

**Appendix 1.** Formulas used for RRT data. See Horvitz et al. (1976) for more detail.

The proportion of the population ( $n$ ) with the non-sensitive attribute ( $\pi_y$ ) is given by (1).

$$\pi_y = \frac{P_2}{(P_2+P_3)} = \frac{P_2}{(1-P_1)} \quad (1)$$

The proportion of the population with the sensitive attribute ( $\hat{\pi}_A$ ) when ( $\pi_y$ ) is known, is estimated by (2); with  $P$  being the probability of selecting the sensitive attribute ( $P=P_1$ ) and  $\hat{\lambda}$  being the observed  $P$  of “yes” in the RRT section.

$$(\hat{\pi}_A|\pi_y) = \frac{\hat{\lambda}-(1-P)\pi_y}{P} \quad (2)$$

The variance is given by (3), with  $\lambda$  being the probability of a “yes” response ( $\lambda = P\pi_A + (1 - P) \pi_y$ )

$$\text{var}(\hat{\pi}_A|\pi_y) = \frac{\lambda(1-\lambda)}{nP^2} \quad (3)$$