



# Quantitative HIA for long term air pollution exposure – an example of an end-to-end health burden study

Ivan Hanigan PhD

Director, WHOCC for CC&HIA

Curtin University School of Population Health



This work is licensed under an  
Attribution 4.0 licence ([NB caveat](#)).

# Masterclass aims and objectives

- Aims:
  - Enhance your understanding of quantitative HIA methods for air pollution,
  - Contribute to the development for future iterations
- Objectives
  - Demonstrate an end-to-end example of health burden
  - A worked example for hands-on “Coached” practice
  - Gather participant feedback using the Slido survey:  
at Slido.com with #3370 651



# Worked example and tools

- To support the masterclass we will use ‘Coached’ practice
- There is a set of our tools we provide from Github
  - [https://github.com/cardat/DatSciTrain\\_HIA\\_Health\\_Impact\\_Functions\\_explained](https://github.com/cardat/DatSciTrain_HIA_Health_Impact_Functions_explained)
- We will start by looking at the 02\_Lifetables\_with\_spreadsheet
  - We will not have time to look at the other tools today
  - We will not have time to troubleshoot your computing environment
  - Feel free to play around with the other tools later, and send any feedback via Github issues or pull requests.

# Quickest way is to click ‘Code’ and Download Zip

Screenshot of a GitHub repository page for "DatSciTrain\_HIA\_Health\_Impact\_Functions\_explained". The repository is public and has 1 branch and 0 tags. The main branch is selected. A pull request by ivanhagan is visible.

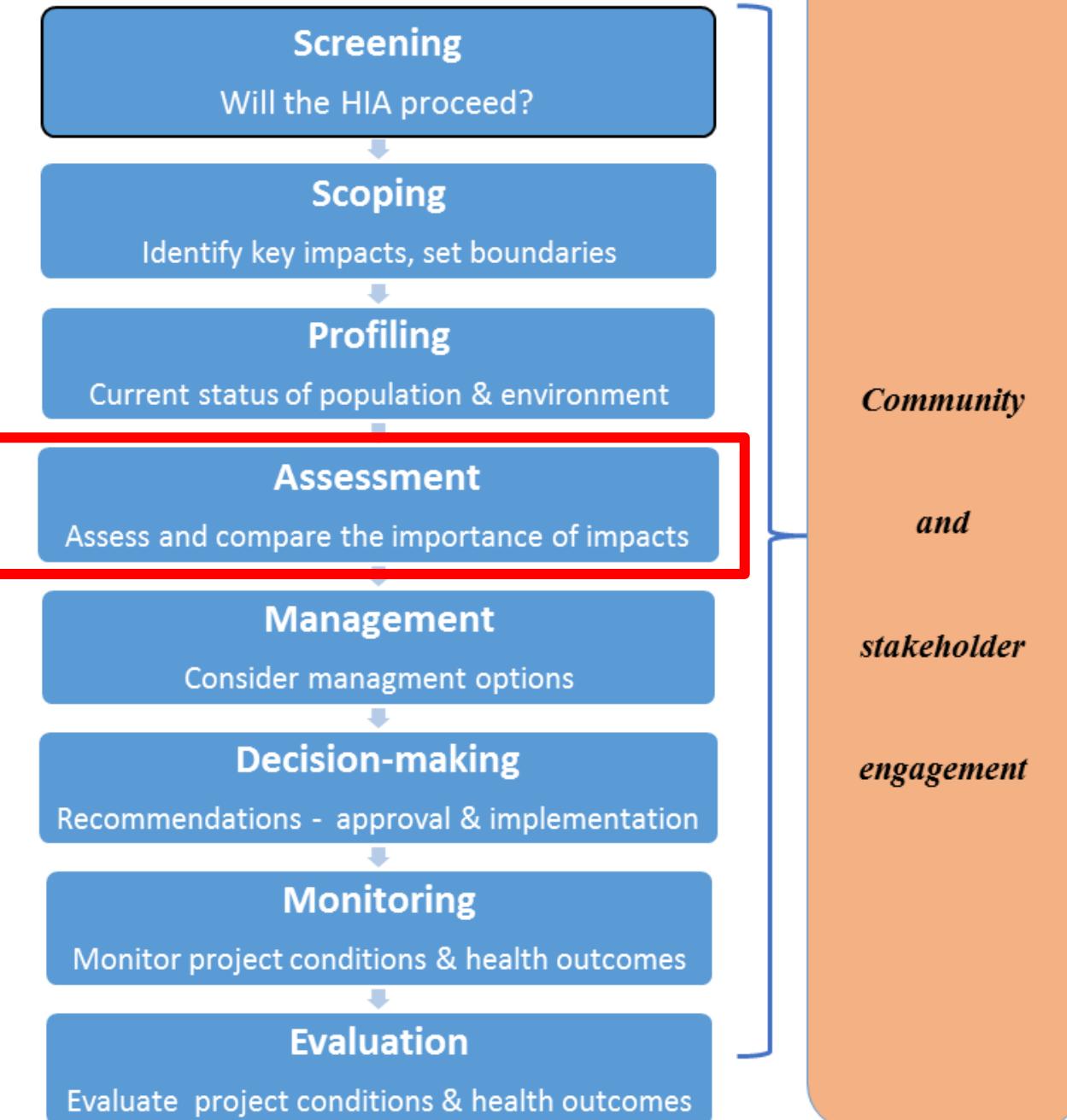
The "Code" button is highlighted in green, and a dropdown menu is open, showing options for cloning the repository via HTTPS, SSH, or GitHub CLI. The SSH link is copied to the clipboard. Other options include "Local" (selected), "Codespaces", "Clone", "Open with GitHub Desktop", and "Download ZIP".

File/Folder	Description	Last Commit
01_HIF_explainer_with_R	restructure and add	36 minutes ago
02_Lifetables_with_spreadsheet	add lifetable with sp	
03_Lifetables_with_R	add the lifetables w	
04_Regression_models_with_R	restructure and add	
05_System_dynamics_with_Vensim	add system dynami	
DatSciTrain_HIA_Health_Impact_Functions_expl...	restructure and add	
LICENSE	restructure and add new examples	36 minutes ago
README.md	restructure and add new examples	36 minutes ago

# The HIA Process

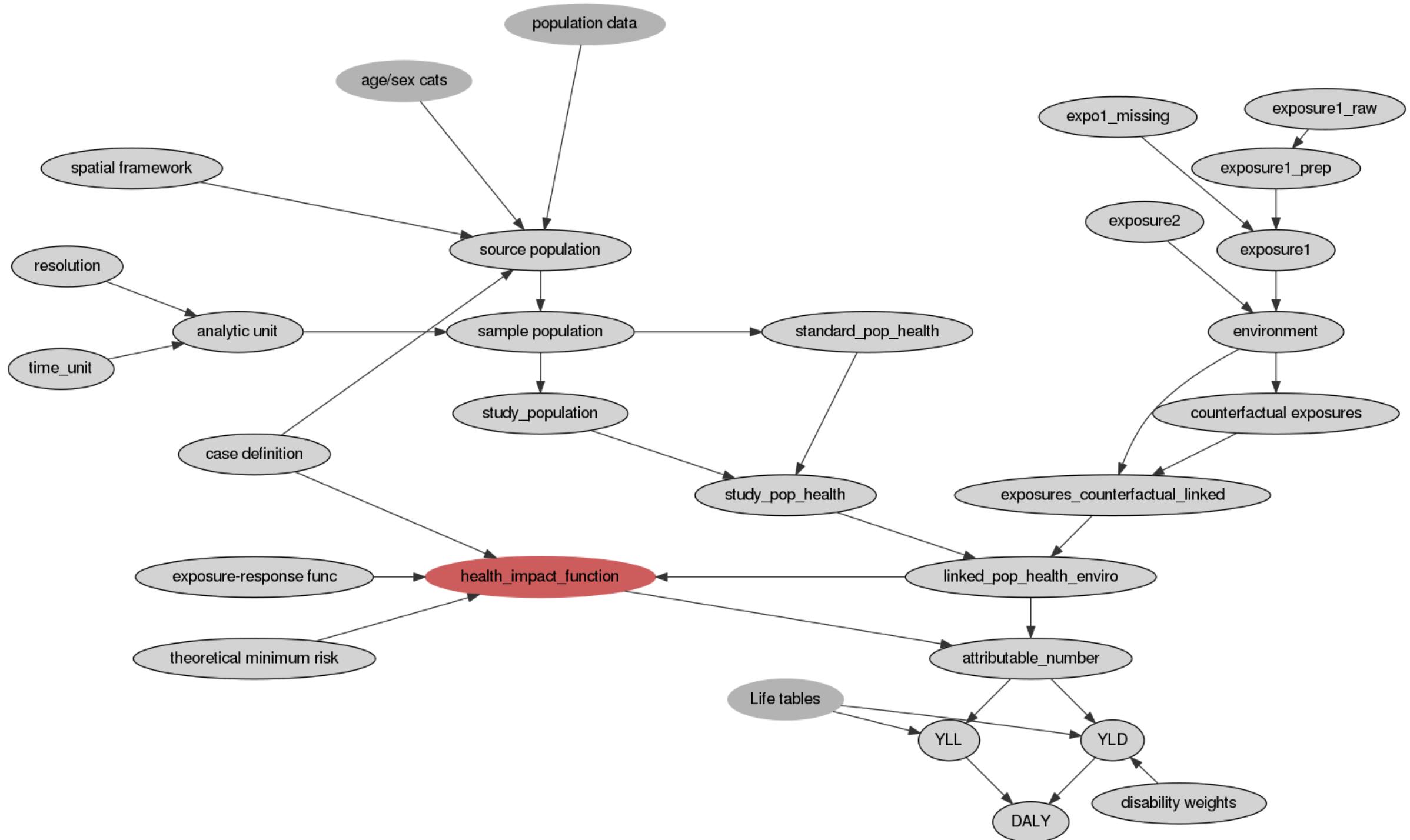
## Aus enHealth 17

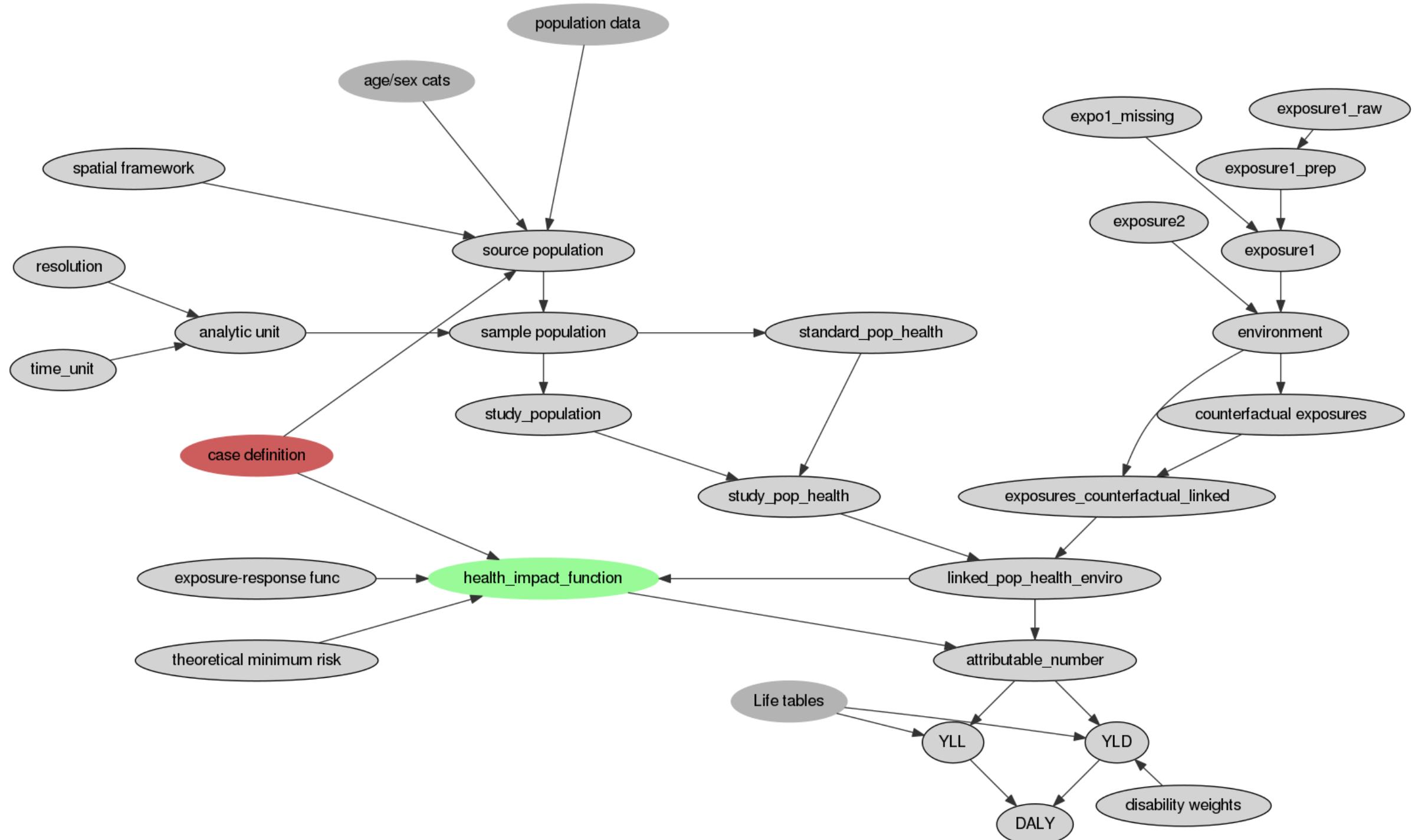
- Similar/same in other countries
  - Principles:
    1. Democracy
    2. Equity
    3. Sustainable Development
    4. Ethical use of evidence
    5. Comprehensive approach to health

