# Quantifying social determinants of health using qualitative inputs and structural equation modelling

Laurens Reumers, Maastricht University
✉ l.reumers@maastrichtuniversity.nl
Niels Hameleers, Maastricht University
Henk Hilderink, RIVM
Marleen Bekker, Wageningen University & Research
Maria Jansen, Maastricht University; South Limburg
Jan-Kees Helderman, Radboud University Nijmegen
Dirk Ruwaard, Maastricht University

Maastricht University

ALLES IS
GEZONDHEID

ZonMw

# Background: our starting point

- Social determinants of health

- Complex interrelationships, 'wicked problems'

**Comments slide 2**

Starting point: our evaluation of Dutch national programme 'All about Health'. Collaborative platform; large number of partners; prevention side.

Nearly all (societal) initiatives affiliated with the programme are concerned with social determinants of health (SDOH).

Problem with quantifying SDOH: there are usually multiple determinants working together, affecting each other. Obtaining relationship estimates (coefficients) from seperate models results inextremely biased model of how variables interact, so one estimation model needed.

1. Which to include? 2. How to quantify?

# Two-step methodological approach

1. Obtaining the conceptual model

2. Quantifying that model

**Comments slide 3**

This presentation: methodological approach to studying health effects from SDOH.

Qualitative part: What variables are relevant and should be included? What is our expectation of the relationships between them?

Then on to quantification.

# Group model building: obtaining a conceptual model

- Method to consult stakeholders

- Experiential knowledge

- Used for modelling complex systems

**Comments slide 4**

Group model building (GMB): participatory method putting a group (8 to 20) stakeholders together; interactively model problem or situation using their experiential/expert knowledge.

Decisions about which variables are and are not deemed relevant are made by a group of actors with deep knowledge of the situation. Still subjective, but not arbitrary.

More information on the GMB method and process itself can be found in literature - it is quite well-documented.

We have also written an article on a practical application that will come out soon:

Reumers LM, Bekker MPM, Hilderink HBM, Jansen MWJ, Helderman JK, Ruwaard D (forthcoming). Qualitative modelling of social determinants of health using group model building: The case of debt, poverty, and health. *International Journal for Equity in Health*.

