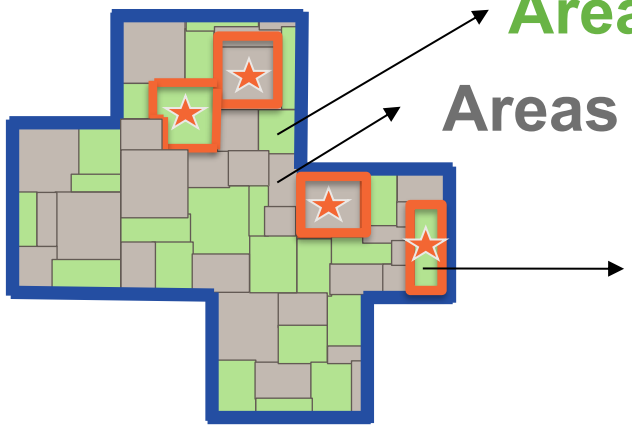
Disentangling Indirect Effects of Vaccine Assignment from Other Causal Pathways in Cluster Designs with Noncompliance

EPI-MALARIA-003: an illustrative case study

Silvia Noirjean, Andrea Callegaro

Cluster Design in the Malaria Vaccine Implementation Programme

- The MVIP was launched in response to the WHO's 2016 recommendation to pilot the RTS,S/AS01 malaria vaccine in routine immunization programs.



Areas introducing the vaccine = *Exposed clusters*

Areas not introducing the vaccine = *Unexposed clusters*

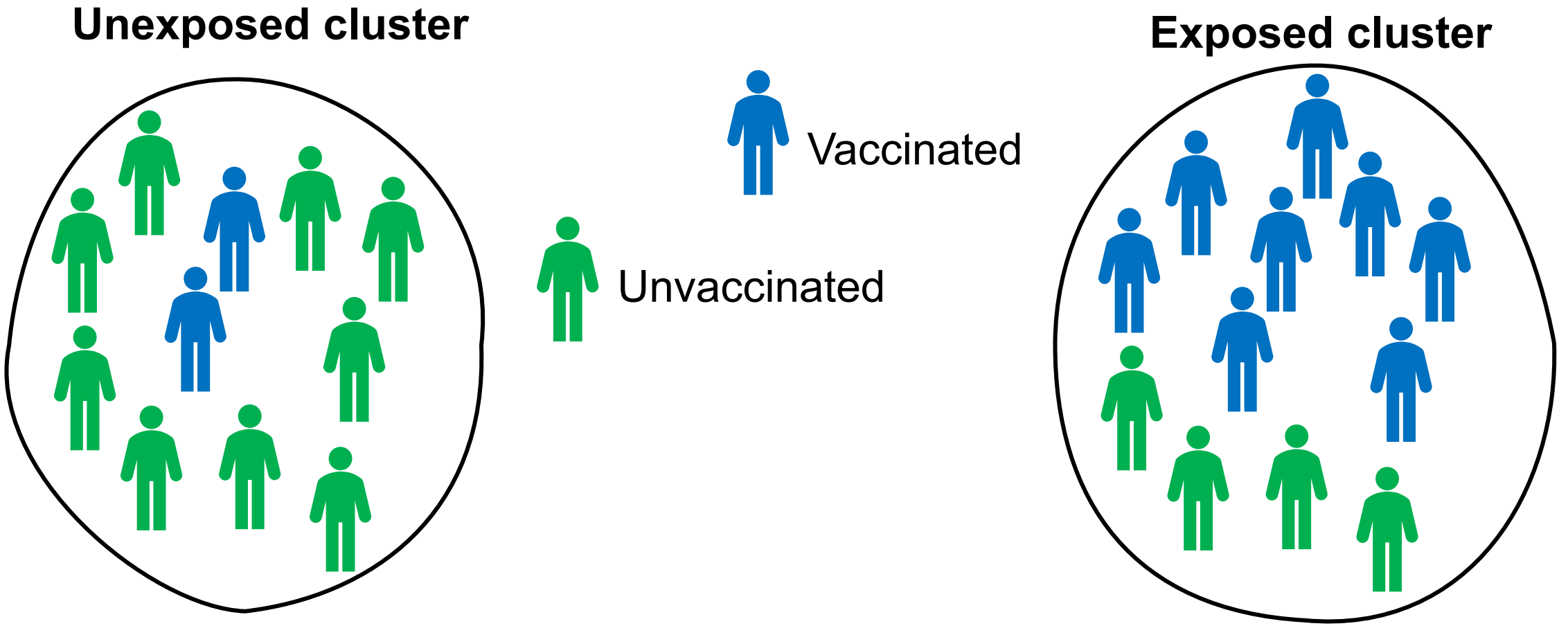
GSK clusters (4 clusters per country)

*Exposed and unexposed clusters were to be comparable in terms of malaria transmission, health facilities level, geographical region and population size.