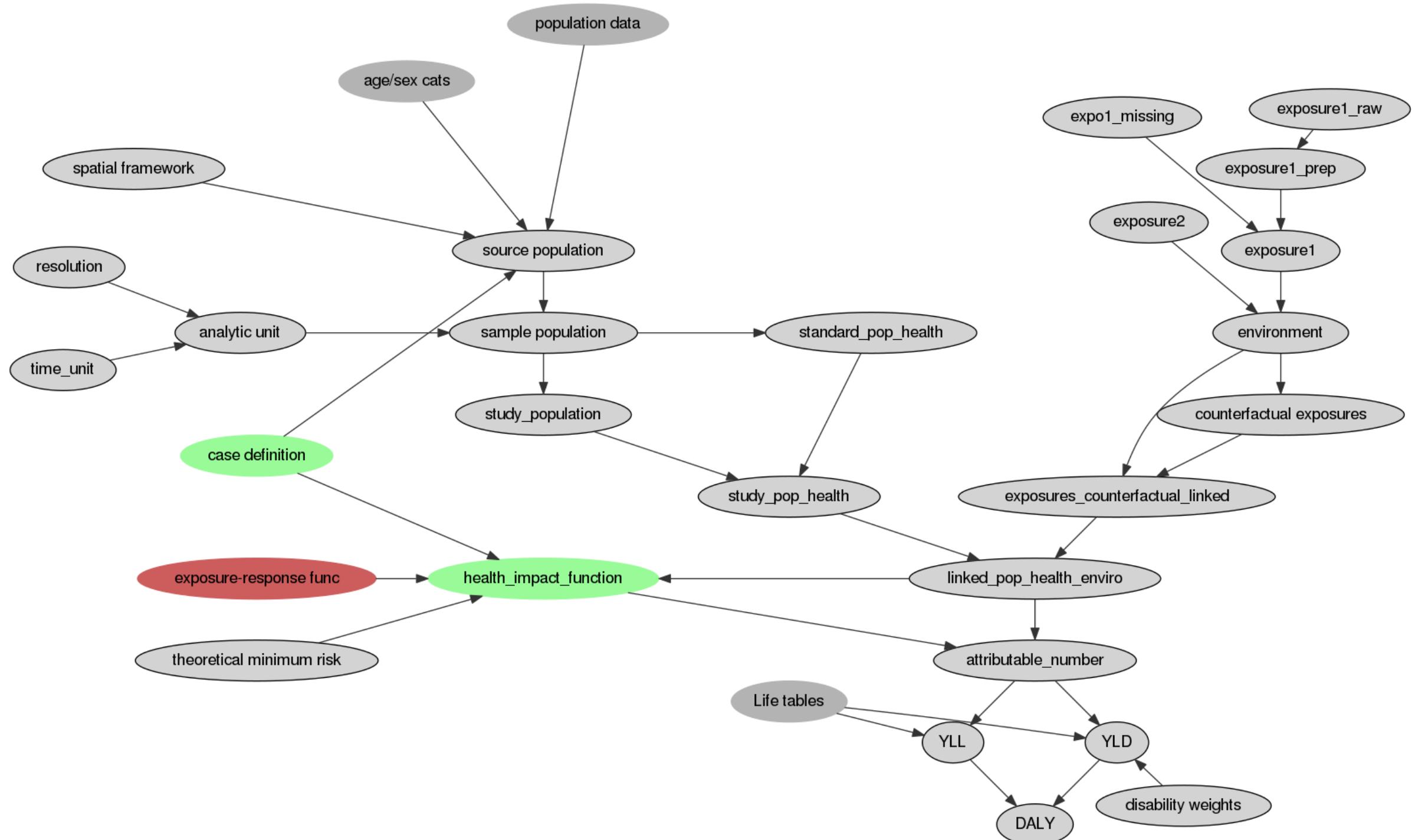


Curtin University

WHO COLLABORATING CENTRE FOR CLIMATE  
CHANGE AND HEALTH IMPACT ASSESSMENT







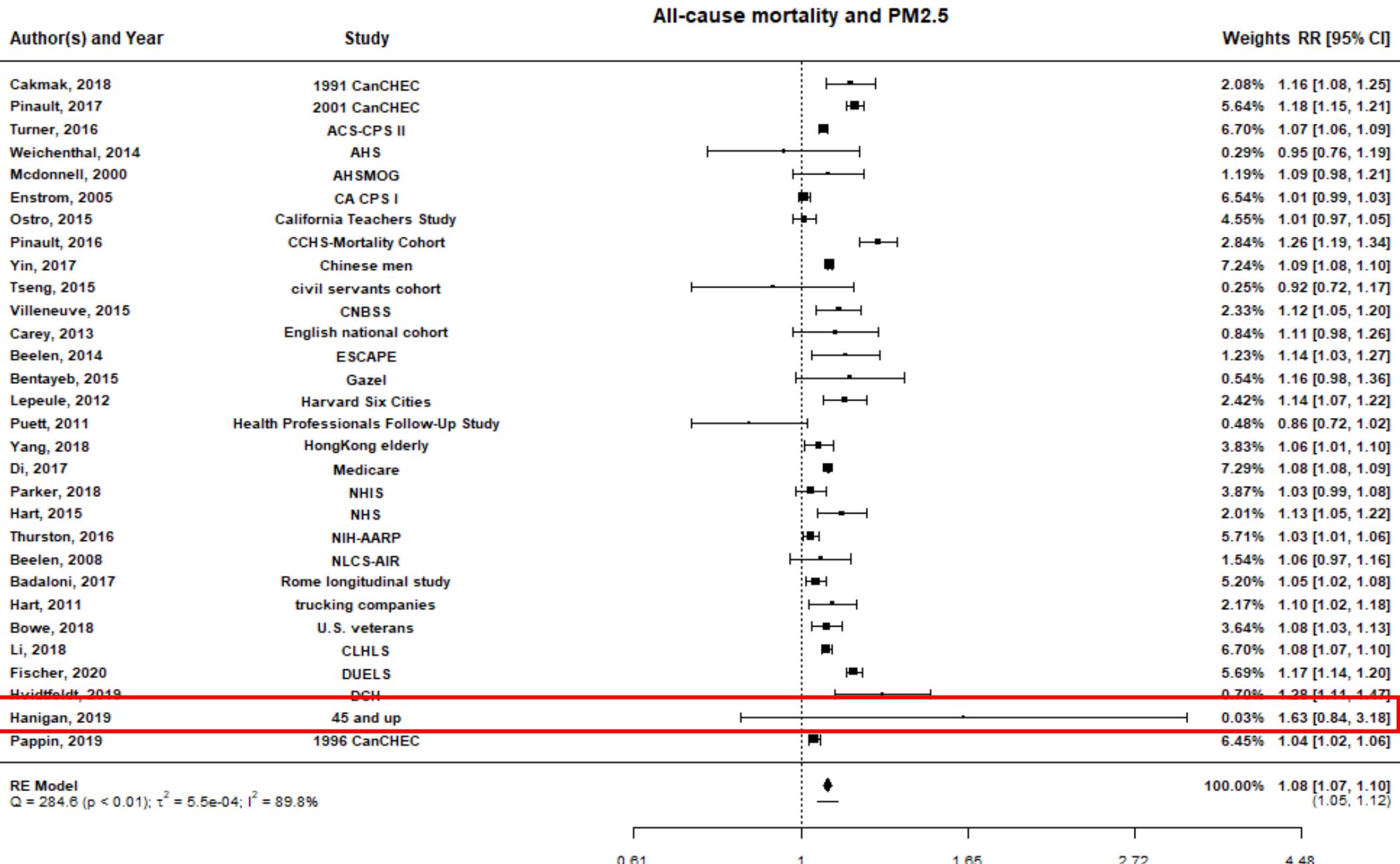
# 2020 WHO commissioned report

For long-term exposure impacts assessment

Pollutant metric	Health outcome	RR (95% CI) per 10 ug/m3	Range of concentrations
PM2.5, annual mean	Mortality, all-cause (natural) adults.	1.08 (95%CI:1.06, 1.09) per 10 µg/m3 All increase in PM2.5	All
PM10, annual mean	Mortality, all-cause (natural) adults.	1.04 (95%CI:1.03, 1.06) per 10 µg/m3 All increase in PM10	

A population consistently exposed to 10 µg/m<sup>3</sup> more PM<sub>2.5</sub> are expected to have ~1.08 times (or ~8%) more “non-injury” deaths compared to a counterfactual population, in the age 30+ group. RR is the exposure-response function. Chen & Hoek. (2020) systematic review and meta-analysis. Environment International <https://doi.org/10.1016/j.envint.2020.105974>

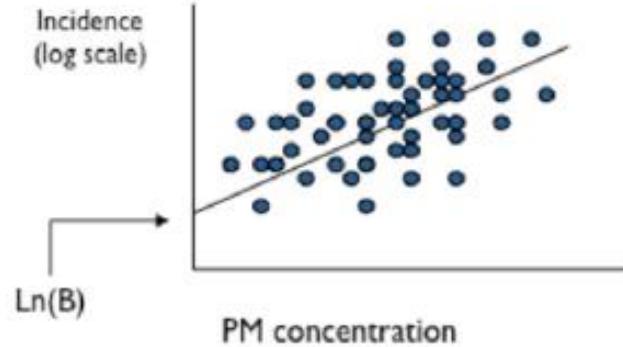
# We class Epidemiology studies and meta-analysis as 'primary', HIA 2nd'ary



Chen and Hoek,  
Environ. Int.  
2020.

