An Introduction to Mediation

Real World Medical Statistics Meeting

NIHR Statistics Group

KING'S College LONDON

Kim Goldsmith, Graeme MacLennan, Richard Emsley
Facilitator: Simon Bond

Institute of Psychiatry, Psychology & Neuroscience, King's College London
Centre for Health Care Randomised Trials (CHaRT) , University of Aberdeen
Cambridge Clinical Trials Unit (CCTU)
Friday, June 22nd , 2018

Tweet about the meeting and breakout group

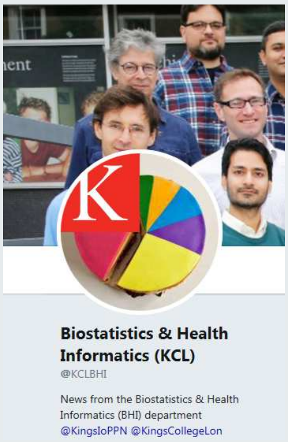
Hashtags **#RealWorldMedStats**
#statistics
#causalmediation

@NihrG – NIHR Statistics Group

@KimberleyGol – Kim
@KCLBHI – Kim and Richard's Department

@gsmaclennan – Graeme
@hsru_aberdeen – Graeme's Department

@richardaemsley – Richard



2

Session Outline

- Introduction to mediation and its application
  Kim ~10 – 15 min
- Mediation in fields other than psychology/psychiatry
  Graeme ~ 10 – 15 min
- Group Discussion
  All ~ 30 – 40 min
  - ❑ Questions about mediation
  - ❑ Mediation in other fields
  - ❑ What are your mediation hypotheses?

3

Mediation

Hyman, 1955:

*"When the analyst interprets a relationship, he determines the **process** through which the assumed **cause** is related to what we take to be its **effect**. **How** did the result come about? What are the 'links' between the two variables? …. Described in formal terms, the interpretation of a statistical relationship between two variables involves the introduction of further variables and an examination of the resulting interrelationships between all of the factors".*

David Kenny (on his website):

*"One reason for testing mediation is trying to understand the **mechanism** through which the **causal variable affects the outcome**".*

In other words, **mediation** allows for **MECHANISM EVALUATION**.

4

Mediation and mediators

A mediator (*M*) is a variable that occurs in the causal pathway from an exposure (*D*) or a randomised treatment (*R*) to an outcome variable (*Y*).

It causes variation in the outcome and itself is caused to vary by the exposure/treatment variable.

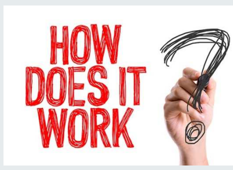
This causal chain implies a temporal relation

- *D* or *R* occurs before *M* and
- *M* occurs before *Y*

Mediating variables are often called **intervening** or **intermediate variables**.
(They have also been called process variables; but we reserve this term for variables that measures aspects of the therapeutic process.)

5

Treatment mechanisms of change: Why are we interested?



- ❑ Further develop or confirm postulated theoretical treatment models
  Such models can and should be evaluated using mediation analysis
- ❑ May provide information about accuracy of the theoretical model and how intervention works
- ❑ May highlight ways in which treatment can be developed, tailored or refined
  Treatment failed to change the hypothesized *M*? Change treatment.
  M failed to influence *Y*? Change target.

6

So we can…

Ask not just whether a treatment works….

But also how it works….
And if it didn't work, why this might have happened

Giving us more information about treatments from trials, which are expensive and time-consuming

