Benchmark for judging the balance of benefits and harms in screening

Akira Ohtsuru (Nagasaki University, Japan)

Eleven years have passed since Fukushima's thyroid cancer screening program began for children and young adults after the nuclear accident. This screening has caused marked potential overdiagnosis several ten times greater than without screening, as shown in the age-specific prevalence in the first round of screening[1)] and the incidence in the second round[2)]. Worldwide, it has become clear that ultrasound-based thyroid cancer screening results in the disadvantage of overdiagnosis, which far outweighs its benefits. Thus, recent guidelines no longer recommend screening[3)]. The thyroid radiation dose in Fukushima residents was much lower, suggesting a detectable excess of thyroid cancer due to radiation was unlikely[4)]. However, some Japanese experts have stated that the increase in thyroid cancer revealed in screening is not owing to overdiagnosis but merely clinical early detection of cancer[5)], confusing the present screening policy.

When planning new screenings, we should first consider adhering to screening principles[6)]. However, some cancer screening programs were initiated without those considerations. Appropriate measures are needed to evaluate screening that has already begun. However, a representative measure, for example a reduction of mortality, may take long years to become apparent, especially in thyroid cancer. Thus, a proxy measure to assess whether overdiagnosis or beneficial early diagnosis predominates would facilitate decisions on whether to continue or discontinue screening.

As a benchmark to allow early assessment when screening studies are initiated, the trajectory pattern of detection rates over time can be used. Suppose the natural history of cancer is likely to follow a linear growth pattern. When the increase of cancer detection rate in the first screening will be within approximately five times, and the trajectory returns to the original level in the subsequent screenings, we cannot deny the benefits of early detection. When the natural history is likely to be a pattern of growth arrest, and the trajectory of the screening results shows a 5-fold or more significant increase in the first round, or a marked increase over the original level in the subsequent rounds, marked overdiagnosis is suspected. Based on the natural history of growth arrest pattern[7,8)], the Fukushima thyroid examinations appear mostly overdiagnosed from the benchmark trajectory. In order to reduce the significant damage, urgent policy measures are necessary.

(References)

1) Katanoda K, et al. *Jpn J Clin Oncol* 2016;46: 284-6. 2) Ohtsuru, et al. *JAMA Otolaryngol Head Neck Surg* 145: 4-11, 2019. 3) Bibbins-Domingo K, et al. JAMA 2017;317:1882-7. 4) UNSCEAR 2020 report. 5) Suzuki S. *J Jpn Thyroid Assoc* 2021;12:139-48. 6) Screening programmes: a short guide WHO EU Regional Office 2020. 7) Midorikawa S, et al. *JAMA Otolaryngol Head Neck Surg* 2018;144:57-63. 8) Miyauchi A et al. *Surgery* 2019;165:25-30.