Composite variables in medical research, generally, and with a focus on occupational stress research

Sigurd Mikkelsen

Department of Occupational and Environmental Medicine Bispebjerg Hospital, Copenhagen University

Background

The background for this meeting is a scientific disagreement on the analyses of combined effects of two (or more) work related psychosocial exposures on an outcome:

1. by a standard regression analysis including the main effects and a multiplicative interaction term, or

2. by a regression analysis which only includes a multiplicative interaction term and no main effects

The multiplicative interaction term is a composite variable

Background

Composite variables are frequently used in medical research, eg.:

- body mass index,
- waist-hip ratio,
- Tiffeneau Index (FEV1/FVC),

- allostatic load index,
- socioeconomic status,
- diagnoses,
- scales
- concentrations of chemicals

Are there any good reasons, a rationale, for analysing effects of only a composite variable instead of the independent main and interaction effects of its components? And what could these reasons be?

Background

A composite variable is a variable derived from two or more other variables, e.g. BMI, job-strain and effort reward-imbalance variables

A basic reason for the construction of a composite

variable is often that it seems to be a simple way of adjusting the effect of one variable for the effect of another variable. The effect of a composite variable is often assumed to measure the combined effect of its components.

Background

Two dominant work stress theories:

Karasek's job strain model: work stress results from high job demands in combination with low job control (job strain)

Siegrists effort-reward-imbalance model: work stress results from high efforts at work in combination with low rewards (ERI)

Both models postulate a combined effect of two variables

Background

The outcome ”work stress” is poorly defined. The effect of job strain and ERI on different outcomes is assumed to be due to stress generated by high levels of these stressors. Outcomes include many different measures, eg. from job satisfaction, exhaustion, depressive symptoms and other self-reported outcomes, - to hypertension, coronary heart disease, clinical depression and other ”hard” outcomes.

.

Job strain

Disease

ERI

What is a combined effect of two factors?

Two factors have a combined effect

- if both factors have an independent effect on the outcome, *or*
- if the two factors interact to produce an effect

Two factors do not have a combined effect

- if one of the two factors has no independent effect on the outcome, *and*
- if the two factors do not interact to produce an effect

Measuring job strain and ERI, usual approach

Job strain variables: continuous demands (A) and control (B) are dichotomized at the median. Combined, we then have 4 cells:

high control and low demands (relaxed or low strain jobs), low control and low demands (passive jobs), high control and high demands (active jobs), and low control and high demands (high strain jobs).

Job strain effects are often examined as effects of high strain against the other combinations of demands and control.

ERI variables: the ratio of continuous efforts/rewards, and then dichotomized, usually at unity.

Statistical models examining a combined effect of two factors, A and B on an outcome Y, adjusted for C:

Standard regression model

$E(Y)= \beta_0 + \beta_A A + \beta_B B + \beta_{AB} AB + \beta_C C + \varepsilon$ (1) $E(Y)= \beta_0 + \beta_A A + \beta_B B + \beta_C C + \varepsilon$ (2)

Composite-variable-only model

$E(Y)= \beta_0 + \beta_{AB} AB + \beta_C C + \varepsilon$ (3) Note: applies also to ERI if B is substituted by B^{-1}

Proponents of the composite-variable-only model:

We are only interested in the combined effect of A and B, not in any separate effects of A and B nor in any

multiplicative effect, only in the combined effect since this is the core of the theory.

The product AB reflects the combined effect. Therefore, we use the composite-variable-only model. If the product term AB has an effect, we will conclude that there is a combined effect, supporting the theory.

$$E(Y)= \beta_0 + \beta_{AB}AB + \beta_C C + \varepsilon$$

What is wrong about this argument?

What is wrong?

The model does not consider confounding from A and B

$$E(Y)= \beta_0 + \beta_{AB}AB + \beta_C C + \varepsilon$$

B

AB Y

A

$$E(Y)= \beta_0 + \beta_{AB}AB + \beta_A A + \beta_B B + \varepsilon$$

What is wrong?

Adjusted for confounding from A and B, the alternative model is identical to the standard model:

$$E(Y)= \beta_0 + \beta_A A + \beta_B B + \beta_{AB}AB + \varepsilon$$

There is only no confounding from A and B if $\beta_A=0$ and $\beta_B=0$

It is not part of the theories that there is no effect of A and B on Y, on the contrary, the ERI-theory requires that there is an effect of both efforts (A) and rewards (B). The job strain theory does not state that there is no effects of both demands(A) and control (B).

What is wrong?

The effect estimate of the composite variable (the multiplicative term) is biased if not adjusted for confounding from the main effects.

The composite-variable-only method of analysing combined effects of demands and control, and of efforts and rewards produces biased estimates of their ”combined” effects

A ”combined” effect according to the composite variable-only method may be due to an effect of only one of the component variables and not to any combined effect.

.

IPD-Work consortium

The IPD-Work consortium is a large collaborative research project on the relation between work-related psychosocial factors and health, using meta-analyses of individual participant data from multiple cohort studies representing a range of countries.

To avoid “ ..the use of post hoc comparisons and selective reporting of findings after multiple testing”*, the IPD work consortium adheres to predefined exposure definitions and analyses.

These are the composite-variable-only models described above, and their effect estimates for the composite variable are therefore most likely biased.

Kivimaki M et al. IPD-Work consortium: pre-defined meta-analyses of individual-participant data strengthen evidence base for a link between psychosocial factors and health. Scand J Work Environ Health 2015 May 1;41(3):312-21.

