

Projecting the yearly mortality reductions due to a cancer screening programme

J Med Screen
20(3) 157–164
© The Author(s) 2013
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0969141313504088
msc.sagepub.com



Zhihui (Amy) Liu, James A Hanley* and Erin C Strumpf

Abstract

The decision on whether to implement a 20-year screening programme for a cancer requires weighing the harms and costs against the health benefits (such as the number of cancer deaths averted every year). The evidence of the benefits is often based on a single-number summary, such as the mortality reduction over the entire follow-up time in a single trial, or an average of such one-number measures from a meta-analysis of several trials. There are several problems associated with using the traditional one-number summaries from trials to deduce the yearly mortality reductions expected from a sustained screening programme. We here propose using a rate ratio curve, and its complement (a mortality reduction curve), to address the mortality impact (timing, magnitude, and duration) of a screening programme. This curve is easy to interpret, as it shows when mortality reductions begin, how big they are, and how long they last. We illustrate when and how such rate ratio curves from screening trials could be computed, and how they could be used to compare reduction patterns expected with different screening regimens. We encourage trialists to report the necessary data to arrive at such projections.

Keywords

cancer screening, yearly mortality reductions, rate ratio curve, time lag

Date received: 27 February 2013; accepted: 15 August 2013

Introduction

Making a decision on whether or not to implement a 20-year screening programme for a cancer requires weighing the harms and costs against the health benefits (such as the number of cancer deaths averted every year). The evidence of the benefits is often based on a single-number summary, such as the mortality reduction over the entire follow-up time in a single trial, or an average of such one-number measures from several trials. We recommend against using such one-number summaries to deduce the yearly mortality reductions expected from a sustained screening programme.

As detailed below, we base this recommendation on several reasons, all stemming from the characteristic time-pattern of the mortality reductions produced by any particular screening programme, and the affected time-window in question. First, the reductions do not begin in year one, and if/when they do reach a ‘constant’ level, they do not remain at this level indefinitely. Thus the full pattern (ie. the timing, magnitude and duration of the reductions) cannot be adequately quantified by one number. In addition, the pattern is specific to the screening regimen (eg. the number of screens and spacing between them) employed. For example, 20 annual screens might produce yearly reductions that start at year 5, and extend over possibly 25 years; 10 annual rounds would produce similar yearly reductions starting at about the same time, but extending over a shorter span, possibly

15 years. Compared with 10 annual screenings, the yearly reductions produced by 10 biennial screenings are expected to be smaller but over a longer period of time.

We here address the task of projecting the mortality impact of a screening programme. In Section 1, we propose using a rate ratio curve, instead of a single-number summary, to fully describe the expected timing, magnitude and duration of this impact. In Section 2, we identify trials that have had sufficient rounds of screening allowing us to estimate the asymptote of the curve for a programme with similar spacing of screens. We also give examples to illustrate how much underestimation is involved in the traditional measure. In Section 3, we show how it is possible to use an existing model (previously used for other purposes), and available trial data to project the programme impact, as well as compare reduction patterns produced by regimens with different spacings from those used in trials. Finally, we call on trialists to report necessary data to compute this rate ratio curve.

McGill University Montreal Quebec H3A 1A2 Canada

Corresponding author:

James Hanley, McGill University, Purvis Hall, 1020 Pine Avenue West, Montreal, QC Quebec H3A 1A2, Canada.
Email: james.hanley@mcgill.ca

I The mortality reduction curve, and its shape

The time lag and the affected age window

Consider a cohort of persons who, beginning at age 50, are invited to be screened annually for a cancer until they reach age 69. The mortality impact of the programme is the difference in the yearly number of cancer deaths in the absence of screening (when no one is invited to be screened) versus in the presence of screening (when everyone is). We graph this impact in the affected age window in Figure 1(a).

The first notable feature is the time lag between when a screening programme starts and when the mortality reduction first manifests. Unlike most medical interventions those produce a virtually immediate effect (within hours, days or weeks), cancer screening generates mortality reductions that only become evident several years after the onset of screening.¹⁻⁵ The first screen (say at age 50) detects, and the resulting earlier therapy eradicates, some cancers that otherwise would have proved fatal several years later (from say 55 to 63). Presumably, the average delay would be longer for cancers of the breast and prostate, and shorter for more aggressive cancers, such as that of the lung. The width of the reduction 'wave' (8 years in our example) reflects the variation in cancer stages at detection and in the rates at which cancers would have progressed otherwise.

