Preventive Medicine USA*. New York: Prodist, 1976.

Collen MF. The cost-effectiveness of health checkups--an illustrative study. *The Western Journal of Medicine* 1984;**141** (6):786–92.

Collen MF, Dales LG, Friedman GD, Flagle CD, Feldman R, Siegelaub AB. Multiphasic checkup evaluation study. 4. Preliminary cost benefit analysis for middle-aged men. *Preventive Medicine* 1973;**2**(2):236–46.

Cutler JL, Ramcharan S, Feldman R, Siegelaub AB, Campbell B, Friedman GD, et al.Multiphasic checkup evaluation study. 1. Methods and population. *Preventive Medicine* 1973;**2**(2):197–206.

Dales LG, Friedman GD, Collen MF. Evaluating periodic multiphasic health checkups: a controlled trial. *Journal of Chronic Diseases* 1979;**32**(5):385–404.

Dales LG, Friedman GD, Collen MF. Evaluation of a periodic multiphasic health checkup. *Methods of Information in Medicine* 1974;**13**(3):140–6.

Dales LG, Friedman GD, Ramcharan S, Siegelaub AB, Campbell BA, Feldman R, et al.Multiphasic Checkup Evaluation Study. 3. Outpatient clinic utilization, hospitalization, and mortality experience after seven years. *Preventive Medicine* 1973;**2**(2):221–35.

* Friedman GD, Collen MF, Fireman BH. Multiphasic Health Checkup Evaluation: a 16-year follow-up. *Journal of Chronic Diseases* 1986;**39**(6):453–63.

Ramcharan S, Cutler JL, Feldman R, Siegelaub AB, Campbell B, Friedman GD, et al.Multiphasic checkup evaluation study. 2. Disabilty and chronic disease after seven years of multiphasic health checkups. *Preventive Medicine* 1973;**2**(2):207–20.

Selby JV, Friedman GD, Collen MF. Sigmoidoscopy and mortality from colorectal cancer: the Kaiser Permanente

Multiphasic Evaluation Study. *Journal of Clinical Epidemiology* 1988;**41**(5):427–34.

Malmö 1969 ***{published data only}***

Isacsson SO, Janzon L. Results of a quit-smoking research project in a randomly selected population. *Scandinavian Journal of Social Medicine* 1976;**4**(1):25–9.

Lannerstad O. *Död och sjukdom bland medelålders män*. Malmö: University of Lund, 1978.

Lannerstad O. Morbidity related to smoking and other risk factors. A population study of disability pension, hospital care and sickness benefit days among middle-aged men in Malmö, Sweden. *Scandinavian Journal of Social Medicine* 1980;**8**(1):25–31.

* Lannerstad O, Sternby NH, Isacsson SO, Lindgren G, Lindell SE. Effects of a health screening on mortality and causes of death in middle-aged men. A prospective study from 1970 to 1974 of mean in Malmö, born 1914. *Scandinavian Journal of Social Medicine* 1977;**5**(3):137–40.

Mankato 1982 ***{published data only}***

* Murray DM, Luepker RV, Pirie PL, Grimm RH, Bloom E, Davis MA, et al.Systematic risk factor screening and education: a community-wide approach to prevention of coronary heart disease. *Preventive Medicine* 1986;**15**(6): 661–72.

New York 1971 ***{published data only}***

Fink R, Shapiro S, Rosenberg C. Social research problems in studies involving multiphasic health testing. *Health Services Research* 1972;**7**(4):314–21.

Shapiro S. Evaluation of two contrasting types of screening programs. *Preventive Medicine* 1973;**2**(2):266–77.

Shapiro S, Fink R, Rosenberg C. A program to measure the impact of multiphasic health testing on health differentials between poverty and non-poverty groups. *Medical Care* 1972;**10**(3):207–14.

Northumberland 1969 ***{published data only}***

* Bennett AE, Fraser IG. Impact of a screening programme in general practice: a randomized controlled trial. *International Journal of Epidemiology* 1972;**1**(1):55–60.

OXCHECK 1989 ***{published and unpublished data}***

* Imperial Cancer Research Fund OXCHECK Study Group. Effectiveness of health checks conducted by nurses in primary care: final results of the OXCHECK study. *BMJ* 1995;**310**(6987):1099–104.

Imperial Cancer Research Fund OXCHECK Study Group. Effectiveness of health checks conducted by nurses in primary care: results of the OXCHECK study after one year. *BMJ* 1994;**308**(6924):308–12.

Imperial Cancer Research Fund OXCHECK Study Group. Prevalence of risk factors for heart disease in OXCHECK trial: implications for screening in primary care. *BMJ* 1991; **302**(6784):1057–60.

Langham S, Thorogood M, Normand C, Muir J, Jones L, Fowler G. Costs and cost effectiveness of health checks conducted by nurses in primary care: the OXCHECK study. *BMJ* 1996;**312**(7041):1265–8.

O'Neill C. The OXCHECK Study: a nursing perspective. Oxford and Collaborators Health Check. *Health Reports* 1994;**6**(1):160–5.

Salt Lake City 1972 ***{published data only}***

* Olsen DM, Kane RL, Proctor PH. A controlled trial of multiphasic screening. *The New England Journal of Medicine* 1976;**294**(17):925–30.

South-East London 1967 ***{published data only}***

D'Souza MF, Swan AV, Shannon DJ. A long-term controlled trial of screening for hypertension in general practice. *Lancet* 1976;**1**(7971):1228–31.

Eastwood MR, Trevelyan MH. Relationship between physical and psychiatric disorder. *Psychological Medicine* 1972;**2**(4):363–72.

South-East London Screening Study Group. A controlled trial of multiphasic screening in middle-age: results of the South-East London Screening Study.1977. *International Journal of Epidemiology* 2001;**30**(5):935–40.

Stone DH, Crisp AH. The effect of multiphasic screening on aspects of psychiatric status in middle age: results of a controlled trial in general practice. *International Journal of Epidemiology* 1978;**7**(4):331–4.

Stone DH, D'Souza MF. Multiphasic screening in middle age: results and implications of a controlled trial in British general practice. *Israel Journal of Medical Sciences* 1981;**17** (2-3):215–21.

* The South-East London Study Group. A controlled trial of multiphasic screening in middle-age: results of the South-East London screening study. *International Journal of Epidemiology* 1977;**6**(4):357–63.

Trevelyan H. Study to evaluate the effects of multiphasic screening within general practice in Britain: design and method. *Preventive Medicine* 1973;**2**(2):278–94.

Stockholm 1969 ***{published data only}***

Theobald H, Bygren LO, Carstensen J, Hauffman M, Engfeldt P. Effects of an assessment of needs for medical and social services on long-term mortality: a randomized controlled study. *International Journal of Epidemiology* 1998;**27**(2):194–8.

Titograd 1971 ***{published data only}***

Thorner RM, Djordjevic D, Vuckmanovic C, Pesic B, Culafic B, Mark F. A study to evaluate the effectiveness of multiphasic screening in Yugoslavia. *Preventive Medicine* 1973;**2**(2):295–301.

WHO 1971 ***{published data only}***

Kornitzer M, Rose G. WHO European collaborative trial of multifactorial prevention of coronary heart disease. *Preventive Medicine* 1985;**14**(3):272–8.

Rose G. European collaborative trial of multifactorial prevention of coronary heart disease. *Lancet* 1987;**1**(8534): 685.

World Health Organization European Collaborative Group. An international controlled trial in the multifactorial

prevention of coronary heart disease. *International Journal of Epidemiology* 1974;**3**(3):219–24.
World Health Organization European Collaborative Group. European collaborative trial of multifactorial prevention of coronary heart disease: final report on the 6-year results. *Lancet* 1986;**1**(8486):869–72.
World Health Organization European Collaborative Group. Multifactorial trial in the prevention of coronary heart disease: 1. recruitment and initial findings. *European Heart Journal* 1980;**1**(1):73–80.
World Health Organization European Collaborative Group. Multifactorial trial in the prevention of coronary heart disease: 2. risk factor changes at two and four years. *European Heart Journal* 1982;**3**(2):184–90.
World Health Organization European Collaborative Group. Multifactorial trial in the prevention of coronary heart disease: 3. incidence and mortality results. *European Heart Journal* 1983;**4**(3):141–7.
World Health Organization European Collaborative Group. WHO European collaborative trial in the multifactorial prevention of coronary heart disease. Copenhagen (Denmark): World Health Organization 1989.
World Health Organization European Collaborative Group. WHO European Collaborative Trial in the multifactorial prevention of coronary heart disease. *IARC Scientific Publications* 1990, (103):123–31.

References to studies excluded from this review

Addition - Cambridge *{published data only}*
Eborall HC, Griffin SJ, Prevost AT, Kinmonth AL, French DP, Sutton S. Psychological impact of screening for type 2 diabetes: controlled trial and comparative study embedded in the ADDITION (Cambridge) randomised controlled trial. *BMJ* 2007;**335**(7618):486.
Echouffo-Tcheugui JB, Simmons RK, Williams KM, Barling RS, Prevost AT, Kinmonth AL, et al.The ADDITION-Cambridge trial protocol: a cluster -- randomised controlled trial of screening for type 2 diabetes and intensive treatment for screen-detected patients. *BMC Public Health* 2009;**9**:136.

A Healthy Future *{published data only}*
Durham ML, Beresford S, Diehr P, Grembowski D, Hecht JA, Patrick DL. Participation of higher users in a randomized trial of Medicare reimbursement for preventive services. *Gerontologist* 1991;**31**(5):603–6.
Patrick DL, Beresford SA, Ehreth J, Diehr P, Picciano J, Durham M, et al.Interpreting excess mortality in a prevention trial for older adults. *International Journal of Epidemiology* 1995;**24 Suppl 1**:S27–33.
Patrick DL, Grembowski D, Durham M, Beresford SA, Diehr P, Ehreth J, et al.Cost and outcomes of Medicare reimbursement for HMO preventive services. *Health Care Financing Review* 1999;**20**(4):25–43.

Aleman 2011 *{published data only}*
Aleman S, Endalib S, Stål P, Lööf L, Lindgren S, Sandberg-Gertzén H, et al.Health check-ups and family screening allow detection of hereditary hemochromatosis with less advanced liver fibrosis and survival comparable with the general population. *Scandinavian Journal of Gastroenterology* 2011;**46**(9):1118–26.

Apkon 2005 *{published data only}*
Apkon M, Mattera JA, Lin Z, Herrin J, Bradley EH, Carbone M, et al.A randomized outpatient trial of a decision-support information technology tool. *Archives of Internal Medicine* 2005;**165**(20):2388–94.

Asshauer 1972 *{published data only}*
Asshauer E, Fenner O, Göpel HG, Ullrich HG, Wachholz P. [Medical check up. Pilot study on its evaluation using electronic data processing units]. *Münchener Medizinische Wochenschrift* 1972;**114**(24):1165–71.

Atthobari 2004 *{published data only}*
Atthobari J, Monster TB, de Jong PE, de Jong-van den Berg LT, PREVEND Study Group. The effect of hypertension and hypercholesterolemia screening with subsequent intervention letter on the use of blood pressure and lipid lowering drugs. *British Journal of Clinical Pharmacology* 2004;**57**(3):328–36.

Atthobari 2007 *{published data only}*
Atthobari J, Gansevoort RT, Visser ST, de Jong PE, de Jong-van den Berg LT, PREVEND Study Group. The effect of screening for cardio-renal risk factors on drug use in the general population. *British Journal of Clinical Pharmacology* 2007;**64**(6):810–8.

Bandinelli 2006 *{published data only}*
Bandinelli S, Lauretani F, Boscherini V, Gandi F, Pozzi M, Corsi AM, et al.A randomized, controlled trial of disability prevention in frail older patients screened in primary care: the FRASI study. Design and baseline evaluation. *Aging Clinical and Experimental Research* 2006;**18**(5):359–66.

Barr 2005 *{published data only}*
Barr RJ, Stewart A, Torgerson DJ, Seymour DG, Reid DM. Screening elderly women for risk of future fractures -- participation rates and impact on incidence of falls and fractures. *Calcified Tissue International* 2005;**76**(4):243–8.

Bekwelem 2012 *{published data only}*
Bekwelem W, VanWormer JJ, Boucher JL, Pereira RF. Cardiovascular risk factor screening satisfaction in the Heart of New Ulm Project. *Clinical Medicine & Research* 2012;**10**(1):1–6.

Belcher 1990 *{published data only}*
Belcher DW. Implementing preventive services. Success and failure in an outpatient trial. *Archives of Internal Medicine* 1990;**150**(12):2533–41.

Bonevski 1999 *{published data only}*
Bonevski B, Sanson-Fisher RW, Campbell E, Carruthers A, Reid AL, Ireland M. Randomized controlled trial of a computer strategy to increase general practitioner preventive care. *Preventive Medicine* 1999;**29**(6 Pt 1):478–86.

Boulware 2007 *{published data only}*
Boulware LE, Barnes GJ, Wilson RF, Phillips K, Maynor K, Hwang C, et al.Value of the Periodic Health Evaluation. Rockville, MD: Agency for Healthcare Research and

Quality. 2006 Apr. Evidence Report/Technology Assessment No. 136. AHRQ Publication No. 06-E011. Boulware LE, Marinopoulos S, Phillips KA, Hwang CW, Maynor K, Merenstein D, et al.Systematic review: the value of the periodic health evaluation. *Annals of Internal Medicine* 2007;**146**(4):289–300. [PUBMED: 17310053]

Brown 2009 *{published data only}*
Brown PM, Wilkinson-Meyers L, Parsons M, Weidenbohm K, McNeill R, Brandt T. Cost of prescribed and delivered health services resulting from a comprehensive geriatric assessment tool in New Zealand. *Health & Social Care in the Community* 2009;**17**(5):514–21.

Bula 1999 *{published data only}*
Büla CJ, Bérod AC, Stuck AE, Alessi CA, Aronow HU, Santos-Eggimann B, et al.Effectiveness of preventive in-home geriatric assessment in well functioning, community-dwelling older people: Secondary analysis of a randomized trial. *Journal of the American Geriatrics Society* 1999;**47**(4): 389–95.

Carpenter 1990 *{published data only}*
Carpenter GI, Demopoulos GR. Screening the elderly in the community: controlled trial of dependency surveillance using a questionnaire administered by volunteers. *BMJ* 1990;**300**(6734):1253–6.

Christensen 2003 *{published data only}*
Christensen KS, Toft T, Frostholm L, Ørnbol E, Fink P, Olesen F. The FIP study: a randomised, controlled trial of screening and recognition of psychiatric disorders. *The British Journal of General Practice* 2003;**53**(495):758–63.

Coulton 2008 *{published data only}*
Coulton S, Watson J, Bland M, Drummond C, Kaner E, Godfrey C, et al.The effectiveness and cost-effectiveness of opportunistic screening and stepped care interventions for older hazardous alcohol users in primary care (AESOPS) - a randomised control trial protocol. *BMC Health Services Research* 2008;**8**:129.

Coustasse-Hencke 2000 *{published data only}*
Coustasse-Hencke A, Schutkowski D, Cervantes Y, Rene A. Community intervention to screen risk of cardiovascular diseases in a rural Texas community. *Texas Medicine* 2000; **96**(7):50–5.

Cowan 1992 *{published data only}*
Cowan JA, Heckerling PS, Parker JB. Effect of a fact sheet reminder on performance of the periodic health examination: a randomized controlled trial. *American Journal of Preventive Medicine* 1992;**8**(2):104–9.

Cutchin 2009 *{published data only}*
Cutchin MP, Coppola S, Talley V, Svihula J, Catellier D, Shank KH. Feasibility and effects of preventive home visits for at-risk older people: design of a randomized controlled trial. *BMC Geriatrics* 2009;**9**:54.

Dalby 2000 *{published data only}*
Dalby DM, Sellors JW, Fraser FD, Fraser C, van Ineveld C, Howard M. Effect of preventive home visits by a nurse on the outcomes of frail elderly people in the community: a randomized controlled trial. *CMAJ* 2000;**162**(4):497–500.

Dietrich 2006 *{published data only}*
Dietrich AJ, Tobin JN, Cassells A, Robinson CM, Greene MA, Sox CH, et al.Telephone care management to improve cancer screening among low-income women - a randomized controlled trial. *Annals of Internal Medicine* 2006;**144**(8): 563–71.

Dubey 2006 *{published data only}*
Dubey V, Mathew R, Iglar K, Moineddin R, Glazier R. Improving preventive service delivery at adult complete health check-ups: the Preventive health Evidence-based Recommendation Form (PERFORM) cluster randomized controlled trial. *BMC Family Practice* 2006;**7**:44.

Eekhof 2000 *{published data only}*
Eekhof J, De Bock G, Schaapveld K, Springer M. Effects of screening for disorders among the elderly: an intervention study in general practice. *Family Practice* 2000;**17**(4): 329–33.

Elley 2003 *{published data only}*
Elley CR, Kerse N, Arroll B, Robinson E. Effectiveness of counselling patients on physical activity in general practice: cluster randomised controlled trial. *BMJ* 2003;**326**(7393): 793.

Enzell 1984 *{published data only}*
Enzell K. Mortality among persons with depressive symptoms and among responders and non-responders in a health check-up. An investigation of persons born in 1905 and followed up from age 66 to 75. *Acta Psychiatrica Scandinavica* 1984;**69**(2):89–102.

Fanaian 2010 *{published data only}*
Fanaian M, Laws RA, Passey M, McKenzie S, Wan Q, Davies GP, et al.Health improvement and prevention study (HIPS) - evaluation of an intervention to prevent vascular disease in general practice. *BMC Family Practice* 2010;**11**: 57.

Fang 1999 *{published data only}*
Fang XH, Kronmal RA, Li SC, Longstreth WT Jr, Cheng XM, Wang WZ, et al.Prevention of stroke in urban China: a community-based intervention trial. *Stroke* 1999;**30**(3): 495–501.

Fitzgerald 1991 *{published data only}*
Fitzgerald ST, Gibbens S, Agnew J. Evaluation of referral completion after a workplace cholesterol screening program. *American Journal of Preventive Medicine* 1991;**7**(6):335–40.

Fitzmaurice 2007 *{published data only}*
Fitzmaurice DA, Hobbs FD, Jowett S, Mant J, Murray ET, Holder R, et al.Screening versus routine practice in detection of atrial fibrillation in patients aged 65 or over: cluster randomised controlled trial. *BMJ* 2007;**335**(7616): 383.

Fleischer 2008 *{published data only}*
Fleischer S, Roling G, Beutner K, Hanns S, Behrens J, Luck T, et al.Growing old at home - a randomized controlled trial to investigate the effectiveness and cost-effectiveness of preventive home visits to reduce nursing home admissions: study protocol [NCT00644826]. *BMC Public Health* 2008; **8**:185.

Fletcher 1977 *{published data only}*
Fletcher SW, Sourkes M, Rabzel M, Fletcher RH. Multiphasic screening. Case-finding tool in a Teaching Hospital Medical Clinic. *JAMA* 1977;**237**(9):887–91.

Fletcher 2002 *{published data only}*
Fletcher AE, Jones DA, Bulpitt CJ, Tulloch AJ. The MRC trial of assessment and management of older people in the community: objectives, design and interventions [ISRCTN23494848]. *BMC Health Services Research* 2002;**2**(1):21.

Fox 1997 *{published data only}*
Fox PJ, Breuer W, Wright JA. Effects of a health promotion program on sustaining health behaviors in older adults. *American Journal of Preventive Medicine* 1997;**13**(4):257–64.

Fullard 1987 *{published data only}*
Fullard E, Fowler G, Gray M. Promoting prevention in primary care: controlled trial of low technology, low cost approach. *British Medical Journal* 1987;**294**(6579):1080–2.

Gemson 1990 *{published data only}*
Gemson DH, Sloan RP, Messeri P, Goldberg IJ. A public health model for cardiovascular risk reduction. Impact of cholesterol screening with brief nonphysician counseling. *Archives of Internal Medicine* 1990;**150**(5):985–9.

Gemson 1995 *{published data only}*
Gemson DH, Sloan RP. Efficacy of computerized health risk appraisal as part of a periodic health examination at the worksite. *American Journal of Health Promotion* 1995;**9**(6):462–6.

Giampaoli 1997 *{published data only}*
Giampaoli S, Poce A, Sciarra F, Lo Noce C, Dima F, Minoprio A, et al.Change in cardiovascular risk factors during a 10-year community intervention program. *Acta Cardiologica* 1997;**52**(5):411–22.

Goodwin 2001 *{published data only}*
Goodwin MA, Zyzanski SJ, Zronek S, Ruhe M, Weyer SM, Konrad N, et al.A clinical trial of tailored office systems for preventive service delivery. The Study to Enhance Prevention by Understanding Practice (STEP-UP). *American Journal of Preventive Medicine* 2001;**21**(1):20–8.

Grover 2007 *{published data only}*
Grover SA, Lowensteyn I, Joseph L, Kaouache M, Marchand S, Coupal L, et al.Patient knowledge of coronary risk profile improves the effectiveness of dyslipidemia therapy: the CHECK-UP study: a randomized controlled trial. *Archives of Internal Medicine* 2007;**167**(21):2296–303.

Gysan 2004 *{published data only}*
Gysan DB, Latsch J, Bjarnason-Wehrens B, Albus C, Falkowski G, Herold G, et al.[The PreFord Study. A prospective cohort study to evaluate the risk of a cardiovascular event (overall-collective) as well as a prospective, randomized, controlled, multicentre clinical intervention study (high-risk-collective) on primary prevention of cardiovascular diseases in the Ford Motor Company employees in Germany]. *Zeitschrift für Kardiologie* 2004;**93**(2):131–6.

Hanlon 1995 *{published data only}*
Hanlon P, McEwen J, Carey L, Gilmour H, Tannahill C, Tannahill A, et al.Health checks and coronary risk: further evidence from a randomised controlled trial. *BMJ* 1995;**311**(7020):1609–13.
Kelly MP, Carey L, Hanlon P, Tannahill C, McEwen J, Tannahill A, et al.Do health checks cause stress?. *Occupational Medicine (Oxford, England)* 1999;**49**(5):319–23.

Harding 2009 *{published data only}*
Harding AH, Darnton A, Wegerdt J, McElvenny D. Mortality among British asbestos workers undergoing regular medical examinations (1971-2005). *Occupational and Environmental Medicine* 2009;**66**(7):487–95.

Hay 1998 *{published data only}*
Hay WI, van Ineveld C, Browne G, Roberts J, Bell B, Mills M, et al.Prospective care of elderly patients in family practice. Is screening effective?. *Canadian Family Physician* 1998;**44**:2677–87.

HEART *{published data only}*
McBride P, Underbakke G, Plane MB, Massoth K, Brown RL, Solberg LI, et al.Improving prevention systems in primary care practices: the Health Education and Research Trial (HEART). *The Journal of Family Practice* 2000;**49**(2):115–25.
McBride PE, Massoth KM, Underbakke G, Solberg LI, Beasley JW, Plane MB. Recruitment of private practices for primary care research: experience in a preventive services clinical trial. *The Journal of Family Practice* 1996;**43**(4):389–95.

Heath 1995 *{published data only}*
Heath GW, Fuchs R, Croft JB, Temple SP, Wheeler FC. Changes in blood cholesterol awareness: final results from the South Carolina Cardiovascular Disease Prevention Project. *American Journal of Preventive Medicine* 1995;**11**(3):190–6.

Hellenius 1999 *{published data only}*
Hellénius ML, Johansson J, de Faire U, Elofsson S, Krakau I. Four years experience of a cardiovascular opportunistic screening and prevention programme in the primary health care in Sollentuna, Sweden. *Scandinavian Journal of Primary Health Care* 1999;**17**(2):111–5.

Hendriksen 2005 *{published data only}*
Hendriksen C, Vass M. Preventive home visits to elderly people in Denmark. *Zeitschrift für Gerontologie und Geriatrie* 2005;**38 Suppl 1**:I31–3.

Hiratsuka 2007 *{published data only}*
Hiratsuka VY, Loo R, Will JC, Oberrecht R, Poindexter P. Cardiovascular disease risk factor screening among Alaska Native women: the traditions of the heart project. *International Journal of Circumpolar Health* 2007;**66 Suppl 1**:39–44.

Hogg 1998 *{published data only}*
Hogg WE, Bass M, Calonge N, Crouch H, Satenstein G. Randomized controlled study of customized preventive

medicine reminder letters in a community practice. *Canadian Family Physician* 1998;**44**:81–8.

Huang 2004 *{published data only}*
Huang TT, Acton GJ. Effectiveness of home visit falls prevention strategy for Taiwanese community-dwelling elders: randomized trial. *Public Health Nursing (Boston, Mass.)* 2004;**21**(3):247–56.

Hunter Mellado 1997 *{published data only}*
Hunter Mellado RF, Negron J, Gomez MA. A strategic supportive model for health prevention in the elderly: profile of a Puerto Rican geriatric population in a public health sector. *Boletín de la Asociación Médica de Puerto Rico* 1997;**89**(7-9):102–10.

Hutchison 1998 *{published data only}*
Hutchison B, Birch S, Evans CE, Goldsmith LJ, Markham BA, Frank J, et al.Screening for hypercholesterolaemia in primary care: randomised controlled trial of postal questionnaire appraising risk of coronary heart disease. *BMJ* 1998;**316**(7139):1208–13.

IMPROVE *{published data only}*
Kottke TE, Solberg LI, Brekke ML, Cabrera A, Marquez MA. Delivery rates for preventive services in 44 midwestern clinics. *Mayo Clinic Proceedings* 1997;**72**(6):515–23.
Kottke TE, Solberg LI, Brekke ML, Magnan S, Amundson GM. Clinician satisfaction with a preventive services implementation trial. The IMPROVE project. *American Journal of Preventive Medicine* 2000;**18**(3):219–24.
Magnan S, Solberg LI, Kottke TE, Nelson AF, Amundson GM, Richards S, et al.IMPROVE: bridge over troubled waters. *The Joint Commission Journal on Quality Improvement* 1998;**24**(10):566–78.

Jilcott 2006 *{published data only}*
Jilcott SB, Keyserling TC, Samuel-Hodge CD, Rosamond W, Garcia B, Will JC, et al.Linking clinical care to community resources for cardiovascular disease prevention: the North Carolina Enhanced WISEWOMAN project. *Journal of Women's Health (2002)* 2006;**15**(5):569–83.

Johansson 1999 *{published data only}*
Johansson J, Hellénius ML, Elofsson S, Krakau I. Self-report as a selection instrument in screening for cardiovascular disease risk. *American Journal of Preventive Medicine* 1999; **16**(4):322–4.

Johansson 2002 *{published data only}*
Johansson J, Viigimaa M, Jensen-Urstad M, Krakau I, Hansson LO. Risk factors for coronary heart disease in 55- and 35-year-oldmen and women in Sweden and Estonia. *Journal of Internal Medicine* 2002;**252**(6):551–60.

Kaczorowski 2011 *{published data only}*
Kaczorowski J, Goldberg O, Mai V. Pay-for-performance incentives for preventive care: views of family physicians before and after participation in a reminder and recall project (P-PROMPT). *Canadian Family Physician* 2011;**57** (6):690–6.

Kahler 1978 *{published data only}*
Kahler HD, Mehne P. [Medical check-ups and selection of subjects: some results from a gerontological field-study]. *Aktuelle Gerontologie* 1978;**8**(10):513–22.

Kneipp 2011 *{published data only}*
Kneipp SM, Kairalla JA, Lutz BJ, Pereira D, Hall AG, Flocks J, et al.Public health nursing case management for women receiving temporary assistance for needy families: a randomized controlled trial using community-based participatory research. *American Journal of Public Health* 2011;**101**(9):1759–68.

Knutsen 1991 *{published data only}*
Knutsen SF, Knutsen R. The Tromso Survey: the Family Intervention study--the effect of intervention on some coronary risk factors and dietary habits, a 6-year follow-up. *Preventive Medicine* 1991;**20**(2):197–212.

Koinberg 2003 *{published data only}*
Koinberg I, Fridlund B, Engholm GB, Holmberg L. Comparison between nurse-led check-ups on demand and follow-ups by a physician after breast cancer surgery [abstract]. *European Journal of Cancer* 2003;**1**(5):S362.

Koivisto 1992 *{published data only}*
Koivisto K, Helkala EL, Reinikainen KJ, Hänninen T, Mykkänen L, Laakso M, et al.Population-based dementia screening program in Kuopio: the effect of education, age, and sex on brief neuropsychological tests. *Journal of Geriatric Psychiatry & Neurology* 1992;**5**(3):162–71.

Kolbe-Alexander 2008 *{published data only}*
Kolbe-Alexander TL, Buckmaster C, Nossel C, Dreyer L, Bull F, Noakes TD, et al.Chronic disease risk factors, healthy days and medical claims in South African employees presenting for health risk screening. *BMC Public Health* 2008;**8**:228.

Kolozsi 1982 *{published data only}*
Kolozsi B, Petrov I, Bartók J, Dumeva N. Self evaluation of aged people who voluntarily required preventive periodical gerontological examination. *Archives of Gerontology and Geriatrics* 1982;**1**(2):117–24.

Kono 2004 *{published data only}*
Kono A, Kai I, Sakato C, Harker JO, Rubenstein LZ. Effect of preventive home visits for ambulatory housebound elders in Japan: a pilot study. *Aging Clinical and Experimental Research* 2004;**16**(4):293–9.

Kono 2009 *{published data only}*
Kono A, Fujita T, Tsumura C, Kondo T, Kushiyama K, Rubenstein LZ. Preventive home visit model targeted to specific care needs of ambulatory frail elders: preliminary report of a randomized trial design. *Aging Clinical and Experimental Research* 2009;**21**(2):167–73.

Kowal 1979 *{published data only}*
Kowal DM, Daniels WL. Recommendations for the screening of military personnel over 35 years of age for physical training programs. *The American Journal of Sports Medicine* 1979;**7**(3):186–90.

Kurata 2006 *{published data only}*
Kurata C. Medical check-up findings characteristic of smokers: aimed at improving smoking cessation interventions by physicians. *Internal Medicine (Tokyo, Japan)* 2006;**45**(18):1027–32.

Landi 2001 *{published data only}*
Landi F, Onder G, Tua E, Carrara B, Zuccala G, Gambassi G, et al.Impact of a new assessment system, the MDS-HC, on function and hospitalization of homebound older people: a controlled clinical trial. *Journal of the American Geriatrics Society* 2001;**49**(10):1288–93.

Lauritzen 2011 *{published data only}*
Lauritzen T, Sandbaek A, Skriver MV, Borch-Johnsen K. HbA1c and cardiovascular risk score identify people who may benefit from preventive interventions: a 7 year follow-up of a high-risk screening programme for diabetes in primary care (ADDITION), Denmark. *Diabetologia* 2011; **54**(6):1318–26.

Lave 1995 *{published data only}*
Lave JR, Ives DG, Traven ND, Kuller LH. Participation in health promotion programs by the rural elderly. *American Journal of Preventive Medicine* 1995;**11**(1):46–53.

Liaw 1996 *{published data only}*
Liaw T, Lawrence M, Rendell J. The effect of a computer-generated patient-held medical record summary and/or a written personal health record on patients' attitudes, knowledge and behaviour concerning health promotion. *Family Practice* 1996;**13**(3):289–93.

Manning 1984 *{published data only}*
Manning WG, Leibowitz A, Goldberg GA, Rogers WH, Newhouse JP. A controlled trial of the effect of a prepaid group practice on use of services. *The New England Journal of Medicine* 1984;**310**(23):1505–10.

Marshall 2008 *{published data only}*
Marshall T, Westerby P, Chen J, Fairfield M, Harding J, Westerby R, et al.The Sandwell Project: a controlled evaluation of a programme of targeted screening for prevention of cardiovascular disease in primary care. *BMC Public Health* 2008;**8**:73.

Mathews 2007 *{published data only}*
Mathews G, Alexander J, Rahemtulla T, Bhopal R. Impact of a cardiovascular risk control project for South Asians (Khush Dil) on motivation, behaviour, obesity, blood pressure and lipids. *Journal of Public Health (Oxford, England)* 2007;**29**(4):388–97.

Maxwell 1983 *{published data only}*
Maxwell VB, Crump JH, Thorp J. The measurement of risk indicators for coronary heart disease in air traffic control officers: A screening study in a healthy population. *Aviation, Space, and Environmental Medicine* 1983;**54**(3): 246–9.

McMahon 2002 *{published data only}*
McMahon A, Hodgins M, Kelleher CC. Feasibility of a men's health promotion programme in Irish primary care. *Irish Journal of Medical Science* 2002;**171**(1):20–3.

Medicare - Baltimore *{published data only}*
Burton LC, Paglia MJ, German PS, Shapiro S, Damiano AM. The effect among older persons of a general preventive visit on three health behaviors: smoking, excessive alcohol drinking, and sedentary lifestyle. *Preventive Medicine* 1995; **24**(5):492–97.
Burton LC, Steinwachs DM, German PS, Shapiro S, Brant LJ, Richards TM, et al.Preventive services for the elderly: would coverage affect utilization and costs under Medicare?. *American Journal of Public Health* 1995;**85**(3):387–91.
German PS, Burton LC, Shapiro S, Steinwachs DM, Tsuji I, Paglia MJ, et al.Extended coverage for preventive services for the elderly: response and results in a demonstration population. *American Journal of Public Health* 1995;**85**(3): 379–86.

Meland 1996 *{published data only}*
Meland E, Laerum E, Maeland JG. Life style intervention in general practice: effects on psychological well-being and patient satisfaction. *Quality of Life Research* 1996;**5**(3): 348–54.
Meland E, Laerum E, Ulvik RJ. Effectiveness of two preventive interventions for coronary heart disease in primary care. *Scandinavian Journal of Primary Health Care* 1997;**15**(1):57–64.

Melis 2005 *{published data only}*
Melis RJ, van Eijken MI, Borm GF, Wensing M, Adang E, van De Lisdonk EH, et al.The design of the Dutch EASYcare study: a randomised controlled trial on the effectiveness of a problem-based community intervention model for frail elderly people [NCT00105378]. *BMC Health Services Research* 2005;**5**:65.

Milisen 2006 *{published data only}*
Milisen K, Dejaeger E, Braes T, Dierickx K, De Bondt K, Smeulders W, et al.Process evaluation of a nurse-led multifactorial intervention protocol for risk screening and assessment of fall problems among community-dwelling older persons: a pilot-study. *The Journal of Nutrition, Health & Aging* 2006;**10**(5):446–52.

Minder 2002 *{published data only}*
Minder CE, Müller T, Gillmann G, Beck JC, Stuck AE. Subgroups of refusers in a disability prevention trial in older adults: baseline and follow-up analysis. *American Journal of Public Health* 2002;**92**(3):445–50.

Mitchell 2005 *{published data only}*
Mitchell E, Sullivan F, Grimshaw JM, Donnan PT, Watt G. Improving management of hypertension in general practice: a randomised controlled trial of feedback derived from electronic patient data. *The British Journal of General Practice* 2005;**55**(511):94–101.

Morrissey 1995 *{published data only}*
Morrissey JP, Harris RP, Kincade-Norburn J, McLaughlin C, Garrett JM, Jackman AM, et al.Medicare reimbursement for preventive care. Changes in performance of services, quality of life, and health care costs. *Medical Care* 1995;**33** (4):315–31.

Naor 1975 *{published data only}*
Naor EM. Early and periodic screening, diagnosis and treatment. A report on the Dane County pilot program, 1973. Wisconsin Department of Health and Social Services 1975.

Näyhä 1988 *{published data only}*
Näyhä S, Laakso N, Hassi J, Kolivuori T. A health examination trial among males in northern Finland. *Arctic Medical Research* 1988;**47 Suppl 1**:442–4.

Newcomer 2004 *{published data only}*
Newcomer R, Maravilla V, Faculjak P, Graves MT. Outcomes of preventive case management among high-risk elderly in three medical groups: a randomized clinical trial. *Evaluation & the Health Professions* 2004;**27**(4):323–48.

Nicolaides-Bouman 2004 *{published data only}*
Nicolaides-Bouman A, van Rossum E, Kempen GI, Knipschild P. Effects of home visits by home nurses to elderly people with health problems: design of a randomised clinical trial in the Netherlands [ISRCTN92017183]. *BMC Health Services Research* 2004;**4**(1):35.

O'Malley 2006 *{published data only}*
O'Malley PG, Kowalczyk C, Bindeman J, Taylor AJ. The impact of cardiovascular risk factor case management on the metabolic syndrome in a primary prevention population: results from a randomized controlled trial. *Journal of the Cardiometabolic Syndrome* 2006;**1**(1):6–12.

O'Rourke 1970 *{published data only}*
O'Rourke A, Sheehan B, Gowen J, Wilson K, O'Sullivan N. Screening--a demonstration project. *Journal of the Irish Medical Association* 1970;**63**(394):231–7.

Parker 2005 *{published data only}*
Parker DR, Evangelou E, Eaton CB. Intraclass correlation coefficients for cluster randomized trials in primary care: the cholesterol education and research trial (CEART). *Contemporary Clinical Trials* 2005;**26**(2):260–7.

Phelan 2007 *{published data only}*
Phelan EA, Balderson B, Levine M, Erro JH, Jordan L, Grothaus L, et al.Delivering effective primary care to older adults: a randomized, controlled trial of the senior resource team at group health cooperative. *Journal of the American Geriatrics Society* 2007;**55**(11):1748–56.

Reid 1995 *{published data only}*
Reid C, McNeil JJ, Williams F, Powles J. Cardiovascular risk reduction: a randomized trial of two health promotion strategies for lowering risk in a community with low socioeconomic status. *Journal of Cardiovascular Risk* 1995;**2**(2):155–63.

Reijneveld 2003 *{published data only}*
Reijneveld SA, Westhoff MH, Hopman-Rock M. Promotion of health and physical activity improves the mental health of elderly immigrants: results of a group randomised controlled trial among Turkish immigrants in the Netherlands aged 45 and over. *Journal of Epidemiology and Community Health* 2003;**57**(6):405–11.

Ritzau 1969 *{published data only}*
Ritzau LC. Health screening examinations. A pilot study. *Occupational Health Nursing* 1969;**17**(1):15–6.

Robertson 1992 *{published data only}*
Robertson I, Phillips A, Mant D, Thorogood M, Fowler G, Fuller A, et al.Motivational effect of cholesterol measurement in general practice health checks. *The British Journal of General Practice* 1992;**42**(364):469–72.

Rodondi 2008 *{published data only}*
Rodondi N, Bovet P, Hayoz D, Cornuz J. The Impact of CAROtid plaque Screening on Smoking (CAROSS) cessation and control of other cardiovascular risk factors: Rationale and design of a randomized controlled trial. *Contemporary Clinical Trials* 2008;**29**(5):767–73.

Romundstad 2003 *{published data only}*
Romundstad S, Holmen J, Kvenild K, Aakervik O, Hallan H. Clinical relevance of microalbuminuria screening in self-reported non-diabetic/non-hypertensive persons identified in a large health screening--the Nord-Trøndelag Health Study (HUNT), Norway. *Clinical Nephrology* 2003;**59**(4):241–51.

Rose 1996 *{published data only}*
Rose G, Bengtsson C. Effects of a health examination programme on quality of life and subjective well-being. *Scandinavian Journal of Social Medicine* 1996;**24**(2):124–31.

Rubenstein 2007 *{published data only}*
Rubenstein LZ, Alessi CA, Josephson KR, Trinidad Hoyl M, Harker JO, Pietruszka FM. A randomized trial of a screening, case finding, and referral system for older veterans in primary care. *Journal of the American Geriatrics Society* 2007;**55**(2):166–74.

Sackett 1974 *{published data only}*
Sackett DL, Spitzer WO, Gent M, Roberts RS. The Burlington randomized trial of the nurse practitioner: health outcomes of patients. *Annals of Internal Medicine* 1974;**80**(2):137–42.

Sahlen 2006 *{published data only}*
Sahlen KG, Dahlgren L, Hellner BM, Stenlund H, Lindholm L. Preventive home visits postpone mortality--a controlled trial with time-limited results. *BMC Public Health* 2006;**6**:220.
Sahlen KG, Löfgren C, Mari Hellner B, Lindholm L. Preventive home visits to older people are cost-effective. *Scandinavian Journal of Public Health* 2008;**36**(3):265–71.

Schweitzer 1994 *{published data only}*
Schweitzer SO, Atchison KA, Lubben JE, Mayer-Oakes SA, De Jong FJ, Matthias RE. Health promotion and disease prevention for older adults: opportunity for change or preaching to the converted?. *American Journal of Preventive Medicine* 1994;**10**(4):223–9.

Shannon 2001 *{published data only}*
Shannon KC, Sinacore JM, Bennett SG, Joshi AM, Sherin KM, Deitrich A. Improving delivery of preventive health care with the comprehensive annotated reminder tool (CART). *The Journal of Family Practice* 2001;**50**(9):767–71.

Snow 1989 *{published data only}*
Snow SG, Calder EA, Taylor LS, Lane TJ, Manu P, Federici CM. Screening for cancer and coronary risk factors through a nurse practitioner-staffed preventive health clinic. *Preventive Medicine* 1989;**18**(6):817–23.

Stange 2003 *{published data only}*
Stange KC, Goodwin MA, Zyzanski SJ, Dietrich AJ. Sustainability of a practice-individualized preventive service delivery intervention. *American Journal of Preventive Medicine* 2003;**25**(4):296–300.

Strandberg 1991 *{published data only}*
Strandberg TE, Salomaa VV, Naukkarinen VA, Vanhanen HT, Sarna SJ, Miettinen TA. Long-term mortality after 5-year multifactorial primary prevention of cardiovascular diseases in middle-aged men. *JAMA* 1991;**266**(9):1225–9.

Stuck 2000 *{published data only}*
Stuck AE, Minder CE, Peter-Wüest L, Gillmann G, Egli C, Kesselring A, et al.A randomized trial of in-home visits for disability prevention in community-dwelling older people at low and high risk for nursing home admission. *Archives of Internal Medicine* 2000;**160**(7):977–86.

Stuck 2007 *{published data only}*
Stuck AE, Kharicha K, Dapp U, Anders J, von Renteln-Kruse W, Meier-Baumgartner HP, et al.The PRO-AGE study: an international randomised controlled study of health risk appraisal for older persons based in general practice. *BMC Medical Research Methodology* 2007;**7**:2.

Taylor 1998 *{published data only}*
Taylor AH, Doust J, Webborn N. Randomised controlled trial to examine the effects of a GP exercise referral programme in Hailsham, East Sussex, on modifiable coronary heart disease risk factors. *Journal of Epidemiology and Community Health* 1998;**52**(9):595–601.

Toth-Pal 2004 *{published data only}*
Toth-Pal E, Nilsson GH, Furhoff AK. Clinical effect of computer generated physician reminders in health screening in primary health care--a controlled clinical trial of preventive services among the elderly. *International Journal of Medical Informatics* 2004;**73**(9-10):695–703.

Tulloch 1979 *{published data only}*
Tulloch AJ, Moore V. A randomized controlled trial of geriatric screening and surveillance in general practice. *The Journal of the Royal College of General Practitioners* 1979;**29**(209):733–40.

Turner 1989 *{published data only}*
Turner BJ, Day SC, Borenstein B. A controlled trial to improve delivery of preventive care: physician or patient reminders?. *Journal of General Internal Medicine* 1989;**4**(5):403–9.

van Haastregt 2000 *{published data only}*
van Haastregt JCM, Diederiks JP, van Rossum E, de Witte LP, Crebolder HF. Effects of preventive home visits to elderly people living in the community: Systematic review. *BMJ* 2000;**320**(7237):754–8.

van Hout 2010 *{published data only}*
van Hout HP, Jansen AP, van Marwijk HW, Pronk M, Frijters DF, Nijpels G. Prevention of adverse health trajectories in a vulnerable elderly population through nurse home visits: a randomized controlled trial [ISRCTN05358495]. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 2010;**65**(7):734–42.
van Hout HP, Nijpels G, van Marwijk HW, Jansen AP, Van't Veer PJ, Tybout W, et al.Design and pilot results of a single blind randomized controlled trial of systematic demand-led home visits by nurses to frail elderly persons in primary care [ISRCTN05358495]. *BMC Geriatrics* 2005;**5**:11.

van Rossum 1993 *{published data only}*
van Rossum E, Frederiks C, Philipsen H, Kil-van Lierop J, Mantel A, Portengen J, et al.Design of a Dutch study to test preventive home visits to the elderly. *Nursing Research* 1991;**40**(3):185–8.
van Rossum E, Frederiks CM, Philipsen H, Portengen K, Wiskerke J, Knipschild P. Effects of preventive home visits to elderly people. *BMJ* 1993;**307**(6895):27–32.

van Weel 2006 *{published data only}*
van Weel C, Bakx C, van den Hoogen H, Thien T, van den Bosch W. Long-term outcome of cardiovascular prevention: a Nijmegen Academic Family Practices Network study. *Journal of the American Board of Family Medicine* 2006;**19**(1):62–8.

Vass 2007 *{published data only}*
Kronborg C, Vass M, Lauridsen J, Avlund K. Cost effectiveness of preventive home visits to the elderly: economic evaluation alongside randomized controlled study. *The European Journal of Health Economics* 2006;**7**(4):238–46.
Vass M, Avlund K, Hendriksen C. Randomized intervention trial on preventive home visits to older people: baseline and follow-up characteristics of participants and non-participants. *Scandinavian Journal of Public Health* 2007;**35**(4):410–7.
Vass M, Avlund K, Hendriksen C, Andersen CK, Keiding N. Preventive home visits to older people in Denmark: methodology of a randomized controlled study. *Aging Clinical and Experimental Research* 2002;**14**(6):509–15.

Waldman 1970 *{published data only}*
Waldman A, Ward M, Yaffe S. Periodic health examinations and mortality in the elderly. *Aging and Human Development* 1970;**1**(4):323–32.
Waldman A, Ward M, Yaffe S. Periodic health examinations and mortality in the elderly. *The Gerontologist* 1970;**10**(3 Pt 2):30–1.

Weinehall 1999 *{published data only}*
Weinehall L, Westman G, Hellsten G, Boman K, Hallmans G, Pearson TA, et al.Shifting the distribution of risk: results of a community intervention in a Swedish programme for the prevention of cardiovascular disease. *Journal of Epidemiology and Community Health* 1999;**53**(4):243–50.

Williams 1997 *{published data only}*
Williams SJ, Elder JP, Seidman RL, Mayer JA. Preventive services in a Medicare managed care environment. *Journal of Community Health* 1997;**22**(6):417–34.

Williams 1998 *{published data only}*
Williams RB, Boles M, Johnson RE. A patient-initiated system for preventive health care. A randomized trial in community-based primary care practices. *Archives of Family Medicine* 1998;**7**(4):338–45.

References to studies awaiting assessment

Brett 2012 *{published data only}*
Brett T, Arnold-Reed D, Phan C, Cadden F, Walker W, Manea-Walley W, et al.The Fremantle Primary Prevention Study: a multicentre randomised trial of absolute cardiovascular risk reduction. *British Journal of General Practice* 2012;**62**(594):e22–8.

Stickler 2000 *{published data only}*
Stickler GB. Are yearly physical examinations in adolescents necessary?. *Journal of the American Board of Family Practice* 2000;**13**(3):172–7.

Additional references

Barger 2006
Barger SD, Muldoon MF. Hypertension labelling was associated with poorer self-rated health in the Third US National Health and Nutrition Examination Survey. *Journal of Human Hypertension* 2006;**20**(2):117–23. [PUBMED: 16267563]

Ben-Schlomo 1994
Ben-Shlomo Y, Smith GD, Shipley MJ, Marmot MG. What determines mortality risk in male former cigarette smokers?. *American Journal of Public Health* 1994;**84**(8):1235–42.

Beswick 2008
Beswick AD, Rees K, Dieppe P, Ayis S, Gooberman-Hill R, Horwood J, et al.Complex interventions to improve physical function and maintain independent living in elderly people: a systematic review and meta-analysis. *Lancet* 2008;**371**(9614):725–35.

Bloom 1981
Bloom JR, Monterossa S. Hypertension labeling and sense of well-being. *American Journal of Public Health* 1981;**71**(11):1228–32.

Boulware 2006
Boulware LE, Barnes GJ, Wilson RF, Phillips K, Maynor K, Hwang C, et al.Value of the Periodic Health Evaluation. Rockville, MD: Agency for Healthcare Research and Quality. 2006 Apr. Evidence Report/Technology Assessment No. 136. AHRQ Publication No. 06-E011.

Brewer 2007
Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Annals of Internal Medicine* 2007;**146**(7):502–10. [PUBMED: 17404352]

Collins 2011
Collins RE, Lopez LM, Marteau TM. Emotional impact of screening: a systematic review and meta-analysis. *BMC Public Health* 2011;**11**:603.

Cook 1986
Cook DG, Shaper AG, Pocock SJ, Kussick SJ. Giving up smoking and the risk of heart attacks. *Lancet* 1986;**2**(8520):1376–80.

D'Souza 1976
D'Souza MF, Swan AV, Shannon DJ. A long-term controlled trial of screening for hypertension in general practice. *Lancet* 1976;**1**(7971):1228–31.

Dales 1979
Dales LG, Friedman GD, Collen MF. Evaluating periodic multiphasic health checkups: a controlled trial. *Journal of Chronic Diseases* 1979;**32**(5):385–404.

Djulbegovic 2010
Djulbegovic M, Beyth RJ, Neuberger MM, Stoffs TL, Vieweg J, Djulbegovic B, et al.Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2010;**341**:c4543. [DOI: 10.1136/bmj.c4543]

Ebrahim 2011
Ebrahim S, Taylor F, Ward K, Beswick A, Burke M, Davey Smith G. Multiple risk factor interventions for primary prevention of coronary heart disease. *Cochrane Database of Systematic Reviews* 2011, Issue 1. [DOI: 10.1002/14651858.CD001561.pub3]

Engberg 2002c
Engberg M, Christensen B, Karlsmose B, Lauritzen T. Re: [Should general health surveys be offered to the population?]. *Ugeskrift for Laeger* 2002;**164**(36):4200–1.

FDA 2011
U.S. Food, Drug Administration. Full-body CT scans - what you need to know. http://www.fda.gov/Radiation-EmittingProducts/RadiationEmittingProductsandProcedures/MedicalImaging/MedicalX-Rays/ucm115340.htm (Accessed 28 December 2011).

Fleming 1996
Fleming TR, DeMets DL. Surrogate end points in clinical trials - are we being misled?. *Annals of Internal Medicine* 1996;**125**(7):605–13.

Friedman 1986
Friedman GD, Collen MF, Fireman BH. Multiphasic Health Checkup Evaluation: a 16-year follow-up. *Journal of Chronic Diseases* 1986;**39**(6):453–63.

Furtado 2005
Furtado CD, Aguirre DA, Sirlin CB, Dang D, Stamato SK, Lee P, et al.Whole-body CT screening: spectrum of findings and recommendations in 1192 patients. *Radiology* 2005;**237**(2):385–94.

Golomb 2012
Golomb BA, Evans MA, Dimsdale JE, White HL. Effects of statins on energy and fatigue with exertion: results from

a randomized controlled trial. *Archives of Internal Medicine* 2012;**172**(15):1180–2.

Grønhøj Larsen 2012
Grønhøj Larsen C, Jørgensen KJ, Gøtzsche PC. Regular health checks: cross sectional survey. *PloS One* 2012;**7**(3): e33694.

Han 1997
Han PK. Historical changes in the objectives of the periodic health examination. *Annals of Internal Medicine* 1997;**127** (10):910–7.

Haynes 1978
Haynes RB, Sackett DL, Taylor DW, Gibson ES, Johnson AL. Increased absenteeism from work after detection and labelling of hypertensive patients. *The New England Journal of Medicine* 1978;**299**(14):741–4. [PUBMED: 692548]

Healy 2004
Healy D. *Let them eat Prozac*. New York and London: New York University Press, 2004.

Hewitson 2007
Hewitson P, Glasziou P, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. *Cochrane Database of Systematic Reviews* 2007, Issue 1. [DOI: 10.1002/ 14651858.CD001216.pub2]

Higgins 2011
Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Holland 2009
Holland WW. Periodic health examination - a brief history and critical assessment. *Eurohealth* 2009;**15**(4):16–20.

Hooper 2011
Hooper L, Summerbell CD, Thompson R, Sills D, Roberts FG, Moore H, et al.Reduced or modified dietary fat for preventing cardiovascular disease. *Cochrane Database of Systematic Reviews* 2012, Issue 5. [DOI: 10.1002/ 14651858.CD002137.pub3]

Jørgensen 2003
Jørgensen T, Borch-Johnsen K, Thomsen TF, Ibsen H, Glümer C, Pisinger C. A randomized non-pharmacological intervention study for prevention of ischaemic heart disease: baseline results Inter99. *European Journal of Cardiovascular Prevention and Rehabilitation* 2003;**10**(5):377–86.

Lauritzen 2012
Lauritzen T. [Helbredsundersøgelser og samtaler reducerer deltagernes risikoscore for hjerte–kar–sygdom]. *Månedsskrift for Almen Praksis* 2012;**90**(3):245–50.

Lehman 2010
Lehman R, Yudkin JS, Krumholz H. Licensing drugs for diabetes. *BMJ* 2010;**341**:c4805.

Nissen 2010
Nissen SE, Wolski K. Rosiglitazone revisited. An updated meta-analysis of risk for myocardial infarction and cardiovascular mortality. *Archives of Internal Medicine* 2010; **170**(14):1191–201.

Norris 2008
Norris SL, Kansagara D, Bougatsos C, Fu R, U.S. Preventive Services Task Force. Screening adults for type 2 diabetes: a review of the evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine* 2008;**148**(11):855–68.

Patrick 1999
Patrick DL, Grembowski D, Durham M, Beresford SA, Diehr P, Ehreth J, et al.Cost and outcomes of Medicare reimbursement for HMO preventive services. *Health Care Financing Review* 1999;**20**(4):25–43.

Pill 1988
Pill R, French J, Harding K, Stott N. Invitation to attend a health check in a general practice setting: comparison of attenders and non-attenders. *The Journal of the Royal College of General Practitioners* 1988;**38**(307):53–6.

Pisinger 2009
Pisinger C, Ladelund S, Glümer C, Toft U, Aadahl M, Jørgensen T. Five years of lifestyle intervention improved self-reported mental and physical health in a general population: the Inter99 study. *Preventive Medicine* 2009;**49** (5):424–8.

Raffle & Gray 2007
Raffle A, Gray M. *Screening: evidence and practice*. New York: Oxford University Press, 2007.

Shaw 1999
Shaw C, Abrams K, Marteau TM. Psychological impact of predicting individuals' risks of illness: a systematic review. *Social Science & Medicine* 1999;**49**(12):1571–98.

Sheridan 2003
Sheridan S, Pignone M, Donahue K. Screening for high blood pressure: a review of the evidence for the U.S. Preventive Services Task Force. *American Journal of Preventive Medicine* 2003;**25**(2):151–8. [PUBMED: 12880884]

Sheridan 2008
Sheridan SL, Crespo E. Does the routine use of global coronary heart disease risk scores translate into clinical benefits or harms? A systematic review of the literature. *BMC Health Services Research* 2008;**8**:60. [DOI: 10.1186/ 1472-6963-8-60]

Sheridan 2010
Sheridan SL, Viera AJ, Krantz MJ, Ice CL, Steinman LE, Peters KE, et al.Cardiovascular Health Intervention Research and Translation Network Work Group on Global Coronary Heart Disease Risk. The effect of giving global coronary risk information to adults: a systematic review. *Archives of Internal Medicine* 2010;**170**(3):230–9.

Singh 2011
Singh S, Loke YK, Enright PL, Furberg CD. Mortality associated with tiotropium mist inhalers in patients with chronic obstructive pulmonary disease: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2011;**342**:d3215. [DOI: 10.1136/bmj.d3215]

South-East London Study Group 1977
The South-East London Study Group. A controlled trial of multiphasic screening in middle-age: results of the South-East London screening study. *International Journal of Epidemiology* 1977;**6**(4):357–63.

Stead 2008
Stead LF, Bergson G, Lancaster T. Physician advice for smoking cessation. *Cochrane Database of Systematic Reviews* 2008, Issue 2. [DOI: 10.1002/14651858.CD000165.pub3]

Stone 1981
Stone D, D'Souza MF. Multiphasic screening in middle age: results and implications of a controlled trial in British general practice. *Israel Journal of Medical Sciences* 1981;**17**(2-3):215–21.

Taylor 2011
Taylor F, Ward K, Moore TH, Burke M, Davey Smith G, Casas JP, et al.Statins for the primary prevention of cardiovascular disease. *Cochrane Database of Systematic Reviews* 2011, Issue 1. [DOI: 10.1002/14651858.CD004816.pub4]

Trevelyan 1973
Trevelyan H. Study to evaluate the effects of multiphasic screening within general practice in Britain: design and method. *Preventive Medicine* 1973;**2**(2):278–94.

Uijen 2007
Uijen AA, Bakx JC, Mokkink HG, van Weel C. Hypertension patients participating in trials differ in many aspects from patients treated in general practices. *Journal of Clinical Epidemiology* 2007;**60**(4):330–5.

UK National Screening Committee 2010
UK National Screening Committee. Programme appraisal criteria. http://www.screening.nhs.uk/criteria (accessed 3 August 2012).

Waller 1990
Waller D, Agass M, Mant D, Coulter A, Fuller A, Jones L. Health checks in general practice: another example of inverse care?. *BMJ* 1990;**300**(6732):1115–8.

Welch 2004
Welch, H. Gilbert. *Should I be tested for cancer? Maybe not and here's why*. Berkeley: University of California Press, 2004.

Welch, Schwartz and Woloshin 2011
Gilbert Welch H, Schwartz LM, Woloshin S. *Overdiagnosed - Making people sick in the pursuit of health*. Boston: Beacon Press, 2011.

WHO 1984
WHO cooperative trial on primary prevention of ischaemic heart disease with clofibrate to lower serum cholesterol: final mortality follow-up. Report of the Committee of Principal Investigators. *Lancet* 1984;**2**(8403):600–4.

Wright 2009
Wright JM, Musini VM. First-line drugs for hypertension. *Cochrane Database of Systematic Reviews* 2009, Issue 3. [DOI: 10.1002/14651858.CD001841.pub2]

Xiong 2005
Xiong T, Richardson M, Woodroffe R, Halligan S, Morton D, Lilford RJ. Incidental lesions found on CT colonography: their nature and frequency. *The British Journal of Radiology* 2005;**78**(925):22–9.

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Ebeltoft 1992

Methods	A random sample (n=2000) was taken from the whole eligible population (n=3464). The sample was sent a short questionnaire, and participants returning the questionnaire and giving consent (n=1507) were included and randomised into three groups. One group was offered screening (n=502), another group was offered screening plus health discussions (n=504), and the third group had usual care (n=501). All included participants were sent a more detailed questionnaire before the intervention. The intervention was repeated after one year. After five years all three groups were mailed questionnaires and invited for a follow-up screening. Participants were also followed in national registers for eight years and two comparisons were made: 1) between the three intervention groups and 2) between the 2000 randomly invited to participate in the trial (plus 30 in whom invitation failed for administrative reasons) and the 1434 not invited
Participants	Men and women aged 30-49 years identified through practice registers Setting: general practice Location: Ebeltoft, Denmark Number randomised: See above.
Interventions	Screening included the following: myocardial infarction risk score (Anggaard) electrocardiogram total cholesterol diastolic blood pressure systolic blood pressure spirometry (FEV, vital capacity, FEV/forced vital capacity) liver tests (gamma glutamyl transpeptidase, aspartate aminotransferase, alanine aminotransferase) creatinine non-fasting blood glucose serum urate urinary dipstick (glucose, albumin, blood) body mass index waist/hip ratio CO concentration in expiratory air physical endurance sight (Snellen test) hearing (screening audiometer) HIV status Participants randomised to additional health discussions were invited to annual 45 minute health talks with their physician regarding lifestyle changes. Participants randomised to screening only were sent a personalised letter explaining the findings and giving recommendations Uptake of screening: first round 90%, second round 81% - 83%

Outcomes	mortality physician visits hospitalisation referrals worry self-reported health
Notes	The screening and the screening + health discussion groups were combined in the reports, as there were no differences in outcomes

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "An Aarhus County statistician performed invitation and intervention randomization by computer, independently of the investigators."
Allocation concealment (selection bias)	Low risk	All participants were allocated at once, independently of the investigators
Blinding of participants and personnel (performance bias) All outcomes	High risk	Lack of blinding of general practitioners and control group may have led to performance bias
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The most important outcomes were assessed using register data and were not subject to detection bias. Self-reported outcomes (self-reported health, worry, medication use) may have been biased by the absence of blinding, and is an exception to the overall rating
Incomplete outcome data (attrition bias) Objective outcomes	Low risk	Public registries were used with 5% loss to follow-up. Characteristics of participants lost were similar between groups
Incomplete outcome data (attrition bias) Subjective outcomes	High risk	Loss to follow-up was between 24% and 31%, which indicate a high risk of bias in the context of an unblinded trial
Selective reporting (reporting bias)	Unclear risk	Unclear.
Other bias	High risk	All participants had returned an initial questionnaire, which limits external validity because non-respondents were not included in some of the analyses. The trial

		was set in a small town, and the authors have reported that the trial had a great influence on the control group

Family Heart 1990

Methods	Thirteen matched pairs of general practices were randomised to either intervention or control (external control group). In the intervention practices, eligible men were randomised to either intervention or control (internal control group) and their partners were included. The intervention group was invited for screening and lifestyle intervention at baseline. After one year both intervention and control groups were invited. Only those participants who attended their first health check were included in the analyses, i.e. at baseline for the intervention group and after one year for the control group
Participants	Men aged 40-59 years, and their partners, regardless of age. Setting: general practice Location: UK Number randomised: Not clear. Only the number of households and persons who attended screening are given. The number of people in each group were 3436 (screening), 3576 (internal control) and 5912 (external control). The number of households in each group was 2373 (screening), 2342 (internal control) and 3890 (external control), with a response rate of 73% (adjusted for 'ghosts')
Interventions	Nurse-led screening for cardiovascular risk factors and a lifestyle intervention. Screening tests used: past medical history family history smoking habit body mass index waist/hip ratio blood pressure total cholesterol random blood glucose Coronary risk score (Dundee) was communicated to each participant. The frequency of follow-up examinations was determined by this score together with other individual risk factors, and ranged between every two months (highest risk quintile) and yearly (lowest risk quintile). Lifestyle advice was given, and personally negotiated lifestyle changes were recorded in a booklet Uptake of screening: 73%
Outcomes	morbidity (prevalence of certain conditions) self-reported health medication use
Notes	We chose to use results from the comparison with the internal control group only. The authors found similar effect sizes when using either control group

Risk of bias

Family Heart 1990 (*Continued*)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Low risk	Quote: "Within each intervention practice, the list of men was randomly divided into two groups: intervention and an internal comparison group" Comment: Allocation was done on the full list all at once.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Not described.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Lack of blinding can cause bias in self-reported outcomes.
Incomplete outcome data (attrition bias) Objective outcomes	Unclear risk	No objective outcomes included.
Incomplete outcome data (attrition bias) Subjective outcomes	High risk	Only participants attending health checks were included in the analysis. For those attending, the authors investigated the possible effect of excluding non-returners at the 1-year screening in the intervention group, and found small differences in baseline morbidity but large differences in baseline smoking
Selective reporting (reporting bias)	Low risk	We do not know what was stated in the protocol, but all outcomes that can reasonably be expected seem to be reported
Other bias	Unclear risk	The authors found similar results using both the internal and external control group. However, since the effects were small and possibly due to bias and acclimatisation to blood pressure measurement, this does does not rule out contamination of the internal control group

Göteborg 1963

Methods	Included all men born in 1913 and living in Göteborg, Sweden, in 1962. Allocation of participants was done according to date of birth before any contact was made. The intervention group was invited for 3 rounds of screening and the control group was not contacted. All participants were followed through registries for mortality over 15 years
Participants	Men aged 50 years. Setting: community Location: Göteborg, Sweden Number of people randomised: 1013 (screening) and 1967 (control). Analyses were based on number of people alive when the intervention started on 1 January 1963, which were 1010 (screening) and 1956 (control)
Interventions	The first screening was performed by staff at a local hospital and used the following tests: questionnaire on social data, smoking, personal and family history qestioning about cardiovascular symptoms and chronic bronchitis questionnaire on cardiovascular symptoms weight, height, skinfold thickness blood pressure electrocardiography urinalysis (protein, glucose, osmolality) blood samples (cholesterol, triglycerides, fasting blood sugar, haematocrit, sedimentation rate, creatinine, serum protein electrophoresis, sodium, potassium, chlorides, blood groups) chest x-ray measurement of heart volume general physical examination examination by an ophtalmologist. Half of the intervention group also had a psychiatric interview. The other half had a psychiatric questionnaire and an examination of lung function In 1967, the examination also included a physical test at maximum load The 1973 examination is unclearly described, but included height, weight, skinfold thickness and questions about morbidity, well-being and utilisation of medical care Uptake of screening at first round: 85%, second round 80%, third round 74%
Outcomes	total mortality cardiovascular mortality cancer mortality
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "All men meeting these criteria who were born on a date divisible by three (the third, sixth, ninth day and so on of each month) comprised the study sample". "The men who were born on other days were

Göteborg 1963 (*Continued*)

		regarded as the control group". Comment: allocation method used is likely to yield comparable groups. All men in the eligible age range and geographical area were included and allocated before any contact was made
Allocation concealment (selection bias)	Low risk	As above.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	The regular physicians of the participants in the intervention group were not involved with the study and the control group was not informed about the trial
Blinding of outcome assessment (detection bias) All outcomes	High risk	Death certificates were assessed, and some were reclassified for cause of death. The persons doing this were not blinded to allocation status (L Welin, pers. comm.)
Incomplete outcome data (attrition bias) Objective outcomes	Low risk	Loss to follow-up for mortality was 0.3% in the intervention group and 1.0% in the control group
Incomplete outcome data (attrition bias) Subjective outcomes	Unclear risk	No subjective outcomes.
Selective reporting (reporting bias)	Low risk	
Other bias	High risk	Conditions discovered at screening were treated at the hospital where the screening was conducted. Thus, the standard of care given to the screening group likely differed from that available to the control group, which might bias the results The control group and their regular physicians were not informed about the trial (L Welin, pers. communication) which gives a low risk of contamination

Göteborg 1970

Methods	Included all men in Gothenburg who were born between 1915 and 1925. These were randomised to an intervention group and two control groups. They were followed in registers for mortality and morbidity until the end of 1983, with a mean follow-up time of 11.8 years
Participants	Men aged 47-55 years at entry. Setting: community

Göteborg 1970 (*Continued*)

	Location: Gothenburg, Sweden Number of people randomised: 10,004 (intervention), 10,011 (control 1) and 10,007 (control 2)
Interventions	The intervention group was invited to two screenings with a four year interval. Screening tests used: questionnaire on family history of cardiovascular disease and risk factors height weight total serum cholesterol blood pressure electrocardiogram interview (not specified) Blood pressure, cholesterol and smoking were treated if they exceeded specified thresholds. Systolic blood pressure > 160 mm Hg or diastolic blood pressure > 95 were followed bienially. Systolic blood pressure > 175 mm Hg or diastolic blood pressure > 115 mm Hg were treated with drugs. People with cholesterol > 6.8 mmol/L were offered dietary advice. Cholesterol > 7.8 mmol/L was re-measured and treated with dietary advice. When necessary, this was supplemented with clofibrate or nicotinic acid. Clofibrate use was stopped when its adverse effects became known. People smoking > 15 cigarettes/day were invited to an anti-smoking clinic In control group 1 a 2% random sample was invited to screening at baseline, and an 11% random sample after 4 years. Control group 2 was not contacted at all. After 10 years, a 20% random sample from the intervention group and control group 1 were invited to re-examination Uptake of screening: 75% at first round
Outcomes	mortality cardiovascular mortality (coronary mortality + stroke mortality) cancer mortality morbidity (fatal and non-fatal coronary heart disease, fatal and non-fatal stroke) prescriptions (self-reported use of antihypertensives)
Notes	We combined fatal coronary heart disease and fatal stroke as cardiovascular mortality

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation was done by computer (L Wilhelmsen, personal comm.)
Allocation concealment (selection bias)	Low risk	All participants were randomised before contact.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	General practitioners and the control group were not contacted

Göteborg 1970 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	Low risk	Cause of death was recorded from death certificates. Use of antihypertensive medication was assessed at a personal interview with a physician (L Wilhelmsen, personal comm.)
Incomplete outcome data (attrition bias) Objective outcomes	Low risk	Complete follow up for total and cause-specific mortality.
Incomplete outcome data (attrition bias) Subjective outcomes	High risk	20% (n=2000) from the intervention group and control group 1 were invited to re-examination after 10 years. In the intervention group, 74% attended. In control group 1, 70% attended. Due to lack of blinding there is a high risk of bias
Selective reporting (reporting bias)	Low risk	Outcomes were pre-specified in an early article.
Other bias	High risk	Hypertensives and smokers were treated and followed in a special clinic, thus getting a different standard of care from the two control groups

Inter99 1999

Methods	All 61,301 persons aged 30, 35, 40, 45, 50, 55 and 60 years and living in 11 municipalities in the south-western part of Copenhagen County on 2 December 1998 were included. A random sample was invited to screening and those remaining constituted the control group. The intervention group and a random subsample of the control group (n=5264) had questionnaires at baseline and after 1, 3 and 5 years of follow-up. All participants were followed up through central registers
Participants	Men and women aged 30-60 years. Setting: community Location: Copenhagen, Denmark Number randomised: 13,016 (screening) and 48,285 (control).
Interventions	The screening included: blood pressure height and weight waist and hip circumference and ratio fasting blood samples (HDL, triglyceride, total cholesterol, VLDL, LDL) glucose tolerance test spirometry electrocardiogram Absolute 10-year risk of ischaemic heart disease was assessed using the PRECARD computer program, with individual counselling on risk factors and adverse health behaviours

	High-risk participants were offered four health checks (years 0,1,3 and 5), low risk participants were offered two (years 0 and 5). The intervention group was further randomised into high or low intensity treatment of risk factors. The high intensity group participants who had a high risk of ischaemic heart disease were offered six sessions of group counselling during a 4-6 month period, and were re-invited for a similar intervention after 1 and 3 years. In the low intensity group no participants were offered group counselling. The control group was not contacted, except for the sample that received questionnaires Uptake of screening: first round 53%
Outcomes	self-reported health (mortality results are not published yet)
Notes	Results on mortality and utilisation of healthcare resources are not published yet The results on self-reported health are based on a comparison between the intervention group and the subsample of the control group who had questionnaires

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "From the study population an age- and sex-stratified random sample comprising 13,016 individuals was drawn". Randomisation was done by computer (T Jørgensen, personal comm.)
Allocation concealment (selection bias)	Low risk	Groups were formed before any participants were contacted.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Blinding of the intervention group was not possible. The control group, including the subsample who received questionnaires, were not informed about the trial (T Jørgensen, personal communication). Medical follow-up of high-risk participants was by the participants' general practitioners, who were informed at the beginning of the study but not otherwise involved
Blinding of outcome assessment (detection bias) All outcomes	High risk	Since intervention group participants could not be blinded, there is a risk of detection bias for self-reported outcomes Mortality results are not yet published and risk of detection bias cannot yet be assessed
Incomplete outcome data (attrition bias) Objective outcomes	Unclear risk	No objective outcomes reported yet.

Incomplete outcome data (attrition bias) Subjective outcomes	High risk	Loss to follow-up for self-reported health was 27% in the intervention group and 20% in the sample of the control group who received questionnaires. There is a risk of attrition bias due to the lack of blinding
Selective reporting (reporting bias)	Unclear risk	Not all outcomes are published yet.
Other bias	Low risk	The control group was not informed about the trial and their regular physicians were not involved with the conduct of the trial

Kaiser Permanente 1965

Methods	In April 1964, a sample of members of the Kaiser-Permanente Health Plan in San Francisco and Oakland aged 35-54 years were divided into an intervention group and a control group using an allocation rule based on membership number. Starting in 1965, people in the intervention group were urged annually, by telephone and letter, to have the multiphasic screening examination offered by the Kaiser Health Plan. The intervention lasted 16 years. Participants were followed using questionnaires and registers
Participants	Men and women aged 35-54 years who were members of a large health plan and thus mainly people with employment Setting: community (healthcare plan members) Location: California, USA Number of people randomised: 5156 (intervention) and 5557 (control). For analyses, the authors included people alive on 1 January 1965, when the intervention started. Thus, the groups analysed were: 5138 (intervention) and 5536 (control)
Interventions	The intervention was annual urging to have a broad medical screening. Screening tests used: electrocardiography blood pressure height and weight chest x-rays breast x-rays visual acuity tonometry audiometry spirometry blood tests (not specified, but included a serum chemistry panel) urine tests (not specified) past medical history (self-administered) present symptoms (self-administered) health habits (self-administered) family history (self-administered), social history (self-administered) physical examination by a physician Women were advised to have a pelvic examination by a gynaecologist. Sigmoidoscopy

Kaiser Permanente 1965 (*Continued*)

	was recommended for all persons aged 40 years and over In early years there was a follow-up visit by a physician, including a physical examination, but in later years (not specified) the follow-up could also be performed by a nurse practitioner supervised by a physician The control group was not urged but could have a health similar health check if they wished, as part of their health plan
Outcomes	mortality cardiovascular mortality cancer mortality morbidity hospitalisation physician visits disability new diagnoses.
Notes	People who left the Kaiser Permanente Health Plan were not followed-up. This led to attrition of about 35% in both groups after 16 years, possibly selected as those who lost their employment. An exception to this is mortality, which was assessed using registers. Participants who were found to have moved too far away to be called for a health check after allocation were excluded. There were also exclusions due to identity mix-ups, i.e. participants having more than one health plan identification number.The exact figure is not given for the intervention group, but is stated to be over 200. However, the discrepancy between the groups is larger. Excluded participants were included in the analysis of mortality after 11 years, without important differences

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "The terminal digit and fourth digit of each member's unique seven-digit medical record number were used to assign participants to the two groups. Those with one particular terminal digit were assigned to the study group and those with another terminal digit were assigned to the control group. Those with a third terminal digit were assigned to the former if they had one of two particular fourth digits and to the latter if they had one of two other fourth digits. Medical record numbers are assigned sequentially to new members and are never reassigned." Comment: The method used is likely to yield comparable groups, and all participants were allocated at the same time, before contact

Kaiser Permanente 1965 (*Continued*)

Allocation concealment (selection bias)	Low risk	See above.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Quote: "Neither the subjects nor their physicians were aware that they were participating in a controlled trial" In the regular mail surveys, the participants were not informed about the trial but told that the survey was about improving health services to members
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "Trained readers, blind to the study or control group membership status of the patients, examined the charts selected and abstracted diagnostic data." Quote: "Specially trained personnel, blind to the study or control group membership status of the hospital patients, coded the diagnostic and operative procedure data according to the system of the Hospital Adaption of the International Classification of Diseases (1968)." Quote: "Death certificate copies received from the State were checked against Kaiser Foundation Health Plan clinical records in order to confirm identification of the decedents as study and control group members. Those death certificates accepted for analysis were coded for underlying cause of death (again by trained persons who were blind to the study or control group membership status of the individuals involved), using the International Classification of Diseases Adapted, Eighth Revision." Comment: blinded adjudication of all objective outcomes. Self-reported disability is an exception to this, and may be biased
Incomplete outcome data (attrition bias) Objective outcomes	Unclear risk	Quote: "Since surveys of the subjects still in the Health Plan indicated they used Kaiser-Permanente facilities for over 80% of their outpatient clinic data were gathered from Kaiser clinical charts and hospital data from Health Plan computerized records" Quote: "In June 1980 3326 or 64,5% of the study group and 3544 or 63,8% of the control group were still members" Quote: "Deaths were ascertained by matching names of subjects no longer active in the Health Plan against State of Califor-

Kaiser Permanente 1965 (*Continued*)

		nia mortality records. Mortality surveillance thereby included subjects who left the Health Plan unless they became residents of another state." Comment: People who left the Kaiser Permanente Health Plan were not followed-up. This led to attrition of about 35% in both groups after 16 years. Only people leaving California were lost to follow-up for mortality, and the authors assessed this to be 8-18%
Incomplete outcome data (attrition bias) Subjective outcomes	High risk	As above. The large attrition combined with a 75% response rate at each survey means that at 16 years of follow-up less than half of the participants randomised were included in the analyses
Selective reporting (reporting bias)	High risk	Data on surgery, reasons for hospitalisation and number of prescriptions were collected but never published
Other bias	High risk	After 16 years of intervention the mean number of health checks was 6.8 in the intervention group and 2.8 in the control group. In the intervention group 16% of the participants had never had a health check, compared to 36% in the control group. Thus, there was contamination of the control group

Malmö 1969

Methods	All men born in 1914 and living in Malmö, Sweden in early 1969 were included in the study. Men born in even-numbered months were invited to screening and men born in uneven-numbered months were not. Five-year follow-up through registries
Participants	Men only, all aged 55 years. Setting: community Location: Malmö, Sweden Numbers randomised: 809 (screening) and 804 (control).
Interventions	The intervention group was invited to one screening. The control group was not contacted Screening tests used: blood pressure blood tests (cholesterol, triglycerides, haematocrit) urinalysis (glucose, albumin)

Malmö 1969 (*Continued*)

	height and weight electrocardiography spirometry nitrogen washout sputum cytology heart and lung radiography venous occlusion plethysmography interview and questionnaire physical examination. Participants with hypertension and impaired lung function were followed and treated at the hospital. Of 178 participants classified as heavy smokers, 51 were offered a group counselling intervention to quit. Of these, 5 were prescribed sedatives Uptake of screening: 87%
Outcomes	total mortality cardiovascular mortality cancer mortality hospitalisation morbidity
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "...all men born in even-numbered months in 1914 were invited to take part in an examination of cardiovascular and pulmonary function" Comment: All persons were allocated at the same time, before contact, and the method used is likely to to yield comparable groups
Allocation concealment (selection bias)	Low risk	See above.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	The control group and their regular physicians were unaware of the trial
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The person assessing cause of death was not aware of the allocation (S Isacsson, pers. comm.). Hospitalisation data were from public registers
Incomplete outcome data (attrition bias) Objective outcomes	Low risk	Loss to follow-up was 1% for mortality. For hospitalisation it was 3.6% (intervention) and 5.6% (control)

Malmö 1969 (*Continued*)

Incomplete outcome data (attrition bias) Subjective outcomes	Unclear risk	No subjective outcomes were reported.
Selective reporting (reporting bias)	Low risk	No indications of selective reporting. Reports on all expected outcomes
Other bias	High risk	Conditions identified at screening were followed and treated at a hospital in contrast to the control group who were followed by general practictioners. Thus, the standard of care was likely different Participants in the control group and their primary care physicians were unaware of the trial, which gives a low risk of contamination

Mankato 1982

Methods	Addresses representing the entire community were randomised. In the intervention group, the whole household was invited for screening, but only one eligible participant from each household, selected randomly, was included in the trial and followed. The control group was not invited. After one year, participants in the intervention group who attended the initial screening were re-invited, and the control group was invited for their first time
Participants	Men and women aged 25-74 years. Setting: community Location: Mankato, Minnesota, USA Number randomised: 1,156 (screening) and 1,167 (control).
Interventions	Screening tests used: height weight blood pressure total serum cholesterol expired air carbon monoxide leisure time physical activity Results of tests were returned during the visit. Participants received health education at each measurement station, either on videotape, printed materials, or both. After measurements each family spent 20 minutes with a health educator to review test results and receive further health advice. The average visit lasted 75 minutes Participants with high blood pressure or high cholesterol were referred to their regular physician Uptake of screening: 50%
Outcomes	Prescriptions (self-reported use of antihypertensive drugs)

Mankato 1982 (*Continued*)

Notes	Simultaneously with the trial a population-based programme to educate about risk factors for coronary heart disease was going on. This programme included an offer of screening tests for coronary heart disease risk at the same centre that also conducted the trial. However, participants in the control group were systematically excluded from attending the screening clinic for the duration of the trial

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was done by computer (D Murray, pers. com.)
Allocation concealment (selection bias)	Low risk	All participants were randomised at the same time, before any contact
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Quote: "Except for the recruitment supervisor, HHC (screening site, our comment) staff members were not informed of the study until its conclusion" Quote: "In addition, participants were not informed of their treatment condition and were scheduled together during the 1983 follow-up. They were identified only through a special code kept secret from the staff." Physicians were not informed about the trial, but patiens with high blood pressure or high cholesterol were referred to their regular physician for treatment (D Murray, pers. com.)
Blinding of outcome assessment (detection bias) All outcomes	High risk	The included outcome (medication use) was self-reported (D Murray, pers. com.) and could be biased due to lack of blinding
Incomplete outcome data (attrition bias) Objective outcomes	Unclear risk	No objective outcomes.
Incomplete outcome data (attrition bias) Subjective outcomes	High risk	Only half of those invited attended their first screening and were included in analyses. In addition, there was a 12% loss to follow-up in the intervention group between the baseline screening and the follow-up screening. Seven per cent of the control group participants moved away before the 1-year screening. In summary, in both groups about 40% of those ran-

Mankato 1982 (*Continued*)

		domised were included in the analyses
Selective reporting (reporting bias)	Low risk	No indications of selective reporting.
Other bias	High risk	A population-based programme to educate about risk factors for coronary heart disease was ongoing during the trial. This may have diminished the effect of the intervention

New York 1971

Methods	A random 80% sample of eligible families was invited for screening and the remaining 20% were not. Sampling was stratified by Medicaid status and the presence of a child 12-18 years of age. The main aim was to assess whether health checks would reduce the health difference between poor and non-poor families. The trial appears planned to have lasted 3-4 years, but the authors note that the follow-up may be prolonged if the results indicate an effect on health differentials between economic groups
Participants	Families with at least one person aged 12-74 years old, enrolled for 1 or more years in the Health Insurance Plan of Greater New York Setting: community Location: New York City, New York, USA Number randomised: not clear. The papers mention an expected number of 7,000 non-poor families in the intervention group and a somewhat smaller number of poor families. The control group would be 20% of this size
Interventions	Screening tests used: electrocardiogram blood pressure pulse rate height, weight and skinfold thickness chest x-ray audiometry dental survey visual acuity tonometry spirometry glucose challenge blood tests (cholesterol, total protein, albumin, calcium, total bilirubin, urea nitrogen, uric acid, haemoglobin, white blood cell count, syphilis test) urine tests (pH, protein, glucose, blood, acetone) sickle cell trait urine culture (women only) instruction in breast self-examination mammography (women aged 40+ years) Pap smear
Outcomes	No outcomes reported. The trial was designed to investigate disability and absence from work. Mortality data were also to be gathered

New York 1971 (*Continued*)

Notes	The AMHT programme was discontinued after the first screening round, but follow-up was planned to continue. We have not found reports of the results

Northumberland 1969

Methods	All eligible men were allocated at the same time before any contact was made, excluding 7% because of serious illness. Participants were allocated by date of birth to one of three groups: questionnaire and full examination, questionnaire and examination if indicated by answers to the questionnaire, and neither questionnaire nor examiniation. We used the first and the last group in our analyses. Outcomes were assessed from medical records
Participants	Men aged 50-59 years. Setting: general practice Location: England Numbers randomised: 242 (intervention) and 291 (control).
Interventions	The examination is not specified, is described in the article as a 'routine health examination', a 'full examination', and 'screening programme'. It took an average of 26 minutes Uptake of screening: 90%
Outcomes	physician visits
Notes	Also reported on the prescription of drugs, use of laboratory investigations, sickness certifications and admissions to hospital, but in a way we could not use

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	The randomisation sequence was based on date of birth. All eligible men were allocated at the same time before any contact was made, excluding 7%, balanced across groups, because of serious illness. Authors found small imbalances in the past medical histories between groups, but also noted that there might have been bias in the assesment of this. All in all, we judge that the method used is likely to have produced comparable groups
Allocation concealment (selection bias)	Low risk	See above.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Physicians were involved in trial conduct, were aware of screening status, and treated both screened and unscreened patients

Northumberland 1969 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "Bias could have been introduced in completing the past history recording as the group that the patient was assigned to was indicated on the front page of the schedule" Comment: All outcomes were abstracted from patient records and therefore susceptible to detection bias
Incomplete outcome data (attrition bias) Objective outcomes	Unclear risk	Not described.
Incomplete outcome data (attrition bias) Subjective outcomes	Unclear risk	No subjective outcomes reported.
Selective reporting (reporting bias)	Unclear risk	Unclear.
Other bias	Unclear risk	Cannot rule out contamination of the control group.

OXCHECK 1989

Methods	People who returned an initial questionnaire were included and randomised by household into four groups: health checks at year 1 and 4; at year 2 and 4; at year 3 and 4; and only at year 4. The first three groups constituted the intervention groups and the last group was a control group. Participants in the first two groups were further randomised to annual re-checks or no re-checks
Participants	Men and women aged 35-64 years. Setting: general practice Location: Luton and Dunstable, UK Number randomised: 2776, 2771 and 2760 (screening groups) and 2783 (control)
Interventions	Cardiovascular screening conducted by specially trained nurses (45-60 minutes) Screening tests used: blood pressure total cholesterol height weight personal and family medical history lifestyle questionnaire dietary assessment exercise rates alcohol consumption Counselling about risk factors. Follow-up visits for risk factors (10-20 minutes). Annual re-checks were similar to initial health check, but briefer (30 minutes) Uptake of screening: first round 80%, re-checks 76% - 79%

OXCHECK 1989 (*Continued*)

Outcomes	mortality cardiovascular mortality cancer mortality morbidity (cancer incidence)
Notes	In the meta-analyses, we combined the three groups invited to screening in year 1, 2 and 3 and compared them with the control group. The results were similar when analysing the results for maximum contrast, i.e. only comparing those screened in year one with those in year 4

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation was done independently of the research team, using a computerised algorithm (D Mant, personal comm.)
Allocation concealment (selection bias)	Low risk	The computer generated a list of names for each practice indicating the intervention group to which each individual patient had been allocated (D Mant, personal comm.)
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "A sticker was attached to the outside of each patient's general practice notes indicating the randomisation group"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Cause of death and cancer incidence were from national statistics and likely unbiased
Incomplete outcome data (attrition bias) Objective outcomes	Low risk	As above.
Incomplete outcome data (attrition bias) Subjective outcomes	Unclear risk	No subjective outcomes.
Selective reporting (reporting bias)	Low risk	No indication of selective reporting.
Other bias	Unclear risk	Only people who returned an initial questionnaire were included, which limits external validity due to self-selection Risk of contamination is unclear.

Salt Lake City 1972

Methods	Randomised by family. Allocation ratio was 3:2 (intervention:control)
Participants	Participants consisted of random samples from three groups: 200 families with a low-income and a pre-paid healthcare programme, 200 families with a low-income and no pre-paid healthcare programme, and 166 middle income families, who had volunteered for a study of health care Age >18 years. Setting: community Location: Salt Lake City, Utah, USA Number randomised: 642 (intervention) and 454 (control).
Interventions	Both groups had a baseline interview measuring health status (Bush index), number of disability days caused by illness, patterns of healthcare utilisation, health knowledge, attitudes toward the healthcare system (Hulka scale) and Pilowsky's scale of hypochondriasis. The intervention group was urged by telephone to obtain a multiphasic screening examination at no cost. Each patient's physician had to give permission for the patient to participate. After screening the results were sent to the physician for interpretation and follow-up. The control group was not urged to be screened Screening tests used: audiometry visual acuity tonometry blood pressure electrocardiogram spirometry chest x-ray urinalysis (specific gravity, glucose, protein, red-cell count, white-cell count, casts) blood tests (globulin, uric acid, urea nitrogen, glucose, alkaline phosphatase, glutamic oxalacetic transaminase, bilirubin (total, direct and indirect), triglycerides, cholesterol, latex fixation for rheumatic arthritis, creatinine, thyroid studies, haematology) breast examination and mammogram cervical cytology Uptake of screening: 60%.
Outcomes	hospitalisation physician visits disability
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No description.
Allocation concealment (selection bias)	Unclear risk	No description.

Salt Lake City 1972 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	High risk	Primary care physicians had to give permission for each person to participate. Lack of blinding of physicians could cause performance bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Outcomes were patient-reported and susceptible to bias due to the lack of blinding
Incomplete outcome data (attrition bias) Objective outcomes	High risk	Those who changed economic status, did not attend for screening, did not consult their physician about screening results, or who did not participate in the 1-year follow-up, were excluded. This resulted in only 49% of the intervention group and 82% of the control group participants being included in analyses
Incomplete outcome data (attrition bias) Subjective outcomes	High risk	As above.
Selective reporting (reporting bias)	Unclear risk	Unclear.
Other bias	Unclear risk	Not enough information to judge the risk of contamination.

South-East London 1967

Methods	Eligible persons were identified through registers and randomised by family to intervention or control. The screening group was invited by letter to two rounds of screening, with a two-year interval. After five years both groups were invited for screening, but the authors state that this screening was "non-prescriptive, in the sense that no therapeutic activity was expected to result from it". Follow-up was continued for a further four years
Participants	Men and women aged 40-64 years. Setting: general practice Number randomised: according to one paper the numbers were 3460 (screening) and 3337 (control) (Trevelyan 1973), whereas another gives 3876 (screening) and 3353 (control) (South-East London Study Group 1977). The mortality analyses were based on 3292 (intervention) and 3132 (control) participants
Interventions	Screening tests used: physical examination (first screening only) history questionnaire on symptoms height and weight vision hearing testing

South-East London 1967 (*Continued*)

	chest x-ray spirometry electrocardiography blood pressure blood chemistry faecal occult blood testing Advice on smoking and weight was given to all for whom it was appropriate. All results were passed on to the patient's general practitioner Uptake of screening: first round 73%, second round 66%
Outcomes	mortality cardiovascular mortality cancer mortality hospitalisation morbidity physician visits self-reported health disability worry
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Eligible participants and couples were listed alphabetically and alternate allocation was used. After randomisation, a matching took place which is unclearly described. It resulted in the exclusion of 276 participants from the control group The sizes of the groups vary between reports.
Allocation concealment (selection bias)	Low risk	Participants were identified and randomised before any contact was made
Blinding of participants and personnel (performance bias) All outcomes	High risk	"All information gathered at both screening sessions was passed on to the general practitioners" Comment: general practitioners were not blinded, which gives a risk of performance bias. Not clear whether the control group was informed about the trial
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No mention of blinding of outcome assessment. Self-reported outcomes are susceptible to bias due to lack of blinding

South-East London 1967 (*Continued*)

Incomplete outcome data (attrition bias) Objective outcomes	Low risk	After 5 years 20% of the participants had migrated from the area and were lost to follow-up for physician visits but not for other objective outcomes. Thus low risk for these outcomes but high risk for the outcome "physician visits"
Incomplete outcome data (attrition bias) Subjective outcomes	High risk	Loss to follow-up for subjective outcomes after five years was 47% (intervention) and 41% (control)
Selective reporting (reporting bias)	High risk	According to an early report data were collected on prescriptions issued, referrals and investigations carried out, but were not reported and are not available
Other bias	High risk	The control group was screened after 5 years, which biased the 9-year results towards no effect A high degree of involvement of general practitioners gives a risk of contamination

Stockholm 1969

Methods	A double sample was drawn from the eligible population and divided into three age groups. From these, a random sample was drawn using sample fractions in the proportions of 3:2:1, with the highest fraction for the youngest age stratum. These people were sent a questionnaire about social and physical difficulties and health needs. Based on this, and on data from the public inpatient register, they were substratified by expected needs for medical services: high need, low need, no need, and unknown need. Randomisation to screening and control groups took place within these strata, but proportionally more were randomised to screening in the two groups with high and low needs for services compared to those with no or unknown needs for services. The authors used regression analysis, in which they controlled for the baseline imbalances introduced by the randomisation scheme. Participants were followed for mortality in registers for 22 years
Participants	Men and women aged 18-65. Setting: community Location: Stockholm, Sweden Number randomised: 3064 (screening) and 29,122 (control).
Interventions	Participants in the intervention group were invited to one screening, while the control group was not Screening tests used: blood pressure social, psychiatric and medical interviews blood tests (not specified) physical examination

Stockholm 1969 (*Continued*)

	electrocardiogram exercise tests (not specified) psychological tests (not specified) eye examination dental examination Participants with identified need for specialist services were directly referred, whereas participants were instructed to contact their primary care physician for other identified issues. Simple services like reassurance and prescription of simple medications (not specified) were provided by the researchers Uptake of screening: 84%
Outcomes	total mortality cardiovascular mortality cancer mortality
Notes	We obtained data on mortality within each of the 12 strata in which randomisation was performed, and treated them as 12 separate trials, each giving an estimate of the effect. We then combined the results with a fixed-effect model meta-analysis, and used this estimate for our meta-analysis. Our result is nearly identical to that of the authors

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation was done by computer (H Theobald, personal communication)
Allocation concealment (selection bias)	Low risk	All participants were randomised at the same time.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The Intervention group could not be blinded. Not clear whether the control group and their general practitioners were aware of the trial
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Cause of death on death certificate was used.
Incomplete outcome data (attrition bias) Objective outcomes	Low risk	Less than 1% missing outcome data.
Incomplete outcome data (attrition bias) Subjective outcomes	Unclear risk	No subjective outcomes.
Selective reporting (reporting bias)	Unclear risk	Data on hospitalisation, operations and cancer incidence have been collected but not yet published (H Theobald, personal com.)

Stockholm 1969 (*Continued*)

Other bias	Unclear risk	Both groups had a questionnaire at baseline. The effect of this is unclear

Titograd 1971

Methods	A random sample was drawn from the eligible population and randomly divided into an intervention and a control group. A 20% random subsample of both groups were interviewed at baseline. Analysis was planned after 6 years, and follow-up would be continued for a further 4 years in case of no effect
Participants	Men and women aged 30-49 years. Setting: community Location: Titograd, former Yugoslavia Number randomised: 6577 (screening) and 6573 (control).
Interventions	The intervention group was invited for screening at baseline and with two-year intervals. Follow-up of positive test results and treatment of identified conditions would be done according to specified regimens. The control group was not invited for screening Screening tests used: height and weight chest x-ray electrocardiogram blood pressure fundus examination spirometry visual acuity blood sedimentation rate red and white blood cell counts haemoglobin blood urea nitrogen latex fixation test (not clear for which antibodies) glucose tolerance serum cholesterol WR (syphilis) urinalysis (not specified) cervical smear
Outcomes	No outcomes were reported. The outcomes studied were mortality, morbidity (from medical records), absence from work, and utilisation of outpatient and inpatient services
Notes	

WHO 1971

Methods	Forty matched pairs of factories were randomised to intervention or control. Follow-up varied between factories, but was between 5 and 6 years
Participants	Men aged 40-59 years at entry. Setting: workplace Location: UK, Belgium, Poland and Italy. Spain was also part of the trial, but was not included in the analyses of events because it started late compared to the other part of the trial. This decision was made before results were available to the investigators Numbers randomised: 30,489 (intervention) and 30,392 (control). A 10% random sample of the control group was screened at baseline and was not included in the analysis of events. Thus, the numbers analysed were: 30,489 (intervention) and 26,971 (control)
Interventions	Screening tests used: blood pressure total serum cholesterol weight questionnaire on smoking, physical activity and symptoms (angina, history of severe pain) The men at highest risk (10-20%, definitions varied between centres) were called for a physician interview and given advice and treatment. All men at the intervention factories were given advice on cholesterol-lowering dietary changes. Individual advice was given when relevant for smoking cessation, weight reduction, excercise, control of hypertension. Patients were treated and followed-up by the research teams Annually, a random 5% sample was re-examined. At the end of follow-up, all in the intervention and control groups were invited to examination Uptake of screening: 86%.
Outcomes	total mortality cardiovascular mortality (only reported coronary mortality, which we used) cancer mortality (only data from the UK, Poland and Italy parts of the trial) morbidity (fatal and non-fatal coronary heart disease)
Notes	Effect estimates from an appropriate analysis, taking clustering into account, was reported for total and coronary heart disease mortality and we used this in our meta-analysis. For cancer mortality, no such estimate was reported, and we thus ignored the clustering in the meta-analysis, but investigated the effect in a pre-specified sensitivity analysis

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	One centre used coin-flips (G De Backer, personal comm.). No description is available for the other centres
Allocation concealment (selection bias)	Low risk	"Twenty-four large industrial groups (mainly factories) were recruited and then paired according to type

WHO 1971 (*Continued*)

		of industry and are. One of each pair was allocated at random to receive the intervention programme while the other served as a control" "[The factories] were required to commit themselves to participation before knowing whether their allocation would be to an active programme of intervention or to a passive control status"
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Primary care physicians and the control group were not informed about the trial
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Blinding of assessment of cause of death is not described in the articles summarising all countries. There was blinded assessment in the UK and Belgium, but we cannot rule out unblinded assessment in other centres
Incomplete outcome data (attrition bias) Objective outcomes	Low risk	"Survival status at end of trial was established in 99.8%." Thus, total and coronary heart disease mortality are at low risk of attrition bias. Cancer mortality is an exception, because it was not reported from the Belgian part of the trial. The reason given for this is that all non-coronary deaths were only categorised as such, without detailing the cause of death, as per the trial's protocol. The risk of bias due to this is unclear
Incomplete outcome data (attrition bias) Subjective outcomes	Unclear risk	No subjective outcomes included.
Selective reporting (reporting bias)	Low risk	Outcomes were pre-specified in early articles.
Other bias	High risk	Participants in the intervention groups were treated and followed by the research team, in contrast to the control group. Thus, the standard of care was different

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
A Healthy Future	Only enrolled medicare beneficiaries, mean age 73. The intervention included promotion of and counselling on advance directives, including living will/power of attorney, housing plans, and care insurance
Addition - Cambridge	Only screening for diabetes.
Aleman 2011	Not randomised.
Apkon 2005	Intervention was a decision support tool, not health checks.
Asshauer 1972	Not randomised.
Atthobari 2004	Not randomised.
Atthobari 2007	Not randomised.
Bandinelli 2006	Intervention was multifactorial disability prevention. All participants were screened at baseline
Barr 2005	Only screening for fracture risk. Vitamin D was prescribed.
Bekwelem 2012	Not randomised.
Belcher 1990	Participants were patients attending a veterans outpatient clinic having had a medical condition diagnosed during active service, and had a high prevalence of CHD, COPD, and diabetes
Bonevski 1999	Intervention was education of general practitioners to increase screening rates. Both intervention and control group received health checks
Boulware 2007	Not randomised (systematic review).
Brown 2009	Compared two different types of health assessment. Both groups received health checks
Bula 1999	Only enrolled persons over 75 years of age. Intervention included screening for extensiveness of social network, social support, home safety and access to environment
Carpenter 1990	Intervention was screening with a single instrument: the Winchester disability rating scale
Christensen 2003	Only screening for psychiatric disorders.
Coulton 2008	Compared two interventions for alcohol abuse.
Coustasse-Hencke 2000	Not randomised.
Cowan 1992	Intervention was a reminder sheet to physicians about preventive services
Cutchin 2009	Intervention did not include medical screening tests.

(*Continued*)

Dalby 2000	Participants were selected for risk.
Dietrich 2006	Intervention was reminders for three cancer screenings in persons overdue for these
Dubey 2006	Intervention was a checklist on preventive services delivered and promoted to staff.
Eekhof 2000	Only enrolled persons over 75 years. Intervention was screening for reduced vision and hearing, urinary incontinence and reduced mobility
Elley 2003	Intervention was promotion of excercise.
Enzell 1984	Not randomised.
Fanaian 2010	Participants were selected for having hypertension or hyperlipidaemia
Fang 1999	Not randomised.
Fitzgerald 1991	The intervention was educational mailing to improve referral following cholesterol screening
Fitzmaurice 2007	Only screening for atril fibrillation.
Fleischer 2008	Participants were selected for risk.
Fletcher 1977	Participants were patients at a hospital polyclinic and did not represent a general population
Fletcher 2002	Compared detailed and targeted assessment. Both groups were screened
Fox 1997	The intervention was a personalised health plan and counselling. No unscreened group
Fullard 1987	Not randomised.
Gemson 1990	Compared two kinds of follow-up for people with borderline high cholesterol
Gemson 1995	The intervention was a computerised report following a health check. No unscreened group
Giampaoli 1997	Not randomised
Goodwin 2001	Intervention was counselling of clinicians to improve delivery of preventive services
Grover 2007	Intervention was sharing coronary risk profile with patients. No unscreened group
Gysan 2004	Intervention was reduction of cardiovascular risk in people at high risk
Hanlon 1995	The intervention was feedback on screening results. No unscreened group
Harding 2009	Participants were selected for risk.

(*Continued*)

Hay 1998	Intervention was placement of screening questionnaires in charts
HEART	Intervention was efforts aimed at staff to improve delivery of services
Heath 1995	Not randomised.
Hellenius 1999	Not randomised.
Hendriksen 2005	Not randomised.
Hiratsuka 2007	Intervention was lifestyle intervention. No unscreened group
Hogg 1998	Intervention was reminders about overdue preventive health measures
Huang 2004	Compared two different kinds of falls prevention.
Hunter Mellado 1997	Intervention was patient education.
Hutchison 1998	Intervention was a questionnaire appraising risk of CHD, with people scoring high being recommended a cholesterol test
IMPROVE	Intervention was efforts directed at clinics to improve delivery of preventive services
Jilcott 2006	Intervention was lifestyle intervention. No unscreened group
Johansson 1999	Not randomised.
Johansson 2002	Not randomised.
Kaczorowski 2011	Not randomised.
Kahler 1978	Not randomised.
Kneipp 2011	Only enrolled people with at least one chronic condition.
Knutsen 1991	Intervention was health promotion. No unscreened group.
Koinberg 2003	Only enrolled breast cancer patients.
Koivisto 1992	Not randomised.
Kolbe-Alexander 2008	Not randomised.
Kolozsi 1982	Not randomised.
Kono 2004	Participants were selected for risk (frailty).
Kono 2009	Participants were selected for risk (frailty).

(*Continued*)

Kowal 1979	Not randomised.
Kurata 2006	Not randomised.
Landi 2001	Not randomised.
Lauritzen 2011	Intervention was intensive treatment of screening-detected diabetes mellitus versus usual care
Lave 1995	Intervention was waivers for free continuous health promotion. Both groups had health checks
Liaw 1996	Intervention was different kinds of feedback on screening results
Manning 1984	Intervention was different systems of payment for all services
Marshall 2008	Not randomised.
Mathews 2007	Not randomised.
Maxwell 1983	Not randomised.
McMahon 2002	No unscreened control group.
Medicare - Baltimore	Only enrolled persons 65 years or older. In addition to screening and lifestyle advice, the intervention included immunizations for influenza, pneumoccus and tetanus, and counselling for medication use, falls prevention, emotional distress, sleep problems and urinary incontinence
Meland 1996	Compared two different styles of follow-up for people with high cardiovascular risk. No unscreened control group
Melis 2005	Participants were selected for risk (frailty).
Milisen 2006	Not randomised.
Minder 2002	Not randomised.
Mitchell 2005	Participants had hypertension.
Morrissey 1995	Only enrolled persons 65 years or older. In addition to screening and lifestyle advice, the intervention received pneumococcal and influenza vaccines, memory improvement intervention, counselling for problem-solving, medication awareness, falls and accidents, help with deciding about life-sustaining treatment and living wills
Naor 1975	Could not be retrieved.
Newcomer 2004	Participants were selected for risk
Nicolaides-Bouman 2004	Participants were selected for poor self-reported health.

(*Continued*)

Näyhä 1988	Not randomised (randomised three groups of communities).
O'Malley 2006	Intervention was two different kinds of counselling to reduce risk factors
O'Rourke 1970	Not randomised.
Parker 2005	Intervention was an effort to improve guideline adherence.
Phelan 2007	Only enrolled persons 75 years or older. In addition to screening, intervention contains comprehensive psycho-social assessment and intervention, medication review and close collaboration between nurse, geriatrician and primary care provider
Reid 1995	Intervention was group counselling compared to a pamphlet. Both groups received screening
Reijneveld 2003	Intervention was health education and excercise, not health checks
Ritzau 1969	Not randomised.
Robertson 1992	Compared immediate and delayed feedback on blood cholesterol during opportunistic screening for CHD risk
Rodondi 2008	Intervention was carotid artery ultrasound screening.
Romundstad 2003	Not randomised.
Rose 1996	Not randomised.
Rubenstein 2007	Participants were selected for risk.
Sackett 1974	Compared different models for providing clinical services in general
Sahlen 2006	Only screened with questionnaires.
Schweitzer 1994	Only enrolled persons 65 years or older. Intervention included assessment of social functioning, psychosocial and rehabilitation needs, preventive social work
Shannon 2001	Intervention was a reminder tool for improving content of health checks
Snow 1989	Not randomised.
Stange 2003	Intervention was efforts directed at clinics to improve delivery of preventive services
Strandberg 1991	Intervention was treatment of cardiovascular risk factors in persons with high risk
Stuck 2000	Only enrolled persons 75 years or older. The intervention was a comprehensive geriatric support system, involving nurse and geriatrician with contact to general practitioner and specialst therapists (e.g. physical therapists)

(*Continued*)

Stuck 2007	In two centers only enrolled person 65 years or older, and in one center only persons 60 years or older. The intervention was use of HRA-O instrument which included an extensive questionnaire on medical, social, and psychological issues, and measurement of cardiovascular risk factors. Different kinds of feed-back to participants and their health care providers were used, and included incorporation into the general practitioners patient record, groups sessions with an interdisciplinary team, and visits by a specially trained nurse who consulted with a geriatrician
Taylor 1998	Intervention was a referral programme for control of cardiovascular risk factors. Both groups were screened at baseline
Toth-Pal 2004	Not randomised.
Tulloch 1979	Only enrolled persons 70 years or older. Intervention was screening and surveillance in a geriatric clinic
Turner 1989	Compared different preventive care reminder systems for physicians
van Haastregt 2000	Not randomised.
van Hout 2010	Only screening with questionnaires.
van Rossum 1993	The intervention was home visits which included discussion of selected topics such as health, functional state, medication, social contacts, and housing conditions. No screening tests were used
van Weel 2006	The intervention was treatment of risk factors in people at high risk
Vass 2007	Intervention was education of professionals, with the aim of improving the home visits. The intervention was a multidimensional assessment and intervention aimed at improving overall functioning. Both groups received preventive home visits
Waldman 1970	Not randomised.
Weinehall 1999	Not randomised.
Williams 1997	Only enrolled persons 65 years or older. The intervention was screening with cholesterol, snellen chart, audiometry and TSH, immunisations for influenza and pneumococcus, and an intensive health promotion effort.
Williams 1998	Intervention was reminders for preventive services.

DATA AND ANALYSES

Comparison 1. Health checks versus control

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Total mortality	9		Risk Ratio (Random, 95% CI)	0.99 [0.95, 1.03]
2 Total mortality - sensitivity analyses	6		Risk Ratio (Random, 95% CI)	0.98 [0.94, 1.03]
2.1 Excluding cluster trials	6		Risk Ratio (Random, 95% CI)	0.98 [0.94, 1.03]
3 Total mortality - no. of health checks	9		Risk Ratio (Random, 95% CI)	0.99 [0.96, 1.03]
3.1 One health check	3		Risk Ratio (Random, 95% CI)	1.00 [0.94, 1.06]
3.2 More than one health check	6		Risk Ratio (Random, 95% CI)	0.99 [0.93, 1.05]
4 Total mortality - lifestyle intervention	9		Risk Ratio (Random, 95% CI)	0.99 [0.96, 1.03]
4.1 Major lifestyle intervention	4		Risk Ratio (Random, 95% CI)	0.99 [0.93, 1.06]
4.2 No major lifestyle intervention	5		Risk Ratio (Random, 95% CI)	1.00 [0.94, 1.06]
5 Total mortality - length of follow-up	9		Risk Ratio (Random, 95% CI)	0.99 [0.96, 1.03]
5.1 Up to five years	2		Risk Ratio (Random, 95% CI)	1.03 [0.66, 1.60]
5.2 More than 5 years	7		Risk Ratio (Random, 95% CI)	0.99 [0.95, 1.03]
6 Total mortality - age of trial	9		Risk Ratio (Random, 95% CI)	0.99 [0.96, 1.03]
6.1 Trial started before 1980	7		Risk Ratio (Random, 95% CI)	0.99 [0.95, 1.03]
6.2 Trial started after 1980	2		Risk Ratio (Random, 95% CI)	1.03 [0.66, 1.62]
7 Total mortality - geographical location	9		Risk Ratio (Random, 95% CI)	0.99 [0.96, 1.03]
7.1 USA	1		Risk Ratio (Random, 95% CI)	0.98 [0.88, 1.09]
7.2 Europe	8		Risk Ratio (Random, 95% CI)	0.99 [0.95, 1.03]
8 Total mortality - examination by physician	9		Risk Ratio (Random, 95% CI)	0.99 [0.96, 1.03]
8.1 Examination by physician	5		Risk Ratio (Random, 95% CI)	1.00 [0.94, 1.06]
8.2 No examination by physician	4		Risk Ratio (Random, 95% CI)	0.99 [0.93, 1.06]
9 Total mortality - selection bias	9		Risk Ratio (Random, 95% CI)	0.99 [0.96, 1.03]
9.1 low risk of selection bias	7		Risk Ratio (Random, 95% CI)	0.99 [0.94, 1.03]
9.2 Unclear risk of selection bias	2		Risk Ratio (Random, 95% CI)	1.00 [0.93, 1.08]
9.3 High risk of selection bias	0		Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
10 Total mortality - performance bias	9		Risk Ratio (Random, 95% CI)	0.99 [0.96, 1.03]
10.1 low risk	5		Risk Ratio (Random, 95% CI)	0.98 [0.94, 1.02]
10.2 Unclear risk	1		Risk Ratio (Random, 95% CI)	1.02 [0.94, 1.11]
10.3 High risk	3		Risk Ratio (Random, 95% CI)	1.08 [0.87, 1.33]
11 Total mortality - detection bias	9		Risk Ratio (Random, 95% CI)	0.99 [0.96, 1.03]

11.1 Low risk	6	Risk Ratio (Random, 95% CI)	0.99 [0.94, 1.04]
11.2 Unclear risk	2	Risk Ratio (Random, 95% CI)	1.00 [0.93, 1.08]
11.3 High risk	1	Risk Ratio (Random, 95% CI)	0.92 [0.77, 1.10]
12 Total mortality - incomplete outcome data	9	Risk Ratio (Random, 95% CI)	0.99 [0.96, 1.03]
12.1 Low risk	8	Risk Ratio (Random, 95% CI)	0.99 [0.95, 1.03]
12.2 Unclear risk	1	Risk Ratio (Random, 95% CI)	0.98 [0.88, 1.09]
12.3 High risk	0	Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
13 Total mortality - contamination	9	Risk Ratio (Random, 95% CI)	0.99 [0.96, 1.03]
13.1 Low risk	5	Risk Ratio (Random, 95% CI)	0.99 [0.95, 1.03]
13.2 Unclear risk	1	Risk Ratio (Random, 95% CI)	1.27 [0.95, 1.70]
13.3 High risk	3	Risk Ratio (Random, 95% CI)	0.99 [0.90, 1.10]
14 Cardiovascular mortality	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
15 Cardiovascular mortality - sensitivity analyses	5	Risk Ratio (Random, 95% CI)	0.99 [0.87, 1.12]
15.1 Excluding cluster trials	5	Risk Ratio (Random, 95% CI)	0.99 [0.87, 1.12]
16 Cardiovascular mortality - no. of health checks	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
16.1 Only one health check	3	Risk Ratio (Random, 95% CI)	0.89 [0.69, 1.14]
16.2 More than one health check	5	Risk Ratio (Random, 95% CI)	1.11 [0.95, 1.30]
17 Cardiovascular mortality lifestyle intervention	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
17.1 Major lifestyle intervention	3	Risk Ratio (Random, 95% CI)	0.99 [0.86, 1.15]
17.2 No major lifestyle intervention	5	Risk Ratio (Random, 95% CI)	1.03 [0.84, 1.27]
18 Cardiovascular mortality - length of follow-up	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
18.1 Up to five years	2	Risk Ratio (Random, 95% CI)	0.84 [0.22, 3.18]
18.2 More than five years	6	Risk Ratio (Random, 95% CI)	1.02 [0.94, 1.12]
19 Cardiovascular mortality - age of trial	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
19.1 Trial started before 1980	7	Risk Ratio (Random, 95% CI)	1.01 [0.90, 1.13]
19.2 Trial started after 1980	1	Risk Ratio (Random, 95% CI)	1.64 [0.97, 2.76]
20 Cardiovascular mortality - geographical location	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
20.1 Europe	7	Risk Ratio (Random, 95% CI)	1.04 [0.90, 1.20]
20.2 USA	1	Risk Ratio (Random, 95% CI)	1.01 [0.85, 1.20]
21 Cardiovascular mortality - examination by physician	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
21.1 Examination by physician	5	Risk Ratio (Random, 95% CI)	1.03 [0.84, 1.27]
21.2 No examination by physician	3	Risk Ratio (Random, 95% CI)	0.99 [0.86, 1.15]
22 Cardiovascular mortality - selection bias	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
22.1 Low risk	6	Risk Ratio (Random, 95% CI)	1.01 [0.88, 1.16]
22.2 Unclear risk	2	Risk Ratio (Random, 95% CI)	1.17 [0.71, 1.91]
22.3 High risk	0	Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]

23 Cardiovascular mortality - performance bias	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
23.1 Low risk	5	Risk Ratio (Random, 95% CI)	0.96 [0.85, 1.08]
23.2 Unclear risk	1	Risk Ratio (Random, 95% CI)	1.05 [0.91, 1.21]
23.3 High risk	2	Risk Ratio (Random, 95% CI)	1.57 [1.18, 2.09]
24 Cardiovascular mortality - detection bias	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
24.1 Low risk	5	Risk Ratio (Random, 95% CI)	1.00 [0.85, 1.17]
24.2 Unclear risk	2	Risk Ratio (Random, 95% CI)	1.17 [0.71, 1.91]
24.3 High risk	1	Risk Ratio (Random, 95% CI)	1.09 [0.83, 1.43]
25 Cardiovascular mortality - incomplete outcome data	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
25.1 Low risk	7	Risk Ratio (Random, 95% CI)	1.04 [0.90, 1.20]
25.2 Unclear risk	1	Risk Ratio (Random, 95% CI)	1.01 [0.85, 1.20]
25.3 High risk	0	Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
26 Cardiovascular mortality - contamination	8	Risk Ratio (Random, 95% CI)	1.03 [0.91, 1.17]
26.1 Low risk	5	Risk Ratio (Random, 95% CI)	0.97 [0.86, 1.09]
26.2 Unclear risk	1	Risk Ratio (Random, 95% CI)	1.64 [0.97, 2.76]
26.3 High risk	2	Risk Ratio (Random, 95% CI)	1.21 [0.81, 1.83]
27 Cancer mortality	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
28 Cancer mortality - sensitivity analyses	5	Risk Ratio (Random, 95% CI)	0.97 [0.85, 1.09]
28.1 Excluding cluster trials	5	Risk Ratio (Random, 95% CI)	0.97 [0.85, 1.09]
29 Cancer mortality - no. of health checks	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
29.1 Only one health check	3	Risk Ratio (Random, 95% CI)	1.10 [1.00, 1.21]
29.2 More than one health check	5	Risk Ratio (Random, 95% CI)	0.92 [0.83, 1.02]
30 Cancer mortality lifestyle intervention	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
30.1 Major lifestyle intervention	3	Risk Ratio (Random, 95% CI)	1.01 [0.82, 1.24]
30.2 No major lifestyle intervention	5	Risk Ratio (Random, 95% CI)	1.02 [0.91, 1.15]
31 Cancer mortality - length of follow-up	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
31.1 Up to five years	2	Risk Ratio (Random, 95% CI)	1.33 [0.89, 1.99]
31.2 More than five years	6	Risk Ratio (Random, 95% CI)	1.00 [0.90, 1.10]
32 Cancer mortality - age of trial	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
32.1 Trial started before 1980	7	Risk Ratio (Random, 95% CI)	1.01 [0.91, 1.12]
32.2 Trial started after 1980	1	Risk Ratio (Random, 95% CI)	1.19 [0.75, 1.89]
33 Cancer mortality - geographical location	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
33.1 Europe	7	Risk Ratio (Random, 95% CI)	1.02 [0.91, 1.15]
33.2 USA	1	Risk Ratio (Random, 95% CI)	0.98 [0.80, 1.20]
34 Cancer mortality - examination by physician	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
34.1 Examination by physician	5	Risk Ratio (Random, 95% CI)	1.02 [0.91, 1.15]

34.2 No examination by physician	3	Risk Ratio (Random, 95% CI)	1.01 [0.82, 1.24]
35 Cancer mortality - selection bias	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
35.1 Low risk	6	Risk Ratio (Random, 95% CI)	0.98 [0.87, 1.10]
35.2 Unclear risk	2	Risk Ratio (Random, 95% CI)	1.10 [0.98, 1.24]
35.3 High risk	0	Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
36 Cancer mortality - performance bias	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
36.1 Low risk	5	Risk Ratio (Random, 95% CI)	1.00 [0.86, 1.16]
36.2 Unclear risk	1	Risk Ratio (Random, 95% CI)	1.05 [0.88, 1.25]
36.3 High risk	2	Risk Ratio (Random, 95% CI)	1.08 [0.80, 1.46]
37 Cancer mortality - detection bias	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
37.1 Low risk	5	Risk Ratio (Random, 95% CI)	0.99 [0.86, 1.13]
37.2 Unclear risk	2	Risk Ratio (Random, 95% CI)	1.10 [0.98, 1.24]
37.3 High risk	1	Risk Ratio (Random, 95% CI)	0.93 [0.63, 1.38]
38 Cancer mortality - incomplete outcome data	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
38.1 Low risk	6	Risk Ratio (Random, 95% CI)	0.98 [0.86, 1.12]
38.2 Unclear risk	2	Risk Ratio (Random, 95% CI)	1.07 [0.96, 1.20]
38.3 High risk	0	Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
39 Cancer mortality - contamination	8	Risk Ratio (Random, 95% CI)	1.01 [0.92, 1.12]
39.1 Low risk	5	Risk Ratio (Random, 95% CI)	1.01 [0.88, 1.17]
39.2 Unclear risk	1	Risk Ratio (Random, 95% CI)	1.19 [0.75, 1.89]
39.3 High risk	2	Risk Ratio (Random, 95% CI)	0.99 [0.82, 1.18]

Analysis 1.1. Comparison 1 Health checks versus control, Outcome 1 Total mortality.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 1 Total mortality

Study or subgroup	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
Göteborg 1963	-0.08338161 (0.09329714)		4.5 %	0.92 [0.77, 1.10]
Kaiser Permanente 1965	-0.02020271 (0.05459466)		13.2 %	0.98 [0.88, 1.09]
South-East London 1967	0.09531018 (0.10343498)		3.7 %	1.10 [0.90, 1.35]
Malmö 1969	-0.21072103 (0.18796486)		1.1 %	0.81 [0.56, 1.17]
Stockholm 1969	0.01980263 (0.04240699)		21.9 %	1.02 [0.94, 1.11]
Göteborg 1970	-0.02020271 (0.0312761)		40.3 %	0.98 [0.92, 1.04]
WHO 1971	-0.05445619 (0.05632343)		12.4 %	0.95 [0.85, 1.06]
OXCHECK 1989	0.2390169 (0.14994558)		1.8 %	1.27 [0.95, 1.70]
Ebeltoft 1992	-0.22314355 (0.20581798)		0.9 %	0.80 [0.53, 1.20]
Total (95% CI)			**100.0 %**	**0.99 [0.95, 1.03]**

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 7.85$, df = 8 (P = 0.45); $I^2 = 0.0\%$

Test for overall effect: Z = 0.69 (P = 0.49)