# Demonstration case: poverty and health

## Qualitative model

Full stakeholder
model
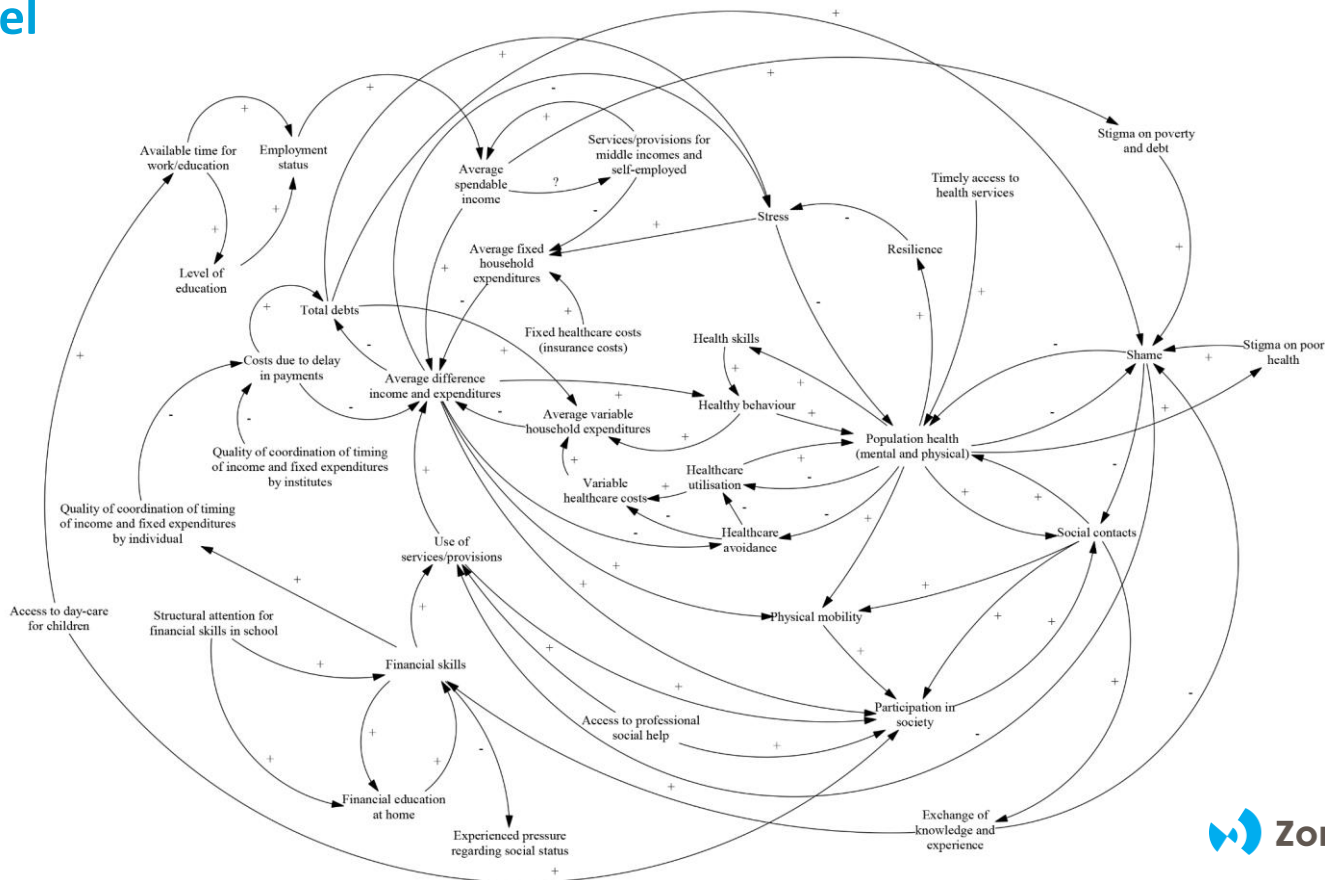
**Comments slide 5**

Causal loop diagram (CLD = a qualitative system dynamics model) made in the practical application mentioned on the previous slide. Model was made by stakeholders in Dutch city of Utrecht, the setting of our (real-world) demonstration case. Good example of how extensive and complex it easily gets.

Can be read as standard conceptual models. The single-headed arrows are hypothesised causal arrows; '+' means