Limitations of the Canadian National Breast Screening Studies

By: Jay Baker, MD, FSBI, Dana Smethermen, MD, and Jessica Leung, MD, FACR, FSBI, and the Screening Leadership Group

Of the randomized controlled trials (RCT) designed to study screening mammography, the Canadian National Breast Screening Study (CNBSS) is certainly the most problematic. The CNBSS, which took place from 1980–1985, is actually two separate studies. CNBSS1 included approximately 50,000 volunteer women age 40–49, and determined the mortality benefit in the experimental group assigned to annual screening mammography plus clinical breast exam (CBE) versus the control group assigned to usual care (1). CNBSS2 had almost 40,000 volunteer women age 50–59, and compared the benefit of annual mammography plus CBE to yearly CBE alone (2).

From the time the results were first published in 1992 (1,2) and again in follow-up in 2000, 2002 and 2014 (3-5), the CNBSS has been controversial, because it is the only RCT that found essentially no decrease in mortality associated with an invitation to screening. In fact, among women in their 40s at the 7-year report, there were 36% more deaths due to breast cancer in the screening group (1).

There are a number of explanations for these counterintuitive findings, most of which relate to vulnerabilities and shortcomings in the execution of the study. The problems in the study execution have been well-documented, and include flaws in the randomization process, lack of statistical power, non-generalizable results, poor quality imaging, suboptimal interpretation, and inconsistent threshold for interpretation (6-9).

The flaws in the randomization process principally arose from three features. First, unlike all other RCTs, potential participants in the Canadian trials initially underwent a careful physical exam. Second, women with physical exam findings including palpable lumps, skin or nipple retraction, and even palpable axillary adenopathy were not excluded from this “screening” trial (10). Finally, the randomization was unblinded and decentralized. Because almost 80% of women with advanced palpable cancers were assigned to the screening arm in the first round of the study, there has been speculation that concerned clinicians did not follow the randomization process, but rather “allocated” some symptomatic women to the study group so that they would get a mammogram. While there is no proof that this occurred, there is circumstantial evidence that it did (11). Moreover, whether the imbalance was due to intentional tampering or occurred by chance alone, the net effect is the same, i.e., a failure to produce two equal cohorts of patients for comparison.

Other problems also contribute to the controversial nature of the study. Although the average 5-year survival for women in the United States and Canada diagnosed with breast cancer in the 1980s was 75–80%, women in the control arm of CNBSS1 had a better than 90% 5-year survival (9). This is likely due to the fact that the study subjects were volunteers and likely to be healthier on average than the average Canadian woman.
Thus, it was a greater challenge for screening to sufficiently improve outcomes for women in the study arm and show a statistically significant advantage of early detection.

The Canadian trial was criticized at the time of the trial for poor quality mammography, even compared to mammographic imaging of that era (6,7). In order to reduce radiation dose, mammography for the trial was performed without the benefit of scatter-reducing grids that were already in routine use. Standard imaging for much of the trial utilized a straight lateral view and not a mediolateral-oblique view which images more tissue. The combination of poor quality imaging and the investigators’ resistance to taking corrective action led two advisors to resign in protest (10).

Finally, technologists participating in the trial received no special training in performing mammograms. Radiologists new to mammography also received no training in interpretation. After a radiologist’s recommendation for biopsy, a surgeon ultimately decided whether to move forward with a biopsy. Fully 25% of recommended biopsies were ultimately not performed (10).

All told, the Canadian trials were a missed opportunity to measure the efficacy of mammography and clinical breast examination in women ages 40–49, and mammography alone in women ages 50–59. The CNBSS trials are an excellent demonstration of the need to carefully consider all facets of a large screening trial before accepting its results as scientifically valid. The numerous design and execution flaws described above explain in large part why the results of the Canadian National Breast Screening Study are dramatically different than all other RCTs.

References: