Submitted Electronically

United State Preventive Services Taskforce

report states that, "…there is broad international consensus that GDGs should be multidisciplinary, with representation from all key stakeholders (ACCF and AHA, 2008; AGREE, 2003; NICE, 2009; SIGN, 2008)" (1). The USPSTF has failed on this account. The USPSTF panel did not include a single expert in breast cancer diagnosis or care. This is unreasonable for a guideline with such important implications, and could have been easily achieved if proper planning had prevailed. Failure to include knowledgeable experts hampered the ability of the USPSTF to understand and review the evidence. The IOM report suggests that such guideline development cannot assess the evidence in the same way that a multidisciplinary group can. Nuances are missed and data is sometimes misunderstood. This

The USPSTF relies almost entirely on randomized trial data to assess the mortality reduction from screening mammography. Given the existence of randomized trials, albeit using obsolete mammography technology in an era when good systemic breast cancer treatment also was unavailable, it is reasonable to cite the trials as demonstrating the *existence* of mortality reduction as a benefit of screening. However, there are numerous reasons why these trials (or any trials) underestimate the *magnitude* of mortality reduction, magnitude being of great consequence in assessing benefits versus harms. Most important among these are: [a] non-compliance in the study cohort (women counted in the study group who do not undergo screening dilute the observed benefit); and [b] contamination of the control group (women in the control group who undergo screening or diagnostic imaging for signs or symptoms) outside of the study actually do experience the mortality reduction of screening but are counted in the control group, thereby further diluting the observed benefit).

The magnitude of the effects of both [a] and [b] is readily demonstrated by cohort (incidence-based mortality) studies and case-control studies, within which the magnitude of mortality reduction observed for invitation to screening (invited versus not invited) is very substantially lower than that observed for exposure to screening (screened versus not screened). This is because the invitation to screening group differs from the exposure to screening group by the effects of [a] and [b]. This has been documented by EUROSCREEN systematic reviews of organized screening programs in Europe: for incidence-based mortality studies, invitation to screening yielded a 25% mortality reduction (RR 0.75, 95% CI 0.69-0.81), while exposure to screening yielded a 38% mortality reduction (RR 0.62, 95% CI 0.56-0.69); for case-control studies after adjustment for self-selection, invitation to screening yielded a 31% mortality reduction (RR 0.69, 95% CI 0.57-0.83), while exposure to screening yielded a 48% mortality reduction (RR 0.52, 95% CI 0.42-0.65) (2).

The USPSTF recommendation statement pays lip service to these data, acknowledging in a single sentence only the invitation to screening (invited versus not invited) data. In the United States, however, screening mammography is opportunistic and not centrally organized, so the effectiveness of screening is based on whether an individual woman actually attends screening. Non-compliance and contamination, meaningful in the context of invitation-to-screen trials, are meaningless in the context of opportunistic screening. Therefore, for the United States, mortality reduction is best measured, and screening guidelines best based, on exposure to screening (screened versus not screened). As stated above, the EUROSCREEN pooled data from 20 incidence-based studies showed a 38% mortality reduction for screened women compared to unscreened women and from 8 case-control studies, a 48% mortality reduction for screened compared to unscreened women after adjustment for self-selection. There also are robust incidence-based mortality-study data for exposure to screening (screened versus not screened) from within North America, based on service screening data from the organized

CISNET estimates of lives saved are more relevant than RCT data not just because they are based on the performance of modern screening mammography in the United States, but also because they reflect benefit to women actually screened compared to unscreened women, while RCT data reflect benefit to women invited to screening compared to uninvited women. Since Draft Table 2 lists "harms" to women actually attending screening based on modern U.S. data, Table 1 should describe the benefits to women screened based on modern data, not simply women invited to screen in outdated RCTs. To do otherwise would be to make the same error the Task Force made in their 2009 recommendations of confusing number needed to invite (NNI) with number needed to screen (NNS) to prevent one breast cancer death.

Women Age 40-

74. The 2015 CISNET models estimate that annual mammography in the 40-49 age decade provides a 58.5% improvement in LYG compared to biennial screening.

