Appendix 2: Detailed results for secondary outcomes of study

Morbidity

Few trials reported on well-defined clinical events. The Göteborg 1970 trial did not find effects on non-fatal CHD, RR 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)[22]. The results from the WHO trial 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).[27] The OXCHECK 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 found that after 7 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.[26] The conditions were not defined, and likely included elevated risk factors like blood pressure or blood glucose.

The South-East London Screening Study did not find effects on the prevalence of angina, ischaemic changes on electrocardiogram or bronchitic symptoms after 5 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).[14] 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 reported reasons for hospitalisations in categories, e.g. ischaemic heart disease, cerebrovascular disease, and neoplasms, and did not find differences between groups.[20] There was low power due to the stratification on disease categories. See results on total hospitalisation below.

The British Family Heart Study investigated the effect on the prevalence of four conditions.[17] 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 are 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 it 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 screened and unscreened groups would allow an assessment of the full effect of screening on diagnostic activity. However, only one trial reported such results, but only for the first six years.[26] 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 participants of 4.3 in the screening group and 3.6 in the control group. This corresponds to a 20% increase. The trial lasted for 16 years, but follow-up for new diagnoses was not continued.

Three 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 found an average of 2.3 diseases per person at the first screening.[14] Of these 53% were not previously known. The Ebeltoft trial reported the percentage of participants with abnormal findings prompting health advice at the inititial screening to be 76%.[18] 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 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 found a total of 2 031 abnormalities in 384 people screened.[23] This trial used very broad biochemical screening.

In summary, health checks likely 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, e.g admission rates, number of people admitted once or more, or number of days in hospital.

The Kaiser Permanente trial reported the mean number of days in hospital over 18 years of follow-up.[26] 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 reported the number of participants admitted to hospital once or more during nine years of follow-up; risk ratio 1.04 (95% CI 0.96 to 1.13).[14] The amount of missing data is unclear, but probably low for this outcome.

The Malmö trial also studied the number admitted once or more and found similar results, risk ratio 1.05 (95% CI 0.92 to 1.20).[20] There was 3-5% missing data.

The Salt Lake City trial 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.[23] The result is unreliable due to biased exclusions after randomisation.

The Ebeltoft trial compared admission rates in the two intervention groups with the control group and did not find an effect after eight years, rate ratio 0.91 (95% CI 0.63 to 1.32).[18] They also compared the random sample invited to participate in the trial with all not invited and found similar results, rate ratio 0.97 (95% CI 0.80 to 1.18). There was five percent 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.

Disabillity

Three trials investigated the effect on disabililty. The Kaiser Permanente trial found that after 16 years 31% of the screening group and 30% of the control reported total or partial disability on a questionnaire.[26] Attrition was roughly one third and response rates around 75%, which leaves only half of the people randomised in this analysis. The South-East London Screening Study found that 2,5% in the screening group and 1.8% in the control group reported major disability after five years.[14] There was between 40% and 50% missing data in this analysis. The Salt Lake City trial compared the number of disability days before and after the intervention and did not find an effect.[23]

In summary, we did not find an effect on disability but the results are 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 used the General Health Questionnaire (GHQ-12) at baseline and after one and five years.[18] 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 used the Middlesex Hospital Questionnaire on a subset of participants after five years.[14] 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[14] 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 (Chi^2=3.274, P=0.07).

The Ebeltoft trial used a 5-point scale at baseline and after five years.[18] 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 are only available in a graph. This shows 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 79.5% of the screening group and 75.7% of the internal control group reported good or excellent health after one year.[17] 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 used SF-12 and found significantly slower deterioration of both physical and mental health components in the intervention group.[25] For mental health, the difference after 5 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 is 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 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 1992-3 (intervention groups screened) and 1997 (intervention and control group screened).[18] 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 1 and 9 years, with missing outcome data ranging between 5%[18] and 51%.[23]

The Kaiser Permanente trial found a mean number of physician visits of 16.0 in both groups after five years, not including the screenings themselves.[26] The results are 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 did not find an effect on the mean annual number of physician visits.[14] It is not clear whether the screening visits were included in this, and we cannot tell whether the results are from the 5 year or 9 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 found an average number of consultations per participant of 5.4 in the screening group and 5.0 in the control group during 1½ years.[15] This did not include the screenings themselves. The type of health check was not specified, and there is a high risk of detection bias.

The Salt Lake City trial did not find effects after one year, but this result is unreliable.[23] The screening visits were not included in the analysis.

The Ebeltoft trial found an increased rate of physician visits after 5 years in the screening plus health discussion group compared to the control group, rate ratio 1.15 (95% CI 1.02 to 1.31), but not in the screening only group compared to controls, rate ratio 1.01 (95% CI 0.89 to 1.15).[18] 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.

Of relevance, the Kaiser Permanente trial reported the mean number of laboratory tests per participant after 5 and 10 years, based on a 20% sample.[26] 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 is 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 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 anti-hypertensive medications compared to 19.6% in the control group (Chi^2=16.41, P<0.0001, our calculation).[22]

The Kaiser Permanente trial 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.[26]

The Ebeltoft trial presented data on self-reported use of selected types of drugs after five years.[18] In the screening groups, 4.8% reported using blood pressure medication compared to 6.8% in the control group (Chi^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 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.[17]

The Mankato trial reported that the proportion of participants on blood pressure medication after one year was 13.8% in the intervention group 9.8% in the control group (P<0.05).[24]

In summary, we cannot conclude on total drug use. Two out of four trials found increased the use of anti-hypertensive medication, but there was is 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.[14,26] Neither trial found an effect, and neither trial reported the exact results but only mentioned their findings in a narrative.