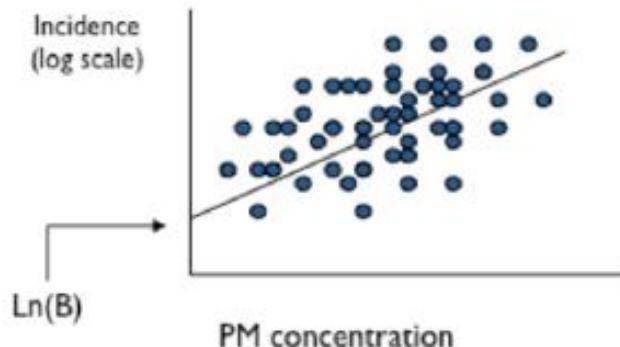
## Epidemiology study

- 1) Exposure  
AND
- 2) Health impact  
function  
AND
- 3) Population  
AND
- 4) Incidence rate



Adapted from BenMap toolkit USA user guide

## Epidemiology study



$$\ln(y) = \ln(B) + \beta(\text{PM})$$

## Health impact function

$$\Delta Y = (1 - e^{-\beta \Delta \text{PM}}) * \text{Pop} * Y_0 - \text{Baseline Incidence}$$

$\beta$  - Effect estimate

$\Delta \text{PM}$  - Air quality change

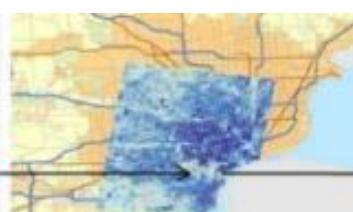
**Pop** - Exposed population

Attributable Fraction =  
$$(RR-1) / RR$$
  
$$1 - (1/RR)$$

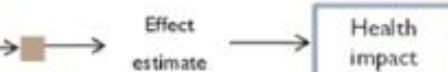
## Exposure



## Population



## Incidence rate

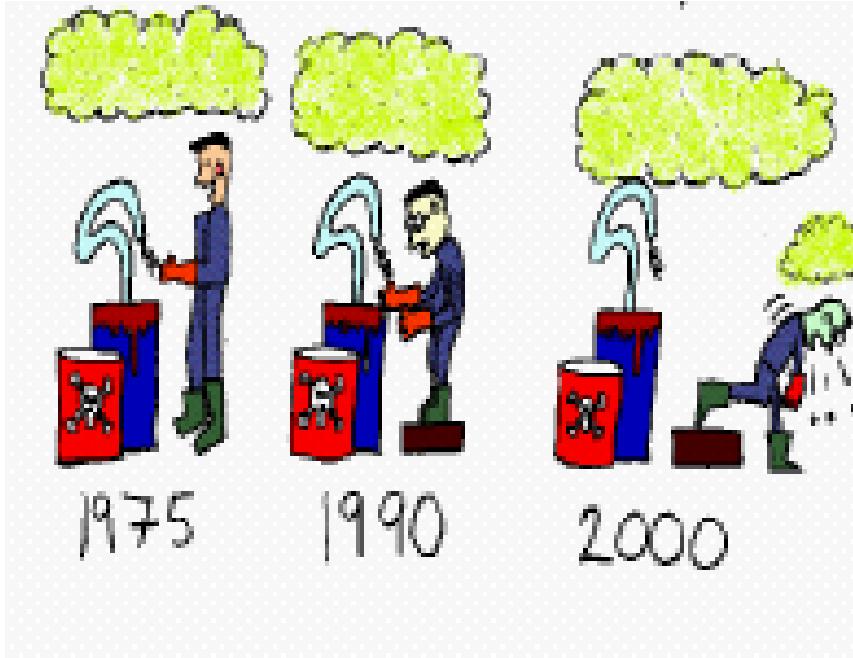


Adapted from BenMap toolkit USA user guide

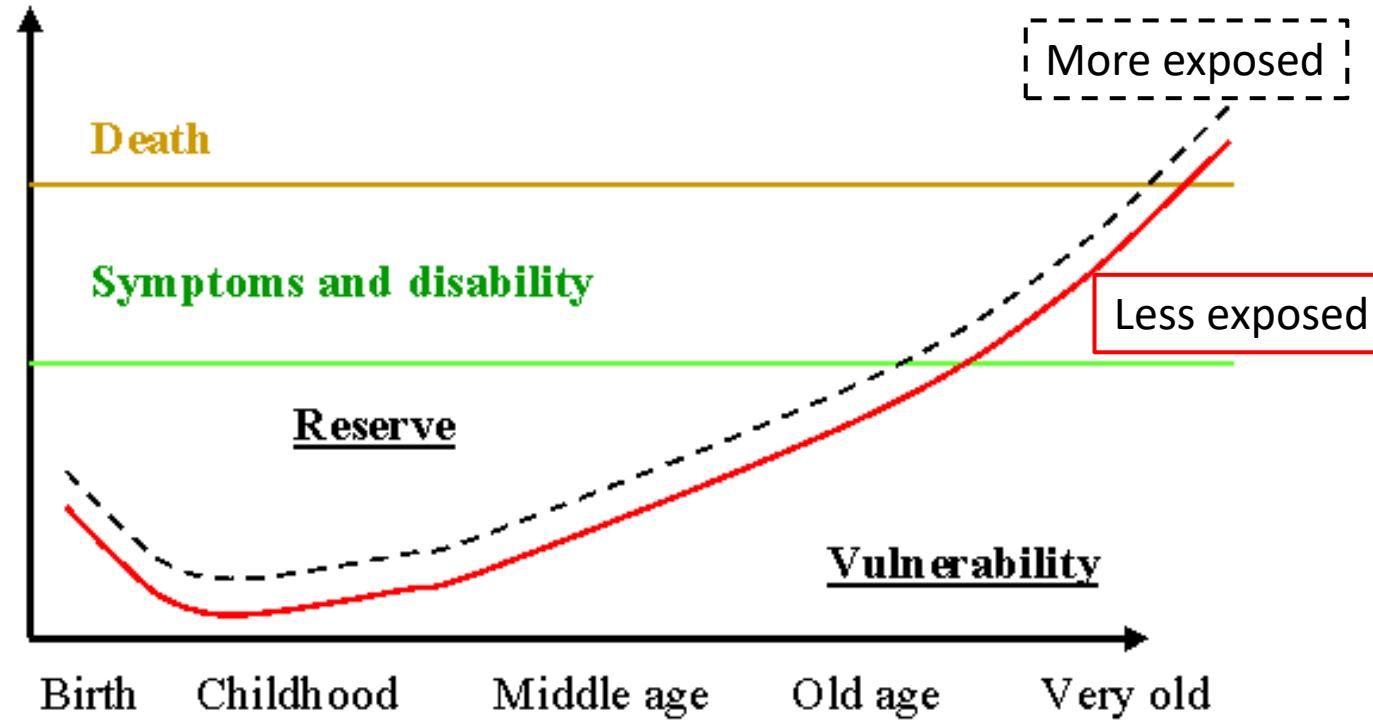
- 1) Exposure  
AND
- 2) Health impact  
function  
AND
- 3) Population  
AND
- 4) Incidence rate

# Long-term exposures

## Chronic Toxicity



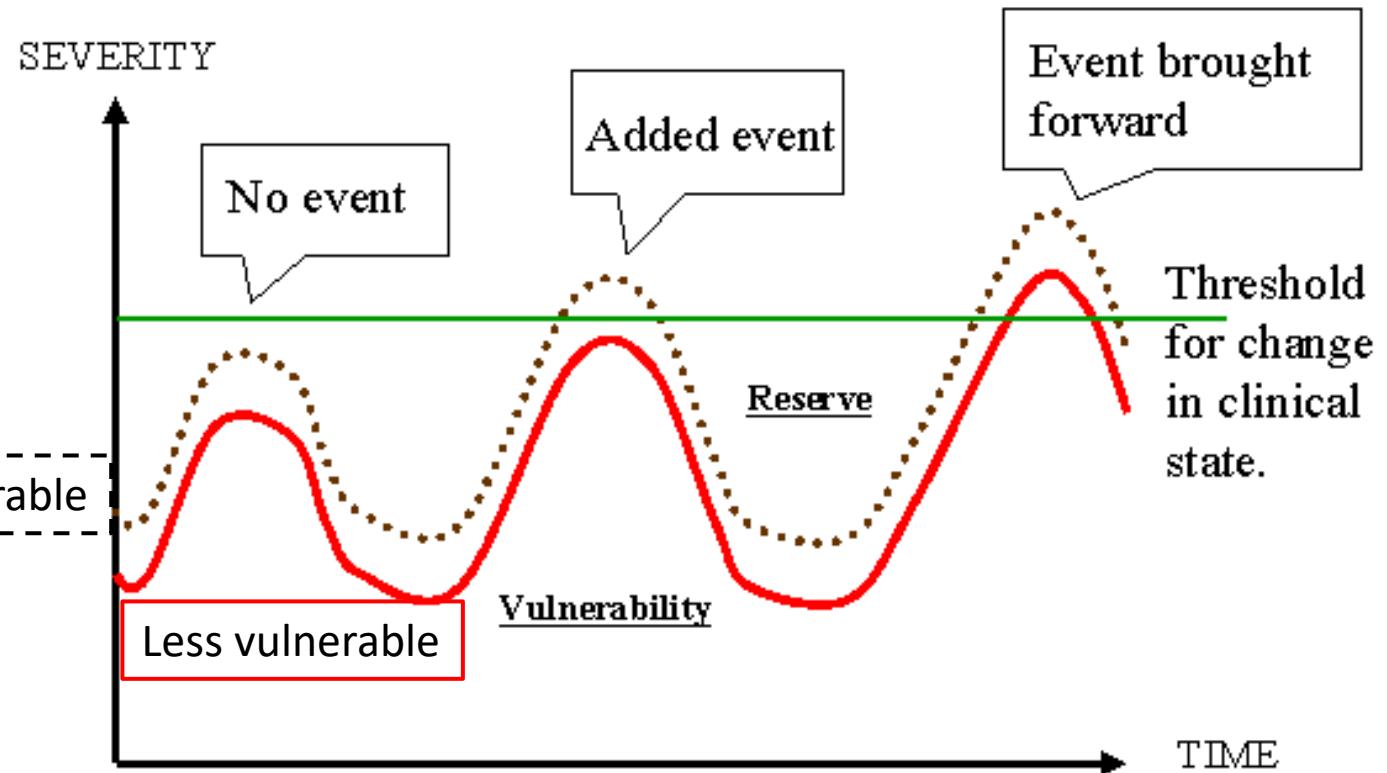
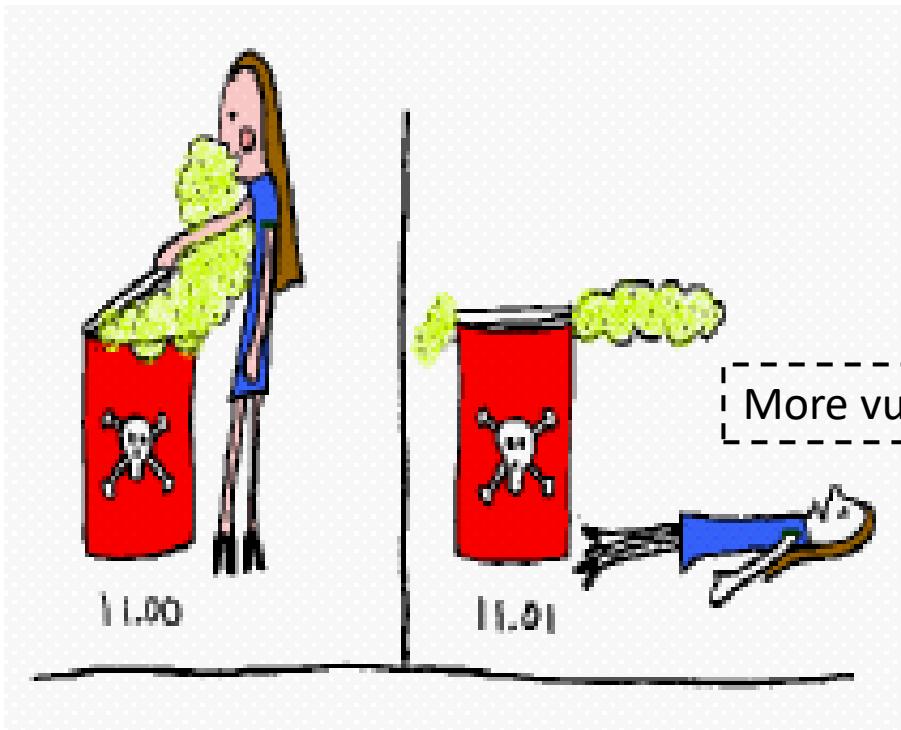
## Role of air pollution in chronic disease



Source: unknown

# Short-term exposures (and other Stressors – e.g. viruses)

## Acute Toxicity



Source: unknown



How are you?