GSK

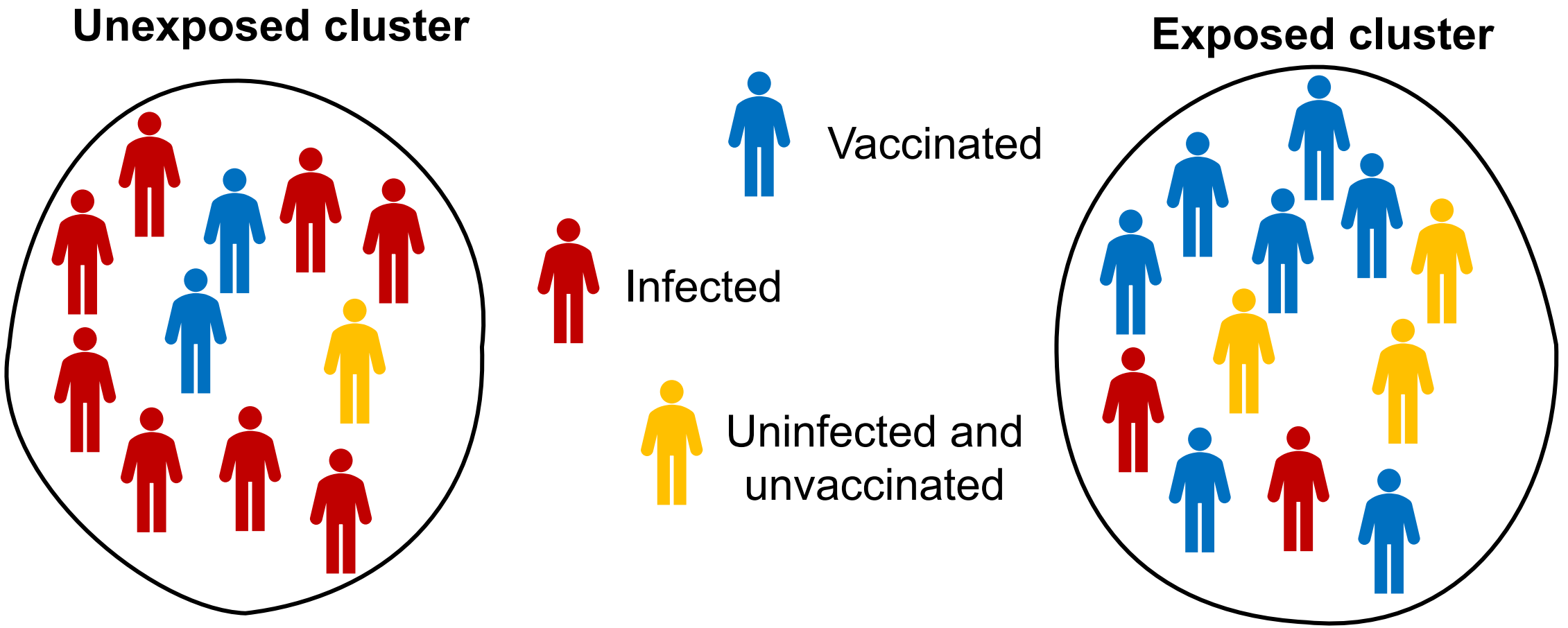
Cluster Design

Noncompliance



Cluster Design

Indirect Effects



Methodological Proposal

Mediation Analysis

- Estimating indirect effects in the presence of **interference within clusters** and **noncompliance with the cluster design** is particularly challenging.