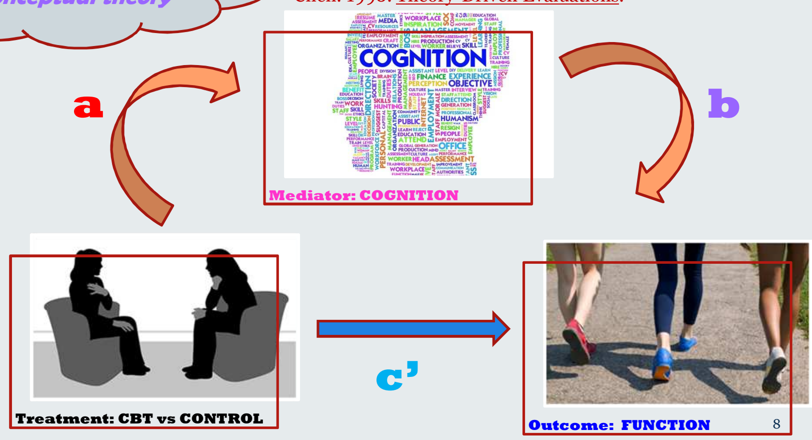
(Explanatory trials)

7

Clinical psychology and psychiatry mediation example

Windgassen, **Goldsmith**, Moss-Morris, Chalder. JMH, 2016; Jan 6:1-7.
MacKinnon. 2008. Introduction to Statistical Mediation Analysis.
Chen. 1990. Theory-Driven Evaluations.

a = action theory
b = conceptual theory



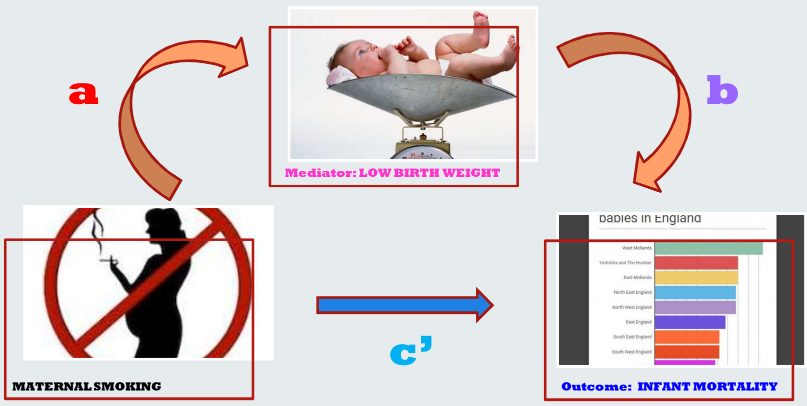
8

Mediation examples in other fields

Epidemiology:

e.g. Hypothesised mechanisms for the transmission of disease.

Often using binary endpoints such as death, disease or injury.

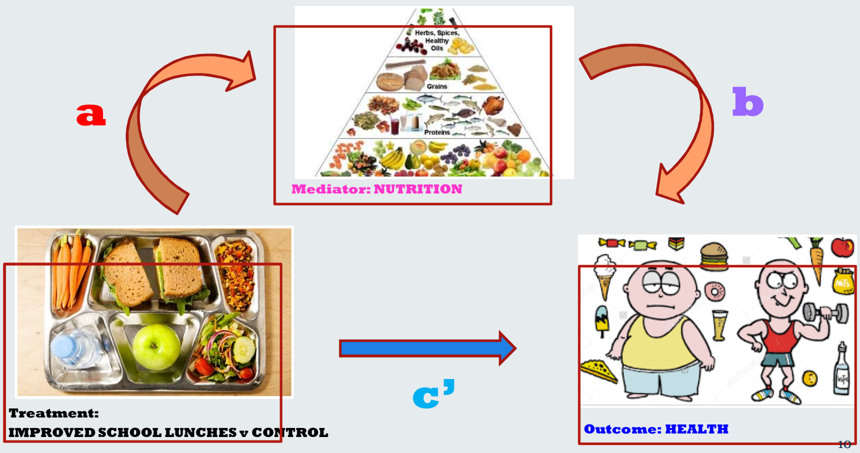


9

Mediation examples in other fields

Prevention research:

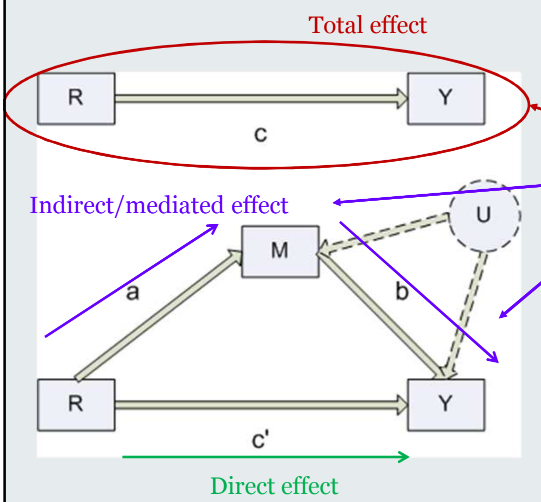
Public health treatments are typically based on a mediation theory:



10

Simple single mediator model (continuous M and Y)

R = treatment
M = mediator
Y = outcome
U = unmeasured confounders



$$Y_i = \alpha_1 + \beta_1 R_i + \varepsilon_{i1}$$
$$M_i = \alpha_2 + \beta_2 R_i + \varepsilon_{i2}$$
$$Y_i = \alpha_3 + \beta_3 R_i + \gamma M_i + \varepsilon_{i3}$$

Total effect = β_1 = c
Direct effect = β_3 = c'
a = β_2
b = γ
Indirect (mediated) effect = $\beta_2\gamma$ = ab
Total effect = c = ab + c'
Indirect effect = ab = c – c'

Note: **not** getting a new estimate of the total effect
<u>Partitioning total effect on the outcome</u> into **indirect effect through mediator** and remaining direct effect

Baron and Kenny Steps Method

Baron and Kenny (1986) (Judd and Kenny (1981)) discussed four steps to establish mediation (see also David Kenny website):

Highly cited & well-known (so mentioned), but:

1. Not as powerful as other methods (MacKinnon et al, 2002)
2. Require significance of total effect
   Now widely agreed not necessary for mediation analysis
   Treatment didn't work? Mediation analysis probably more important
3. Does **not calculate *ab***, which quantifies the indirect effect

Instead: Product of coefficients [POC, see previous slide] (or causal inference methods)

MacKinnon. 2001. Mediating variable. *International Encyclopedia of the Social and Behavioural Sciences*, p. 9503-7.
MacKinnon. 2008. *Introduction to Statistical Mediation Analysis*, Chps 3 and 4.

12

Linear Structural Equation Models (LSEM) for mediation

Can implement product of coefficients approach in this framework

Are useful because they allow for:

- Simultaneous fitting of multiple regressions
- Measurement error by modelling latent variables
- Use full information maximum likelihood, so account for missing data under a missing at random assumption
- (Allow longitudinal/repeated measures modelling, see Goldsmith et al, 2016 and 2017)

13

Other considerations

- Mediation is longitudinal
  - Hypothesises causal chain - by definition a longitudinal process
  - Study design and analysis should respect this
- Confounding of mediator – outcome relationship could = bias
  - And of all relationships in observational studies
  - Consider potential confounders at design stage, measure during study, include in models
  - Measure and adjust for baseline mediator and outcome (Pickles et al 2015, Landau et al 2018)
- Measurement error in mediator could = bias
  - Repeated measures or other designs -> use of structural equation models
- Reporting
  - Various effect sizes (MacKinnon, 2008)
  - Report both ***a*** and ***b paths*** as well as indirect effect (***a*** x ***b***)
  - Appropriate confidence interval (percentile bootstrap, Fritz et al 2012)
  - Stata sem example: ***https://stats.idre.ucla.edu/stata/faq/how-can-i-do-mediation-analysis-with-the-sem-command/***

14

Implications of results – treatment refinement

***a* path – action theory – answers important question:**
Did the treatment affect the mediator?

If ***a* is not significant** :

Suggests treatment not acting as expected (does not change the targeted mediator)
Suggests the **treatment needs to be modified so that it either**:

1. affects that mediator
2. affects a different mediator that is related to the outcome, or
3. both

***b* path – conceptual theory – answers important question:**
Is there a relationship between the mediator and outcome?

If ***b* is not significant** :

Suggests outcome cannot be changed by affecting that mediator
Suggests **need to modify treatment** to affect mediator **related to the outcome**

