Screening for Breast Cancer: What Truly Is the Benefit?

O.S. Miettinen, MD, MPH, PhD

In an article in this issue of the CJPH, Hanley et al.¹ respond to the most recent review by the Canadian Task Force on Preventive Health Care aimed at quantification of the benefit from screening for breast cancer.²

These authors have much in common with the Task Force. They, like the TF: are concerned to quantify the aggregate benefit from the screening for some defined type(s) of population for which public policy is, or might be, in favour of the screening; think of this benefit for any such population in terms of the proportional reduction in mortality from the cancer; think that a measure of this reduction can be derived from the studies that have addressed the counterpart of this reduction in experimental cohorts; are not concerned to judge whether the diagnostic work-ups and treatments in a given trial, even dating from a half-century ago, have relevance to practices at present and beyond; do not view it relevant to take note of the protocols for and practices of the diagnostic work-ups and treatments in those studies; and finally, do not think it necessary to judge the validity of the trials used for quantification of the mortality reduction, even if major problems of validity in them are well known.

In this aggregate of principles, the inattention to the well-known problems of validity in the trials is particularly surprising. Suffice it to note that while much emphasis is placed on the trials being randomized, even this feature of them has been prone to be invalid. Thus, S. Mukherjee, in his “biography of cancer”,³ points out and explains how the HIP trial was “instantly a logistical nightmare” (p. 295); and how, in this trial, “The unscreened group had been mistakenly *overloaded* with patients with prior breast cancer” (p. 297; italics in the original). He also points out and explains how “The [Canadian National Breast Screening Study] faltered, ..., by succumbing to the opposite sin: by selectively *enriching* the mammography group with high-risk women” (p. 299; italics in the original).

On the other hand, Hanley et al. diverge, profoundly, from the Task Force in the way they think of, and derive, the results from the trials; and in this, these authors take major exception also to the culture that the TF in this respect shares with the trialists themselves.

Different from the TF, Hanley et al. appreciate a fundamental truism that has been obvious to many others before them: that in trials on screening for a cancer, the proportional reduction in mortality from the cancer – in their screened subcohorts, insofar as the reduction occurs at all – cannot be constant over successive intervals of time after the screening’s initiation; that it is initially nil, then increases and later declines, and ultimately totally vanishes. Despite these understandings by others, the TF, like trialists themselves, draws from any given trial a single-value result for the

proportional reduction in mortality from the cancer; and this is done with no regard for the arbitrarily set durations for the screening and the follow-up for deaths from the cancer, the follow-up starting from the initiation of the screening.

These authors appreciate, also, that the proportional reduction in mortality from the cancer in a screened cohort reaches, under certain conditions, its maximal, asymptotic level, which prevails for a certain duration even after the screening’s discontinuation. They do not, however, elaborate on this asymptote, nor do they specify where this has been done (which is their reference no. 4).

Hanley et al. take a keen interest in this experimental asymptote. They take this asymptote to represent what they, like the TF, want to know; that is (to say it again), the proportional reduction in mortality from the cancer resulting from the screening’s introduction, after the experimentations, as an available service, to a population (dynamic rather than cohort-type).

They therefore set out to assess this asymptotic level of the reduction on the basis of five of the six trials that the TF made use of, the five from which they could derive the mortality ratios specific to particular subintervals of time since the screening’s inception. Their core “finding” was that the asymptotic reduction in those experimentally screened cohorts was at least 40%, and they took this to be the reduction also for a population to which the screening has been, or might be, introduced as a service – thus estimating the proportional reduction in deaths from the cancer in such a population to be much higher than what the TF inferred from the same trials.

Sadly, Hanley et al., like the TF, were mistaken in their goal, i.e., in what they set out to quantify. There is no need for “measuring the mortality impact of breast cancer screening” for the entire population for which it has been, or might be, made available.