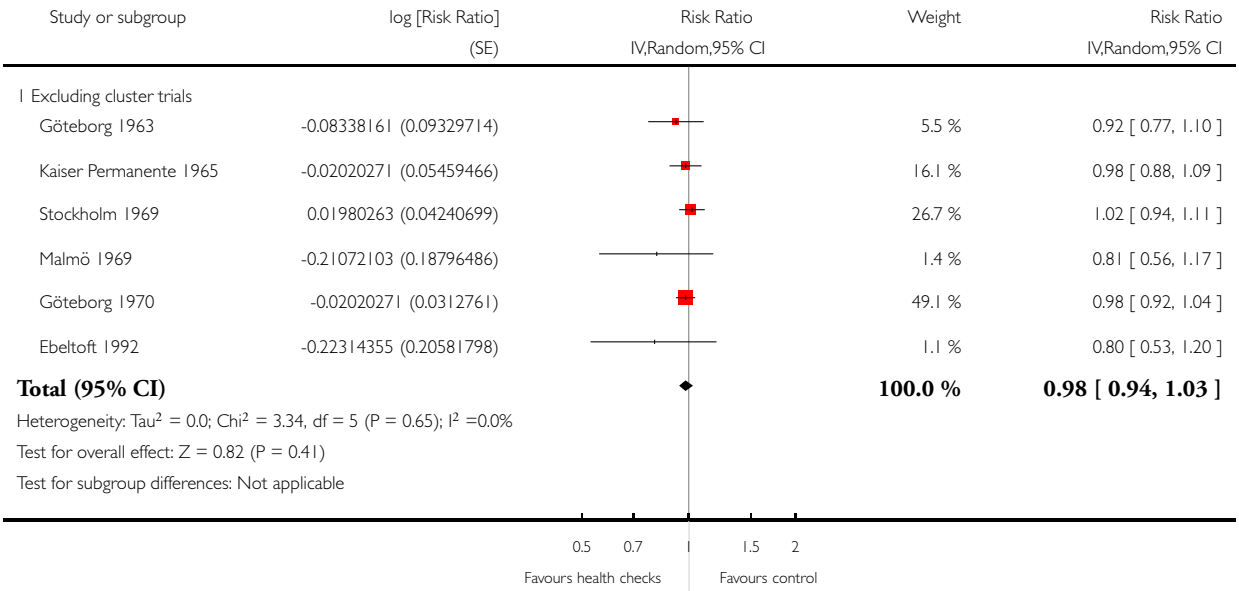
Test for subgroup differences: Not applicable

Analysis 1.2. Comparison 1 Health checks versus control, Outcome 2 Total mortality - sensitivity analyses.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 2 Total mortality - sensitivity analyses

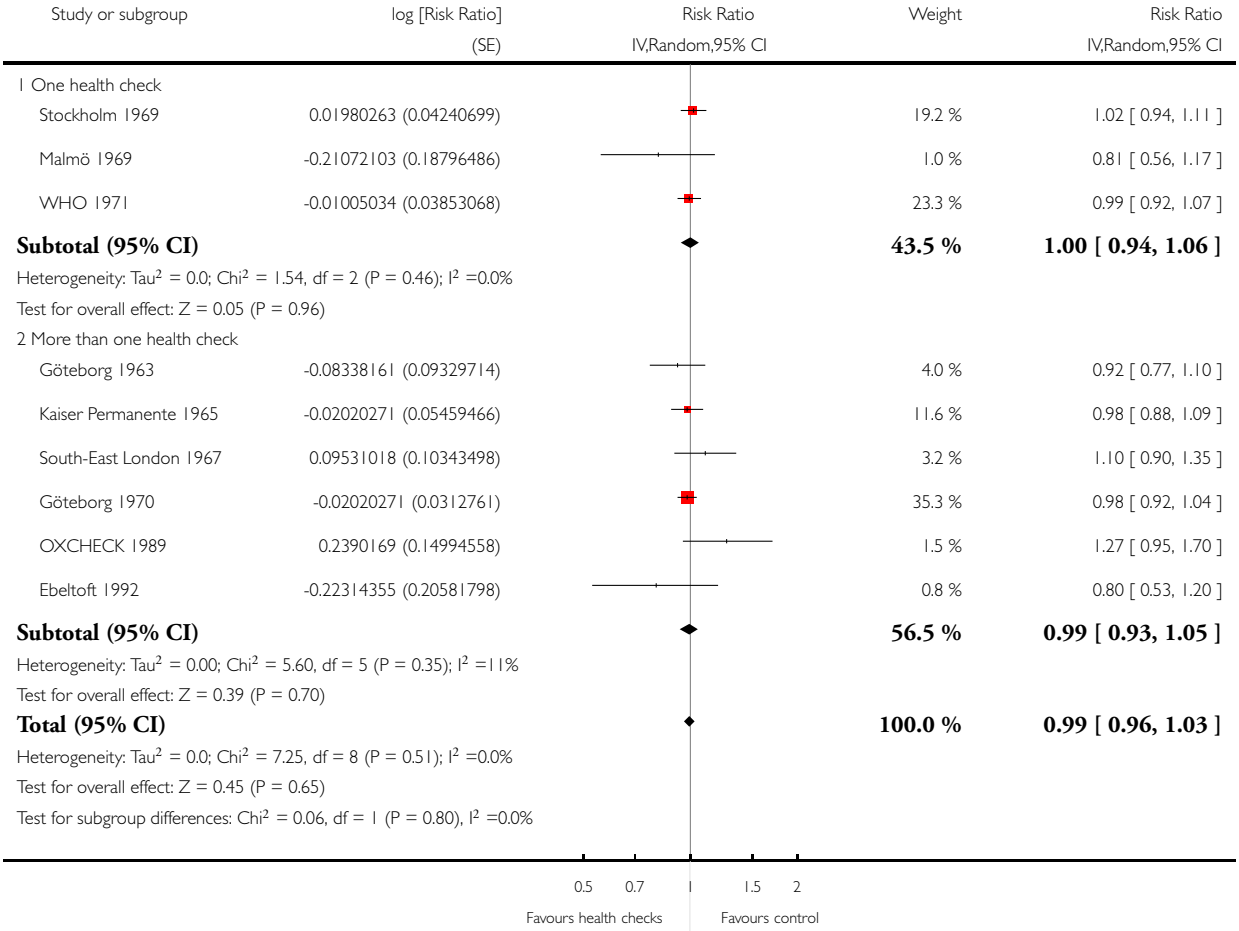


Analysis 1.3. Comparison 1 Health checks versus control, Outcome 3 Total mortality - no. of health checks.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 3 Total mortality - no. of health checks

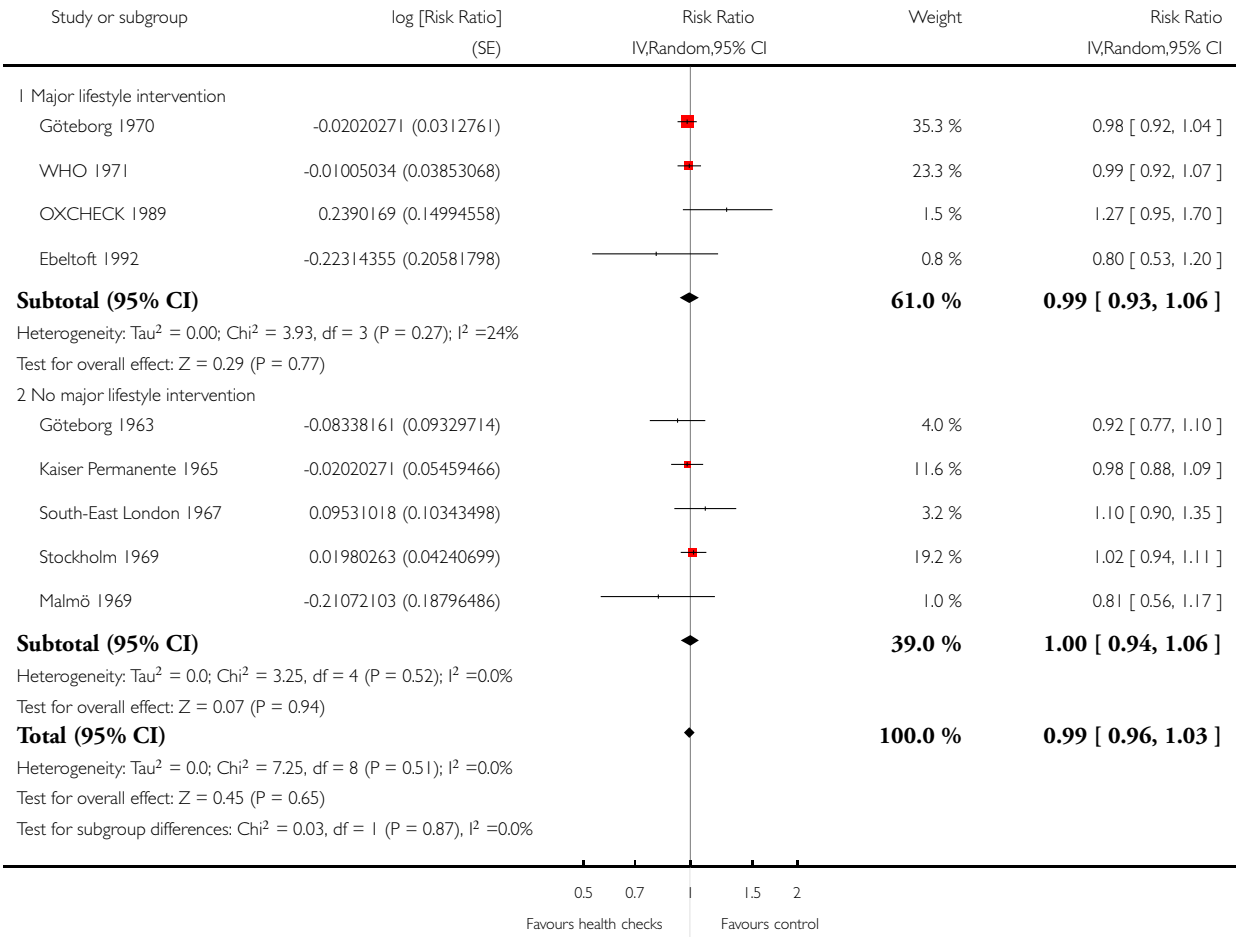


Analysis 1.4. Comparison 1 Health checks versus control, Outcome 4 Total mortality - lifestyle intervention.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 4 Total mortality - lifestyle intervention

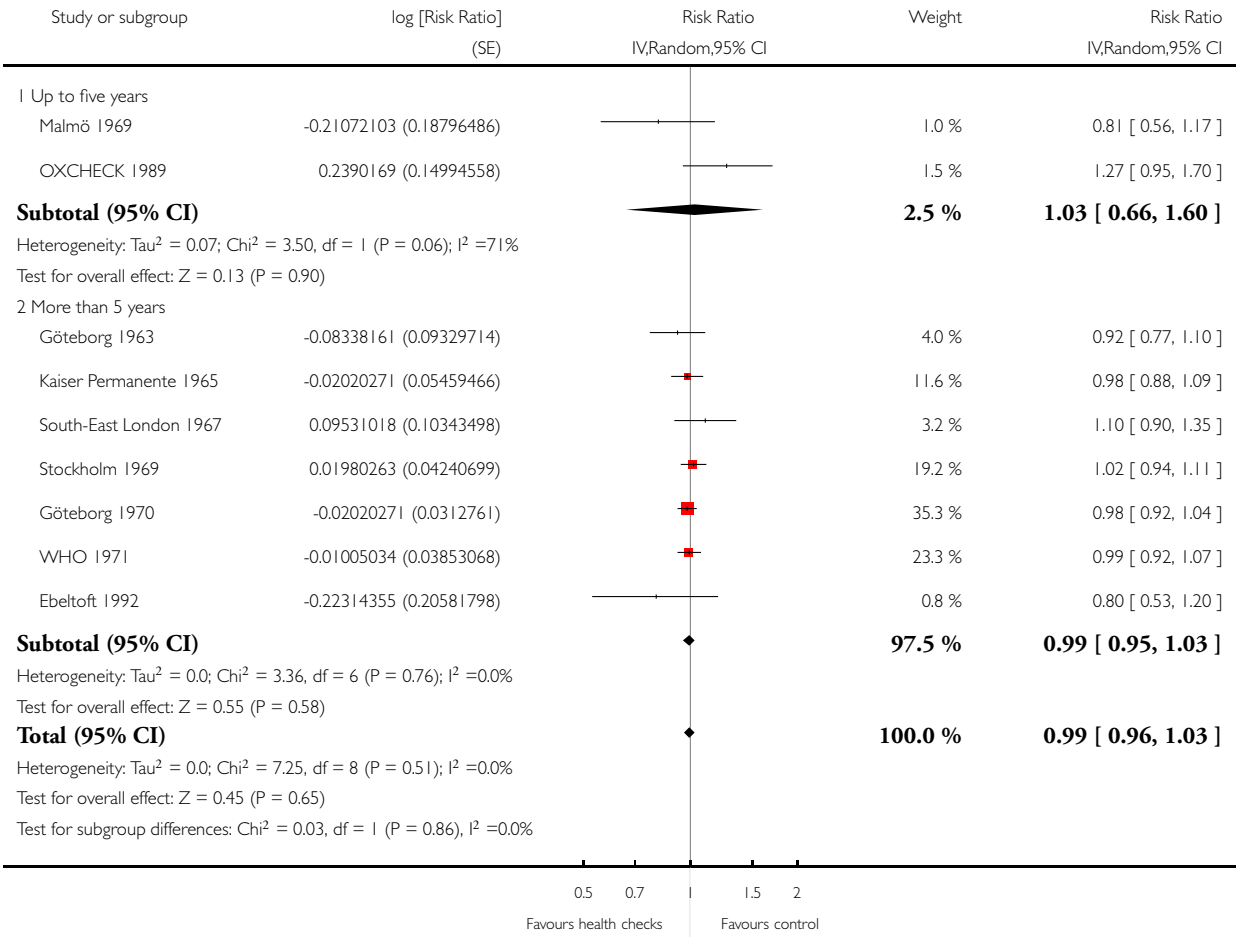


Analysis 1.5. Comparison 1 Health checks versus control, Outcome 5 Total mortality - length of follow-up.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 5 Total mortality - length of follow-up

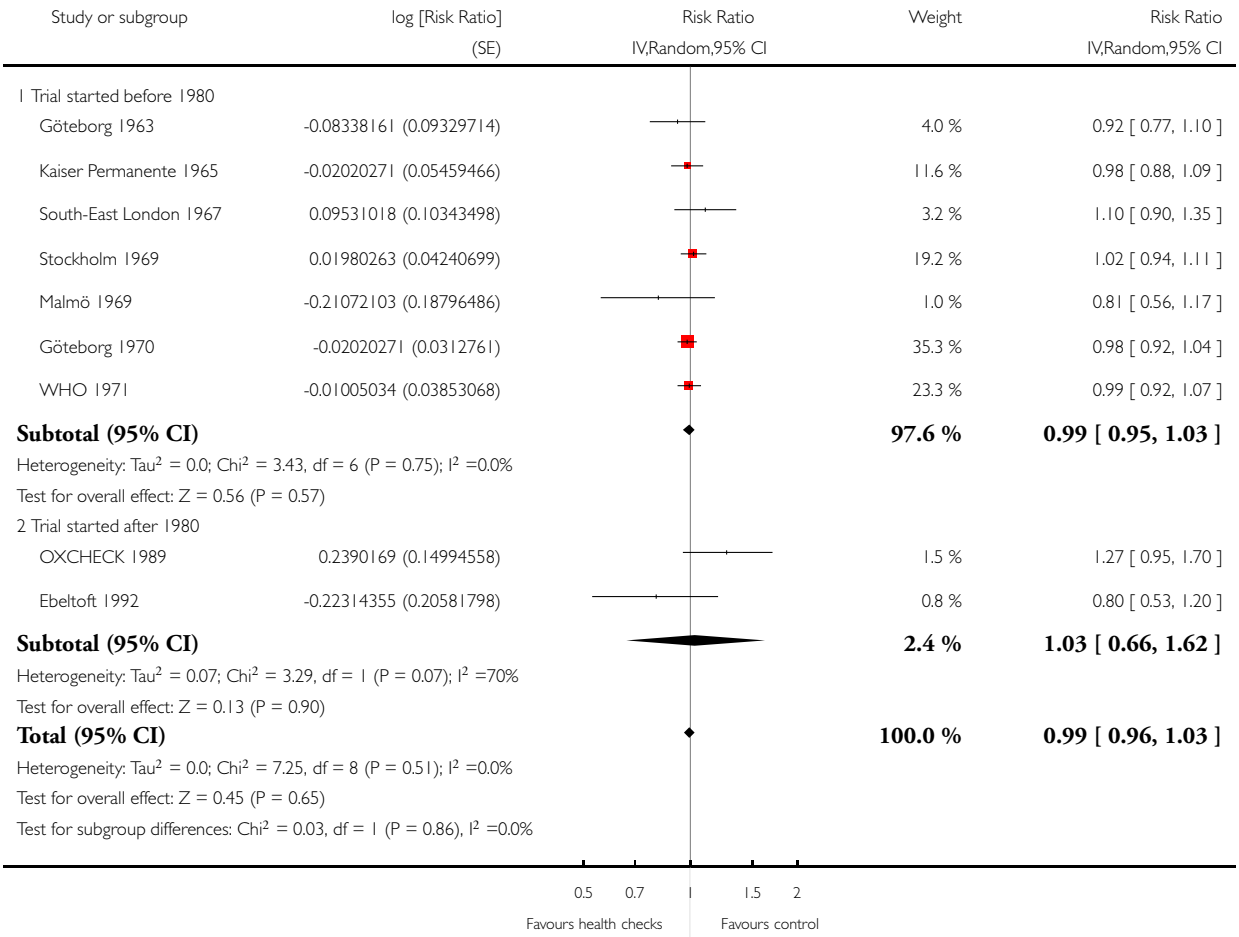


Analysis 1.6. Comparison 1 Health checks versus control, Outcome 6 Total mortality - age of trial.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 6 Total mortality - age of trial

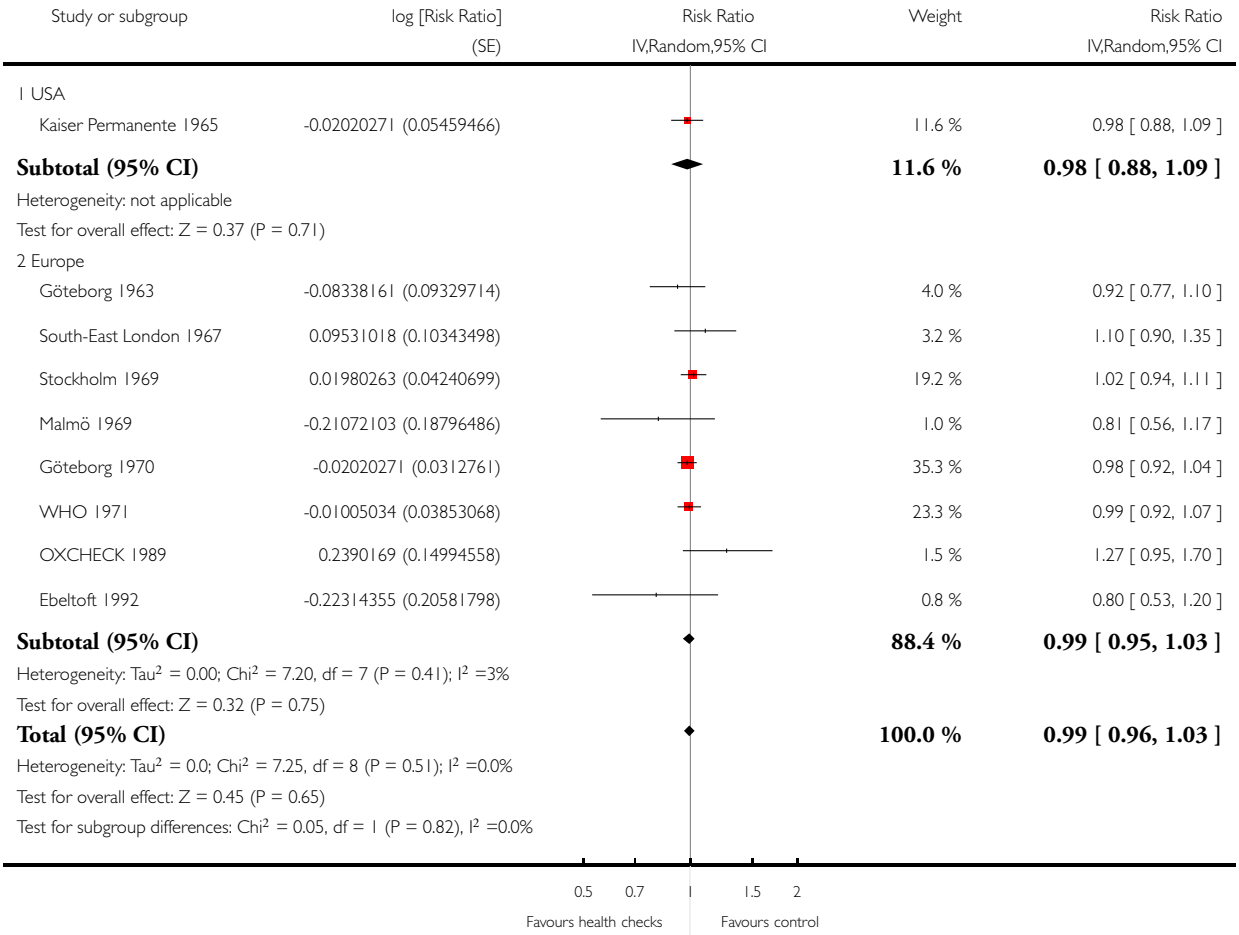


Analysis 1.7. Comparison 1 Health checks versus control, Outcome 7 Total mortality - geographical location.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 7 Total mortality - geographical location

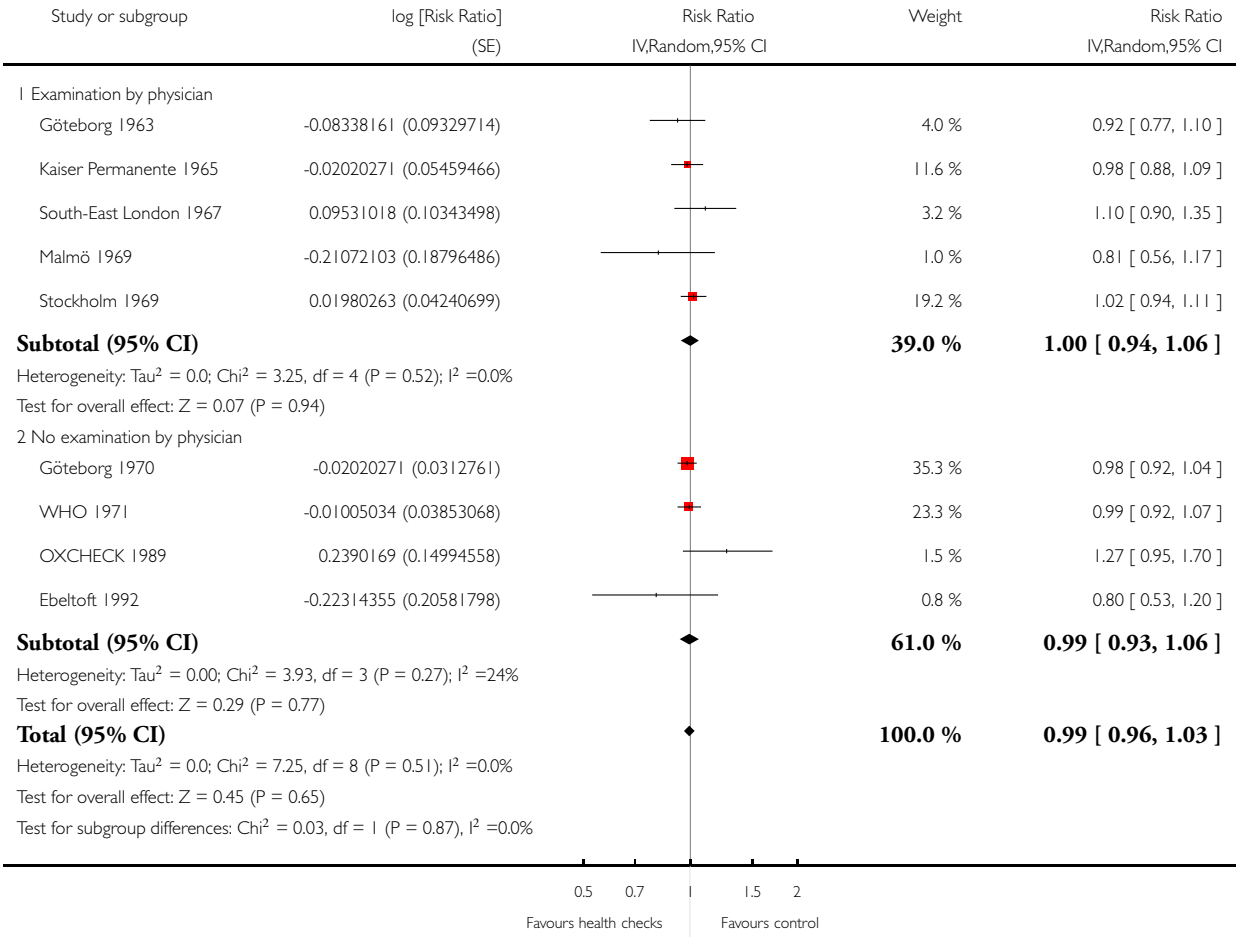


Analysis 1.8. Comparison 1 Health checks versus control, Outcome 8 Total mortality - examination by physician.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 8 Total mortality - examination by physician

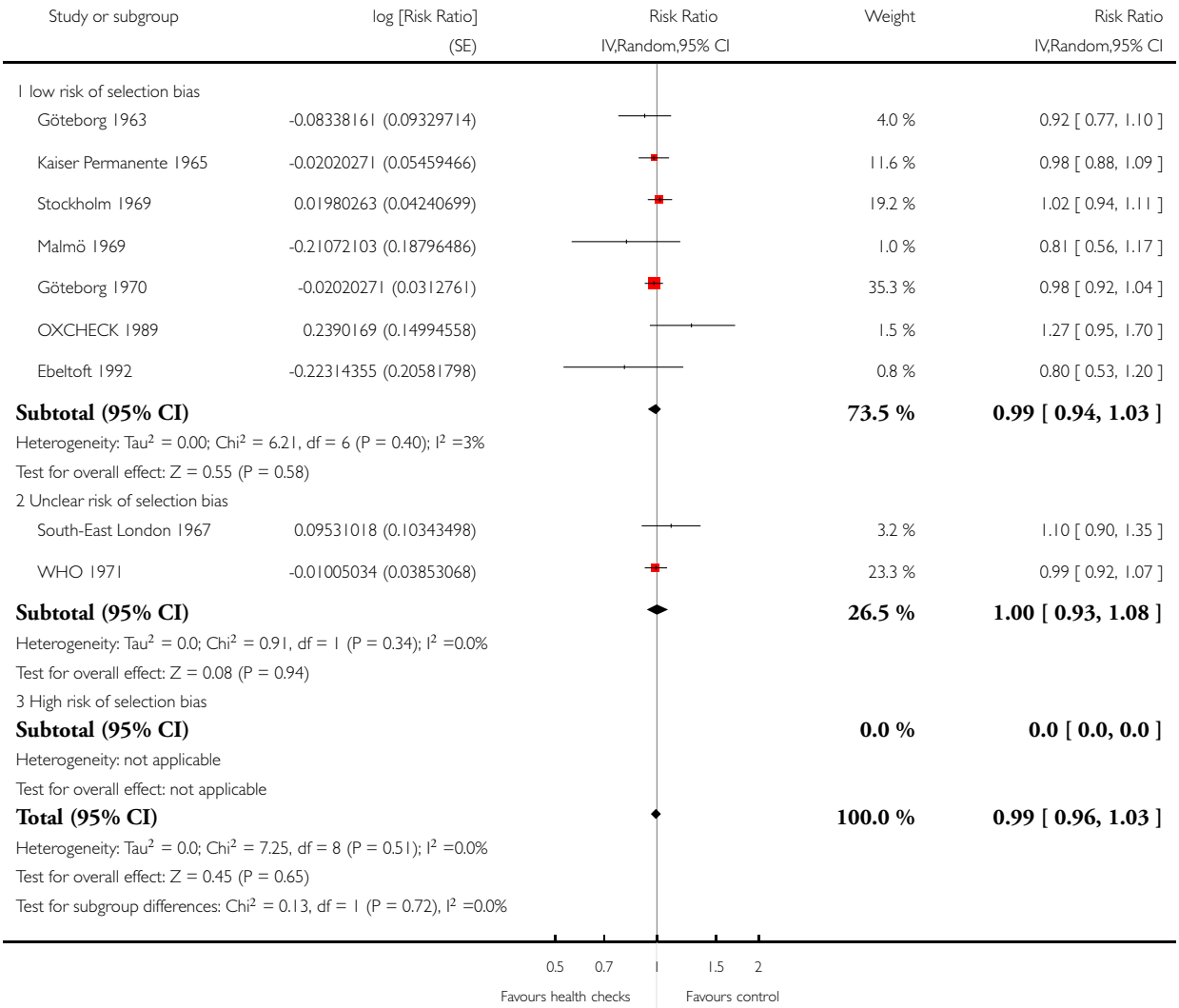


Analysis 1.9. Comparison 1 Health checks versus control, Outcome 9 Total mortality - selection bias.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 9 Total mortality - selection bias



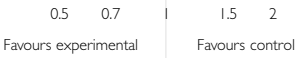
Analysis 1.10. Comparison 1 Health checks versus control, Outcome 10 Total mortality - performance bias.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 10 Total mortality - performance bias

Study or subgroup	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
1 low risk				
Göteborg 1963	-0.08338161 (0.09329714)		4.0 %	0.92 [0.77, 1.10]
Kaiser Permanente 1965	-0.02020271 (0.05459466)		11.6 %	0.98 [0.88, 1.09]
Malmö 1969	-0.21072103 (0.18796486)		1.0 %	0.81 [0.56, 1.17]
Göteborg 1970	-0.02020271 (0.0312761)		35.3 %	0.98 [0.92, 1.04]
WHO 1971	-0.01005034 (0.03853068)		23.3 %	0.99 [0.92, 1.07]
Subtotal (95% CI)			**75.2 %**	**0.98 [0.94, 1.02]**
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 1.54$, df = 4 (P = 0.82); $I^2 = 0.0\%$				
Test for overall effect: Z = 1.07 (P = 0.29)				
2 Unclear risk				
Stockholm 1969	0.01980263 (0.04240699)		19.2 %	1.02 [0.94, 1.11]
Subtotal (95% CI)			**19.2 %**	**1.02 [0.94, 1.11]**
Heterogeneity: not applicable				
Test for overall effect: Z = 0.47 (P = 0.64)				
3 High risk				
South-East London 1967	0.09531018 (0.10343498)		3.2 %	1.10 [0.90, 1.35]
OXCHECK 1989	0.2390169 (0.14994558)		1.5 %	1.27 [0.95, 1.70]
Ebeltoft 1992	-0.22314355 (0.20581798)		0.8 %	0.80 [0.53, 1.20]
Subtotal (95% CI)			**5.6 %**	**1.08 [0.87, 1.33]**
Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 3.30$, df = 2 (P = 0.19); $I^2 = 39\%$				
Test for overall effect: Z = 0.69 (P = 0.49)				
Total (95% CI)			**100.0 %**	**0.99 [0.96, 1.03]**
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 7.25$, df = 8 (P = 0.51); $I^2 = 0.0\%$				
Test for overall effect: Z = 0.45 (P = 0.65)				
Test for subgroup differences: $Chi^2 = 1.47$, df = 2 (P = 0.48), $I^2 = 0.0\%$				

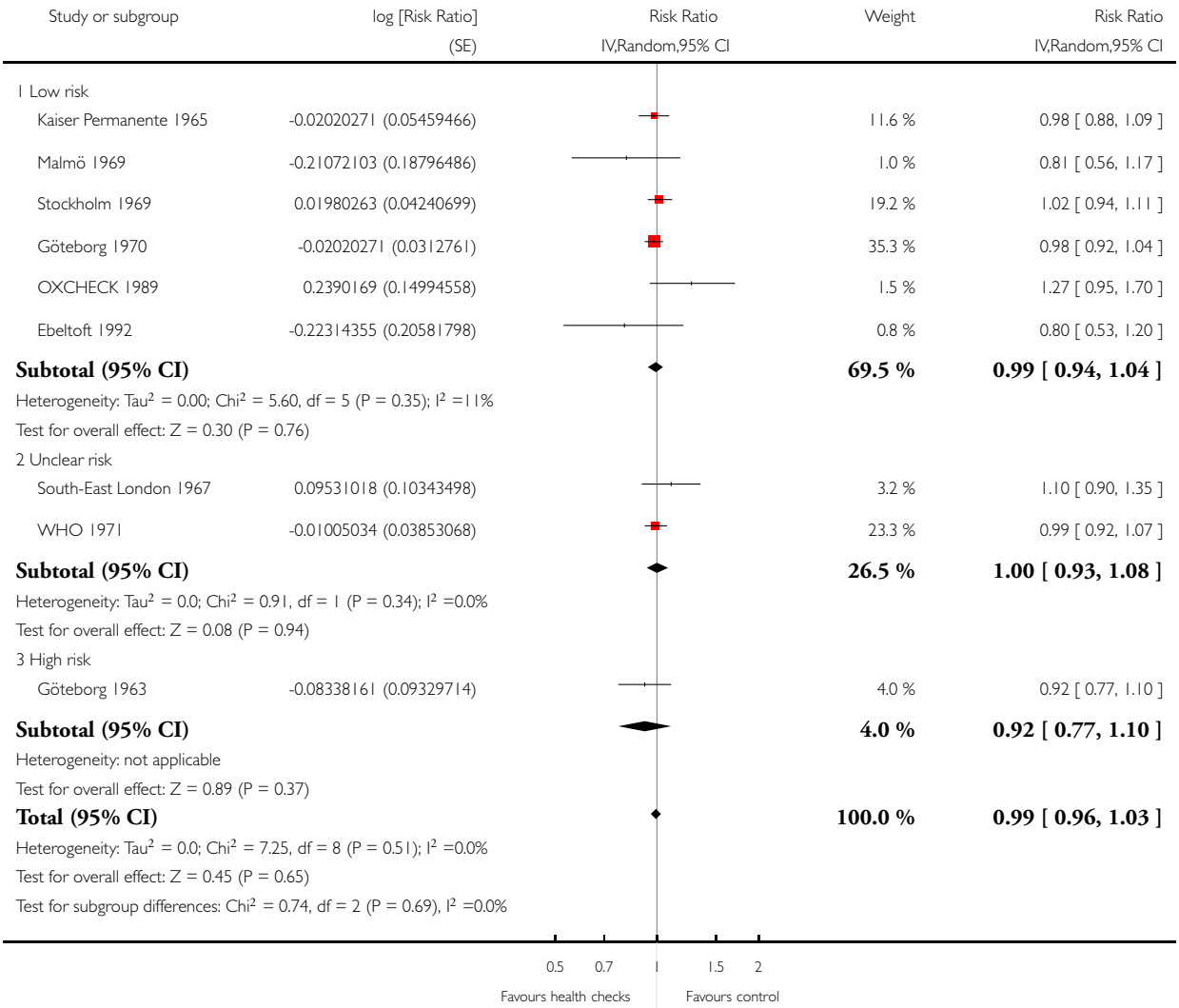


Analysis 1.11. Comparison 1 Health checks versus control, Outcome 11 Total mortality - detection bias.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 11 Total mortality - detection bias

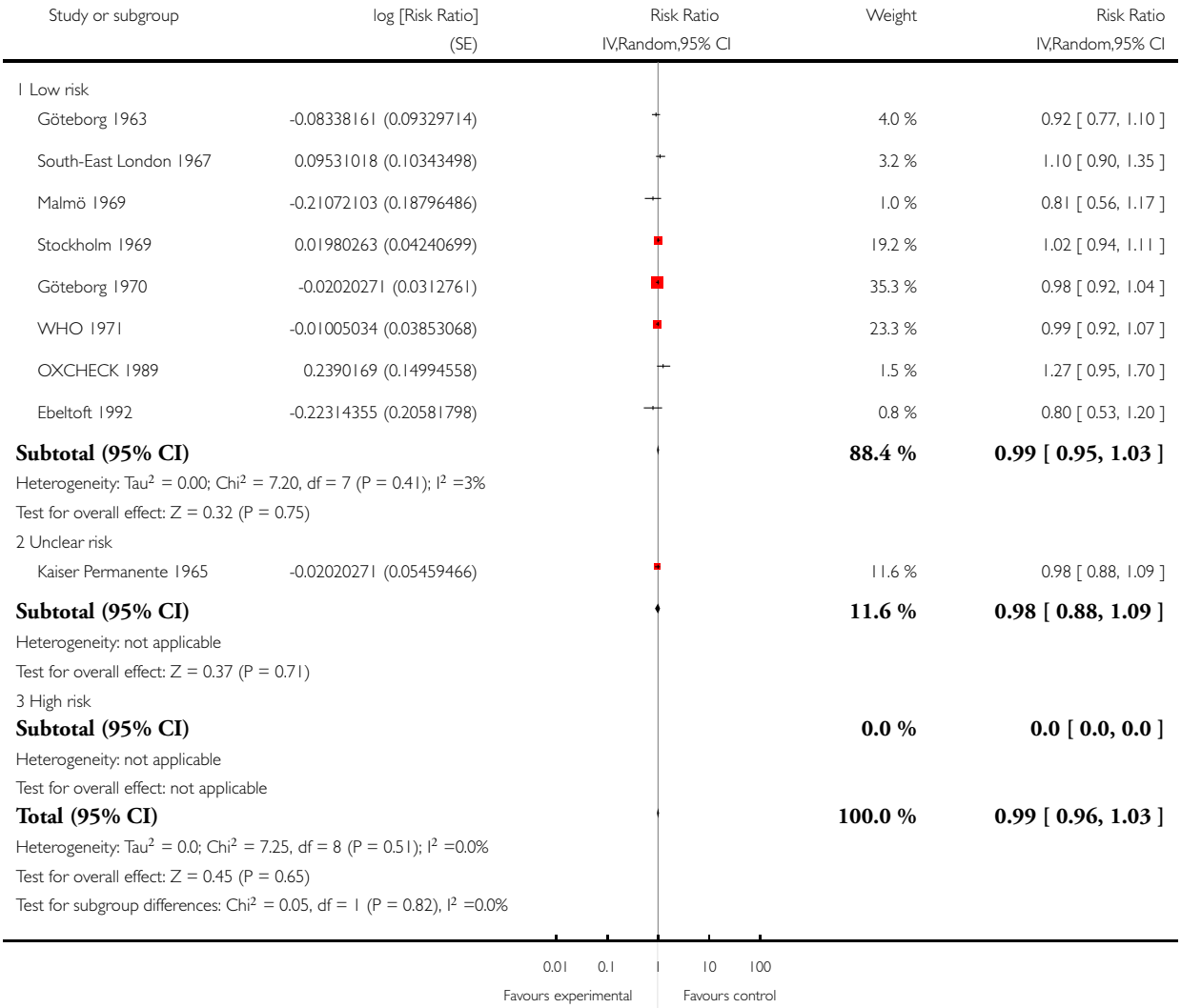


Analysis 1.12. Comparison 1 Health checks versus control, Outcome 12 Total mortality - incomplete outcome data.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 12 Total mortality - incomplete outcome data

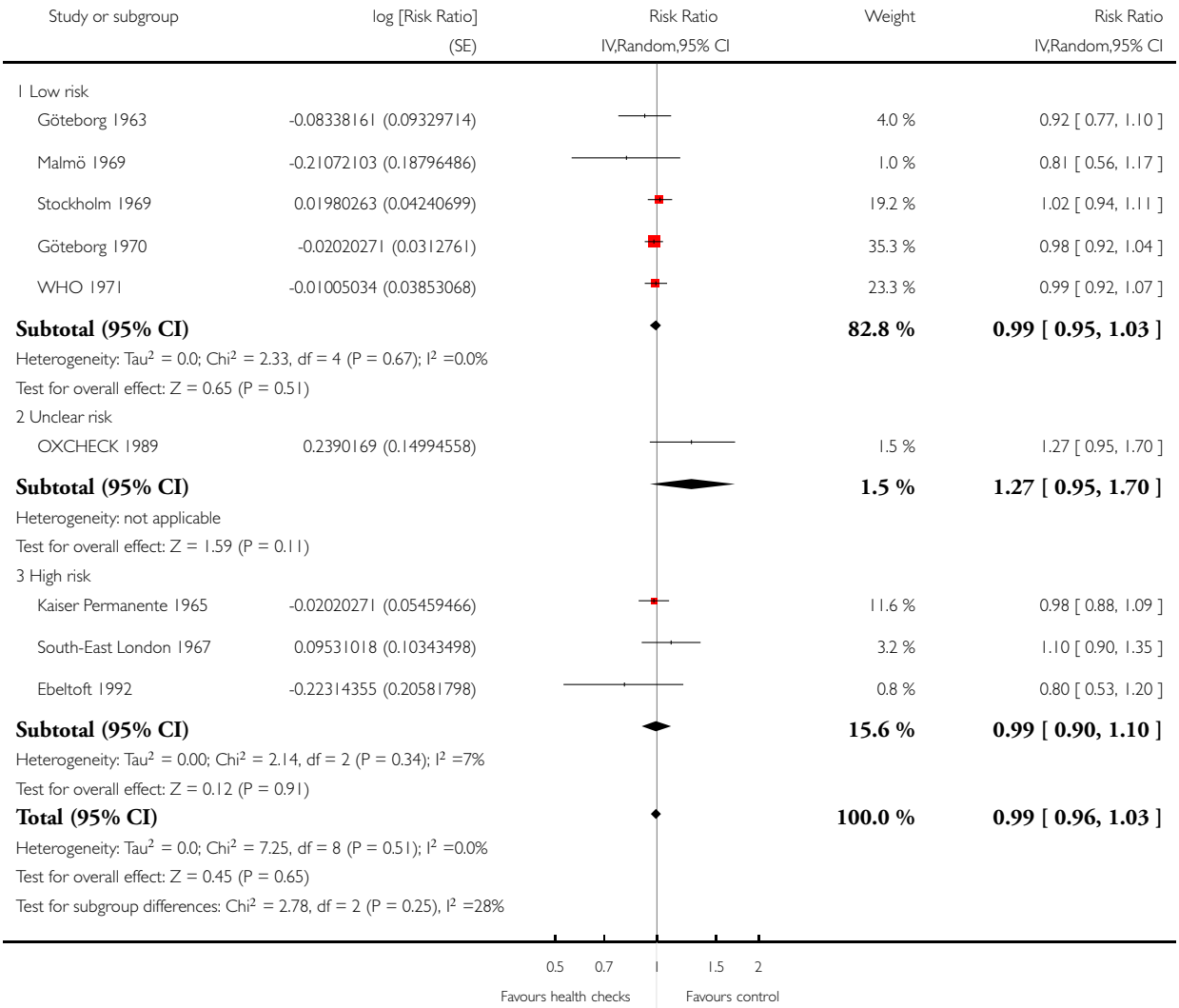


Analysis 1.13. Comparison 1 Health checks versus control, Outcome 13 Total mortality - contamination.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 13 Total mortality - contamination



Analysis 1.14. Comparison 1 Health checks versus control, Outcome 14 Cardiovascular mortality.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

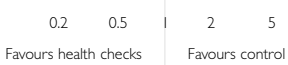
Outcome: 14 Cardiovascular mortality

Study or subgroup	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
Göteborg 1963	0.0861777 (0.13877654)		10.9 %	1.09 [0.83, 1.43]
Kaiser Permanente 1965	0.00995033 (0.08796951)		16.2 %	1.01 [0.85, 1.20]
South-East London 1967	0.43178242 (0.17447207)		8.3 %	1.54 [1.09, 2.17]
Malmö 1969	-0.86750057 (0.31153434)		3.4 %	0.42 [0.23, 0.77]
Stockholm 1969	0.04879016 (0.07268649)		18.0 %	1.05 [0.91, 1.21]
Göteborg 1970	-0.02020271 (0.05224347)		20.5 %	0.98 [0.88, 1.09]
WHO 1971	-0.071496 (0.07101522)		18.2 %	0.93 [0.81, 1.07]
OXCHECK 1989	0.49469624 (0.26583166)		4.5 %	1.64 [0.97, 2.76]
Total (95% CI)			**100.0 %**	**1.03 [0.91, 1.17]**

Heterogeneity: $Tau^2 = 0.02$; $Chi^2 = 19.32$, df = 7 (P = 0.01); $I^2 = 64\%$

Test for overall effect: Z = 0.50 (P = 0.62)

Test for subgroup differences: Not applicable



Analysis 1.15. Comparison 1 Health checks versus control, Outcome 15 Cardiovascular mortality - sensitivity analyses.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

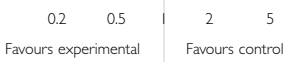
Outcome: 15 Cardiovascular mortality - sensitivity analyses

Study or subgroup	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
1 Excluding cluster trials				
Göteborg 1963	0.0861777 (0.13877654)		13.9 %	1.09 [0.83, 1.43]
Kaiser Permanente 1965	0.00995033 (0.08796951)		23.1 %	1.01 [0.85, 1.20]
Stockholm 1969	0.04879016 (0.07268649)		26.9 %	1.05 [0.91, 1.21]
Malmö 1969	-0.86750057 (0.31153434)		3.8 %	0.42 [0.23, 0.77]
Göteborg 1970	-0.02020271 (0.05224347)		32.4 %	0.98 [0.88, 1.09]
Total (95% CI)			**100.0 %**	**0.99 [0.87, 1.12]**

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 8.75$, df = 4 (P = 0.07); $I^2 = 54\%$

Test for overall effect: Z = 0.19 (P = 0.85)

Test for subgroup differences: Not applicable

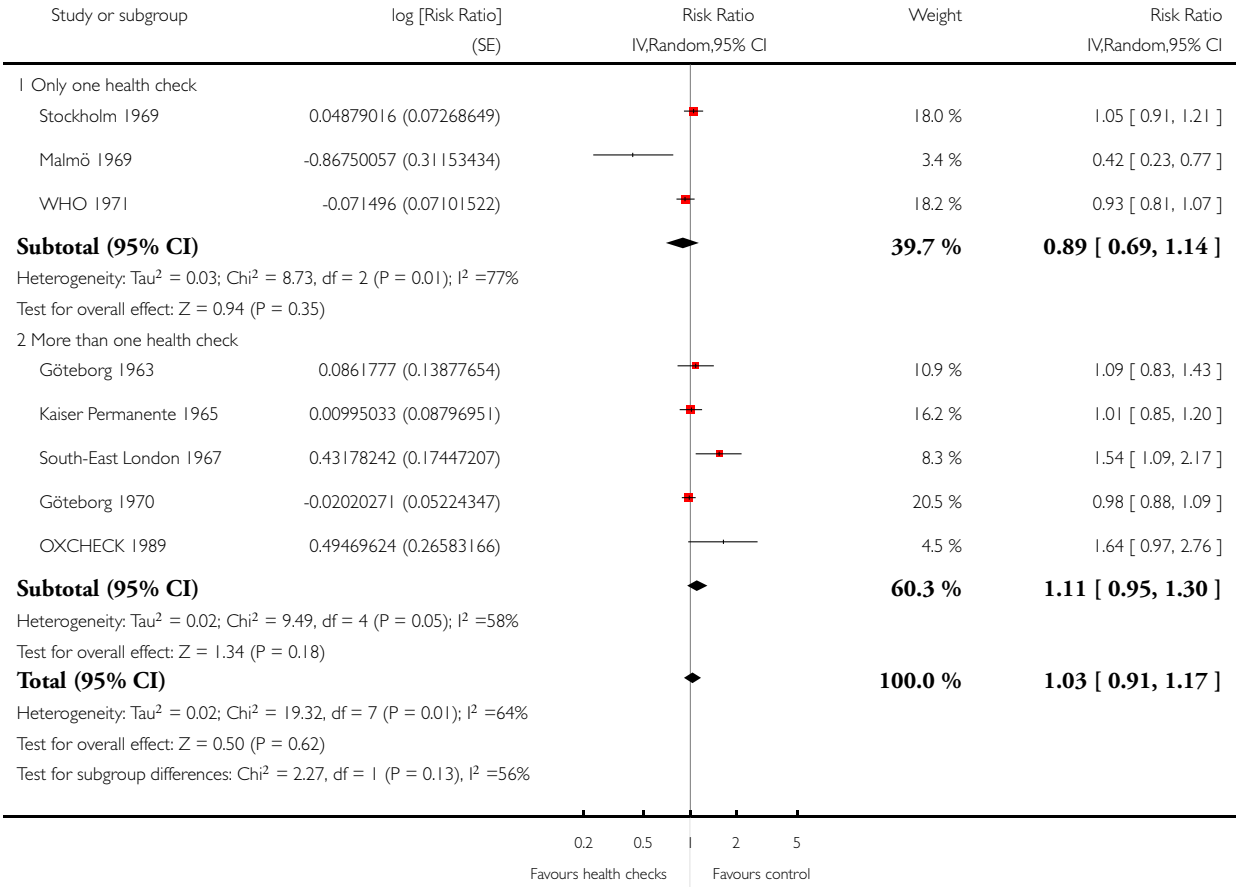


Analysis 1.16. Comparison 1 Health checks versus control, Outcome 16 Cardiovascular mortality - no. of health checks.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 16 Cardiovascular mortality - no. of health checks

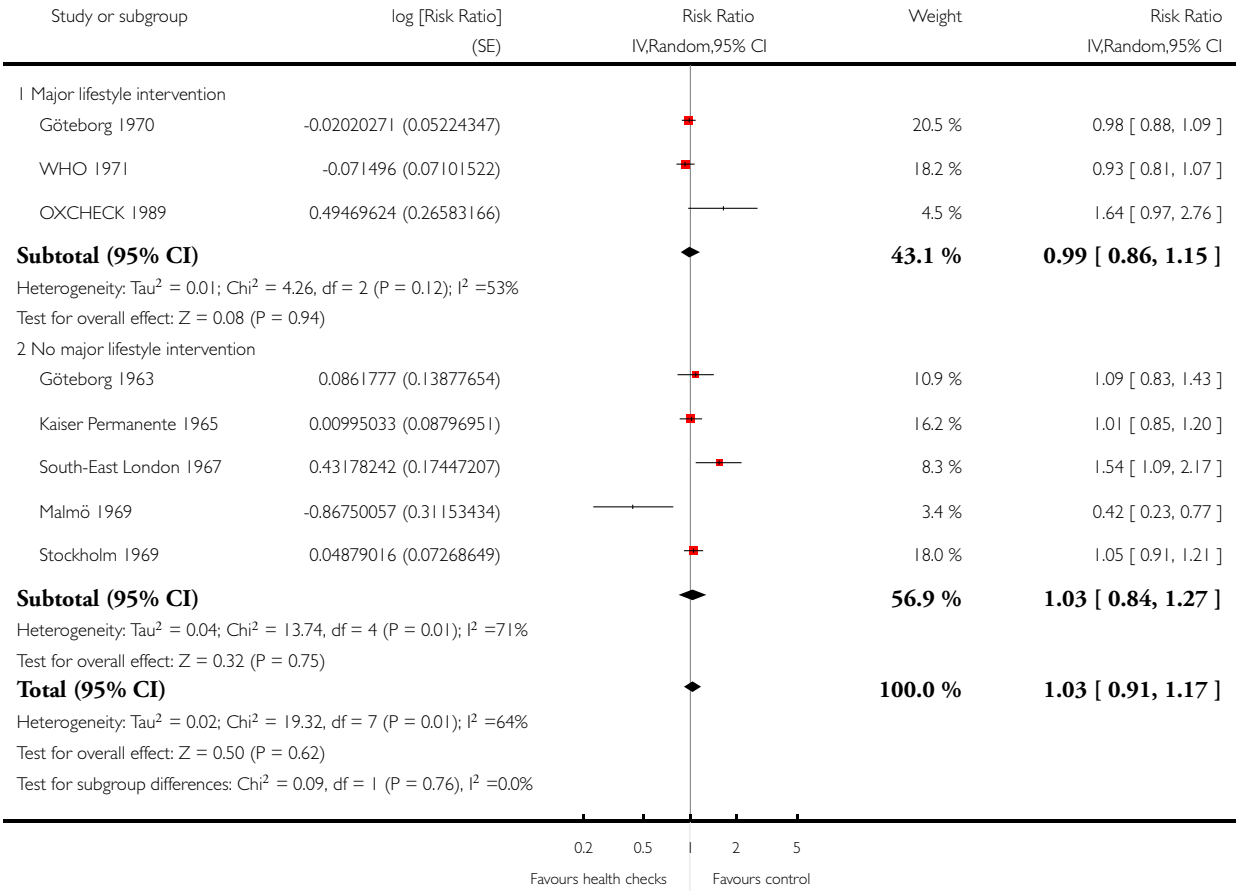


Analysis 1.17. Comparison 1 Health checks versus control, Outcome 17 Cardiovascular mortality lifestyle intervention.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 17 Cardiovascular mortality lifestyle intervention

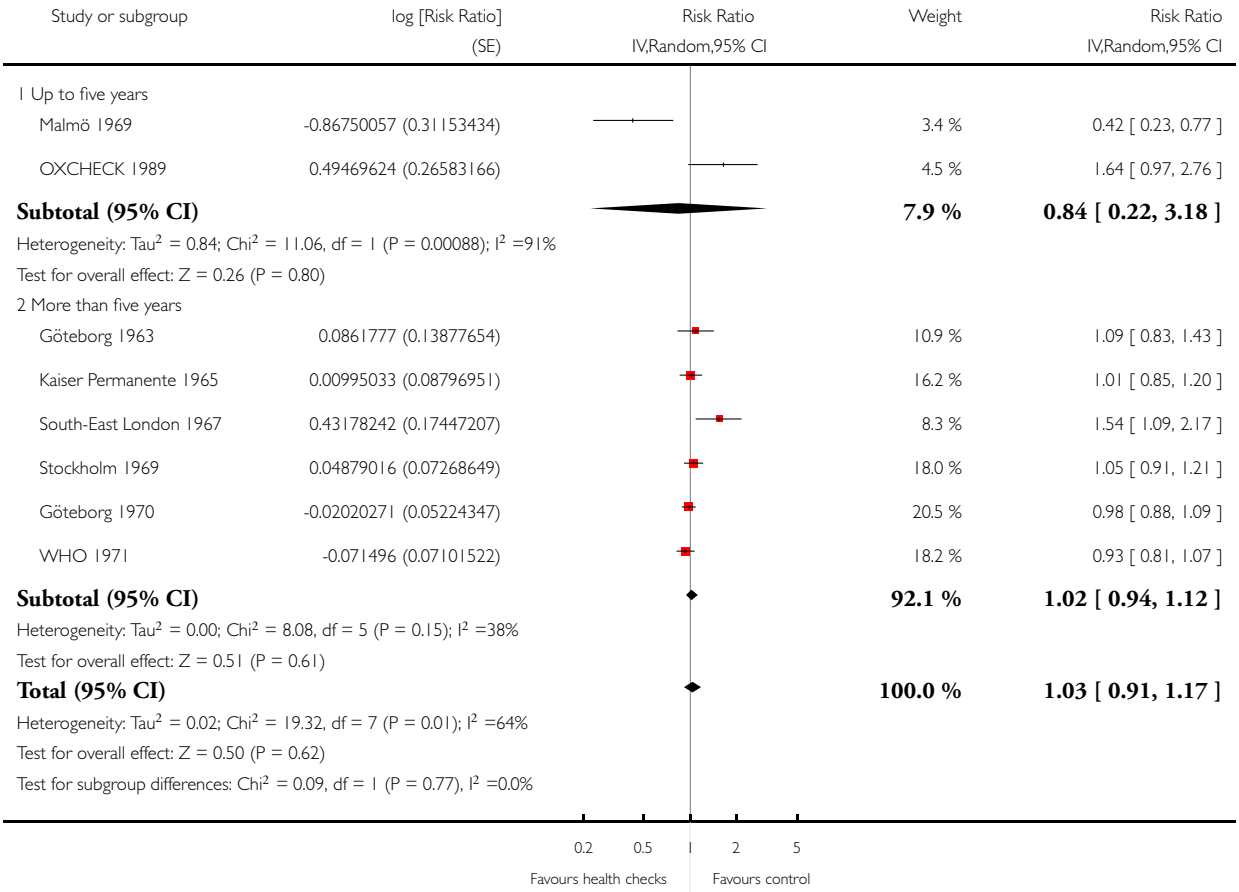


Analysis 1.18. Comparison 1 Health checks versus control, Outcome 18 Cardiovascular mortality - length of follow-up.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 18 Cardiovascular mortality - length of follow-up

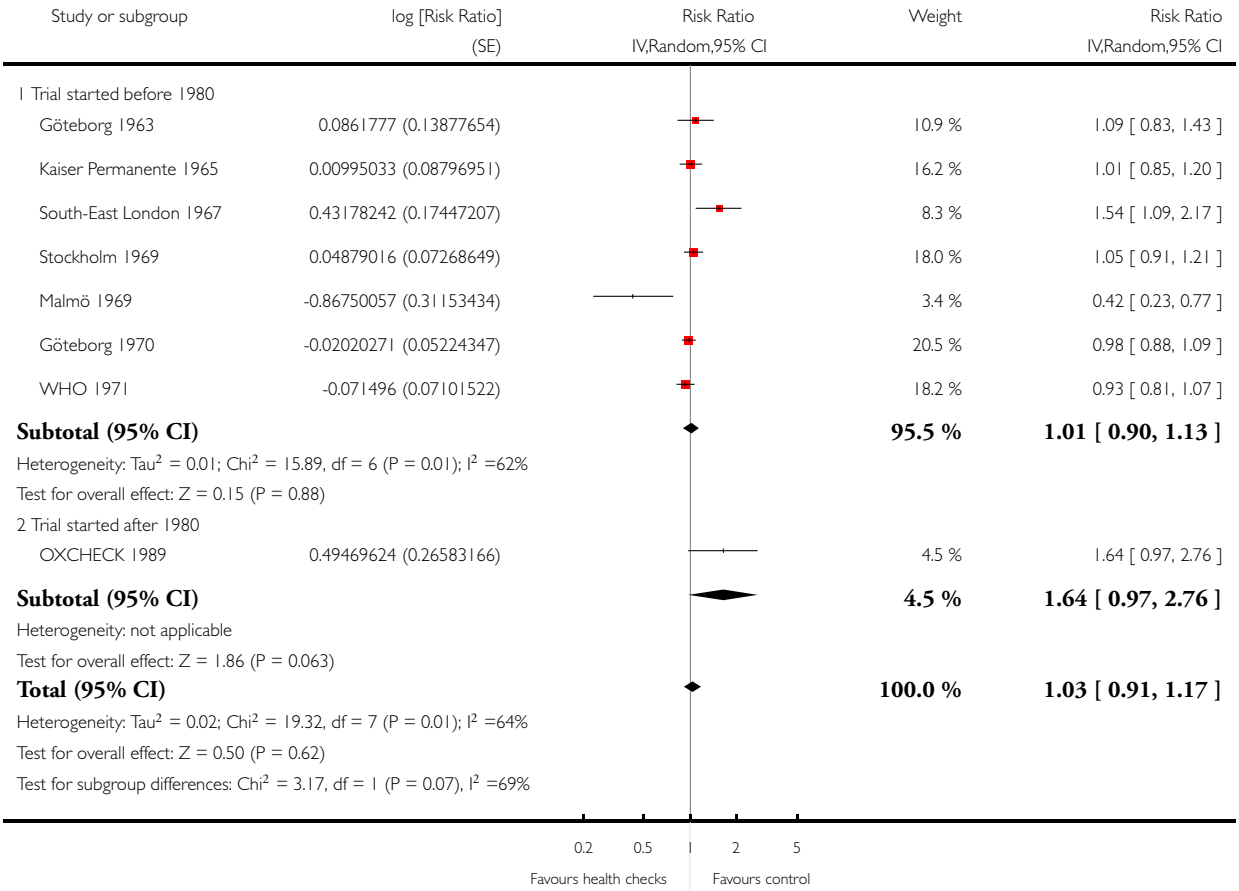


Analysis 1.19. Comparison 1 Health checks versus control, Outcome 19 Cardiovascular mortality - age of trial.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 19 Cardiovascular mortality - age of trial