15

Recommendations for triallists interested in mediation

Key point is to build in study of mediators in at the design stage:

- Measure mediators and outcome at baseline
- Measure potential confounders of M – Y relationship at baseline
- Measure mediators at important intermediate time points, to respect temporality
- Consider multiple arm designs where possible
  True for trials in general, but also in terms of mediators, provides rich data information about treatments in shorter amount of time
- Plan in time/funding for such analyses

16

Acknowledgements

Some slides/ideas courtesy of:

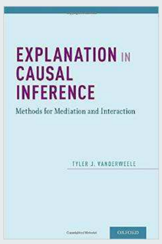
- ❑ **David MacKinnon**
- ❑ **Sabine Landau**
- ❑ **Richard Emsley**
- ❑ **Mizan Khondoker**
- ❑ **Artemis Koukounari**

17

Causal mediation analysis

Statistical mediation as described here has four main problems:

1. Unmeasured confounding between mediator and outcome
2. No interactions between exposure and mediator on outcome
3. Doesn't easily extend to non-linear models
4. Assumes correctly specified models



Causal mediation analysis has arisen from the causal inference literature, and addressed these problems.

Formally defines the causal mediation parameters.

18

Mediation analysis in Stata

Sobel test
- **sgmediation**
- Could run individual regressions and code to obtain indirect effect/to provide 95% percentile bootstrap CI

LSEM approach
- **sem** with additional code to provide 95% percentile bootstrap CI
  - Need to take care in obtaining estimates for categorical/count variables or in the presence of interactions

Causal inference approach
When used with certain settings, give same results as other approaches, but also more flexible
- **paramed**
  - Outcome variable Y: binary, continuous, or count
  - Treatment variable T: binary or continuous
  - Mediator variable M: binary or continuous
  - Covariates: categorical or continuous
  - Proper estimates in the presence of exposure-mediator interaction

19

Selected References

Buis. Direct and indirect effects in a logit model . *The Stata Journal* 2010 *10*(1), 11–29.

Chalder, Goldsmith et al. Rehabilitative therapies for chronic fatigue syndrome: a secondary mediation analysis of the PACE trial. *Lancet Psychiatry.* 2015 *2*(2):141-52.

Chen. 1990. Action theory and conceptual theory: summatively diagnosing the intervention program. *Theory-DrivenEvaluations.* Newbury Park, CA: Sage Publications.

Cheung. Comparison of methods for constructing confidence intervals of standardized indirect effects. *Behavior Research Method.* 2009 *41*, 425-438.

Collins and Graham. The effect of the timing and spacing of observations in longitudinal studies of tobacco and other drug use: temporal design considerations. *Drug Alcohol Depen* 2002 *68*: S85 – S96.

Fleiss, Shrout. Effects of Measurement Errors on Some Multivariate Procedures. *American Journal of Public Health,. 1977 7*(12), 1188-1191.

Fritz, Taylor, MacKinnon. Explanation of Two Anomalous Results in Statistical Mediation Analysis. *Multivariate Behav Res.* 2012 *47*(1), 61-87.

Goldsmith, Chalder, White, Sharpe, Pickles. Measurement error, time lag, unmeasured confounding: considerations for longitudinal estimation of the effect of a mediator in randomised clinical trials. *Stat Meth Med Res*, 2016; Sep 19. pii: 0962280216666111. [Epub ahead of print].

Goldsmith, Chalder, White, Sharpe, Pickles. Tutorial: Simplex, latent growth and latent change structural equation models for longitudinal mediation in the PACE trial of treatments for chronic fatigue syndrome. Psychological Methods, doi: 10.1037/met0000154 [Epub ahead of print].

Hoyle and Kenny. 1999. Statistical power and tests of mediation. In: *Statistical strategies for smallsample research.* Newbury Park: Sage.

Hyman. 1955. The Introduction of Additional Variables and the Elaboration of the Analysis. *Survey Design and Analysis* (pp. 275-329). New York, NY: The Free Press.

