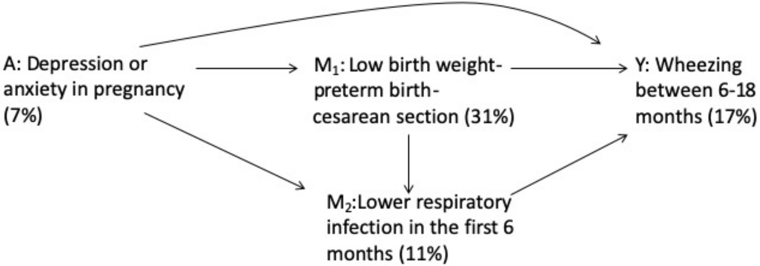
Comparison of approaches to model sequential multiple mediators

D Zugna, F Fasanelli, M Popovic, L Richiardi
Dip. Scienze Mediche, Univ. Torino

SISMEC

Rome, 20 September 2019

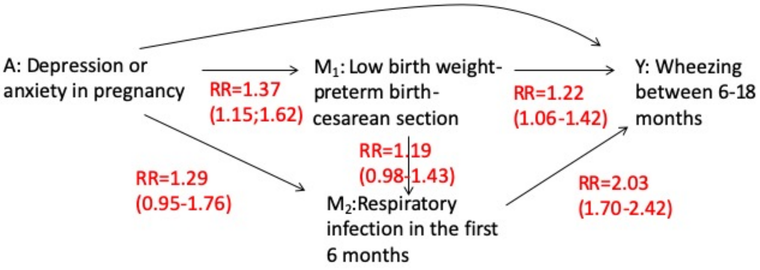
Case study: Ninfea cohort, N=4797



Confounders (C): region, maternal age, education and BMI, parity, child sex

Standard regression analysis

Total effect of A on Y: RR=1.64 (1.31;2.05)



Adjusted for C: $A - Y$, $A - M_1$, $A - M_2$
Adjusted for A, C: $M_1 - Y$,, $M_1 - M_2$
Adjusted for A, C, M_1: $M_2 - Y$

Assumptions

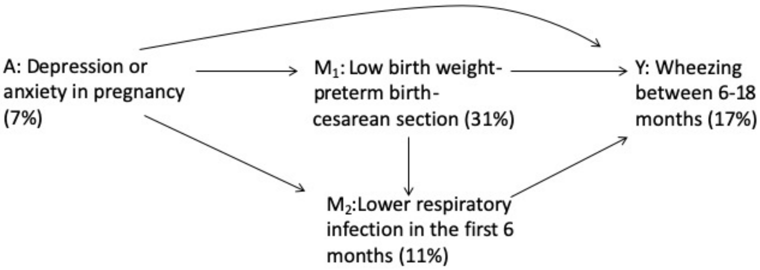
The identification of direct (unmediated) and indirect (mediated) effects requires:

1. No unmeasured $A - Y$ confounding
2. No unmeasured $M_1 - Y$ and $M_2 - Y$ confounding
3. No unmeasured $A - M_1$ and $A - M_2$ confounding
4. No measured $M_1 - Y$ and $M_2 - Y$ confounding affected by A (intermediate confounding)

Multiple mediation: Why?

- Interest in estimating the mediated effect of A on Y through different mediation pathways: $A \to M_1 \to Y$, $A \to M_1 \to M_2 \to Y$ and $A \to M_2 \to Y$
- Allows dealing with intermediate confounding

Case study: Ninfea cohort, N=4797



Confounders (C): region, maternal age, education and BMI, parity, child sex

Sequential mediators: $M_1 \Rightarrow M_2$

1. To model M_1 and estimate the portion of the mediated effect through M_1
2. To model M_1 and M_2 jointly and estimate the portion of the mediated effect through M_1 and M_2 considered together

- This allows to assess the additional contribution of M_2 beyond M_1 alone
- It is not trivial to estimate the effect mediated through M_2 alone because M_1 and M_2 share common pathways

Some definitions

- M_a is the counterfactual value of M if exposure A were set to the value a
- $Y_{aM_{a^*}}$ is the counterfactual value for Y if A were set to a and M were set to M_{a^*}, the level it would have been for each individual had exposure been a^*
- Marginal natural direct effect: $E[Y(a, M_1(a^*), M_2(a^*, M_1(a^*))]/E[Y(a^*, M_1(a^*), M_2(a^*, M_1(a^*))]$
- Marginal natural indirect effect: $E[Y(a, M_1(a), M_2(a, M_1(a))]/E[Y(a, M_1(a^*), M_2(a^*, M_1(a^*))]$
- Conditional natural direct and indirect effects can also be defined and estimated

Methods

- Inverse odds ratio weighting approach (IORW) [1]
  It applies appropriate weights to render the exposure and the mediators independent, deactivating the indirect pathways. The weights are the inverse of the exposure-mediators odds ratio conditional on the covariates and are used in the weighted regression analysis for the outcome to estimate the direct effect.
- Weighting approach [2]
  To estimate for example $E[Y(a, M_1(a^*), M_2(a^*, M_1(a^*))]$, it generates a pseudo-population for the exposure group $A = a^*$ using the individuals' own values of mediators and confounders and with the outcome that would have been observed if each subject had been a member of the exposure group $A = a$. Further it applies appropriate weights to render the exposure and the covariates independent.