There is no population-level benefit from the availability of screening for breast cancer in any meaning other than the sum of the *individual benefits* from the availability of this service to the women constituting the population at issue. These individuals are not concerned with the epidemiological topic of the rate of mortality from the cancer in that population, nor with its subordinate, equally esoteric topic of proportional reduction in this mortality consequent to the screening having become available. The benefit to these individuals, if any, is an instance in which their undergoing a round of the screening leads to detection of a (latent) case of the cancer and the ensuing treatment results in cure of the cancer while otherwise – upon diagnosis prompted by the cancer’s clinical manifestations – the disease would no longer be curable. The *value*

Author’s Affiliation

Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC

of this individual benefit they understand to be quite individualistic, dependent on their age at the time, among other factors.

The proportional reduction in mortality from the cancer in a screening-eligible population, while thus irrelevant to the members of the population, is also unrealistic to try to quantify, whether in terms of the Hanley et al. approach or that of the TF. One reason for this is the unjustifiability of the premise in these attempts that once the screening is available, the eligible women avail themselves of it, for decades, at the same rate as those in the trials did, for a few years.

The interest that Hanley et al. took in the asymptote of the proportional reduction in mortality from the cancer, as it on certain conditions (which they did not appreciate) is estimable from screened cohorts, would have been justifiable, and highly so, for a reason very different from that which motivated them.

The asymptotic level of the proportional reduction in mortality from the cancer in screening experiments equals something that is critically important to individual women in the population at issue. It equals the proportional reduction in the cancer's rate/probability of incurability, or in its case-fatality rate, attendant to its detection under the screening, when not considering whether the diagnosis is due to the screening or to symptoms emerging between two successively scheduled rounds of the screening. (This is explained in the authors' reference no. 4.)

In this individual-centered, clinical-type framework of thought, the population-level benefit – the sum of the individual benefits (cf. above) – from the screening's availability in a given span of calendar time (e.g., the first year of its availability) is, in plain numerical terms (when not accounting for the valuations of the cures), the total number of otherwise incurable cases that, in the population in that period, were cured by screening-afforded early treatments. This is the period-specific number of detections of the cancer consequent to the screening multiplied by three probabilities: the probability of the case being one of a genuine, life-threatening cancer (rather than overdiagnosed as such); the probability of a screen-diagnosed genuine case of the cancer being incurable by treatment delayed to the time when the cancer already would be clinically manifest; and the probability of undelayed treatment upon screen-diagnosis being curative of such an other-

wise incurable case (i.e., the proportional reduction in incurability addressed above, though adjusted for it to be specific to screen-diagnosed cases).

All of the clinical-type probabilities in this calculation of the population-level benefit from the screening – should the latter be of interest – are relevant in themselves: they are germane to *knowledge-based screening* for breast cancer and for women's decisions to avail themselves of it (while the proportional reduction in the rate of mortality from the cancer, in the population at issue, is not; cf. above).

Laudably, Hanley et al. set out to help correct a major flaw in the still-common way of thinking about, and estimating, the magnitude of the benefit from screening for breast cancer, the benefit as it takes place in trials on the screening, this flaw being a major reason why the extensive research on this topic has resulted in a very high degree of confusion and controversy about the extent of the benefit in those trials, and secondarily about it in actual practice of the screening. These authors make a compelling case for the need to correct this flaw, even though this point of theirs is not new but only routinely ignored. They also call attention to, and illustrate, an alternative measure of the benefit in those trials but, regrettably, fail to grasp the true meaning of this measure, as they too do not proceed from tenable, genuinely first principles.

I remain almost as pessimistic about progress in the research as I have been before⁴ – except if the CJPH should now proceed to arrange for public discourse aimed at correcting the various prevailing, ingrained anomalies of the research culture surrounding screening for breast cancer, among other types of cancer.

REFERENCES

1. Hanley JA, McGregor M, Liu Z, Liu Z, Strumpf EC, Dendukuri N. Measuring the mortality impact of breast cancer screening. *Can J Public Health* 2013;104(7):e437-e442.
2. The Canadian Task Force on Preventive Health Care. Recommendations on screening for breast cancer in average-risk women aged 40-74 years. *Can Med Assoc J* 2011. DOI: 10.1503/cmaj. 110334.
3. Mukherjee S. *The Emperor of All Maladies. A Biography of Cancer*. New York, NY: Scribner, 2010.
4. Miettinen OS. Screening for a cancer: A sad chapter in epidemiology. *Eur J Epidemiol* 2008;23:647-53.