Compared to  
what?



Curtin University

WHO COLLABORATING CENTRE FOR CLIMATE  
CHANGE AND HEALTH IMPACT ASSESSMENT

# Health Impact Function for ‘excess risk’ or ‘attributable fraction’

Using formulae to describe this we can present the two approaches as below. First the ‘excess risk formula’ is:

$$\frac{\text{Observed}}{\text{Exposed}} = RR \times \frac{\text{Expected}}{\text{Un-exposed}}$$

Where the RR is calculated for the level of exposure observed in the study population.

The attributable number (AN) can be calculated as:

$$AN = (RR - 1) \times \text{Expected}$$

This is sometimes expanded when the HIF is estimated using log-linear regression models of continuous exposure levels and applied to multiple subpopulations into this equivalent formula:

$$AN_i = \sum(e^{(\beta \times X_i)} - 1) \times \text{Expected}_i$$

Where  $X_i$  is the exposure for  $\text{studypopulation}_i$ .

The alternative formulation is the ‘attributable fraction formula’:

$$AN = (RR - 1)/RR \times \text{Observed}$$

Which is often simplified to

$$AN = (1 - (1/RR)) \times \text{Observed}$$

Equivalent to

$$AN = \sum(1 - e^{(-\beta \times X_i)}) \times \text{Observed}_i$$

## FURTHER READING

Appendix 1: Hanigan, et al. (2020).  
Environmental health indicators.  
Human Health and Social Impacts  
(HHSI) Node and the NSW DPIE.  
DOI:10.17605/OSF.IO/YJ98D

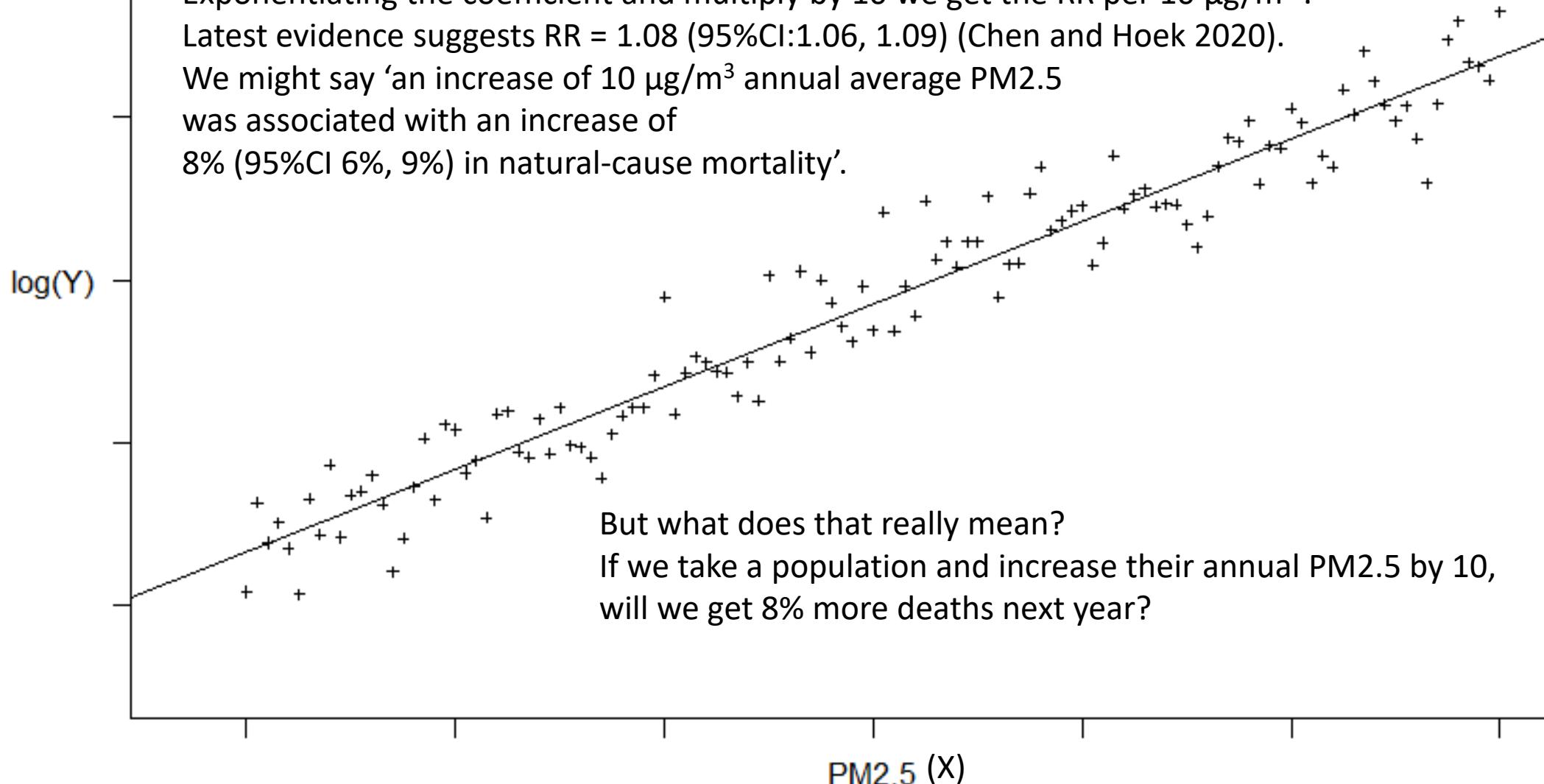
## Regression model $Y = \beta_0 + \beta_1 X + \beta_2 Z + [...]$

$\beta_1$  is the log-linear coefficient representing the slope of association between  $Y$  (e.g. mortality) with  $X$  (in this case PM2.5).

Exponentiating the coefficient and multiply by 10 we get the RR per 10  $\mu\text{g}/\text{m}^3$ .

Latest evidence suggests RR = 1.08 (95%CI:1.06, 1.09) (Chen and Hoek 2020).

We might say 'an increase of 10  $\mu\text{g}/\text{m}^3$  annual average PM2.5 was associated with an increase of 8% (95%CI 6%, 9%) in natural-cause mortality'.

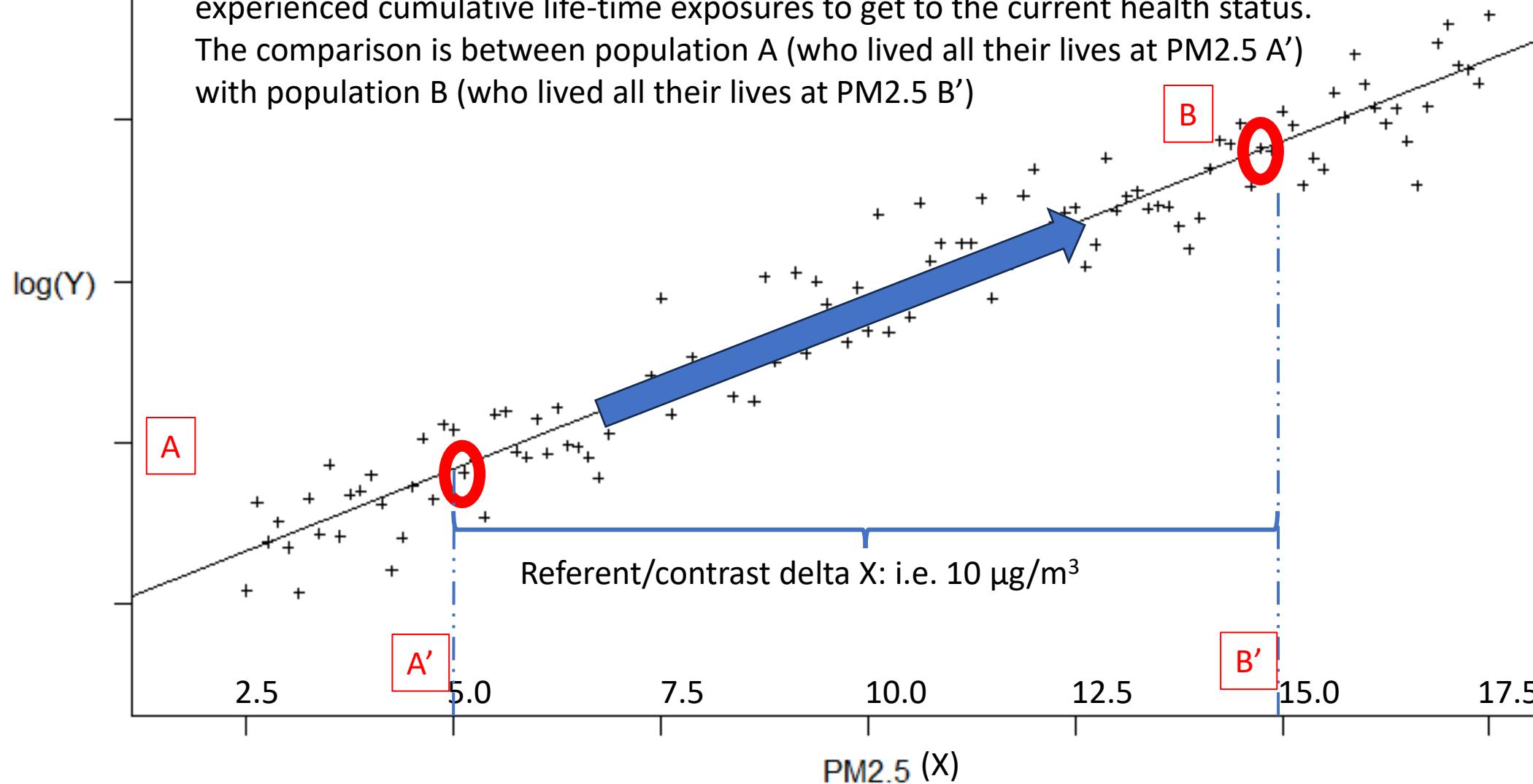


$$\text{Regression model } Y = \beta_0 + \beta_1 X + \beta_2 Z + [...]$$

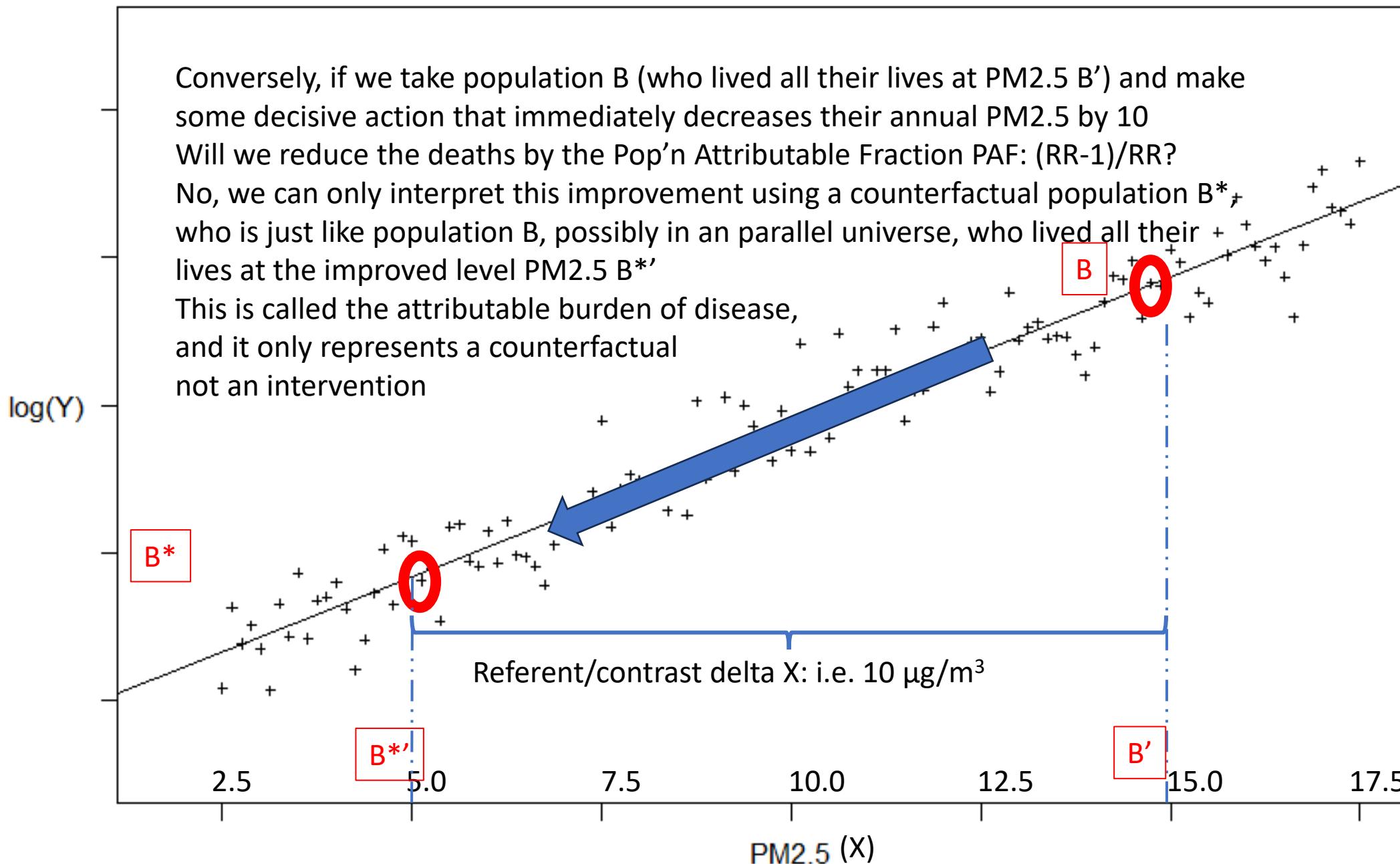
No, that is not the correct interpretation.