Mortality reductions produced by subsequent annual screens (at ages 51, 52, ..., 69) occur even later (from

say 56 to 64, 57 to 65, ..., 74 to 82). After the effect of the last screen disappears, cancer mortality rates return gradually (from say 78 to 85) to those in the absence of screening. Thus, the 20 screens affect possibly 35 age-bins in the age-span 50 to 85.

The total number of deaths averted in that span is shown as the white area in Figure 1(a). The total number of years gained is the sum of the products of the age-specific number of deaths averted and the age-specific remaining life expectancies. For costing purposes, this total can be averaged over the number averted, invited, or screened.

The mortality rate ratio curve

Another way to display the same mortality reductions in Figure 1(a) is through a rate ratio curve, as in Figure 1(b). The yearly ratio is calculated as the yearly number (or rate) of cancer deaths in the presence of screening divided by the yearly number (or rate) of cancer deaths in the absence of screening. Each yearly ratio can be thought of as the fraction of fatal cancers that could not be helped by screening. Their complements, usually expressed as percentages, represent the yearly mortality reductions.

If the yearly number of fatal cancers remains constant throughout the screening programme, the rate ratio curve should exhibit a bathtub shape: it would be close to constant for a large portion of the age-window where the effect of sustained screening is manifest. Little mortality impact is expected in the early portion, ie. before the

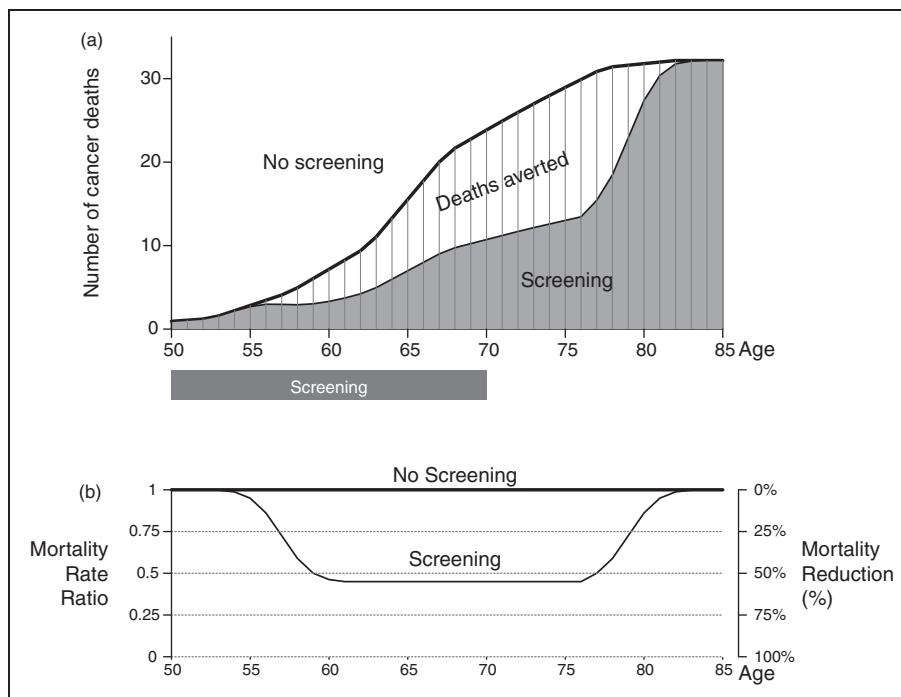


Figure 1. Impact of a hypothetical 20-year screening programme measured (a) in absolute numbers of cancer-specific deaths averted and (b) as rate ratios and as percentage reductions.

deaths averted by the first screen would have otherwise occurred, and again in the late portion, ie. long after the deaths averted by the last screen would have otherwise occurred. By describing the timing, magnitude and duration of the yearly reductions over the full time window that would be affected by a screening programme, the curve shows when reductions begin, how big they are, and how long they last.