[1] Tchetgen Tchetgen, *Stat Med*, 2013
[2] VanderWeele and Vansteelandt, *Epidemiol Method*, 2014

Methods

- Imputation-based approach [3]
  To estimate, for example, $E[Y(a, M_1(a^*), M_2(a^*, M_1(a^*))]$ it standardises the mean outcome in each stratum defined by mediators M_1, M_2 among individuals exposed at level $A = a$, to the mediator distribution of individuals exposed at level $A = a^*$. This is obtained through an imputation procedure where the observed data are complemented with imputed data in which the same individual is evaluated at different exposure levels, a and a^*, but corresponding to the observed mediator levels.

i	A_i	a	a^*	$Y_{aM_{a^*}}$
1	1	1	1	Y_{1M_1}
1	1	0	1	Y_{0M_1}
2	0	0	0	Y_{0M_0}
2	0	1	0	Y_{1M_0}

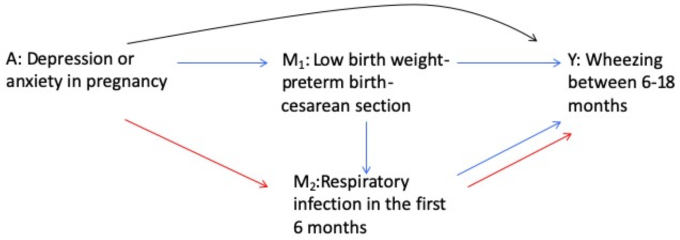
- A regression model is performed on imputed data to estimate direct and indirect effects

[3] Vansteelandt et al, *Epidemiologic Methods*, 2012

Methods

- Extension of the imputation-based approach [4]

Natural indirect effect=Natural indirect effect with respect to M_1+
Partial indirect effects with respect to M_2



Two further assumptions need to be satisfied, the absence of unmeasured confounding of $M_1 - M_2$ association and the confounders of this association need not to be affected by the exposure.

[4] Steen et al, *Am J Epi*, 2017

Methods

The validity of these methods is subject to the correct specification of the following models:

	Method			
Models for	IORW	Weighting	Imputation	Extended imp
Outcome	$\surd^*$	$\surd$	$\surd$	$\surd$
Mediators				$\surd$
Exposure	$\surd$	$\surd$		
Nested counterfactual			$\surd$	$\surd$

*: Unlike the other methods, the mediators are never entered into the regression model for the outcome and is only used to calculate the weight

Results

	Through M_1		Through M_1 and M_2	
	Estimate	95% CI*	Estimate	95% CI*
Marginal effect	Weighting approach			
NDE	1.60	1.32-1.94	1.57	1.28-1.96
NIE	1.01	0.99-1.04	1.04	0.99-1.09
TE	1.63	1.34-2.00	1.62	1.32-2.06
Conditional effect	IORW approach			
NDE	1.60	1.26-1.93	1.57	1.23-1.90
NIE	1.02	0.94-1.12	1.04	0.95-1.15
TE	1.64	1.35-1.99	1.64	1.35-1.99
Conditional effect	Imputation approach			
NDE	1.60	1.32-1.95	1.57	1.29-1.90
NIE	1.02	1.00-1.04	1.05	1.01-1.09
TE	1.64	1.35-1.99	1.64	1.35-1.99

*95% CI: calculated by bootstrap

Results

	Extended imputation approach	
Conditional effect	Estimate	95% CI*
Joint natural direct effect	1.57	1.26-1.97
Joint natural indirect effect	1.05	1.01-1.09
Natural indirect effect by M_1	1.00	0.99-1.00
Partial indirect effect by M_2	1.05	1.01-1.09

*95% CI: calculated by bootstrap

Conclusions

- All models are based on counterfactual definitions
- The described approaches give similar results
- Their application requires: (i) glm, (ii) weighting (excluding the imputation-based approach), (iii) estimate of the predicted values, (iv) bootstrap procedures
- R library has been developed for the imputation-based approach (*medflex*)

References

1. Tchetgen Tchetgen EJ. Inverse odds ratio-weighted estimation for causal mediation analysis. *Stat Med.* 2013 Nov 20;32(26):4567-80.
2. VanderWeele TJ, Vansteelandt S. Mediation analysis with multiple mediators. *Epidemiol Methods*. 2014;2(1):95-115.
3. Vansteelandt S, Bekaert M, Lange T.Imputation strategies for the estimation of natural direct and indirect effects. *Epidemiol Methods*. 2012;1(1):131-158.
4. Steen J, Loeys T, Moerkerke B, and Vansteelandt S. Flexible Mediation Analysis With Multiple Mediators. *Am J Epidemiol.* 2017;186(2):184-193.

Aknowledgements

This work was supported by the European Union's Horizon 2020 research and innovation programme (LifeCycle Project) and by the programme "Dipartimenti di Eccellenza 2018−2022" (Teseo Project).