Web Appendix

Manuscript: Preventable fraction in the context of disease progression

Table of contents

eAppendix 1. Derivation of the preventable fraction using parameters in eTable 1

$$\frac{\Pr(Y=1)-\Pr(Y^1=1)}{\Pr(Y=1)}$$

$$= \frac{\{\Pr(Y=1|A=0)\Pr(A=0)+\Pr(Y=1|A=1)\Pr(A=1)\} - \{\Pr(Y^1=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)\}}{\Pr(Y=1|A=0)\Pr(A=0)+\Pr(Y=1|A=1)\Pr(A=1)}$$

$$= \frac{\{\Pr(Y^0=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)\} - \{\Pr(Y^1=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)\}}{\Pr(Y^0=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)}$$

$$= \frac{\Pr(A=0)\{\Pr(Y^0=1|A=0)-\Pr(Y^1=1|A=0)\}}{\Pr(Y^0=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)}$$

$$= \frac{\Pr(A=0)\{\Pr(Y^1=0,Y^0=1|A=0)-\Pr(Y^1=1,Y^0=0|A=0)\}}{\Pr(Y^0=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)}$$

$$= \frac{(1-\pi)\{\theta_{11}^0(\emptyset_{01}^0-\emptyset_{10}^0)+\theta_{01}^0\gamma^0\}}{(1-\pi)\{\theta_{11}^0(\emptyset_{11}^0+\emptyset_{01}^0)+\theta_{01}^0\gamma^0\}+\pi\{\theta_{11}^1(\emptyset_{11}^1+\emptyset_{10}^1)\}},$$

which is the same as the preventable fraction formula in the main text. Note that in the derivation in the main text, $\theta_{11}^0\emptyset_{01}^0$ corresponded to $\Pr(Y^1=0,Y^0=1|A=0)$ and $\theta_{01}^0\gamma^0$, to $\Pr(Y^1=\mathrm{Undefined},Y^0=1|A=0)$. Here, $\theta_{11}^0\emptyset_{01}^0+\theta_{01}^0\gamma^0$ corresponds to $\Pr(Y^1=0,Y^0=1|A=0)$.

eAppendix 2. Preventable proportion

In line with the article by Suzuki et al. [1], the notion of preventable proportion is defined by Yamada and Kuroki (see equation 13 in [2]), and, in words, corresponds to the proportion of individuals with disease who would not develop the condition had they all been exposed to the protective intervention. The preventable proportion can be expressed as:

$$\frac{\Pr(Y=1)-\Pr(Y=1,Y^1=1)}{\Pr(Y=1)}$$

$$=\frac{\{\Pr(Y=1|A=0)\Pr(A=0)+\Pr(Y=1|A=1)\Pr(A=1)\}-\{\Pr(Y=1,Y^1=1|A=0)\Pr(A=0)+\Pr(Y=1,Y^1=1|A=1)\Pr(A=1)\}}{\Pr(Y=1|A=0)\Pr(A=0)+\Pr(Y=1|A=1)\Pr(A=1)}$$

$$=\frac{\{\Pr(Y^0=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)\}-\{\Pr(Y^0=1,Y^1=1|A=0)\Pr(A=0)+\Pr(Y^1=1,Y^1=1|A=1)\Pr(A=1)\}}{\Pr(Y^0=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)}$$

$$=\frac{\Pr(A=0)\{\Pr(Y^0=1|A=0)-\Pr(Y^0=1,Y^1=1|A=0)\}}{\Pr(Y^0=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)}$$

$$=\frac{(1-\pi)\{\theta_{11}^0\emptyset_{01}^0+\theta_{01}^0\gamma^0\}}{(1-\pi)\{\theta_{11}^0(\emptyset_{11}^0+\emptyset_{01}^0)+\theta_{01}^0\gamma^0\}+\pi\{\theta_{11}^1(\emptyset_{11}^1+\emptyset_{10}^1)\}}.$$

Note that, as described in **Table 1** of the main text, the term $\theta_{01}^0\gamma^0$corresponds to the proportion of the *SY* response type 5 in the unexposed population. By definition, the preventable proportion ranges from 0 to 1.

eTable 1. Response types and potential outcomes for the variables *S* and *Y*, disease and disease progression respectively, under the alternative outcome definition that corresponds to "occurrence of severe disease", rather than "severity of disease". The parameters corresponding to the probabilities of the response types are the same as in **Table 1**. We present parameters for the exposed, unexposed, and total populations; the superscripts for θ, $\emptyset$ and γ, or their absence, indicate the population and should not be confused with potential outcomes notation.

		S^a		Probabilities of response types (*S*)				Y^a		Probabilities of response types (*Y*)		
***SY* type**	**Response type of *S***	S^1	S^0	**Exposed (*A* = 1)**	**Unexposed (*A* = 0)**	**Total population**	**Response type of *Y***	Y^1	Y^0	**Exposed (*A* = 1)**	**Unexposed (*A* = 0)**	**Total population**
1	Doomed	1	1	θ_{11}^1	θ_{11}^0	θ_{11}	Doomed	1	1	$\theta_{11}^1\emptyset_{11}^1$	$\theta_{11}^0\emptyset_{11}^0$	$\theta_{11}\emptyset_{11}$
2	Doomed	1	1				Causal	1	0	$\theta_{11}^1\emptyset_{10}^1$	$\theta_{11}^0\emptyset_{10}^0$	$\theta_{11}\emptyset_{10}$
3	Doomed	1	1				Preventive	0	1	$\theta_{11}^1\emptyset_{01}^1$	$\theta_{11}^0\emptyset_{01}^0$	$\theta_{11}\emptyset_{01}$
4	Doomed	1	1				Immune	0	0	$\theta_{11}^1\emptyset_{00}^1$	$\theta_{11}^0\emptyset_{00}^0$	$\theta_{11}\emptyset_{00}$
5	Preventive	0	1	θ_{01}^1	θ_{01}^0	θ_{01}	Preventive	0	1	$\theta_{01}^1\gamma^1$	$\theta_{01}^0\gamma^0$	$\theta_{01}\gamma$
6	Preventive	0	1				Immune	0	0	$\theta_{01}^1(1-\gamma^1)$	$\theta_{01}^0(1-\gamma^0)$	$\theta_{01}(1-\gamma)$
7	Immune	0	0	θ_{00}^1	θ_{00}^0	θ_{00}	Immune	0	0	θ_{00}^1	θ_{00}^0	θ_{00}
Total				1	1	1				1	1	1

References

1. Suzuki, E., E. Yamamoto, and T. Tsuda, *On the relations between excess fraction, attributable fraction, and etiologic fraction.* Am J Epidemiol, 2012. **175**(6): p. 567-75.
2. Yamada, K. and M. Kuroki, *Counterfactual-Based Prevented and Preventable Proportions.* Journal of Causal Inference, 2017. **5**(2): 20160020.