Principal Stratification

Principal Stratification (PS)

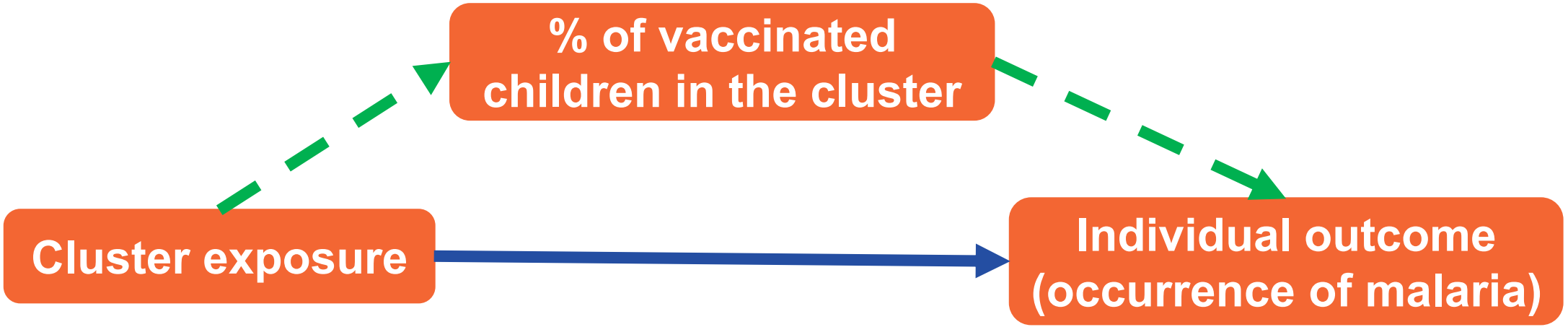
- PS classifying individuals in the following (latent) strata:

If the cluster were exposed	If the cluster were not exposed	Stratum label
Would receive the vaccine	Would receive the vaccine	*Always Taker*
Would receive the vaccine	Would not receive the vaccine	*Complier*
Would not receive the vaccine	Would receive the vaccine	*Defier*
Would not receive the vaccine	Would not receive the vaccine	*Never Taker*

- **Goal**: to estimate a total effect for each strata (heterogenous causal effects).

GSK

Mediation Analysis

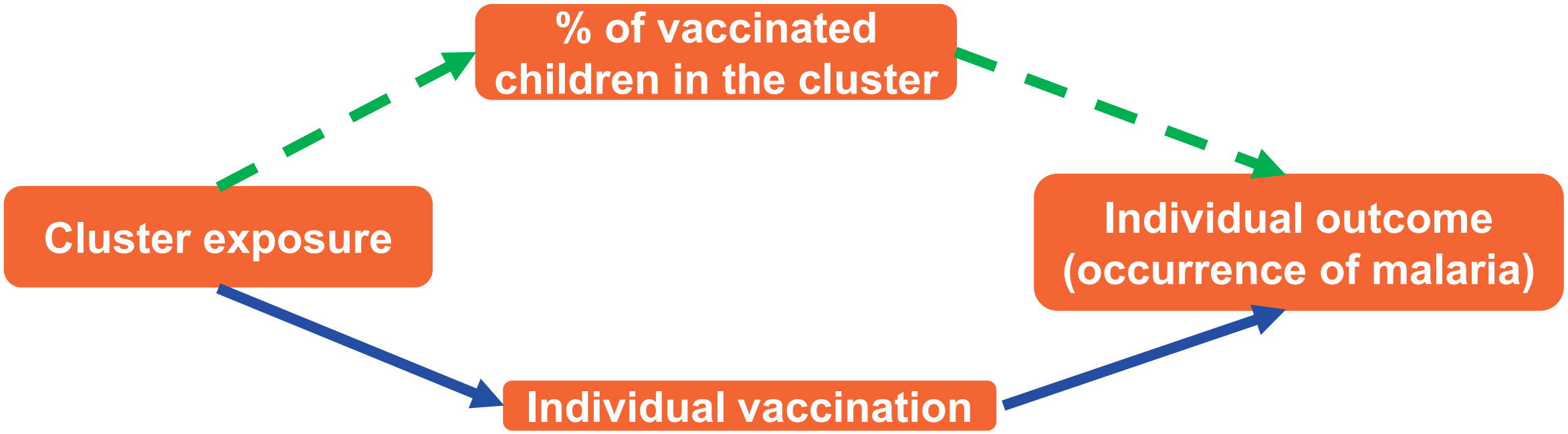


- Each of stratum-specific total effects can be decomposed in
  - **Indirect effect**: the effect on the individual outcome of having a given % of vaccinated children in the same cluster.
  - **Direct effect**: any residual effects of the cluster exposure on the outcome.

