Appendix

Under our formulation $bias = \Sigma_s \psi_s \, \epsilon_s \, w_s$, where ψ_s is the maximum possible bias if there were complete confounding of the unobserved covariate and treatment, w_s is the known fraction of subjects in stratum s, and

$$\epsilon_s = pr(X=1|1,s,R=1) - pr(X=0|1,s,R=1). \tag{A1}$$

By invoking the randomization, we can obtain an upper bound on ϵ_s. Let

$$\gamma_s = pr(X=1|\, s)$$
$$= pr(X=1|z,s) \text{ as a consequence of the randomization.} \tag{A2}$$

In other words the randomization guarantees that the distribution of X does not depend on treatment assignment. Let

$$\kappa_{zs} = \frac{pr(R=1|\; z,s,X=0)}{pr(R=1|z,s,X=1)}, \tag{A3}$$

namely, the relative risk of missingness for $X=1$ versus $X=0$ among subjects randomized to z in stratum s. Writing (A1) as

$$\epsilon_s = \frac{pr(X=1,R=1\,|Z=1,s)}{\Sigma_x\, pr(x,R=1\,|Z=1,s)} - \frac{pr(X=1,R=1|Z=0,s)}{\Sigma_x\, pr(x,R=1|Z=0,s)}$$

$$= \frac{pr(R=1\,|Z=1,s,X=1)pr(X=1|\,Z=1,s)}{\Sigma_x\, pr(R=1|Z=1,s,x)\, pr(x|Z=1,s)} - \frac{pr(R=1\,|Z=0,s,X=1)pr(X=1|\,Z=0,s)}{\Sigma_x\, pr(R=1|Z=0,s,x)\, pr(x|Z=0,s)},$$

and substituting (A2) and (A3) gives

$$\epsilon_s = \frac{\gamma_s}{\gamma_s + \kappa_{1s}\,(1-\gamma_s)} - \frac{\gamma_s}{\gamma_s + \kappa_{0s}\,(1-\gamma_s)} \,. \tag{A4}$$

If $\kappa_{1s} = \kappa_{0s}$, namely the relative risk for missingness between $X = 0$ and $X = 1$ is the same in the two treatment groups, $\epsilon_s = 0$ and there is no bias. This is a weaker condition than MAR which requires that $\kappa_{zs} = 1$, namely that missingness does not depend on X.

Remarkably, it is possible to obtain an upper bound on ϵ_s given only

$$\tau_s = \frac{\kappa_{0s}}{\kappa_{1s}}, \tag{A5}$$

which is the ratio between treatment groups of the relative risk of a missing outcome for $X = 0$ versus $X = 1$. To find the maximum of ϵ_s given τ_s, we substitute (A5) into (A4) with $\kappa_{0s} = \tau_s \kappa_{1s}$ and set $\partial\epsilon_s/\partial\gamma_s = 0$ and $\partial\epsilon_s/\partial\kappa_{1s} = 0$. This gives

$$g_s = \sqrt{\tau_s}\,\kappa_{1s} \,/\, (1+ \sqrt{\tau_s}\,\kappa_{1s}) \text{ for } \kappa_{1s} > 0. \tag{A6}$$

Substituting (A6) into (A4) gives the unconstrained maximum for ϵ_s,

$$\epsilon_{Ms} = \frac{\sqrt{\tau_s} - 1}{\sqrt{\tau_s} + 1}. \tag{A7}$$

Because of constraints on some probabilities the maximum in (A7) is not always attained. The following discussion elucidates the constraints. To simplify the notation we drop the subscript s. Define $\theta_{zx} = pr(R = 1 |\, z, x)$ so $\kappa_z = \theta_{z0}/\theta_{z1}$. The maximum value of ϵ in (A7) implicitly assumes $\theta_{zx} \leq 1$. To investigate violations of this constraint we begin with the identity

$$\pi_z = pr(R = 1\,|z),$$
$$= pr(R = 1|z,\, X = 1)pr(X = 1|\, z) + pr(R = 1|z,\, X = 0)\, pr(X = 0|z)$$
$$= \theta_{z1}\,(\gamma + \kappa_z\,(1-\gamma)). \qquad \text{(A8)}$$