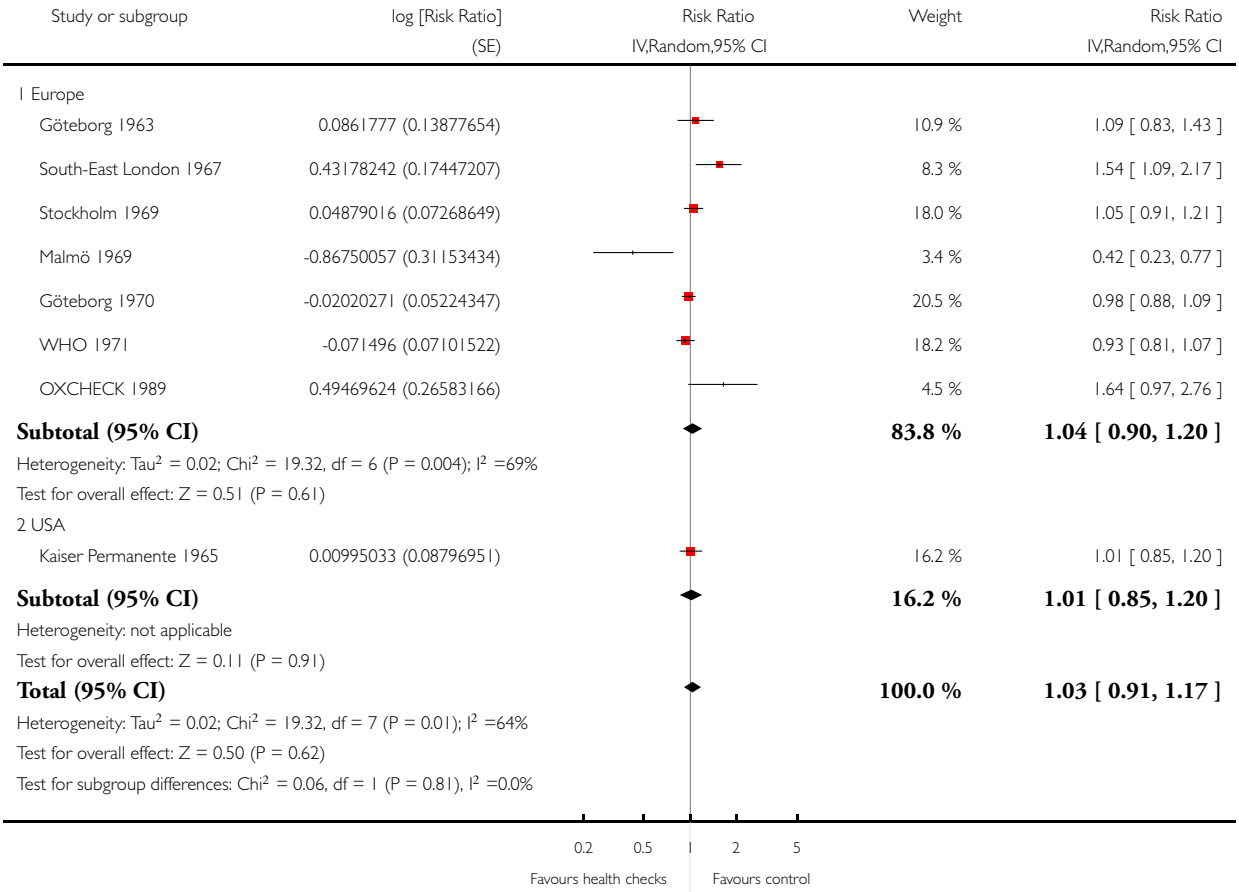


Analysis 1.20. Comparison 1 Health checks versus control, Outcome 20 Cardiovascular mortality - geographical location.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 20 Cardiovascular mortality - geographical location

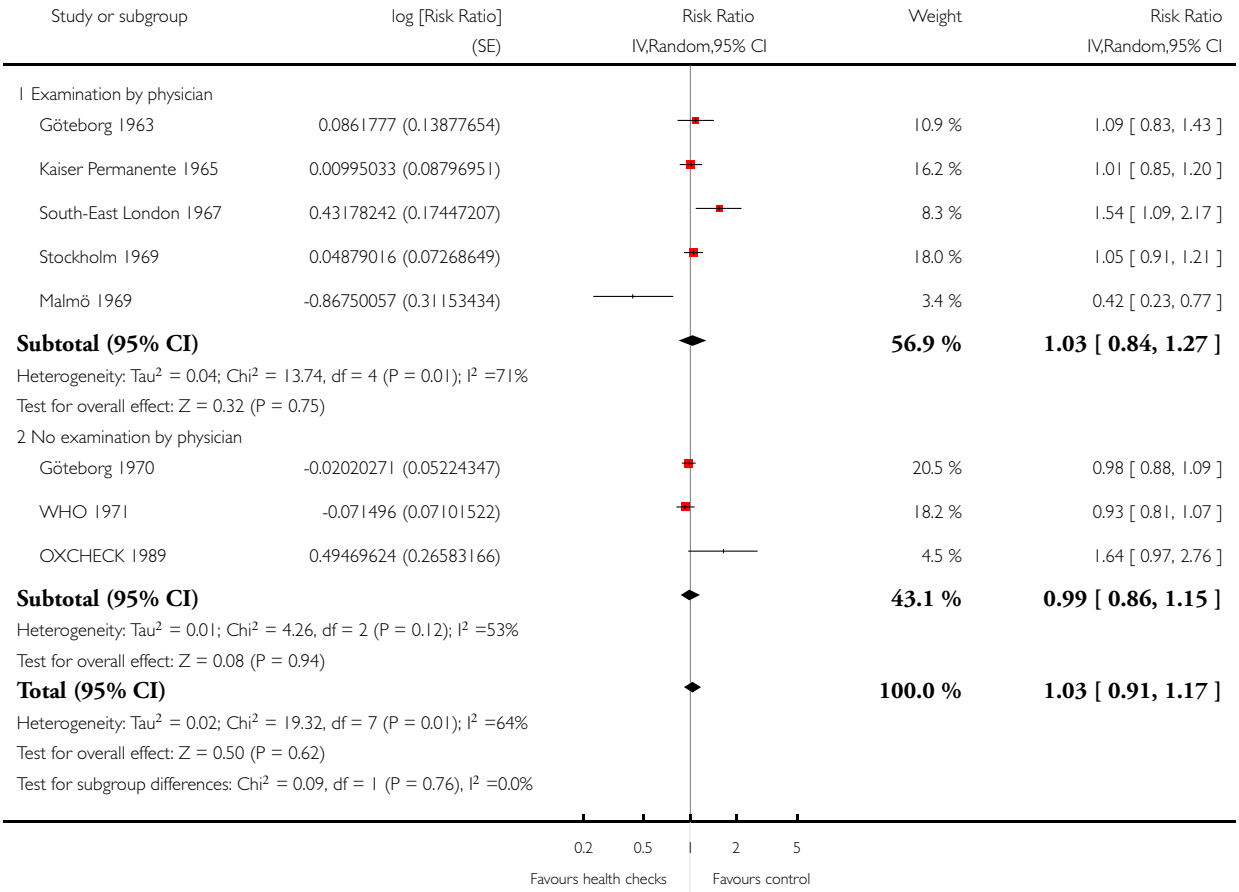


Analysis 1.21. Comparison 1 Health checks versus control, Outcome 21 Cardiovascular mortality - examination by physician.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 21 Cardiovascular mortality - examination by physician

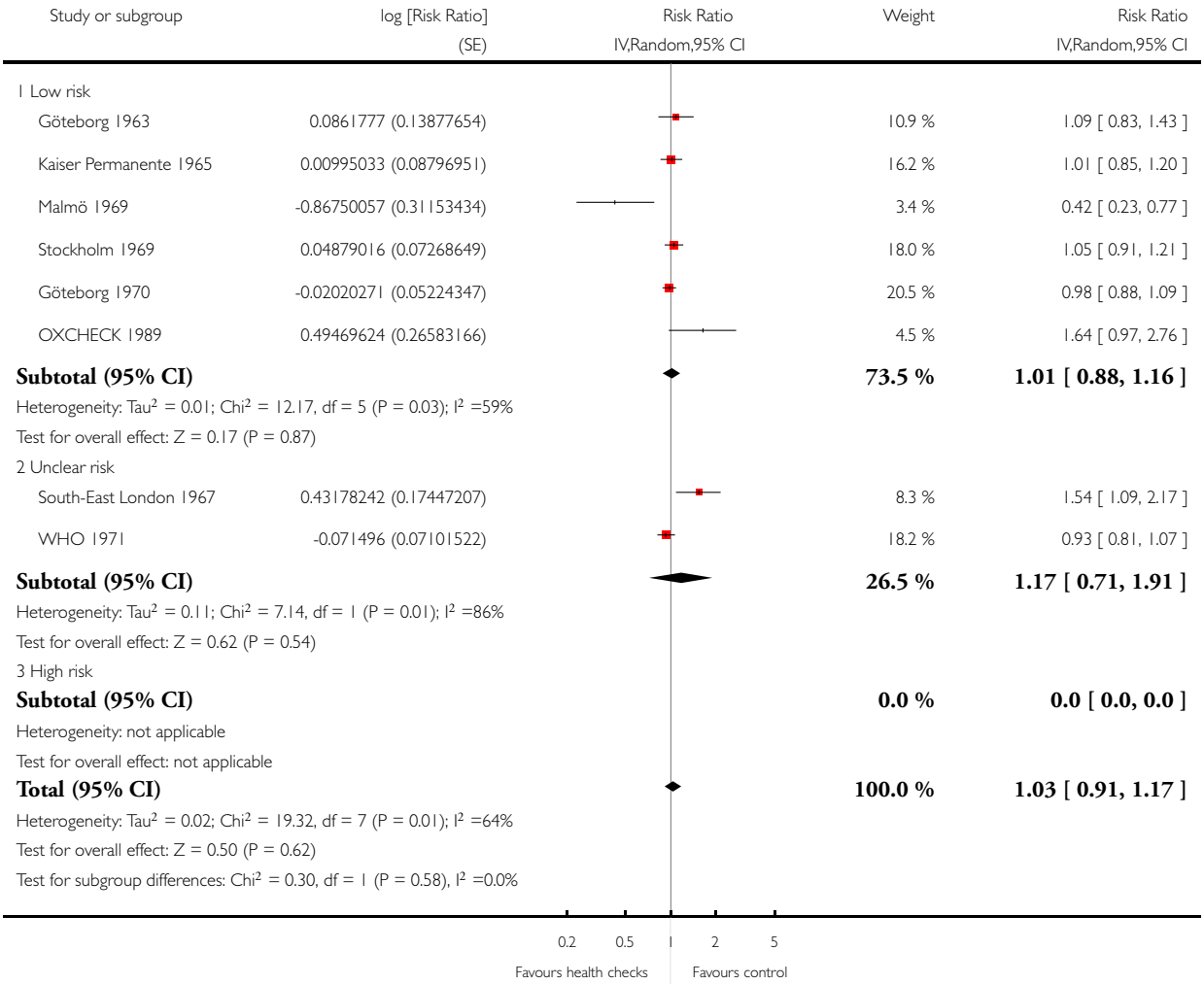


Analysis 1.22. Comparison 1 Health checks versus control, Outcome 22 Cardiovascular mortality - selection bias.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 22 Cardiovascular mortality - selection bias

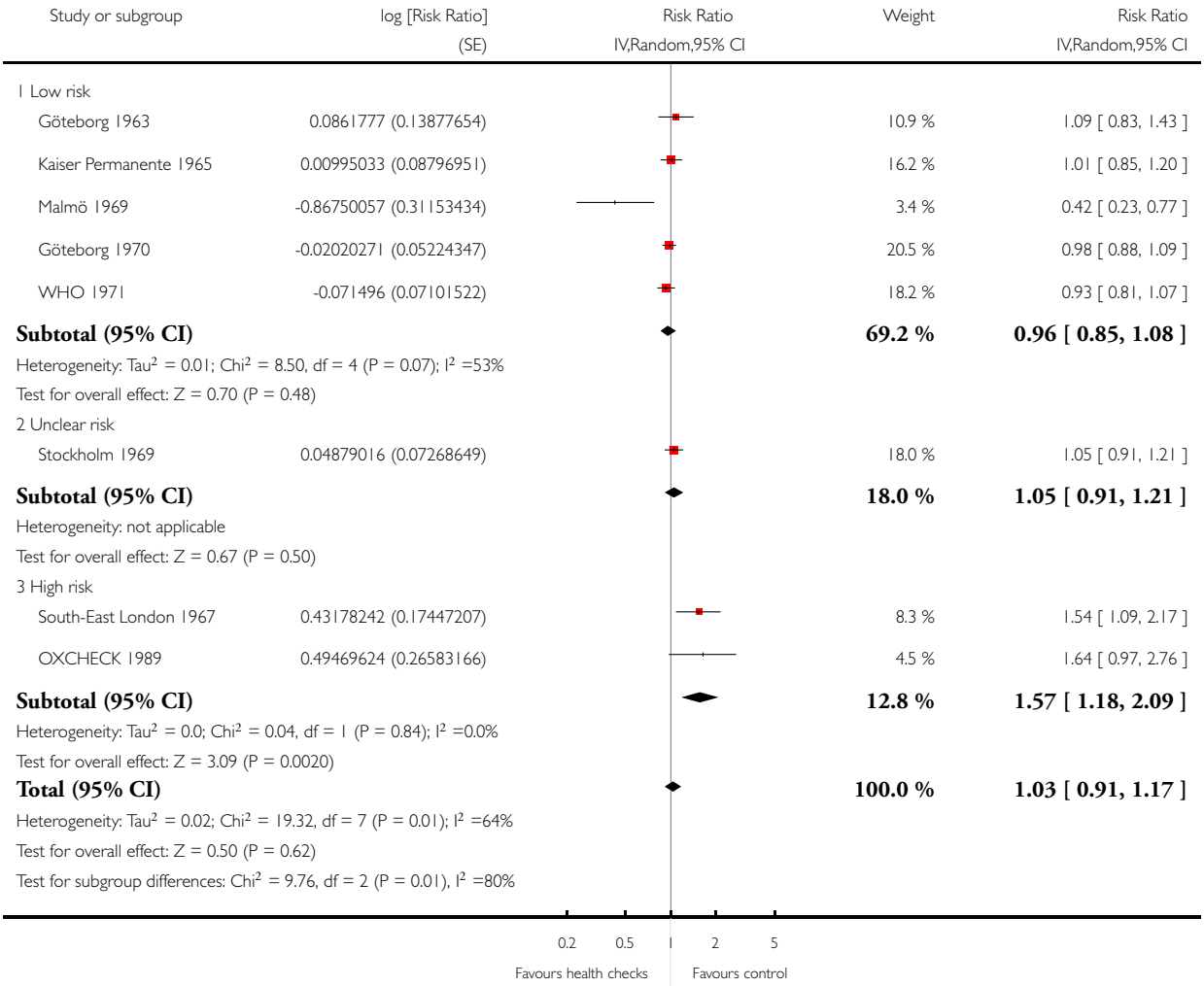


Analysis 1.23. Comparison 1 Health checks versus control, Outcome 23 Cardiovascular mortality - performance bias.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 23 Cardiovascular mortality - performance bias

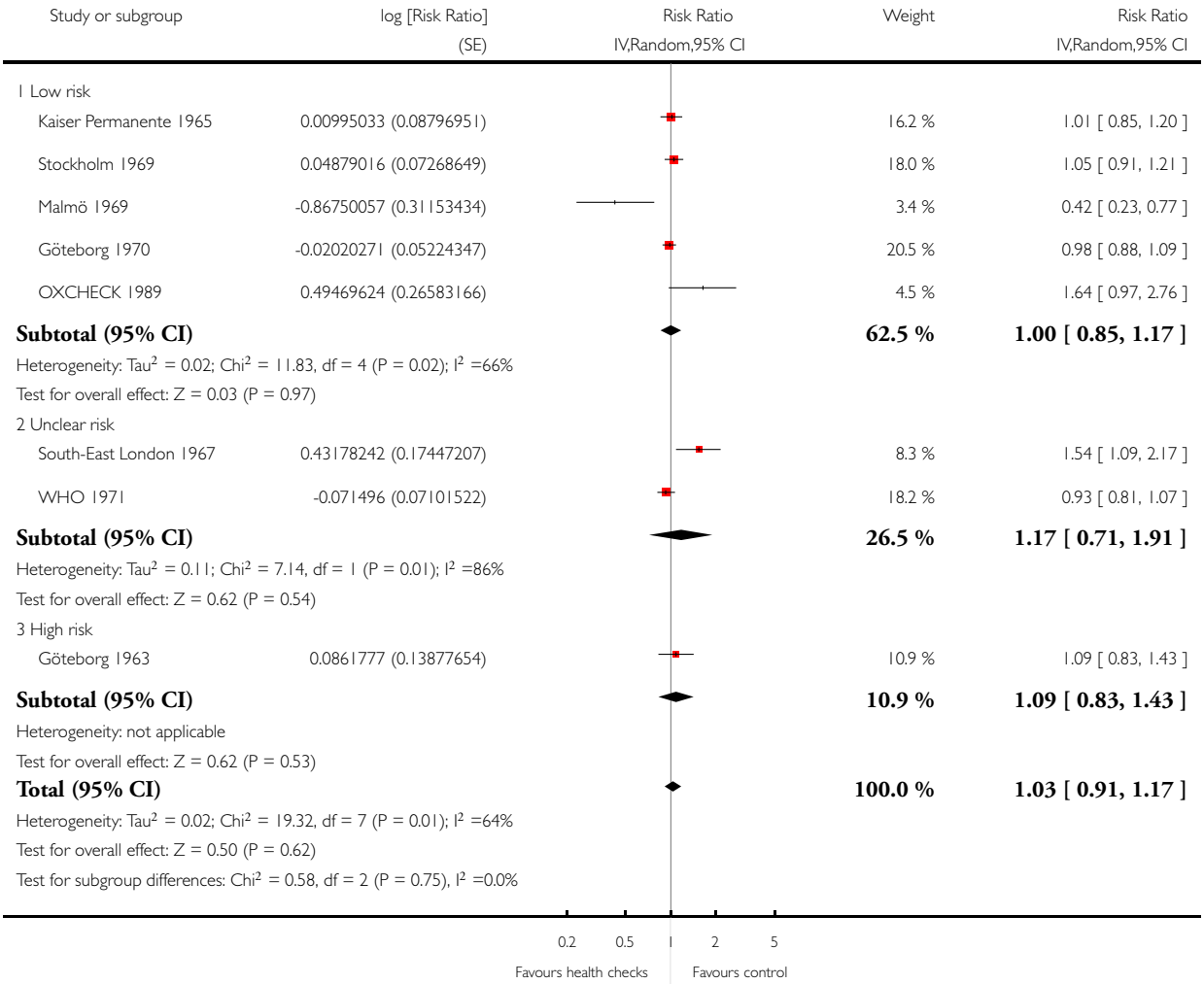


Analysis 1.24. Comparison 1 Health checks versus control, Outcome 24 Cardiovascular mortality - detection bias.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 24 Cardiovascular mortality - detection bias

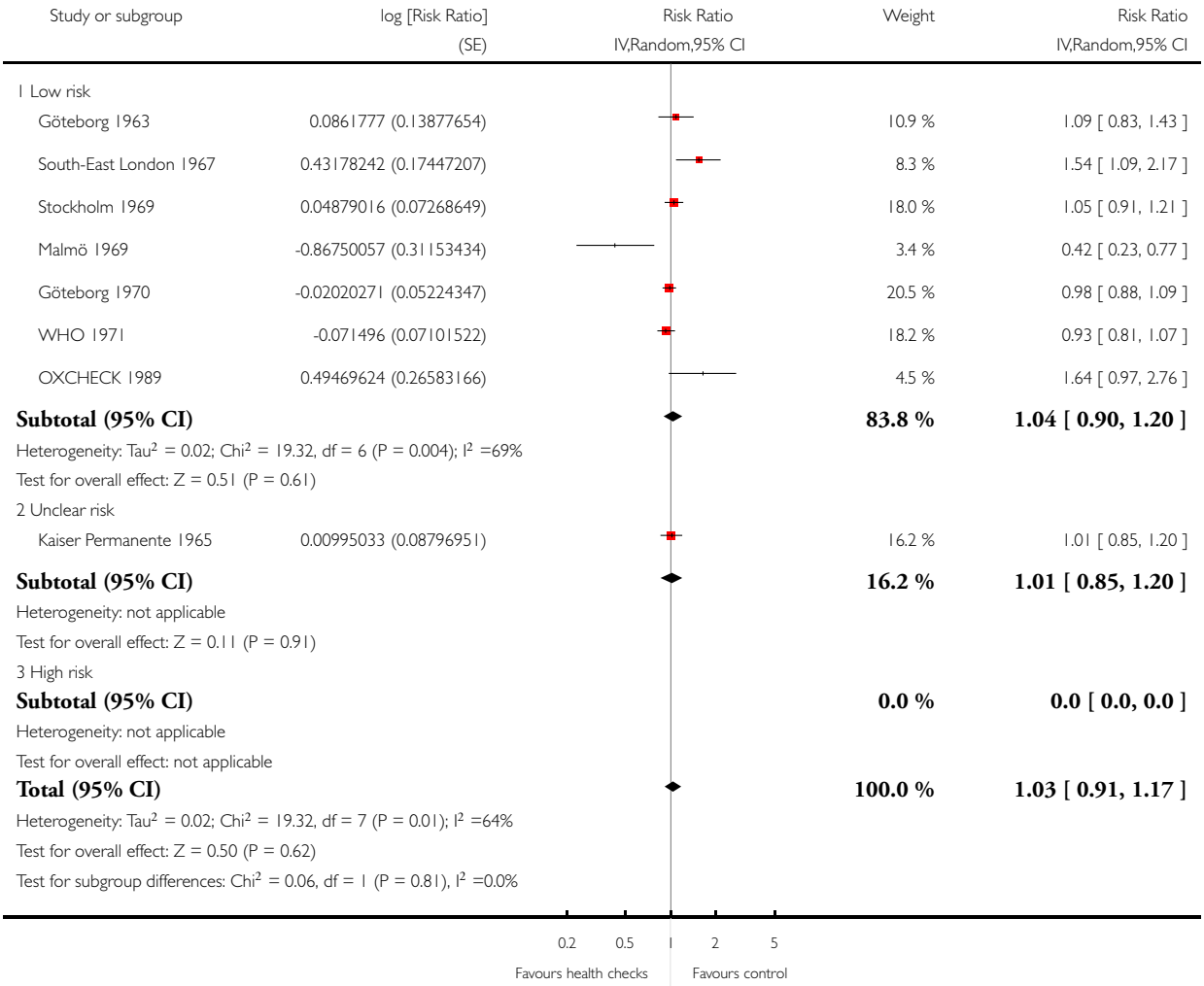


Analysis 1.25. Comparison 1 Health checks versus control, Outcome 25 Cardiovascular mortality - incomplete outcome data.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 25 Cardiovascular mortality - incomplete outcome data

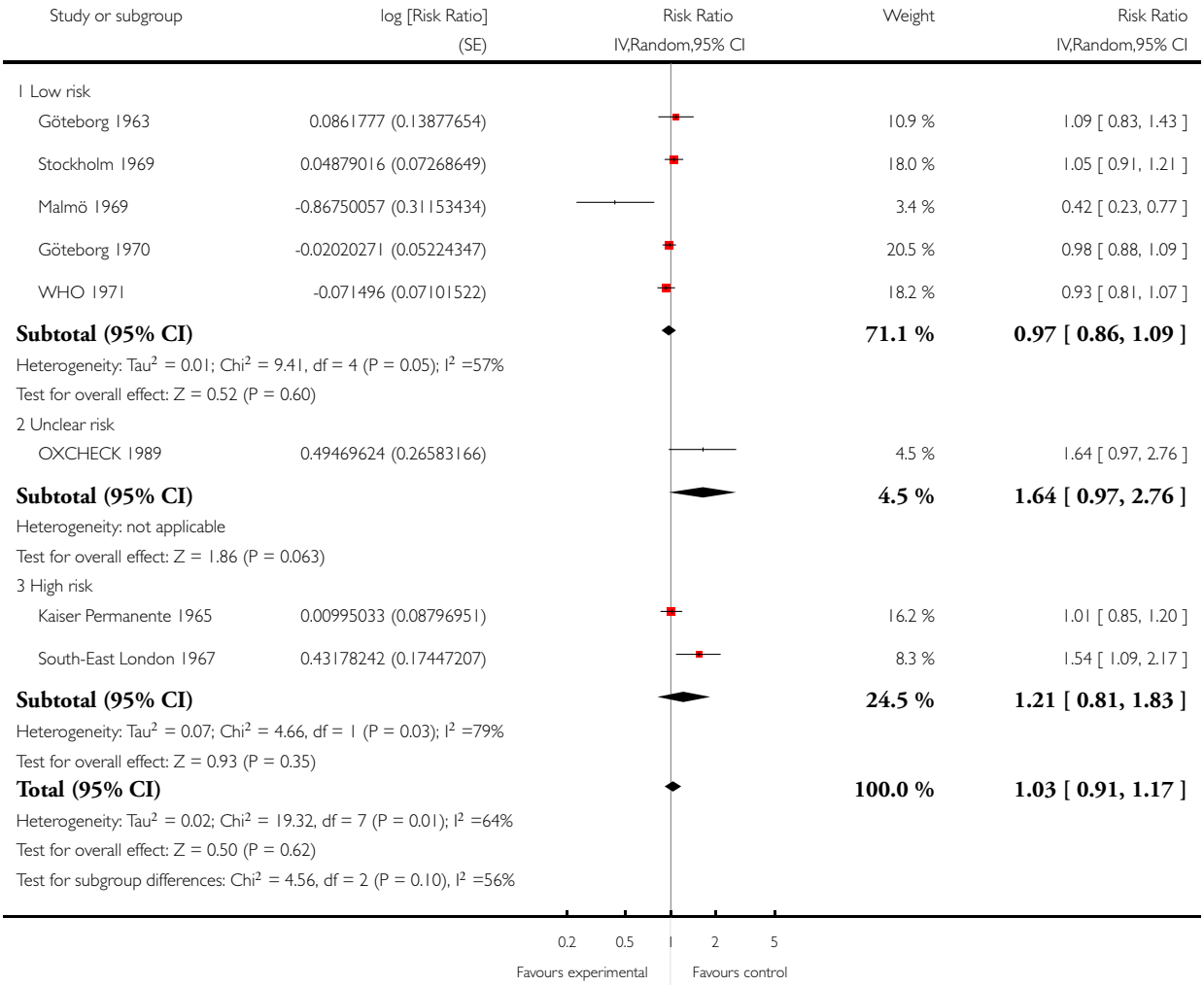


Analysis 1.26. Comparison 1 Health checks versus control, Outcome 26 Cardiovascular mortality - contamination.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 26 Cardiovascular mortality - contamination



Analysis 1.27. Comparison 1 Health checks versus control, Outcome 27 Cancer mortality.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 27 Cancer mortality

Study or subgroup	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
Göteborg 1963	-0.07257069 (0.20003035)		5.5 %	0.93 [0.63, 1.38]
Kaiser Permanente 1965	-0.02020271 (0.10343498)		15.2 %	0.98 [0.80, 1.20]
South-East London 1967	0.00995033 (0.20181826)		5.4 %	1.01 [0.68, 1.50]
Malmö 1969	0.63127178 (0.40996279)		1.4 %	1.88 [0.84, 4.20]
Stockholm 1969	0.04879016 (0.08953493)		18.2 %	1.05 [0.88, 1.25]
Göteborg 1970	-0.13926207 (0.06744554)		24.2 %	0.87 [0.76, 0.99]
WHO 1971	0.10408971 (0.06207813)		26.0 %	1.11 [0.98, 1.25]
OXCHECK 1989	0.17395331 (0.23578033)		4.1 %	1.19 [0.75, 1.89]
Total (95% CI)			**100.0 %**	**1.01 [0.92, 1.12]**

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 10.41$, df = 7 (P = 0.17); $I^2 = 33\%$

Test for overall effect: Z = 0.24 (P = 0.81)

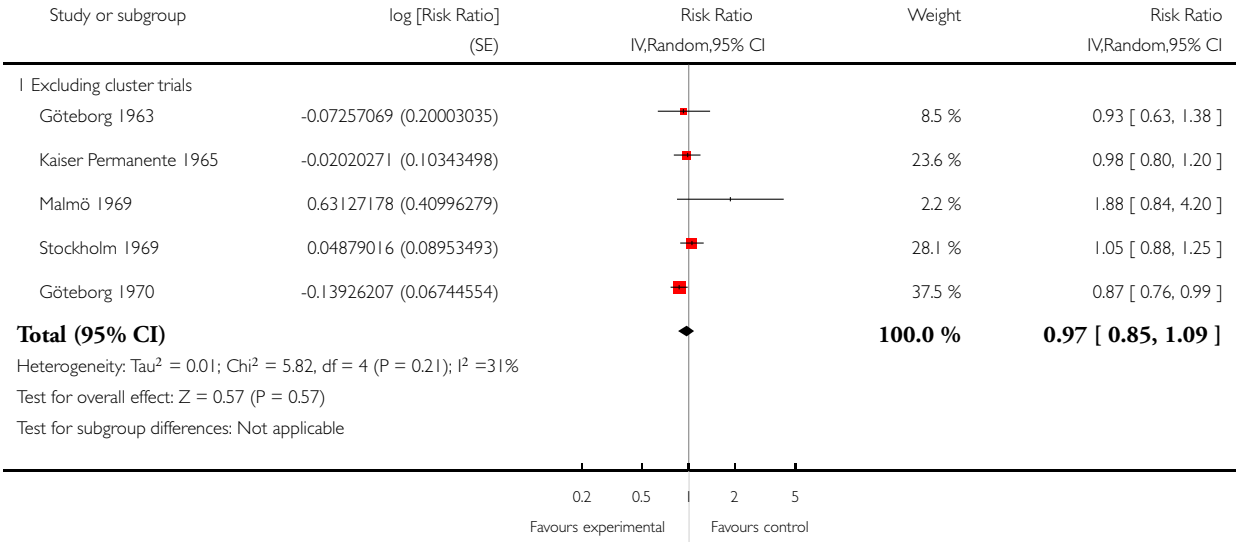
Test for subgroup differences: Not applicable

Analysis 1.28. Comparison 1 Health checks versus control, Outcome 28 Cancer mortality - sensitivity analyses.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 28 Cancer mortality - sensitivity analyses

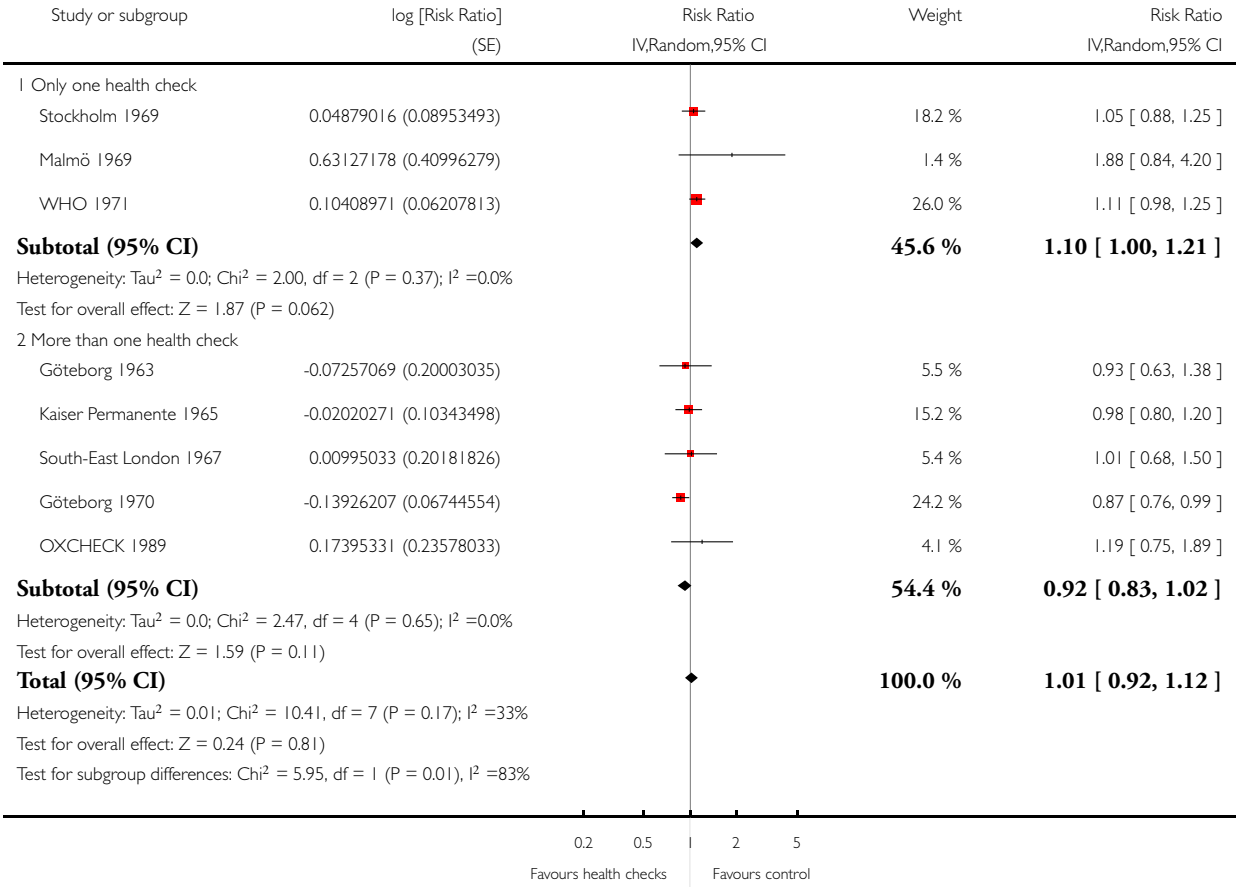


Analysis 1.29. Comparison 1 Health checks versus control, Outcome 29 Cancer mortality - no. of health checks.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 29 Cancer mortality - no. of health checks



Analysis 1.30. Comparison 1 Health checks versus control, Outcome 30 Cancer mortality lifestyle intervention.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 30 Cancer mortality lifestyle intervention

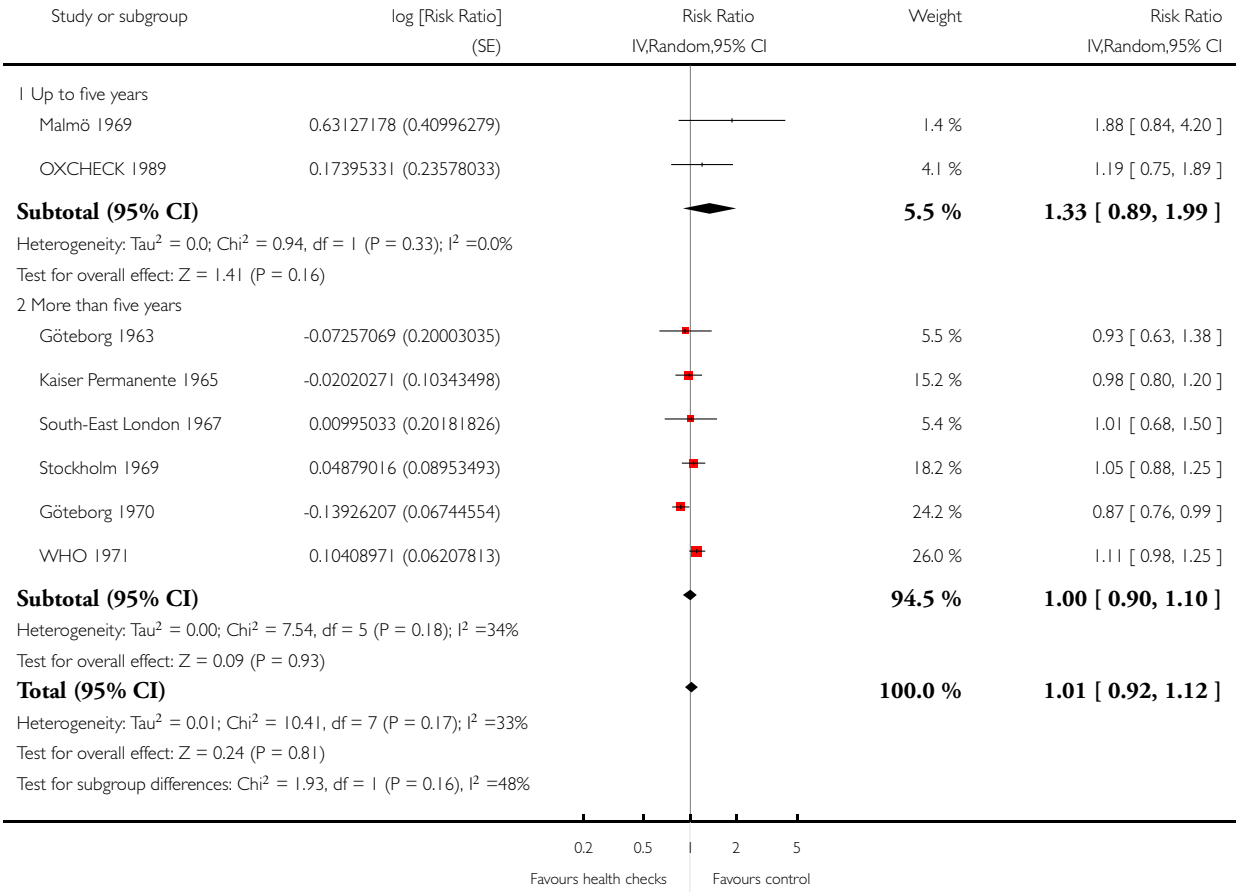
Study or subgroup	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
1 Major lifestyle intervention				
Göteborg 1970	-0.13926207 (0.06744554)		24.2 %	0.87 [0.76, 0.99]
WHO 1971	0.10408971 (0.06207813)		26.0 %	1.11 [0.98, 1.25]
OXCHECK 1989	0.17395331 (0.23578033)		4.1 %	1.19 [0.75, 1.89]
Subtotal (95% CI)			**54.3 %**	**1.01 [0.82, 1.24]**
Heterogeneity: $Tau^2 = 0.02$; $Chi^2 = 7.62$, df = 2 (P = 0.02); $I^2 = 74\%$				
Test for overall effect: Z = 0.11 (P = 0.92)				
2 No major lifestyle intervention				
Göteborg 1963	-0.07257069 (0.20003035)		5.5 %	0.93 [0.63, 1.38]
Kaiser Permanente 1965	-0.02020271 (0.10343498)		15.2 %	0.98 [0.80, 1.20]
South-East London 1967	0.00995033 (0.20181826)		5.4 %	1.01 [0.68, 1.50]
Malmö 1969	0.63127178 (0.40996279)		1.4 %	1.88 [0.84, 4.20]
Stockholm 1969	0.04879016 (0.08953493)		18.2 %	1.05 [0.88, 1.25]
Subtotal (95% CI)			**45.7 %**	**1.02 [0.91, 1.15]**
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 2.69$, df = 4 (P = 0.61); $I^2 = 0.0\%$				
Test for overall effect: Z = 0.39 (P = 0.70)				
Total (95% CI)			**100.0 %**	**1.01 [0.92, 1.12]**
Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 10.41$, df = 7 (P = 0.17); $I^2 = 33\%$				
Test for overall effect: Z = 0.24 (P = 0.81)				
Test for subgroup differences: $Chi^2 = 0.01$, df = 1 (P = 0.92), $I^2 = 0.0\%$				

Analysis 1.31. Comparison 1 Health checks versus control, Outcome 31 Cancer mortality - length of follow-up.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 31 Cancer mortality - length of follow-up

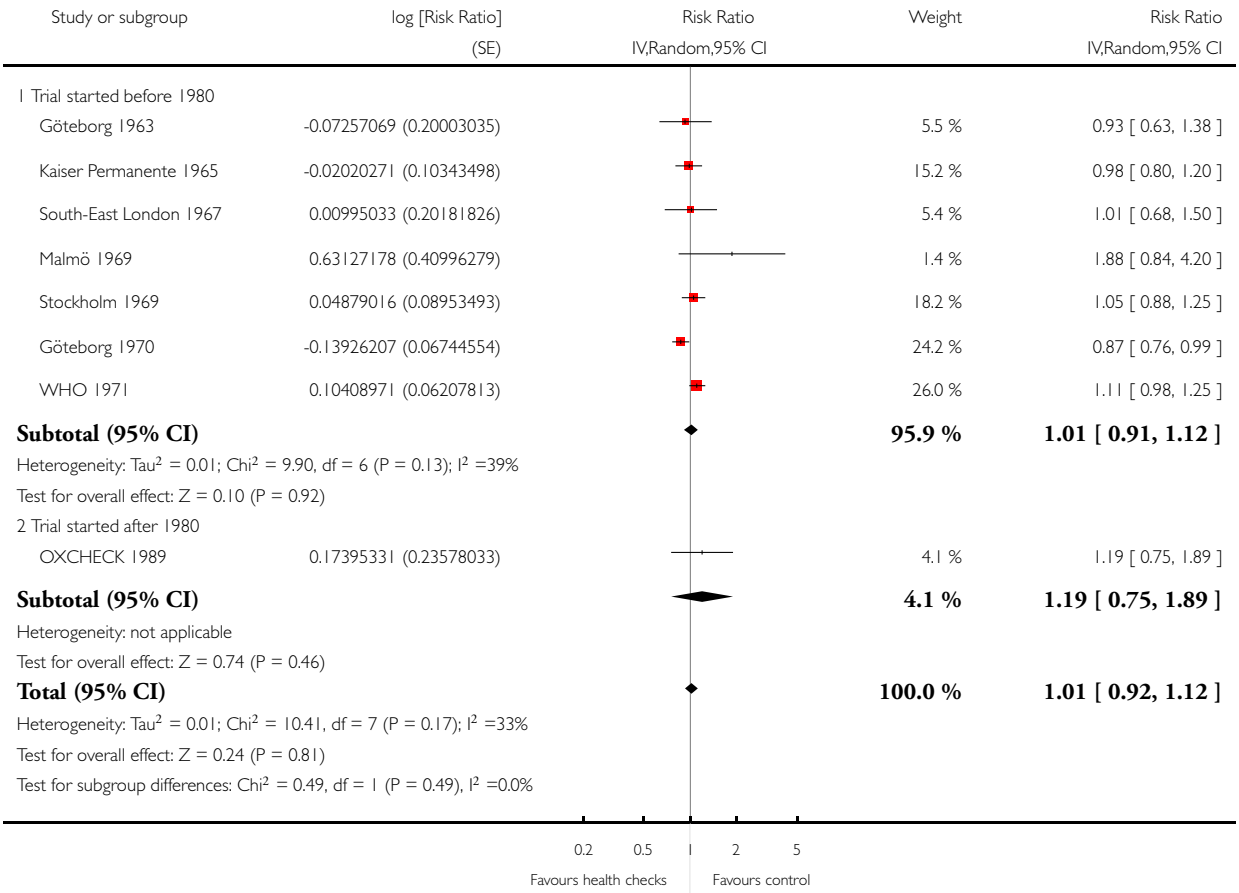


Analysis 1.32. Comparison 1 Health checks versus control, Outcome 32 Cancer mortality - age of trial.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 32 Cancer mortality - age of trial

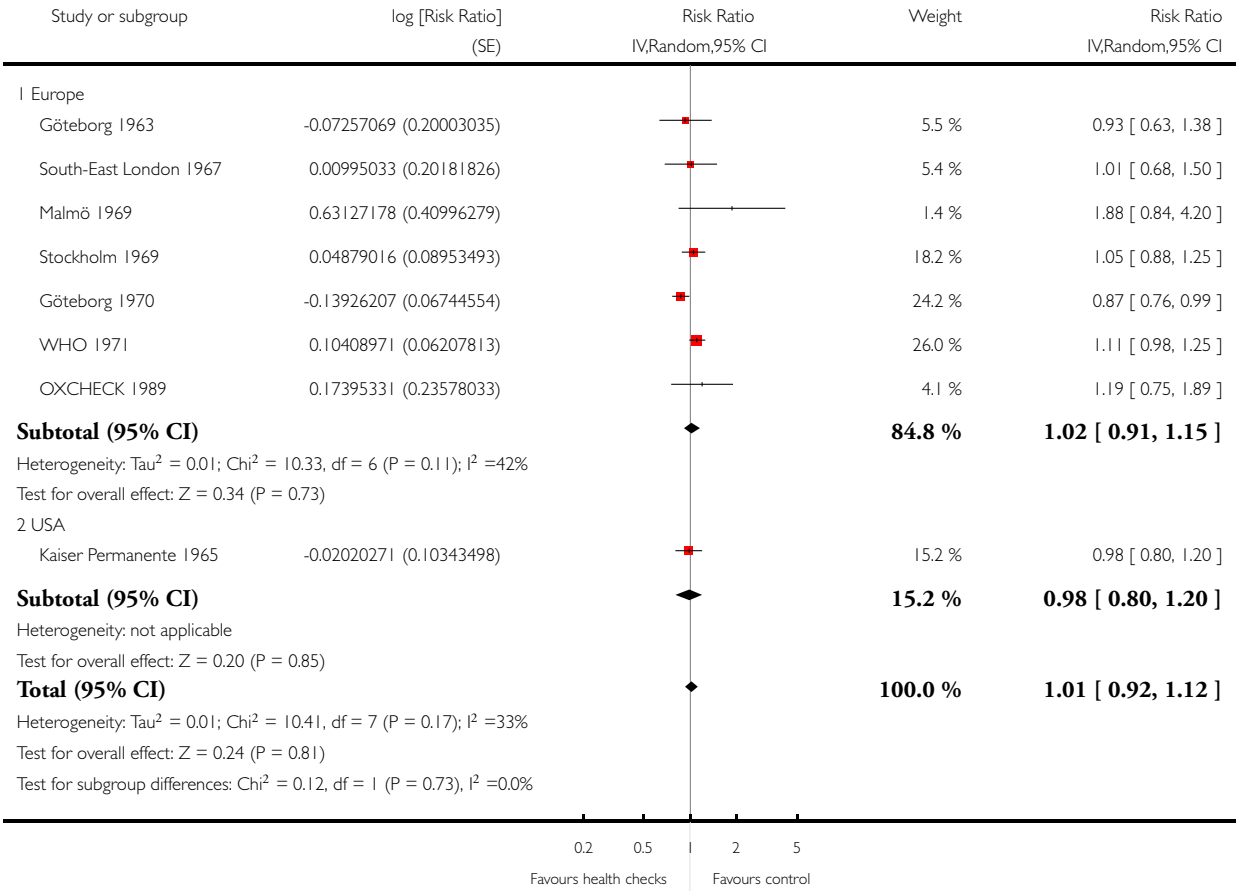


Analysis 1.33. Comparison 1 Health checks versus control, Outcome 33 Cancer mortality - geographical location.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 33 Cancer mortality - geographical location

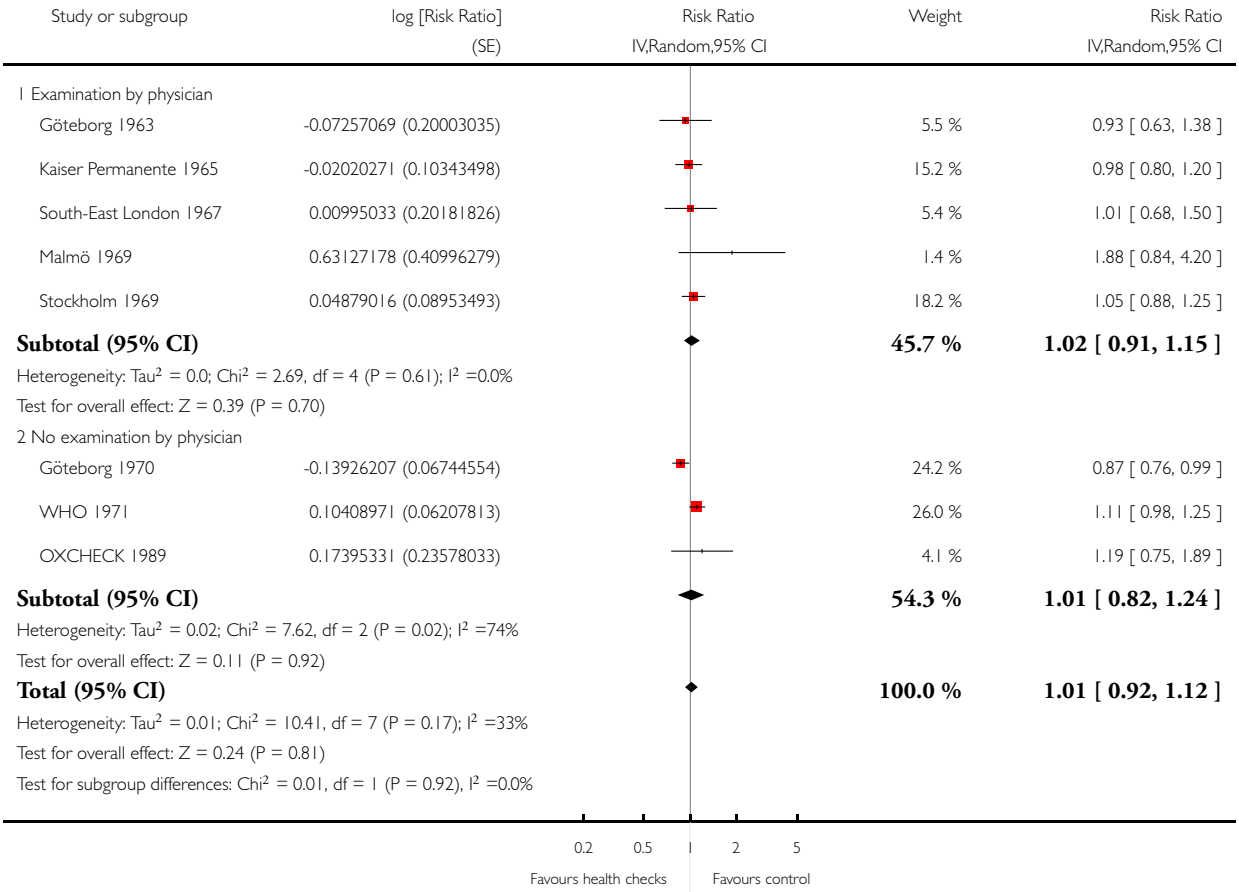


Analysis 1.34. Comparison 1 Health checks versus control, Outcome 34 Cancer mortality - examination by physician.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 34 Cancer mortality - examination by physician

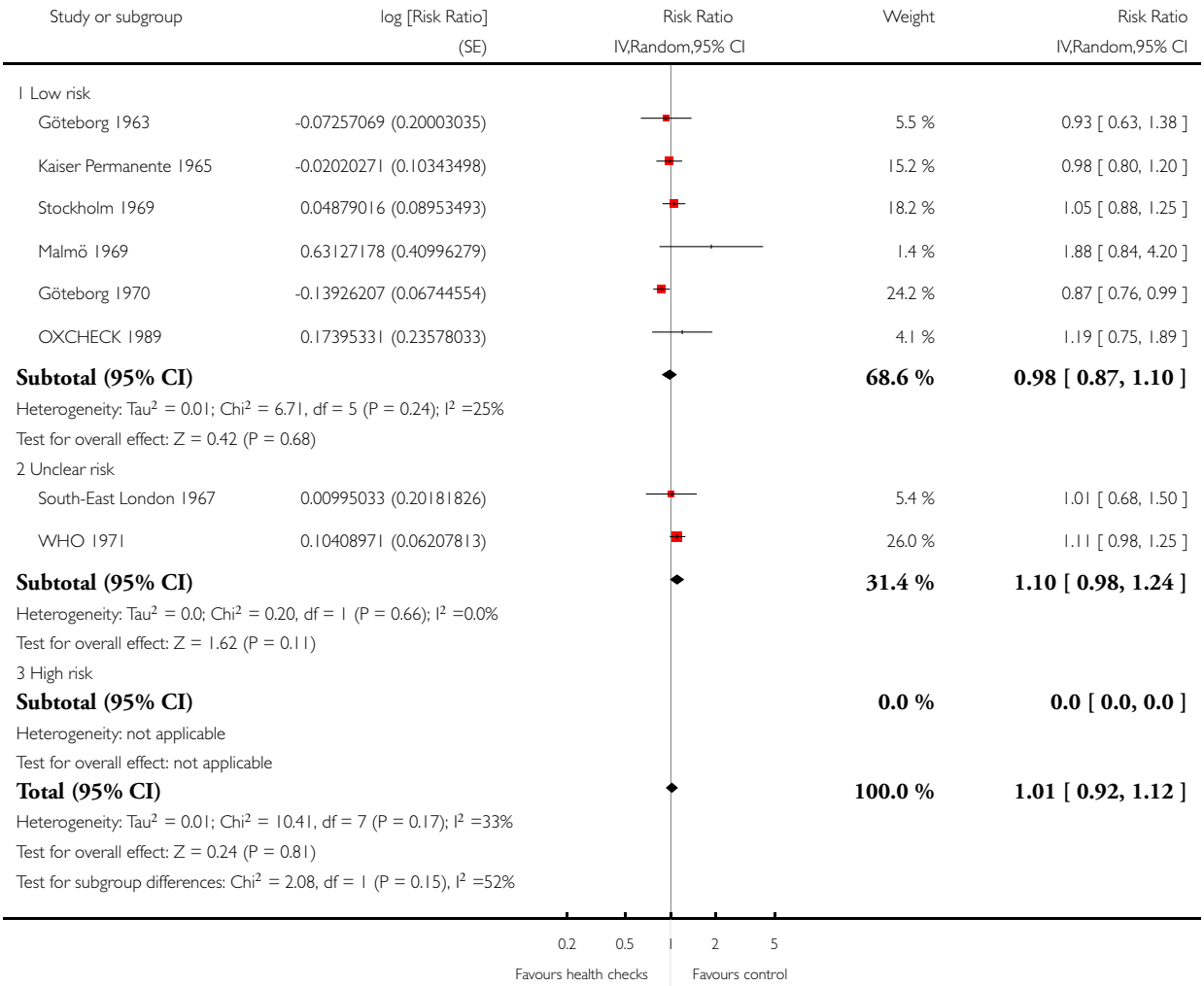


Analysis 1.35. Comparison 1 Health checks versus control, Outcome 35 Cancer mortality - selection bias.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 35 Cancer mortality - selection bias

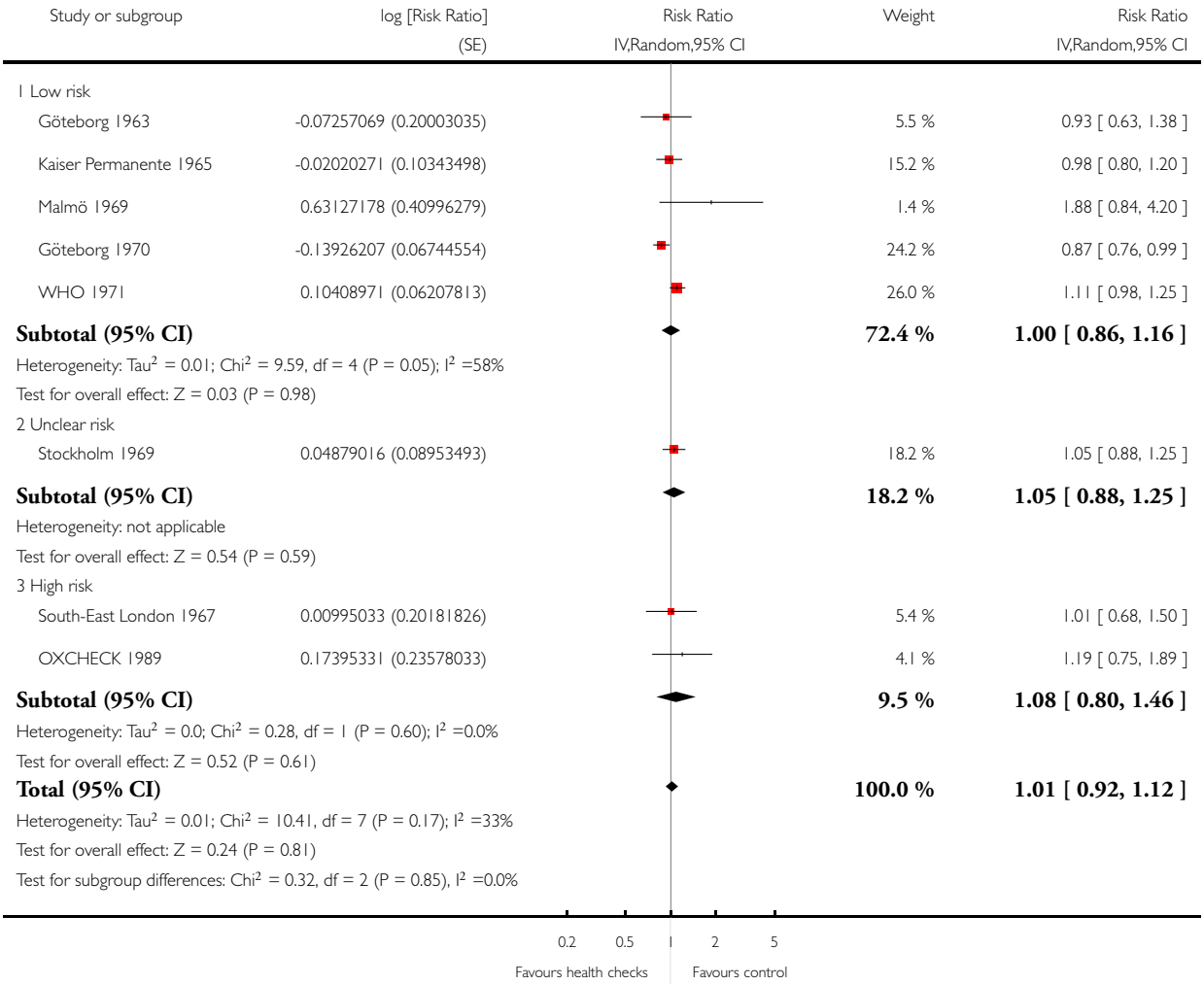


Analysis 1.36. Comparison 1 Health checks versus control, Outcome 36 Cancer mortality - performance bias.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 36 Cancer mortality - performance bias

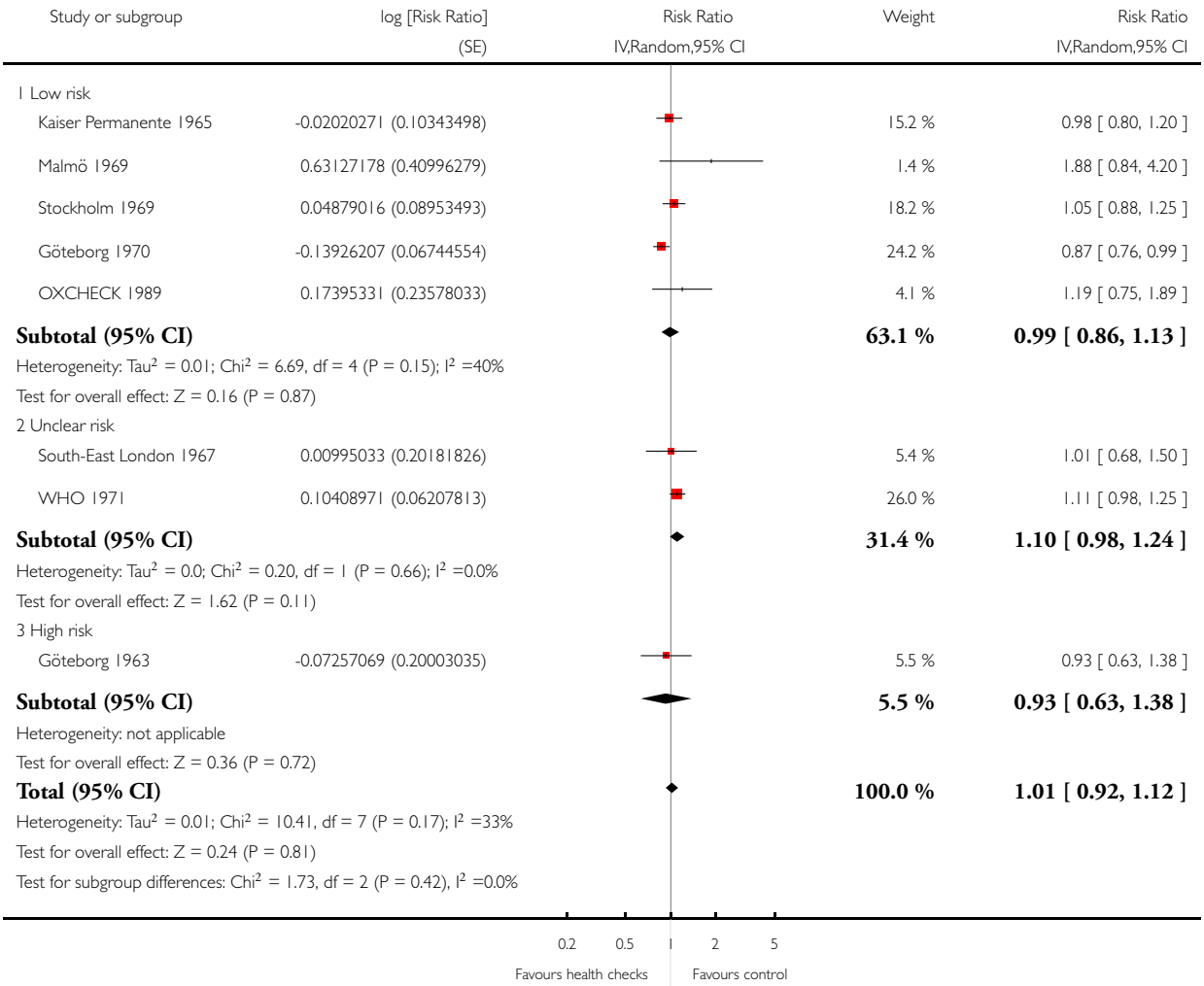


Analysis 1.37. Comparison 1 Health checks versus control, Outcome 37 Cancer mortality - detection bias.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 37 Cancer mortality - detection bias

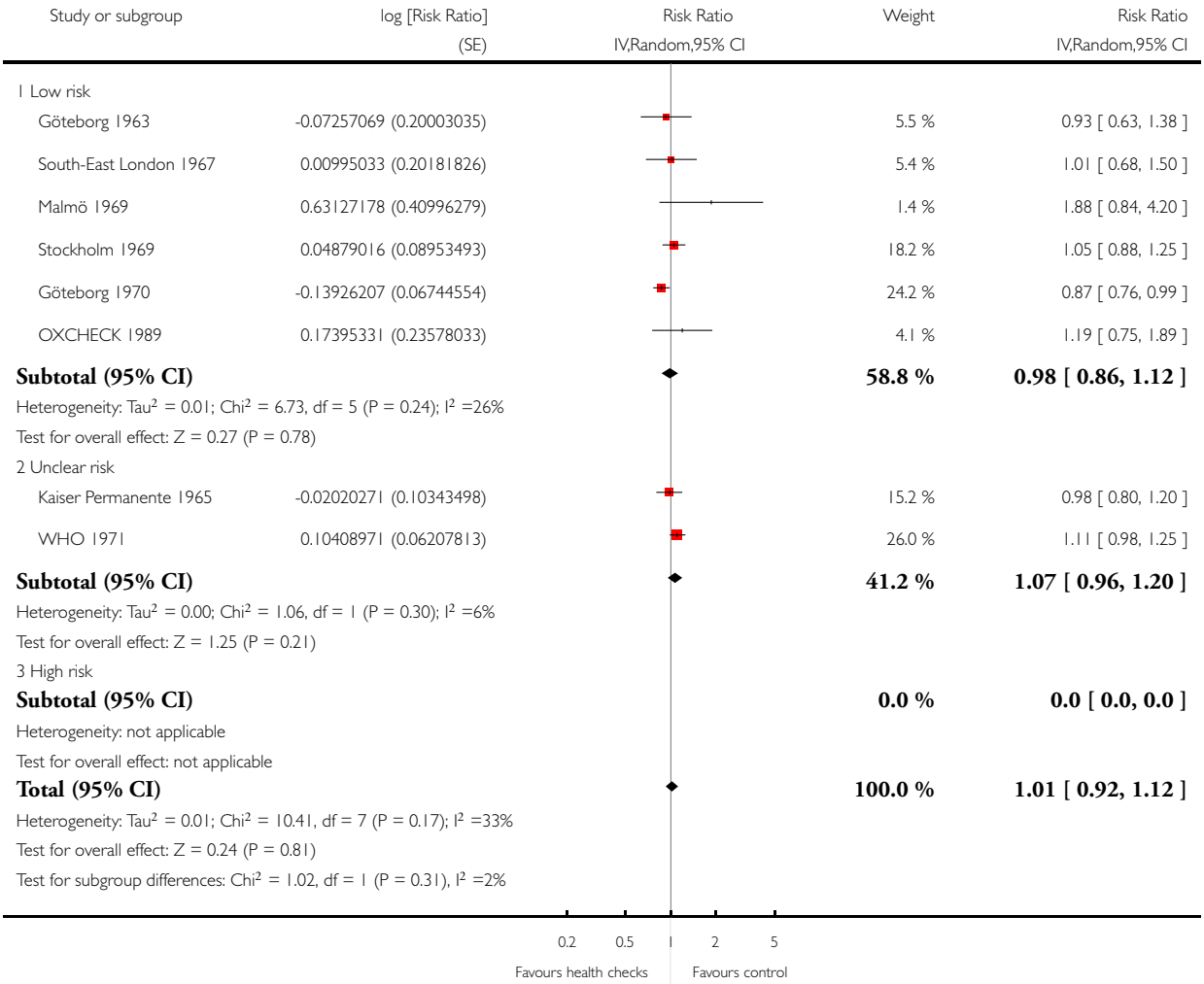


Analysis 1.38. Comparison 1 Health checks versus control, Outcome 38 Cancer mortality - incomplete outcome data.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 38 Cancer mortality - incomplete outcome data

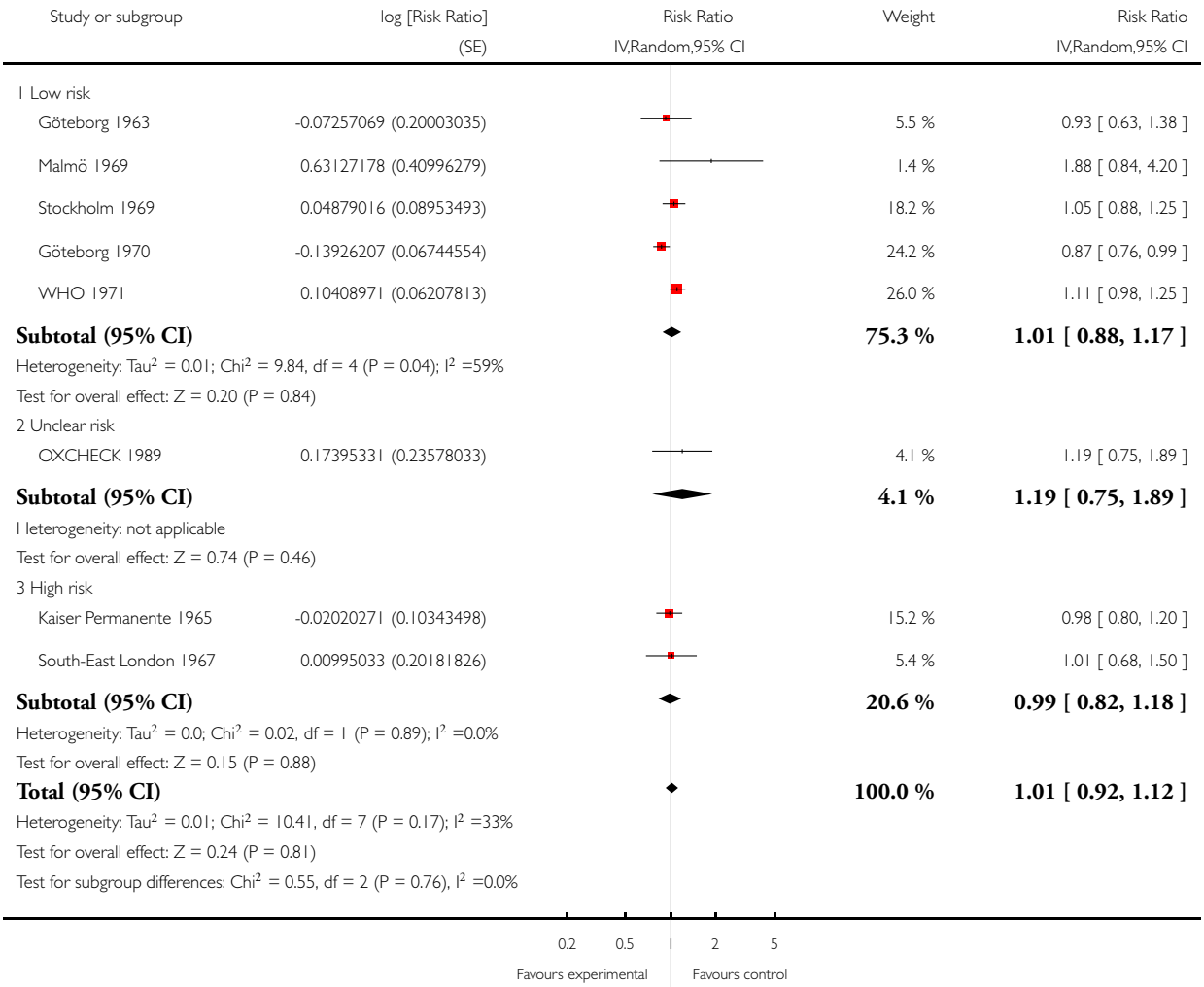


Analysis 1.39. Comparison 1 Health checks versus control, Outcome 39 Cancer mortality - contamination.

Review: General health checks in adults for reducing morbidity and mortality from disease

Comparison: 1 Health checks versus control

Outcome: 39 Cancer mortality - contamination



ADDITIONAL TABLES

Table 1. Overview of tests used in the trials

	Blood pressure	Cholesterol	Height and weight	Risk score	Electrocardiogram	Biochemistry panel	History	Spirometry	Urine analyses	Diabetes	Clinical examination	Vision and/or hearing	Cancer screening
Götebor 1963	x	x	x		x	x	current symptoms, personal and family history		x	fasting blood sugar	x	x	chest X-ray
Kaiser Permanente 1965	x	probably	x		x	x	current symptoms, personal and family history	x	x		x	x	chest X-ray, mammography, pelvic exam, sigmoidoscopy
South-East London 1967	x	probably	x		x	x	current symptoms, personal history	x			x	x	chest X-ray, faecal occult blood
Malmö 1969	x	x	x		x	haematocrit, triglycerides, cholesterol	interview and questionnaire, not specified	x	x		x		chest X-ray
Northun berland 1969	?	?	?	?	?	?	current symptoms	?	?	?	?	?	?

Table 1. Overview of tests used in the trials (*Continued*)

Stock-holm 1969	x	proba-bly			x	x	current symp-toms, per-sonal history				x	x	
Götebor 1970	x	x	x		x		family history						
WHO 1971	x	x	x				current symp-toms						
Salt Lake City 1972	x	x			x	x		x	x			x	chest X-ray, mam-mogra-phy, cervi-cal smear
Mankato 1982	x	x	x										
OX-CHECK 1989	x	x	x				per-sonal and family history						
Family Heart 1990	x	x	x	Dundee			per-sonal and family history			ran-dom capil-lary glu-cose			
Ebeltoft 1992	x	x	x	Anggaard	x	x		x	x	non-fasting blood glu-cose		x	

Table 1. Overview of tests used in the trials (*Continued*)

Inter99 1999	x	x	x	PRECARD	x			x		oral glucose tolerance test			

Not all screening tests used are shown; see Characteristics of included studies for full details. The Kaiser Permanente 1965, South-East London 1967, and Stockholm 1969 trials did not specify the contents of their biochemical screening. It seems unlikely that cholesterol was not included.

APPENDICES

Appendix 1. Medline Strategy A

Ovid MEDLINE(R) <1950 to August Week 1 2010>

1 (randomized controlled trial or controlled clinical trial).pt. or randomized.ab. or placebo.ab. or clinical trials as topic.sh. or randomly.ab. or trial.ti. (687779)
2 exp animals/ not humans.sh. (3516522)
3 1 not 2 [Cochrane RCT Filter 6.4.d Sens/Precision Maximizing] (636290)
4 (Periodic physical examination or Periodic physical examinations or Periodic health examination or Periodic health examinations or Periodic health evaluations or periodic health evaluation or Periodic screening or Periodic check up or Periodic checkup or Annual physical examination).ti,ab. (977)
5 (Annual physical examinations or Annual health examination or Annual health examinations or Annual screen or Annual screening or Annual health check up or Annual check up or Annual checkup).ti,ab. (877)
6 (Multiphasic health examination or Multiphasic screening or Multiphasic checkup or Multiphasic Health testing or Preventive health examinations or Preventive screening or primary care screening or Initial physical examination).ti,ab. (802)
7 (Initial screen or Initial screening or Initial check up or preventive services delivery or preventive service delivery or preventive service or preventive services or well care visit or well care visits or general health screening or preventive health screening).ti,ab. (5485)
8 exp Mass screening/ or screen$.ti,ab. (356136)
9 (primary care or community or communities or general practice$ or general practices).ti,ab. (291709)
10 (Kidney or renal or cardiovascular or vascular or cardiac or cardiovascular risk or coronary or heart or respiratory or pulmonary or lung).ti,ab. (2213994)
11 8 and 9 and 10 (2508)
12 or/4-7,11 (10406)
13 12 and 3 (1073)
14 adult/ or aged/ or "aged, 80 and over"/ or frail elderly/ or middle aged/ or young adult/ (4671530)
15 elderly.ti,ab. (132226)
16 middle age?.ti,ab. (24751)
17 old age.ti,ab. (14831)
18 adult/ or aged/ or "aged, 80 and over"/ or frail elderly/ or middle aged/ [ML] (4656253)
19 middle aged.ti,ab. (20562)
20 or/14-19 [Adult] (4693593)
21 13 and 20 (768)

22 exp infants/ (823182)
23 exp child/ (1341196)
24 13 not (or/22-23) [not children] (979)
25 21 or 24 (1007)

Appendix 2. Medline strategy B

Ovid Healthstar <1999 to October 2010>, Ovid Healthstar <1966 to 1998>, Ovid OLDMEDLINE(R) <1947 to 1965>, Ovid MEDLINE(R) <1996 to November Week 3 2010>, Ovid MEDLINE(R) Daily Update <17 November 2010>

1 Physical examination/ and ((annual or GP or periodic or yearly or routine).ti. or ((primary adj2 (care or healthcare)) or primary health$ or general practitioner? or general practice or family doctor? or family practice? or family physician?).ti,ab.) (2529)
2 (health check$ or healthcheck$ or annual physical? or annual medical or medical check$ or primary care check$ or wellness check$ well care or wellcare or well woman or well visit?).ti. (915)
3 ((annual or periodic or regular or routine or yearly) and (check$ or check-up? or health$ exam$ or health evaluation? or medical exam$ or physical? exam$ or wellness check$ or GP visit? or physician? visit? or doctor? visit? or office visit?)).ti. (856)
4 ((annual or yearly) adj2 (medical? or physical?)).ti. (211)
5 ((annual or yearly) and visit?).ti. (29)
6 (preventive? and (care check$ or checkup? or check-up? or visit? or exam$ or family doctor? or GP or family physician? or general practitioner?)).ti. (748)
7 or/1-6 [Annual Checkups --Combine with filters only] (4479)
8 Physical examination/ (30998)
9 (check-up? or checkup?).ti,ab. (9904)
10 (annual medical or yearly medical or annual physical).ab. (841)
11 ((annual or periodic or (primary adj2 (care or healthcare)) or primary health$ or general practitioner? or general practice or GP or family doctor? or family practice? or family physician? or regular or routine or yearly) adj3 (healthcheck? or health$ exam$ or health evaluation? or medical exam$ or office visit? or GP visit? or physical? exam$ or wellness check$)).ab. (3760)
12 ((annual or yearly) adj3 (physician? visit? or doctor? visit? or office visit?)).ab. (94)
13 "well care".ti,ab. (111)
14 (prevent$ and (screen$ or visit?)).ti. or (prevent$ adj3 (screen$ or visit?)).ab. (7461)
15 or/8-14 [Checkups general] (51400)
16 Mass screening/ (108483)
17 Multiphasic screening/ [ML] (1149)
18 ((community$ or program? or multiphasic or multi-phasic or (primary adj2 care) or "office visit?" or GP or general practice or care or healthcare or routine or annual) adj2 screening).ab. (23661)
19 screening.ti. (105590)
20 or/16-19 [Screening] (173357)
21 Primary prevention/ [ML] (19810)
22 exp Preventive Health Services/ (484698)
23 Health promotion/ or Healthy People Programs/ (68750)
24 (prevention or preventive or preventative).ti. (143300)
25 Risk assessment/ (227898)
26 or/21-25 [Prevention/Risk Assessment] (804059)
27 Risk factors/ (693977)
28 or/21-25,27 [Prevention/Risk Assessment/Risk Factors] (1387612)
29 exp Primary health care/ or Family practice/ or Physicians, family/ (188352)
30 ((family or general) adj (doctor? or practice? or practitioner? or physician$)).ti. (39510)
31 (primary adj2 (care or health care or healthcare or medical care or patient care)).ti. (41400)
32 Community Health services/ or Community mental Health Services/ or Community Pharmacy Services/ or Mobile Health units/ or Community Health Centers/ or Community health nursing/ (86820)
33 community$.ti. (95974)
34 or/29-33 [Primary/Community Care] (341710)

35 exp Aged/ [Elderly as group are unique whereas Adult is often not mentioned in indexing] (2371071)
36 (exp Cardiovascular Diseases/ or exp Digestive System Diseases/ or exp Endocrine System Diseases/ or exp Musculoskeletal Diseases/ or exp Lung Diseases, Obstructive/) and (pc or di).fs. (1059958)
37 disease?.hw. and (pc or di).fs. (785758)
38 (diabet$ or cardio$ or heart or disease or copd).ti. (993308)
39 or/36-38 [Diseases--selected] (2210888)
40 (randomized controlled trial or controlled clinical trial).pt. or randomized.ab. or placebo.ab. or clinical trials as topic.sh. or randomly.ab. or trial.ti. (948770)
41 exp animals/ not humans.sh. (1416010)
42 "comment on".cm. or systematic review.ti. or literature review.ti. or editorial.pt. or letter.pt. or meta-analysis.pt. or news.pt. or review.pt. [to exclude irrelevant publication types] (3277119)
43 40 not (or/41-42) [Modified Cochrane RCT Filter 6.4.d Sens/Precision Maximizing] (762774)
44 16 and (or/25,27) [Screening & Risk Factors/Assessment] (21559)
45 intervention?.ti. or (intervention? adj6 (clinician? or collaborat$ or community or complex or DESIGN$ or doctor? or educational or family doctor? or family physician? or family practitioner? or financial or GP or general practice? or hospital? or impact? or improv$ or individuali?e? or individuali?ing or interdisciplin$ or multicomponent or multi-component or multidisciplin$ or multi-disciplin$ or multifacet$ or multi-facet$ or multimodal$ or multi-modal$ or personali?e? or personali?ing or pharmacies or pharmacist? or pharmacy or physician? or practitioner? or prescrib$ or prescription? or primary care or professional$ or provider? or regulatory or regulatory or tailor$ or target$ or team$ or usual care)).ab. (166969)
46 (collaborativ$ or collaboration? or tailored or personali?ed).ti,ab. [added v2.0] (106136)
47 (exp hospitals/ or exp Hospitalization/ or exp Patients/ or exp Nurses/ or exp Nursing/) and (study.ti. or evaluation studies as topic/) [changed for v2.0 based on analysis of Mesh found on CBA & ITS not found by Filter 1.6] (39255)
48 demonstration project?.ti,ab. (2426)
49 (pre-post or "pre test$" or pretest$ or posttest$ or "post test$" or (pre adj5 post)).ti,ab. (59947)
50 (pre-workshop or post-workshop or (before adj3 workshop) or (after adj3 workshop)).ti,ab. (627)
51 trial.ti. or ((study adj3 aim?) or "our study").ab. (588608)
52 (before adj10 (after or during)).ti,ab. (313147)
53 ("quasi-experiment$" or quasiexperiment$ or "quasi random$" or quasirandom$ or "quasi control$" or quasicontrol$ or ((quasi$ or experimental) adj3 (method$ or study or trial or design$))).ti,ab,hw. [ML] (73957)
54 ("time series" adj2 interrupt$).ti,ab,hw. [ML] (936)
55 (time points adj3 (over or multiple or three or four or five or six or seven or eight or nine or ten or eleven or twelve or month$ or hour? or day? or "more than")).ab. (7327)
56 pilot.ti. (39707)
57 Pilot projects/ [ML] (97983)
58 (clinical trial or multicenter study).pt. [ML removed RCT--redundant v2.0] (854060)
59 (multicentre or multicenter or multi-centre or multi-center).ti. (33079)
60 random$.ti,ab. or controlled.ti. (770410)
61 (control adj3 (area or cohort? or compar? or condition or group? or intervention? or participant? or study)).ab. not (controlled clinical trial or randomized controlled trial).pt. [ML remove DESIGN changed truncation on Compare] (267909)
62 "comment on".cm. or systematic review.ti. or literature review.ti. or editorial.pt. or letter.pt. or meta-analysis.pt. or news.pt. or review.pt. [to exclude irrelevant publication types] (3277119)
63 exp animals/ not humans.sh. [ML] (1416010)
64 *experimental design/ or *pilot study/ or quasi experimental study/ [EM] (24724)
65 ("quasi-experiment$" or quasiexperiment$ or "quasi random$" or quasirandom$ or "quasi control$" or quasicontrol$ or ((quasi$ or experimental) adj3 (method$ or study or trial or design$))).ti,ab. [EM] (73957)
66 ("time series" adj2 interrupt$).ti,ab. [EM] (936)
67 (animal/ or animal.hw.) not ((animal/ or animal?.kw,hw.) and (human/ or human?.hw,kw.)) [EM] (1392461)
68 (book or letter).pt. [EM] (761941)
69 (or/45-52,55-56,59-61,64-66) not (or/67-68) [EPOC Methods Filter EM 2.0] (1863049)
70 (or/45-61) not (or/62-63) [EPOC Methods Filter ML 2.0] (2099569)
71 exp Drug Therapy/ and irrational.ti,ab. [ML] (241)
72 (rational adj4 (drug therapy or "drug use" or prescribing)).ti,ab. (890)
73 (rational or irrational).ti. and drug therapy.hw. (360)

74 ((promote or prefer) adj5 generic).ti,ab. (55)
75 prescribing habits.ti,ab. (734)
76 ((physician? or doctor? or nurse?) adj4 compliance).ti,ab. (1841)
77 (promoting.ti. and (health$ or care or education or nurse? or nursing or patient? or hospital$).ti,hw.) or (promoting and (doctor? or physician? or pharmacist?)).ti. (7096)
78 (fund-hold$ or fundhold$ or capitation or capitated or copay$ or co-pay$).ti,ab. (6998)
79 intervention?.ti. (74491)
80 (intervention? adj6 (clinician? or collaborat$ or community or complex or doctor? or educational or family doctor? or family physician? or family practitioner? or financial or GP or general practice? or hospital? or impact? or improv$ or individuali?e? or individuali?ing or interdisciplin$ or multicomponent or multi-component or multidisciplin$ or multi-disciplin$ or multifacet$ or multi-facet$ or multimodal$ or multi-modal$ or personali?e? or personali?ing or pharmacies or pharmacist? or pharmacy or physician? or practitioner? or prescrib$ or prescription? or primary care or professional$ or provider? or regulatory or regulatory or tailor$ or target$ or team$ or usual care)).ab. [ADDED Usual Care, GP & Increased ADJ] (97337)
81 (impact or effect$ or change? or changing).ti. [Added for 1.8] (1224865)
82 (communit$ or team$ or interdisciplinar$ or multidisciplinar$).ti,ab. [Added for 1.8] (483983)
83 (implementation or implementing).ti. and (care or healthcare).ti,hw. (8591)
84 ((effect? or effectiveness or chang$ or improv$ or impact) adj3 practice).ti,ab. (24465)
85 (Improv$ adj3 (diagnosis or treatment? or prescribing)).ti,ab. (58294)
86 (((evidence or evidence-based) adj4 intervention) or evidence-driven).ti,ab. (2364)
87 "practice-based".ti,ab. (4241)
88 (improv$ adj3 (decision$ or implement$ or health care or healthcare or initiative? or management or multifacet$ or multi-facet$ or multi-component or practi?e? or practitioner? or prescrib$ or prescription? or professional? or program? or programme? or provider?)).ti. [DELETED Effectiv$] (9636)
89 (improv$ adj2 (patient-care or family practice or ((family or general) adj2 (practi?e or practitioner? or doctor?)) or primary care)).ab. [DELETED Effectiv$] (5665)
90 recommended practice?.ti,ab. (853)
91 ((information or evidence) adj2 uptake).ti,ab. (321)
92 ((knowledge adj2 (application or broke$ or creation or diffus$ or disseminat$ or exchang$ or implement$ or management or mobili$ or translat$ or transfer$ or uptake or utili$)) or (evidence$ adj2 (exchang$ or translat$ or transfer$))).ti,ab. (8210)
93 (KT adj2 (application or broke$ or diffus$ or disseminat$ or decision$ or exchang$ or implement$ or intervent$ or mobili$ or plan$ or policy or policies or strateg$ or translat$ or transfer$ or uptake or utili$)).ti,ab. (108)
94 ((computer-tailored or individuali?ing or individuali?ed or personali?e? or personali?ing or tailor$) adj2 (feedback or intervention? or information or plan?)).ti,ab. (5569)
95 ((conventional or evidence-based or pattern or regular or routine or standard or traditional or usual) adj2 (care or healthcare or patient care or practice)).ti,ab. (67068)
96 (collaborative? or interdisciplin$ or inter-disciplin$ or multidisciplin$ or multi-disciplin$ or team? or team-based or skill-mix).ti. (42073)
97 ((collaborative or multidisciplinary or interdisciplinary) adj2 (care or healthcare or patient care or team?)).ab. (16629)
98 (skill? adj2 (mix or mixes)).ti,ab. (790)
99 (doctor-driven or doctor-led or GP-LED or nurse-led or nurse-driven or pharmacist-led or pharmacist-driven or physician-led or physician-driven).ti,ab. (2914)
100 physician directed.ti,ab. (355)
101 (BOOKLET? or leaflet$ or pamphlet$ or "written information").ti. or ((BOOKLET? or leaflet$ or pamphlet$ or "written information") adj5 (intervention? or care or healthcare or physician? or practitioner? or provider?)).ab. (3595)
102 (academic detailing or e-detailing or (opinion? adj2 leader?)).ti,ab. (1400)
103 ("audit and feedback" or ((physician? or doctor? or practitioner? or nurse? or provider?) adj feedback)).ti,ab. (731)
104 reminder?.ti. (1352)
105 (reminder? adj2 (clinician? or physician? or practitioner? or nurse? or doctor? or provider?)).ab. (406)
106 ((clinician? or physician?) adj2 (prompt or prompts or prompting)).ti,ab. (534)
107 ((doctor? or nurse? or pharmacist? or physician? or practitioner?) adj2 behavio?r?).ti,ab. (3572)
108 (nurse? adj4 substitut$).ti,ab. (147)
109 (practice pattern? or ((change? or changing) adj2 practice)).ti,ab. (12136)
110 Physician's Practice Patterns/ [DELETED change and other kw] (57122)

111 (nurse-practitioner? or physician? assistant?).ti. (5859)
112 ((doctor? or nurse? or pharmacist? or physician?) adj2 role?).ab. (7056)
113 ((nurse? or physician? or pharmacist? or provider?) adj2 initiative?).ti,ab. (420)
114 (virtual reality or VR Training or VR simulat$ or (simulat$ adj2 skill?)).ti,ab. (5040)
115 (blog$ or wiki$ or PDA or "palm pilot?" or blackberr$ or Twitter or tweet or tweeting or facebook or social networking or social marketing or youtube).ti,ab. or blogging/ (7937)
116 (health 20 or healthcare 20 or health care 20 or web 20).ti,ab. (365)
117 Guidelines as topic/ (26467)
118 (((individuali$ or integrated) adj2 (care or healthcare or medical care)) or patient-centred or patient-centered or patient-control$).ti,ab. (20092)
119 quality improvement.ti,ab. (17211)
120 *Patient satisfaction/ (28388)
121 (algorithm? and (care or healthcare or patient?)).ti,hw. (18173)
122 Education, Pharmacy, Continuing/ or Education, Medical, Continuing/ or Education, Nursing, Continuing/ or Education, Professional, Continuing/ (50829)
123 (continuing adj2 education adj3 (physician? or nurse? or nursing or practitioner? or doctor? or family physician? or general practitioner? or family doctor? or primary care or primary healthcare)).ab. or (continuing adj3 education).ti. (8728)
124 (((continuing or "on the job" or "off the job" or postgrad$ or post-grad$ or resident? or intern? or internship? or workplace) adj2 (education$ or training)) or (skill? adj (education or training))).ti,ab. (31787)
125 (reminder? adj2 (clinician? or physician? or practitioner? or nurse? or doctor? or provider?)).ab. (406)
126 (Referral? adj3 (early or increase? or primary care or specialist? or general practitioner? or optimi?e? or optimal or reduce? or reducing)).ab. (6487)
127 referral?.ti. (10562)
128 (specialist? and (primary care or primary healthcare or GP or general practitioner? or family doctor)).ti. (678)
129 (specialist? adj3 (primary care or primary healthcare or GP or general practitioner? or family doctor)).ab. (2886)
130 Reminder systems/ (2895)
131 Guideline adherence/ or (guideline? adj3 (adherence or compliance or concordance or implement$ or UPTAKE)).ti,ab. [Increased adj] (33357)
132 "Referral and Consultation"/ (64067)
133 "comment on".cm. or systematic review.ti. or literature review.ti. or editorial.pt. or letter.pt. or meta-analysis.pt. or news.pt. or review.pt. (3277119)
134 exp animals/ not humans.sh. (1416010)
135 (or/71-132) not (or/133-134) [Interventions ML 1.9] (1653094)
136 15 and 34 [Checkups (general) & Primary Care/Community Care] (4467)
137 15 and (or/20,26,39) [Checkup(general) & Screening/Prevention/Disease pc] (25132)
138 7 and 70 [Annual Checkups and EPOC Filter] (749)
139 7 and 43 [Annual Checkups and RCT Filter] (253)
140 15 and 34 and 43 [Checkups (general) & Primary/Community Care RCT Filter] (354)
141 15 and 34 and 70 [Checkups (general) & Primary/Community Care EPOC Filter] (1142)
142 15 and 34 and 135 [Checkups (general) & Primary/Community Care & Intervention Filter] (1929)
143 142 not (140 or 141) (1090) [Checkups (general) & Primary/Community Care & Intervention Filter - dupes removed]
144 (or/16-17) and (or/25,27) [Screening & Risk Factors/Assessment] (21618)
145 138 not 139 [Unique Annual Checkups EPOC Filter] (502)
146 remove duplicates from 145 (294)
147 remove duplicates from 139 (154)
148 from 147 keep 49-154 [Ann Checkups RCT ML] (106)
149 from 147 keep 1-48 [Ann Checkups RCT HS] (48)
150 remove duplicates from 140 (202)
151 from 148 keep 1-106 (106)
152 from 149 keep 1-48 (48)
153 remove duplicates from 138 (444)

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations <17 December 2010>

---

1 health check$.ti. (15)
2 check-up?.ti. (27)
3 1 or 2 (35)

Appendix 3. Strategies: Cochrane, EMBASE, EPOC Register, CINAHL

The Cochrane Library
Search 1
("family practice* OR "family doctor*" OR "family physician*" OR "general practice" OR "general practitioner*"):ti AND Physical Examination[Mesh] (44)

Search 2:

#1 ("primary care" or "primary healthcare" or "primary health care" or "family doctor*" or "family physician*" or "general practice*" or "general practitioner*"):ti,ab (8459)
#2 MeSH descriptor Mass Screening (3629)
#3 MeSH descriptor Multiphasic Screening (17)
#4 (#1 AND (#2 OR #3)) (323 total [278 = trials])

Search 3:
#7 MeSH descriptor Physical Examination (676)
#8 (annual* OR yearly):ti (1072)
#9 (#7 AND #8) (1 trial)
Search 4:
#7 MeSH descriptor Physical Examination (676)
#8 (annual* OR yearly):ti (1072)
#9 (#7 AND #8) (3)
#10 MeSH descriptor Family Practice (2201)
#11 MeSH descriptor Physicians, Family (465)
#12 MeSH descriptor Primary Health Care (2387)
#13 (#7 AND (#10 OR #11 OR #12)) (28 trials)
EMBASE
EMBASE Classic+EMBASE <1947 to 2010 December 16>

1 (checkup? or check-up? or health check$).ti. (1597)
2 (healthcheck$ or annual physical? or annual medical or medical check$ or primary care check$ or well care or wellcare or well woman or well visit?).ti. (412)
3 "wellness check$".ti,ab. (11)
4 Physical examination/ and (annual or periodic or yearly or regular).ti. [EM] (477)
5 (*medical examination/ or *clinical examination/ or *functional assessment/ or periodic medical examination/) and (annual or yearly or regular).ti. [EM] (71)
6 (healthcheck$ or annual physical? or annual medical or medical check$ or primary care check$ or well care or wellcare or well woman or well visit?).ti. (412)
7 ((annual or periodic or regular or routine or yearly) and (GP visit? or physician? visit? or doctor? visit? or office visit?)).ti. (10)
8 ((annual or periodic or yearly) adj3 (GP visit? or physician? visit? or doctor? visit? or office visit?)).ab. (81)
9 (((primary adj2 (care or healthcare)) or general practitioner? or general practice or GP or family doctor? or family practice? or family physician?) adj3 ((physical or medical or health$) adj exam$)).ti,ab. (75)
10 ((annual or yearly) adj2 (medical? or physical?)).ti. (227)
11 ((annual or yearly) and visit?).ti. (23)
12 or/1-11 [Checkups --Combine with filters only EM] (2539)
13 *Physical examination/ [EM] (6179)
14 *medical examination/ or *clinical examination/ or *functional assessment/ or *periodic medical examination/ [EM] (7674)

15 ((annual or periodic or regular or routine or yearly) adj3 (medical assessment? or health assessment? or check$ or check-up? or health$ exam$ or health evaluation? or medical exam$ or physical? exam$ or wellness check$)).ab. (7423)
16 (annual medical or yearly medical or annual physical).ab. (696)
17 ((annual or periodic or regular or routine or yearly) adj3 (healthcheck? or health$ exam$ or health evaluation? or medical exam$ or office visit? or GP visit? or physical? exam$ or wellness check$)).ab. (3444)
18 ((annual or yearly) adj3 (physician? visit? or doctor? visit? or office visit?)).ab. (74)
19 "well care".ti,ab. (80)
20 (prevent$ and visit?).ti. or (prevent$ adj3 visit?).ab. (704)
21 ((annual or periodic or regular or routine or yearly) adj3 (preventive or preventative)).ab. [ADDED] (540)
22 (preventive? adj3 (care check$ or visit? or exam$)).ab. or (preventive? and (care check$ or visit? or exam$)).ti. (2122)
23 or/13-22 [Checkups general EM] (24077)
24 *mass screening/ [includes multiphasic EM] (20315)
25 *screening/ [EM] (9856)
26 *screening test/ [EM] (4541)
27 ((community$ or program? or multiphasic or multi-phasic or (primary adj2 care) or "office visit?" or GP or general practice or care or healthcare or routine or annual) adj2 screening).ab. (20713)
28 screening.ti. (96634)
29 or/24,27-28 [Screening Narrow EM] (117685)
30 *primary prevention/ or *preventive health service/ [EM Focussed] (12840)
31 *health promotion/ [used for healthy people programs EM Focussed] (23098)
32 (prevention or preventive or preventative).ti. (162581)
33 *Risk assessment/ [EM] (19282)
34 or/30-33 [Prevention/Risk Assessment EM Focussed] (209876)
35 *primary health care/ or *primary medical care/ or *general practitioner/ [EM focussed] (43797)
36 ((family or general) adj (doctor? or practice? or practitioner? or physician$)).ti. (42380)
37 (primary adj2 (care or health care or healthcare or medical care or patient care)).ti. (30926)
38 community care/ or community health nursing/ or community psychiatric nursing/ or health center/ [EM broad] (81230)
39 *community care/ or *community health nursing/ or *community psychiatric nursing/ or *health center/ [EM broad] (39861)
40 community$.ti. (83113)
41 or/35-37,39-40 [Primary/Community Care EM Focussed] (196537)
42 (elderly or geriatric?).hw. (99410)
43 *aged/ [EM focussed] (39537)
44 (exp *cardiovascular disease/ or exp *digestive system disease/ or exp *musculoskeletal disease/ or chronic obstructive lung disease/ or exp *asthma/) and (pc or di).fs. [EM] (1035530)
45 disease?.ti. and (pc or di).fs. [EM did not use hw] (211164)
46 (diabet$ or cardio$ or heart or disease or copd).ti. (1266353)
47 or/44-46 [Diseases--selected EM] (2179709)
48 randomized controlled trial/ or controlled study/ or major clinical study/ or random$.ti,ab. or ((control or controlled) adj3 (trial? or study or group? or cohort?)).ti,ab. [EM] (4999915)
49 (clinical trial/ or clinical study/) and (control or controlled).ti. (53389)
50 (editorial or letter or note or "review" or trade or survey).pt. [to exclude irrelevant publication types EM] (3469302)
51 (animal model? or animal experiment? or animal study? or animal trial? or canine or feline or bovine or cow or cows or mice or dog? or cat or cats or rabbit? or rat or rats or veterinar$).ti. or (animal or veterinary).hw. (4530797)
52 (or/48-49) not (or/50-51) [RCT for EM] (3465031)
53 intervention?.ti. or (intervention? adj6 (clinician? or collaborat$ or community or complex or DESIGN$ or doctor? or educational or family doctor? or family physician? or family practitioner? or financial or GP or general practice? or hospital? or impact? or improv$ or individuali?e? or individuali?ing or interdisciplin$ or multicomponent or multi-component or multidisciplin$ or multi-disciplin$ or multifacet$ or multi-facet$ or multimodal$ or multi-modal$ or personali?e? or personali?ing or pharmacies or pharmacist? or pharmacy or physician? or practitioner? or prescrib$ or prescription? or primary care or professional$ or provider? or regulatory or regulatory or tailor$ or target$ or team$ or usual care)).ab. (131021)
54 (collaborativ$ or collaboration? or tailored or personali?ed).ti,ab. [added v2.0] (91299)
55 (exp *hospital/ or *hospitalization/ or *patient/ or *outpatient/ or *exp hospital patient/) and (study.ti. or *evaluation/) [EM] (10478)

56 (exp *nurse/ or exp *nursing/) and (study.ti. or *evaulation/) [EM] (4478)
57 demonstration project?.ti,ab. (1989)
58 (pre-post or "pre test$" or pretest$ or posttest$ or "post test$" or (pre adj5 post)).ti,ab. (58602)
59 (pre-workshop or post-workshop or (before adj3 workshop) or (after adj3 workshop)).ti,ab. (491)
60 trial.ti. or ((study adj3 aim?) or "our study").ab. (526804)
61 (before adj10 (after or during)).ti,ab. (372526)
62 ("quasi-experiment$" or quasiexperiment$ or "quasi random$" or quasirandom$ or "quasi control$" or quasicontrol$ or ((quasi$ or experimental) adj3 (method$ or study or trial or design$))).ti,ab,hw. [ML] (119667)
63 ("time series" adj2 interrupt$).ti,ab,hw. [ML] (624)
64 (time points adj3 (over or multiple or three or four or five or six or seven or eight or nine or ten or eleven or twelve or month$ or hour? or day? or "more than")).ab. (6874)
65 pilot.ti. (33592)
66 (multicentre or multicenter or multi-centre or multi-center).ti. (24946)
67 random$.ti,ab. or controlled.ti. (686009)
68 *experimental design/ or *pilot study/ or quasi experimental study/ [EM] (2719)
69 ("quasi-experiment$" or quasiexperiment$ or "quasi random$" or quasirandom$ or "quasi control$" or quasicontrol$ or ((quasi$ or experimental) adj3 (method$ or study or trial or design$))).ti,ab. [EM] (103508)
70 ("time series" adj2 interrupt$).ti,ab. [EM] (624)
71 (control adj3 (area or cohort? or compar? or condition or group? or intervention? or participant? or study)).ab. not (randomized controlled trial/ or controlled study/ or major clinical study/) [EM] (139311)
72 (animal model? or animal experiment? or animal study? or animal trial? or canine or feline or bovine or cow or cows or mice or dog? or cat or cats or rabbit? or rat or rats or veterinar$).ti. or (animal or veterinary).hw. [EM] (4530797)
73 (editorial or letter or note or "review" or trade or survey).pt. [EM] (3469302)
74 (or/53-71) not (or/72-73) [EPOC Filter 2.1 EM] (1446917)
75 "Public Health, Social Medicine and Epidemiology ".ec. (1499319)
76 12 and 52 [Annual Checkups & RCT] (430)
77 12 and 74 [Annual Checkups & EPOC] (213)
78 23 and 29 [Physical Exams & Screening] (1408)
79 23 and 34 [Physical Exams & Prevention/Risk Assessment] (1814)
80 23 and 41 [Physical Exams & Primary Care] (914)
81 23 and (or/42-43) [Physical Exams & Aged/geriatric/elderly] (647)
82 23 and 47 [Physical Exams & Diseases these do not seem relevant] (3308)
83 (or/78-81) and 52 [RCT Results 2] (1154)
84 (or/78-81) and 74 [EPOC Results 2] (701)
85 83 not (or/76-77) [Remaining RCT results dupes removed] (1103)
86 84 not (or/76-77,83) [Remaining EPOC results dupes removed] (168)

EPOC Register (Reference Manager)

health check* OR check up* or physical exam* or annual physical [all fields] 59 results

CINAHL (EBSCOhost)

#	**Query** Friday, 17 December 2010 4:07:39 PM	**Results**
S34	(S1 OR S11 OR S14 OR S15 OR S17) AND (S8 OR S12 OR S33)	380
S33	S18 or S19 or S20 or S21 or S22 or S23 or S24 or S25 or S26 or S28 or S29 or S30 or S31 or S32	Display
S32	TI pilot	Display

(*Continued*)

S31	(MH "Pilot Studies")	Display
S30	AB "before-and-after"	Display
S29	AB time series	Display
S28	TI time series	Display
S27	AB (before* n7 during or before n3 after) or AU (before* n7 during or before n3 after)	Display
S26	TI ((time point*) or (period* n4 interrupted) or (period* n4 multiple) or (period* n4 time) or (period* n4 various) or (period* n4 varying) or (period* n4 week*) or (period* n4 month*) or (period* n4 year*)) or AB ((time point*) or (period* n4 interrupted) or (period* n4 multiple) or (period* n4 time) or (period* n4 various) or (period* n4 varying) or (period* n4 week*) or (period* n4 month*) or (period* n4 year*))	Display
S25	TI ((quasi-experiment* or quasiexperiment* or quasi-random* or quasirandom* or quasi control* or quasicontrol* or "quasi* W3 method*" or "quasi* W3 study" or "quasi* W3 studies" or "quasi* W3 trial" or "quasi* W3 design*" or "experimental W3 method*" or "experimental W3 study" or "experimental W3 studies" or "experimental W3 trial" or "experimental W3 design*")) or AB ((quasi-experiment* or quasi-experiment* or quasi-random* or quasirandom* or quasi control* or quasicontrol* or "quasi* W3 method*" or "quasi* W3 study" or "quasi* W3 studies" or "quasi* W3 trial" or "quasi* W3 design*" or "experimental W3 method*" or "experimental W3 study" or "experimental W3 studies" or "experimental W3 trial" or "experimental W3 design*"))	Display
S24	TI pre w7 post or AB pre w7 post	Display
S23	MH "Multiple Time Series" or MH "Time Series"	Display
S22	TI ((comparative N2 study) or (comparative N2 studies) or "evaluation study" or "evaluation studies") or AB ((comparative N2 study) or (comparative N2 studies) or "evaluation study" or "evaluation studies")	Display
S21	MH Experimental Studies or Community Trials or Community Trials or Pretest-Posttest Design + or Quasi-Experimental Studies + Pilot Studies or Policy Studies + Multicenter Studies	Display

(*Continued*)

S20	TI ("pre test*" or pretest* or posttest* or "post test*") or AB ("pre test*" or pretest* or posttest* or "post test*")	Display
S19	TI (intervention* or multiintervention* or multi-intervention* or postintervention* or post-intervention* or preintervention* or pre-intervention*) or AB (intervention* or multiintervention* or multi-intervention* or postintervention* or post-intervention* or preintervention* or pre-intervention*)	Display
S18	(MH "Quasi-Experimental Studies")	Display
S17	(S10 AND S16) AND prevent*	652
S16	MW screening or "multiphasic* screen*"	24631
S15	TI Physical Exam* and TI (annual* OR yearly)	15
S14	S9 and S13	64
S13	TI annual or yearly	13826
S12	TI (intervention* OR collaborat* or team* or efficacy or effectiveness) or AB (intervention* OR collaborat* or team* or efficacy or effectiveness) or MW (intervention* OR collaborat* or team* or efficacy or effectiveness)	234766
S11	S9 and S10	409
S10	MH Family practice or MH Physicians, family or MH primary health care	31992
S9	MH physical examination	12098
S8	S2 or S3 or S4 or S5 or S6 or S7	Display
S7	TI ("control* N1 clinical" or "control* N1 group*" or "control* N1 trial*" or "control* N1 study" or "control* N1 studies" or "control* N1 design*" or "control* N1 method*") or AB ("control* N1 clinical" or "control* N1 group*" or "control* N1 trial*" or "control* N1 study" or "control* N1 studies" or "control* N1 design*" or "control* N1 method*")	Display
S6	TI controlled or AB controlled	Display
S5	TI random* or AB random*	Display
S4	TI ("clinical study" or "clinical studies") or AB ("clinical study" or "clinical studies")	Display

(*Continued*)

S3	(MM "Clinical Trials+")	Display
S2	TI ((multicent* n2 design*) or (multicent* n2 study) or (multicent* n2 studies) or (multicent* n2 trial*)) or AB ((multicent* n2 design*) or (multicent* n2 study) or (multicent* n2 studies) or (multicent* n2 trial*))	Display
S1	TI "checkup*" OR "check up*" or "health check*"	487

Appendix 4. Strategies for July 2012 update

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>; search date: 4 July 2012

1 Physical examination/ and ((annual or GP or periodic or yearly or routine).ti. or ((primary adj2 (care or healthcare)) or primary health$ or general practitioner? or general practice or family doctor? or family practice? or family physician?).ti,ab.) (2073)
2 (health check$ or healthcheck$ or annual physical? or annual medical or medical check$ or primary care check$ or wellness check$ well care or wellcare or well woman or well visit?).ti. (806)
3 ((annual or periodic or regular or routine or yearly) and (check$ or check-up? or health$ exam$ or health evaluation? or medical exam$ or physical? exam$ or wellness check$ or GP visit? or physician? visit? or doctor? visit? or office visit?)).ti. (886)
4 ((annual or yearly) adj2 (medical? or physical?)).ti. (267)
5 ((annual or yearly) and visit?).ti. (28)
6 (preventive? and (care check$ or checkup? or check-up? or visit? or exam$ or family doctor? or GP or family physician? or general practitioner?)).ti. (769)
7 or/1-6 [Annual Checkups --Combine with filters only] (4074)
8 7 not (cannibis or alcohol$ or abuse or narcotics or addiction?).ti. (4019)
9 Physical examination/ (27194)
10 (check-up? or checkup?).ti,ab. (8444)
11 (annual medical or yearly medical or annual physical).ab. (616)
12 ((annual or periodic or (primary adj2 (care or healthcare)) or primary health$ or general practitioner? or general practice or GP or family doctor? or family practice? or family physician? or regular or routine or yearly) adj3 (healthcheck? or health$ exam$ or health evaluation? or medical exam$ or office visit? or GP visit? or physical? exam$ or wellness check$)).ab. (3056)
13 ((annual or yearly) adj3 (physician? visit? or doctor? visit? or office visit?)).ab. (57)
14 "well care".ti,ab. (74)
15 (prevent$ and (screen$ or visit?)).ti. or (prevent$ adj3 (screen$ or visit?)).ab. (5502)
16 or/9-15 [Checkups general] (43592)
17 Mass screening/ (74782)
18 Multiphasic screening/ [ML] (1032)
19 ((community$ or program? or multiphasic or multi-phasic or (primary adj2 care) or "office visit?" or GP or general practice or care or healthcare or routine or annual) adj2 screening).ab. (18736)
20 screening.ti. (90599)
21 or/17-20 [Screening] (139691)
22 Primary prevention/ [ML] (12669)
23 exp Preventive Health Services/ (388517)
24 Health promotion/ or Healthy People Programs/ (46825)
25 (prevention or preventive or preventative).ti. (141654)
26 Risk assessment/ (147905)
27 or/22-26 [Prevention/Risk Assessment] (641516)
28 Risk factors/ (488641)

29 or/22-26,28 [Prevention/Risk Assessment/Risk Factors] (1054069)
30 exp Primary health care/ or Family practice/ or Physicians, family/ (132932)
31 ((family or general) adj (doctor? or practice? or practitioner? or physician$)).ti. (36432)
32 (primary adj2 (care or health care or healthcare or medical care or patient care)).ti. (30347)
33 Community Health services/ or Community mental Health Services/ or Community Pharmacy Services/ or Mobile Health units/ or Community Health Centers/ or Community health nursing/ (67284)
34 community$.ti. (79276)
35 or/30-34 [Primary/Community Care] (265615)
36 exp Aged/ [Elderly as group are unique whereas Adult is often not mentioned in indexing] (2111749)
37 (exp Cardiovascular Diseases/ or exp Digestive System Diseases/ or exp Endocrine System Diseases/ or exp Musculoskeletal Diseases/ or exp Lung Diseases, Obstructive/) and (pc or di).fs. (1042974)
38 disease?.hw. and (pc or di).fs. (773829)
39 (diabet$ or cardio$ or heart or disease or copd).ti. (1104250)
40 or/37-39 [Diseases--selected] (2326288)
41 (randomized controlled trial or controlled clinical trial).pt. or randomized.ab. or placebo.ab. or clinical trials as topic.sh. or randomly.ab. or trial.ti. (792755)
42 exp animals/ not humans.sh. (3736637)
43 "comment on".cm. or systematic review.ti. or literature review.ti. or editorial.pt. or letter.pt. or meta-analysis.pt. or news.pt. or review.pt. [to exclude irrelevant publication types] (3035036)
44 41 not (or/42-43) [Cochrane RCT Filter 6.4.d Sens/Precision Maximizing] (604530)
45 17 and (or/26,28) [Screening & Risk Factors/Assessment] (13587)
46 intervention?.ti. or (intervention? adj6 (clinician? or collaborat$ or community or complex or DESIGN$ or doctor? or educational or family doctor? or family physician? or family practitioner? or financial or GP or general practice? or hospital? or impact? or improv$ or individuali?e? or individuali?ing or interdisciplin$ or multicomponent or multi-component or multidisciplin$ or multi-disciplin$ or multifacet$ or multi-facet$ or multimodal$ or multi-modal$ or personali?e? or personali?ing or pharmacies or pharmacist? or pharmacy or physician? or practitioner? or prescrib$ or prescription? or primary care or professional$ or provider? or regulatory or regulatory or tailor$ or target$ or team$ or usual care)).ab. (129705)
47 (collaborativ$ or collaboration? or tailored or personali?ed).ti,ab. [added v2.0] (88248)
48 (exp hospitals/ or exp Hospitalization/ or exp Patients/ or exp Nurses/ or exp Nursing/) and (study.ti. or evaluation studies as topic/) [changed for v2.0 based on analysis of Mesh found on CBA & ITS not found by Filter 1.6] (33815)
49 demonstration project?.ti,ab. (1765)
50 (pre-post or "pre test$" or pretest$ or posttest$ or "post test$" or (pre adj5 post)).ti,ab. (53471)
51 (pre-workshop or post-workshop or (before adj3 workshop) or (after adj3 workshop)).ti,ab. (482)
52 trial.ti. or ((study adj3 aim?) or "our study").ab. (505273)
53 (before adj10 (after or during)).ti,ab. (318191)
54 ("quasi-experiment$" or quasiexperiment$ or "quasi random$" or quasirandom$ or "quasi control$" or quasicontrol$ or ((quasi$ or experimental) adj3 (method$ or study or trial or design$))).ti,ab,hw. [ML] (88695)
55 ("time series" adj2 interrupt$).ti,ab,hw. [ML] (720)
56 (time points adj3 (over or multiple or three or four or five or six or seven or eight or nine or ten or eleven or twelve or month$ or hour? or day? or "more than")).ab. (7030)
57 pilot.ti. (32625)
58 Pilot projects/ [ML] (71559)
59 (clinical trial or multicenter study).pt. [ML removed RCT--redundant v2.0] (569629)
60 (multicentre or multicenter or multi-centre or multi-center).ti. (24273)
61 random$.ti,ab. or controlled.ti. (643784)
62 (control adj3 (area or cohort? or compar? or condition or group? or intervention? or participant? or study)).ab. not (controlled clinical trial or randomized controlled trial).pt. [ML remove DESIGN changed truncation on Compare] (283159)
63 "comment on".cm. or systematic review.ti. or literature review.ti. or editorial.pt. or letter.pt. or meta-analysis.pt. or news.pt. or review.pt. [to exclude irrelevant publication types] (3035036)
64 exp animals/ not humans.sh. [ML] (3736637)
65 *experimental design/ or *pilot study/ or quasi experimental study/ [EM] (18154)
66 ("quasi-experiment$" or quasiexperiment$ or "quasi random$" or quasirandom$ or "quasi control$" or quasicontrol$ or ((quasi$ or experimental) adj3 (method$ or study or trial or design$))).ti,ab. [EM] (88695)

67 ("time series" adj2 interrupt$).ti,ab. [EM] (720)
68 (animal/ or animal.hw.) not ((animal/ or animal?.kw,hw.) and (human/ or human?.hw,kw.)) [EM] (3640992)
69 (book or letter).pt. [EM] (768605)
70 (or/46-53,56-57,60-62,65-67) not (or/68-69) [EPOC Methods Filter EM 2.0] (1564756)
71 (or/46-62) not (or/63-64) [EPOC Methods Filter ML 2.0] (1712551)
72 16 and 35 [Checkups (general) & Primary Care/Community Care] (3163)
73 16 and (or/21,27,40) [Checkup(general) & Screening/Prevention/Disease pc] (20424)
74 7 and 71 [Annual Checkups and EPOC Filter] (502)
75 7 and 44 [Annual Checkups and RCT Filter] (167)
76 16 and 35 and 44 [Checkups (general) & Primary/Community Care RCT Filter] (258)
77 16 and 35 and 71 [Checkups (general) & Primary/Community Care EPOC Filter] (774)
78 (or/17-18) and (or/26,28) [Screening & Risk Factors/Assessment] (13635)
79 74 not 75 [Unique Annual Checkups EPOC Filter] (341)
80 16 and (or/21,27,40) and 44 [Checkup(general) & Screening/Prev/Disease pc & RCT] (984)
81 (or/17-18) and (or/26,28) and 44 [Screening & Risk Factors/Assessment &RCT] (775)
82 or/75-76,80-81 [RCT Results] (1842)
83 (or/74,77,79) not 82 [EPOC Results] (738)
84 (2011$ or "2012" or 201012$).ed. [Entry Date Limit Dec 2010; 2011- July 4, 2012] (1078390)
85 82 and 84 [RCT Results limited by Entry Date] (173)
86 83 and 84 [EPOC Results limited by Entry Date] (60)
87 limit 82 to yr="2011-current" [RCT Results limited by year] (189)
88 limit 83 to yr="2011-Current" [EPOC Results limited by year] (72)
89 85 or 87 [RCT Results Dec.2010- Jul.2012] (249)
90 86 or 88 [EPOC Results Dec.2010-Jul.2012] (91)

EMBASE

Embase Classic+Embase <1947 to 2012 July 03>; search date 3 July 2012

1 (checkup? or check-up? or health check$).ti. (1770)
2 (healthcheck$ or annual physical? or annual medical or medical check$ or primary care check$ or well care or wellcare or well woman or well visit?).ti. (448)
3 "wellness check$".ti,ab. (11)
4 Physical examination/ and (annual or periodic or yearly or regular).ti. [EM] (518)
5 (*medical examination/ or *clinical examination/ or *functional assessment/ or periodic medical examination/) and (annual or yearly or regular).ti. [EM] (80)
6 (healthcheck$ or annual physical? or annual medical or medical check$ or primary care check$ or well care or wellcare or well woman or well visit?).ti. (448)
7 ((annual or periodic or regular or routine or yearly) and (GP visit? or physician? visit? or doctor? visit? or office visit?)).ti. (10)
8 ((annual or periodic or yearly) adj3 (GP visit? or physician? visit? or doctor? visit? or office visit?)).ab. (89)
9 (((primary adj2 (care or healthcare)) or general practitioner? or general practice or GP or family doctor? or family practice? or family physician?) adj3 ((physical or medical or health$) adj exam$)).ti,ab. (90)
10 ((annual or yearly) adj2 (medical? or physical?)).ti. (248)
11 ((annual or yearly) and visit?).ti. (33)
12 or/1-11 [Checkups --Combine with filters only EM] (2809)
13 *Physical examination/ [EM] (6781)
14 *medical examination/ or *clinical examination/ or *functional assessment/ or *periodic medical examination/ [EM] (8194)
15 ((annual or periodic or regular or routine or yearly) adj3 (medical assessment? or health assessment? or check$ or check-up? or health$ exam$ or health evaluation? or medical exam$ or physical? exam$ or wellness check$)).ab. (8662)
16 (annual medical or yearly medical or annual physical).ab. (839)
17 ((annual or periodic or regular or routine or yearly) adj3 (healthcheck? or health$ exam$ or health evaluation? or medical exam$ or office visit? or GP visit? or physical? exam$ or wellness check$)).ab. (3946)
18 ((annual or yearly) adj3 (physician? visit? or doctor? visit? or office visit?)).ab. (80)
19 "well care".ti,ab. (92)

20 (prevent$ and visit?).ti. or (prevent$ adj3 visit?).ab. (851)
21 ((annual or periodic or regular or routine or yearly) adj3 (preventive or preventative)).ab. [ADDED] (647)
22 (preventive? adj3 (care check$ or visit? or exam$)).ab. or (preventive? and (care check$ or visit? or exam$)).ti. (2378)
23 or/13-22 [Checkups general EM] (26945)
24 *mass screening/ [includes multiphasic EM] (22720)
25 *screening/ [EM] (13506)
26 *screening test/ [EM] (5316)
27 ((community$ or program? or multiphasic or multi-phasic or (primary adj2 care) or "office visit?" or GP or general practice or care or healthcare or routine or annual) adj2 screening).ab. (25115)
28 screening.ti. (116308)
29 or/24,27-28 [Screening Narrow EM] (140282)
30 *primary prevention/ or *preventive health service/ [EM Focussed] (14694)
31 *health promotion/ [used for healthy people programs EM Focussed] (26478)
32 (prevention or preventive or preventative).ti. (181121)
33 *Risk assessment/ [EM] (22751)
34 or/30-33 [Prevention/Risk Assessment EM Focussed] (235714)
35 *primary health care/ or *primary medical care/ or *general practitioner/ [EM focussed] (49809)
36 ((family or general) adj (doctor? or practice? or practitioner? or physician$)).ti. (45333)
37 (primary adj2 (care or health care or healthcare or medical care or patient care)).ti. (36546)
38 community care/ or community health nursing/ or community psychiatric nursing/ or health center/ [EM broad] (87000)
39 *community care/ or *community health nursing/ or *community psychiatric nursing/ or *health center/ [EM broad] (42258)
40 community$.ti. (95980)
41 or/35-37,39-40 [Primary/Community Care EM Focussed] (219858)
42 (elderly or geriatric?).hw. (110920)
43 *aged/ [EM focussed] (42581)
44 (exp *cardiovascular disease/ or exp *digestive system disease/ or exp *musculoskeletal disease/ or chronic obstructive lung disease/ or exp *asthma/) and (pc or di).fs. [EM] (1133883)
45 disease?.ti. and (pc or di).fs. [EM did not use hw] (231436)
46 (diabet$ or cardio$ or heart or disease or copd).ti. (1468913)
47 or/44-46 [Diseases--selected EM] (2466604)
48 randomized controlled trial/ or controlled study/ or major clinical study/ or random$.ti,ab. or ((control or controlled) adj3 (trial? or study or group? or cohort?)).ti,ab. [EM] (5589117)
49 (clinical trial/ or clinical study/) and (control or controlled).ti. (56250)
50 (editorial or letter or note or "review" or trade or survey).pt. [to exclude irrelevant publication types EM] (3890278)
51 (animal model? or animal experiment? or animal study? or animal trial? or canine or feline or bovine or cow or cows or mice or dog? or cat or cats or rabbit? or rat or rats or veterinar$).ti. or (animal or veterinary).hw. (4884807)
52 (or/48-49) not (or/50-51) [RCT for EM] (3879863)
53 intervention?.ti. or (intervention? adj6 (clinician? or collaborat$ or community or complex or DESIGN$ or doctor? or educational or family doctor? or family physician? or family practitioner? or financial or GP or general practice? or hospital? or impact? or improv$ or individuali?e? or individuali?ing or interdisciplin$ or multicomponent or multi-component or multidisciplin$ or multi-disciplin$ or multifacet$ or multi-facet$ or multimodal$ or multi-modal$ or personali?e? or personali?ing or pharmacies or pharmacist? or pharmacy or physician? or practitioner? or prescrib$ or prescription? or primary care or professional$ or provider? or regulatory or regulatory or tailor$ or target$ or team$ or usual care)).ab. (167374)
54 (collaborativ$ or collaboration? or tailored or personali?ed).ti,ab. [added v2.0] (116023)
55 (exp *hospital/ or *hospitalization/ or *patient/ or *outpatient/ or *exp hospital patient/) and (study.ti. or *evaluation/) [EM] (24253)
56 (exp *nurse/ or exp *nursing/) and (study.ti. or *evaulation/) [EM] (5137)
57 demonstration project?.ti,ab. (2186)
58 (pre-post or "pre test$" or pretest$ or posttest$ or "post test$" or (pre adj5 post)).ti,ab. (76162)
59 (pre-workshop or post-workshop or (before adj3 workshop) or (after adj3 workshop)).ti,ab. (637)
60 trial.ti. or ((study adj3 aim?) or "our study").ab. (688547)
61 (before adj10 (after or during)).ti,ab. (427106)