Mediation Analysis

Direct Effects

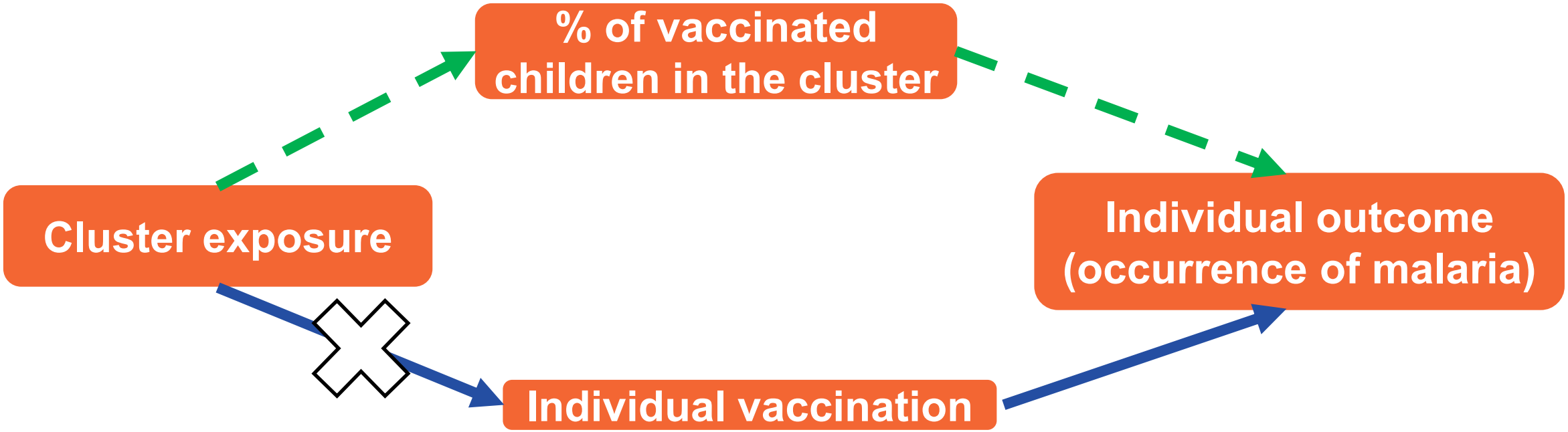
- Direct effects captures the effect of individual vaccination.



Mediation Analysis

Direct Effects

- Always Takers <u>always</u> receive vaccination, Never Takers <u>never</u> receive vaccination regardless of the cluster exposure (**by definition**).
  - No direct effects for Always Takers and Never Takers.



GSK

Blending Mediation Analysis and Principal Stratification

- Direct effects can only exist for Compliers.

If the cluster were exposed	If the cluster were not exposed	Stratum label	Estimands
Would receive the vaccine	Would receive the vaccine	Always Taker	Total and Indirect Effects
Would receive the vaccine	Would not receive the vaccine	Complier	Total, **Direct** and Indirect Effects
Would not receive the vaccine	Would not receive the vaccine	Never Taker	Total and Indirect Effects

Parametric Models

- Two nested models:

1. A model for the **principal stratum membership** conditional on covariates (e.g., multinomial logistic regression).

2. A model for the **potential outcome*** conditional on the principal stratum and covariates.

   *Outcome = number of malaria episodes within a fixed follow-up time; a zero-inflated poisson model was used.

- Bayesian Inference is recommended [Stan code available].

