Review of *Benefits and Risks of Screening Mammography for Women in Their Forties: a Statistical Appraisal*

Grace Yi Chen

There has been a controversy about whether or not to recommend regular mammography to women aged 40–49 years. Dr. Donald A. Berry did meta-analysis of mortality data from 8 randomized clinical trials (RCT) and evaluated their evidence about potential benefits and risks of screening. The authors discussed about the similarity, difference and quality of the 8 RCTs. Usually, a standard statistical analysis method: Mantel-Haenszel test could be used for mammography screening efficacy assessment. 18% reduction in breast cancer mortality on average was estimated using this method. The author proposed to use hierarchical Bayesian model which would account for the heterogeneity among populations. The estimated reduction in breast cancer mortality is similar with the one from Mantel-Haenszel test but has wider interval. Later the author discussed about the potential risk of doing screening. One problem is that the benefit of screening is not immediate. As the study shows that the 18% mortality reduction estimation on average is based on 15 years after randomization. There is also possibility for false-positive, false-negative findings as well as psychological burden. Clinicians should effectively communicate the benefits and risks to their patients. In the second paper, Dr. Daniel B. Kopans and Dr. Elkan Halpern has some disagreements with Dr. Berry on his analysis and conclusions. Dr. Berry also gave his response to the criticism.

I think Dr. Berry gave a good analysis and discussion about the potential benefit and risk for women aged 40-49 years doing mammography screening. The benefits of doing mammography screening depend on the development of breast cancer, the treatment, and the methods of screening. With the development of technology, nowadays we should have lower false-positive and false negative rates in breast cancer screening. This agrees with what Kopans said that we should recommend patients to have mammography with better quality. In addition, I learnt a lot from Berry and Kopans' discussion about the meaning of p-values and statistical method selection. Berry is a statistician famous for his achievement in Bayesian analysis so he would think using a Bayesian method here to be appropriate accounting for population heterogeneity. I think either using frequentist methods like GEE or Bayesian method should both work here. A statistically significant result could be clinically meaningless, but this might be up to the clinicians to consider. Moreover, doing meta-analysis is hard because we would like to include as many studies as possible but different studies might have different designs and implementations. Meta-analyses that include biased studies will be at risk of being misleading. In this case, we know that the randomization of Canadian trial lacks blinding, and the study is under powered. Including this study in the meta-analysis might lead to a biased conclusion.

Questions:

1. In the case of Canadian trial which lacks blinding and is under-powered, should we include it in the meta-analysis?
2. Berry mentioned that the efficacy result from a meta-analysis is usually on average, but he wants to relate it to an individual's decision-making. I am wondering would using

Bayesian models that account for population heterogeneity better than using mixed model or GEE with p-values for this purpose in general?