USPSTF requested CISNET to model starting ages of 40 and 45 (Tables 10a, b). Data in these tables, however, erroneously show 28 data points that are identical for LYG and QALY for starting ages 40 (10a) and 45 (10b). This is erroneous based upon the differences of mortality reduction in the same tables. We are very concerned that incorrect information was used in formulating the "C' level recommendation for women ages 40-49, given that this apparent error was not noted by a single member of the Task Force or CISNET.

Women Age 40-49: "C" Recommendation

The USPSTF explains its "C" recommendation for women ages 40-49 by indicating that women should individually decide whether they will undergo screening based on an informed *personal* decision of whether the benefits of screening exceed the harms. If the final recommendation retains the "C" rating, however, the recommendation will likely limit patient choice, not empower it. The Affordable Care Act requires private insurers to cover screening tests with a USPSTF grade of "B" or above at no cost to the patient. There is no such requirement for screening tests with a "C" grade. If the draft recommendations are adopted as final, 17 million women ages 40-49 could be forced to make a financial decision about breast cancer screening and many will not be able to benefit from the shared decision making process with their physicians, as recommended by the Task Force.(6) We strongly believe that the USPSTF's rating should not become a barrier to a woman's access to care or limit her informed choice about breast cancer screening.

To be clear, the "C" rating is not dictated by the evidence; it is a value judgment based on the Task Force's opinion of what constitutes benefits and harms of mammography and its subjective weighting of the net benefits of screening for this population. There is ample support, not only in the evidence contained in these comments but also in the draft recommendations, on which thd p s Mя я þ has f "Cậ s ratiɀ r rce M al } Mos' r

ages 40-79. Failure to provide a comparison benefit analysis while providing harms data demonstrates bias and non-transparency. In addition, CISNET modeling of the UK Age Trial of annual screening in women 40-49 with 100% compliance and 13 years follow-up yielded a median mortality reduction of 28% (range 25% to 35%), in good agreement with the observed mortality reduction of 24% for screened versus unscreened women.

Although unstated by the USPSTF, their strategy is clearly to maximize lives saved per mammogram performed, not to save the most women's lives. While the USPSTF is careful to mask this decision in terms of benefits versus harms, it amounts to benefit (in terms of lives or life-years saved) versus cost, with number of mammograms performed as a surrogate for cost. In this vein, the CISNET modelers go so far as to explicitly list "number of mammograms" as a "harm" of screening (Collaborative Modeling, p. iv). This comes perilously close to resembling the "death panel" approach to health care access that critics of the Affordable Care Act (and general government involvement in health care administration) fear.

The USPSTF relied on an obsolete method for analyzing outcomes data for biennial (2-year) screening, incorrectly using a 1-year period instead of 2-year period for cancer ascertainment. Use of a 2-year ascertainment period, the current standard approach in the United States (7),

mammography registries that participates in the BCSC (9). These population-based data involve screening mammography from 1996-2006, representing actual United States practice.

Performance of Screening Mammography

Age	Exams	Cancer Detection Rate	Sensitivity	Positive Predictive Value
50-59	186,944	3.7 per 1000	77.3%	22.2%
60-69	116,362	4.9 per 1000	80.1%	29.3%
70-79	75,692	6.2 per 1000	80.4%	37.6%
80-89	23,409	7.9 per 1000	83.4%	40.7%
90-101	1,041	14.1 per 1000	93.8%	55.6%

USPSTF has chosen to refer to these events using the pejorative term of “false positives”, they are formally considered “incomplete”. The harms of false-positive results are discussed, concentrating on the psychological harms (anxiety/apprehension) that may occur after learning of the need for additional testing, especially when this involves a biopsy. However, the USPSTF recommendation statement omits current evidence indicating that there is no long-term anxiety and no measurable health utility decrement. The recommendation statement also omits current evidence indicating that screened women who experience false-

underlying cancer incidence trends, causing these studies to overestimate, often to a great extent, the true magnitude of overdiagnosis.(23)

The USPSTF recommendation statement estimates the frequency of overdiagnosis at 19% based on data reported from three of the randomized trials as provided in the accompanying Systematic Review by Nelson et al,

overdiagnosis, as an endpoint. There was no data collected beyond the intervention (screening) period on the frequency of breast cancer screening outside the trial in either invited or uninvited (control) groups. It is well known that service screening started in several Canadian provinces shortly after the end of the CNBSS trials. It is also well documented that the older population studied to estimate overdiagnosis in the Malmo trial had other-cause mortality that may have affected the accuracy of overdiagnosis estimates. Hence, these RCTs may be no more accurate in estimating overdiagnosis than non-RCT-based estimates.