GSK

Results

Principal Strata Size

DUMMY DATA

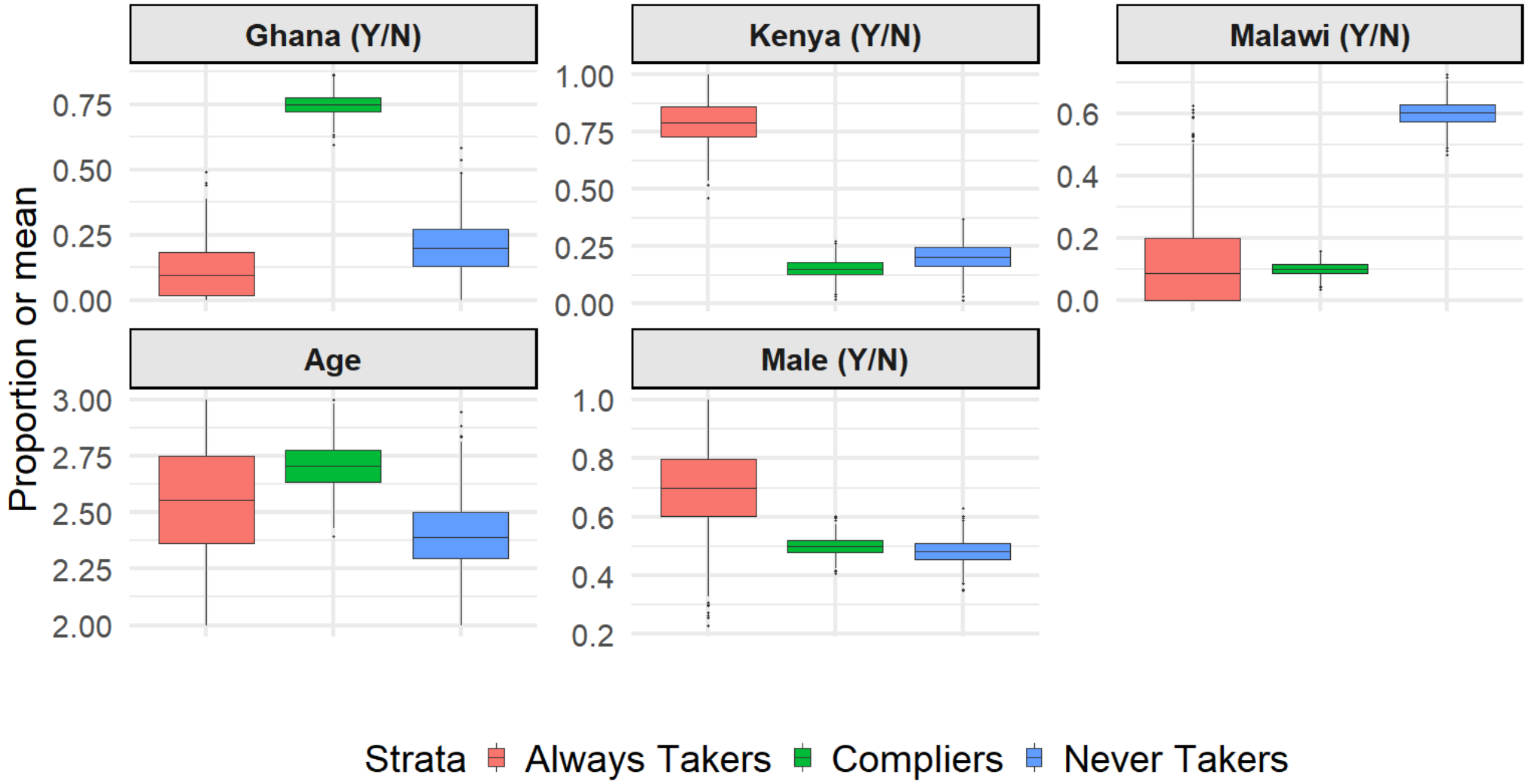
	Mean	*SD*	*2.5%*	*97.5%*
Always Takers	0.06	0.002	0.056	0.064
Never Takers	0.10	0.002	0.096	0.104
Compliers	0.84	0.001	0.838	0.842