62 ("quasi-experiment$" or quasiexperiment$ or "quasi random$" or quasirandom$ or "quasi control$" or quasicontrol$ or ((quasi$ or experimental) adj3 (method$ or study or trial or design$))).ti,ab,hw. [ML] (204291)
63 ("time series" adj2 interrupt$).ti,ab,hw. [ML] (859)
64 (time points adj3 (over or multiple or three or four or five or six or seven or eight or nine or ten or eleven or twelve or month$ or hour? or day? or "more than")).ab. (9331)
65 pilot.ti. (42375)
66 (multicentre or multicenter or multi-centre or multi-center).ti. (33074)
67 random$.ti,ab. or controlled.ti. (811498)
68 *experimental design/ or *pilot study/ or quasi experimental study/ [EM] (4792)
69 ("quasi-experiment$" or quasiexperiment$ or "quasi random$" or quasirandom$ or "quasi control$" or quasicontrol$ or ((quasi$ or experimental) adj3 (method$ or study or trial or design$))).ti,ab. [EM] (117089)
70 ("time series" adj2 interrupt$).ti,ab. [EM] (859)
71 (control adj3 (area or cohort? or compar? or condition or group? or intervention? or participant? or study)).ab. not (randomized controlled trial/ or controlled study/ or major clinical study/) [EM] (178450)
72 (animal model? or animal experiment? or animal study? or animal trial? or canine or feline or bovine or cow or cows or mice or dog? or cat or cats or rabbit? or rat or rats or veterinar$).ti. or (animal or veterinary).hw. [EM] (4884807)
73 (editorial or letter or note or "review" or trade or survey).pt. [EM] (3890278)
74 (or/53-71) not (or/72-73) [EPOC Filter 2.1 EM] (1833355)
75 "Public Health, Social Medicine and Epidemiology ".ec. (1649983)
76 12 and 52 [Annual Checkups & RCT] (492)
77 12 and 74 [Annual Checkups & EPOC] (268)
78 23 and 29 [Physical Exams & Screening] (1585)
79 23 and 34 [Physical Exams & Prevention/Risk Assessment] (2015)
80 23 and 41 [Physical Exams & Primary Care] (1055)
81 23 and (or/42-43) [Physical Exams & Aged/geriatric/elderly] (708)
82 23 and 47 [Physical Exams & Diseases these do not seem relevant] (3962)
83 (or/78-81) and 52 [RCT Results 2] (1304)
84 (or/78-81) and 74 [EPOC Results 2] (846)
85 83 not (or/76-77) [Remaining RCT results dupes removed] (1244)
86 84 not (or/76-77,83) [Remaining EPOC results dupes removed] (226)
87 76 or 83 [RCT Results] (1736)
88 77 or 84 [EPOC Results] (1070)
89 limit 87 to em="201049-201227" [RCT limited by Entry Week Dec2010 - Jul2012] (207)
90 limit 88 to em="201049-201227" [EPOC limited by Entry Week Dec2010-Jul2012] (161)
91 limit 87 to yr="2011-Current" [RCT results limited by year] (163)
92 limit 88 to yr="2011-Current" [EPOC results limited by year] (137)
93 89 or 91 [RCT results Dec2010 - Jul2012] (217)
94 90 or 92 [EPOC results Dec2010 - Jul2012] (179)

Cochrane Library

Cochrane Library, Issue 7, 2012; search date 9 July 2012
#1 (Checkup* or "check-up*" or "health check*"):ti (44)
#2 ("family practice"):ti or ("family doctor"):ti or ("family physician"):ti or ("general practice"):ti or ("general practitioner"):ti (1801)
#3 MeSH descriptor Physical Examination explode all trees (51357)
#4 (#2 AND #3) (103)
#5 ("primary care" or "primary healthcare" or "primary health care" or "family doctor" or "family physician" or "general practice" or "general practitioner"):ti,ab,kw (9577)
#6 MeSH descriptor Mass Screening (3567)
#7 MeSH descriptor Multiphasic Screening explode all trees (16)
#8 (annual or annually or yearly):ti (1014)
#9 MeSH descriptor General Practice explode all trees (2114)
#10 MeSH descriptor General Practitioners (31)

#11 MeSH descriptor Physicians, Family (445)
#12 MeSH descriptor Physicians, Primary Care (21)
#13 MeSH descriptor Family Practice (2055)
#14 (#3 AND #8) (9)
#15 ((#5 AND (#6 OR #7)) AND NOT #14) (401)
#16 (#9 OR #10 OR #11 OR #12 OR #13) (2539)
#17 ("check up" or checkup or "health check"):ab,kw (316)
#18 (#8 AND #16) (4)
#19 (#1 OR #4 OR #14 OR #15 OR #18) (553)
#20 ((#17 AND (#2 OR #5 OR #16)) AND NOT #19) (30)

CINAHL

	Search date: 4 July 2012	
#	**Query**	**Results**
S37	s35 or s36 **[results from Dec 2010 - July 2012]**	73
S36	S34 AND EM 20101217-20120704	73
S35	S34 AND DT 20101217-20120704	49
S34	(S1 OR S11 OR S14 OR S15 OR S17) AND (S8 OR S12 OR S33)	463
S33	S18 or S19 or S20 or S21 or S22 or S23 or S24 or S25 or S26 or S28 or S29 or S30 or S31 or S32	229043
S32	TI pilot	9895
S31	(MH "Pilot Studies")	25656
S30	AB "before-and-after"	14831
S29	AB time series	1513
S28	TI time series	210
S27	AB (before* n7 during or before n3 after) or AU (before* n7 during or before n3 after)	23111
S26	TI ((time point*) or (period* n4 interrupted) or (period* n4 multiple) or (period* n4 time) or (period* n4 various) or (period* n4 varying) or (period* n4 week*) or (period* n4 month*) or (period* n4 year*)) or AB ((time point*) or (period* n4 interrupted) or (period* n4 multiple) or (period* n4 time) or (period* n4 various) or (period* n4 varying) or (period* n4 week*) or (period* n4 month*) or (period* n4 year*)	42975

(*Continued*)

	)	
S25	TI ((quasi-experiment* or quasiexperiment* or quasi-random* or quasirandom* or quasi control* or quasicontrol* or "quasi* W3 method*" or "quasi* W3 study" or "quasi* W3 studies" or "quasi* W3 trial" or "quasi* W3 design*" or "experimental W3 method*" or "experimental W3 study" or "experimental W3 studies" or "experimental W3 trial" or "experimental W3 design*")) or AB ((quasi-experiment* or quasi-experiment* or quasi-random* or quasirandom* or quasi control* or quasicontrol* or "quasi* W3 method*" or "quasi* W3 study" or "quasi* W3 studies" or "quasi* W3 trial" or "quasi* W3 design*" or "experimental W3 method*" or "experimental W3 study" or "experimental W3 studies" or "experimental W3 trial" or "experimental W3 design*"))	10542
S24	TI pre w7 post or AB pre w7 post	7697
S23	MH "Multiple Time Series" or MH "Time Series"	1158
S22	TI ((comparative N2 study) or (comparative N2 studies) or "evaluation study" or "evaluation studies") or AB ((comparative N2 study) or (comparative N2 studies) or "evaluation study" or "evaluation studies")	9014
S21	MH Experimental Studies or Community Trials or Community Trials or Pretest-Posttest Design + or Quasi-Experimental Studies + Pilot Studies or Policy Studies + Multicenter Studies	29721
S20	TI ("pre test*" or pretest* or posttest* or "post test*") or AB ("pre test*" or pretest* or posttest* or "post test*")	7053
S19	TI (intervention* or multiintervention* or multi-intervention* or postintervention* or post-intervention* or preintervention* or pre-intervention*) or AB (intervention* or multiintervention* or multi-intervention* or postintervention* or post-intervention* or preintervention* or pre-intervention*)	127480
S18	(MH "Quasi-Experimental Studies")	5077
S17	(S10 AND S16) AND prevent*	787
S16	MW screening or "multiphasic* screen*"	29176
S15	TI Physical Exam* and TI (annual* OR yearly)	17
S14	S9 and S13	76

(*Continued*)

S13	TI annual or yearly	14847
S12	TI (intervention* OR collaborat* or team* or efficacy or effectiveness) or AB (intervention* OR collaborat* or team* or efficacy or effectiveness) or MW (intervention* OR collaborat* or team* or efficacy or effectiveness)	275161
S11	S9 and S10	469
S10	MH Family practice or MH Physicians, family or MH primary health care	38050
S9	MH physical examination	13968
S8	S2 or S3 or S4 or S5 or S6 or S7	159942
S7	TI ("control* N1 clinical" or "control* N1 group*" or "control* N1 trial*" or "control* N1 study" or "control* N1 studies" or "control* N1 design*" or "control* N1 method*") or AB ("control* N1 clinical" or "control* N1 group*" or "control* N1 trial*" or "control* N1 study" or "control* N1 studies" or "control* N1 design*" or "control* N1 method*")	70716
S6	TI controlled or AB controlled	54265
S5	TI random* or AB random*	94245
S4	TI ("clinical study" or "clinical studies") or AB ("clinical study" or "clinical studies")	22384
S3	(MM "Clinical Trials+")	7295
S2	TI ((multicent* n2 design*) or (multicent* n2 study) or (multicent* n2 studies) or (multicent* n2 trial*)) or AB ((multicent* n2 design*) or (multicent* n2 study) or (multicent* n2 studies) or (multicent* n2 trial*))	6772
S1	TI "checkup*" OR "check up*" or "health check*"	564

Trial Registries

WHO ICTRP (14 Results) 31 July 2012

health checks OR checkup OR check up (2 Results)
health checks OR checkup OR check up OR physical exam OR annual physical (9 Results)
health check (3 Results)

Physical exam (0 Results)
Annual physical (0 results)
Periodic health evaluation (0 Results)
Periodic health examination (0 Results)
Yearly physical (0 Results)

ClinicalTrials.Gov (94 (July 31) + 74 (Aug 2) Results-duplicates could not be removed)

July 31, 2012:
("health check" OR "check up*" OR "checkup*") AND NOT (child OR adolescent OR teen OR adolescents OR teens OR teenager OR teenagers OR elderly OR children) (18 Results)
NOTE: The asterisk in the above search line appears to prevent the term from being searched. Consequently, reran the search string without * on August 2, 2012 and found 74 trials. These are very likely duplicates but results were sent to authors for review-M. Fiander
31 July 2012
"Periodic health evaluation" (2 Results)
"Periodic health examination" (7 Results)
"yearly physical" (1 Result)
("general health check" OR "health check up") AND NOT (child OR adolescent OR elderly OR "recovery management") (61 Results)
(checkup NOT (child OR adolescent OR elderly)) AND Adult (4 Results). Searched Intervention field
("health check" NOT (child OR adolescent OR elderly)) AND Adult (1 Result). Searched Intervention field
2 August 2 2012:
("health check" OR "health checks" OR "check ups" OR "check up" OR "checkups" OR "checkup") NOT (child OR infant OR elderly OR children OR teen OR teens OR teenager OR teenagers OR adolescent*) = 74

HISTORY

Protocol first published: Issue 2, 2011

Review first published: Issue 10, 2012

CONTRIBUTIONS OF AUTHORS

PCG initiated the project, LTK drafted the protocol and KJJ and PCG provided comments. LTK, KJJ and CGL screened titles and abstracts and made decisions about inclusion of trials. LTK and KJJ extracted data, LTK analysed data and drafted the review, and KJJ, PCG and CGL contributed to the revisions.

DECLARATIONS OF INTEREST

None

SOURCES OF SUPPORT

Internal sources

- Nordic Cochrane Centre, Denmark.

Salary and facilities

External sources

- Trygfonden, Denmark.

Part of salary for LTK

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We originally expected to include trials of geriatric screening but found that the intervention in most of these differed too much from our perception of what consitutes a health check. The actual medical screening was usually a minor component in a complex intervention involving other important interventions, for example screening for functional status; social, financial or legal needs; or home safety; or interventions such as specialist revision of individual medication or falls prevention. Consequently, it would not be possible to isolate the effect of the screening and we therefore chose not to include trials which were described as targeting an elderly population or which only included people over 65 years of age. Complex interventions directed at elderly people, including geriatric assessments, have been reviewed by Beswick and colleagues (Beswick 2008).

Guidelines for screening with urinary dipsticks differ substantially – a systematic review

Lasse T. Krogsbøll

ABSTRACT

INTRODUCTION: Urinary dipsticks are frequently used for screening as part of health checks and at hospital admission, but the benefits and harms of this are unknown.

METHODS: Health authorities and a selection of specialist societies in nine countries were identified through internet searches. Recommendations on dipstick screening at health checks or hospital admission were sought on websites as well as by email contact. Other relevant organisations encountered were also included. Recommendations were summarised narratively.

RESULTS: A total of 67 organisations were included. No positive or negative recommendations were found regarding screening with combined dipsticks. Screening for bacteriuria in non-pregnant persons was discouraged, while guidance on screening with dipsticks for haemoglobin, glucose and protein was uncommon and often unclear.

CONCLUSION: Useful guidance was rare. Practitioners are largely left to themselves when deciding whether or not to offer screening with urinary dipsticks. This situation needs to be remedied as benefit has not been shown and because screening with dipsticks can cause harm.

SYSTEMATIC REVIEW

The Nordic Cochrane Centre, Rigshospitalet

Dan Med J 2014;61(2):A4781

A frequently used component of general health checks is analysis of the urine [1, 2], which is often performed as a urinary dipstick test [3]. Patients admitted to hospital are also often routinely screened with a urinary dipstick, but the prevalence of this practice is unknown and likely varies between countries and regions. Use of urinary dipsticks may lead to detection of a wide array of serious conditions, e.g. urological cancers or glomerulonephritis. Early detection through screening could lead to improved prognosis, but it could also lead to unnecessary follow-up investigations such as kidney biopsies, cystoscopies, unnecessary antibiotic treatment, long-term follow-up of inconsequential abnormalities and psychological stress in healthy persons.

Dipsticks frequently combines testing for multiple substances, e.g. protein, glucose, blood, nitrite and leukocytes, which complicates the assessment of such testing. Screening for protein or albumin has been recommended for persons with certain risk factors [4-6] and is common in some countries, although there have been no trials on this [7]. In Japan, the general population has been systematically screened for proteinuria and haematuria with dipsticks for decades [8]. Enthusiasm for screening for asymptomatic microscopic haematuria has declined [9, 10], although not entirely [11, 12]. Screening asymptomatic non-pregnant persons for leukocytes, nitrite and glucose in the urine has fallen out of favour and it is unclear how often dipsticks are used for that purpose. However, it can be difficult to avoid as leukocytes and nitrite are frequently included in commonly used combined dipsticks.

There are no trials on screening for haemoglobin or protein in the urine [7, 10] and probably none on screening for glucose, leukocytes and nitrite. In other types of screening, trials have sometimes shown the benefits to be smaller than expected [13-16], and the harms greater [13, 14, 16]. In light of this lack of robust evidence, it is puzzling why screening with dipsticks is prevalent. One possible explanation may be that they are easy to use and are perceived as harmless. Furthermore, the idea that any early detection of disease is beneficial is widespread among clinicians and patients alike, despite evidence of over-diagnosis and other harms with several forms of screening [17].

It is the task of health authorities to provide recommendations on which interventions to use, both in sick and healthy people. Specialist societies also provide recommendations. The purpose of the present study was to find and describe existing recommendations on screening with urinary dipsticks, focusing on two types of screening: general health checks and routine screening of patients admitted to hospital.

METHODS

The search strategy was defined a priori, with the aim of limiting the workload while increasing the chance of finding the most important recommendations.

Six types of organisations were pre-specified: the main national health authority issuing guidance to health professionals and national professional societies for nephrology, urology, clinical biochemistry, general internal medicine, and general practice/family practice. Nine countries were pre-specified, based on the official language and on the likelihood of finding recommendations: Australia, Canada, Denmark, Ireland, New Zealand, Norway, Sweden, the United Kingdom and USA.

The internet was searched with Google to identify

the relevant organisations. When two organisations of the same kind from one country appeared equally important, they were both included. When online collections of guidelines were found, e.g. the National Guideline Clearinghouse (USA) or Helsebiblioteket (Norway), these were searched, too. Other organisations were also included when judged to be important, e.g. international organisations or charities, without first looking at the contents of their website.

The website of each included organisation was browsed for guidelines or recommendations on the topic and searched using relevant pre-specified keywords, when possible. The search terms were: urinary dipstick, dipstick, urinalysis, urine strip test, urine screening, routine urinalysis, routine dipstick, routine testing, routine admission testing, admission testing, bladder cancer AND screening, (haematuria OR haematuria) AND screening, kidney disease AND screening, renal disease AND screening, proteinuria AND screening, glomerulonephritis, diabetes AND screening, bacteriuria AND screening, cystitis AND screening, health check, health evaluation, health examination, albumin. The terms were modified to suit the individual search engines and were translated when needed.

Longer documents that might have contained guidance were also searched, e.g. health technology assessments. Finally, all included organisations were e-mailed and asked whether they knew of relevant guidelines, also guidelines issued by other organisations. Recommendations were sought regarding screening with combined dipsticks and common individual components: haemoglobin, protein or albumin, leukocytes and nitrite and glucose. Recommendations for screening of specific risk groups, e.g. people with diabetes or pregnant women, were not specifically sought out. When guidance on population-based screening programmes was found, it was included as such recommendations have relevance for screening in health checks.

Relevant text, including the reference, was copied into an Excel sheet. Information on whether the included websites linked to guidelines from other organisations was also recorded along with an indication of whether the organisation explicitly endorsed that guideline. The data collection was done in November and December 2010, and in January 2013 the websites were revisited to check for new guidelines and updates.

The results were summarised in tables and in narrative. No statistics were used.

RESULTS

A total of 67 organisations were included (**Figure 1**, **Table 1**). In six cases, more than one type of organisation from a country was included, in one case two websites from the same organisation were included, and in four cases two countries shared a specialist society. Three international specialist organisations, three charities and one guideline-producing network were also included because they appeared to be important sources of guidance. Of these, five were in nephrology, one in urology and one was general.

Health checks

Combined dipsticks

No recommendations were found on screening with combined dipsticks.

Haemoglobin

Only one organisation, the UK National Screening Committee, gave a recommendation regarding screening with dipsticks for haemoglobin, recommending against using them (**Table 2**) [18]. Nephrological and urological societies from the UK had a joint statement recommending against testing for haematuria in the absence of identifiable clinical reasons, but did not explicitly mention dipsticks [19].

Other organisations mentioned the topic without giving recommendations. Two stated that the evidence behind screening for bladder cancer was insufficient to determine the balance between benefits and harms [20, 21], two urological societies discussed the course of action when asymptomatic microscopic haematuria had been identified [22, 23], and a list of policy positions from one public authority stated “No policy” under screening for bladder cancer, while at the same time noting that it is “very common in general practice and often part of a routine medical examination” [24].

Leukocytes/nitrite

No organisations explicitly mentioned screening with dipsticks for leukocytes or nitrite, but four organisations

FIGURE 1

Overview of process. See Table 1 for names of included organisations and website addresses.

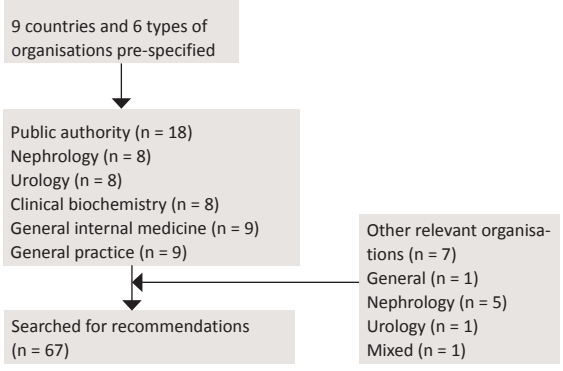


TABLE 1

List of organisations searched for recommendations.

Country	Organisation
Public authority	
Denmark	National Board of Health (www.sst.dk)
Sweden	Socialstyrelsen (www.socialstyrelsen.se), Statens Beredning för medicinsk Utvärdering (www.sbu.se)
Norway	The Norwegian Knowledge Centre for the Health Services, (Kunnskapssenteret, www.kunnskapssenteret.no/ and www.helsebiblioteket.no/ Retningslinjer)
UK	UK National Screening Committee (www.screening.nhs.uk), National Institute for Health and Clinical Excellence (www.nice.org.uk)
Ireland	Health Service Executive (www.hse.ie)
USA	United States Preventive Services Task Force (www.uspreventiveservicestaskforce.org/), Agency for Healthcare Research and Quality (www.ahrq.gov), National Guideline Clearinghouse (www.guideline.gov)
Canada	Canadian Task Force on Preventive Health Care (www.canadiantaskforce.ca), Public Health Agency of Canada (www.phac-aspc.gc.ca)
Australia	National Health and Medical Research Council (www.nhmrc.gov.au), Clinical Practice Guidelines Portal (www.clinicalguidelines.gov.au)
New Zealand	National Screening Unit (www.nsu.nz) (under the National Health Board), Ministry of Health (www.health.govt.nz/), New Zealand Guidelines Group (www.health.govt.nz/about-ministry/ministry-health-websites/new-zealand-guidelines-group)
Nephrology	
Denmark	Dansk Nefrologisk Selskab (www. nephrology.dk)
Sweden	Svensk Njurmedicinsk Förening (www.njur.se)
Norway	Norsk Nyremedisinsk Forening (www.nephro.no)
UK	The Renal Association (www.renal.org)
Ireland	Irish Nephrology Society (www.nephrology.ie)
USA	American Society of Nephrology (www.asn-online.org)
Canada	Canadian Society of Nephrology (www.csnscn.ca)
Australia	Australian and New Zealand Society of Nephrology (www.nephrology.edu.au)
New Zealand	Australian and New Zealand Society of Nephrology (www.nephrology.edu.au)
Urology	
Denmark	Dansk Urologisk Selskab (www.urologi.dk)
Sweden	Svensk Urologisk Förening (www.urologi.org)
Norway	Norsk Urologisk Forening (www.legeforeningen.no/nuf)
UK	British Association of Urological Surgeons (www.baus.org.uk)
Ireland	Irish Society of Urology (at the website of the Royal College of Surgeons in Ireland, www.rcsi.ie)
USA	American Urological Association (www.auanet.org)
Canada	Canadian Urological Association (www.cua.org)
Australia	Urological Society of Australia and New Zealand (www.usanz.org.au)
New Zealand	Urological Society of Australia and New Zealand (www.usanz.org.au)
Clinical biochemistry	
Denmark	Dansk Selskab for Klinisk Biokemi (www.dskb.dk)
Sweden	Svensk Förening för Klinisk kemi (www.kliniskkemi.org)
Norway	Norsk Forening for Medisinsk Biokjemi (legeforeningen.no/Fagmed/Norsk-forening-for-medisinsk-biokjemi)
UK	Association for Clinical Biochemistry (www.acb.org.uk)
Ireland	Association of Clinical Biochemists in Ireland (www.acbi.ie)
USA	American Association for Clinical Chemistry (www.aacc.org)
Canada	Canadian Society of Clinical Chemists (www.cscc.ca)
Australia	Australasian Association of Clinical Biochemists (www.aacb.asn.au)
New Zealand	Australasian Association of Clinical Biochemists (www.aacb.asn.au)
General internal medicine	
Denmark	Dansk Selskab for Intern Medicin (www.dsim.dk)
Sweden	Svensk Internmedicinsk Förening (www.sim.nu/sv)
Norway	Norsk Indremedisinsk Forening (legeforeningen.no/Fagmed/Norsk-indremedisinsk-forening)
UK	The Royal College of Physicians in London. (www.rcplondon.ac.uk)
Ireland	Irish Association of Internal Medicine (www.internalmedicine.ie)
USA	American College of Physicians (www.acponline.org), Society of General Internal Medicine (www.sgim.org)
Canada	Canadian Society of Internal Medicine (www.csimonline.com)
Australia	Internal Medicine Society of Australia and New Zealand (www.imsanz.org.au)
New Zealand	Internal Medicine Society of Australia and New Zealand (www.imsanz.org.au)
General practice	
Denmark	Dansk Selskab for Almen Medicin (www.dsam.dk)
Sweden	Svensk Förening för Allmänmedicin (www.sfam.se)
Norway	Norsk Forening for Allmennmedisin (www.legeforeningen.no/Fagmed/Norsk-forening-for-allmennmedisin)
UK	Royal College of General Practitioners (www.rcgp.org.uk)
Ireland	Irish College of General Practitioners (www.icgp.ie)
USA	American Academy of Family Physicians (www.aafp.org)
Canada	The College of Family Physicians of Canada (www.cfpc.ca)
Australia	The Royal Australian College of General Practitioners (www.racgp.org.au)
New Zealand	The Royal New Zealand College of General Practitioners (www.rnzcgp.org.nz)
Other	National Kidney Foundation (www.kidney.org/professionals/kdoqi), Kidney Disease: Improving Global Outcomes (www.kdigo.org), International Society of Nephrology (www.theisn.org), European Association of Urology (www.uroweb.org), Caring for Australasians with Renal Impairment (www.cari.org.au), Scottish Intercollegiate Guidelines Network (www.sign.ac.uk). European Renal Association – European Dialysis and Transplant Association (www.european-renal-best-practice.org)

TABLE 2

Summary of relevant identified content.

	Recommendation on dipsticks	Other relevant content
Combined	None	None
Haemoglobin	*The UK National Screening Committee [18]* "Screening for bladder cancer should not be offered" "Screening by urine dipstick testing for protein and blood is not recommended and should no longer take place" *Joint Statement by the Renal Association (UK) and the British Association of Urological Surgeons' joint statement [19]* "Urine testing for haematuria should only be performed for identifiable clinical reasons; there is currently no evidence to support opportunistic screening of the general population"	*The United States Preventive Services Task Force [20]* Concluded that the evidence is insufficient to determine the balance of benefits and harms of screening for bladder cancer in asymptomatic adults. *American Academy of family physicians [21]* "The AAFP concludes that the current evidence is insufficient to assess the balance of benefits and harms of screening for bladder cancer in asymptomatic adults" *American Urological Association [22]* Guideline on management on asymptomatic microscopic haematuria mentions that there is limited evidence behind screening for haematuria, but does not recommend against screening *Danish Urological Society [23]* In a guideline on bladder cancer, it is discussed what should be done if asymptomatic microscopic haematuria is identified. No recommendations about screening *The New Zealand National Screening Unit [24]* In an overview of policy positions, screening for bladder cancer is listed as "No policy"
Leukocytes/nitrite	*European Association of Urology [25]* Recommends that screening for asymptomatic bacteriuria should only be done in pregnant women and before invasive genitourinary procedures *United States Preventive Services Task Force [26]* "The available evidence continues to support screening for asymptomatic bacteriuria in pregnant women, but not in other groups of adults" *American Academy of Family Physicians [27]* "The AAFP recommends against screening for asymptomatic bacteriuria in men and nonpregnant women" *Royal Australian College of General Practitioners [28]* "Identifying and treating non-pregnant adults with asymptomatic bacteriuria does not improve outcomes and may increase antibiotic resistance"	None
Glucose	*Danish Health and Medicines Authority [29]* "Examination of the possible use of urine test strips for screening has not been included in this HTA report as it is regarded as an obsolete analysis in this connection"	*UK National Screening Committee [30]* "Policy position: General population screening should not be offered. Whole population screening has been assessed against the UK NSC criteria and does not meet a number of the criteria" "The UK National Screening Committee has identified the need for a Vascular Risk Management Programme, however, which includes diabetes." This refers to the NHS Health Check programme, which does not use dipsticks for glucose. *Joint statement from the Danish Society for Clinical Biochemistry, Danish College of General Practitioners and Danish Endocrinological Society [31]* "The working group recommends an intensified effort in detecting persons with unrecognised diabetes, but does not recommend general screening." No specific mention of dipsticks, but the rejection of general screening must also encompass dipsticks
Protein/albumin	*UK National Screening Committee [32]* "Policy position: A national screening programme for kidney disease is not recommended" "Screening by urine dip stick testing for protein and blood is not recommended and should no longer take place." (Found on website relating to screening for bladder cancer [18]) *Canadian Society of Nephrology [33]* Recommends against mass screening with dipsticks, but recommends screening high-risk groups using ACR or PCR	*The Royal Australian College of General Practitioners [28]* Recommends screening high risk people with BP, ACR and eGFR. "Dipstick urine test is not adequate to identify microalbuminuria" *Kidney Disease: Improving Global Outcomes (KDIGO) [34]* No recommendation, but makes a note that there appears to be no evidence supporting screening unselected populations *Scottish Intercollegiate Guidelines Network [35]* "Dipstick proteinuria (≥ 1+) can be used to identify patients at risk of subsequent end-stage renal disease and cardiovascular disease" "Urine dipstick testing cannot be used reliably in isolation to diagnose the presence or absence of proteinuria" *New Zealand National Screening Advisory Committee [36]* States that current policy is "Opportunistic screening and self-testing using a urine dip-stick" *U.S. Preventive Services Task Force [46]* "Concludes that the evidence is insufficient to assess the balance of benefits and harms of routine screening for chronic kidney disease (CKD) in asymptomatic adults. Mentions urine testing for albuminuria" *American Academy of Family Physicians [47]* "The AAFP concludes that the evidence is insufficient to assess the balance of benefits and harms for routine screening for chronic kidney disease (CKD) in asymptomatic adults. Common tests considered for CKD screening include creatinine-derived estimates of glomerular filtration rate (GFR) and urine testing for albumin"

ACR = albumin-creatinine ratio; BP = blood pressure; eGFR = estimated glomerular filtration rate; HTA = Health Technology Assessment; NHS = National Health Service; NSC = National Screening Committee; PCR = protein-creatinine ratio.

offered guidance on screening of healthy people for asymptomatic bacteriuria. All recommendations went against screening of non-pregnant asymptomatic persons [25-28].

Glucose

The only mention of screening for glucose with urinary dipsticks was in a health technology assessment report which noted that this technique was considered obsolete and would not be included in the report [29]. The UK National Screening Committee and a joint statement from three Danish specialist societies recommended that population screening for diabetes be avoided, without mentioning dipsticks, but both highlighted a need for increased detection of unrecognised diabetes [30, 31].

Protein/albumin

Two organisations unequivocally recommended avoiding screening with dipsticks for protein. One of these was the UK National Screening Committee, but the recommendation was found on the web page relating to screening for bladder cancer [18], while the page about screening for kidney disease did not mention dipsticks [32]. A 2008 guideline from the Canadian Society of Nephrology also recommended against mass screening with dipsticks for protein [33].

Other organisations touched on the subject without giving relevant recommendations. Kidney Disease: Improving Global Outcomes noted that there appears to be no evidence for screening unselected populations with reagent strips [34].

The Scottish Intercollegiate Guidelines Network noted that dipstick testing can be used to identify persons at risk of subsequent end-stage renal disease and cardiovascular disease, but also noted that "urine dipstick testing cannot be used reliably in isolation to diagnose the presence or absence of proteinuria" [35].
A New Zealand public authority gave its policy regarding screening for chronic kidney disease as "opportunistic screening and self-testing using a urinary dipstick" [36].

Several other organisations, including the influential National Kidney Foundation K/DOQI guideline, gave no recommendations for or against general screening, but recommended screening high-risk groups for chronic kidney disease, with varying definitions of what constituted high risk [37-42]. The recommended tests were typically measurement of the albumin-creatinine ratio (ACR) or an albumin-specific dipstick in combination with the estimated glomerular filtration rate. The topic of ACR dipsticks was mentioned by the National Institute of Health and Clinical Excellence [37], stating that dipsticks should only be used if they are capable of measuring albumin at low concentrations and of expressing the results as an ACR.

Admission to hospital

No recommendations were found on any kind of routine dipstick screening on admission to hospital.

DISCUSSION

Recommendations on the use of urinary dipsticks for screening purposes were scarce and often unclear. Despite a thorough search of websites from health authorities and medical societies in nine countries, no recommendations were found on the use of combined dipsticks in health checks or at admission to hospital.

Only one clear statement was found on screening for microscopic haematuria with dipsticks, recommending against their use. Surprisingly, only one urological society gave clear guidance on screening for microscopic haematuria, recommending against, but did not mention dipsticks. Other organisations discussed the topic without giving recommendations. The scarcity of clear guidance may be related to the fact that the literature seems to be in a stalemate, with some observational studies hinting at a possibly important beneficial effect [8, 11], but with no trials to confirm or refute this.

No clear recommendations were found on screening for urinary glucose with dipsticks, but, as was stated in one health technology assessment report, this technique is considered obsolete. It is likely that some experts consider it self-evident that it should not be used, but it is unlikely that all practitioners – including nurses who perform the tests in hospitals – know this.

Regarding screening for bacteriuria, only four recommendations were found, and they all clearly discouraged this practice, except in pregnant women. However, none of the recommendations specifically mentioned dipsticks as the screening method.

Screening for chronic kidney disease was frequently mentioned, and some organisations discussed limitations of dipstick testing for protein, but clear recommendations were scant. As with glucose, it is possible that some experts simply consider dipstick screening for proteinuria an obsolete technique not worth recommending against in guidelines. Assessing the albumin-creatinine ratio in high-risk persons was often recommended, but although this is a better measure than proteinuria, and although a high-risk only strategy likely reduces over-diagnosis and overtreatment, it is still not clear whether screening is beneficial or not. Albumin-creatinine ratio and dipstick proteinuria are predictors for total and cardiovascular mortality [43], but ACR only adds minimally to traditional cardiovascular risk prediction methods [44]. Treatment with angiotensin-converting enzyme inhibitors appears to reduce end-stage renal

disease in persons with chronic kidney disease, macroalbuminuria and diabetes [7], but has not been proven effective for non-diabetic chronic kidney disease stage 1-3, which constitute the majority of cases [45]. Screening trials have not been conducted and information on the harms of diagnosis, treatment, and follow-up is scarce [7].

The comprehensive and systematic search used in this study far exceeds what can be expected from a clinician looking for guidance. However, it is possible that some guidance has been overlooked or misinterpreted. The language limitations and the selection of certain medical fields probably reduced the number of recommendations found. Also, the choice of not searching regional and local authorities may mean that some guidance has been missed. However, such guidance, if it exists, will not necessarily reflect any national or international consensus. Four hospitals were contacted and none of them had any policy on the topic.

The combined dipsticks in common use in health checks and at admission to hospital have a potential to do harm, as do all medical interventions. Even when used for non-screening purposes, they give redundant information that may initiate a diagnostic cascade, and from this viewpoint their existence can be questioned. Using them for screening purposes without solid knowledge from randomised trials that the benefits exceed the harms is unethical, and guidance from authorities and specialist societies should reflect this. There is a need for clear and pragmatic "Do not use" lists regarding tests, helping practitioners avoid subjecting their patients to possibly useless and potentially harmful tests.

CORRESPONDENCE: *Lasse T. Krogsbøll,* The Nordic Cochrane Centre, Rigshospitalet, 2100 Copenhagen Ø, Denmark. E-mail: ltk@cochrane.dk.

ACCEPTED: 3 December 2013.

CONFLICTS OF INTEREST: Disclosure forms provided by the authors are available with the full text of this article at www.danmedj.dk.

ACKNOWLEDGEMENTS: The project arose from discussions with *Peter C. Gøtzsche* and *Karsten Juhl Jørgensen*, who also commented on the protocol and the manuscript.

LITERATURE

1. Merenstein D, Daumit GL, Powe NR. Use and costs of non-recommended tests during routine preventive health exams. Am J Prev Med 2006;30:521-7.
2. Prochazka AV, Lundahl K, Pearson W et al. Support of evidence-based guidelines for the annual physical examination: Division of primary care providers. Arch Intern Med 2005;165:1347-52.
3. Grønhøj Larsen C, Juhl Jørgensen K, Gøtzsche PC. Regular health checks: cross sectional survey. PloS One 2012;7(3):e33694.
4. Powe NR, Boulware LE. Population-based screening for CKD. Am J Kidney Dis 2009;53(3 suppl 3):S64-S70.
5. Hallan SI, Stevens P. Screening for chronic kidney disease: which strategy? J Nephrol 2010;23:147-55.
6. Boulware LE, Jaar BG, Tarver-Carr ME et al. Screening for proteinuria in US adults: a cost-effectiveness analysis. JAMA 2003;290:3101-14.
7. Fink HA, Ishani A, Taylor BC et al. Screening for, monitoring, and treatment of chronic kidney disease stages 1 to 3: a systematic review for the U.S. Preventive Services Task Force and for an American College of Physicians Clinical Practice Guideline. Ann Intern Med 2012;156:570-81.
8. Yamagata K, Iseki K, Nitta K et al. Chronic kidney disease perspectives in Japan and the importance of urinalysis screening. Clin Exp Nephrol 2008;12:1-8.
9. Malmström PU. Time to abandon testing for microscopic haematuria in adults? BMJ 2003;326:813-5.
10. Chou R, Dana T. Screening adults for bladder cancer : a review of the evidence for the U.S. preventive services task force. Ann Intern Med 2010;153:461-8.
11. Messing EM, Madeb R, Young T et al. Long-term outcome of hematuria home screening for bladder cancer in men. Cancer 2006;107:2173-9.
12. Brown RS. Has the time come to include urine dipstick testing in screening asymptomatic young adults? JAMA 2011;306:764-5.
13. Chou R, Croswell JM, Dana T et al. Screening for prostate cancer: a review of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2011;155:762-71.
14. Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2011;(1):CD001877.
15. 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;380:1741-8.
16. Schilling FH, Spix C, Berthold F et al. Neuroblastoma screening at one year of age. N Engl J Med 2002;346:1047-53.
17. Welch HG, Schwartz L, Woloshin S. Overdiagnosed – making people sick in the pursuit of health. Boston: Beacon Press, 2011.
18. United Kingdom National Screening Committee. The UK NSC policy on Bladder Cancer screening in adults. www.screening.nhs.uk/bladdercancer (8 May 2013).
19. Joint consensus statement on the initial assessment of haematuria prepared on behalf of the Renal Association and British Association of Urological Surgeons. www.renal.org/Libraries/Other_Guidlines/Haematuria_-_RA-BAUS_consensus_guideline_2008.sflb.ashx (8 May 2013).
20. U.S. Preventive Services Task Force. Screening for bladder cancer in adults, topic page. August 2011. www.uspreventiveservicestaskforce.org/uspstf/uspsblad.htm (8 May 2013).
21. American Academy of Family Physicians. www.aafp.org/online/en/home/clinical/exam/bladdercancer.html (8 May 2013).
22. Davis R, Jones JS, Barocas DA et al. Diagnosis, evaluation and follow-up of asymptomatic microhematuria (AMH) in adults: AUA guideline. J Urol 2012;188(6 suppl):2473-81. www.auanet.org/education/asymptomatic-microhematuria.cfm (8 May 2013).
23. Nationale kliniske retningsliner for behandling af blæretumorer i Danmark. Dansk BlæreCancer Gruppe, DaBlaCa. www.skejby.net/Webudgaven/DaBlaCa2010.htm (8 May 2013).
24. National Screening Advisory Committee. Screening policy positions and practice in New Zealand. www.nsu.govt.nz/files/Strategic_Screening/Screening_Policy_Positions_Table.pdf (8 May 8 2013).
25. Grabe M, Bjerklund-Johansen TE, Botto H et al. Guidelines on urological infections. European Association of Urology 2012. www.uroweb.org/gls/pdf/17_Urological%20infections_LR%20II.pdf (8 May 2013).
26. Lin K, Fajardo K. Screening for asymptomatic bacteriuria in adults: evidence for the U.S. Preventive Services Task Force reaffirmation recommendation statement. Ann Intern Med 2008;149(1):W20-W24.
27. American Academy of Family Physicians. Bacteriuria, asymptomatic. www.aafp.org/online/en/home/clinical/exam/bacteriuriaasmptomatic.html (8 May 2013).
28. Royal Australian College of General Practitioners. Guidelines for preventive activities in general practice, 8th edition. www.racgp.org.au/your-practice/guidelines/redbook/ (8 May 2013).
29. Type 2 diabetes. Health technology assessment of screening, diagnosis and treatment. Danish Health Technology Assessment 2005;7(1). www.sst.dk/publ/publ2005/CEMTV/diabetes_uk/diabetes_uk.pdf (7 May 2013).
30. The UK NSC policy on diabetes screening in adults. www.screening.nhs.uk/diabetes (8 May, 2013).
31. Borch-Johnsen K, Beck-Nielsen H, Christiansen JS et al. Guidelines on diagnosis of type 2 diabetes. Ugeskr Læger 2003;165:1558-61.
32. The UK NSC policy on kidney disease screening in adults. www.screening.nhs.uk/kidneydisease (8 May 2013).
33. Levin A, Hemmelgarn B, Culleton B et al. Guidelines for the management of chronic kidney disease. CMAJ 2008;179:1154-62.
34. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management

of chronic kidney disease. Kidney Int Suppl 2013; 3: 1-150. www.kdigo.org/clinical_practice_guidelines/ckd.php (8 May 2013).
35. Diagnosis and management of chronic kidney disease. SIGN guideline 103, 2008. www.sign.ac.uk/guidelines/fulltext/103/ (8 May 2013).
36. National Screening Advisory Committee. Screening policy positions and practice in New Zealand. www.nsu.govt.nz/files/Strategic_Screening/Screening_Policy_Positions_Table.pdf (8 May 2013).
37. National Institute for Health and Clinical Excellence. NICE clinical guideline 73. Chronic kidney disease: early identification and management of chronic kidney disease in adults in primary and secondary care. www.nice.org.uk/CG73 (8 May 2013).
38. Tidig upptäckt och initial utredning av kronisk njursjukdom – Riktlinjer för Södra Sjukvårdsregionen. www.njur.se/Filer/Behandlingsriktlinjer/Riktlinjer_for_tidig_upptackt_och_initial_utredning.pdf (8 May 2013).
39. The Renal Association. Clinical practice guidelines for the detection, monitoring and care of patients with chronic kidney disease. www.renal.org/Clinical/GuidelinesSection/Detection-Monitoring-and-Care-of-Patients-with-CKD.aspx#downloads (8 May 2013).
40. Caring for Australasians with Renal Impairment. Screening for early chronic kidney disease - prepublication draft. www.cari.org.au/CKD_early_CKD/Screening_of_CKD.pdf (8 May, 2013).
41. National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification and stratification. Am J Kidney Dis 2002;39:S1-S266. www.kidney.org/professionals/KDOQI/guidelines_ckd/toc.htm (8 May 2013).
42. Kidney Health Australia. Chronic Kidney Disease (CKD) Management in general practice (2nd edition). www.kidney.org.au//LinkClick.aspx?fileticket=vfDcA4sEUMs%3d&tabid=635&mid=1584 (8 May 2013).
43. Chronic Kidney Disease Prognosis Consortium. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. Lancet 2010;375:2073-81.
44. Chang A, Kramer H. Should eGFR and albuminuria be added to the Framingham risk score? Chronic kidney disease and cardiovascular disease risk prediction. Nephron Clin Pract 2011;119(2):c171-c177.
45. Sharma P, Blackburn RC, Parke CL et al. Angiotensin-converting enzyme inhibitors and angiotensin receptor blockers for adults with early (stage 1 to 3) non-diabetic chronic kidney disease. Cochrane Database Syst Rev 2011;(10):CD007751.
46. Screening for chronic kidney disease: clinical summary of U.S. Preventive Services Task Force recommendation. AHRQ Publication No. 12-05166-EF-3, August 2012. www.uspreventiveservicestaskforce.org/uspstf12/kidney/ckdsumm.htm (8 May 2013).
47. American Academy of Family Physicians. Screening chronic kidney disease. www.aafp.org/online/en/home/clinical/exam/kidneydisease.html (8 May 2013).

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	separate abstract file
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	2
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	2
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	2 (mentioned that several things were pre-specified)
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	2
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	2-3
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	NA
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	3
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	4
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	NA
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	NA
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	NA
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	NA

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	NA
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	NA
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	Figure
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	NA (but table 1 shows

			information on organisations)
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	NA
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	4-6
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	NA
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	NA
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	7-8
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	8
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	9
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	Abstract

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Screening with urinary dipsticks for reducing morbidity and mortality - Cochrane review

Lasse T Krogsbøll, Karsten Juhl Jørgensen, Peter C Gøtzsche

The Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark

Accepted for publication in the Cochrane Library.

Abstract

Background

Urinary dipsticks are sometimes used for screening asymptomatic people, and for case-finding among inpatients or outpatients without genito-urinary symptoms. Abnormalities identified on screening sometimes lead to additional investigations, which may identify serious disease, such as bladder cancer and chronic kidney disease (CKD). Urinary dipstick screening could improve prognoses due to earlier detection, but could also lead to unnecessary potentially invasive follow-up testing and unnecessary treatment.

Objectives

We aimed to quantify the benefits and harms of screening with urinary dipsticks in general populations and patients in hospitals.

Search methods

We searched the Cochrane Renal Group's Specialised Register to 8 September 2014 through contact with the Trials Search Co-ordinator using search terms relevant to this review.

Selection criteria

Randomised controlled trials (RCTs) and other study types that compared urinary dipstick screening with no dipstick screening were eligible for inclusion. We searched for studies that investigated the use of urinary dipsticks for detecting haemoglobin, protein, albumin, albumin-creatinine ratio, leukocytes, nitrite, or glucose, alone or in any combination, and in any setting. We planned to exclude studies conducted in patients with urinary disorders.

Data collection and analysis

It was planned that two reviewers would independently extract data from included studies and assess risk of bias using the Cochrane risk of bias tool. However, no studies met our inclusion criteria.

Main results

Literature searches to 8 September 2014 yielded 4298 records, of which 4249 were excluded following title and abstract assessment. There were 49 records (44 studies) eligible for full text assessment; of these 18 studies were not RCTs and 26 studies compared interventions or controls that were irrelevant to this review. Thus, no studies were eligible for inclusion in this review.

Authors' conclusions

We found no evidence to assess the benefits and harms of screening with urinary dipsticks, which remain unknown.

Plain language summary

Screening with urinary dipsticks for reducing morbidity and mortality.

Urinary dipsticks are sometimes used for screening healthy people and patients that do not have symptoms of urinary disease. Urinary dipsticks can be used to test for several different substances, such as blood, sugar, protein, white blood cells and nitrite in the urine, which may indicate the presence of disease. Identified abnormalities sometimes lead to additional investigations, which may identify serious disease, such as bladder cancer and chronic kidney disease. Detection could

improve health outcomes from finding disease at earlier stages, but could also lead to unnecessary follow-up testing, which may be invasive, and lead to unnecessary treatment.

We searched the literature to September 2014 to identify studies that compared urinary dipstick screening with no dipstick screening. However, we found no studies that met our inclusion criteria. We were unable to determine benefits and harms associated with urinary dipstick screening.

Background

Urinary dipstick testing is widely used to screen for the presence of disease with the aim of reducing morbidity and mortality in both healthy people and patients (Grønhøj Larsen 2012; Merenstein 2006; Prochazka 2005). Dipsticks can test for either single or multiple substances in urine, and are sometimes used in general health checks.

Urinary dipstick testing is recommended for screening pregnant women to detect bacteria in the urine (bacteriuria) (Lin 2008; NICE 2008a), and for people with diabetes to detect a specific protein (albumin) in the urine (albuminuria) (NICE 2008b). Another potential screening population is people with hypertension. However, there is a lack of consensus on these recommendations, and most guidelines recommend use of albumin-creatinine or protein-creatinine ratios rather than dipsticks to detect proteinuria or albuminuria. However, dipsticks are less expensive than these tests and dipstick proteinuria is strongly related to total and cardiovascular mortality (CKDPC 2010)

Since the 1970s, school children and employed adults in Japan have been offered urinary dipstick screening for blood and protein; from 1983, this was extended to all adults aged 40 years and over (Imai 2007). Taiwan implemented dipstick screening for children in 1990, and Korea in 1998 (Hogg 2009).

There appear to be no recommendations for population-based screening with urinary dipsticks, and the scientific debate persists about screening for CKD with other methods (Brown 2011). Opportunistic screening is often recommended, but only for high risk groups (Krogsbøll 2014). A systematic review of screening for CKD with any method found no RCTs and concluded that the role of screening was uncertain (Fink 2012).

There is, however, discrepancy between recommendations and practice. Ease of use, low cost, and the perceived test safety may contribute to this discrepancy. Although screening can work,(Holme 2013; Raffle 2003; Thomason 1998) it is noteworthy that experience with other screening interventions for diseases such as prostate cancer (Djulbegovic 2010), breast cancer (Gøtzsche 2013), and neuroblastoma (Schilling 2002; Woods 2002) have indicated that screening benefits can be fewer than expected, and that screening can cause more harm than good.

Dipstick testing is routinely used for case-finding among persons with a condition that increases the risk of nephropathy, such as diabetes and hypertension. Both are conditions with a wide a spectrum of severity, do often not cause symptoms, and encompass a large proportion of adults. The definitions of these conditions are made through consensus, and have been the subject of debate as they have been broadened over the decades. Thus, case-finding in such broad categories borders on screening, but randomised trials are unlikely to be performed and as the question must be informed by detailed analysis of observational studies, which is outside the scope of this review.

Description of the condition

Microscopic blood in urine (haematuria) can be caused by urological cancers of any kind, but because bladder cancer is relatively common, and haematuria a frequent sign, most research has centred on this. The prognosis for people with bladder cancer is highly dependent on the extent of invasion into the bladder wall; non-muscle-invasive lesions often have a favourable prognosis following minimally invasive treatment, in contrast to muscle-invasive lesions. However, unexpected post-mortem findings are less common than for other urological cancers, such as prostate and kidney cancers (Avgerinos 2001; Karwinski 1990), which suggests that bladder cancer

may have a short preclinical phase, possibly rendering it a poor target for screening. Furthermore, microscopic haematuria is not a robust marker for bladder cancer because haematuria can be associated with a plethora of benign conditions (Malmström 2003), and novel markers have not yet been tested sufficiently.

CKD is a major health problem with a long preclinical phase. Staging is based on estimated glomerular filtration rate (eGFR) and evidence of kidney damage, such as proteinuria or pathological findings with ultrasound imaging (NKF 2002). When this staging formula was applied to the adult population in the United States, CKD prevalence was found to be 13%, and more than 45% among people over 70 years of age (Coresh 2007). Although most people with CKD do not go on to develop end-stage kidney disease (ESKD), its prevalence is increasing (Hemmelgarn 2006). It has been argued that the current staging system is inappropriate because many people with CKD have low eGFR but no evidence of kidney damage (Moynihan 2013). Low eGFR could be considered normal, particularly among older people, most of whom are unlikely to develop symptomatic kidney disease (Bauer 2008).

It has been reported that proteinuria detected using urinary dipsticks was associated with subsequent ESKD (Iseki 2003) and the test can identify some people who are at risk of rapid decline in kidney function (Clark 2011). It has also been reported that asymptomatic microscopic haematuria is associated with ESKD (Iseki 2003; Vivante 2011). Both low eGFR and proteinuria are risk factors for cardiovascular and all-cause mortality (Hillege 2001; Matsushita 2010), although they do not seem to substantially improve traditional prediction tools (Chang 2011).

Diabetes mellitus can cause glycosuria, and early detection may prevent or postpone complications such as blindness, neuropathy or cardiovascular disease through early treatment and weight loss. A trial of screening for diabetes using other methods than dipsticks did not find beneficial effects (Simmons 2012).

Asymptomatic bacteriuria may be detected with dipstick testing for leukocytes and nitrite, but detection in urine are common findings, particularly among older people, and treatment is not recommended in the absence of symptoms. On the other hand, urinary tract infection can present with vague and uncharacteristic symptoms. Screening for asymptomatic bacteriuria is recommended only for pregnant women and before genitourinary procedures (EAU bacteriuria 2012; Lin 2008).

How the intervention might work

Although many variations exist, the urinary dipsticks commonly used in general health checks usually test for at least haemoglobin, protein or albumin, leukocytes, nitrite and glucose. In screening, the use of a combined urinary dipstick is in some ways comparable with a general health check, which also includes components with different potentials for benefits and harms for a range of very different diseases. Likewise, there is a wide range of relatively harmless conditions that can result in an abnormal test.

Benefits

The potential benefits from dipstick screening are well known. Many diseases screened for using urinary dipsticks are both common and serious. Early diagnosis of diabetes mellitus and appropriate interventions and lifestyle changes may reduce common comorbidities such as blindness, neuropathy, nephropathy,or cardiovascular disease. Identification of chronic kidney disease may allow early therapy to reduce CKD morbidity and mortality, although a Cochrane review concluded that the value of treating CKD stages 1 to 3 with angiotensin-converting enzyme inhibitors remains unclear (Sharma 2011). Glomerulonephritis often responds to treatment but it is has not been shown whether extent treatment of subclinical glomerulonephritis improves prognosis. If detection of microscopic haematuria enables earlier detection of bladder cancer, morbidity, mortality and harmful effects of invasive treatments for advanced disease may be reduced.

Harms

Harms from dipstick screening mainly relate to superfluous follow-up tests and therapeutic interventions, and not the screening itself. Harms include discomfort and anxiety related to non-invasive follow-up tests such as kidney ultrasound, and from concerns about possible health issues, but most importantly, the possibility of morbidity related to unnecessary invasive investigations.

Investigations for persistent microscopic haematuria often include flexible cystoscopy in local anaesthesia, computed tomography imaging (CT scan) of the urinary tract, and urine cytology. In some instances, rigid cystoscopy and biopsy under general anaesthesia may be required, which carries a risk of complications such as bladder perforation, bleeding, and infection. The initial

nephrological work-up of patients with proteinuria or microscopic haematuria may suggest the need for kidney biopsy, which carries a risk of serious complications such as haemorrhage.

Imaging of the abdomen may reveal unexpected abnormalities, which can lead to further investigations (Furtado 2005). A study of CT colonography reported that the prevalence of incidental abnormalities was very high, around 40%, which led to additional investigations in 14% of cases (Xiong 2005). Even when serious abnormalities such as cancer are found incidentally, there is no guarantee that this improves prognosis (Welch 2004).

Common to all conditions that may be detected using urinary dipstick testing is a risk that the identified condition would never have caused symptoms in the person's remaining lifetime (over-diagnosis), and that the diagnosis therefore will not improve prognosis, but instead lead to unnecessary worry and over-treatment with inherent harms. Over-diagnosis and over-treatment are documented in screening for breast cancer (Gøtzsche 2013; Jørgensen 2009), prostate cancer (Djulbegovic 2010), lung cancer, melanoma and thyroid cancer (Welch 2010). Although the concepts of over-diagnosis and over-treatment are most familiar in cancer screening, they also apply to screening for other conditions such as hypertension, hypercholesterolaemia, diabetes (Welch 2011), and CKD (Moynihan 2013).

The possibility of adverse psychological effects associated with diagnostic tests and treatment must also be considered as a potential harm, as well as the impact of negative screening results on providing a false sense of security, with the possibility of some people ignoring important symptoms.

Why it is important to do this review

The possible benefits of urinary dipstick screening must be weighed against possible harms. The main question is whether screening reduces morbidity and mortality and if the harms are acceptable.

Our review investigated the use of urinary dipsticks to screen healthy people and hospital in- or outpatients for the presence of disease. Our main interest was the effects of combined dipsticks use, but we anticipated that the existing literature would be scant and therefore also planned to assess RCTs of screening for individual components of dipsticks, such as blood or protein.

This review focused on clinical outcomes that are relevant to people, such as mortality and ESKD.

Objectives

We aimed to quantify the benefits and harms of screening with urinary dipsticks in general populations and in patients at hospitals.

Methods

Criteria for considering studies for this review

Types of studies

All RCTs and other studies in which allocation to screening with urinary dipsticks or no screening was obtained using alternation (e.g. alternate medical records), date of birth or similar methods were eligible for inclusion.

Types of participants

Inclusion criteria

We did not impose age limitations and included studies from both general and patient populations. We included studies of screening using urinary dipsticks performed as part of a health check, such as in general practice or at the community level, as well as studies of screening hospital in- or outpatients, and patients in non-hospital specialist clinics.

Exclusion criteria

We excluded studies where dipstick testing was done on indication, e.g. in people with suspected urinary tract infection, as well as studies conducted exclusively in populations of patients with urinary diseases because the pretest probability of disease would be high. Dipstick testing in these situations could also be viewed more properly as diagnostic testing rather than screening.

Types of interventions

We included studies of single or repeat use of urinary dipstick screening that tested for one or more of the following: haemoglobin, protein, albumin, albumin-creatinine ratio, glucose, leukocytes and nitrite. We included studies regardless of who performed the test, such as healthcare professionals or study participants (following instruction).

Types of outcome measures

Primary outcomes

- All-cause mortality

- Cardiovascular mortality
- Cancer mortality
- ESKD (patients requiring renal replacement therapy, i.e. dialysis or kidney transplantation).

Secondary outcomes

- Admission to hospital
- Drug therapy
- Surgery
- New diagnoses (cancers, including cancer stages, urolithiasis, CKD (stages 1 to 3), CKD (stages 4 and 5), diabetes mellitus, bacteriuria)
- Follow-up investigations resulting from a positive test
- Complications to follow-up investigations
- Self-reported health
- Quality of life
- Disability.

Search methods for identification of studies

Electronic searches

We searched the Cochrane Renal Group's Specialised Register to 8 September 2014 through contact with the Trials Search Co-ordinator using search terms relevant to this review.

The Cochrane Renal Group’s Specialised Register contains studies identified from sources.

Quarterly searches of the Cochrane Central Register of Controlled Trials CENTRALWeekly searches of MEDLINE OVID SPHandsearching of renal-related journals and the proceedings of major renal conferencesSearching of the current year of EMBASE OVID SPWeekly current awareness alerts for selected renal journalsSearches of the International Clinical Trials Register (ICTRP) Search Portal and ClinicalTrials.gov. Studies contained in the specialised register are identified through search strategies for CENTRAL, MEDLINE, and EMBASE based on the scope of the Cochrane Renal Group. Details of these strategies as well as a list of handsearched journals, conference proceedings and current awareness alerts are available in the Specialised Register section of information about the Cochrane Renal Group.

Searching other resources

Reference lists of review articles, relevant studies and clinical practice guideline

Data collection and analysis

Selection of studies

Titles and abstracts were screened independently by two authors who discarded studies that were clearly not eligible and assessed the full text of potentially eligible studies to determine which satisfied inclusion criteria. Disagreements were to be resolved through discussion, with the third author as arbiter when necessary.

Data extraction and management

Data extraction was to be carried out independently by two authors using standard data extraction forms. Studies reported in non-English language journals were to be translated before assessment. Where more than one publication of one study existed, reports were to be grouped together and the

publication with the most complete data used in the analyses. Where relevant outcomes were only published in earlier versions these data were to be used. Any discrepancy between published versions was to be highlighted.

Assessment of risk of bias in included studies

We planned to assess risk of bias using the Cochrane risk of bias assessment tool (Higgins 2011).