positive relationship ('the more X, the more Y'), '-' means negative relationship ('the more X, the less Y', and 'the less X, the more Y').

The system structure indicates that any initiative that does an intervention on any of the variables in the model is likely to indirectly affect a lot of the other variables as well. Showing how this works is also the purpose of the model here.

# Demonstration case: poverty and health

## Qualitative model

Simplified model
based on stakeholder
model
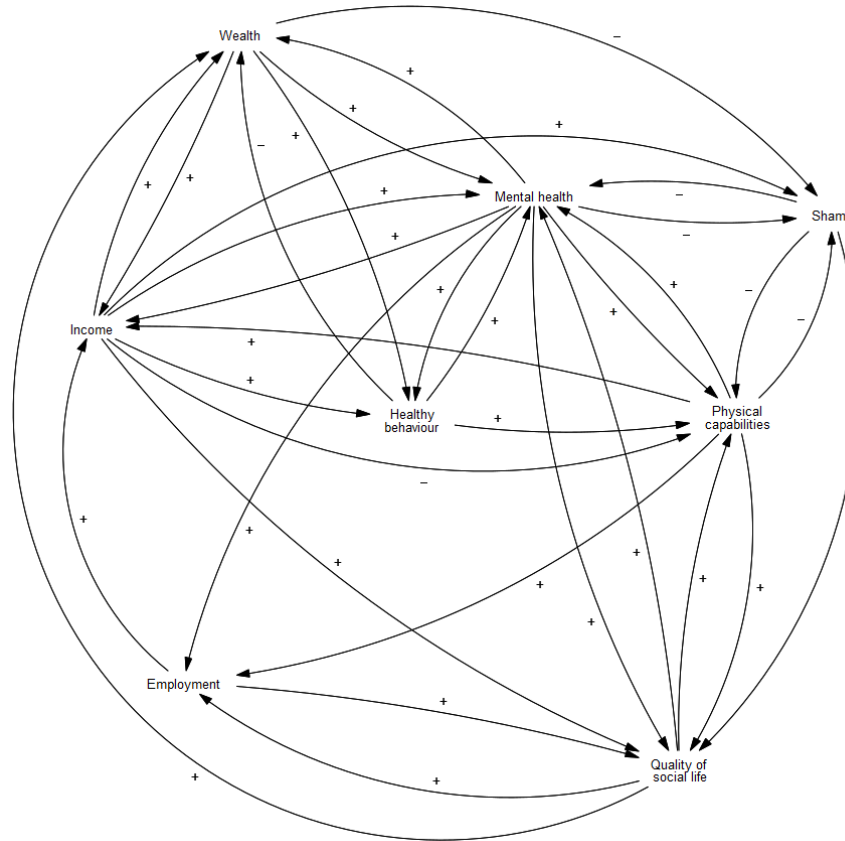
**Comments slide 6**