We can rewrite (A8) as

$$\theta_{z1} = \pi_z/(\gamma + \kappa_z(1-\gamma)). \qquad \text{(A9)}$$

Substituting (A9) into the constraint that $\theta_{z0} = \theta_{z1}\kappa_z \le 1,$ rewriting in terms of κ_1 and $\kappa_0 = \tau\,\kappa_1$, and solving for κ_1 gives

$$\kappa_1 \ge \begin{cases} \frac{\pi_1-\gamma}{1-\gamma}\,, & \text{if } \gamma < \pi_1, \\ 0, & \text{if } \gamma \ge \pi_1, \end{cases} \qquad \text{(A10a)}$$

$$\kappa_1 \ge \begin{cases} \frac{\pi_0-\gamma}{\tau(1-\gamma)}\,, & \text{if } \gamma < \pi_{0,} \\ 0, & \text{if } \gamma \ge \pi_{0,} \end{cases} \qquad \text{(A10b)}$$

Substituting (A9) into the constraint that $\theta_{z1} \le 1$, rewriting in terms of κ_1 and $\kappa_0 = \tau\,\kappa_1$, and solving for κ_1 gives

$$\kappa_1 \le \begin{cases} \frac{\gamma}{\pi_1-(1-\gamma)}\,, & \text{if } \gamma > 1-\pi_1, \\ \infty, & \text{if } \gamma \le 1-\pi_1, \end{cases} \qquad \text{(A11a)}$$

$$\kappa_1 \le \begin{cases} \frac{\gamma}{\tau(\pi_0-(1-\gamma))}\,, & \text{if } \gamma > 1-\pi_{0,} \\ \infty, & \text{if } \gamma \le 1-\pi_{0.} \end{cases} \qquad \text{(A11b)}$$

Equations (A10a) and (A10b) represent lower bounds and (A11a) and (A11b) represent upper bounds on κ_1. For these bounds to be operable, the following conditions must hold:

Condition 1. (A10a) $\leq$ (A11a). If $\gamma \leq 1 - \pi_1$, (A11a) is infinity so the condition holds. If $\gamma > 1 - \pi_1,$ this condition reduces to $(1 - \pi_1)\,\pi_1 \geq 0$, which always holds.

Condition 2. (A10b) $\leq$ (A11b). If $\gamma \leq 1 - \pi_0$, (A11b) is infinity, so the condition always holds. If $\gamma > 1 - \pi_0$, this condition reduces to $\tau\,(1 - \pi_0)\,\pi_0 \geq 0$, which always holds because $\tau \geq 0$.

Condition 3. (A10a) $\leq$ (A11b) If $\gamma \leq 1 - \pi_0$, (A11b) is infinity so the condition always holds. If $\gamma > 1 - \pi_1$, we compute the bound on τ when (A6) and hence (A7) hold. Substituting (A7) into (A10a) and (A11b) and solving for κ_1 gives

$$\frac{\pi_1}{1+\sqrt{\tau}\,(1-\pi_1)} \leq \kappa_1 \leq \frac{\sqrt{\tau}\;(1-\pi_0)+1}{\tau\;\pi_0} \tag{A12}$$

Setting the lower bound of (A12) less than or equal to the upper bound of (A12) gives the requirement for (A6) and hence (A7) to hold, namely, $1/\sqrt{\tau} \geq \pi_0 + \pi_1 - 1$. Thus if $\pi_0 + \pi_1 \leq 1$ or $\tau \leq 1/\,(\pi_0 + \pi_1 - 1)^2$, we can obtain the maximum value of ϵ in (A7). If instead $\tau > 1/\,(1 - \pi_0 - \pi_1)^2$ and $\pi_0 + \pi_1 > 1$, we find the largest value of ϵ by computing the two possible values of γ, which are the solutions to a quadratic equation, at the boundary condition (A10a) = (A11b),

$$\gamma_U = \frac{\zeta \pm\sqrt{\zeta^2 - 4\,\pi_1\,(1-\pi_0)\tau\,(\tau-1)}}{2\,(\tau-1)},\ \text{where } \zeta = -1 - \pi_0\tau + \pi_1\tau + \tau. \tag{A13}$$

Substituting (A13) into (A10a) gives $\kappa_{1U} = (\pi_1 - \gamma_U)/(1 - \gamma_U)$. Substituting (A13) and κ_{1U} into (A4) gives

$$\epsilon_U = \begin{cases} max(\epsilon_U^*), & \text{if } \tau \geq 1, \\ min(\epsilon_U^*), & \text{if } \tau < 1, \end{cases} \quad \text{where} \quad \epsilon_u^* = \frac{\gamma_U}{\gamma_U + \kappa_{1U}\,(1-\gamma_U)} - \frac{\gamma_U}{\gamma_U + \tau\ \kappa_{1U}\,(1-\gamma_U)} \tag{A14}$$

Because g_U represents two possible solutions, the notation in (A14) indicates that we select the g_U solution that gives the largest ϵ_U when $\tau \geq 1$ and the smallest ϵ_U when $\tau < 1$.