The rate-ratio curve in Figure 1(b) is not new: Morrison¹ introduced a schematic version, entitled “changes in the disease-specific mortality rate”, to graphically illustrate and emphasize the time lag between the first screen and the beginning and end of the mortality reductions. Early trialists⁶ were also keenly aware of the waning effect after the termination of screening. A more comprehensive version, showing what affects the shape, is presented in a theoretical piece by Miettinen et al², and then in an application to mammography with the asymptote as the ‘estimand’. Hanley³ showed how a rate ratio curve could arise as the convolution of the effects of 10 annual rounds of screening, and also studied the asymptote in colon cancer screening; Baker et al⁵ simulated rate ratio curves under screening of large, moderate and little effect. These four versions are shown in Figure 2.

Much of the statistical work that has addressed this non-proportional hazards time pattern has focused on statistical tests applied to data from screening trials, and thus on maximizing statistical power^{7,8} dealing with the non-proportionality⁹, and selecting the optimal time at

which the analysis of trial data should be carried out.¹⁰ The data analysis in each actual trial tested a regimen-specific null hypothesis over some (un-predetermined) follow-up period: “does the amount and spacing of screening used in this trial have a non-zero impact on cancer mortality?” There has been much less focus on deducing the impact of a sustained screening programme.

2 Distinction between nadir in a trial and asymptote in a programme

Trial nadir and programme asymptote

Our focus is on identifying the *asymptote* of the rate ratio curve, as it represents the sustained reduction that could be expected from a screening programme. In the following, we describe how it is possible – but only in some instances – to estimate the programme asymptote from trial data.

Figure 3 shows the distinctive patterns produced by a trial of 3 annual screenings versus by a programme of 20 annual screenings. If each round of screening reduces mortality over 5 future years, then three rounds would produce 3 waves of such reductions. The affected time window spans over a total of 7 years, with a maximum reduction of 35% in year 6. In contrast, a programme of 20 screenings would produce 20 such waves, affecting many more years, with a sustained reduction of 46% for

