Why are estimates of overdiagnosis in routine breast cancer screening so profoundly different?

For ICSN conference 2017

Sisse Helle Njor

Dept. of Clinical Epidemiology

Aarhus University

Denmark

Eugenio Paci

ISPO Cancer Prevention and Research Institute

Florence

Italy

Matejka Rebolj

Centre for Cancer Prevention

Wolfson Institute of Preventive Medicine

Barts & The London School of Medicine and Dentistry

Queen Mary University of London

United Kingdom

Background

Several studies estimated overdiagnosis in mammography screening at ≥20%, whereas other studies reported far lower estimates. We investigated the role of study designs in explaining this difference.

Methods

We used data from Funen, Denmark (Njor et al. BMJ 2013) which suggested 1% overdiagnosis. The study included birth cohorts offered mammography ≤7 times at age 59-70 years and followed for 5-14 years thereafter to allow for observation of a compensatory dip. The background incidence was estimated using similar unscreened cohorts. We identified five published, highly-cited, studies that continued the observation of breast cancer incidence at post-screening ages but estimated overdiagnosis at ≥20% (Zahl et al. BMJ 2004; Jørgensen and Gøtzsche BMJ 2009; Jørgensen et al. BMC Womens Health 2009; Zahl and Mæhlen Tidsskr Nor Laegeforen 2012; Kalager et al. Ann Intern Med 2012). We recalculated estimates of overdiagnosis in Funen by applying each of these studies' designs, i.e. using the same age groups, calendar years and assumptions on the background incidence. This approach was suggested previously by Etzioni and Gulati (JNCI 2016).

Results

The estimates ranged from 15% overdiagnosis at the lower to 53% at the higher end, and were all remarkable close to the estimates from the original papers.

Conclusions

The different study designs produced a wide range of estimates of overdiagnosis on the same data. High estimates were due to either a) insufficient correction for background breast cancer risk, b) inclusion of women who could not contribute to compensatory dip, c) insufficient follow-up time or a mixture of these. This demonstrates that the highly cited studies using the plain age-period approach are based on unmet assumptions. A common methodological background in observational epidemiology could avoid invalid results, false controversies and confusion in the communication with women.

Words: 288