GSK

Results

Distribution of Covariates by Principal Strata

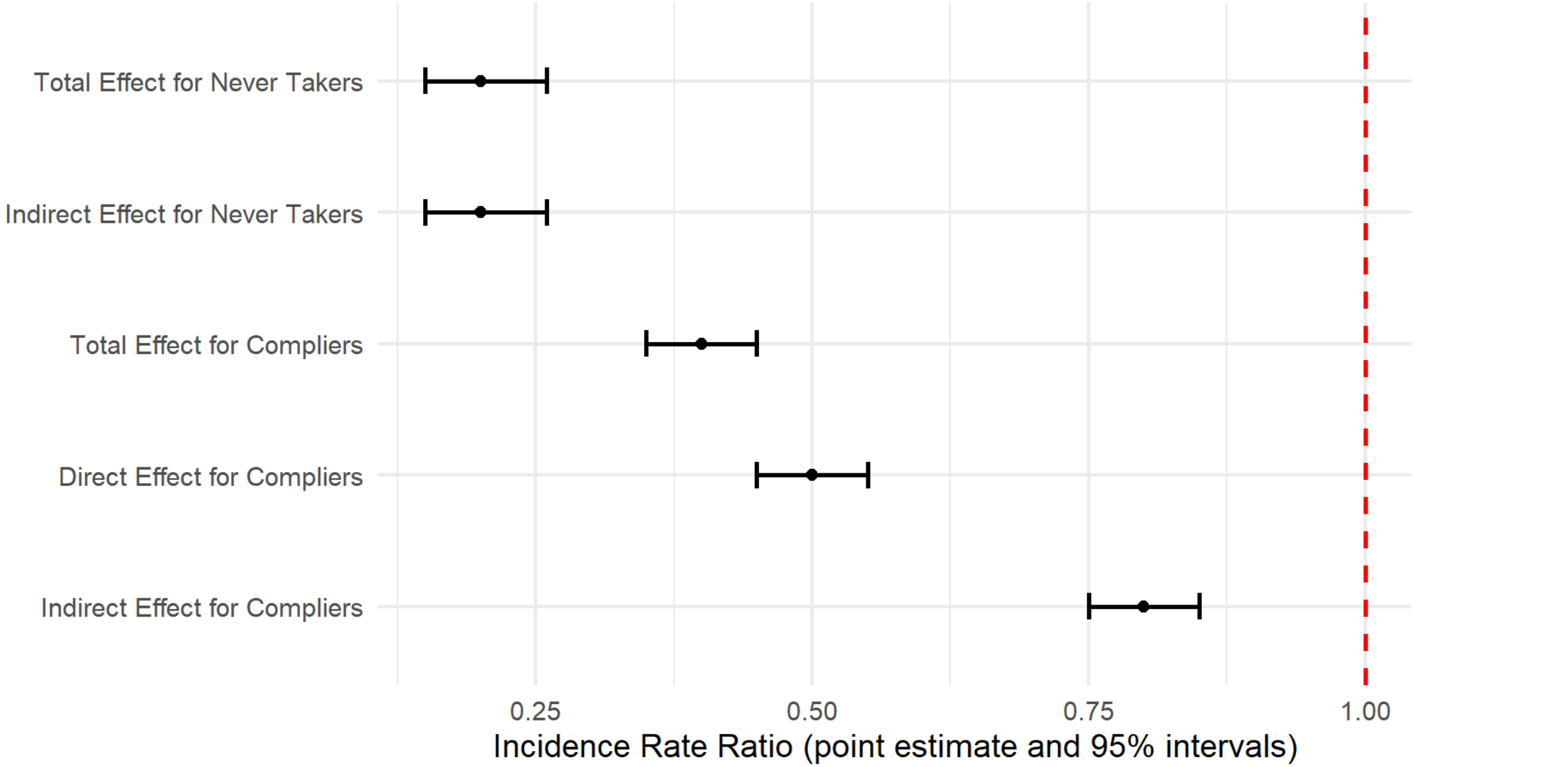
DUMMY DATA



Results

Estimated Causal Effects

DUMMY DATA



Conclusions

- The proposed methodology allows disentangling **direct** and **indirect effects** from the total effect.

  - **Relaxing several assumptions** required by standard approaches.
  - **Accommodating heterogeneous effects** across strata.
  - **Offering deeper insight** into factors that drive decisions to seek vaccination.

Noirjean, S., Mariani, M., Mattei, A., & Mealli, F. (2025). Exploiting network information to disentangle spillover effects in a field experiment on teens' museum attendance. *Journal of Educational and Behavioral Statistics*, *50*(4), 567-603.

Disclosures

- **Conflict of interest**: SN and AC are employees of, and hold shares in, GSK.
- **Funding**: This work was funded by GSK.
- **Acknowledgements:** We thank Raghavendra Devadiga, who served as study statistician for the EPI-MALARIA-003 study, for his valuable contribution to this work.

GSK

gsk.com

Notation

N children, indexed by $i, \ldots, N$.
J clusters, indexed by $j = 1, \ldots, J$.

- Z_j: indicator for the exposure of cluster j ($Z_j = 1$ exposed, $Z_j = 0$ unexposed).
- S_j: proportion of vaccinated children in cluster j.
- Y_{ij}: count of episodes of (any) malaria for child i in cluster j.
- T_{ij}: follow-up time for child i in cluster j (years).
- $\boldsymbol{X}_{ij}$: observations on K baseline covariates for child i in cluster j.