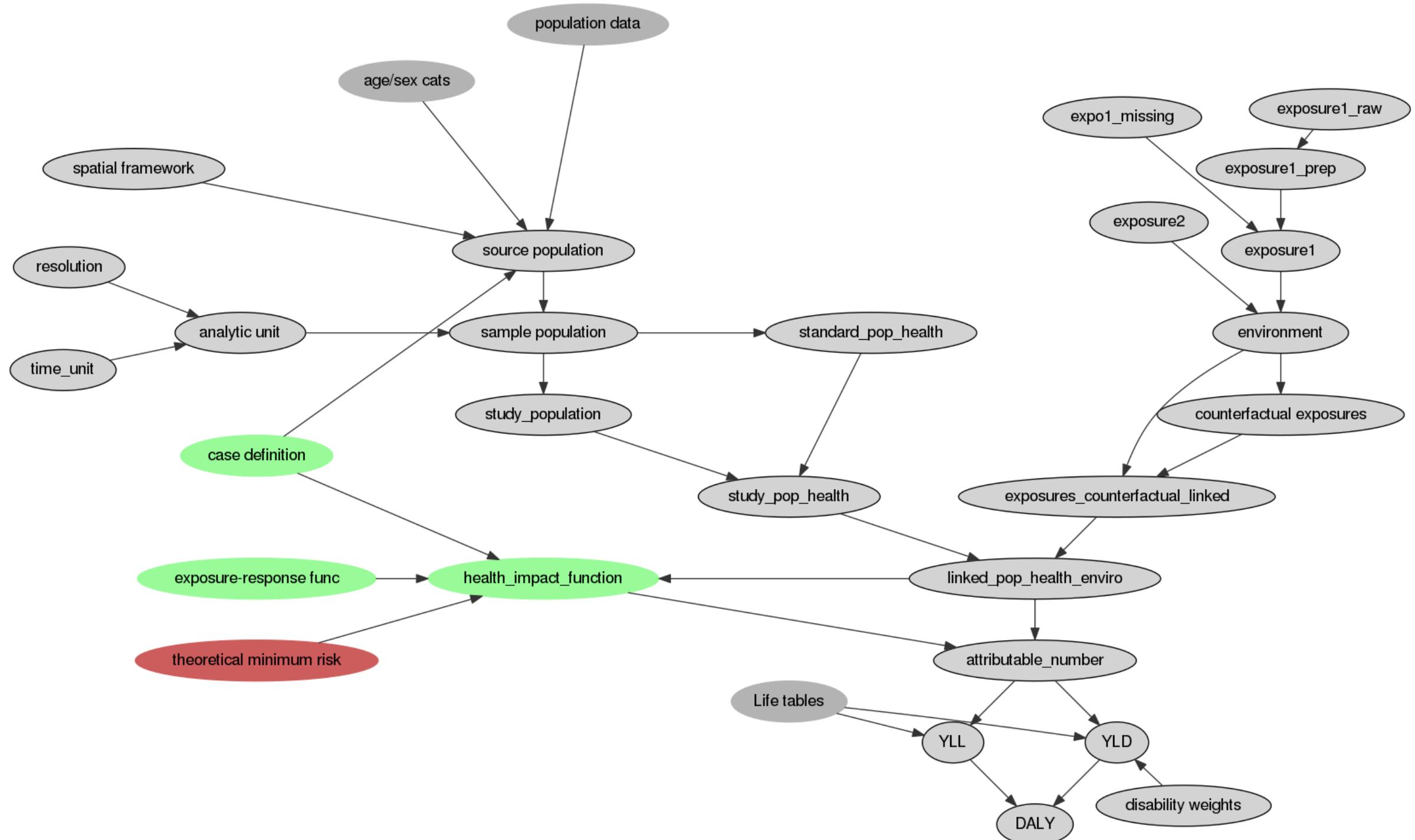
We have derived the RR from a comparison between populations who have all experienced cumulative life-time exposures to get to the current health status.

The comparison is between population A (who lived all their lives at PM2.5 A') with population B (who lived all their lives at PM2.5 B')