Some simplifications had to be made for quantification. Also made use of review session with scientific experts and of scientific literature.

Some necessary changes made for the sake of operationalisation. E.g., health not one concept, so split up into mental health and physical capabilities components.

Resulting model still heavily based on the full stakeholder model, but contains fewer variables (from 39 to 8 variables). From 71 relationships in the full model to 34.

A difficulty with quantification is precisely one of the strengths of qualitative model: it has feedback loops. [continue to next slides]

# The problem with feedback loops
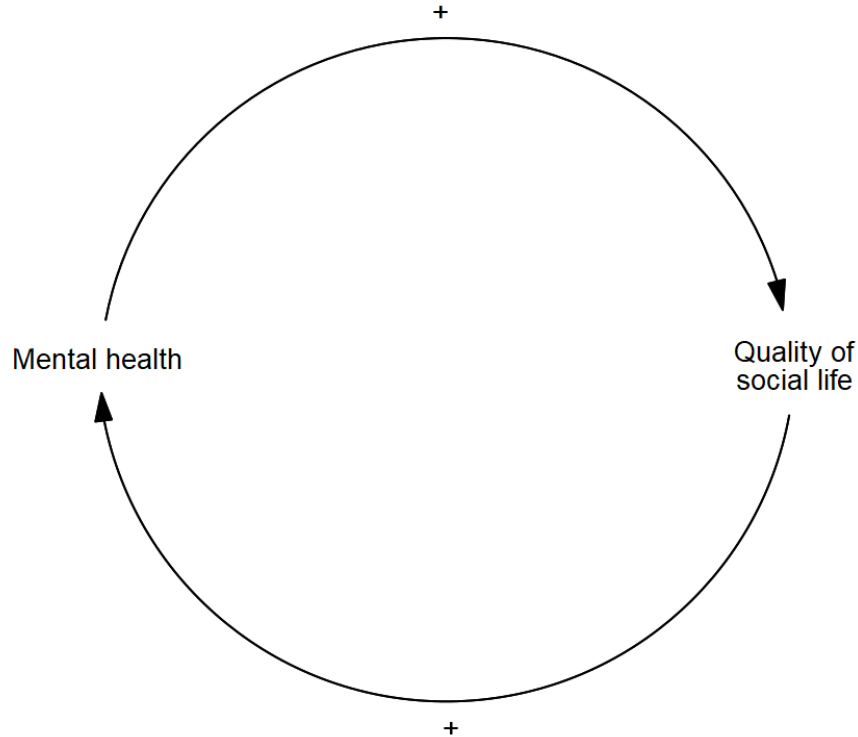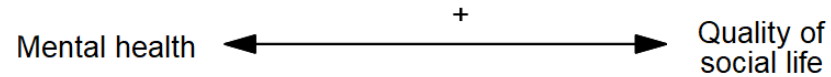
**Comments slide 7**

Simplest feedback loop possible.

Better mental health leads to a better social life and a better social life improves your mental health.

To be more careful with causality: mental health *predicts* quality of social life and vice versa. A feedback loop can also include more variables.

In this figure: two directional arrows = problematic to quantify with ordinary regression analysis. Instead of two separate causal arrows, a regression analysis would do: [see next slide]

# The problem with feedback loops



Mental health ⟵————— + —————⟶ Quality of social life

**Comments slide 8**

A regression analysis would do this.

Gives you only one coefficient, with the strength of the correlation, but no directionality.

Longitudinal designs (data over multiple time points) can help here. If event A occurs before event B, B could not have caused A. Temporality enables some sense of directionality.

We use structural equation models (SEMs). A lot of different sorts of SEMs exist, each with their strengths and flaws (well-documented in literature).

We specifically use an autoregressive cross-lagged panel model with latent components. A relatively older method, had its criticisms over the years, but probably best-suited for obtaining useful single coefficients for each of the arrows in the CLD (qualitative model)

# Longitudinal design: autoregressive cross-lagged panel model
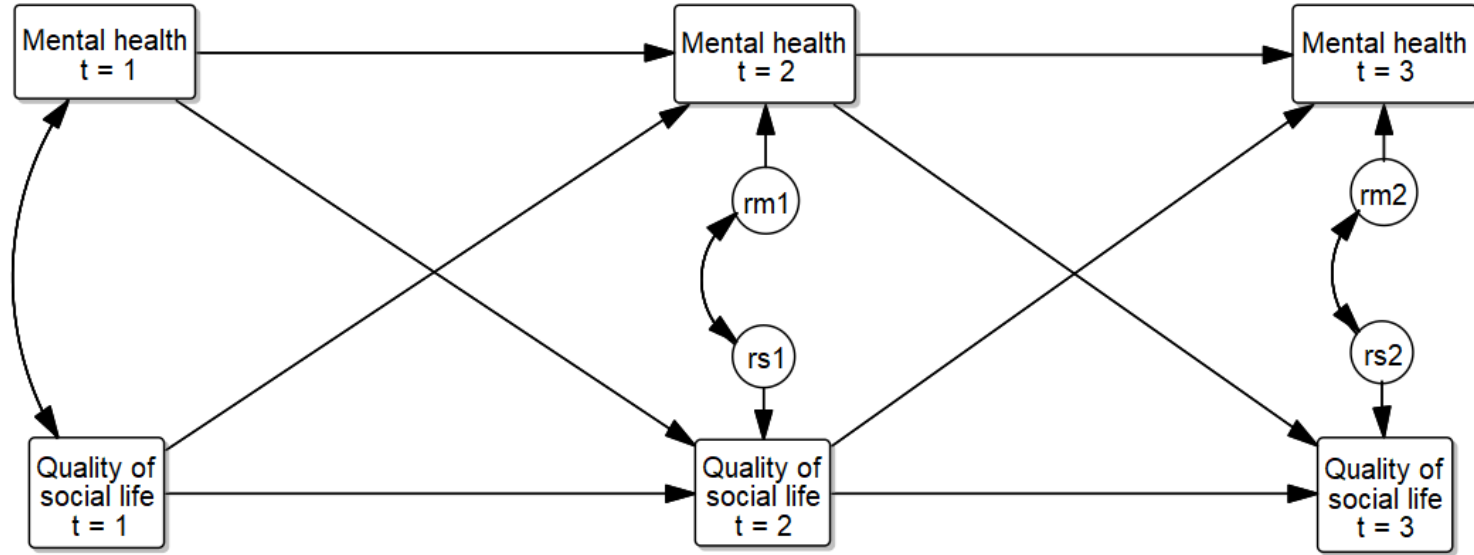
**Comments slide 9**

Simple autoregressive cross-lagged panel model. The same two variables, mental health and quality of social life, now measured at three different time points (denoted with 't'). In some uses, two time points may be enough, and more than three time points is generally even better, although more difficult to gather data on.

Called an '*auto*regressive' model because the variables at time points t=2 and t=3 are regressed on their previous value. The cross-lags are the crossed arrows in the middle – the effects we are most interested in: Does mental health at one time predict quality of social life at the next? And vice versa?

Important: even though we have directional arrows this does not mean that they are necessarily causal arrows. They show Granger-causality: can be seen as indication of directionality, but still a danger of confounding omitted variables that influence the relationships. Very important to keep in mind.

The covariance arrows between variables at the same time are necessary to control for things that happen around the same time and where any temporal precedence cannot be ascertained.

Can also add latent variable constructs in the model, which we have also done. This means variables are measured with multiple indicators (with confirmatory factor analysis, CFA), which reduces measurement error. A great advantage of SEMs.

Subsequent figues will leave out covariance arrows and residuals, for simplicity.

# Demonstration case: poverty and health

## Qualitative model

Simplified model
based on stakeholder
model



Maastricht University
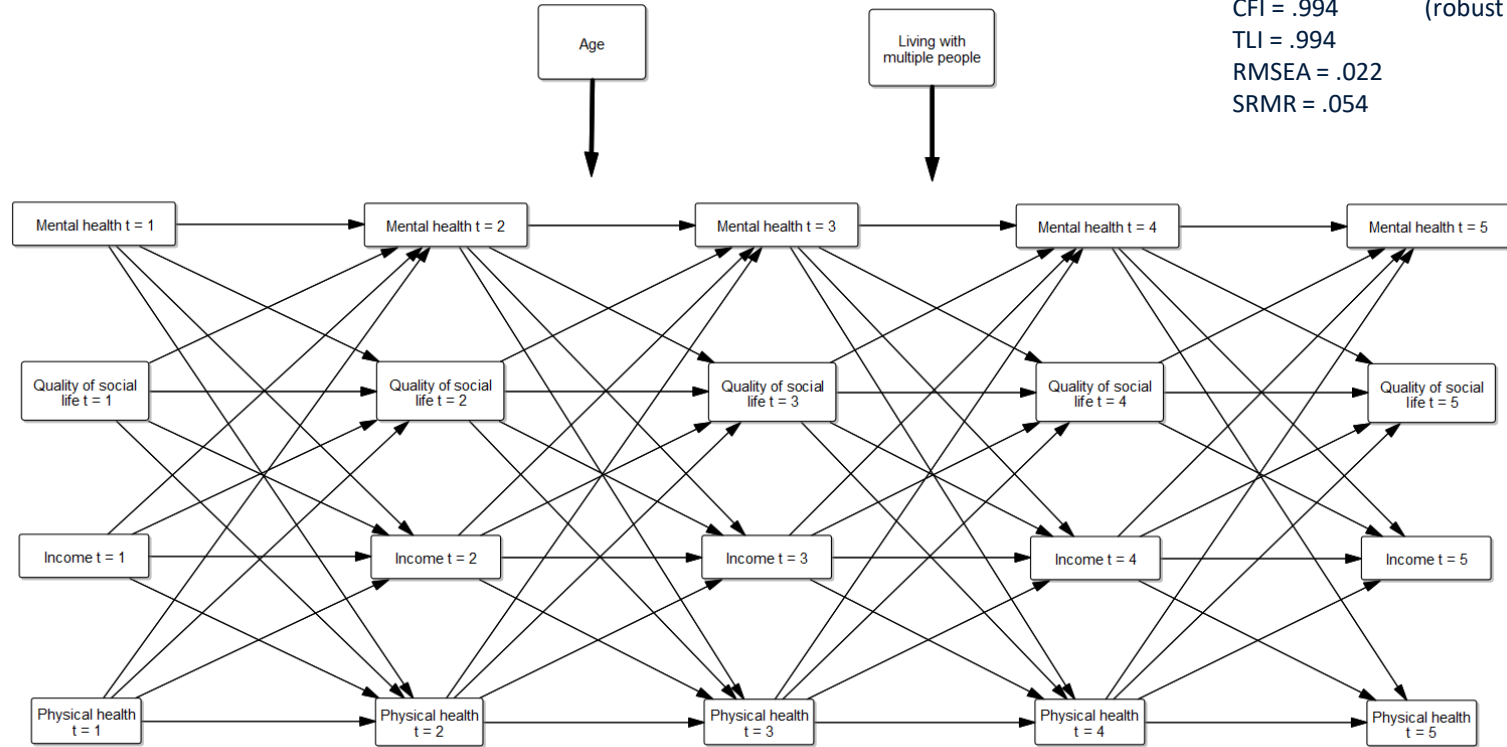
ZonMw
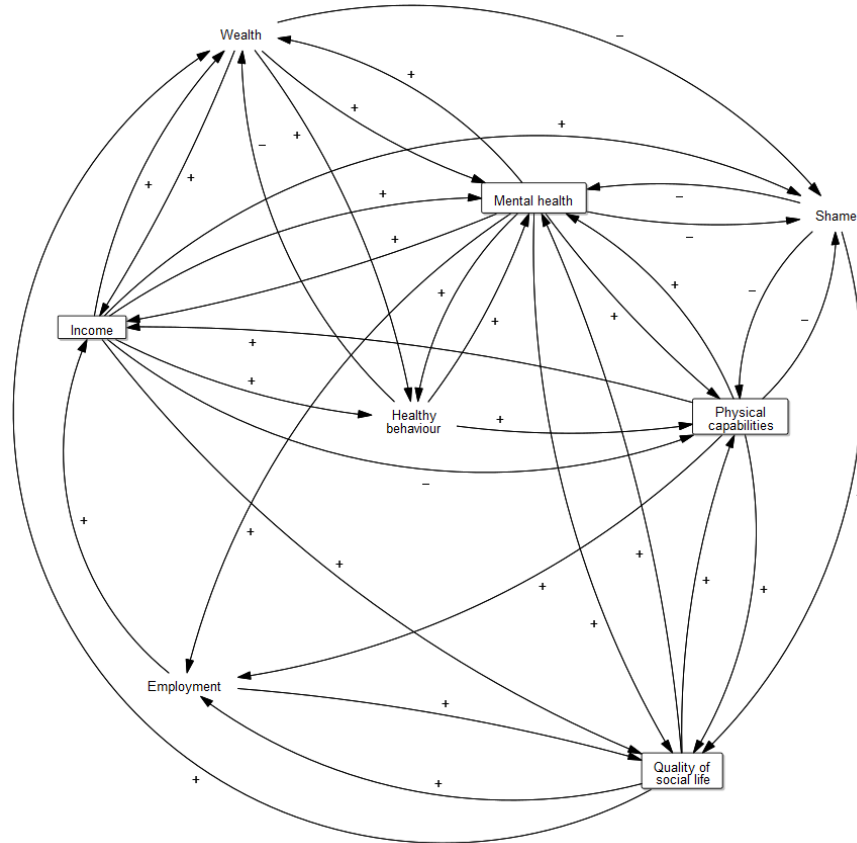
**Comments slide 10**

For this qualitative CLD, we have so far quantified four variables (indicated by the boxes) together, along with two control variables.

Data requirements quite high: dataset with measurements of same individuals repeated at multiple time points. More variables = more individuals (N) needed.

Our N here was over 15,000, which is a lot more than needed. Several hundred is sometimes enough, provided a smaller number of variables. Several thousand should usually be enough.

# Demonstration case: poverty and health

## Quantitative model structure

N = 16045
CFI = .994          (robust fit statistics)
TLI = .994
RMSEA = .022
SRMR = .054



Maastricht University

ZonMw

**Comments slide 11**

Some very preliminary outputs, quickly exported specifically to show in this presentation.

Model was constrained so that the effects were equal over time (from year 1 to year 2, from year 2 to year 3, etc.), so each effect from one variable on another (and on itself) was denoted with one coefficient.

Better way to visualise this on next slide.

# Demonstration case: poverty and health

## Quantitative model results

N = 16045
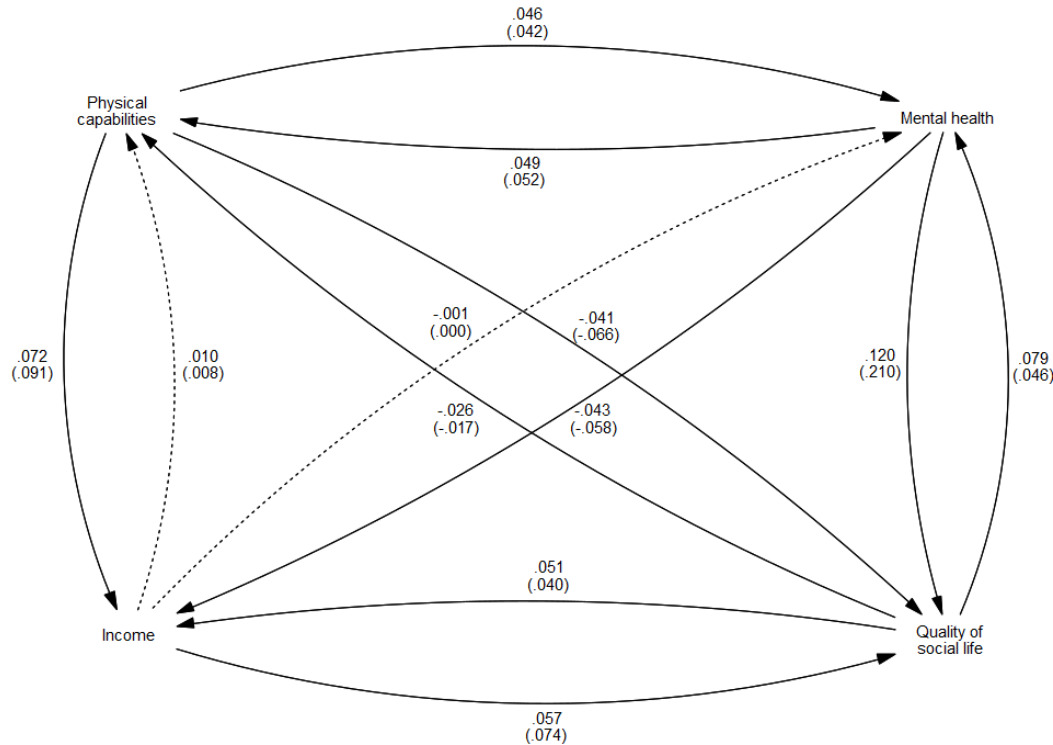CFI = .994        (robust fit statistics)
TLI = .994
RMSEA = .022
SRMR = .054

**Comments slide 12**

CLD notation with four variables and coefficients. We did also control for age and whether someone lives with other people. Coefficients can be interpreted as linear regression coefficients. The standardised ones (the top ones) can be compared to each other, to compare effect sizes.

Some things we would expect, specifically all arrows on the outside. Mental health and quality of social life seem to affect each other the most. Largest effect that of mental health on social contacts. Physical capabilities have clear effect on income, and both income and social contacts, and physical capabilities and mental health seem to affirm each other as well.

The arrows in the middle appear pretty counter-intuitive. With some speculation, there may be a confounder at play concerning employment. Adding something about employment, and specifically work hours or pressure may change the relationship between income and mental health. Adding the variables of health behaviour and wealth (present in the CLD) may also change results some.

In any case, thinking hard about possible confounders is still important for the validity and accuracy of the results. But for now, surprising results are also results; we should not try to ignore them, of course.

As said: preliminary results, so prone to change as model is expanded, but this may give some idea of what such a model could look like.

# In conclusion

**What can we do with this?**

- Better understand how complex interrelationships work

- Estimate indirect effects of activities

**Comments slide 13**

Connection with health impact assessment? Building the model and obtaining the coefficients is not HIA itself.

Both the qualitative model-building and subsequent quantification may improve understanding about the complex relationships in the system of variables under study. Valuable in itself.

Additionally, such models can be used as tools for estimating indirect effects of activities.

E.g., intervention focused on poverty with a known effect size on household income in a population. Model from our demonstration case can give an estimate of how that is expected to affect all the variables in the model over time.

Not very easy to construct a model like; takes quite some time and effort. But, upside: can be re-used for multiple assessments on different activities in different settings (feeding the model local data, where available).

Currently also developing this for our model: possibility to be applied to other settings. Should not require any modelling knowledge, but could be done by just filling in some baseline population characteristics. Such tools could be quite useful for helping with HIA.