Example 1

IPD-work study on coronary heart disease and job strain (Lancet 2012**):**

Job strain: HR = **1.23 (**95% CI **1.08–1.39)**

Based on information in an appendix Michael Ingre showed that there was
no multiplicative interaction,
demands had no effect, while

control had a significant effect

Thus, there was no combined effect of demands and control, only an effect of control. This information was not discussed and was not mentioned in the conclusion. The effect of control could be explained by SES.

Example 2

IPD-work study on clinical depression and job strain (Psychological Medicine 2017):

Job strain: HR = **1.27** (95% CI **1.04–1.55**)

Information in an appendix (sensitivity analyses), using the standard analysis model showed that there was no multiplicative interaction,
demands had no effect, while
control had a significant effect.

Thus, there was no combined effect of demands and control, only an effect of control. This effect could be due to an effect of SES

Example 3

IPD-work study on clinical depression and Effort Reward-Imbalance (Epidemiology 2017)

ERI: **1.16** (95%CI **1.01 - 1.34**)

Effort: 0.99 (95% CI 0.87-1.13)
Reward: **1.18** (95% CI **1.04-1.33**)

These figures do not point to any interaction between effort and reward and only one of the two factors was associated with CHD. There was no combined effect. These results were published in the text, but they were not discussed and not

included in the conclusion. The effect of rewards could possibly be explained by SES.

Dragano N et al. Effort-Reward Imbalance at Work and Incident Coronary Heart Disease: A Multicohort Study of 90,164 Individuals. Epidemiology 2017 Jul;28(4):619-26.

Conclusion

Composite multiplicative or ratio variables should only be used in the context of a full linear model in which the variables that make up the composite variable are included.

Kronmal RA. Spurious Correlation and the Fallacy of the Ratio Standard Revisited. Journal of the Royal Statistical Society Series A (Statistics in Society) 1993;156(3):379-92.

Brambor T, Clark WR, Golder M. Understanding Interaction Models:Improving Empirical Analyses. Political Analysis: An Annual Publication of theMethodology Section of the American Political Science Association 2017;14(1):63-82.

Thank you for your attention

Reference List

(1) Brambor T, Clark WR, Golder M. Understanding Interaction Models:Improving Empirical Analyses. Political Analysis: An Annual Publication of theMethodology Section of the American Political Science Association 2017;14(1):63-82.

(2) Dragano N, Siegrist J, Nyberg ST, Lunau T, Fransson EI, Alfredsson L, et al. Effort-Reward Imbalance at Work and Incident Coronary Heart Disease: A Multicohort Study of 90,164 Individuals. Epidemiology 2017 Jul;28(4):619-26.

(3) Dragano N, Siegrist J, Kivimaki M. Bias in occupational stress research - Authors' reply 59. Epidemiology 2018 Mar 21.

(4) Ingre M. Excuse me, but did the IPD-work consortium just "falsify" the job-strain model? Scand J Work Environ Health 2015 Sep 1;41(5):504-5.

(5) Ingre M. P-hacking in academic research: a critical review of the job strain model and of the association between night work and breast cancer in women Department of Psychology, Stockholm University; 2017.

(6) Ingre M, Andersen JH, Mikkelsen S. Re: Re: Effort-Reward Imbalance at Work and Incident Coronary Heart Disease 1. Epidemiology 2018 Jul 25.

(7) Kivimaki M, Nyberg ST, Batty GD, Fransson EI, Heikkila K, Alfredsson L, et al. Job strain as a risk factor for coronary heart disease: a collaborative meta-analysis of individual participant data. Lancet 2012

Oct 27;380(9852):1491-7.

(8) Kivimaki M, Singh-Manoux A, Virtanen M, Ferrie JE, Batty GD, Rugulies R. IPD-Work consortium: pre defined meta-analyses of individual-participant data strengthen evidence base for a link between psychosocial factors and health. Scand J Work Environ Health 2015 May 1;41(3):312-21.

Reference List (continued)

(9) Kivimaki M, Nyberg ST, Kawachi I. Authors' reply: Calculation of population attributable risk should to be based on robust estimates. Scand J Work Environ Health 2015 Sep 1;41(5):506-7.

(10) Kronmal RA. Spurious Correlation and the Fallacy of the Ratio Standard Revisited. Journal of the Royal Statistical Society Series A (Statistics in Society) 1993;156(3):379-92.

(11) Madsen IE, Nyberg ST, Magnusson Hanson LL, Ferrie JE, Ahola K, Alfredsson L, et al. Job strain as a risk factor for clinical depression: systematic review and meta-analysis with additional individual participant data. Psychol Med 2017 Jan 26;1-15.

(12) Madsen IH, Rugulies R, Kivimäki M. Job strain and clinical depression – authors' reply. Psychological Medicine 2017;doi:10.1017/S0033291717001647:1-2.

(13) Mikkelsen S, Bonde JP, Andersen JH. Analysis of job strain effects. Occup Environ Med 2011 Oct;68(10):786.

(14) Mikkelsen S, Andersen JH, Bonde JP, Hansen ÅM, Kolstad H, Thomsen JF. Job strain and clinical depression. Psychological Medicine 2017;19(Jun):1-2

(15) Mikkelsen S, Andersen JH, Ingre M. Artefacts in occupational stress research. 58. Epidemiology 2018 Mar 21.