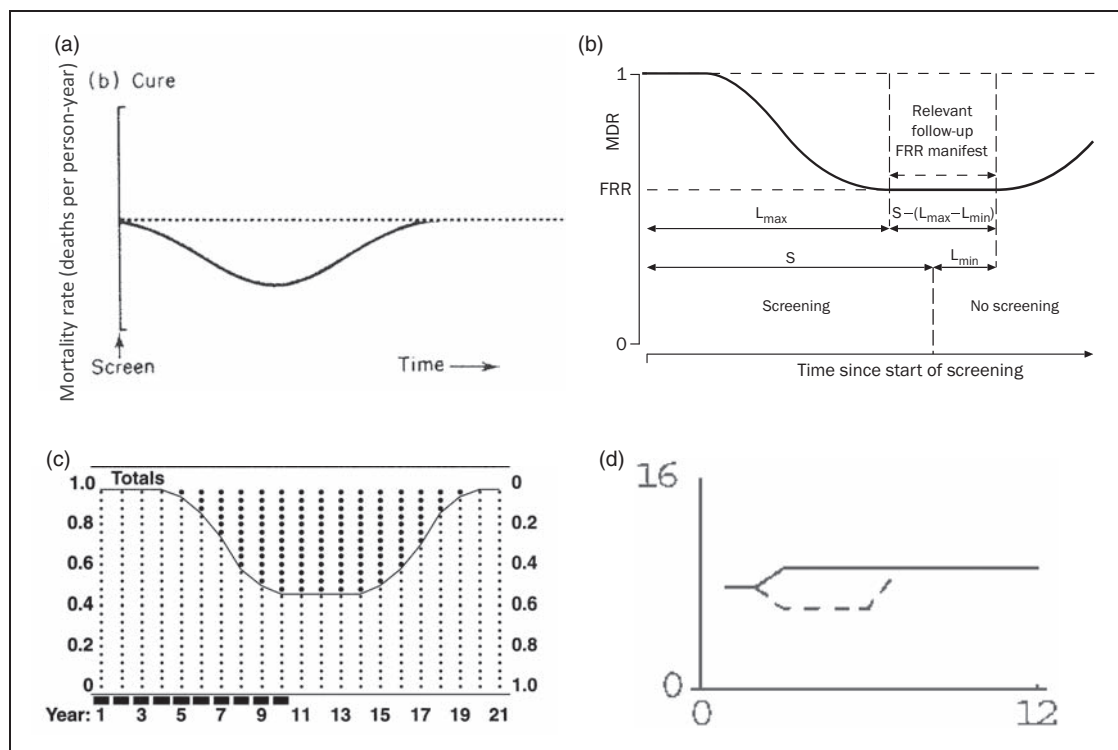


Figure 2. Hypothetical rate-ratio curves, as depicted in textbooks and other publications. (a), (b) and (d) invoke the bathtub shape, while (c) derives it from the convolution of the separate effects of 10 annual rounds of screening. The 4 panels (a–d) correspond to references #1, 2, 3 and 5.