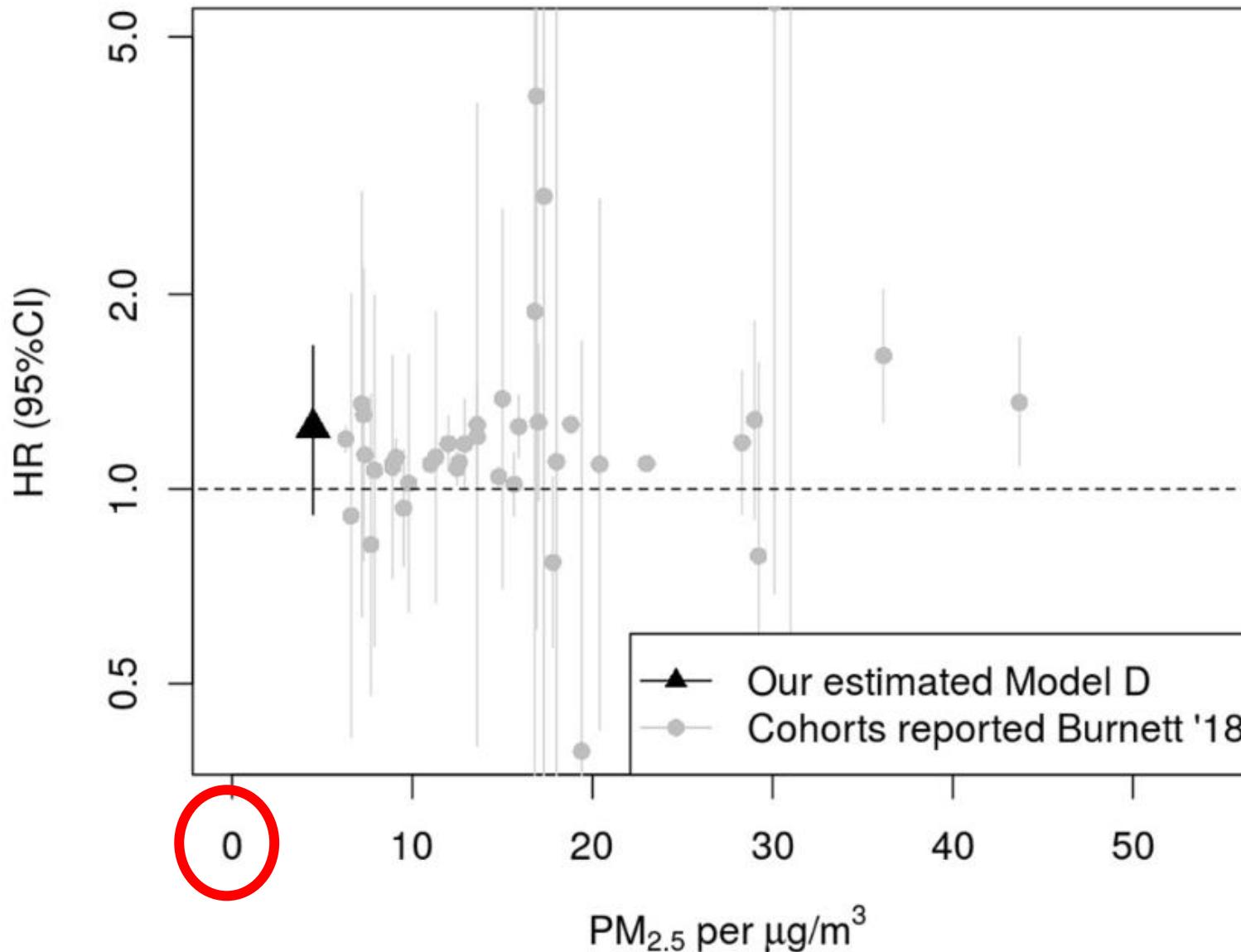


$$\text{Regression model } Y = \beta_0 + \beta_1 X + \beta_2 Z + [...]$$

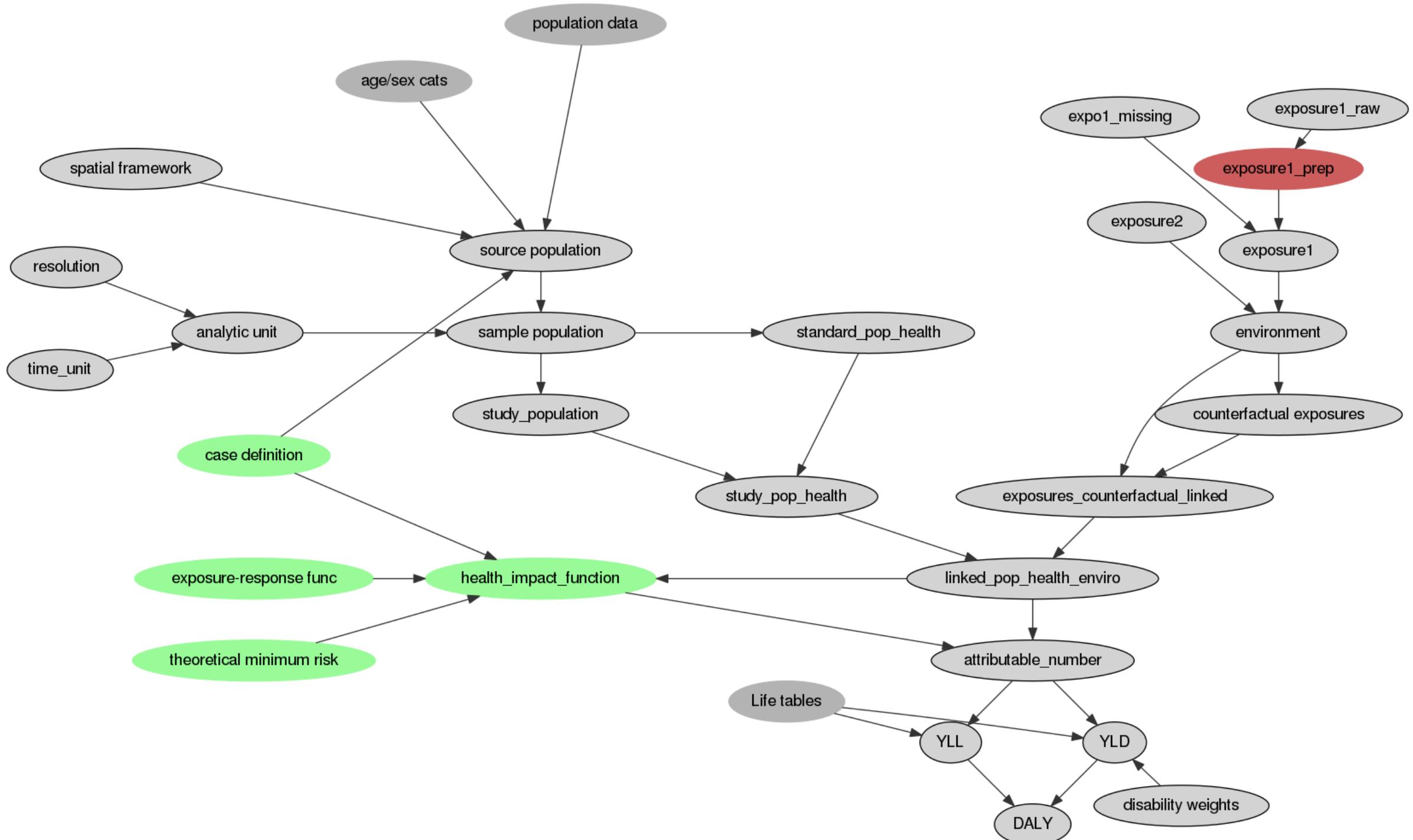




# Theoretical minimum risk level (TMREL)?

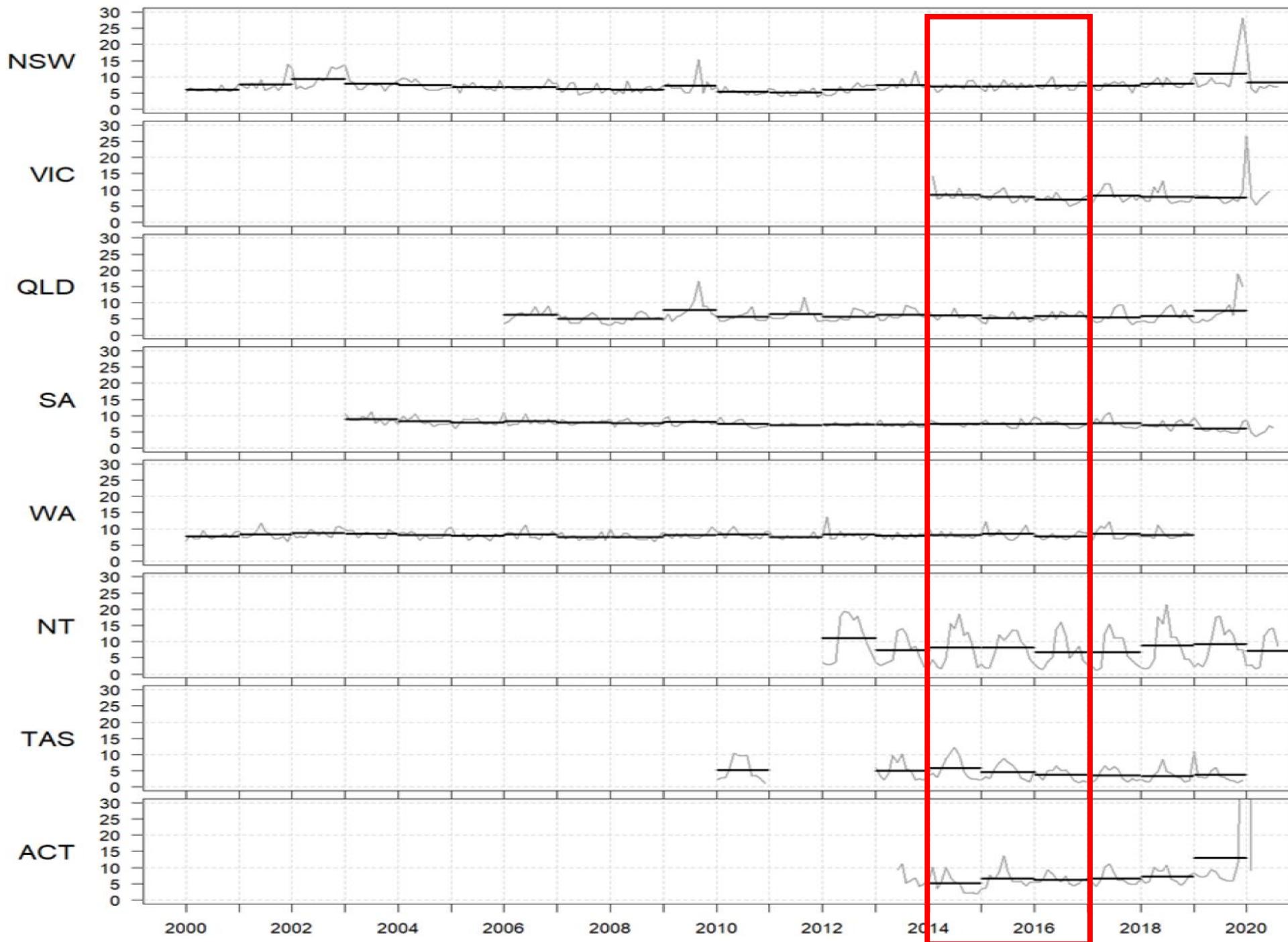


Hanigan et al 2019.  
Environment  
International.  
Study population:  
45andUp in Sydney  
(85K)



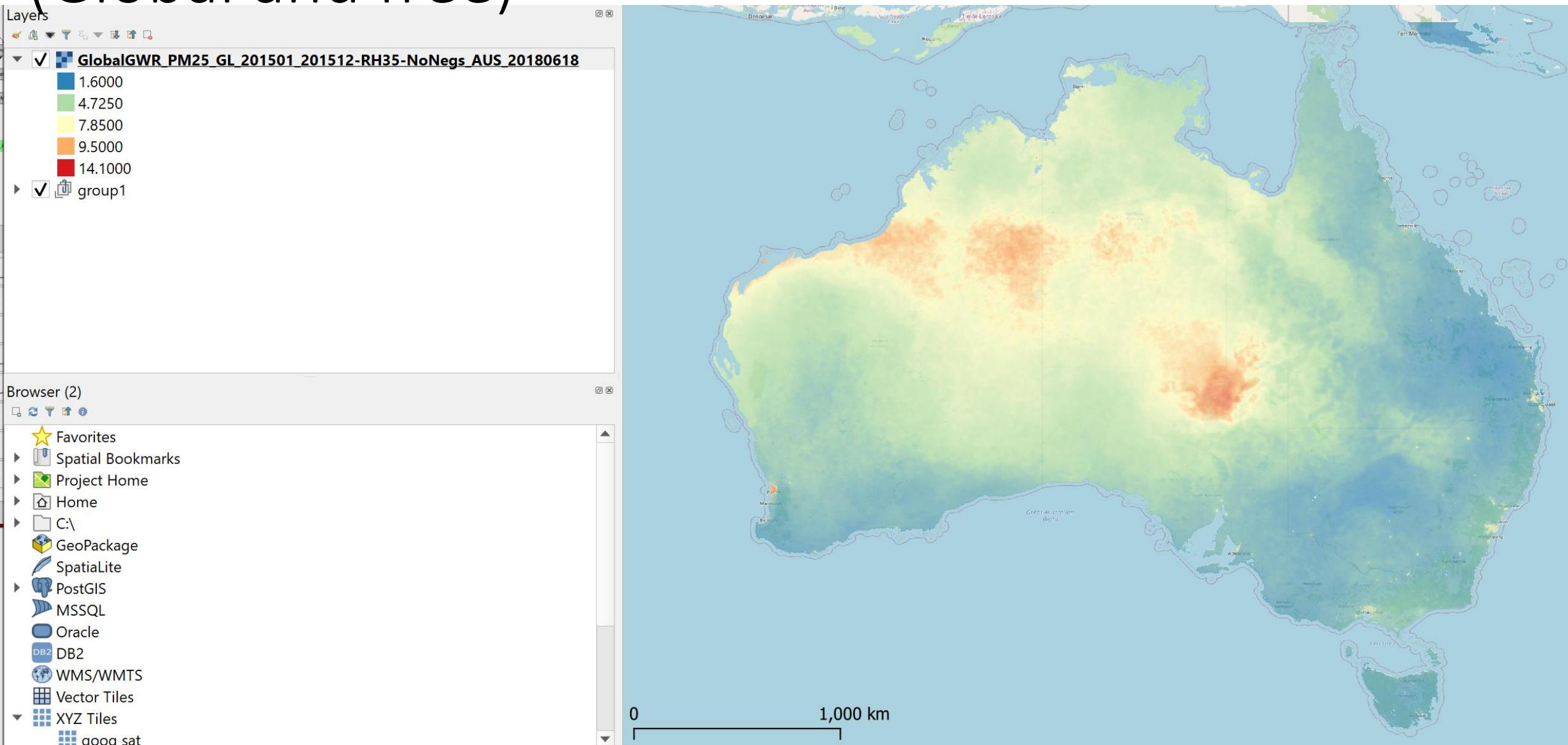
PM2.5  
population  
weighted  
annual  
averages