Potential Outcomes

For each child i in cluster j, we define:

- $Y_{ij}(z)$: the potential count of episodes of (any) malaria if cluster exposure were z.
- $Y_{ij}\left(z, S_j(z')\right)$: the potential count of episodes of (any) malaria if cluster exposure were z and the proportion of children vaccinated in the same cluster were the one that would be observed under cluster exposure z'.

GSK

Estimands [I]

- G_{ij}: stratum membership of child i in cluster j (Always Taker, Never Taker or Complier).
- For each stratum g, **Total Effect**: $TE_g = \frac{\sum_{ij:G_{ij}=g} Y_{ij}\left(1,S_j(1)\right)}{\sum_{ij:G_{ij}=g} T_{ij}} \Big/ \frac{\sum_{ij:G_{ij}=g} Y_{ij}\left(0,S_j(0)\right)}{\sum_{ij:G_{ij}=g} T_{ij}}$

  - We are interested in the Vaccine Efficacy: $VE_g = 1 - TE_g$
  - From the stratum-specific Total Effects, it is possible to re-obtain the **average Total Effect**:

$$TE = \sum_g TE_g \times p(G_{ij} = g)$$

GSK

Estimands [II]

Natural Direct Effect: $$NDE_g(z) = \frac{\sum_{ij:G_{ij}=g} Y_{ij}\left(1, S_j(z)\right)}{\sum_{ij:G_{ij}=g} T_{ij}} \Big/ \frac{\sum_{ij:G_{ij}=g} Y_{ij}\left(0, S_j(z)\right)}{\sum_{ij:G_{ij}=g} T_{ij}}$$

Natural Indirect Effect: $$NIE_g(z') = \frac{\sum_{ij:G_{ij}=g} Y_{ij}\left(z', S_j(1)\right)}{\sum_{ij:G_{ij}=g} T_{ij}} \Big/ \frac{\sum_{ij:G_{ij}=g} Y_{ij}\left(z', S_j(0)\right)}{\sum_{ij:G_{ij}=g} T_{ij}}$$

- From Natural Direct and Indirect Effects, it is possible to re-obtain the **stratum-specific Total Effect**:

$$TE_g = NDE_g(z) \times NIE_g(z')$$

GSK

Target Parameters

- **Direct effect**: The effect of receving the vaccine in an exposed cluster. $NDE_{01}(1) = IR_{01}^{(1,1)}/IR_{01}^{(0,1)}$
  - $NDE_{01}(1)$: the direct effect for Compliers when the proportion of vaccinated children is the one that would be observed if the cluster were exposed.

- **Indirect effect**: The effect of other children in the same cluster being vaccinated if the child is not vaccinated.

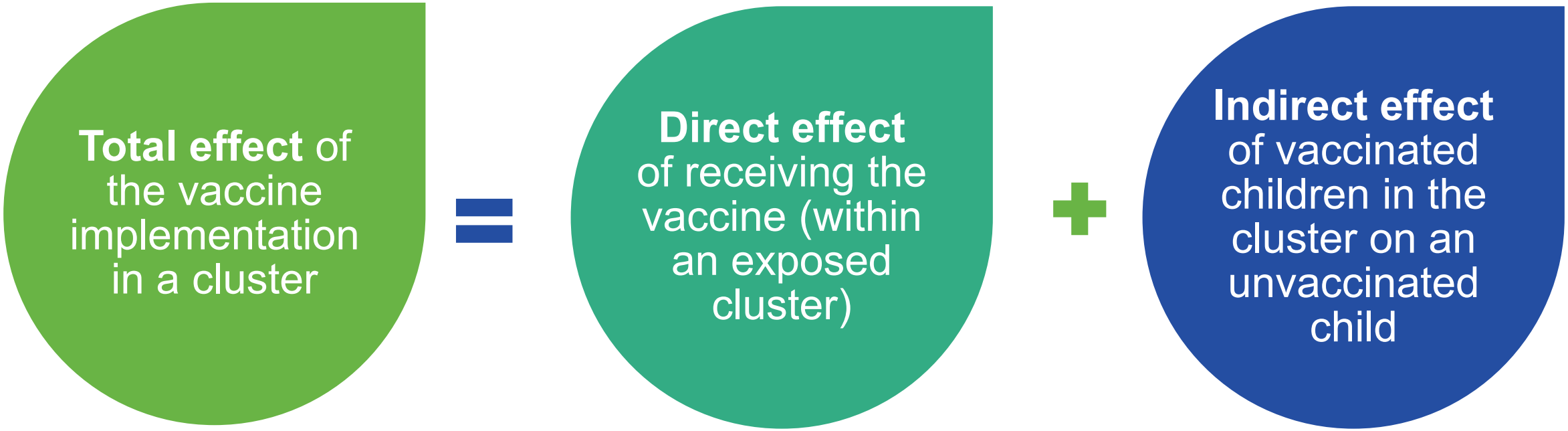
$$NIE_{00}(z) = IR_{00}^{(z,1)}/IR_{00}^{(z,0)}$$
$$NIE_{01}(z) = IR_{01}^{(0,1)}/IR_{01}^{(0,0)}$$