Kennedy. 2008. *Guide to Econometrics.* Blackwell Publishing, Malden, MA, p158.

Klein. 2011. *Principles and Practice of Structural Equation Modelling.* New York, NY: The Guildford Press.

Landau, Emsley, Dunn. (2018). Beyond total treatment effects in RCTs: Baseline measurement of intermediate outcomes needed to reduce confounding in mediation investigations. *Clinical Trials,* 2018, 15(3) 247–256, doi: 10.1177/1740774518760300. (Controlling for baseline in mediation models)

le Cessie, Debeij, Rosendaal, Cannegieter, Vandenbroucke. Quantification of bias in direct effects estimates due to different types of measurement error in the mediator. *Epidemiology.* 2012 23(4):551-60.

MacKinnon and Dwyer. Estimating Mediated Effects in Prevention Studies. *Evaluation Review.* 1993 *17*(2), 144-158.

MacKinnon, Warsi and Dwyer. A Simulation Study of Mediated Effect Measures. *Multivariate Behav Res.* 1995 *30*(1), 41.

20

Selected References

MacKinnon. 2001. Mediating variable. *International Encyclopedia of the Social and Behavioural Sciences* (pp. 9503-7). Oxford, UK: Elsevier Science, Ltd.

MacKinnon, Goldberg, Clarke, Elliot, Cheong, Lapin, ... & Krull. 2001. Mediating mechanisms in a program to reduce intentions to use anabolic steroids and improve exercise self-efficacy and dietary behavior. *Prevention Science*, *2*(1), 15-28.

MacKinnon et al. A comparison of methods to test mediation and other intervening variable effects. *Psychol Methods.* 2002 *7*(1):83-104.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. 2007. Mediation analysis. *Annual review of psychology*, *58*, 593.

MacKinnon. 2008. *Introduction to Statistical Mediation Analysis.* Taylor & Francis Group: New York, NY.

Muthén and Asparouhov. Causal Effects in Mediation Modeling: An Introduction With Applications to Latent Variables. *Structural Equation Modeling.* 2015 *22*(1): 12-23.

O'Rourke and MacKinnon. (2015). When the test of mediation is more powerful than the test of the total effect. *Behavioral Research Methods*, *47*(2): 424–442.

Pickles, A., Harris, V., Green, J., Aldred, C., McConachie, H., Slonims, V., ... & Charman, T. (2015). Treatment mechanism in the MRC preschool autism communication trial: implications for study design and parent-focussed therapy for children. *Journal of Child Psychology and Psychiatry*, *56*(2), 162-170. (Controlling for baseline in mediation models)

Preacher and Kelley. Effect size measures for mediation models: quantitative strategies for communicating indirect effects. *Psychol Methods.* 2011 *16*(2), 93-115.

Tang and DeRubeis. Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *J Consult Clin Psych.* 1999 *67*: 894-904.

Vanderweele and Vansteelandt. Odds Ratios for Mediation Analysis for a Dichotomous Outcome. *Am J Epi.* 2010 *172*(2): 1339–1348.

Vanderweele, Valeri, Ogburn. The role of measurement error and misclassification in mediation analysis: mediation and measurement error. *Epidemiology.* 2012 23(4):561-4.

Valeri and VanderWeele. 2013. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*, 18:137-150. (GOOD STARTING POINT FOR CAUSAL INFERENCE)

Valeri, Lin, Vanderweele. Mediation analysis when a continuous mediator is measured with error and the outcome follows a generalized linear model. *Statistics in Medicine.* 2014 10;33(28):4875-90.

Windgassen, Goldsmith et al. Establishing how psychological therapies work: the importance of mediation analysis. 2016 *JMH*, *25*(2):93-9.

Wright. Correlation and causation Part I. Method of path coefficients. *Journal of Agricultural Research.* 1920a *20*, 0557-0585.

Wright. The Relative Importance of Heredity and Environment in Determining the Piebald Pattern of Guinea-Pigs. *Proc Natl Acad Sci U S A*, 1920b *6*(6), 320-332.

21