eAppendix

How to translate prevalence to incidence

Applying the model of prevalence pool in steady state[1] to latent duration in screening, we can obtain the following equations to translate prevalence to incidence, where the latent duration denotes the time from the date when cancer became detectable by screening and cytology to the date when it could be diagnosed in clinical settings without screening or the date of operation.

If a steady state population has the incidence I to enter a latent (prevalence) pool, the prevalence p, the mean latent duration D of the pool, and the incidence J to be diagnosed in clinical setting in the population, the following equations are met;

$$\frac{p}{1-p} = I \times D \text{ and } p = J \times D. \tag{A.1}$$

To show the proof, we let N denote the size of people in the population, P denote the size of the latent pool, and Δt denote any time interval. In steady state, the inflow and the outflow of the latent pool are balanced as:

$$\text{Inflow} = (N-P) \times \Delta t \times I = P \times \Delta t \times (1/D) = \text{Outflow}.$$

This proves the first equation of (A.1). The incidence J is defined as the number of outflow divided by person-time at risk in the population $N \times \Delta t$, which is described as

$$J = \frac{P \times \Delta t \times (1/D)}{N \times \Delta t} = p \times (1/D).$$

This proves the second equation of (A.1).

The unexposed comparison group in Fukushima, whose incidence rate is comparable with the Japanese annual incidence rate, can be assumed to be a steady state dynamic population. Therefore, like (A.1), the following equations are met;

$$\frac{p_0}{1-p_0} = I_0 \times D_0 \text{ and } p_0 = J_0 \times D_0,$$

where I_0, p_0, D_0, and J_0 denote the corresponding parameters among the unexposed group. Then, the prevalence ratio (PR) and the prevalence odds ratio (POR) between the exposed and unexposed groups in Fukushima can be described as

$$\text{PR} = \frac{p_1}{p_0} = \frac{p_1}{J_0 \times D_0} = \frac{p_1/D_0}{J_0} \approx \frac{J_1}{J_0} \tag{A.2}$$

and

$$\text{POR} = \frac{p_1/(1-p_1)}{p_0/(1-p_0)} = \frac{p_1/(1-p_1)}{I_0 \times D_0} = \frac{\{p_1/(1-p_1)\}/D_0}{I_0} \approx \frac{I_1}{I_0}, \tag{A.3}$$

where I_1, p_1, and J_1 denote the corresponding parameters among the exposed group. In the nearly equal signs of (A.2) and (A.3), we use the steady state assumption in the exposed group, and assume that D_0 is equal to the mean latent duration in the exposed group. If these

assumptions are met, PR and POR are equivalent to the clinical incidence ratio and the screening incidence ratio, respectively. When the mean latent duration in the exposed group is shorter than that in the unexposed group, PR and POR are conservative estimates of the clinical incidence ratio and the screening incidence ratio, respectively.

How to correct the effect of preclinical tumors

In screening, the latent prevalence pool can be divided into a group of true cancers (TCs) and a group of preclinical tumors (PCTs) which do not grow up to cancers. If the ratio of TCs to PCTs is assumed to be $(1-\alpha)/\alpha$, then the prevalence of TCs is given as

$$p_{\mathrm{TC}} = (1-\alpha)\, p.$$

We let $J_{1\mathrm{TC}}$ denote the clinical incidence of TCs in the exposed group in screening. We also let $J_{0\mathrm{TC}}$ denote the clinical incidence of TCs in the national data of thyroid cancer incidence (i.e. unexposed group). Note that the assumption of steady state dynamic population is likely to be met in the national data, and it is composed of only the TCs. Then, the ratio of $J_{1\mathrm{TC}}$ to $J_{0\mathrm{TC}}$ is described as

$$\frac{J_{1\mathrm{TC}}}{J_{0\mathrm{TC}}} \approx \frac{p_{1\mathrm{TC}}/D_{0\mathrm{TC}}}{J_{0\mathrm{TC}}} = \frac{p_1/D_{0\mathrm{TC}}}{J_{0\mathrm{TC}}} \times (1-\alpha), \qquad \text{(A.4)}$$

where $p_{1\mathrm{TC}}$ denotes the prevalence of TCs in the exposed group and $D_{0\mathrm{TC}}$ denotes the mean latent duration of the pool in the unexposed group. Note again that we use $D_{0\mathrm{TC}}$ as the mean latent duration in the exposed group had they not been exposed. In conclusion, we can obtain the incidence rate ratio of TCs by multiplying the coefficient $(1-\alpha)$ to the equation (A.2).

Reference

1. Greenland S, Rothman KJ. Measures of occurrence. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:32–50.