Measures of treatment effect

For dichotomous outcomes (mortality, ESKD), we planned to use the risk ratio (RR) with 95% confidence intervals (CI). For measurement scale outcomes (self-reported health, quality of life, disability), we planned to use the mean difference (MD), or the standardised mean difference (SMD) if different scales were used. Some outcomes may have been reported in various ways (admission to hospital, drug therapy, surgery, new diagnoses, follow-up investigations, complications to follow-up investigations), such as rates, continuous, or dichotomous outcomes. We planned to choose the format that would have informed the best synthesis of available results.

For measurement scale outcomes, we planned to extract both change from baseline and final means when available. Missing standard deviations were planned to have been estimated from similar studies, when possible. Time-to-event data were planned to be treated as dichotomous data, because the relevant outcomes (mortality, ESKD) were likely to have been ascertained for all participants.

Unit of analysis issues

We planned to include cluster RCTs, and when possible, extract effect measures and standard error rates from an analysis that takes clustering into account. If that was not possible, we planned to

extract the number of clusters and estimate the intra-cluster correlation coefficient to inform a reliable analysis. If this was not possible, we planned to disregard the clustering and investigate the effect of this in a sensitivity analysis.

Dealing with missing data

We planned to extract data for intention-to-treat analyses (ITT) and contact authors if required information was missing. Where ITT analysis was not possible, we planned to extract data from an available case analysis and assess the risk of bias from attrition.

Assessment of heterogeneity

We planned to analyse heterogeneity using a Chi^2 test on N-1 degrees of freedom, with an alpha of 0.05 used for statistical significance, and the I^2 statistic (Higgins 2003).

Assessment of reporting biases

We did not expect that a sufficient number of studies would be identified to create a useful funnel plot. Assessing reporting bias is difficult, but we planned to note whether outcomes that we considered important were reported. We planned to contact authors about possible unpublished outcomes.

Data synthesis

We planned to use a random-effects model and to express the results as both relative risks and number-needed-to-screen to achieve the relevant outcomes, both beneficial and harmful.

Subgroup analysis and investigation of heterogeneity

We planned to perform the following subgroup analyses:

- risk of bias
- substances tested for (e.g. haemoglobin, protein/albumin or albumin-creatinine ratio, glycosuria, leukocytes/nitrite, or combinations of substances)
- population type (general populations, pregnant women, patients)
- age of participants.

Sensitivity analysis

If possible, we planned to perform sensitivity analyses to explore the influence of the following factors on effect size:

- excluding cluster RCTs
- repeating the analysis excluding unpublished studies
- repeating the analysis excluding any very long or large studies to establish how much they dominate the results.

Results

Searches yielded 4298 records, of which 4249 were excluded based on title and abstract (Figure 1). We identified 49 records (44 studies) for possible inclusion and full-text assessment. These were either not RCTs (n = 18) or compared interventions or controls that were not relevant to this review (n = 31) (Characteristics of excluded studies). Thus, no studies could be included in this review.

Risk of bias in included studies

Risk of bias assessment could not be conducted.

Effects of interventions

No studies met the inclusion criteria.

Discussion

Summary of main results

We found no studies that compared screening with urinary dipsticks with no screening. Screening with urinary dipsticks for haemoglobin, protein, albumin, albumin-creatinine ratio, leukocytes, nitrite, or glucose, alone or in any combination, has unknown benefits and harms.

Agreements and disagreements with other studies or reviews

General populations

The older observational literature is mainly concerned with assessing the diagnostic yield, or exploring the feasibility and cost of screening programs, tacitly implying that any discovery of asymptomatic illness is beneficial. Given knowledge about overdiagnosis and overtreatment associated with several types of screening tests (Black 2010; Independent UK Panel 2012; Welch 2011a) and their sometimes disappointing benefits (Djulbegovic 2010; Gøtzsche 2013; Krogsbøll 2012; Schilling 2002; Simmons 2012; Woods 2002) such an assumption is not warranted.

Some studies avoided this assumption, but used methods that did not enable reliable conclusions to be made about benefits and harms. Japanese urine screening programs for children and adults that used dipstick testing for haemoglobin and protein were implemented in the 1970s . An analysis of incidence rates of ESKD in Japan, using data from a nationwide dialysis registry from 1983 to 2000, found steadily increasing incidence rates during the entire period (Wakai 2004). This rise was

mainly due to diabetic nephropathy, nephrosclerosis, and unknown causes, while ESKD due to glomerulonephritis rose until the mid 1990s, where it started to decline. This observation is compatible with the hypothesis that screening caused the decline, given the expected latency of effect, but other explanations are also possible, such as a decrease in incidence of glomerulonephritis or the implementation of possibly useful treatments for glomerulonephritis (Reid 2011; Samuels 2003). We have not found studies that compared the incidence of glomerulonephritis before and after the introduction of the Japanese screening programs in the 1970s, and an increase in incidence caused by detection of asymptomatic cases is also possible, as glomerulonephritis, particularly immunoglobulin A (IgA) nephropathy, can remain subclinical.

Comparisons between countries are difficult to interpret because of variations in biopsy policies, and a systematic review of glomerulonephritis incidence found very large variations (McGrogan 2011). For example, five studies reported the incidence of IgA nephropathy in children: four non-screening studies and one screening study. In the non-screening studies, the incidence ranged between 0.03/100,000/year and 0.57/100,000/year and the screening study found 4.5/100,000/year.

The prognosis of children with screening-detected glomerular disease appears to be good (Ito 1990), and better than for symptomatic cases (Takebayashi 1992). This could be due to effective treatments arresting or slowing the disease at an early stage, but it could also reflect over diagnosis of subclinical cases, that is, cases that would not have become symptomatic if not discovered by screening, or length bias (screening preferentially detects less aggressive disease as there is more time to detect these cases).

Messing 2006 compared long-term outcomes of bladder cancer detected through screening with outcomes of clinically detected bladder cancer and found dramatic differences in mortality between men with screen-detected cancers and men with cancers not detected by screening. However, the populations were probably not comparable because the risk of death from other causes than bladder cancer was smaller among the men with screen detected cancers, which suggests selection bias. Furthermore, the possibility of over-diagnosis of less aggressive tumours has not been ruled out, and this would also confer a spurious survival advantage to the screened group.

The observations that proteinuria and eGFR are clearly and consistently associated with the the risk of ESKD, myocardial infarction, acute kidney injury, and death suggest that screening could be beneficial.(Hemmelgarn 2010; James 2010) However, such observations do not resolve the classic screening-related questions, e.g. whether the efficacy of treating screening-detected disease is similar to what is observed in trials of disease not detected in screening, whether the compliance with both screening and preventive treatments is adequate in asymptomatic persons, how much overdiagnosis the screening causes, and whether the benefits outweigh the harms. A simulation study of screening for proteinuria found that it was cost-effective when targeted to people with hypertension, those aged over 60 years, or when conducted at the infrequent interval of 10 years (Boulware 2003). However, many assumptions are needed for simulation studies, and they cannot constitute proof.

A review of general health checks (Krogsbøll 2012) included five studies (19,813 participants) that contained screening with a urinary dipstick as part of the intervention (Engberg 2002; Friedman 1986; Lannerstad 1977; Olsen 1976; Tibblin 1982). These studies did not find beneficial effects on morbidity or mortality. Friedman 1986 (which included 10,674 participants and 16 years of follow-

up) reported cause-specific mortality in detail and did not find effects on deaths from genitourinary disease, or in other cancers, which included cancers of the bladder, kidney and ureter. The studies were likely underpowered to detect small beneficial effects of dipstick screening.

Hospitalised patients and clinic patients

A cohort study of dipstick testing in medical outpatients without relevant symptoms, found that 17% had an abnormal result, but that management was changed as a result in only 0.7%. (Rüttimann 1994) Three older cohorts assessed routine urine microscopy and similarly found many abnormal test results but few consequences for management (Boland 1995, Boland 1996, Kroenke 1986).

High risk patients

A special issue is case-finding in people with known conditions that are strongly associated with nephropathy. Although we found no trial evidence regarding this, the practice may be justified when the association is very strong, such as with diabetes mellitus. Also, this practice is so ubiquitus that trials are unlikely to be conducted. Exactly what level of risk justifies case-finding in high-risk groups is not clear and is not the topic of this review.

Authors' conclusions

Implications for practice

We found no trials that investigated dipstick screening versus no dipstick screening, and therefore the benefits and harms remain unknown. Because there are potential harms related to dipstick screening, and since any screening programme entails financial and opportunity costs, urinary dipsticks should not be used for screening purposes (i.e. without an indication) in non-pregnant persons, except in the study setting. This conclusion does not address dipstick testing done on clinical indications such as fever, or in very high risk patients such as people with diabetes.

Implications for research

Conduct of RCTs are feasible for routine screening of healthy people, hospital inpatients and outpatients because the intervention is widespread but not standard of care. Careful consideration to ensure that studies are adequately powered is needed for future studies.

Acknowledgements

We would like to thank:

- the Cochrane Renal Group Trials Search Coordinator, Ruth Mitchell, for help with the search strategy
- the referees for their comments and feedback during the preparation of this review.

Contributions of authors

Drafting the protocol: LTK, KJ, PCG

Study selection: LTK, KJ

Extract data from studies: LTK, KJ

Enter data into RevMan: LTK

Carry out the analysis: there was none

Interpret the analysis: there was none

Draft the final review: LTK, KJ, PCG

Disagreement resolution: there was none

Update the review: LTK

Declarations of interest

Lasse T Krogsbøll: None known

Karsten Juhl Jørgensen: None known

Peter C Gøtzsche: None known.

Differences between protocol and review

None

Characteristics of included studies

None

Characteristics of excluded studies

Allen 1991

Reason for exclusion: Wrong intervention. Compared urinary glucose monitoring in diabetes with blood glucose monitoring.

Apoola 2009

Reason for exclusion: Wrong intervention. Compared partner notification with urethral swab and urine antigen testing.

Balogun 2011

Reason for exclusion: Not RCT.

Battelino 2011

Reason for exclusion: Wrong intervention. Compared blood glucose meter with continuous blood glucose monitoring for children with type 1 diabetes.

Beatty 1994

Reason for exclusion: Not RCT.

Bubner 2009

Reason for exclusion: Wrong intervention. Compared point-of-care testing with laboratory testing in general practice.

Calderon-Margalit 2005

Reason for exclusion: Not RCT.

Calero 2011

Reason for exclusion: Not RCT.

Charles 2009

Reason for exclusion: Wrong intervention. Compared intensified treatment of people with screen-detected diabetes with usual care.

Dallosso 2012

Reason for exclusion: Wrong intervention. Compared monitoring with blood glucose or urine testing for people with type 2 diabetes.

Davies 1991

Reason for exclusion: Wrong intervention. No unscreened control group. Compared two different methods of screening for glycosuria.

Davies 1993

Reason for exclusion: Not RCT.

Davies 1999

Reason for exclusion: Not RCT.

Diercks 2002

Reason for exclusion: Wrong intervention. Factorial design that compared fosinopril, pravastatin and placebo in people with elevated urinary albumin excretion.

Dolan 1987

Reason for exclusion: Wrong intervention. Compared urinary glucose monitoring by dipstick with urine glucose monitoring by tablet system.

Downing 2012

Reason for exclusion: Not RCT.

DPPRG 2005

Reason for exclusion: Wrong intervention. Compared lifestyle intervention, metformin, and placebo for prevention of diabetes in people with elevated fasting glucose and impaired glucose tolerance.

Falguera 2010

Reason for exclusion: Wrong intervention. Compared empirical treatment of pneumonia with targeted treatment based on urine antigen testing.

Fulcher 1991

Reason for exclusion: Not RCT.

Gallichan 1994

Reason for exclusion: Wrong intervention. Compared blood glucose monitoring with urine dipstick monitoring in patients with type 2 diabetes.

Goldby 2011

Reason for exclusion: Not RCT.

Grimm 1997

Reason for exclusion: Wrong comparison. Both groups had dipstick testing.

Jolic 2011

Reason for exclusion: Not RCT.

Jou 1998

Reason for exclusion: Not RCT.

Kazemier 2012

Reason for exclusion: Wrong intervention. Compared antibiotic treatment with no treatment in pregnant women with asymptomatic bacteriuria.

Kenealy 2005

Reason for exclusion: Wrong intervention. Compared patient reminders, computer reminders, both reminders, and usual care, for screening for diabetes.

Koschinsky 1984

Reason for exclusion: Wrong intervention. Compared urinary and blood glucose testing in diabetics. Both groups tested for urinary glucose for 4 weeks and then for blood glucose for 4 weeks.

Lauritzen 1994

Reason for exclusion: Wrong intervention. Described variation in albumin-creatinine ratio in the RCT screened arms.

Lauritzen 2008

Reason for exclusion: Wrong intervention. Compared health checks, health checks and lifestyle conversations, and usual care.

Lenz 2002

Reason for exclusion: Wrong comparison. Compared nurse practitioner and physician treatment of diabetes.

Little 2009

Reason for exclusion: Wrong intervention. Compared five different management strategies for suspected urinary tract infection.

McEwan 1990

Reason for exclusion: Wrong intervention. Dipstick screening was part of a complex intervention.

McGhee 1997

Reason for exclusion: Not RCT.

Messing 1995

Reason for exclusion: Not RCT.

Morris 2012

Reason for exclusion: Not RCT. Systematic review of diagnostic studies comparing spot protein-creatinine ratio with 24 hour protein-creatinine ratio for screening pregnant women.

Naimark 2001

Wrong intervention. Tested education directed towards physicians to increase their microalbuminuria testing pattern among people with type 2 diabetes.

Neumann 2008

Reason for exclusion: Wrong intervention. Compared urine microscopy with a malaria dipstick.

Nevedomskaya 2011

Reason for exclusion: Not RCT.

Ochoa 2007

Reason for exclusion: Not RCT.

Reyburn 2007

Reason for exclusion: Wrong intervention. Compared urine microscopy and a rapid diagnostic test for malaria.

Schwab 1992

Reason for exclusion: Not RCT.

Simmons 2012

Reason for exclusion: Wrong intervention. Compared screening for diabetes with no screening for diabetes (no dipsticks).

Tissot 2001

Reason for exclusion: Not RCT.

Worth 1982

Reason for exclusion: Wrong intervention. Compared dipstick testing for glycosuria with dipstick testing for glycosuria combined with one of two methods for blood glucose measurements. Used cross-over design. Included people with known diabetes.

References

Excluded studies

Allen 1991

Allen BT, DeLong ER, Feussner JR. Impact of glucose self-monitoring on non-insulin-treated patients with type II diabetes mellitus. Randomized controlled trial comparing blood and urine testing. Diabetes Care 1990;13(10):1044-50.

Apoola 2009

Apoola A, Beardsley J. Does the addition of a urine testing kit to use of contact slips increase the partner notification rates for genital chlamydial infection? International Journal of STD & AIDS 2009;20(11):775-7.

Balogun 2011

Balogun W, Abbiyesuku FM. Excess renal insufficiency among type 2 diabetic patients with dip-stick positive proteinuria in a tertiary hospital. African Journal of Medicine and Medical Sciences 2011;40(4):399-403.

Battelino 2011

Battelino T, Phillip M, Bratina N, Nimri R, Oskarsson P, Bolinder J. Effect of continuous glucose monitoring on hypoglycemia in type 1 diabetes. Diabetes Care 2011;34(4):795-800.

Beatty 1994

Beatty OL, Ritchie CM, Hadden DR, Kennedy L, Bell PM, Atkinson AB. Is a random urinary albumin concentration a useful screening test in insulin-treated diabetic patients? Irish Journal of Medical Science 1994;163(9):406-9. [MEDLINE: 7982765]

Bubner 2009

Bubner TK, Laurence CO, Gialamas A, Yelland LN, Ryan P, Willson KJ et al. Effectiveness of point-of-care testing for therapeutic control of chronic conditions: results from the PoCT in General Practice Trial. Medical Journal of Australia 2009;190(11):624-6.

Gialamas A, Yelland LN, Ryan P, Willson K, Laurence CO, Bubner TK, et al. Does point-of-care testing lead to the same or better adherence to medication? A randomised controlled trial: the PoCT in General Practice Trial. Medical Journal of Australia 2009;191(9):487-91.

Calderon-Margalit 2005

Calderon-Margalit R, Mor-Yosef S, Mayer M, Adler B, Shapira SC. An administrative intervention to improve the utilization of laboratory tests within a university hospital. International Journal for Quality in Health Care 2005;17(3):243-8.

Calero 2011

Calero A, Katz S, Thomas P, Rosenberg J, Fagan D. Use of screening urine dipsticks in well-child care to detect asymptomatic proteinuria without proper interpretation leads to unnecessary referral of pediatric patients [abstract no: 816]. In: Pediatric Academic Societies Annual Meeting; 2011 Apr 25-28; San Diego, CA. 2011.

Charles 2009

Charles M, Ejskjaer N, Witte DR, Sandbaek A. Neuropathy in a population with screen-detected type 2 diabetes [abstract no:94]. Journal of the Peripheral Nervous System 2009;14(3):254.

Dallosso 2012

Dallosso HM, Eborall HC, Daly H, Martin-Stacey L, Speight J, Realf K, et al. The DESMOND SMBG randomised controlled trial protocol. BMC Family Practice 2012;13(-):18.

Davies 1991

Davies M, Alban-Davies H, Cook C, Day J. Self testing for diabetes mellitus. BMJ 1991;303(6804):696-8.

Davies 1993

Davies MJ, Williams DR, Metcalfe J, Day JL. Community screening for non-insulin-dependent diabetes mellitus: self-testing for post-prandial glycosuria. Quarterly journal of Medicine 1993;86(10):677-84.

Davies 1999

Davies MJ, Ammari F, Sherriff C, Burden ML, Gujral J, Burden AC. Screening for Type 2 diabetes mellitus in the UK Indo-Asian population. Diabetic Medicine 1999;16(2):131-7.

Diercks 2002

Diercks GF, Stroes ES, van Boven AJ, van Roon AM, Hillege HL, de Jong PE. Urinary albumin excretion is related to cardiovascular risk indicators, not to flow-mediated vasodilation, in apparently healthy subjects. Atherosclerosis 2002;163(1):121-6.

Dolan 1987

Dolan M, Clarke H, Heffernan A, McKenna TJ. Assessment of home urine sugar tests: sensitivity, frequency of use and effect on diabetes control. Practical Diabetes 1987;4(4):161-2.

Downing 2012

Downing H, Thomas-Jones E, Gal M, Waldron CA, Sterne J, Hollingworth W, et al. The diagnosis of urinary tract infections in young children (DUTY): protocol for a diagnostic and prospective observational study to derive and validate a clinical algorithm for the diagnosis of UTI in children presenting to primary care with an acute illness. BMC Infectious Diseases 2012;12:158.

DPPRG 2005

Diabetes Prevention Program Research Group. Strategies to identify adults at high risk for type 2 diabetes: the Diabetes Prevention Program. Diabetes Care 2005;28(1):138-44.

Falguera 2010

Falguera M, Ruiz-González A, Schoenenberger JA, Touzón C, Gázquez I, Galindo C et al. Prospective, randomised study to compare empirical treatment versus targeted treatment on the basis of the urine antigen results in hospitalised patients with community-acquired pneumonia. Thorax 2010;65(2):101-6.

Fulcher 1991

Fulcher RA, Maisey SP. Evaluation of dipstick tests and reflectance meter for screening for bacteriuria in elderly patients. British Journal of Clinical Practice 1991;45(4):245-6.

Gallichan 1994

Gallichan MJ. Self-monitoring by patients receiving oral hypoglycaemic agents: A survey and a comparative trial. Practical Diabetes 1994;11(1):28-30.

Goldby 2011

Goldby S, Gray LJ, Khunti K, Williams S, Sheppard D, Taub N, et al. Using the Leicester practice risk score to identify people with Type 2 diabetes and impaired glucose regulation [abstract no: P331]. Diabetic Medicine 2011;28(Suppl 1):131.

Grimm 1997

Grimm RH, Svendsen KH, Kasiske B, Keane WF, Wahi MM. Proteinuria is a risk factor for mortality over 10 years of follow-up, MRFIT Research Group. Multiple Risk Factor Intervention Trial. Kidney International 1997;Suppl 63:S10-4.

Jolic 2011

Jolic M, Lim YL, Hamblin S, Kinney S. A retrospective audit of screening practices used to detect abnormal glucose regulation in a cohort of AMI patients admitted to a coronary care unit - An Australian study. European Heart Journal 2011;32(Suppl 1):66-7.

Jou 1998

Jou W, Powers RD. Utility of dipstick urinalysis as a guide to management of adults with suspected infection or hematuria. Southern Medical Journal 1998;91(3):266-9.

Kazemier 2012

Kazemier BM, Schneeberger C, De Miranda E, Van Wassenaer A, Bossuyt PM, Vogelvang TE, et al. Costs and effects of screening and treating low risk women with a singleton pregnancy for asymptomatic bacteriuria, the ASB study. BMC Pregnancy & Childbirth 2012;12(-):52.

Kenealy 2005

Kenealy T, Arroll B, Petrie KJ. Patients and computers as reminders to screen for diabetes in family practice. Randomized-controlled trial. Journal of General Internal Medicine 2005;20(10):916-21.

Koschinsky 1984

Koschinsky T, Toeller M, Berger H, Dannehl K, Dopstadt K, Yousefian M et al. [Improved metabolic state in insulin-tested diabetes. Self-testing of blood or urine glucose on 2 days per week]. Deutsche Medizinische Wochenschrift 1984;109(5):174-9.

Lauritzen 1994

Lauritzen T, Christiansen JS, Brock A, Mogensen CE. Repeated screening for albumin-creatinine ratio in an unselected population. The Ebeltoft Health Promotion Study, a randomized, population-based intervention trial on health test and health conversations with general practitioners. Journal of Diabetes & its Complications 1994;8(3):146-9.

Lauritzen 2008

Lauritzen T, Jensen MS, Thomsen JL, Christensen B, Engberg M. Health tests and health consultations reduced cardiovascular risk without psychological strain, increased healthcare utilization or increased costs. An overview of the results from a 5-year randomized trial in primary care. The Ebeltoft Health Promotion Project (EHPP). Scandinavian Journal of Public Health 2008;36(6):650-61.

Lenz 2002

Lenz ER, Mundinger MO, Hopkins SC, Lin SX, Smolowitz JL. Diabetes care processes and outcomes in patients treated by nurse practitioners or physicians. Diabetes Educator 2002;28(4):590-8.

Little 2009

Leydon GM, Turner S, Smith H, Little P, UTIS team. Women's views about management and cause of urinary tract infection: qualitative interview study. BMJ 2010;340:c279. [MEDLINE: 20139217]

Little P, Moore MV, Turner S, Rumsby K, Warner G, Lowes JA, et al. Effectiveness of five different approaches in management of urinary tract infection: randomised controlled trial. BMJ 2010;340:c199. [MEDLINE: 20139214]

Little P, Turner S, Rumsby K, Warner G, Moore M, Lowes JA, et al. Dipsticks and diagnostic algorithms in urinary tract infection: development and validation, randomised trial, economic analysis, observational cohort and qualitative study. Health Technology Assessment 2009;13(19):1-73. [MEDLINE: 19364448]

Turner D, Little P, Raftery J, Turner S, Smith H, Rumsby K, et al. Cost effectiveness of management strategies for urinary tract infections: results from randomised controlled trial. BMJ 2010;340:c346. [MEDLINE: 20139218]

McEwan 1990

McEwan RT, Davison N, Forster DP, Pearson P, Stirling E. Screening elderly people in primary care: a randomized controlled trial. British Journal of General Practice 1990;40(332):94-7.

McGhee 1997

McGhee M, O'Neill K, Major K, Twaddle S. Evaluation of a nurse-led continence service in the south-west of Glasgow, Scotland. Journal of Advanced Nursing 1997;26(4):723-8.

Messing 1995

Messing EM, Young TB, Hunt VB, Gilchrist KW, Newton MA, Bram LL et al. Comparison of bladder cancer outcome in men undergoing hematuria home screening versus those with standard clinical presentations. Urology 1995;45(3):387-96.

Morris 2012

Morris RK, Riley RD, Doug M, Deeks JJ, Kilby MD. Diagnostic accuracy of spot urinary protein and albumin to creatinine ratios for detection of significant proteinuria or adverse pregnancy outcome in patients with suspected pre-eclampsia: systematic review and meta-analysis. BMJ 2012;345:e4342.

Naimark 2001

Naimark DMJ, Bott MT, Tobe SW, Reznick RK, David D. Promotion of urine microalbuminuria screening among primary care physicians: a randomized, controlled, educational intervention trial [abstract]. Journal of the American Society of Nephrology 2001;12(Progam & abstracts):231a.

Naimark DMJ, Bott MT, Tobe SW. Facilitating the adoption of microalbuminuria (mau) screening among type II diabetic patients in primary care: preliminary results of a randomized educational

intervention trial [abstract no:SU-PO826]. Journal of the American Society of Nephrology 2002;13(September, Program & abstracts):638a.

Neumann 2008

Neumann CG, Bwibo NO, Siekmann JH, McLean ED, Browdy B, Drorbaugh N. Comparison of blood smear microscopy to a rapid diagnostic test for in-vitro testing for P. falciparum malaria in Kenyan school children. East African Medical Journal 2008;85(11):544-9.

Nevedomskaya 2011

Nevedomskaya E, Mayboroda OA, Deelder AM. Cross-platform analysis of longitudinal data in metabolomics. Molecular BioSystems 2011;7(12):3214-22.

Ochoa 2007

Ochoa Sangrador C, Conde Redondo F; Grupo Investigador del Proyecto. Utility of distinct urinalysis parameters in the diagnosis of urinary tract infections. Anales de Pediatria 2007;67(5):450-60.

Reyburn 2007

Reyburn H, Mbakilwa H, Mwangi R, Mwerinde O, Olomi R, Drakeley C et al. Rapid diagnostic tests compared with malaria microscopy for guiding outpatient treatment of febrile illness in Tanzania: randomised trial. BMJ 2007;334(7590):403.

Schwab 1992

Schwab SJ, Dunn FL, Feinglos MN. Screening for microalbuminuria: A comparison of single sample methods of collection and techniques of albumin analysis. Diabetes Care 1992;15(11):1581-4.

Simmons 2012

Simmons RK, Echouffo-Tcheugui JB, Sharp SJ, Sargeant LA, Williams KM, Prevost AT et al. Screening for type 2 diabetes and population mortality over 10 years (ADDITION-Cambridge): a cluster-randomised controlled trial. Lancet 2012;380(9855):1741-8.

Tissot 2001

Tissot E, Woronoff-Lemsi MC, Cornette C, Plesiat P, Jacquet M, Capellier G. Cost-effectiveness of urinary dipsticks to screen asymptomatic catheter-associated urinary infections in an intensive care unit. Intensive Care Medicine 2001;27(12):1842-7.

Worth 1982

Worth R, Home PD, Johnston DG, Anderson J, Ashworth L, Burrin JM et al. Intensive attention improves glycaemic control in insulin-dependent diabetes without further advantage from home blood glucose monitoring: results of a controlled trial. BMJ 1982;285(6350):1233-40.

Other references

Avgerinos 2001

Avgerinos DV, Björnsson J. Malignant neoplasms: discordance between clinical diagnoses and autopsy findings in 3118 cases. APMIS 2001;109(11):774-80. [MEDLINE: 11900057]

Bauer 2008

Bauer C, Melamed ML, Hostetter TH. Staging of chronic kidney disease: time for a course correction. Journal of the American Society of Nephrology 2008;19(5):844-6. [MEDLINE: 18385419]

Black 2010

Black WC, Welch HG. Overdiagnosis in cancer. Journal of the National Cancer Institute 2010;102(9):605-13.

Boulware 2003

Boulware LE, Jaar BG, Tarver-Carr ME, Brancati FL, Powe NR. Screening for proteinuria in US adults: a cost-effectiveness analysis. JAMA 2003;290(23):3101-14.

Brown 2011

Brown RS. Has the time come to include urine dipstick testing in screening asymptomatic young adults? JAMA 2011;306(7):764-5. [MEDLINE: 21846861]

Chang 2011

Chang A, Kramer H. Should eGFR and albuminuria be added to the Framingham risk score? Chronic kidney disease and cardiovascular disease risk prediction. Nephron Clin Pract 2011;119(2):c171-7.

CKDPC 2010

Chronic Kidney Disease Prognosis Consortium. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. Lancet 2010;375(9731):2073-81.

Clark 2011

Clark WF, Macnab JJ, Sontrop JM, Jain AK, Moist L, Salvadori M, et al. Dipstick proteinuria as a screening strategy to identify rapid renal decline. Journal of the American Society of Nephrology 2011;22(9):1729-36. [MEDLINE: 21807890]

Coresh 2007

Coresh J, Selvin E, Stevens LA, Manzi J, Kusek JW, Eggers P, et al. Prevalence of chronic kidney disease in the United States. JAMA 2007;298(17):2038-47. [MEDLINE: 17986697]

Djulbegovic 2010

Djulbegovic M, Beyth RJ, Neuberger MM, Stoffs TL, Vieweg J, Djulbegovic B, et al. Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials. BMJ 2010;341:c4543. [MEDLINE: 20843937]

EAU bacteriuria 2012

Grabe M, Bjerklund-Johansen TE, Botto H, Wullt B, Çek M, et al. Guidelines on urological infections. European Association of Urology. www.uroweb.org/gls/pdf/17_Urological%20infections_LR%20II.pdf (Accessed May 8, 2013).

Engberg 2002

Engberg M, Christensen B, Karlsmose B, Lous J, Lauritzen T. General health screenings to improve cardiovascular risk profiles: a randomized controlled trial in general practice with 5-year follow-up. The Journal of Family Practice 2002;51(6):546-52.

Fink 2012

Fink HA, Ishani A, Taylor BC, Greer NL, MacDonald R, Rossini D, et al. Screening for, monitoring, and treatment of chronic kidney disease stages 1 to 3: a systematic review for the US Preventive Services Task Force and for an American College of Physicians clinical practice guideline. Annals of Internal Medicine 2012;156(8):570-81. [MEDLINE: 22508734]

Friedman 1986

Friedman GD, Collen MF, Fireman BH. Multiphasic Health Checkup Evaluation: a 16-year follow-up. Journal of Chronic Diseases 1986;39(6):453-63.

Furtado 2005

Furtado CD, Aguirre DA, Sirlin CB, Dang D, Stamato SK, Lee P, et al. Whole-body CT screening: spectrum of findings and recommendations in 1192 patients. Radiology 2005;237(2):385-94. [MEDLINE: 16170016]

Grønhøj Larsen 2012

Grønhøj Larsen C, Jørgensen KJ, Gøtzsche PC. Regular health checks: cross-sectional survey. PLoS One 2012;7(3):e33694.

Gøtzsche 2013

Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. Cochrane Database of Systematic Reviews 2013;6:CD001877.

Hemmelgarn 2006

Hemmelgarn BR, Zhang J, Manns BJ, Tonelli M, Larsen E, Ghali WA, et al. Progression of kidney dysfunction in the community-dwelling elderly. Kidney International 2006;69(12):2155-61. [MEDLINE: 16531986]

Hemmelgarn 2010

Hemmelgarn BR, Manns BJ, Lloyd A, James MT, Klarenbach S, Quinn RR, et al. Relation between kidney function, proteinuria, and adverse outcomes. JAMA 2010;303(5):423-9.

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-60. [MEDLINE: 12958120]

Higgins 2011

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Hillege 2001

Hillege HL, Janssen WM, Bak AA, Diercks GFH, Grobbee DE, Crijns HJ, et al. Microalbuminuria is common, also in a nondiabetic, nonhypertensive population, and an independent indicator of cardiovascular risk factors and cardiovascular morbidity. Journal of Internal Medicine 2001;249(6):519-26. [MEDLINE: 11422658]

Hogg 2009

Hogg RJ. Screening for CKD in children: a global controversy. Clinical Journal of the American Society of Nephrology 2009;4(2):509-15.

Holme 2013

Holme Ø, Bretthauer M, Fretheim A, Odgaard-Jensen J, Hoff G. Flexible sigmoidoscopy versus faecal occult blood testing for colorectal cancer screening in asymptomatic individuals. Cochrane database of systematic reviews 2013;9(CD009259).

Imai 2007

Imai E, Yamagata K, Iseki K, Iso H, Horio M, Mkino H, Hishida A, et al. Kidney disease screening program in Japan: history, outcome, and perspectives. Clinical Journal of The American Society of Nephrology: CJASN 2007;2(6):1360-6. [MEDLINE: 17942780]

Independent UK Panel 2012

Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. Lancet 2012;380(9855):1778-86.

Iseki 2003

Iseki K, Ikemiya Y, Iseki C, Takishita S. Proteinuria and the risk of developing end-stage renal disease. Kidney International 2003;63(4):1468-74. [MEDLINE: 12631363]

Ito 1990

Ito K, Kawaguchi H, Hattori M. Screening for proteinuria and hematuria in school children--is it possible to reduce the incidence of chronic renal failure in children and adolescents? Acta paediatrica Japonica 1990;32(6):710-5.

James 2010

James MT, Hemmelgarn BR, Wiebe N, Pannu N, Manns BJ, Klarenbach SW, et al. Glomerular filtration rate, proteinuria, and the incidence and consequences of acute kidney injury: a cohort study. Lancet 2010;376(9758):2096-103.

Jørgensen 2009

Jørgensen KJ, Gøtzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. BMJ 2009;339:b2587. [MEDLINE: 19589821]

Karwinski 1990

Karwinski B, Svendsen E, Hartveit F. Clinically undiagnosed malignant tumours found at autopsy. APMIS 1990;98(6):496-500. [MEDLINE: 2383394]

Krogsbøll 2012

Krogsbøll LT, Jørgensen KJ, Grønhøj Larsen C, Gøtzsche PC. General health checks in adults for reducing morbidity and mortality from disease. Cochrane Database of Systematic Reviews 2012;10:CD009009.

Krogsbøll 2014

Krogsbøll LT. Guidelines for screening with urinary dipsticks differ substantially - a systematic review. Danish Medical Journal 2014;61(2):A4781.

Lannerstad 1977

Lannerstad O, Sternby NH, Isacsson SO, Lindgren G, Lindell SE. Effects of a health screening on mortality and causes of death in middle-aged men. A prospective study from 1970 to 1974 of mean in Malmö, born 1914. Scandinavian Journal of Social Medicine 1977;5(3):137-40.

Lin 2008

Lin K, Fajardo K. Screening for asymptomatic bacteriuria in adults: evidence for the U.S. Preventive Services Task Force reaffirmation recommendation statement. Annals of Internal Medicine 2008;149(1):W20-4. [MEDLINE: 18591632]

Malmström 2003

Malmström PU. Time to abandon testing for microscopic haematuria in adults? BMJ 2003;326(7393):813-5. [MEDLINE: 12689982]

Matsushita 2010

Chronic Kidney Disease Prognosis Consortium. Matsushita K, van der Velde M, Astor BC, Woodward M, Levey AS, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. Lancet 2010;375(9731):2073-81. [MEDLINE: 20483451]

McGrogan 2011

McGrogan A, Franssen CF, de Vries CS. The incidence of primary glomerulonephritis worldwide: a systematic review of the literature. Nephrology, dialysis, transplantation 2011;26(2):414-30.

Merenstein 2006

Merenstein D, Daumit GL, Powe NR. Use and costs of nonrecommended tests during routine preventive health exams. American Journal of Preventive Medicine 2006;30(6):521-7. [MEDLINE: 16704947]

Messing 2006

Messing EM, Madeb R, Young T, Gilchrist KW, Bram L, Greenberg EB et al. Long-term outcome of hematuria home screening for bladder cancer in men. Cancer 2006;107(9):2173-9.

Moynihan 2013

Moynihan R, Glassock R, Doust J. Chronic kidney disease controversy: how expanding definitions are unnecessarily labelling many people as diseased. BMJ 2013;347:f4298.

NICE 2008a

National Institute of Health and Clinical Excellence. Antenatal care. Routine care for the healthy pregnant woman. NICE Clinical Guideline 62. 2008. http://www.nice.org.uk/nicemedia/pdf/CG062NICEguideline.pdf (accessed 3 May 2012).

NICE 2008b

National Institute for Health and Clinical Excellence. Chronic Kidney Disease. Early identification and management of chronic kidney disease in adults in primary and secondary care. NICE Clinical Guideline 72. 2008. http://www.nice.org.uk/CG73 (accessed 3 May 2012).

NKF 2002

National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. American Journal of Kidney Diseases 2002;39(2 Suppl 1):S1-266. [MEDLINE: 11904577]

Olsen 1976

Olsen DM, Kane RL, Proctor PH. A controlled trial of multiphasic screening. New England Journal of Medicine 1976;294(17):925-30.

Prochazka 2005

Prochazka AV, Lundahl K, Pearson W, Oboler SK, Anderson RJ. Support of evidence-based guidelines for the annual physical examination: a survey of primary care providers. Arch Intern Med 2005;165(12):1347-52.

Raffle 2003

Raffle AE, Alden B, Quinn M, Babb PJ, Brett MT. Outcomes of screening to prevent cancer: analysis of cumulative incidence of cervical abnormality and modelling of cases and deaths prevented. BMJ 2003;326(7395):901.

Reid 2011

Reid S, Cawthon PM, Craig JC, Samuels JA, Molony DA, Strippoli GF. Non-immunosuppressive treatment for IgA nephropathy. Cochrane Database of Systematic Reviews 2011, Issue 3. Art. No.: CD003962. DOI: 10.1002/14651858.CD003962.pub2.

Samuels 2003

Samuels JA, Strippoli GF, Craig JC, Schena FP, Molony DA. Immunosuppressive agents for treating IgA nephropathy. Cochrane Database of Systematic Reviews 2003, Issue 4. Art. No.: CD003965. DOI: 10.1002/14651858.CD003965.

Schilling 2002

Schilling FH, Spix C, Berthold F, Erttmann R, Fehse N, Hero B et al. Neuroblastoma screening at one year of age. New England Journal of Medicine 2002;346(14):1047-53.

Sharma 2011

Sharma P, Blackburn RC, Parke CL, McCullough K, Marks A, Black C, et al. Angiotensin-converting enzyme inhibitors and angiotensin receptor blockers for adults with early (stage 1 to 3) non-diabetic chronic kidney disease. Cochrane Database of Systematic Reviews 2011, Issue 10. Art. No.: CD007751. DOI: 10.1002/14651858.CD007751.pub2.

Simmons 2012

Simmons RK, Echouffo-Tcheugui JB, Sharp SJ, Sargeant LA, Williams KM, Prevost AT, et al. Screening for type 2 diabetes and population mortality over 10 years (ADDITION-Cambridge): a cluster-randomised controlled trial. Lancet 2012;380(9855):1741-8.

Takebayashi 1992

Takebayashi S, Yanase K. Asymptomatic urinary abnormalities found via the Japanese school screening program: a clinical, morphological and prognostic analysis. Nephron 1992;61(1):82-8.

Thomason 1998

Thomason MJ, Lord J, Bain MD, Chalmers RA, Littlejohns P, Addison GM, et al. A systematic review of evidence for the appropriateness of neonatal screening programmes for inborn errors of metabolism. Journal of Public Health Medicine 1998;20(3):331-43.

Tibblin 1982

Tibblin G, Welin L, Larsson B, Ljungberg IL, Svärdsudd K. The influence of repeated health examinations on mortality in a prospective cohort study, with a comment on the autopsy frequency. The study of men born in 1913. Scandinavian Journal of Social Medicine 1982;10(1):27-32.

Vivante 2011

Vivante A, Afek A, Frenkel-Nir Y, Tsur D, Farfel A, Golan A, et al. Persistent asymptomatic isolated microscopic hematuria in Israeli adolescents and young adults and risk for end-stage renal disease. JAMA 2011;306(7):729-36. [MEDLINE: 21846854]

Wakai 2004

Wakai K, Nakai S, Kikuchi K, Iseki K, Miwa N, Masakane I, et al. Trends in incidence of end-stage renal disease in Japan, 1983-2000: age-adjusted and age-specific rates by gender and cause. Nephrology, dialysis, transplantation 2004;19(8):2044-52.

Welch 2004

Welch HG. Should I be tested for cancer? Maybe not and here's why. Berkeley: University of California Press, 2004.

Welch 2010

Welch HG, Black WC. Overdiagnosis in cancer. Journal of the National Cancer Institute 2010;102(9):605-13. [MEDLINE: 20413742]

Welch 2011

Welch HG, Schwartz LM, Woloshin S. Overdiagnosed - making people sick in the pursuit of health. Boston: Beacon Press, 2011. [Other: ISBN: 978-080702200-9]

Welch 2011a

Welch HG, Schwartz LM, Woloshin S. Overdiagnosed - making people sick in the pursuit of health. Boston: Beacon Press, 2011.

Woods 2002

Woods WG, Gao RN, Shuster JJ, Robison LL, Bernstein M, Weitzman S et al. Screening of infants and mortality due to neuroblastoma. New England Journal of Medicine 2002;346(14):1041-6.

Xiong 2005

Xiong T, Richardson M, Woodroffe R, Halligan S, Morton D, Lilford RJ. Incidental lesions found on CT colonography: their nature and frequency. British Journal of Radiology 2005;78(925):22-9. [MEDLINE: 15673525]

Appendix

Search terms

CENTRAL

1. reagent next strip*:ti,ab,kw
2. urinalysis:ti,ab,kw
3. MeSH descriptor Urine, this term only with qualifier: AN
4. test next strip*:ti,ab,kw
5. dipstick*:ti,ab,kw
6. (urin* near/2 (strip or strips or stick or sticks)):ti,ab,kw
7. urine next test*:ti,ab,kw

8. stick next test*:ti,ab,kw
9. multistix:ti,ab,kw
10. (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9)
11. ((leukocyte* or esterase* or nitrite* or haemoglobin or hemoglobin or protein or glucose or ketones) near/3 (screen or screened or screening or test or tests or tested or testing)):ti,ab,kw
12. (leukocyte* or esterase* or nitrite* or haemoglobin or hemoglobin or protein or glucose or ketones):ti,ab and screening:kw
13. (#10 OR #11 OR #12)
14. asymptomatic:ti,ab,kw
15. healthy:ti,ab,kw
16. (patients or patient or inpatient* or outpatient*):ti,ab,kw
17. ((general next practic*) or (family next practic*) or (family next physician* or general physician*)):ti,ab,kw
18. ((community next health) or (community next nurs*)):ti,ab,kw
19. (#14 OR #15 OR #16 OR #17 OR #18)
20. (#13 AND #19)

MEDLINE

1. Reagent Strips/
2. Urinalysis/
3. Urine/an
4. dipstick*.tw.
5. (urin* adj2 (strip or strips or stick or sticks)).tw.
6. urinalysis.tw.

7. urine test*.tw.
8. stick test*.tw.
9. multistix.tw.
10. or/1-9
11. ((leukocyte* or esterase* or nitrite* or haemoglobin or hemoglobin or protein or glucose or ketones) adj5 (screen or screened or screening or test or tests or tested or testing)).tw.
12. (leukocyte* or esterase* or nitrite* or haemoglobin or hemoglobin or protein or glucose or ketones).tw. and Mass Screening/
13. or/10-12
14. Asymptomatic Diseases/
15. Asymptomatic Infections/
16. (asymptomatic or healthy).tw.
17. Patients/
18. Inpatients/
19. Outpatients/
20. (patients or patient or inpatient* or outpatient*).tw.
21. exp General Practice/
22. Community Health Services/
23. Community Health Nursing/
24. (general practic* or family practice* or family physician* or general physician*).tw.
25. (community health or community nurs*).tw.
26. or/14-25
27. and/13,26
28. volunteers.tw.

29. 28 not 29

EMBASE

1. test strip/
2. urinalysis/
3. dipstick*.tw.
4. (urin* adj2 (strip or strips or stick or sticks)).tw.
5. urinalysis.tw.
6. urine test*.tw.
7. stick test*.tw.
8. multistix.tw.
9. or/1-8
10. ((leukocyte* or esterase* or nitrite* or haemoglobin or hemoglobin or protein or glucose or ketones) adj5 (screen or screened or screening or test or tests or tested or testing)).tw.
11. (leukocyte* or esterase* or nitrite* or haemoglobin or hemoglobin or protein or glucose or ketones).tw. and (screening/ or screening test/ or mass screening/)
12. or/9-11
13. asymptomatic disease/
14. asymptomatic infection/
15. (asymptomatic or healthy).tw.
16. patient/
17. outpatient/
18. hospital patient/
19. aged hospital patient/

20. (patients or patient or inpatient* or outpatient*).tw.

21. general practice/

22. general practitioner/

23. health center/

24. community health nursing/

25. (general practic* or family practice* or family physician* or general physician*).tw.

26. (community health or community nurs*).tw.

27. or/13-26

28. and/12,27

29. limit 28 to human

30. volunteers.tw.

31. 29 not 30

Downstream consequences of screening with urinary dipsticks: systematic review of observational studies

Lasse T. Krogsbøll and Karsten Juhl Jørgensen

The Nordic Cochrane Centre, Rigshospitalet.

Not yet submitted for publication.

Abstract

Background: Urinary dipsticks are often used for screening purposes, but benefits have not been documented. Dipstick testing can result in invasive diagnostic procedures and other potentially harmful consequences.

Objectives: To quantify the frequency of potentially harmful downstream consequences following dipstick screening.

Design: Systematic review.

Data sources: PubMed and reference lists of included articles.

Selection criteria: Studies that reported on the frequency of our outcomes in cohorts of persons screened with urinary dipsticks for protein or albumin, haemoglobin, leukocytes, nitrite, or glucose, alone or in combination. We did not include studies in persons with specific diseases, or studies of screening for one particular disease.

Data collection: One observer screened titles and abstract identified in the search. Two observers assessed full text papers for eligibility and two observers extracted data.

Results: Thirty-two studies were included, with highly varying methods and populations and often a high risk of bias. Our outcomes were infrequently reported and numbers were small. The frequency of outcomes ranged as follows: positive dipstick result: 0.5% to 34.1%, receiving a

diagnosis: 0% to 11.6%, receiving a cancer diagnosis: 0% to 3.4%, false positive dipstick result: 34.6% to 99.5%, renal biopsy: 0.02% to 0.73%, cystoscopy: 0% to 14.4%, any type of imaging: 0% to 14.4%, surgery: 0% to 4.7%, prescription of drugs: 0% to 2.4%, long-term drug treatment (≥ 1 year): 0% to 1.3%. The estimated need for long-term follow-up (≥ 1 year) ranged from 0% to 3.8%.

Conclusions: The potentially harmful consequences of screening with urinary dipsticks have been inadequately studied and cannot be quantified reliably. Depending on population and type of follow-up of positive dipstick results, the harms varied from almost complete absence to very high frequencies. Further research is needed, given the widespread use of screening with urinary dipsticks.

Background

Urinary dipsticks are frequently used as a screening tool for detecting various conditions in general health checks[1–5] or on admission to hospital. They test for one or several substances, typically protein, haemoglobin, glucose, leukocytes and nitrite, and can indicate the presence of a range of conditions that may have a better prognosis at earlier stages, e.g. kidney disease, diabetes, and urological malignancies. Urinary dipsticks are non-invasive and cheap, but may lead to harm through invasive follow-up investigations, such as cystoscopy and renal biopsy, and through labelling, worry, and the need for long-term follow-up and medication.

The possible benefits of screening with urinary dipsticks have not been studied in randomised trials, and few cohort studies have attempted to quantify them. Similarly, the harms are largely unexplored. Several other types of screening have been thoroughly investigated in randomised trials, and it is not a given that benefits exceed harms. For example, screening for prostate cancer with prostate specific antigen testing may not reduce mortality, but it causes substantial harm in the form overtreatment.[6] Also, large randomised trials of general health checks have shown a lack of benefit.[7]

Determining the balance between benefits and harms of screening with urinary dipsticks requires large randomised trials as these effects will likely be small in healthy populations. Since there are none, and given the widespread use of dipstick screening, a quantification of the potentially harmful downstream clinical consequences through a systematic review of observational studies is pertinent.

Methods

This review was done according to a protocol that prespecified aims, eligibility criteria, methods, and outcomes.

Types of studies

We included cohorts that were screened or invited to be screened with dipsticks, regardless of whether data collection was done prospectively or retrospectively, and regardless of the design label used by the authors. For example, we accepted simple studies of random samples of persons invited for dipstick screening, as well as studies that retrospectively distinguished between routine tests (screening) and tests performed on indication, for example in record review studies of screening in-patients. We did not include studies in which participants were selected on the basis of a disease related outcome, i.e. case-control studies,[8] or studies in people referred because of previously identified urinary abnormalities, e.g. asymptomatic microscopic haematuria, unless such studies also presented the number of persons screened.

Types of intervention

Screening with urinary dipsticks for protein, albumin, haemoglobin, leukocytes, nitrite, or glucose, alone or in any combination. We excluded studies in which simultaneous screening tests were used that would obscure the effect of dipstick screening, e.g. urine microscopy or measurement of serum creatinine.

Types of participants

To focus our review on general use of dipsticks, we excluded studies of dipstick screening for single, specific diseases and studies in populations already diagnosed with a specific disease, including urological or nephrological patients, where dipsticks could be viewed as a tool for monitoring. We accepted studies in populations selected for common risk factors such as

hypertension or type 2 diabetes since these categories encompass large proportions of adults who are often specifically targeted for screening.

Types of outcomes

Our outcomes were: number of persons with a positive dipstick result, new diagnoses, cancer diagnoses, false positives, biopsies (kidney, bladder, prostate), cystoscopies, follow-up imaging tests, operations, drug treatment, need for long-term drug treatment (>1 year), need for long-term follow-up (>1 year).

During the review process, we decided to exclude studies that only reported on the prevalence of positive dipstick results, as these studies were overwhelmingly common and not of primary interest to us.

Data extraction

Search

We iteratively developed a PubMed search using index references for assessing sensitivity and specificity. Our final strategy was:

1. urine/analysis[mesh] OR urinalysis[mesh] OR dipstick* OR (reagent strip)
2. screening OR diagnosis
3. 1 AND 2
4. animals NOT (animals AND humans)
5. 3 NOT 4

One observer (LTK) screened titles and abstracts and selected articles for full-text assessment, using a low threshold for retrieval. Two observers (LTK, KJJ) independently assessed these for eligibility, and disagreements were resolved through discussion.

Two authors (LTK, KJJ) independently extracted data from each included study into a data extraction sheet. We extracted information on study design, years of study conduct, funding source, length of follow-up (when applicable), authors' conclusions, number of participants invited (when relevant), number of participants included in analyses, number of participants excluded and the reasons, age and gender, risk profile (e.g. hypertension), country, method and place of recruitment (e.g. general practice or community), medical specialty and information on patient mix (when relevant), substances tested for, name of dipstick used, whether other screening tests that could obscure the effect of dipstick screening were performed simultaneously, number of screenings, whether positive results were verified by repeat dipstick testing before further investigation, whether trace dipstick results were counted, method of data collection for dipstick results and outcome data, information on the investigational programme for positive results (usual care, or according to a detailed protocol), and information on risk of bias. Risk of bias was assessed on four domains: selection bias (representativeness of sample), incomplete outcome data, outcome reporting bias, and other bias. Reference lists of included studies were searched for potentially relevant studies.

For our outcomes, we noted the number of people with events, distinguishing between the number of people with an event and the total number of events. We extracted details about the events, e.g. diagnoses made or types of operations. For the number of persons requiring long-term follow-up (≥1 year), we estimated this based on the diagnoses made in each study. In cases of doubt, we conferred with a specialist in either nephrology or urology, as relevant.

Data synthesis

To allow comparisons between studies, we calculated percentages although the numbers were often small. Since we studied harm, we calculated percentages based on the number screened rather than the number invited. This means that the percentages express the risk of the event for the persons screened, rather than the population effect, which is analogous to doing a per protocol analysis for harms in a randomised trial, instead of an intention to treat analysis. When enough data points were available to allow a meaningful summary statistic, a median was calculated.

For estimating the percentage of false positive test results, we subtracted the number of persons receiving a diagnosis from the number with a positive test result and divided with the number of persons with a positive test result.

Figures were made with R [2.15.1], using the Lattice package, and Microsoft Word.

Results

The search yielded 7584 references and 749 articles were selected for full-text assessment (figure 1). Of these, we included 26 studies reported in 33 articles. A further 6 studies reported in 6 articles were identified through reference lists. Thus, we included a total of 32 studies reported in 39 articles. Sixteen were studies of combined dipsticks, 5 only tested for haemoglobin, 5 for glucose, 2 for nitrite, 3 for albumin, and 1 for protein. No studies tested for leukocytes alone.

The study populations and settings were highly diverse. Twenty studies were in general populations, 6 were in primary care, 4 were in hospital settings, 1 was in a specialist clinic, and 1 included healthy newborn children at a hospital (table 1). Seventeen studies included adults, 12 included children, and 3 included all ages.

Four of the included studies analysed data from existing screening programmes with annual dipstick screening,[9–12] but only one of these reported data in a way that allowed calculation of a

percentage.[9] In the other 3 studies, the number of events for the entire period was reported, with the participants having been in the programme for varying lengths of time. We present these results as well, but do not calculate percentages.[10–12] In two of these studies, the populations partly overlap.[11] [12]

The number of dipstick tests varied, as did the type of follow-up for persons with positive dipstick results. In 14 studies the diagnostic follow-up likely reflected usual care at the time, whereas in 16 studies a specific diagnostic algorithm was used, and in 2 it was unclear. Four studies used particularly aggressive protocols for screening and diagnostic work-up, with multiple dipstick testing for haemoglobin and cystoscopy and imaging for all with a single positive test, including trace findings.[13–16]

For most of our outcomes, the number of events was small (Table 2). Furthermore, our outcomes were infrequently reported. In several cases it was clear that at least some events had occurred, without this being explicitly stated and quantified. For example, some articles listed biopsy-verified renal diseases, but not the full number of biopsies performed. We present these data as they contain important information, but with an indication that they are underestimates.

The methods were generally poorly reported. The risk of selection bias was high in 10 studies, unclear in 13 studies, and low in 9 studies. The risk of bias due to incomplete outcome data was high in 5 studies, unclear in 17 studies, and low in 10 studies, and was always in the direction of underestimation of event frequencies. The risk of outcome reporting bias was always unclear, as most studies did not focus on our outcomes, and as protocols or prespecified outcomes were rarely mentioned. In table 1 we highlight the major bias issues identified.

Positive dipstick results and diagnoses

Twenty-three studies reported the number of persons who tested positive at least once, and rates ranged between 0.5% and 34.1%, with a median of 9.7% (figure 2, table 2). Studies of combined dipsticks often had higher positive rates than studies testing for single substances. Thirteen studies reported on the number of people in whom a urinary abnormality was confirmed by subsequent dipstick tests. The rates ranged between 0.1% and 4.1% with a median of 0.9%, disregarding one study that found a rate of 20.7%, clearly due to bias (table 1).[17]

In 25 studies, the percentage of participants who received a new diagnosis ranged from 0.0% to 11.6%, median 0.5% (figure 2, table 2). When removing two studies that used aggressive protocols, the rates ranged between 0.0% and 3.6%, with the same median. In 3 studies, percentages could not be calculated.[10–12]

Nine studies reported a total of 57 cancers (figure 2, table 2), of which 38 were cancers of the bladder or ureter, 12 were prostate cancers, 6 were renal cancers, and 1 was a leukaemia metastasis to the bladder (the leukaemia was already known). In two studies percentages could not be calculated. The four studies that enrolled high-risk participants and used aggressive protocols had cancer detection rates of 0.9%, 0.9%, 1.2%, and 3.4%,[13–16] while studies with less testing and less rigorous follow-up protocols found rates close to 0%.[11 12 18–21] The median was 0.2%.

False positives

The false positive rates could be calculated for 21 studies (table 2). The range was 34.6% to 99.5% and the median 95.5%.

Biopsies

Five studies clearly reported the number of renal biopsies (figure 2, table 2),[22–27] and in 4 other studies minimum figures were available.[10 18 28–30] Three additional studies reported the number of

biopsies in a way that did not allow calculation of percentages.[10–12] Of the 9 studies in which percentages could be calculated, six used combined dipsticks, two used dipsticks for albumin, and one used dipsticks for protein. The 2 studies in in-patients found rates of 0%, and the 7 in general populations found 0.02%, 0.04%, 0.07%, 0.10%, 0.14%, 0.68%, and 0.73%.

Criteria for performing a renal biopsy were defined in 4 studies (table 1).[24 27–29] Of the remaining 5 studies, 2 were retrospective analyses of routine screening at hospital admission[22 31] and 3 were prospective studies in general populations.[18 26 30] Biopsy criteria in these 5 studies probably reflected the local standard of care at the time.

We did not find data on prostate biopsies or bladder biopsies as a result of dipstick testing.

Cystoscopies

Seven studies clearly reported the number of cystoscopies resulting from dipstick screening, and in 4 additional studies incomplete information was available (figure 2, table 2). Four studies with aggressive protocols found cystoscopy rates between 11.2% and 14.4%.[13–16] The remaining studies investigated positive test results in more realistic ways, and found rates between 0% and 3.1%.[18–23 28] In a study of university students, 6 were offered cystoscopy, but none attended.[28]

Imaging

Nine studies gave complete information on the use of imaging tests resulting from dipstick testing (figure 2, table 2).[13 15 18 21 23 25 31–33] In an additional 3, the results were incompletely reported.[14 16 27] The rates were highly varying, depending on population and type of dipstick test. Four studies using aggressive protocols found rates between 11% and 14.4%,[13–16] while the remaining studies found rates between 0% and 3.2%. In adults, the most commonly used tests were intravenous pyelography

and abdominal ultrasound, and in children it was intravenous pyelography and voiding cystourethrography.

Surgery

Eight studies reported the amount of surgery resulting from the dipstick testing (figure 2, table 2).[15 21–25 30 32 33] In three studies, no operations resulted from dipstick screening, two studies reported one operation each, and one reported 2 operations, all leading to rates of zero or close to zero. One study used an aggressive protocol, leading to a surgery rate of 4.7%,[15] and one study screened for nitrite in the urine of children, resulting in a surgery rate of 0.6%, based on 11 events.[32] The operations were: 1 ureter neo-implantation due to vesico-ureteral reflux, 5 urethral dilatations due to a "tight urethra", and 5 urethrotomies. The authors of that article remarked that they had no influence on the decisions to operate.

Drug treatment

Eight studies reported on the number of people that were prescribed drugs following dipstick screening (figure 2, table 2). Drug treatment appeared to be uncommon, with rates ranging between 0% and 2.4%. The higher rates were found in a study of adult hospital in-patients[22] and in a study using an aggressive protocol,[15] while the remaining studies found rates below 1%. The drugs prescribed were antibiotics for asymptomatic bacteriuria, angiotensin converting enzyme inhibitors, and drugs for prevention of urolithiasis.

Long-term drug treatment

Data on the number needing long-term medication as a result of the dipstick screening were available from 5 studies (figure 2, table 2), and minimum figures were available from 1 study. Rates

were close to 0%, except in one study using an aggressive protocol,[15] in which the rate was 1.3%. All numbers were small and underreporting is likely. The drugs used were angiotensin-converting enzyme inhibitors, drugs for prevention of urolithiasis, and unspecified.

Long-term follow-up

No studies explicitly reported the number of persons requiring long-term follow-up. The number could be clearly deduced from the text in one study, with a rate of 0.3%.[28] For the other studies, we calculated minimum rates based on diagnoses considered very likely to require long-term follow-up (tables 2 and 3). In 7 studies percentages could not be calculated, but the diagnoses are presented. The rates in 25 studies ranged between 0.0% and 3.8%, with a median of 0.3%. We did not count persistent asymptomatic urinary abnormalities, since follow-up could have been brief in some cases, and we always judged conservatively in cases of doubt. Thus, the figures are underestimates.

Discussion

Summary of main findings

Reliable data were scarce and the studies were often poorly reported. Rates of positive dipstick tests were nonetheless high, and combined dipsticks generally had higher rates than dipsticks for single substances, as expected. New diagnoses resulting from the screening were generally uncommon, but two studies with aggressive protocols for screening and follow-up procedures had rates of 11% and 12%. Several of these diagnoses were of unclear relevance to patients, e.g. benign prostatic hyperplasia or asymptomatic urolithiasis. Most positive dipstick results were false positives, in the sense that they did not result in a diagnosis. Rates of renal biopsy, cystoscopy, imaging, and surgery were below 1%, except in studies with aggressive protocols, as were rates of drug prescriptions and

of long-term drug use. However, underreporting was likely in many studies, as these outcomes were usually not of primary interest to the authors.