CISNET modeling of invasive cancer overdiagnosis (Table 11) shows median values of only 2-3%, almost all overdiagnosis is attributed to DCIS. The extreme range of invasive cancer overdiagnosis of “1.4% to 24.9%” (note, no median value provided) undermines confidence in this assessment. The upper range occurred in a single model which assumed no temporal incidence change since the 1970s, in contrast to all other models and to known worldwide breast cancer incidence increases over the last 60 years.

In several parts of its text, the USPSTF recommendation statement emphasizes the harm of overdiagnosis as being most important among all harms, but there is no parallel emphasis on the frequency with which a screened woman may experience overdiagnosis. The only indication of frequency is buried in Table 3, at the very end of the recommendation statement, in which the USPSTF calculates that 20 women per 1000 who undergo a lifetime of screening will

detected cancers and screening-associated breast cancer deaths averted. These observations are true for both a single screening mammography examination and for multiple screening examinations over a lifetime from age 40 years.

The American Association of Physicists in Medicine has stated that “Risks of medical imaging at effective doses below 50 mSv for single procedures or 100 mSv for multiple procedures over short time periods are too low to be detectable and may be nonexistent(28). Predictions of hypothetical cancer incidence and deaths in patient populations exposed to such low doses are highly speculative and should be discouraged. These predictions are harmful because they lead to sensationalistic articles in the public media that cause some patients and parents to refuse medical imaging procedures, placing them at substantial risk by not receiving the clinical benefits of the prescribed procedures. For reference, the mean effective dose of the typical mammography exam (consisting of two views of each breast) is about 0.5 mSv, so even 40 years of annual screening exams does not approach the effective dose at which the the relationship between radiation exposure to the breast and cancer risk is a significant concern.

Harms of Not Screening

Women at any age who choose not to be screened, as well as women who are unable to be screened if constrained by personal cost considerations that may flow from the “C” recommendation for women ages 40-49 or the “I” recommendation for women above age 74, will forego both the benefits and harms of screening. However, malignancies still will be diagnosed in non-screened women, detected by palpation instead of screening. The USPSTF recommendation statement does not (but should) include discussion of the harms of cancer detection by palpation relative to the harms of cancer detection by screening mammography.

The harms analysis of “false-positive tests “(recalls) and "unnecessary" biopsy recommendations is seriously flawed. Harms of screening have not been compared to harms of non-screened women as stated. The harms analysis incorrectly assumes non-screened women will not undergo false positive tests (such as clinical physical exam), diagnostic breast imaging or “unnecessary” breast biopsies independent of screening. In fact, non-screened women frequently present to their clinician for diagnostic evaluation and biopsy of what eventually proves to be a benign finding. Barton showed that 23% of women (32% of women in their 40s) had a clinical visit for a breast problem in a 10 year period (29). In addition, 6.5% underwent an invasive procedure; nearly all proved benign (or in USPSTF terminology, “unnecessary”). More germane to screening harms judgment was the observation that screened women had significantly fewer symptomatic visits and subsequent work-ups. In addition, Blanchard (in a different study, showed annually screened women’s risk of undergoing

The draft recommendations of the USPSTF utilize data included in the article by Miller et al - Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial (24). Numerous criticisms of the two Canadian National Breast Screening Studies (CNBSS) were published at the time of release (31-36). Its use by the USPSTF in estimating overdiagnosis has been discussed in an earlier portion of this document. In addition, publications dispute the authors' conclusion that annual mammography in women 40-59 does not reduce mortality from breast cancer beyond that of physical examination or usual care, when adjuvant therapy for breast cancer is freely available. The major problems with the CNBSS studies, including inadequate quality of mammogra

References

1. Developing Trustworthy Clinical Practice Guidelines, Board of Health Care Clinical practice guidelines we can trust / Committee on Standards for Services, Institute of Medicine of the National Academies; R. Graham, et al. 2011.

