

Appendix 3

Technical details underpinning the three mediation analyses

Analysis one: effect of deprivation on survival status, with stage as mediator

Let $S_{most|C}^*$ be a random draw from the distribution of *Stage* given baseline confounders C (i.e. calendar year of diagnosis and region) amongst the most deprived patients, and let $S_{least|C}^*$ be a random draw from the distribution of *Stage* given C amongst the least deprived patients. Then, following Vanderweele, Vansteelandt and Robins (Epidemiology, 2014),(1) we define the **randomized interventional analogue of the natural indirect effect (RIANIE)** as the log odds ratio of death comparing:

- setting stage to random draws (given baseline confounders) of the distribution of stage amongst the most deprived, to
- setting stage to random draws (given baseline confounders) of the distribution of stage amongst the least deprived,

if, under both scenarios, everyone were set to be most deprived. Written mathematically:

$$RIANIE = \log \left(\text{odds} \left(Y(\text{most}, S_{most|C}^*) \right) \right) - \log \left(\text{odds} \left(Y(\text{most}, S_{least|C}^*) \right) \right)$$

where

- Y is the survival status, $Y=1$ for death, 0 for alive
- $\text{Odds}(Y) = \text{Pr}(Y = 1) / \text{Pr}(Y = 0)$
- $Y(\text{most}, s)$ is the potential value of Y when deprivation is set to *most deprived*, and *Stage* is set to s .

The **randomized interventional analogue of natural direct effect (RIANDE)** is the log odds ratio of death, if everyone had their stage set to be a random draw (given confounders) of the distribution of stage amongst the least deprived, comparing setting their deprivation level to most versus least deprived. Written mathematically:

$$RIANDE = \log \left(\text{odds} \left(Y(\text{most}, S_{least|C}^*) \right) \right) - \log \left(\text{odds} \left(Y(\text{least}, S_{least|C}^*) \right) \right)$$

The **randomised intervention analogue of the total causal effect (RIATCE)** is the log odds ratio of death comparing:

- setting deprivation to most deprived and stage to a random draw (given baseline confounders) of the distribution of stage amongst the most deprived, versus
- setting deprivation to least deprived and stage to a random draw (given baseline confounders) of the distribution of stage amongst the least deprived.

Written mathematically:

$$RIATCE = \log \left(\text{odds} \left(Y(\text{most}, S_{\text{most}|C}^*) \right) \right) - \log \left(\text{odds} \left(Y(\text{least}, S_{\text{least}|C}^*) \right) \right)$$

This is the sum of the *randomized interventional analogues of the natural direct and indirect effects*.

Similarly, for analysis two: effect of deprivation on survival status, with treatment as mediator

Let $T_{\text{most}|C}^*$ be a random draw from the distribution of *Treatment* given baseline confounders C (i.e. calendar year of diagnosis and region) amongst the most deprived patients, and let $T_{\text{least}|C}^*$ be a random draw from the distribution of *Treatment* given C amongst the least deprived patients. Then we have:

$$RIANIE = \log \left(\text{odds} \left(Y(\text{most}, T_{\text{most}|C}^*) \right) \right) - \log \left(\text{odds} \left(Y(\text{most}, T_{\text{least}|C}^*) \right) \right)$$

$$RIANDE = \log \left(\text{odds} \left(Y(\text{most}, T_{\text{least}|C}^*) \right) \right) - \log \left(\text{odds} \left(Y(\text{least}, T_{\text{least}|C}^*) \right) \right)$$

$$RIATCE = \log \left(\text{odds} \left(Y(\text{most}, T_{\text{most}|C}^*) \right) \right) - \log \left(\text{odds} \left(Y(\text{least}, T_{\text{least}|C}^*) \right) \right)$$

where now $Y(\text{most}, t)$ is the potential value of Y when deprivation is set to *most deprived*, and *Treatment* is set to t .

For analysis three: effect of deprivation on treatment, with stage as mediator

$$RIANIE = \log \left(\text{odds} \left(T(\text{most}, S_{\text{most}|C}^*) \right) \right) - \log \left(\text{odds} \left(T(\text{most}, S_{\text{least}|C}^*) \right) \right)$$

$$RIANDE = \log \left(\text{odds} \left(T(\text{most}, S_{\text{least}|C}^*) \right) \right) - \log \left(\text{odds} \left(T(\text{least}, S_{\text{least}|C}^*) \right) \right)$$

$$RIATCE = \log \left(\text{odds} \left(T(\text{most}, S_{\text{most}|C}^*) \right) \right) - \log \left(\text{odds} \left(T(\text{least}, S_{\text{least}|C}^*) \right) \right)$$

where

- T is surgical treatment, $T=1$ for major surgery, 0 for minor/no surgical procedures, and
- $T(\text{most}, s)$ is the potential value of T when deprivation is set to *most deprived*, and *Stage* is set to s .

The *proportion mediated (PM)* is defined as

$$PM = \frac{RIANIE}{RIATCE}$$

1. VanderWeele TJ, Vansteelandt S, Robins JM. Effect Decomposition in the Presence of an Exposure-Induced Mediator-Outcome Confounder. *Epidemiology*. 2014;25(2):300-6.