  - $NIE_{00}(z)$: the indirect effect for Never Takers when the cluster exposure is z (for $z = 0,1$).
  - $NIE_{01}(0)$: the indirect effect for Compliers when the cluster exposure is $z = 0$.

$$TE_g = IR_g^{(1,1)}/IR_g^{(0,0)}$$

- **Total effect**: The effect of living in an exposed cluster (versus an unexposed one)
  - TE_g: the total effect for stratum g ($g = 00,11,01$). Total effects for the different strata can be combined to obtain an overall ITT estimate.

Causal Effects



Parametric Models

- Two (nested) models must be specified:
  1. A model for the principal stratum membership conditional on covariates → *Multinomial logit model*

$$\log\left(\frac{p(G_{ij}=g|\boldsymbol{X}_{ij})}{p(G_{ij}=Complier|\boldsymbol{X}_{ij})}\right)=\gamma_g+\boldsymbol{X}_{ij}\boldsymbol{\delta}'_g \quad \text{for } g\in\{Always\ Taker, Never\ Taker\}.$$

Stratum membership

Covariates

Parametric Models

2. A model for the potential outcomes conditional on the principal stratum and covariates
   → *Zero-inflated Poisson model*

$$p\left(Y_{ij}\left(z, S_j(z)\right) = y | G_{ij} = g, \boldsymbol{X}_{ij}\right) = \begin{cases} \phi_{g,z} + \left(1 - \phi_{g,z}\right) \cdot Pois\left(y;\ \mu_{ij,g,z}\right) & \text{if } y = 0 \\ \left(1 - \phi_{g,z}\right) \cdot Pois\left(y;\ \mu_{ij,g,z}\right) & \text{if } y > 0 \end{cases}$$

Cluster exposure

Potential outcome for the count of malaria episodes within a given follow-up period

Potential outcome for the proportion of vaccinated children in the cluster (i.e., the **mediator**)

$$\log\left(\mu_{ij,g,z}\right) = \log\left(T_{ij}\right) + \alpha_{g,z} + \beta_g^{(S)} S_j(z) + \boldsymbol{X}_{ij}\boldsymbol{\beta}^{(X)\prime}$$

$$\text{for } g \in \{\text{Always Taker}, \text{Never Taker}, \text{Complier}\} \text{ and } z \in \{0,1\}$$

Follow-up time

Assumptions [I]

- **Assumption 1** [*Cluster-level SUTVA*]
  - *No hidden variation of the exposure and mediator across clusters*.
  - *Cluster-level interference:* no interference between children in different clusters

    → interference between children in the same cluster can occurr.

- **Assumption 2** [*Monotonicity*]: the subgroup of «Defiers» does not exist.

- **Assumption 3** [*Unconfoundedness*]: Exposed and unexposed clusters are conditionally exchangeable.

- **Assumption 4** [*Exclusion Restriction for all the principal strata*]: Direct effects are only attributable to vaccine receipt.

GSK

Assumptions [II]

- **Assumption 5** [*Latent ignorability of the mediator (I)*]: After conditioning on the cluster exposure, the principal stratum membership and the observed covariates*, there are no other factors affecting both the individual outcome (the occurrence of malaria) and the % of vaccinated children in the cluster.

- **Assumption 6** [*Latent ignorability of the mediator (II)*]: After conditioning on the principal stratum membership and the observed covariates, there are no other factors affecting both the individual outcome (the occurrence of malaria) and the % of vaccinated children in the cluster which are affected by the cluster exposure.

Weaker than the classical mediation analysis assumptions

*Both subject-level and cluster-level covariates can be considered.

GSK