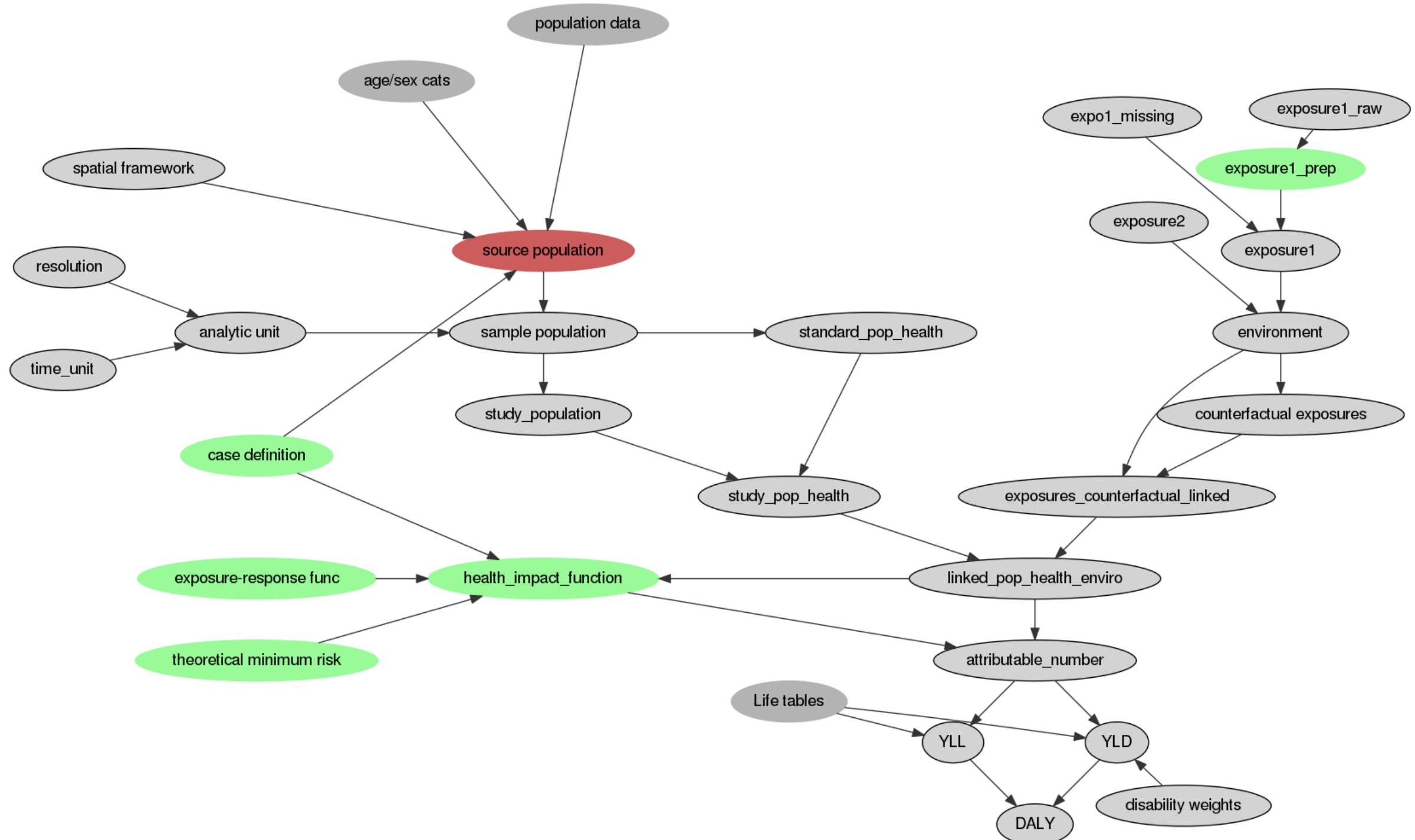
Study  
period?  
2014-2016

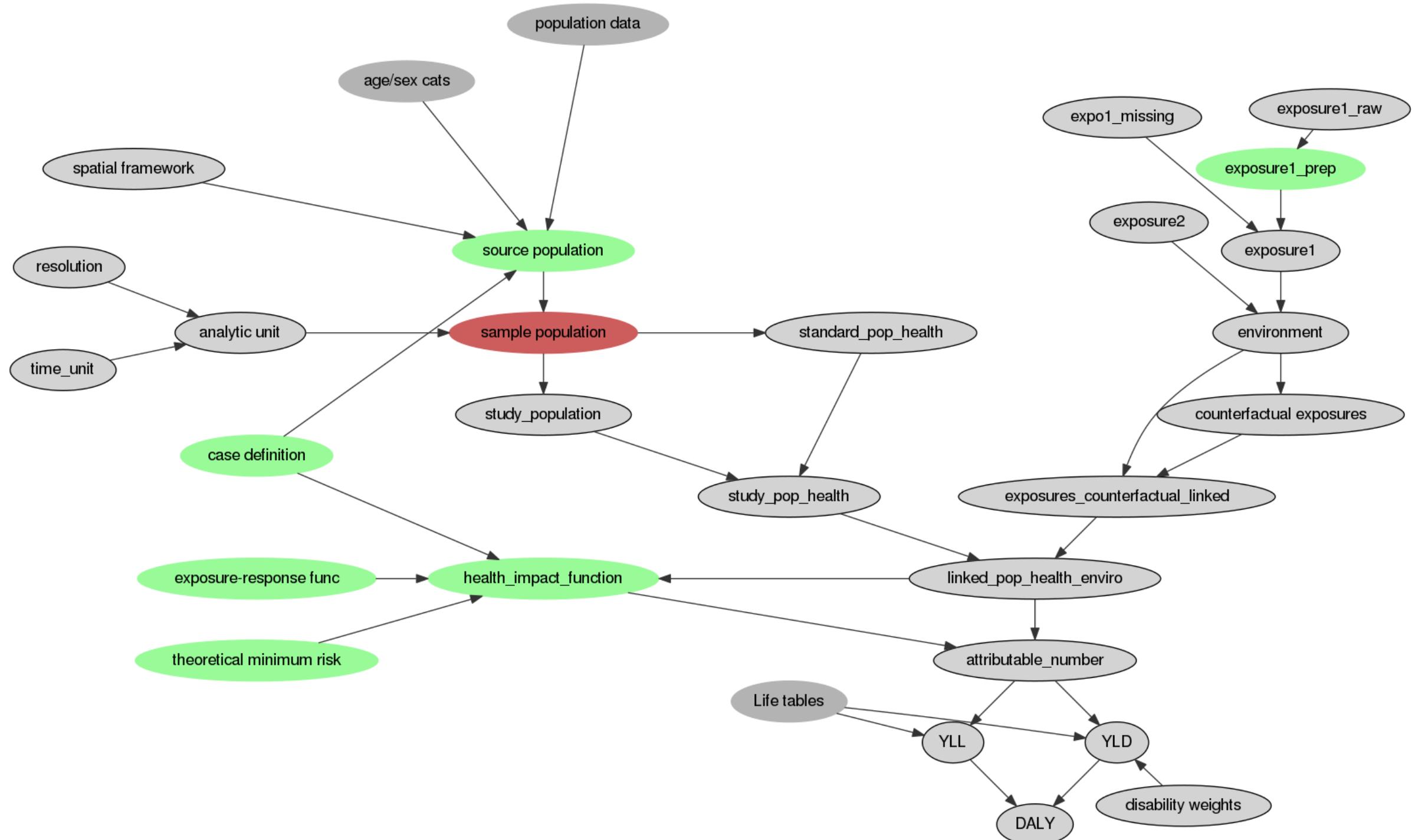


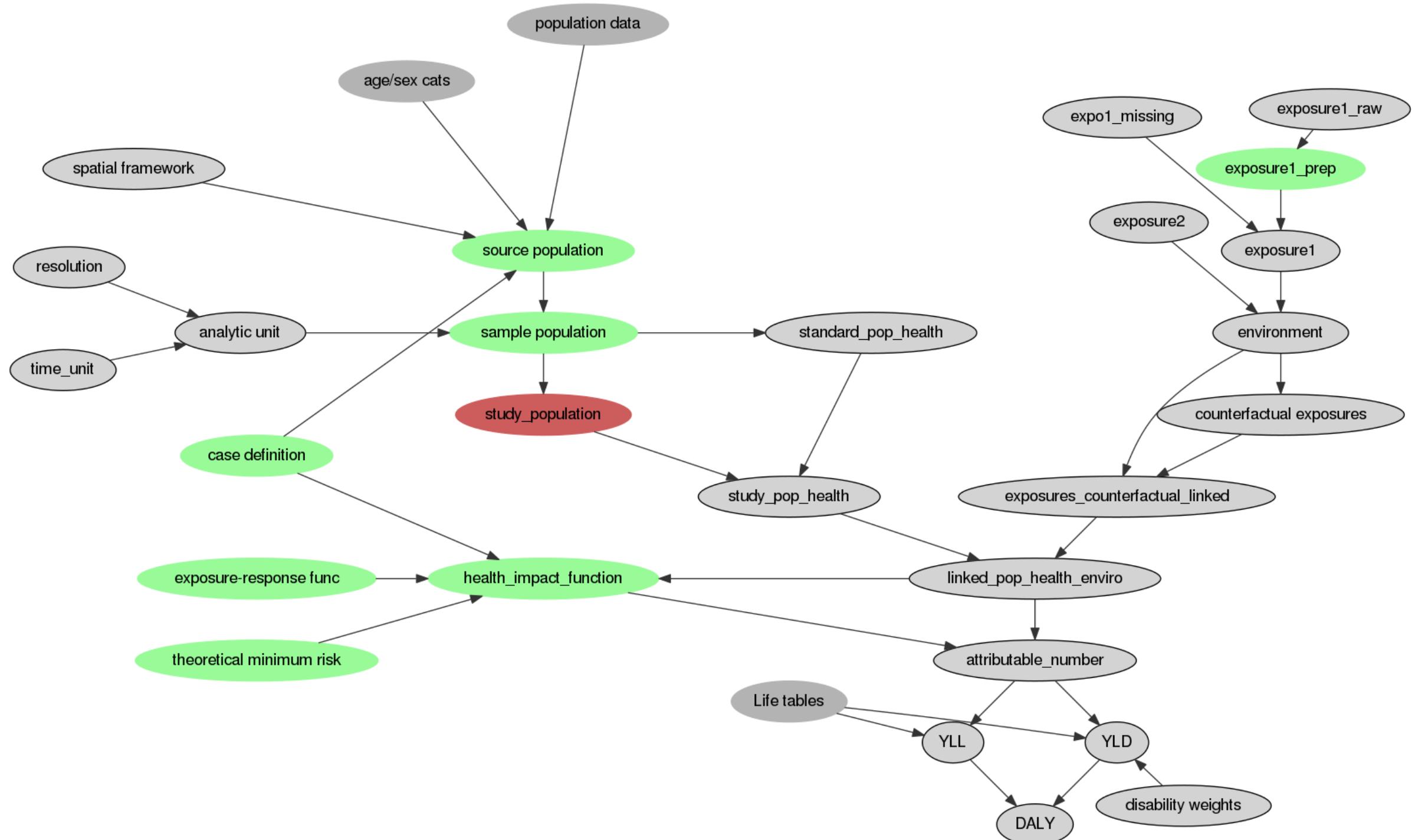
National Air  
Pollution  
Monitoring  
Database,  
derived from  
regulatory  
monitor data  
from NSW DPIE,  
Vic EPA, Qld  
DES, SA EPA,  
WA DEWR, Tas  
EPA, NT EPA,  
and ACT Health.  
Downloaded  
from the Centre  
for Air  
pollution,  
energy and  
health Research  
DOI  
10.17605/OSF.I  
O/JXD98