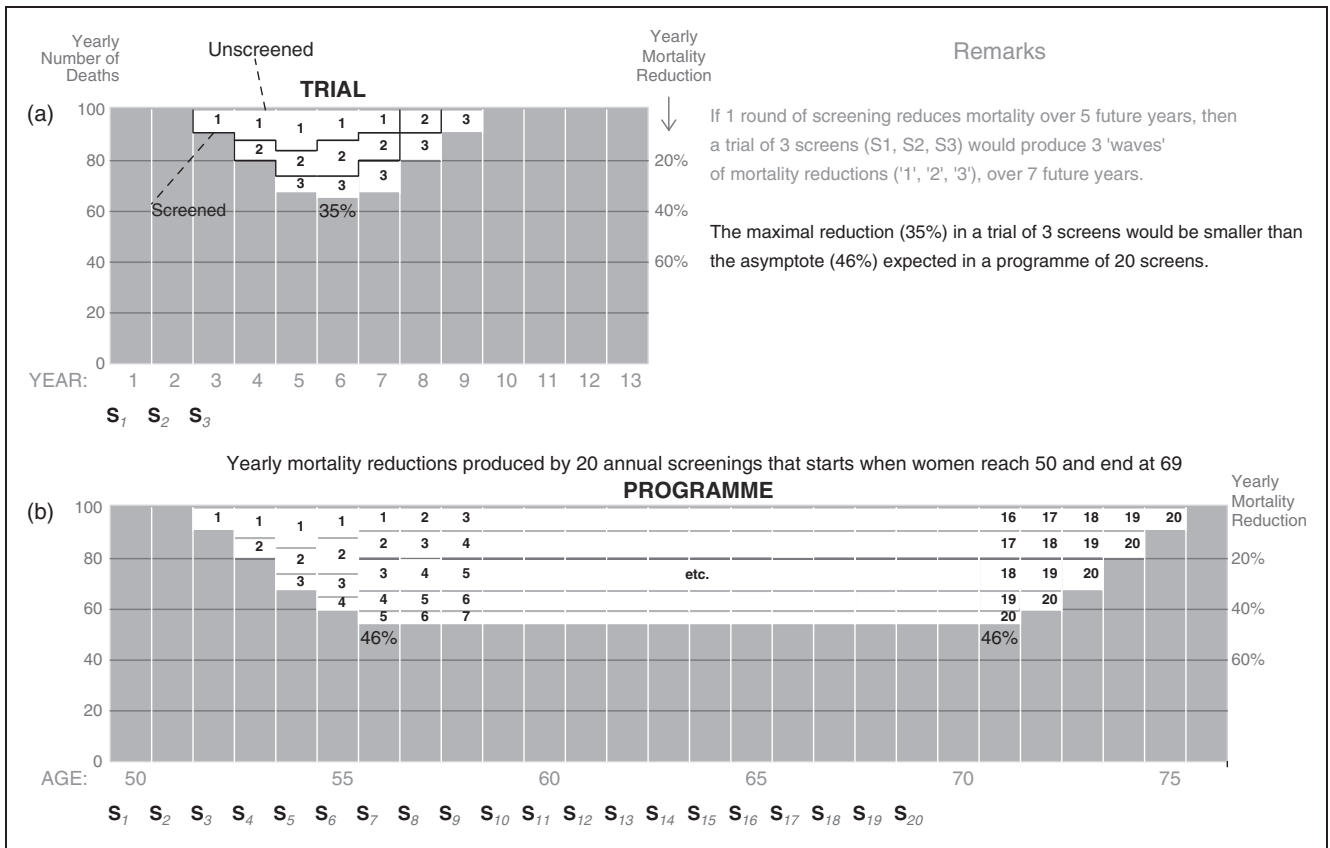


Figure 3. The 35% maximal mortality reduction produced by a (hypothetical) trial of 3 annual screenings (a) does not necessarily reach the 46% asymptote produced by a programme of 20 annual screenings (b), particularly if the impact of each round is spread over more than 3 years. Shown in (a) is a hypothetical trial of 3 annual rounds of cancer screening (S₁, S₂, S₃) compared with no screening. The depth of the white rectangle in each year represents the percentage mortality reduction, relative to an unscreened group, for the year shown on the horizontal axis. Annual mortality reductions produced by screening only begin to be expressed in year three (when the first effect of S₁ is discernible); they are greater in years 4 and 5, reaching a maximum of 35% in year 6 (when the combined effect of S₁, S₂ and S₃, denoted by '1', '2' and '3' respectively, is maximal); in year 7 the combined effects begin to wear off, and the mortality in the screening arm begins to revert to that in the non-screening arm; in year 9, the last effect of S₃ is discernible. Thus the maximum reduction is 35% and it would have been greater than if screening had not been discontinued at year three. By contrast the average effect of screening over the 13 years of observation (the metric used by task forces) would be 12%. Shown in (b) is a hypothetical screening programme with annual screening beginning at age 50 and continuing until age 69, compared with no screening. Again, the depth of the white rectangle represents the percentage mortality reduction for the age shown on the horizontal axis. The mortality reduction reaches 46% at age 56 and is maintained at that level for many ages – until three years after the last screen when it starts to decrease again.

16 years, much longer and deeper than the width and the maximum depth of the reductions seen in a trial. As is seen by comparing panels (a) and (b), the nadir seen in a trial usually underestimates the asymptote in a programme. However, even if all that was required was to measure the nadir carefully by, for example, smoothing¹¹ to avoid overestimation resulting from the yearly statistical fluctuations, few trials have provided yearly data that might allow this to be done. Instead, the universal practice is to report an averaged reduction, computed over the entire follow-up time of the trial. Because this average includes the almost-zero reductions outside the affected time window, it is even smaller than the nadir, and thus an even greater underestimate of the programme asymptote of interest.

The report of the National Lung Screening Trial (NLST)¹², presented in Table 1, illustrates the difference between evidence based on a few screenings which produce some reductions in lung cancer mortality over a short time-window, and the level of data needed to project what would occur if 50-year-old people were offered regular screenings until they reached age 69. The deficit of 88 deaths in part (a) of the table is clearly statistically significant, and expectedly shows that 3 CT screenings would reduce lung cancer mortality by some non-zero amount. But the pattern of the yearly deficits in part (b) is incomplete and puzzling. If the 42% deficit in year 6 were to be followed by two similarly large deficits in years 7 and 8, then it would suggest that a screening programme could achieve an asymptote twice the size of the reported

Table 1. Lung cancer deaths in the NLST report.

(a) What was reported in NEJM (August 4, 2011)									
Follow-up Year	1	2	3	4	5	6	7	8	All
Screens	S1	S2	S3						
X-ray Arm									442
CT Arm									354
Reduction									20%
(b) Year-specific data extracted from graph in that report									
X-ray Arm	37	68	82	95	84	73	4	?	
CT Arm	31	57	67	84	72	42	3	?	
Reduction	16%	16%	18%	12%	14%	42%	?	?	

20% reduction. If instead the deficit in year 6 were to be followed by diminishingly small deficits of the sizes seen in years 1-5, it would suggest that the deficit in year 6 was merely a statistical aberration, and that the asymptote in a programme would be much smaller than the reported 20%.

The additional numbers of cancer deaths in years 7 and 8 were unknown at the time of the report, because the causes of the deaths that occurred in these latter years had not all been adjudicated by the time the overall mortality reduction became statistically significant. This is a striking example of the distinction between getting a statistical significant result with just 3 screens, and providing evidence on what a screening programme (of possibly many more screens) would achieve.