Condition 4. (A2b) $\leq$ *(A3a)* If $\gamma \leq 1 - \pi_1$, (A3a) is infinity so (A2b) $\leq$ (A3a). If $\gamma > 1 - \pi_1$, we compute the bound on τ when (A6) and hence (A7) holds. Substituting (A6) into (A10b) and (A11a) and solving for κ_1 gives

$$\frac{\pi_0}{\tau + \sqrt{\tau}\,(1-\pi_0)} \leq \kappa_1 \leq \frac{\sqrt{\tau} + (1-\pi_1)}{\pi_1 \sqrt{\tau}}. \tag{A15}$$

Setting the lower bound of (A15) less than or equal to the upper bound of (A14) gives the requirement that for (A6) and hence (A7) to hold, $\sqrt{\tau} \geq \pi_0 + \pi_1 - 1$. Thus if $\pi_0 + \pi_1 \leq 1$ or $\tau \geq (1 - \pi_0 - \pi_1)^2$, we can obtain the maximum value of ϵ in (A7). If instead, $\tau < (1 - \pi_0 - \pi_1)^2$ and $\pi_0 + \pi_1 > 1$, we find the largest value of ϵ by computing the two possible values of γ at the boundary condition (A10b) = (A11a),

$$\gamma_L = \frac{\beta \pm \sqrt{\beta^2 - 4\,\pi_0 (1-\pi_1)\,(1-\tau)}}{2\,(1-\tau)}, \qquad \text{where } \beta = 1 + \pi_0 - \pi_1 - \tau\,. \tag{A16}$$

Substituting (A16) into (A11a) gives $\kappa_{1L} = \gamma_L / \left(\pi_1 - (1 - \gamma_L\,)\right)$. Substituting (A16) and κ_{1L} into (A4) gives

$$\epsilon_L = \begin{cases} max(\epsilon_L^*), & \text{if } \tau \geq 1, \\ min(\epsilon_L^*), & \text{if } \tau < 1, \end{cases} \quad \text{where} \quad \epsilon_L^* = \frac{\gamma_L}{\gamma_L + \kappa_{1L}\,(1-\gamma_L)} - \frac{\gamma_L}{\gamma_L + \tau\ \kappa_{1L}\,(1-\gamma_L)}. \tag{A17}$$

Because γ_L represents two possible solutions, the notation in (A17) indicates that we select the γ_L solution that gives the largest ϵ_L when $\tau \geq 1$ and the smallest ϵ_L when $\tau < 1$.

Reintroducing the subscript s, combining (A7), (A13) (A14), (A16), (A17), the attained maximum of ϵ_s, which depends only on τ_s and π_{zs}, is

$$\epsilon_s^*(\tau_s;\pi_{zs}) = \begin{cases} \epsilon_{Ms}, & \text{if } (1-\pi_{0s}-\pi_{1s})^2 \leq \tau_s \leq 1/(1-\pi_{0s}-\pi_{1s})^2 \text{ or } \pi_{0s}+\pi_{1s} \leq 1, \\ \epsilon_{Ls}, & \text{if } \tau_s < (1-\pi_{0s}-\pi_{1s})^2 \text{ and } \pi_{0s}+\pi_{1s} > 1, \\ \epsilon_{Us}, & \text{if } \tau_s > 1/(1-\pi_{0s}-\pi_{1s})^2 \text{ and } \pi_{0s}+\pi_{1s} > 1. \end{cases} \quad \text{(A18)}$$

Because the choice of τ_s versus $1/\tau_s$ is arbitrary, as either 0 or 1 can be assigned to X, it is reassuring that $\epsilon_s^*(\tau_s;\pi_{zs}) = -\epsilon_s^*(1/\tau_s;\pi_{zs})$. If we take the limit as τ_s approaches 0 or infinity then ϵ_{Ls} and ϵ_{Us} apply, giving

$$\epsilon_{(max)s} = max\left(\frac{1-\pi_{0s}}{\pi_{1s}}, \frac{1-\pi_{1s}}{\pi_{0s}}\right). \quad \text{(A19)}$$

Numerically we checked that $\epsilon_{(max)s}$ is greater than or equal to ϵ_{Ms}.