

From prediabetic individuals we have acquired and used data at each age during follow-up that they were disease-free (i.e. diabetes).<sup>1,2</sup> When an individual was free of disease reaching an age  $j$  at some point during follow-up, he or she constituted the population at risk for any age  $j$  (risk set,  $R_j$ ). When someone developed diabetes, deceased or was censored at age  $j$ , this individual was then removed from the risk set for age  $j+1$  and older. When someone entered the study at age  $j+1$ , this individual was added to the risk set for age  $j+1$ .<sup>3</sup> For instance, for lifetime risk at 45, hazards ( $h_j$ ), age-specific incidences ( $f_j$ ), cumulative incidences ( $F_j$ ), and survival probabilities ( $S_j$ ) were calculated according to the standard Kaplan-Meier methods for each age  $j$  (assuming  $F_{44} = 0$  and  $S_{44} = 1$ ):

$$h_j = e_j / R_j \quad (e_j = \# \text{ of events at age } j)$$

$$f_j = h_j \Delta S_j - 1$$

$$F_j = \sum_{i=1}^j f_i$$

$$S_j = 1 - F_j$$

So the cumulative incidence of the disease (diabetes) is  $F_j$ , which applies to individuals that survive through age  $j-1$ . The competing risk of death from another cause is not taken into account in this cumulative incidence.

Which means that when an individual decease it is counted as a withdrawal in which it is assumed that this individual has the same risk of the disease compared to the individuals that are alive at censoring. However, since these individuals die before age  $j$ , they have a zero future risk of the disease. This competing risk of death will result in overestimation of the lifetime risk.<sup>4</sup> Therefore, a separate survival function was used ( $U_j$ ) with death included as an event alongside diabetes in order to adjust for the competing risk of death. The adjusted incidence and true lifetime risk was calculated as follows:

$$f_j^* = h_j \Delta U_j - 1$$

$$F_j^* = \sum_{i=1}^j f_i^*$$

$$S_j^* = 1 - F_j^*$$

To calculate lifetime risks at the starting age 55, 65, 75 and 85 we used similar methods. We set the  $F_{T-1}$  and  $U_{T-1}$  to 0 for every index age  $T$  and used the original hazard ( $h_j$ ) to calculate  $U_j$  for  $j \leq T$ .

1. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 2007; 26(11): 2389-430.
2. Lloyd-Jones DM, Larson MG, Beiser A, Levy D. Lifetime risk of developing coronary heart disease. *Lancet* 1999; 353(9147): 89-92.
3. Meister R, Schaefer C. Statistical methods for estimating the probability of spontaneous abortion in observational studies--analyzing pregnancies exposed to coumarin derivatives. *Reprod Toxicol* 2008; 26(1): 31-
4. Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. *Br J Cancer* 2004; 91(7): 1229-35.