The importance of using time-specific rates to pursue the asymptote of the curve was also highlighted in a recent review of screening trials in colon and prostate cancer. Whereas the overall reduction in the largest colon trial been reported to be 20%, the re-analysis, which took account of the timing of, and interruptions in, screening, found that an uninterrupted programme would yield reductions with an asymptote of 40%.³ In screening trials for prostate cancer, where the time lag between screening and when the mortality deficits manifest are even longer, the deficits produced by the first screen would not be expected for at least six years; however the majority of the follow-up has only extended to about year 11 in the European Randomized Study of Screening for Prostate Cancer (ERSPC).¹³ A re-analysis¹⁴ showed that the reductions only began in year 7, and reached an asymptote of approximately 50% by year 12. One commentator¹⁵ put it well: “perhaps a better summary of the European trial result is not the 20% overall reduction in prostate cancer mortality, but the combination of no reduction in the first seven or so years and a reduction of about 50% after 10 years”.

Several task forces have examined screening programmes for breast, lung, colon and prostate cancers. Although their stated purpose was to estimate what a sustained programme would do, all of the meta-analyses they used merely averaged the overall reductions seen in

different trials. Thus they all greatly underestimated the asymptotes that would characterize the programmes they considered.⁴

A few authors have explicitly dealt with the delay, either by using the hazard ratio from a certain time point onwards¹⁶, or (in those trials with a sufficiently long duration of screening), by ‘letting the data speak for themselves’ as to when the asymptote begins.^{2,13}

An alternative metric

An alternative approach, that indirectly addresses the asymptote and directly acknowledges the time-pattern of the reductions produced by a limited number of rounds of screening, is to examine the mortality impact *only in cancers diagnosed during the screening period*. This avoids the dilution, which Baker⁵ refers to as “post screening noise”, described above: cancers that arise long after the screening is discontinued could not have been affected by the screening carried out in the trial. In one version¹⁷ of this alternative approach, where the cumulative incidence of cancers deaths - in those diagnosed in this screening period - in the two study arms are compared, it is assumed that there is no over-diagnosis in the screening arm. The other version¹⁸ avoids having to make this assumption by using the number of cancers that were diagnosed in the non-screening arm during the screening period. The efficacy of the 3 rounds of CT screening is then determined by calculating the ‘deficit’ of (442-354 =) 88 cancer deaths, and expressing this 88 as a percentage, not of 442, but of the number that *could possibly have been helped* by screening (the 88 who were, and the xxx whose cancers, despite being diagnosed in the screening period in the screening arm, proved fatal nevertheless). Unfortunately, as of the time of writing, this number xxx is not known.

The approaches described above do not allow projections to be made for a programme that uses a *different spacing* of screening examinations than was used in a trial. We therefore here describe some (necessarily-model-based) that do. This round-by-round approach also makes it possible to deal with trials in which the nadir may not have reached the asymptote.

3. Projecting the reduction patterns that would be produced by different regimens from those used in trials

Approaches

Because a trial usually does not contain sufficient rounds of screening, the nadir observed in it would underestimate the asymptote expected in a sustained programme with the same spacing of screenings. Thus, modeling assumptions are required to extrapolate from a trial of say 3 annual screens to a programme with say 20 annual screens. The ‘round by round approach’ we have described in Figure 3 can also be immediately applied to programmes with different durations and spacings (eg. 20 annual screens versus 10 biennial screens).

Several projections of the mortality reductions due to cancer screening have been based on extensive modeling of the natural histories of cancers and how their progress is altered by earlier detection and therapy. Many of these efforts^{19–21} have also quantified the associated costs and use very sophisticated simulation modeling to examine the impact of prevention, screening, and treatment on cancer incidence and mortality at the population level. These approaches usually require a very large number of parameter inputs, obtained from diverse data sources (such as trials, registries and surveys).

We first illustrate a round-by-round approach, using the model proposed by Hu and Zelen.¹⁰ Previously, it has mostly been used for planning early-detection trials, including the recent NLST, where the yearly numbers were aggregated for the power calculation for the interim and ultimate statistical tests performed during and at the end of the trial. We use it here to generate and display the rate ratio curve proposed in Section 1, to show the projected timing, magnitude and duration of the yearly reductions in a *programme* (the yearly numbers that the software aggregates for power calculations do not appear to have been previously used for this purpose). Hu and Zelen model the mortality in each year under the screening and no-screening scenarios via a total of seven parameters (see Figure 4) quantifying the sensitivity of the screening test, the natural (and altered) course of cancer from initiation to normal clinical diagnosis and post clinical diagnosis.

Illustration

As sufficient information to fit new parameter values has not yet been extracted from the completed NLST, we will use some modifications of the input values²² used to plan the trial. Rather than use the FORTRAN software the trial statisticians used to implement the Hu-Zelen integrals, we re-programmed them in R. The only modifications we made were to two of the input parameters, to better represent how the cancer deaths are averted. In the planning, the authors assumed the ‘average’ CT sensitivity would be 85%, and that those whose cancers were detected by screening would have their (counterfactual)

post-clinical-diagnosis survival altered from an exponential distribution with a median 1.53 or 1.74 years to one where the median was 2.42 or 2.21 years: (the planning calculations assumed that all would eventually die of their cancer; moreover, there was no possibility of a ‘cure’, unless by a ‘cure’ one means that one dies of another cause). Instead, in light of the very rapid progression of many lung cancers, and the possibility of over-diagnosis, we assumed that the ‘real’ sensitivity was much less, and that the possibility of cure (rather than a very short extension of a few months of life) was confined to subgroup of screen-detected cancers; the remainder, even if detected by screening, would continue to have virtually the same mortality rates as their counterparts who were not screened. Thus, we set the ‘sensitivity’ at 25% rather than 85%, and the median survival of 30 years (‘cure’) for those whose otherwise fatal cancers were found at a curable stage.

Figure 4(a) shows the resulting 35-year projection for a programme of 20 annual screenings. With the exception of the slightly unrealistic (but numerically inconsequential) pattern at the front end (see below), the rate ratio curve, and its complement the reduction curve, resemble the anticipated bathtub-shape presented in Figure 1. The curve stays close constant for the middle part where there was sustained screening, and it gradually tails off after screening was stopped. The ‘excess’ deaths after years 25 are a consequence of the assumed exponential survival model in which cancer deaths are merely delayed, not averted – in keeping with the corresponding pattern shown in version (b) of the Figure in Morrison’s textbook.

Figure 4(b) shows the projection for a biennial programme; it is a little shallower than the annual one, but the reductions persist for almost the same duration. The oscillations in the ‘round by round’ waves are more prominent than in (a), and reflect the local effects of variations in the progression rates of different cancers together with the intra-individual variability in their stages at each examination time. The considerably smaller mortality reductions than in (a) emphasize the fact that two year screening intervals allow many more lung cancers to progress to the incurable stage in the interim.

Possible reasons why the early portion of the projected curve does not show the anticipated time lag more clearly may include (i) the numbers of cancer-specific deaths are expected to be very small in the first few years, which lead to large uncertainty in the early portion of the rate ratio curve; (ii) the exponential form, assumed for the sojourn time distribution, does not take into account the time lag between screenings and their induced mortality reductions, (iii) the assumption of independence between an individual’s sojourn time and their post-clinical diagnosis survival time: we would expect a strong correlation, that is, a relatively fast-growing cancer would be aggressive both pre- and post-detection; and (iv) the mortality rates do not explicitly accommodate cures from cancer nor deaths from other causes.

In order to deal with these front-end and back-end issues, considerably more refinements would need to be incorporated into the model, such as stage-specific sensitivities, transition rates, and survival distributions, as well as age-specific competing risks. While Zelen and colleagues, and other CISNET investigators, have indeed incorporated such refinements, they now face the reality of having to deal with the over-diagnosis that accompanies the newer screening tools, and the added model complexity and uncertainty. Instead, we are currently exploring a minimalist model that focuses only on the mortality reductions.

Conclusion

Unlike therapeutic trials in patients, cancer screening trials in asymptomatic persons generate mortality reductions that can only manifest several years after the onset of screening. The often reported single-number cumulative mortality reduction, in either a trial or a meta-analysis of trials, is of limited use in projecting the timing, duration and magnitude of the mortality reductions that would be expected from a sustained screening programme, of longer duration and possibly with a different screening regimen.