2. Broeders M, Moss S, Nyström L, et al. The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. J Med Screen 2012; 19 Suppl 1:14-25.

3. Coldman A, Phillips N, Wilson C, et al. Pan-Canadian study of mammography screening and mortality

18. Zorzi M, Puliti D, Vettorazzi M, et al. Mastectomy rates are decreasing in the era of service screening: a population-based study in Italy (1997-2001). Brit J Cancer 2006; 95(9):1265-1268.

19. Spillane AJ, Kennedy CW, Gillett DJ, et al. Screen-detected breast cancer compared to symptomatic presentation: an analysis of surgical treatment and end-points of effective mammographic screening. ANZ J Surg 2001; 71(7):398-402.

20. Tosteson ANA, Fryback DG, Hammond CS, et al. Consequences of false-positive screening mammograms. JAMA Intern Med 2014; 174(6):954-961.

21. Ponti A, Lynge E, James T, et al. International variation in management of screen-detected ductal carcinoma in situ of the breast. Eur J Cancer 2014; 50(15):2695-2704.

22. Puliti D, Duffy SW, Miccinesi G, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. J Med Screen 2012; 19 Suppl 1:42-56.

23. Bleyer A, Welch HG. Effective Screening Mammography on Breast Cancer Incidence. N Engl J Med 2013; 368:679.

24. Miller AB, Wall C, Baines CJ, Sun P, To T, Narod SA. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomized screening trial. BMJ 2014; 348:g366 doi: http://dx.doi.org/10.1136/bmj.g366.

25. Marmot MG, Altman DG, Cameron DA et. al British Journal of Cancer (2013 108, 2205-2240

26. Esserman LJ, Thompson IM Jr, Reid B. Overdiagnosis and overtreatment in cancer. JAMA 2013; 310(8):797-798.

27. Flowers CI, O'Donoghue C, Moore D, et al. Reducing false-positive biopsies: a pilot study to reduce biopsy rates for BI-RADS 4A/B assessments through testing risk stratification and new thresholds for intervention. Breast Cancer Res Treat 2013; 139(3):769-777.

28. AAPM Position Statement on Radiation Risks from Medical Imaging Procedures. PP 25-A12/13/2011 through 12/31/2016. Available on-line at: http://www.aapm.org/org/policies/details.asp?id=318&type=PP&current=true

29. Barton MB, Elmore JG, Fletcher SW. Breast symptoms among women enrolled in a health maintenance organization: frequency, evaluation, and outcome. Ann Intern Med. 1999;130: 651-657.

30. Blanchard, Karen, et al. &rJET3oA MCI839[(-)]l48(>515.23 Tm2 4 273.h -0.05 319.73 Tm[(l)-6(i)5(n Tc[(56)4(en)4(r)-3e J

35. Kopans DB. The Canadian Screening Program: a different perspective. AJR Am J Roentgenol 1990; 155(4):748-749.

36. Yaffe, MJ “Correction: Canada Study.” J Natl Cancer Inst 85, no. 2 (1993): 94.

37. Tabar L, Gad A, Holmberg LH, et al. Reduction in Mortality from Breast Cancer after Mass Screening with Mammography. The Lancet, Volume 325, Issue 8433, 13 April 1985, pages 829-832.

38. Rose SL, Tidwell AL, Gujnoch LJ, et al.Implementation of breast tomosynthesis in a routine screening practice: an observational study. AJR Am J Roentgenol 2013; 200(6):1401-1408.

39. Skaane P, Bandos AI, Gullien R, et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. Radiology 2013; 267(1):47-56.

40. Friedewald SM, Rafferty EA, Rose SL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. JAMA 2014; 311(24):2499-2507.

41. Ciatto S, Houssami N, Bernardi D, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. Lancet Oncol 2013; 14(7):583-589.