The kind of diagnostic follow-up after positive dipstick results differed and affected the results. Several studies aimed to assess the true prevalence of disease in a population, and used ambitious regimens, e.g. frequently repeated screening and cystoscopy and intravenous pyelography for all with a single trace positive result, [13–16] making the results unrepresentative of clinical practice or any realistic screening programme. It is doubtful that these studies in any meaningful way can be said to estimate the "true" level of disease in a population, as many diseases have a reservoir of cases that are of no clinical relevance.[34] Some studies used more cautious follow-up protocols, resulting in less use of invasive tests and treatments, while other studies referred to primary care physicians for planning follow-up. One such study in which follow-up was usual care found that several children with asymptomatic bacteriuria were operated on debatable indications.[32] This study design may be suitable for reflecting the harms of today's disorganised urine screening. At the same time, it is likely that prudent management of screen-detected abnormalities in an organised screening programme may result in less unneeded surgery, drug treatment, or invasive testing. We found several analyses of established screening programmes in various Asian countries,[9 10 35–38] but they rarely reported on our outcomes and few could be included.

Other literature

Several systematic reviews have assessed the value of screening for haematuria, proteinuria and bacteriuria, but were not focussed on dipsticks, and also included urine microscopy or laboratory measurements. In 1989, a review found that screening for haematuria and proteinuria was not justified based on its low positive predictive value in most populations.[39] A 2010 review of screening for bladder cancer found no trials and no high-quality cohort studies.[40] A 2008 review of

screening for asymptomatic bacteriuria found that the available evidence supported screening in pregnant women, but not in other groups.[41] Screening for diabetes with dipsticks is rarely mentioned in guidelines and is considered obsolete[42] but glucose is still included in many combined dipsticks. A 2012 review found that the effect of screening for chronic kidney disease (CKD) with any method is uncertain as no randomised trials could be identified.[43]

Our review included studies using only dipsticks, but large scale studies of dipsticks in conjunction with other screening tests have also found many with positive test results, few with serious illness and some use of invasive testing.[44–47] Studies of routine urine microscopy in in-patients and outpatients have found that it increased cost without adding important benefit.[48 49]

Strengths and limitations

We believe this is the first systematic review to assess the potentially harmful downstream consequences of screening with urinary dipsticks. We comprehensively searched the most relevant database, using a strategy designed to be sensitive, and thoroughly assessed the papers for relevant outcome data. To prevent errors, we used double and independent assessment of eligibility, data extraction and risk of bias assessment.

The most important limitation is the lack of control groups in the studies, which makes exact estimation of the effects of dipstick screening impossible. There was too much heterogeneity in methods, populations, and settings to make reliable generalisations about the frequencies of events. Our choice of limiting the search to one database (PubMed) means that our review is probably not exhaustive, as will usually be the case with a review of observational studies, but our search of reference lists did not suggest this was a big problem. The outcomes of interest to us were rarely the main focus of the included articles, and our outcomes were often haphazardly and incompletely reported. Conversely, the absence of an event was rarely described and only in few cases were we

able to extract this information with certainty. Thus, it is likely that more studies had zero events, particularly small studies.

Many studies were old, and diagnostic work-up have changed in some respects. For example, intravenous pyelography has in many countries been abandoned in favour of computed tomography with intravenous contrast, which involves greater radiation doses.

Meaning of the study

The main implication of our results is that the frequency of potentially harmful downstream events resulting from screening with urinary dipsticks is inadequately studied and cannot be quantified precisely based on the existing literature. Although we document that dipstick screening can lead to invasive procedures, radiation exposure, and long-term follow-up and drug treatment, the estimates of their frequency were usually based on small numbers and were highly variable depending on setting and design. The adequately powered studies often did not report on our outcomes.

Many articles estimated the prevalence of unidentified disease in a population, aiming to assess the potential benefit of screening. Others investigated the diagnostic yield of screening programmes which were already established and interpreted the identified pathology as evidence of success. Such approaches ignore the harm from the diagnostic process, overdiagnosis and overtreatment, the uncertainty surrounding the efficacy of treating screening-detected conditions, and the uncertainty of the magnitude of effects on a population level. For estimating the balance between benefits and harms of screening, knowledge of the frequency of identified pathology is insufficient and several authors of included studies in fact called for controlled trials. However, none have been performed and the potential benefits remain unclear, while the harms are certain to exist.

Contributorship

LTK and KJJ coined the idea. LTK drafted the protocol and KJJ contributed to the revisions. LTK performed the search and screened titles and abstracts. LTK and KJJ assessed studies for inclusion and extracted data. LTK did the analyses and drafted the manuscript and KJJ contributed to the revisions.

Acknowledgements

We would like to thank:

- Inge Eidemak and Hans-Erik Wittendorff for helping with assessing the need for long-term follow-up
- Peter C Gøtzsche for commenting on the protocol and manuscript.

References

1 Merenstein D, Daumit GL, Powe NR. Use and costs of nonrecommended tests during routine preventive health exams. *Am J Prev Med* 2006;**30**:521–7.

2 Grønhøj Larsen C, Jørgensen KJ, Gøtzsche PC. Regular health checks: cross-sectional survey. *PLoS One* 2012;**7**:e33694.

3 Lemelin J, Hogg W, Baskerville N. Evidence to action: a tailored multifaceted approach to changing family physician practice patterns and improving preventive care. *CMAJ* 2001;**164**:757–63.

4 Prochazka A V, Lundahl K, Pearson W, *et al.* Support of evidence-based guidelines for the annual physical examination: a survey of primary care providers. *Arch Intern Med* 2005;**165**:1347–52.

5 Panou C, Alevizaki P, Mauri D, *et al.* Urinary test use for cancer screening: an underestimated health economics pitfall? *J Lab Clin Med* 2004;**143**:366–7.

6 Moyer VA. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2012;**157**:120–34.

7 Krogsbøll LT, Jørgensen KJ, Grønhøj Larsen C, *et al.* General health checks in adults for reducing morbidity and mortality from disease. *Cochrane database Syst Rev* 2012;**10**:CD009009.

8 Dekkers OM, Egger M, Altman DG, *et al.* Distinguishing case series from cohort studies. *Ann Intern Med* 2012;**156**:37–40.

9 Murakami M. Screening for proteinuria and hematuria in school children--methods and results. *Acta Paediatr Jpn* 1990;**32**:682–9.

10 Utsunomiya Y, Koda T, Kado T, *et al.* Incidence of pediatric IgA nephropathy. *Pediatr Nephrol* 2003;**18**:511–5.

11 Yamagata K, Yamagata Y, Kobayashi M, *et al.* A long-term follow-up study of asymptomatic hematuria and/or proteinuria in adults. *Clin Nephrol* 1996;**45**:281–8.

12 Yamagata K, Takahashi H, Tomida C, *et al.* Prognosis of asymptomatic hematuria and/or proteinuria in men. High prevalence of IgA nephropathy among proteinuric patients found in mass screening. *Nephron* 2002;**91**:34–42.

13 Britton JP, Dowell AC, Whelan P. Dipstick haematuria and bladder cancer in men over 60: results of a community study. *BMJ* 1989;**299**:1010–2.

14 Britton JP, Dowell AC, Whelan P, *et al.* A community study of bladder cancer screening by the detection of occult urinary bleeding. *J Urol* 1992;**148**:788–90.

15 Messing EM, Young TB, Hunt VB, *et al.* The significance of asymptomatic microhematuria in men 50 or more years old: findings of a home screening study using urinary dipsticks. *J Urol* 1987;**137**:919–22.

16 Messing EM, Young TB, Hunt VB, *et al.* Home screening for hematuria: results of a multiclinic study. J.Urol. 1992;**148**:289–92.

17 Nielen MMJ, Schellevis FG, Verheij RA. The usefulness of a free self-test for screening albuminuria in the general population: a cross-sectional survey. *BMC Public Health* 2009;**9**:381.

18 Haug K, Bakke A, Daae LN, *et al.* Screening for hematuria, glucosuria and proteinuria in people aged 55-64. Technical, clinical and cost-benefit experience from a pilot study. *Scand J Prim Health Care* 1985;**3**:31–4.

19 Macleod M. Routine urine tests in general practice. *Lancet* 1970;**1**:1167–8.

20 Gawkrodger DJ, Davies M. Routine urinalysis in a general dermatology clinic. *Br J Dermatol* 1995;**132**:167–8.

21 Wakui M, Shiigai T. Urinary tract cancer screening through analysis of urinary red blood cell volume distribution. *Int J Urol* 2000;**7**:248–53.

22 Akin B V, Hubbell FA, Frye EB, *et al.* Efficacy of the routine admission urinalysis. *Am J Med* 1987;**82**:719–22.

23 Mitchell N, Stapleton FB. Routine admission urinalysis examination in pediatric patients: a poor value. *Pediatrics* 1990;**86**:345–9.

24 Vehaskari VM, Rapola J. Isolated proteinuria: analysis of a school-age population. *J Pediatr* 1982;**101**:661–8.

25 Vehaskari VM, Rapola J, Koskimies O, *et al.* Microscopic hematuria in school children: epidemiology and clinicopathologic evaluation. *J Pediatr* 1979;**95**:676–84.

26 Von Bonsdorff M, Koskenvuo K, Salmi HA, *et al.* Prevalence and causes of proteinuria in 20-year-old Finnish men. *Scand J Urol Nephrol* 1981;**15**:285–90.

27 Ahmed I, John GT, Kirubakaran MG, *et al.* Prevalence of proteinuria in rural adult population in Tamil Nadu. *Indian J Med Res* 2006;**124**:185–8.

28 Topham PS, Jethwa A, Watkins M, *et al.* The value of urine screening in a young adult population. *Fam Pract* 2004;**21**:18–21.

29 Zainal D, Baba A, Mustaffa BE. Screening proteinuria and hematuria in Malaysian children. *Southeast Asian J Trop Med Public Health* 1995;**26**:785–8.

30 Heidland A, Bahner U, Deetjen A, *et al.* Mass-screening for early detection of renal disease: benefits and limitations of self-testing for proteinuria. *J Nephrol*;**22**:249–54.

31 Hermansen MC, Blodgett FM. Prospective evaluation of routine admission urinalyses. *Am J Dis Child* 1981;**135**:126–30.

32 Kunin CM, DeGroot JE, Uehling D, *et al.* Detection of urinary tract infections in 3- to 5-year-old girls by mothers using a nitrite indicator strip. *Pediatrics* 1976;**57**:829–35.

33 Iitaka K, Oyama K, Sakai T, *et al.* Screening for bacteriuria in healthy Japanese school children. *Int J Pediatr Nephrol* 1984;**5**:159–62.

34 Welch HG. *Should I Be Tested for Cancer? Maybe Not and Here's Why*. University of California Press 2004.

35 Lin CY, Sheng CC, Chen CH, *et al.* The prevalence of heavy proteinuria and progression risk factors in children undergoing urinary screening. *Pediatr Nephrol* 2000;**14**:953–9.

36 Kawamura T, Ohta T, Ohno Y, *et al.* Significance of urinalysis for subsequent kidney and urinary tract disorders in mass screening of adults. *Intern Med* 1995;**34**:475–80.

37 Wakai K, Nakai S, Kikuchi K, *et al.* Trends in incidence of end-stage renal disease in Japan, 1983-2000: age-adjusted and age-specific rates by gender and cause. *Nephrol Dial Transplant* 2004;**19**:2044–52.

38 Cho BS, Kim SD, Choi YM, *et al.* School urinalysis screening in Korea: prevalence of chronic renal disease. *Pediatr Nephrol* 2001;**16**:1126–8.

39 Woolhandler S, Pels RJ, Bor DH, *et al.* Dipstick urinalysis screening of asymptomatic adults for urinary tract disorders. I. Hematuria and proteinuria. *JAMA* 1989;**262**:1221–4.

40 Chou R, Dana T. Screening Adults for Bladder Cancer : A Review of the Evidence for the U . S . Preventive Services Task Force. *Ann Intern Med* 2010;**153**:461–8.

41 Lin K, Fajardo K. Screening for asymptomatic bacteriuria in adults: evidence for the U.S. Preventive Services Task Force reaffirmation recommendation statement. Ann.Intern.Med. ;**149**:W20–4.

42 Krogsbøll LT. Guidelines for screening with urinary dipsticks differ substantially--a systematic review. *Dan Med J* 2014;**61**:A4781.

43 Fink HA, Ishani A, Taylor BC, *et al.* Screening for, monitoring, and treatment of chronic kidney disease stages 1 to 3: a systematic review for the U.S. Preventive Services Task Force and for an American College of Physicians Clinical Practice Guideline. *Ann Intern Med* 2012;**156**:570–81.

44 Asscher AW, McLachlan MS, Jones R V, *et al.* Screening for asymptomatic urinary-tract infection in schoolgirls. A two-centre feasibility study. *Lancet* 1973;**2**:1–4.

45 Silverberg DS, Allard MJ, Ulan RA, *et al.* City-wide screening for urinary abnormalities in schoolgirls. *Can Med Assoc J* 1973;**109**:981–5.

46 Silverberg DS. City-wide screening for urinary abnormalities in schoolboys. *Can Med Assoc J* 1974;**111**:410–2.

47 Dodge WF, West EF, Smith EH, *et al.* Proteinuria and hematuria in schoolchildren: epidemiology and early natural history. *J Pediatr* 1976;**88**:327–47.

48 Boland BJ, Wollan PC, Silverstein MD. Yield of laboratory tests for case-finding in the ambulatory general medical examination. *Am J Med* 1996;**101**:142–52.

49 Kroenke K, Hanley JF, Copley JB, *et al.* The admission urinalysis: impact on patient care. *J Gen Intern Med*;**1**:238–42.

50 Akinlaja-Mayer G. [Strip testing of urine in school medical examinations. Methods and results]. *Munch Med Wochenschr* 1971;**113**:261–4.

51 Davies M, Alban-Davies H, Cook C, *et al.* Self testing for diabetes mellitus. *BMJ* 1991;**303**:696–8.

52 Falakaflaki B, Mousavinasab SN, Mazloomzadeh S. Dipstick urinalysis screening of healthy neonates. *Pediatr Neonatol* 2011;**52**:161–4.

53 Friderichsen B, Maunsbach M. Glycosuric tests should not be employed in population screenings for NIDDM. *J Public Health Med* 1997;**19**:55–60.

54 Griffiths EP. Letter: Testing factory workers for glycosuria. *Br Med J* 1974;**1**:200.

55 Rüttimann S, Clémençon D. Usefulness of routine urine analysis in medical outpatients. *J Med Screen* 1994;**1**:84–7.

56 Van der Sande MA, Walraven GE, Bailey R, *et al.* Is there a role for glycosuria testing in sub-Saharan Africa? *Trop Med Int Health* 1999;**4**:506–13.

57 Wei J-N, Sung F-C, Lin C-C, *et al.* National surveillance for type 2 diabetes mellitus in Taiwanese children. *JAMA* 2003;**290**:1345–50.

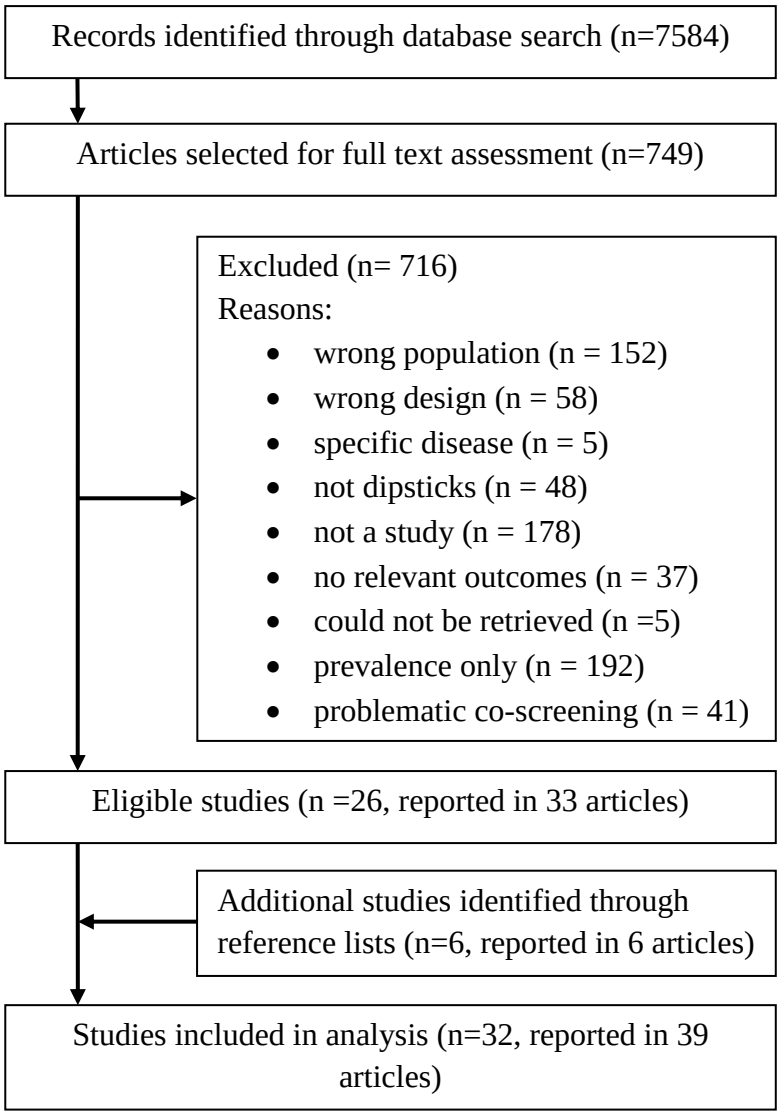


Figure 1 Flowchart describing study selection.

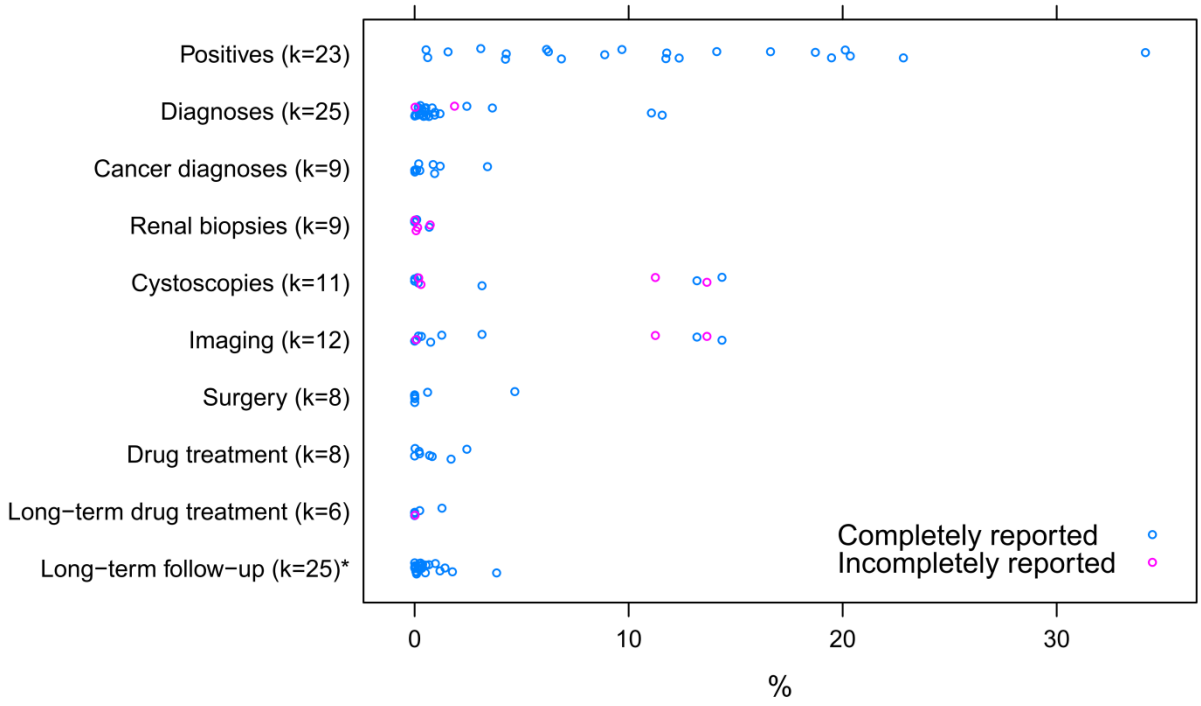


Figure 2 Overview of results. Percentage is the number of persons with the event divided with the number screened. K = number of studies.

*) 24 of 25 data points estimated by us based on diagnoses; see text.

Table 1 Characteristics of included studies.

Study	Methods
Ahmed 2006[27]	**Participants**: India. General population (age > 40 years, mean 51). **Methods**: 'Convenience sample' as part of a health survey (n=5043). Dipstick testing performed in participants home. Excluded persons with acute illness, non-ambulatory persons and menstruating women. **Follow-up**: Persons with persistent albuminuria on the second dipstick examination underwent further evaluation. Ultrasound of the abdomen was done in patients with renal failure (serum creatinine >1.4 mg/dl). Renal biopsy was performed in patients with proteinuria >1 g/day or proteinuria with an active urinary sediment or with renal failure. **Risk of bias**: High risk of selection bias as sample is described as a convenience sample. Other domains unclear.
Akin 1987[22]	**Participants**: USA. Medical in-patients. Mean age 48. **Methods**: Record review of 301 consecutively admitted patients. Eighty percent (n=243) had a dipstick test performed. Some also had urine microscopy, but the number is not specified. Three faculty members judged which tests were indicated and which were routine, and a consensus was reached. We included the results of the routine tests (51%, n=123). **Follow-up**: Usual care.

	Risk of bias: Low risk of selection bias. Included consecutive patients and only excluded three. Other domains unclear.
Akinlaja-Majer 1971[50]	**Participants**: Germany. School children aged 5.5 – 13.5 years. **Methods**: All pupils in selected schools and age-ranges were included. For analysis, we combined results from two cohorts: one with both boys and girls (n=1013), and one with only girls (n=852). **Follow-up**: Usual care. Pupils with abnormal dipstick results were given the result in writing for follow-up with their general practitioner. **Risk of bias**: Unclear
Bonsdorff 1981[26]	**Participants**: Finland. Military recruits (age 20 years). **Methods**: Included military recruits examined in 1975 (n=36,147). Those not exempted from service were re-examined in 1976 (n=29,673). **Follow-up**: Several follow-up examinations are described, but it is not clear whether they performed according to a fixed algorithm or according to clinical judgement. **Risk of bias**: Low risk of selection bias if target population is considered young men. 100 % participation. Low risk of bias due to incomplete outcome data. Other domains unclear.
Britton 1989[13]	**Participants**: UK. Age 60-85. General practice. **Methods**: Eligible men were identified from practice registers (n=942). Men under the care of a urologist or

	deemed unfit to participate were excluded (n=87), leaving 855 men. These were invited, and 578 participated. Dipstick screening was performed at the clinic, and subsequently at home, once a week for 10 weeks. **Follow-up**: All with one or more positive result were offered a full investigation, including cystoscopy, cytology and intravenous pyelography or ultrasound. **Risk of bias**: Low risk of bias from incomplete outcome data. Other domains unclear.
Britton 1992[14]	**Participants**: UK. Age > 60 years. General practice. **Methods**: All eligible men registered with 1 of 5 practices were invited for a health check (n=3152). Those who attended (n=2356) had their urine tested and subsequently self-tested their urine either once a week for 10 weeks (n=1604) or daily for 10 days (n=752). **Follow-up**: All with a positive result were offered full investigation included cystoscopy and intravenous pyelography and/or renal ultrasound. **Risk of bias**: Low risk of bias from incomplete outcome data. Other domains unclear.
Davies 1991[51]	**Participants**: UK. General population (age 45-70 years). **Methods**: Non-diabetic patients in the right age range, registered in one general practice, were invited to participate (n=2984). Half (randomised) were asked to test their urine one hour after the main meal, and the other half both before and after the main meal. We disregard this randomisation. **Follow-up**: All participants with a positive dipstick test were offered a glucose tolerance test.

	Risk of bias: Low risk of selection bias. Invited entire population in eligible age-range in selected area.79% participation. Low risk of bias from incomplete outcome data. All with positive tests were evaluated by researchers at a hospital. Other domains unclear.
Falakaflaki 2011[52]	**Participants**: Iran. Newborn children. **Methods**: Dipstick testing was performed on bag-collected urine samples from 400 healthy, full-term, breast-fed neonates. **Follow-up**: Repeat dipstick after 1 week. Children with persistent abnormalities were referred to nephrology clinic: usual care. **Risk of bias**: Unclear.
Friderichsen 1997[53]	**Participants**: Denmark. General population, age 45-76 years. **Methods**: All eligible persons in one municipality (n=3041) were invited by letter to perform self-testing 1-2 hours after a solid meal, and return a card with the results. **Follow-up**: Fasting blood glucose (capillary ear blood). Persons with glucose between 5.0 and 7.0 mmol/L were offered an oral glucose tolerance test, while persons with glucose ≥ 7.0 mmol/L on two consecutive days were diagnosed with diabetes. Persons diagnosed with diabetes were adviced to contact their family physician for treatment and follow-up. **Risk of bias**: Low risk of selection bias. Invited all persons in eligible age range in study municipality.74%

	participation. Other domains unclear.
Gawkrodger 1995[20]	**Participants:** UK. Genereral dermatology clinic. All ages (not further specified). **Methods:** All new outpatients in a 17-month period were asked to bring a urine sample (n=546), and 525 did so. Results on new diagnoses were presented. **Follow-up:** Usual care. "When an abnormal result was found, the patient was questioned about any relevant symptoms, and the test was repeated, or other appropriate tests or referrals were made." **Risk of bias**: High risk of bias from incomplete outcome data as patients do not appear to have been followed rigorously until examination programme was finished. Other domains unclear.
Griffiths 1974[54]	**Participants**: UK. Employees at a large factory (age not reported). **Methods**: All employees at a large factory (n=13,466) were sent dipsticks and 53% participated. **Follow-up**: Persons with a positive dipstick test were retested with quantitative measurement. Persons with more than 1% sugar in the urine were referred to their own doctors for further investigation. **Risk of bias**: High risk of selection bias due to low response rate. Other domains unclear.
Haug 1985[18]	**Participants**: Norway. General population. Age 55-64 years. **Methods**: Population-based random sample was sent an invitation and a collection tube to be returned by mail. Participation was 55%. **Follow-up**: Repeat dipstick, urinary microscopy, blood tests (hemoglobin, erythrocyte sedimentation rate, plasma

	creatinine), and blood pressure. Persistent haematuria: cystoscopy, X-ray and ultrasound of the kidneys and nephrological examination. Persistent proteinuria: nephrological examination (not described). **Risk of bias**: Low risk of bias due to incomplete outcome data, as the method of outcome data collection is described and appears adequate. Other domains unclear.
Heidland 2009[30]	**Participants**: Germany. Public campaign directed at the general population (all ages invited). **Methods**: A media campaign encouraged the citizens in a city (roughly 125,000 persons) to order a self-testing kit. About 100,000 kits were distributed, and 21,741 reported the results, of which 19,887 could be included in the analyses. **Follow-up**: Persons with a positive test were encouraged to contact their GP for follow-up. GPs were instructed in relevant follow-up tests. **Risk of bias**: High risk of selection bias. Self-testing with 16% participation. High risk of bias from incomplete outcome data. Outcome data were gathered through contact with GPs, who arranged follow-up investigations. Only succeeded in getting data from 35% of GPs. Other domains unclear.
Hermansen 1981[31]	**Participants**: USA. Paediatric in-patients. Age not specified. **Methods**: Included 1,553 consecutively admitted children. Children with an indication for dipstick were excluded from analysis: surgical (n=387), known renal disease (n=46), renal symptoms (n=19), and fever, suspicion of sepsis, failure to thrive, abdominal pain, abdominal trauma, or seizure with fever (n=126). Dipstick testing was

	not performed on 21, leaving 954 children in the analysis. **Follow-up**: Usual care. Clinicians were reminded about positive dipstick results, but the reaction was left to their discretion. **Risk of bias**: Low risk of selection bias and incomplete outcome data: included consecutive patients, dipstick screening done on nearly all eligible, outcome data were gathered from patient records after discharge. Other domains unclear.
Iitaka 1984[33]	**Participants**: Japan. School children (age 6-15 years). **Methods**: Sampling method is not described. First void morning urine, mid-stream, sampled at home and brought to school. Urine from all children (n=28,202) were tested with nitrite dipstick, but cloudy urines were also tested with dipslide culture and regular culture. **Follow-up**: Positive dipsticks were investigated with culture and recalled for two additional screenings. Children who still had bacteriuria after 9 months of follow-up were offered radiological investigations. Choice of investigation is not described. **Risk of bias**: Unclear
Kunin 1976[32]	**Participants**: USA. Public campaign (girls age 3-5 were targeted, but ages 1-16 were included). **Methods**: Out of 7549 eligible girls in the county, 1573 participated. Included an additional 243 girls aged 1-16 who were either outside the age range or lived outside the county. Results cannot be separated. Self-testing was

	done for three consecutive days and results were sent to researchers. **Follow-up**: Results were communicated to the participants usual physician, who was asked to investigate with "intravenous pyelogram and cystourethrogram according to his usual referral pattern". **Risk of bias**: High risk of selection bias due to self-testing. High risk of bias from incomplete outcome data. Not described how outcomes were ascertained. Other domains unclear.
Macleod 1970[19]	**Participants**: UK. General practice. Age > 40 years. **Methods**: Eligible persons attending the practice were encouraged to deliver a urine sample for screening. Persons not attending the practice were sent invitations. **Follow-up**: Usual care. **Risk of bias**: High risk of bias due to discrepancies between figure in table and total number screened. Other domains unclear.
Messing 1986[15]	**Participants**: USA. University health maintenance organisation and internal medicine clinic. Men aged > 50 years. **Methods**: All eligible men were invited, except those judged unlikely to comply, those with known unexplained haematuria, urologic disease or glomerulonephritis within one years, or those who had had urological instrumentation within 3 months. Daily testing for 5 consecutive days, followed by weekly testing for one year. **Follow-up**: All with a single positive test were offered full investigation, including cystoscopy and imaging.

	Risk of bias: High risk of selection bias. Participation rate was 38%, and a later paper found a higher morbidity rate among non-participants. Other domains unclear.
Messing 1989[16]	**Participants**: USA. Primary care (fee-for-service, health maintenance organisation, private multi-specialty clinic, and full-time academic practice). Age > 50 years. **Methods**: All eligible men were invited, except those with fulfilled exclusion criteria (similar to Messing 1986). Daily self-testing for 14 days. **Follow-up**: All with a single positive test were offered full investigation, including cystoscopy and imaging. **Risk of bias**: High risk of selection bias. Participation rate was 45%, and a later paper found higher morbidity rate among non-participants. Other domains unclear.
Mitchell 1990[23]	**Participants**: USA. Paediatric in-patients and daycare patients (53% medical, 47% surgical). Age not specified. **Methods**: Record review of 2,695 consecutive admissions. Patients were excluded if they had an indication for dipstick testing: history of hypertension, abdominal or pelvic pain, abnormally coloured urine, fever of unknown origin, urinary tract symptoms, nephrotoxic drug use, diabetes, sickle cell disease, or other systemic disease with known renal involvement. Also excluded nephrology and urology patients, and patients undergoing genitourinary surgery. Out of the 2,152 remaining children, 732 had a routine dipstick performed. Records of these were analysed (n=732). **Follow-up**: Usual care.

	Risk of bias: High risk of selection bias. Large proportion did not receive dipstick screening, leading to possible bias. Other domains unclear.
Murakami 1990[9]	**Participants**: Japan. School children aged 6-14 years. **Methods**: Analysis of existing screening programme. Gives prevalence estimates over a 13 year period. **Follow-up**: Repeat dipstick and microscopy. Quantitative urine analysis, sulphosalicylic acid test, blood pressure, personal and family history, and blood tests. Based on this, individualised plans were made. **Risk of bias**: Unclear.
Nielen 2009[17]	**Participants**: Netherlands. Public campaign directed at the general population (adults, mean age 53). **Methods**: A public campaign directed at the entire Dutch population invited adults to order a home-testing kit on the internet. 996,927 kits were ordered. The kit contained three dipsticks, and participants were told to seek out their general practitioner if two out of the three dipsticks were positive. Only 71,714 eligible persons answered the subsequent internet questionnaire. **Follow-up**: Persons with at least two positive tests out three were advised to contact their general practitioner for follow-up. **Risk of bias**: High risk of selection bias, since only 7% of those who ordered a test could be included in analyses. High risk of bias from incomplete outcome data. Outcome data were gathered through an internet questionnaire. Other domains unclear.

Ruttiman 1994[55]	**Participants**: Switzerland. Hospital out-patients (medical). Adults (mean age 41 years). **Methods**: 629 consecutive new patients entered the study. The treating physician ordered a dipstick test if they found an indication, but dipstick testing was done for all patients. 427 patients had a non-indicated dipstick test. Records were reviewed after three months for results. **Follow-up**: Usual care. **Risk of bias**: Unclear
van der Sande 1999[56]	**Participants**: Gambia. General population (age > 15 years, mean 35 years). **Methods**: A random sample of the population in the eligible age range were invited to provide a urine sample on the spot. Participation not described. **Follow-up**: Persons with a positive test were offered a fasting capillary blood glucose measurement. Diabetes was diagnosed if this was ≥6.7 mmol/L. **Risk of bias**: Unclear
Topham 2004[28]	**Participants**: UK. Student health service. Age 18-59 years. **Methods**: All students having a health check in a 2-year period were included. **Follow-up**: Repeat dipstick on two specimens, plus culture. Persistent proteinuria (0.5 g/24 h), not orthostatic, normal imaging: renal biopsy. Persistent haematuria (excluding infection, menstruation and strenuous excercise): renal biopsy and flexible cystourethroscopy. Persistent haematuria and proteinuria(0.15 g/24 h): renal biopsy.

	Risk of bias: High risk of selection bias, as only persons who attended a health check were invited. Other domains unclear.
Utsunomiya 2003[10]	**Participants**: Japan. School children aged 6-15 years. **Methods**: Annual screening. Included in analyses all children enrolled in school in Yonago city in 1983-1999. Thus, number of times screened differed between children. **Follow-up**: Repeat dipstick. If still positive, examination by primary care physician according to set criteria. Results sent to hospital based urinary screening committee, who decided on further actions. Biopsy criteria: (1) both persistent hematuria and proteinuria; (2) only persistent hematuria with more than 20 red blood cells per microscopic field and abnormal sediment; (3) only proteinuria other than orthostatic proteinuria; or (4) persistent hypocomplementemia. **Risk of bias**: Low risk of selection bias; 99.5% participation. Other domains unclear.
Vehaskari 1979 + 1982[24 25]	**Participants**: Finland. School children (ages 8, 10, 13, and 15). **Methods**: Randomly sampled 2nd, 4th, 7th and 9th grade classes in Helsinki schools. 2 morning and 2 evening samples were collected at home. Excluded girls with current menstruation. Participation 87%. **Follow-up**: Haematuria or proteinuria in 2 or more samples, or combined haematuria and proteinuria in 1 or more sample, led to investigation (poorly described). If abnormalities were still present after 3-6 months, hospitalisation, intravenous pyelogram, renal biopsy, and creatinine clearance measurement were performed.

	Risk of bias: Low risk of selection bias and incomplete outcome data. Other domains unclear.
Wakui 2000[21]	**Participants**: Japan. Age 20-79 years. General population. **Methods**: Persons participating at a health screening at five hospitals between November 1989 and October 1990 were included. Urine was tested once. **Follow-up**: Persons who tested positive were investigated with urinary red blood cell volume distribution curve analysis. If the pattern was normal or mixed, they were offered full examination, including cystoscopy, excretory urography, abdominal ultrasound, and, if necessary, computed x-ray tomography. Persons with a microcytic pattern were not investigated, but were followed-up after 3 years by telephone. Person with known urological disease or women who were menstruating at the time of testing were not investigated. **Risk of bias**: Low risk of incomplete outcome data. Follow-up investigations done at same centres as screening. Other domains unclear.
Wei 2003[57]	**Participants**: Taiwan. School children (age 6-18 years). **Methods**: All children in grades 1to 9 since 1992, and since 1993, also grades 10-12. First morning urine after 8 hours of fasting brought to school after instruction. **Follow-up**: Positives were re-tested after 2 weeks. If still positive, urine, blood, and blood pressure were examined. Parents received a report of results, and were advised to seek follow-up care from their physicians. In 2002, telephone follow-up was done to distinguish type 1 and type 2 diabetes.

	Risk of bias: Low risk of selection bias. All children in all schools in selected area were eligible and participation was 98%. Other domains unclear.
Yamagata 1996[11]	**Participants**: Japan. Men and women (95% between ages 20 and 60). Workers in a large company. **Methods**: Included all employees who had been dipstick tested as part of a company health check between 1 Jan 1983 and 31 Dec 1992. Thus, the number of times screened must have varied between participants, as must the length of follow-up. **Follow-up**: Follow-up of persons with positive dipstick tests was not done according to a protocol, but tests were ordered "as needed". Persons investigated for urinary abnormalities without identified causes were further followed, and diagnoses on those were also registered, including a syndromal diagnosis of 'chronic nephritis' meaning persistent proteinuria without identified causes. **Risk of bias**: High risk of selection bias as only employees who had a health check were included, and as the study was done at a single company. Other domains unclear.
Yamagata 2002[12]	**Participants**: Japan. Men only, age 20-59. Workers in a large company. **Methods**: Methods similar to Yamagata 1996, with partly overlapping populations. Included all male employees who had been dipstick tested as part of a company health check between 1 Jan 1983 and 31 Dec 1996. **Follow-up**: Follow-up of persons with positive dipstick tests was not done according to a protocol, but tests were ordered "as needed". Persons investigated for urinary abnormalities without identified causes were further

	followed, and diagnoses on those were also registered, including a syndromal diagnosis of 'chronic nephritis' meaning persistent proteinuria without identified causes. **Risk of bias**: High risk of selection bias as only employees who had a health check were included, and as the study was done at a single company. Other domains unclear.
Zainal 1995[29]	**Participants**: Malaysia. School children (7-12 years). **Methods**: Urine samples collected from children in eligible age range at school. **Follow-up**: Haematuria investigated with microscopy. Repeat dipstick after 1-2 months. Persistent abnormality: history, 'urinalysis', 24-hour urine protein excretion, blood tests (urea,creatinine, electrolytes, haemoglobin, haematocrit, albumin, anti-streptolysin titer, anti-nuclear antibodies, and hepatities B surface antigen). Nephritis classified as minor, moderate, or severe. Minor nephritis were put on surveillance, while children with moderate or severe nephritis were offered renal biopsy. **Risk of bias**: Unclear.

Table 2 Overview of results.

Results are presented as number of events (%). Abbreviations: a – albumin, b – bilirubin, g – glucose, h – haemoglobin, l – leucocytes, k – ketones, n – nitrite, p – protein, u – urobilinogen. a) Six were invited but all declined. b) Deduced from diagnoses. See appendix for details. c) Screening was annual, with participants screened varying numbers of times. Incidence rates not reported. d) Partly overlapping populations.

Study	Dipstick	Age	Follow-up	n	≥ one positive test	New diagnosis	FP rate	Cancer diagnosis
Hospital or specialist clinic								
Akin 1987[22]	g, h, k, l, p, u	adults	usual care	123	42 (34.1%)	3 (2.4%)	92,9%	-
Gawkrodger 1995[20]	b, g, h, k, p	adults	usual care	525	36 (6.9%)	5 (1.0%)	86,1%	1 (0.2%)
Hermansen 1981[31]	g, h, p	children	usual care	954	112 (11.7%)	4 (0.4%)	96,4%	-
Mitchell 1990[23]	g, h, p	children	usual care	732	149 (20.3%)	6 (0.8%)	96,0%	-
Ruttiman 1994[55]	b, g, h, k, l, p	all ages	usual care	427	71 (16.2%)	2 (0.5%)	97,2%	-
General population or primary care								
Akinlaja-Majer 1971[50]	a, g, n	children	usual care	1865	–	–	–	-
Falakaflaki 2011[52]	g, h, k, n, p	newborns	usual care	400	25 (6.3%)	2 (0.5%)	92,0%	-
Haug 1985[18]	g, h, p	adults	algorithm	413	40 (9.7%)	15 (3.6%)	62,5%	1 (0.2%)
Macleod 1970[19]	a, g, h	adults	usual care	1019	–	≥ 19 (1.9%)	–	1 (0.1%)
Murakami 1990[9]	g, h, p	children	usual care	7,349,928	311,864 (4.2%)	≥ 1449 (0.0%)	99,5%	-
Topham 2004[28]	h, p	adults	algorithm	3570	220 (6.2%)	10 (0.3%)	95,5%	-
Utsunomiya 2003[10]	h, p	children	algorithm	270,902 in 17 years[c]	–	38 in 17 years[c]	–	-
Vehaskari 1979[24 25]	h, p	children	algorithm	8954	1264 (14.1%)	18 (0.2%)	98,6%	-
Yamagata 1996[11]	h, p	adults	usual care	56,269 in 10 years[c,d]	–	339 in 10 years[c,d]	–	1 in 10 years[c,d]
Yamagata 2002[12]	h, p	adults	usual care	50,501 in 14 years[c,d]	–	402 in 14 years[c,d]	–	2 in 10 years[c,d]
Zainal 1995[29]	h, p	children	algorithm	45,149	4010 (8.9%)	76 (0.2%)	98,1%	-
Britton 1989[13]	h	adults	algorithm	578	132 (22.3%)	–	–	5 (0.9%)
Britton 1992[14]	h	adults	algorithm	2356	474 (20.1%)	–	–	22 (0.9%)
Messing 1986[15]	h	adults	algorithm	235	44 (18.7%)	26 (11.1%)	40,9%	8 (3.4%)
Messing 1989[16]	h	adults	algorithm	1340	261 (19.5%)	155 (11.6%)	40,6%	16 (1.2%)
Wakui 2000[21]	h	adults	algorithm	21,307	912 (4.3%)	16 (0.1%)	98,2%	1 (0.1%)
Davies 1991[51]	g	adults	algorithm	2363	73 (3.1%)	28 (1.2%)	61,6%	-
Friderichsen 1997[53]	g	adults	algorithm	2242	35 (1.6%)	15 (0.7%)	57,1%	-
Griffiths 1974[54]	g	adults	usual care	7169	–	38 (0.5%)	–	-
van der Sande 1999[56]	g	adults	algorithm	5898	36 (0.6%)	14 (0.2%)	61,1%	-
Wei 2003[57]	g	children	algorithm	2,862,083	15,271 (0.53%)	183 (0.0%)	97,8%	-
Iitaka 1984[33]	n	children	usual care	28,202	–	9 (0.0%)	–	-
Kunin 1976[32]	n	children	unclear	1816	26 (1.43%)	17 (0.9%)	34,6%	-
Ahmed 2006[27]	a	adults	algorithm	5043	594 (11.8%)	17 (0.3%)	97,1%	-
Bonsdorff 1981[26]	p	children	unclear	36,147	–	–	–	-
Heidland 2009[30]	a	all ages	algorithm	19,887	2458 (12.3%)	104 (0.5%)	95,8%	2 (0.0%)
Nielen 2009[17]	a	adults	usual care	71,714	–	185 (0.3%)	–	-

Table 2 (continued)

Results are presented as number of events (%). Abbreviations: a – albumin, b – bilirubin, g – glucose, h – haemoglobin, l – leucocytes, k – ketones, n – nitrite, p – protein, u – urobilinogen.

a) Six were invited but all declined. b) Deduced from diagnoses. See appendix for details. c) Screening was annual, with participants screened varying numbers of times. Incidence rates not reported. d) Partly overlapping populations.

Study	Dipstick	Age	Follow-up	n	Renal biopsy	Cystoscopy	Imaging	Surgery	Drug treatment	Long-term drug treatment	Long-term follow-up[b]
Hospital or specialist clinic											
Akin 1987[22]	g, h, k, l, p, u	adults	usual care	123	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (2.4%)	0 (0%)	0 (0%)
Gawkrodger 1995[20]	b, g, h, k, p	adults	usual care	525	–	≥ 1 (0.2%)	–	–	–	–	1 (0.2%)
Hermansen 1981[31]	g, h, p	children	usual care	954	–	–	3 (0.3%)	–	2 (0.2%)	–	1 (0.1%)
Mitchell 1990[23]	g, h, p	children	usual care	732	0 (0%)	1 (0.1%)	–	0 (0%)	6 (0.8%)	0 (0%)	0 (0%)
Ruttiman 1994[55]	b, g, h, k, l, p	all ages	usual care	427	–	–	–	–	1 (0.2%)	–	1 (0.2%)
General population or primary care											
Akinlaja-Majer 1971[50]	a, g, n	children	usual care	1865	–	–	–	–	13 (0.7%)	–	–
Falakaflaki 2011[52]	g, h, k, n, p	newborns	usual care	400	–	–	–	–	–	–	2 (0.5%)
Haug 1985[18]	g, h, p	adults	algorithm	413	≥ 3 (0.73%)	13 (3.2%)	13 (3.2%)	–	–	–	4 (0.4%)
Macleod 1970[19]	a, g, h	adults	usual care	1019	–	≥ 3 (0.3%)	–	–	–	–	18 (1.8%)
Murakami 1990[9]	g, h, p	children	usual care	7,349,928	–	–	–	–	–	–	1024 (0.0%)
Topham 2004[28]	h, p	adults	algorithm	3570	≥ 5 (0.14%)	0 (0%)[a]	–	–	–	–	2 (0.1%)
Utsunomiya 2003[10]	h, p	children	algorithm	270,902	≥ 29 in 17 years[c]	–	–	–	–	≥ 14 in 17 years[c]	27 in 17 years [c]
Vehaskari 1979[24 25]	h, p	children	algorithm	8954	61 (0.68%)	–	67 (0.7%)	0 (0%)	0 (0%)	–	8 (0.1%)
Yamagata 1996[11]	h, p	adults	usual care	56,269	151 in 10 years[c,d]	–	–	–	–	–	277 in 10 years[c,d]
Yamagata 2002[12]	h, p	adults	usual care	50,501	168 in 14 years[c,d]	–	–	–	–	–	321 in 14 years[c,d]
Zainal 1995[29]	h, p	children	algorithm	45,149	≥ 31 (0.07%)	–	–	–	–	–	76 (0.2%)
Britton 1989[13]	h	adults	algorithm	578	–	83 (14.4%)	83 (14.4%)	–	–	–	–
Britton 1992[14]	h	adults	algorithm	2356	–	≥ 265 (11.3%)	≥ 265 (11.2%)	–	–	–	–
Messing 1986[15]	h	adults	algorithm	235	–	31 (13.2%)	31 (13.2%)	11 (4.7%)	4 (1.7%)	3 (1.3%)	9 (3.8%)
Messing 1989[16]	h	adults	algorithm	1340	–	≥ 183 (13.7%)	≥ 183 (13.7%)	–	–	–	19 (1.4%)
Wakui 2000[21]	h	adults	algorithm	21,307	–	36 (0.2%)	38 (0.2%)	1 (0.0%)	–	–	1 (0.0%)
Davies 1991[51]	g	adults	algorithm	2363	–	–	–	–	–	–	28 (1.2%)
Friderichsen 1997[53]	g	adults	algorithm	2242	–	–	–	–	–	0 (0%)	15 (0.7%)
Griffiths 1974[54]	g	adults	usual care	7169	–	–	–	–	–	17 (0.2%)	38 (0.5%)
van der Sande 1999[56]	g	adults	algorithm	5898	–	–	–	–	–	–	14 (0.2%)
Wei 2003[57]	g	children	algorithm	2,862,083	–	–	–	–	–	–	183 (0.0%)
Iitaka 1984[33]	n	children	usual care	28,202	–	–	26 (0.1%)	1 (0.0%)	–	–	6 (0.02%)
Kunin 1976[32]	n	children	unclear	1816	–	–	23 (1.3%)	11 (0.6%)	–	–	5 (0.3%)
Ahmed 2006[27]	a	adults	algorithm	5043	2 (0.04%)	–	≥ 3 (0.1%)	–	≥ 1 (0.0%)	≥ 1 (0.0%)	17 (0.3%)
Bonsdorff 1981[26]	p	children	unclear	36,147	35 (0.10%)	–	–	–	–	–	–
Heidland 2009[30]	a	all ages	algorithm	19,887	≥ 4 (0.02%)	–	–	2 (0.0%)	–	–	74 (0.4%)
Nielen 2009[17]	a	adults	usual care	71,714	–	–	–	–	–	–	208 (0.3%)

Table 3 Diagnoses made in the studies, and diagnoses judged by us to lead to long-term follow-up (≥ 1 year); see text. Studies ordered alphabetically.

Study	New diagnoses	Diagnoses judged as leading to long-term follow-up
Ahmed 2006[27]	chronic renal failure (3), diabetic nephropathy (12), focal segmental glomerulosclerosis (1), biopsy-proven diabetic nephropathy (1)	chronic renal failure (3), diabetic nephropathy (12), focal segmental glomerulosclerosis (1), biopsy-proven diabetic nephropathy (1)
Akin 1987[22]	urinary tract infection (3)	–
Akinlaja-Majer 1971[50]	–	–
Bonsdorff 1981[26]	–	–
Britton 1989[13]	participants may have more than one diagnosis. No. of participants with ≥ 1 new diagnosis is not available. Bladder tumour (4), prostatic cancer (1), epithelial dysplasia (7), inverted papilloma (1), chronic bladder inflammation (3), urine infection (3), bladder stones (6), glomerulonephropathy (9), renal calcification (6), renal cyst (6), pelviureteric junction obstruction (1), bulbar urethral stricture (4), severe outflow onstruction (1), bladder haemangioma (1), chronic retention with bilateral hydronephrosis (1), severe phimosis (1), urethral diverticulum (1).	bladder tumour (4), prostatic cancer (1), epithelial dysplasia (7), inverted papilloma (1), glomerulonephropathy (8), pelviureteric junction obstruction (1), chronic retention with bilateral hydronephrosis (1)
Britton 1992[14]	participants may have more than one diagnosis. No. of participants with ≥ new diagnosis is not available. Bladder tumour (17), prostate cancer (5), epithelial dysplasia (28), inverted papilloma of bladder (1), chronic cystitis (16), urine infection (19), bladder stones (19), bladder haemangioma (1), amyloid of bladder (1), glomerulonephropathy (13), renal calcification (11), renal cyst (36), UPJ obstruction (1), ureteral stone (1), urethral stricture (8), bladder neck stenosis (1), urethral diverticulum (2), chronic retention (2), severe outflow obstruction (1), severe phimosis (1)	bladder tumour (17), prostate cancer (5), epithelial dysplasia (28), inverted papilloma of bladder (1), amyloid of bladder (1), glomerulonephropathy (13), UPJ obstruction (1)
Davies 1991[51]	diabetes mellitus (28)	diabetes mellitus (28)
Falakaflaki 2011[52]	ureteropelvic junction obstruction (1), vesicoureteral reflux (1)	ureteropelvic junction obstruction (1), vesicoureteral reflux (1)
Friderichsen 1997[53]	diabetes mellitus (15)	diabetes mellitus (15)
Gawkrodger 1995[20]	transitional cell carcinoma of ureter (1), renal stones (1), urinary tract infection (3)	transitional cell carcinoma of ureter (1)
Griffiths 1974[54]	diabetes mellitus (38)	diabetes mellitus (38)
Haug 1985[18]	leukaemia of bladder (1), focal GN (3), cystitis (2), urethrotrigonitis (3), cystocele (1), asymptomatic prostatic hyperplasia (1), renal calculi (2), urethral caruncle (1), glomerulonephritis (1)	focal GN (3), glomerulonephritis (1)
Heidland 2009[30]	essential hypertension (47), pyelo/interstitial nephritis (26), diabetic nephropathy (20), chronic glomerulonephritis (4), nephrolithiasis (4), hypernephroma (2), polycystic kidney disease (1)	essential hypertension (47), diabetic nephropathy (20), chronic glomerulonephritis (4), hypernephroma (2), polycystic kidney disease (1)
Hermansen 1981[31]	type 1 diabetes mellitus (1), pelvic kidney (1), sickle cell trait (1), asymptomatic bacteriuria (1)	type 1 diabetes mellitus (1)
Iitaka 1984[33]	UPJ obstruction and hydrophrosis (1), atrophic left kidney with grade III vesico-ureteral reflux and ectopic orifice (1), blunt left upper calyx with cortical thinning and small kidney(1), multiple calyectasies of left kidney and small kidney (1), blunt right upper calyx without cortical thinning (2), diverticulum of left upper calyx (1), unilateral grade I vesicouretereal reflux (1), bilateral grade II vesicoureteral reflux (1)	UPJ obstruction and hydrophrosis (1), atrophic left kidney with grade III vesico-ureteral reflux and ectopic orifice (1), multiple calyectasies of left kidney and small kidney (1), blunt left upper calyx with cortical thinning and small kidney(1), unilateral grade I vesicouretereal reflux (1), bilateral grade II vesicoureteral reflux (1)
Kunin 1976[32]	caliectasy (2), vesicoureteral reflux (5), "tight" urethra (5), meatal stenosis (5)	vesicoureteral reflux (5)
Macleod 1970[19]	diabetes (15), bladder cancer (1), bladder papilloma (2), renal stone (1)	diabetes (15), bladder cancer (1), bladder papilloma (2)
Messing 1986[15]	transitional cell carcinoma (5), renal cell carcinoma (3), calculi (5), benign prostatic hyperplasia (9), "urinary retention"(post-void volume>200 ml) (3), glomerulonephritis (1)	transitional cell carcinoma (5), renal cell carcinoma (3), glomerulonephritis (1)
Messing 1989[16]	serious and non-serious conditions reported separately, with overlap, so the numbers do not add up to 155. Serious: malignancy (16), calculus (4), calculus and urethral stricture (1), infection (6), bladder outlet obstruction (33), nephropathy (2), non-malignant bladder tumour (1). Non-serious: Infection or inflammation (2), calculi (12), benign prostatic hyperplasia (70), mild urethral stricture (3), other obstruction (2), other causes (11)	malignancy (16), nephropathy (2), non-malignant bladder tumour (1)
Mitchell 1990[23]	asymptomatic bacteriuria (6)	–
Murakami 1990[9]	'nephritis': 203, 'nephritis, suspected': 821, urinary tract infection: 425	'nephritis' (203), 'nephritis, suspected' (821)
Nielen 2009[17]	kidney disease (25), hypertension (152), DM (31)	kidney disease (25), hypertension (152), DM (31)
Ruttiman 1994[55]	type 2 diabetes mellitus (1), asymptomatic bacteriuria (1)	type 2 diabetes mellitus (1)
Topham 2004[28]	upper pole scar (1), orthostatic proteinuria (1), congenital single kidney (1), acute interstitial nephritis (1), lupus nephritis (1), familial renal disease (1), thin membrane disease (2), urinary tract infection (2)	lupus nephritis (1), familial renal disease (1)
Utsunomiya 2003[10]	urinary tract anomaly (1), urinary tract infection (8), nephrotic syndrome (2), IgA nephropathy (14), other	IgA nephropathy (14), other nephropathies (13)

	nephropathies (13)	
van der Sande 1999[56]	diabetes melitus (14)	diabetes melitus (14)
Vehaskari 1979[24 25]	urinary tract infection (9), IgA nephropathy (2), ureteropelvic stenosis (2), hereditary nephritis (1), polyarteritis (1), focal segmental sclerosis (1), hydronephrosis (1), mildly dysplastic kidney (1)	IgA nephropathy (2), ureterpelvic stenosis (2), hereditary nephritis (1), polyarteritis (1), focal segmental sclerosis (1), hydronephrosis (1)
Wakui 2000[21]	urolithiasis (8), urinary tract infection (6), benign prostatic hyperplasia (1), bladder cancer (1)	bladder cancer (1)
Wei 2003[57]	diabetes mellitus type 1 (24), diabetes mellitus type 2 (137), drug-induced diabetes mellitus (22)	diabetes mellitus type 1 (24), diabetes mellitus type 2 (137), drug-induced diabetes mellitus (22)
Yamagata 1996[11]	chronic renal failure (42), urolithiasis (48), polycystic kidney disease (4), chronic prostatitis (3), diabetic nephropathy (3), congenital renal anomaly (2), renal tuberculosis (1), medullary sponge kidney (1), bladder carcinoma (1), chronic cystitis (1), urolithiasis (7), 'chronic nephritis' (226)	chronic renal failure (42), polycystic kidney disease (4), diabetic nephropathy (3), medullary sponge kidney (1), bladder carcinoma (1), 'chronic nephritis' (226)
Yamagata 2002[12]	urolithiasis (64), polycystic kidney disease (3), congenital renal anomaly (1), renal tuberculosis (1), chronic prostatitis (1), medullary sponge kidney (1), nutcracker phenomenon (1), bladder carcinoma (2), chronic prostatitis (2), renal sarcoidosis (1), diabetic nephropathy (2), polycystic kidney (2), chronic nephritic syndrome (310), others (11)	polycystic kidney disease (3), medullary sponge kidney (1), bladder carcinoma (2), renal sarcoidosis (1), diabetic nephropathy (2), polycystic kidney (2), chronic nephritic syndrome (310)
Zainal 1995[29]	'minor nephritis' (45), 'moderate nephritis' (27), 'severe nephritis (4)	'minor nephritis' (45), 'moderate nephritis' (27), 'severe nephritis (4)

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Lasse T Krogsbøll
E-mail	l.t.krogsboll@gmail.com
Date of birth	18 June 1980
Work place	The Nordic Cochrane Centre
Principal supervisor	Peter C Gøtzsche

Title of PhD thesis:
Benefits and harms of general health checks and screening with urinary dipsticks

This declaration concerns the following article:
Krogsbøll LT, Jørgensen KJ, Grønhøj Larsen C, Gøtzsche PC. **General health checks in adults for reducing morbidity and mortality from disease. Cochrane Database of Systematic Reviews 2012;issue 10:CD009009**

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	(A,B,C)
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	C

2. Planning of the experiments and methodology design, including selection of methods and method development	C
3. Involvement in the experimental work	C
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

*Benchmark scale of the PhD student's contribution to the article		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	Name:	Title:	Signature:
	Karsten Juhl Jørgensen	MD, DMSc	
	Christian Grønhøj Larsen	MD	
	Peter C Gøtzsche	Professor, MD, DMSC	

Signature of the PhD student and the principal supervisor:	
Date: 3/12-19	Date: 3 Dec 2019
PhD student: Lasse Ryg Pedersen	Principal supervisor:

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Lasse T Krogsbøll
E-mail	l.t.krogsboll@gmail.com
Date of birth	18 June 2014
Work place	The Nordic Cochrane Centre
Principal supervisor	Peter C Gøtzsche

Title of PhD thesis:
Benefits and harms of general health checks and screening with urinary dipsticks

This declaration concerns the following article:
Krogsbøll LT, Jørgensen KJ, Gøtzsche PC. Screening with urinary dipsticks for reducing morbidity and mortality. Cochrane Database of Systematic Reviews. Accepted for publication.

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	(A,B,C)
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	C
2. Planning of the experiments and methodology design, including selection of methods and method development	C

3. Involvement in the experimental work	C
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

Benchmark scale of the PhD student's contribution to the article		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	Name:	Title:	Signature:
	Karsten Juhl Jørgensen	MD, DMSc	
	Peter C Gøtzsche	Professor, MD, DMSc	

Signature of the PhD student and the principal supervisor:	
Date: 3/12-14	Date: 3 Dec 2014
PhD student: Lasse Krogsbøll	Principal supervisor:

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Lasse T Krogsbøll
E-mail	l.t.krogsboll@gmail.com
Date of birth	18 June 2014
Work place	The Nordic Cochrane Centre
Principal supervisor	Peter C Gøtzsche

Title of PhD thesis:
Benefits and harms of general health checks and screening with urinary dipsticks

This declaration concerns the following article:
Krogsbøll LT, Jørgensen KJ. **Downstream consequences of screening with urinary dipsticks: systematic review of observational studies. Not submitted for publication.**

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	(A,B,C)
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	C
2. Planning of the experiments and methodology design, including selection of methods and method development	C

3. Involvement in the experimental work	C
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

Benchmark scale of the PhD student's contribution to the article		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	Name:	Title:	Signature:
	Karsten Juhl Jørgensen	MD, DMSc	

Signature of the PhD student and the principal supervisor:	
Date: 3/12-14	Date: 3 Dec 2014
PhD student: Lise Krogsbøll	Principal supervisor: