Should we use total mortality rather than cancer specific mortality to judge cancer screening programmes?

James Penston believes all cause mortality is a more reliable measure of the effectiveness of screening, but **Robert Steele** and **David Brewster** think it is too stringent

James Penston consultant physician, Scunthorpe General Hospital, Scunthorpe DN15 7BH, UK
james.penston@nhs.net

YES Cancer screening is a source of much dispute—in the case of breast cancer, arguments have raged for more than a decade.[1] One major concern is how the effects are to be measured. Disease specific mortality is used extensively in trials of cancer screening,[2 3] and as the aim of screening is to reduce deaths from the target disease,[2 4] this might seem to be a suitable end point. But the arguments against using disease specific mortality weigh heavily, and all cause mortality is a better measure.

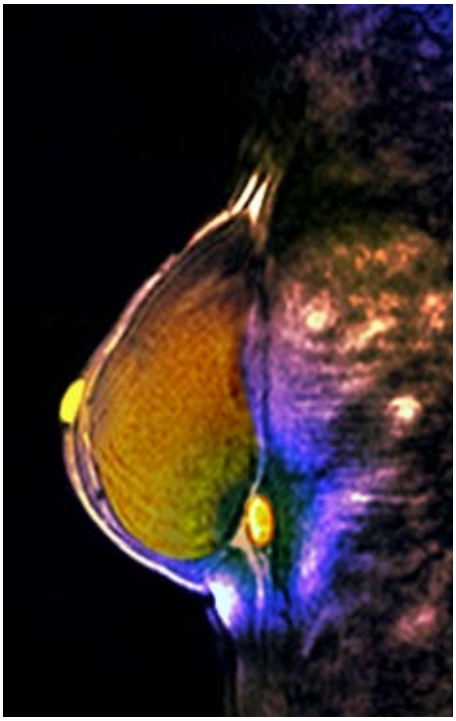
Uncertainties relating to cause of death

Clearly, the accuracy of disease specific mortality depends on correctly identifying the cause of death. However, this is often unreliable,[2] and it entails decisions that can introduce biases, either for or against screening.[2 3] Claims that bias favouring screening predominates have been disputed.[2 4] Nevertheless, one thing is sure: the accuracy of all cause mortality depends solely on the number of deaths identified and is not subject to bias. It is therefore a more reliable end point.

Disease specific mortality also ignores the fact that screening for cancer causes harm. Invasive procedures may have fatal complications, while overdiagnosis—that is, the identification and treatment of tumours that otherwise would have caused no disease—may also result in death.[1-3] A review of 12 trials of screening for breast, lung, and bowel cancer raised doubts about both the identification of screening related deaths and their inclusion in disease specific mortality.[2] If screening related deaths are not included in the mortality figures, the results will be skewed in favour of screening. In contrast, all cause mortality balances the benefits and harms of screening in a single measure.

Shifting definition

Although disease specific mortality is unproblematic when used descriptively, difficulties arise when it is used as an end point in randomised controlled trials. Should the figure include death occurring in a case of overdiagnosis? And what about someone who does not have colorectal cancer but who dies from a perforation due to screening colonoscopy? Although such deaths are not



ZEPHYR/SPL

Robert J C Steele clinical director , Scottish Bowel Screening Programme r.j.c.steele@dundee.ac.uk
David H Brewster director, Scottish Cancer Registry, Edinburgh, UK

NO All medical interventions have the potential to cause harm. This is particularly important in the case of cancer screening because the intervention is offered to people who are, or at least believe themselves to be, in good health, and the tolerance limit of harm must accordingly be low.

Screening may cause harm in several ways.[1] If the screening test is not highly sensitive, false negative results may induce reassurance and create a "certificate of health effect." In other words, people who have received a false negative test result may ignore symptoms or continue to engage in risky behaviour. Then the test itself can cause harm—for example, colonoscopy as a consequence of colorectal screening may lead to colonic perforation or other complications, and this must be monitored. Given that most people who are screened will not have disease, unnecessary psychological morbidity may also be created. Finally, screening inevitably leads to a degree of overdiagnosis —that is, people will be found to have disease that was not destined to become symptomatic in their lifetime. If they are harmed by the treatment of that disease, they will have been disadvantaged by participating in screening. Indeed, if a patient dies as a result of the treatment of screen detected disease, their life will have been shortened by screening.

The question of how we should judge screening programmes is therefore extremely important. We fully concur that any screening programme that causes a demonstrable increase in total mortality, regardless of its effect on cancer specific mortality, is unsustainable. However, stopping a screening programme that does not show a decrease in total mortality is not justifiable.

Because of the biases inherent in screening, the only robust method of proving efficacy is by population based randomised controlled trials. If such trials show that screening reduces disease specific mortality, we can be sure that early detection has a true effect on the natural course of the disease and that the effect is not solely due to lead time or self selection by a particularly healthy population. Randomised trials carried out for both breast[2-7] and colorectal cancer[8-11] screening have consistently shown reductions in disease specific mortality in the region of 20%.

All references are in the version on bmj.com

WHERE DO YOU STAND ON THE ISSUE?
Tell us on bmj.com

strictly linked with disease specific mortality, they are obviously relevant.

We can accommodate all screening related deaths in a randomised trial only by changing disease specific mortality into a vague and arbitrary end point. Alternatively, we could use all cause mortality which is untouched by these problems.

The unfeasibility argument

Advocates of disease specific mortality have a fall-back position.[4] The target cancer, they argue, contributes little to total mortality; trials would have to recruit millions of people to show a statistically significant reduction in all cause mortality; such trials are not feasible; hence, we have to rely on disease specific mortality.

This argument does not show that disease specific mortality is better than all cause mortality; indeed, it seems to concede the opposite point. It also assumes that huge trials would show a reduction in all cause mortality, whereas this is precisely what is in question. And it ignores the existing data that strongly support an absence of any effect of screening on all cause mortality, as, for example, in the case of bowel cancer screening.

Specific case against disease specific mortality

The NHS Bowel Cancer Screening Programme was implemented on the basis of four large randomised trials.[5] Meta-analysis showed a reduction in disease specific mortality of 15% in the screening group compared with controls (odds ratio 0.85, 95% confidence interval 0.78 to 0.92).[5] The odds ratio for all cause mortality, however, was 1.0 (0.99 to 1.02). Thus, in more than 300 000 people included in the four trials, there was no difference in survival[5 6]; nor was there anything to indicate that a larger trial would be worthwhile.

Given that the absolute reduction in disease specific mortality was only 0.1% over 10 years,[7] it would require an enormous trial to detect a difference in all cause mortality between the screening group and controls. But is this really necessary? Taking the above criticisms of disease specific mortality into account and considering the robust nature of all cause mortality in this example, surely we should accept that screening for colorectal cancer has no effect on overall survival?

Conclusions

All cause mortality is a hard end point that is free from bias, produces a robust estimate of the effect, and answers the crucial question of whether cancer screening improves overall survival. In contrast, disease specific mortality requires decisions which introduce bias and fails to deal effectively with deaths from screening. Unsurprisingly, the two measures often support opposing conclusions.[2]

"If screening related deaths are not included in the mortality figures, the results will be skewed in favour of screening"

Disease specific mortality is used in cancer screening trials primarily because it allows the identification of very small reductions in mortality from the target disease. Without it, there would be nothing to support current cancer screening programmes. But this is no reason to accept flawed data. On the contrary, we should prefer the evidence of all cause mortality, recognise that bowel and breast cancer screening do not improve overall survival, and question whether these programmes should continue.

Competing interests: None declared.

Provenance and peer review: Commissioned; not externally peer reviewed.

Cite this as: *BMJ* 2011;343:d6395w

Screening may cause harm in several ways. If the screening test is not highly sensitive, false negative results may induce reassurance

Too stringent

Demonstrating a reduction in all cause or total mortality, however, is a different matter. As even common cancers account for only a small proportion of total deaths (for example, in the United Kingdom, colorectal cancer accounts for 3% of all deaths), to show a reduction in disease specific mortality being translated into a reduction in total mortality would require trials that are too large to be feasible. Furthermore, it is inappropriate to use disease specific mortality as a surrogate for all cause mortality; cancer screening is not designed to reduce all cause mortality but the number of people dying prematurely, in a particularly unpleasant manner.

Proponents of using all cause mortality as an outcome indicator argue that it avoids the bias inherent in the determination of causes of death. In a frequently quoted paper that examined all cause mortality in randomised trials of cancer screening,[12] the point was made that the effect on all cause mortality was often in the opposite direction from the effect on disease specific mortality. However, close examination of the data shows that the confidence intervals around the differences in all cause mortality figures were much wider than those around the disease specific figures and did not reach anywhere near statistical significance, with the exception of a beneficial effect on all cause mortality in the Edinburgh mammography trial. In the correspondence that followed this article, even the original authors concede that a significant reduction in all cause mortality is too stringent a requirement for the determination of the efficacy of screening.[13]

Thus, if a trial shows a reduction in disease specific mortality, even though it has no demonstrable effect on total mortality, it has provided sufficient evidence to offer such screening to the population that has been studied in the trial. To insist that a trial should show a reduction in all cause mortality would deny society the opportunity to engage in screening that, on balance, is more likely to prevent cancer death than cause harm. On the other hand, it is reasonable to insist that potential participants are provided with adequate information and that risks are expressed in absolute terms, with the proviso that such information is understandable to the majority of the population. The Scottish bowel screening programme leaflet, for example, states explicitly that one bowel cancer death is prevented for roughly every 650 people invited for regular screening[14]; even so, uptake is currently 53%.[15]

We emphasise again that it is both appropriate and necessary for the effect of screening on all cause mortality to be assessed and that any screening programme that has a significant adverse effect on this measure must not be supported. However, it is unreasonable to single out screening, which is an intervention with a specific aim, as having to prove a reduction in all cause mortality at a population level. If all medical interventions were similarly constrained, then much of what health professionals do would be deemed inappropriate.

Competing interests: None declared.

Cite this as: *BMJ* 2011;343:d6397

bmj.com

A recent poll on bmj.com asked:

"Should we use total mortality rather than cancer specific mortality to judge cancer screening programmes?"

53% voted yes, out of a total 232 votes cast