Instead, we propose using a rate ratio curve, and its complement, the mortality reduction curve, to address the mortality impact (timing, magnitude, and duration) of a screening programme. This curve is easy to interpret, as it shows when reductions begin, how big they are, and how long they last. We illustrate, using an existing model, how such rate ratio curves could be computed and how it is possible to quantitatively compare the impact of different screening regimens over the appropriate time-window.

Our message is two-fold: we (1) recommend against using one-number summaries to deduce the yearly mortality reductions expected from a sustained screening programme, and (2) call on trialists to report necessary time-specific mortality data to allow the appropriate computation of rate ratio curves that allow the mortality impacts of different screening programmes to be compared over the appropriate time horizon.

Funding

This work was funded by the Canadian Institutes for Health Research.

References

- Morrison AS. *Screening in Chronic Disease*. New York: Oxford University Press, 1985.
- Miettinen OS, Henschke CI, and Pasmantier MW. Mammographic screening: no reliable supporting evidence? *Lancet*, 359:404–405, 2002. <http://image.thelancet.com/extras/1093web.pdf>.
- Hanley JA. Analysis of mortality data from cancer screening studies: Looking in the right window. *Epidemiology* 2005;16:786–790.
- Hanley JA. Measuring mortality reductions in cancer screening trials. *Epidemiologic Reviews* 2011;33:36–45.
- Baker S, Kramer BS, Prorok PC. Early reporting for cancer screening trials. *Journal of medical screening* 2008;15:122–129.
- Shapiro S. Evidence on screening for breast cancer from a randomized trial. *Cancer* 1977;39:2772–2782.
- Zucker DM, Lakatos E. Weighted log rank type statistics for comparing survival curves when there is a time lag in the effectiveness of treatment. *Biometrika* 1990;77:853–864.
- Self SG, Etzioni R. A likelihood ratio test for cancer screening trials. *Biometrics* 1995;51:44–50.
- Self SG. An adaptive weighted log-rank test with application to cancer prevention and screening trials. *Biometrics* 1991;47:975–986.
- Hu P, Zelen M. Planning clinical trials to evaluate early detection programmes. *Biometrika* 1997;84:817–830.
- Hanley JA, McGregor M, Liu Z, Strumpf EC, Dendukuri N. Measuring the Mortality Impact of Breast Cancer Screening. Under review, 2013.
- The National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine* 2011;365:395–409.
- Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized european study. *New England Journal of Medicine* 2009;360:1320–1328.
- Hanley JA. Mortality reductions produced by sustained prostate cancer screening have been underestimated. *Journal of Medical Screening* 2010;147–151.
- Law M. What now on screening for prostate cancer? *Journal of Medical Screening* 2009;16:109–111.
- Caro JJ. *Screening for breast cancer in Quebec: estimates of health effects and of costs*. Montreal: CETS, 1990.
- Tabár L, Vitak B, Chen TH, Yen AM, Cohen A, Tot T, Chiu SY, Chen SL, Fann JC, Rosell J, Fohlin H, Smith RA, Duffy SW. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology* 2011; 260: 658–663.
- Yankelevitz DF, Smith R. Understanding the Core Result of the National Lung Screening Trial. *N Engl J Med* 368;15 april 11, 2013, 1460–1461.
- Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al. Effects of mammography screening under different screening schedules: Model estimates of potential benefits and harms. *Annals of Internal Medicine* 2009;151:738–747.
- Mandelblatt JS, Cronin KA, Berry DA, Chang Y, de Koning HJ, Lee SJ, et al. Modeling the impact of population screening on breast cancer mortality in the United States. *The Breast* 2011;3:S75–81.
- Heijnsdijk EAM, Wever EM, Auvinen, et al. Quality-of-Life Effects of Prostate-Specific Antigen Screening. *N Engl J Med* 2012;367:595–605.
- National Lung Screening Trial Research Team, Aberle DR, Berg CD, Black WC, Church TR, Fagerstrom RM, Galen B, Gareen IF, Gatsonis C, Goldin J, Gohagan JK, Hillman B, Jaffe C, Kramer BS, Lynch D, Marcus PM, Schnall M, Sullivan DC, Sullivan D, Zylak CJ. The National Lung Screening Trial: overview and study design. *Radiology* 2011; 258: 243–253.