# Example of annual PM2.5 2015 from van Donkelaar (Global and free)

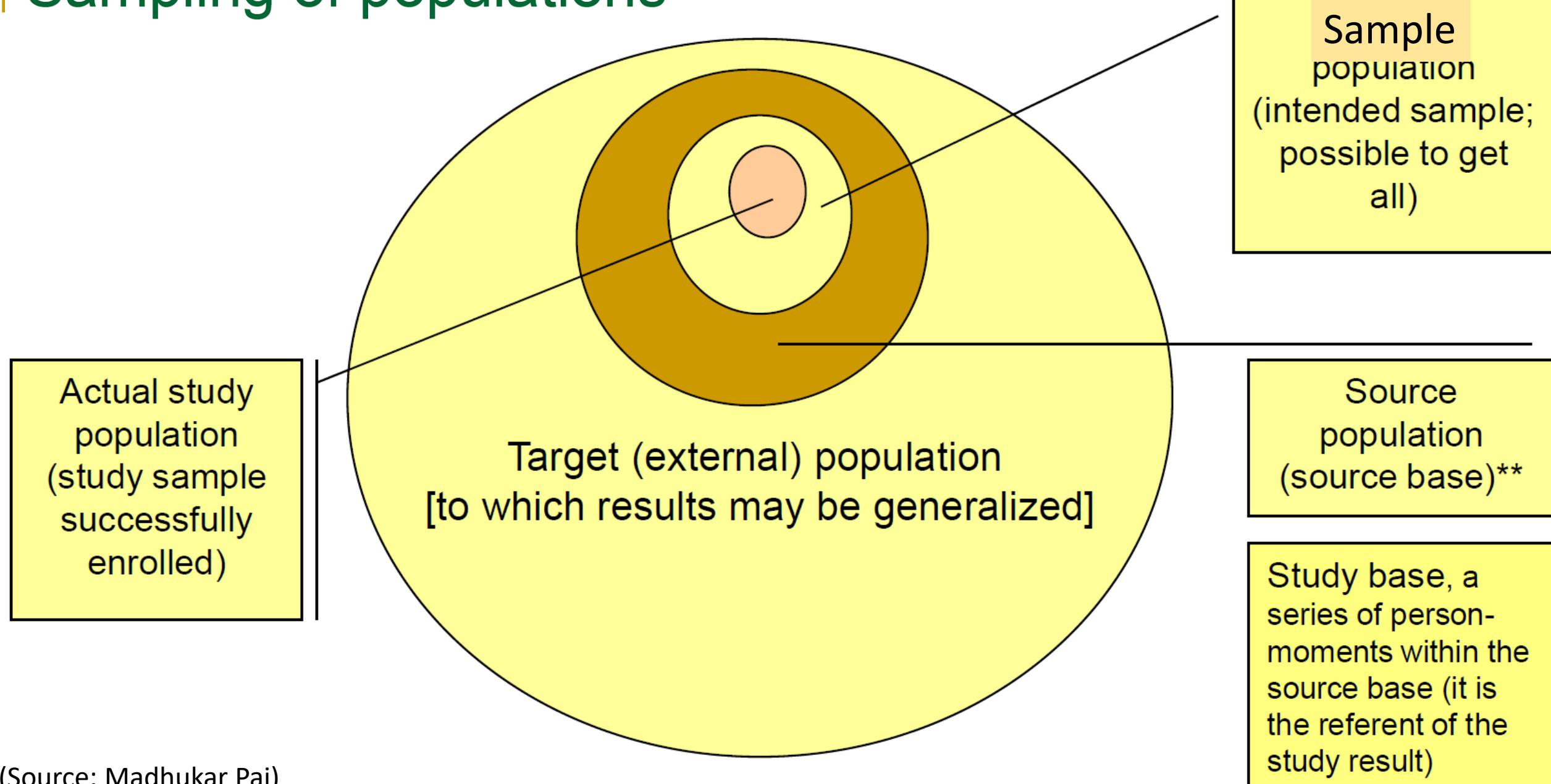






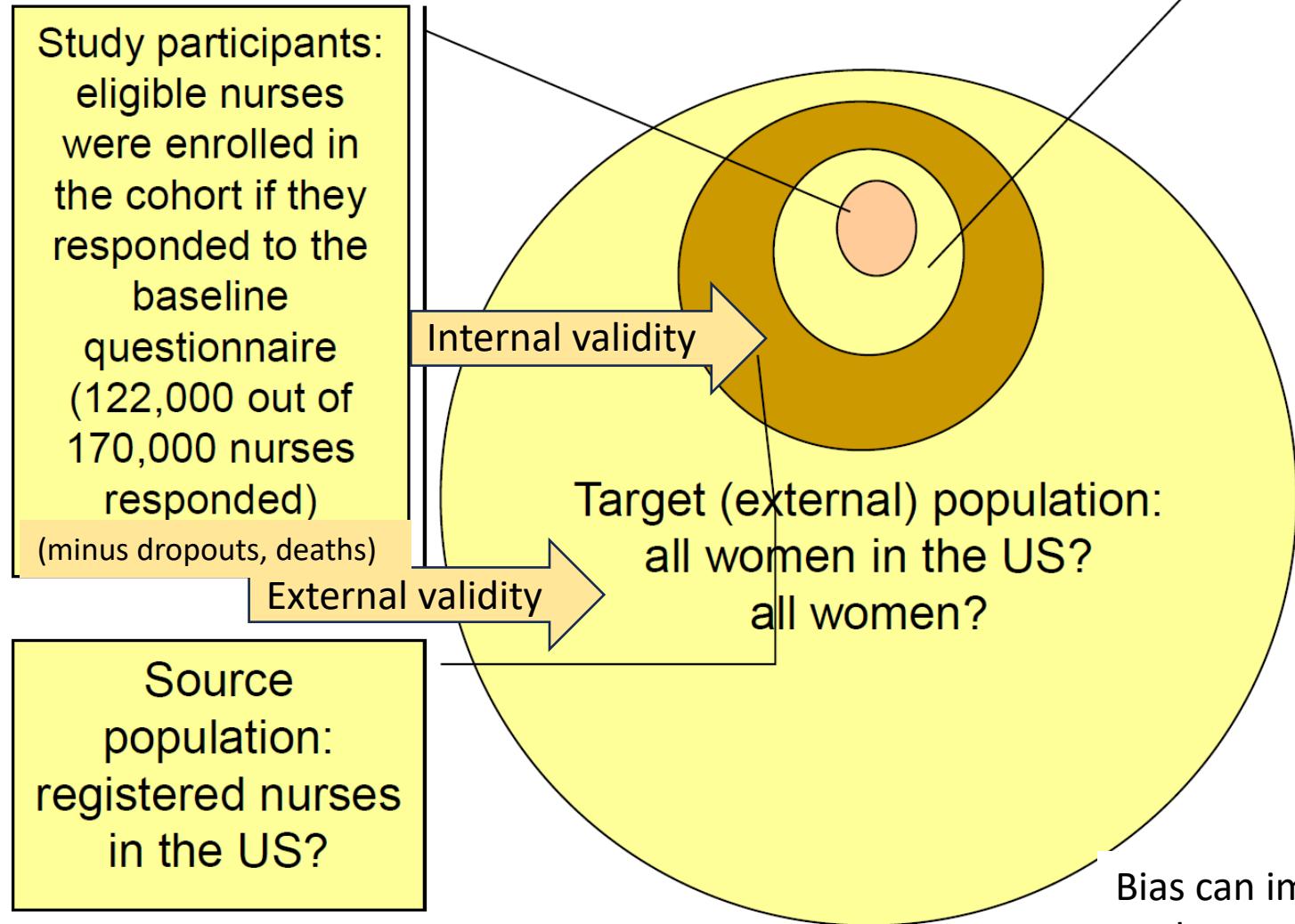


# Sampling of populations



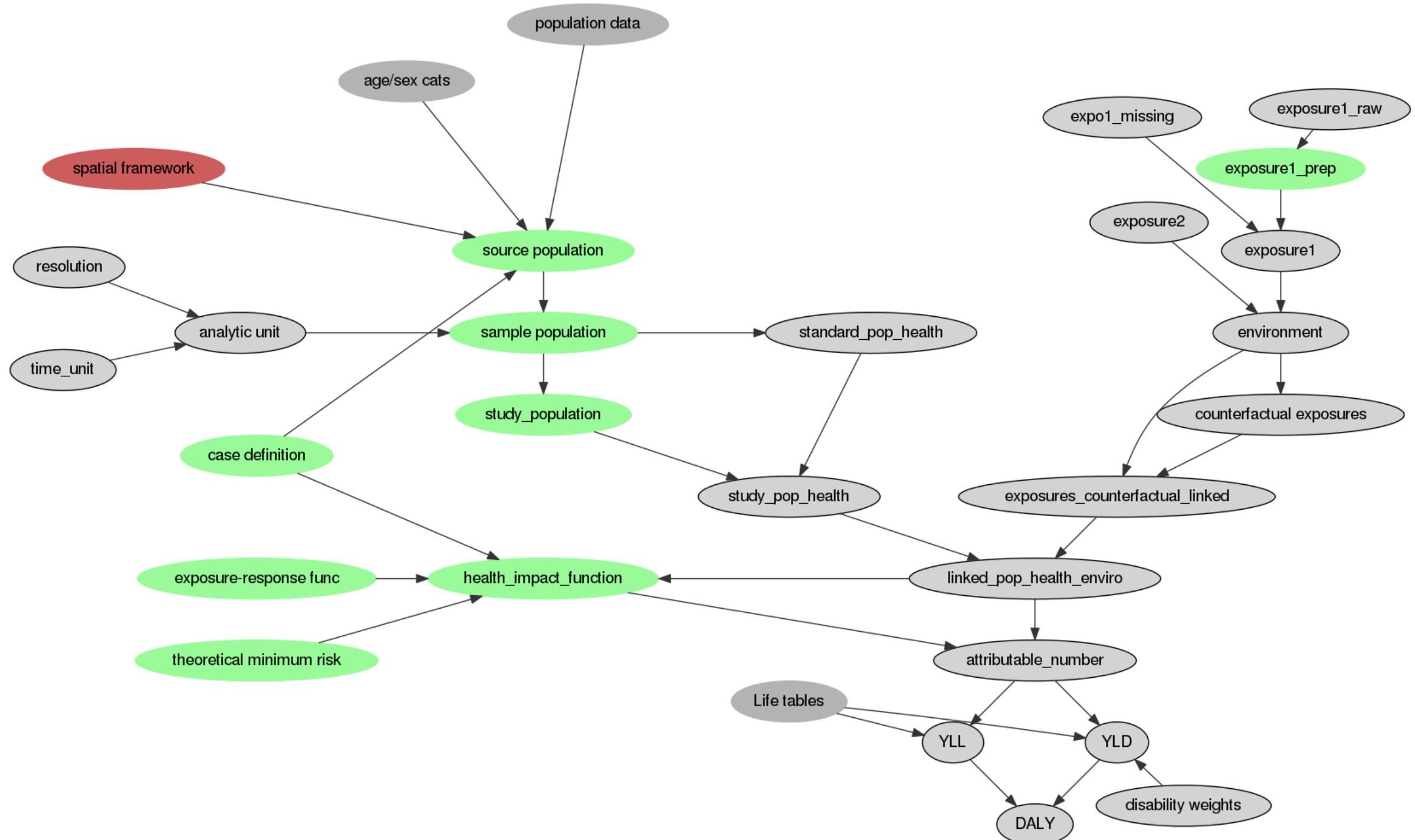
# Sampling of populations

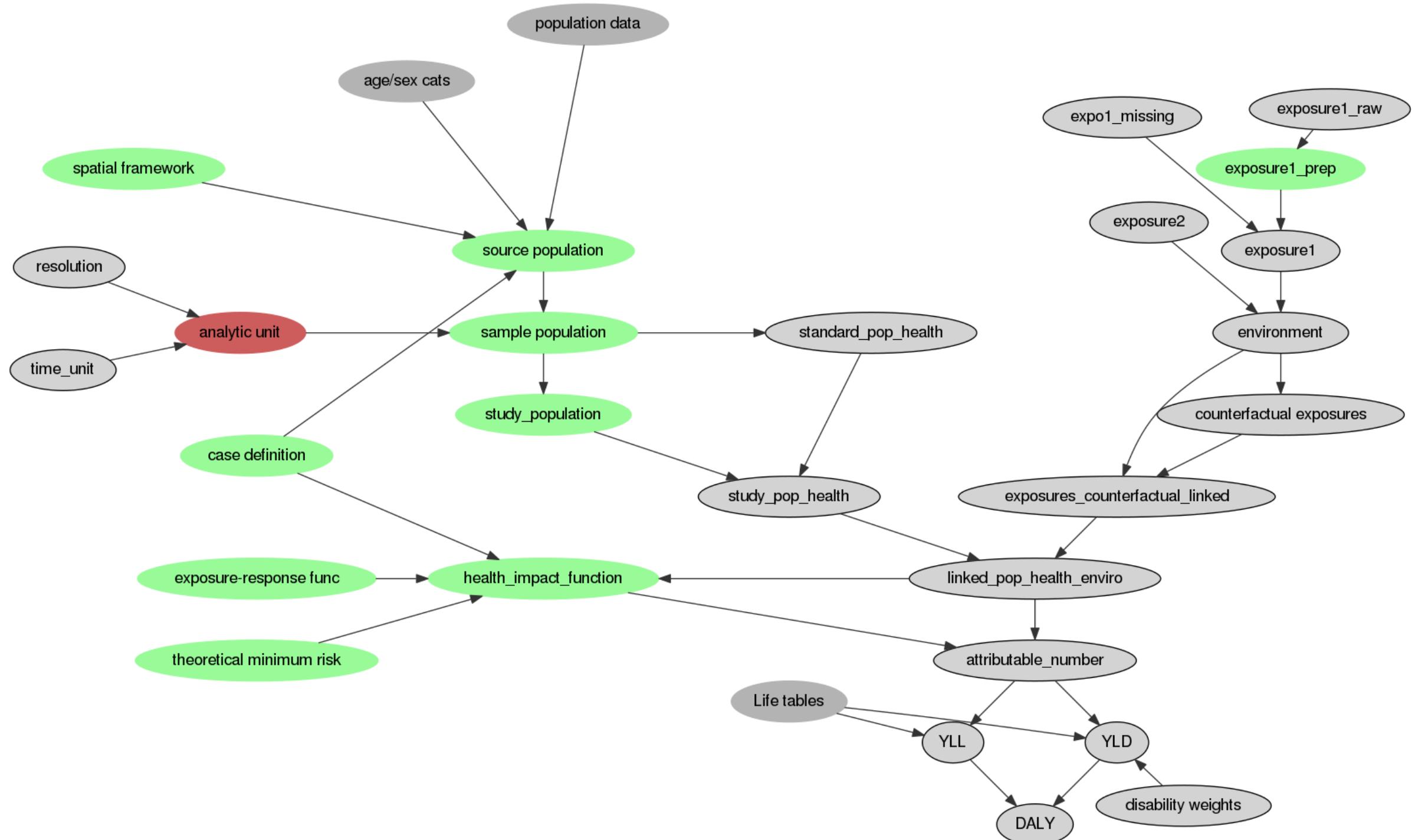
Example: Nurses Health Study

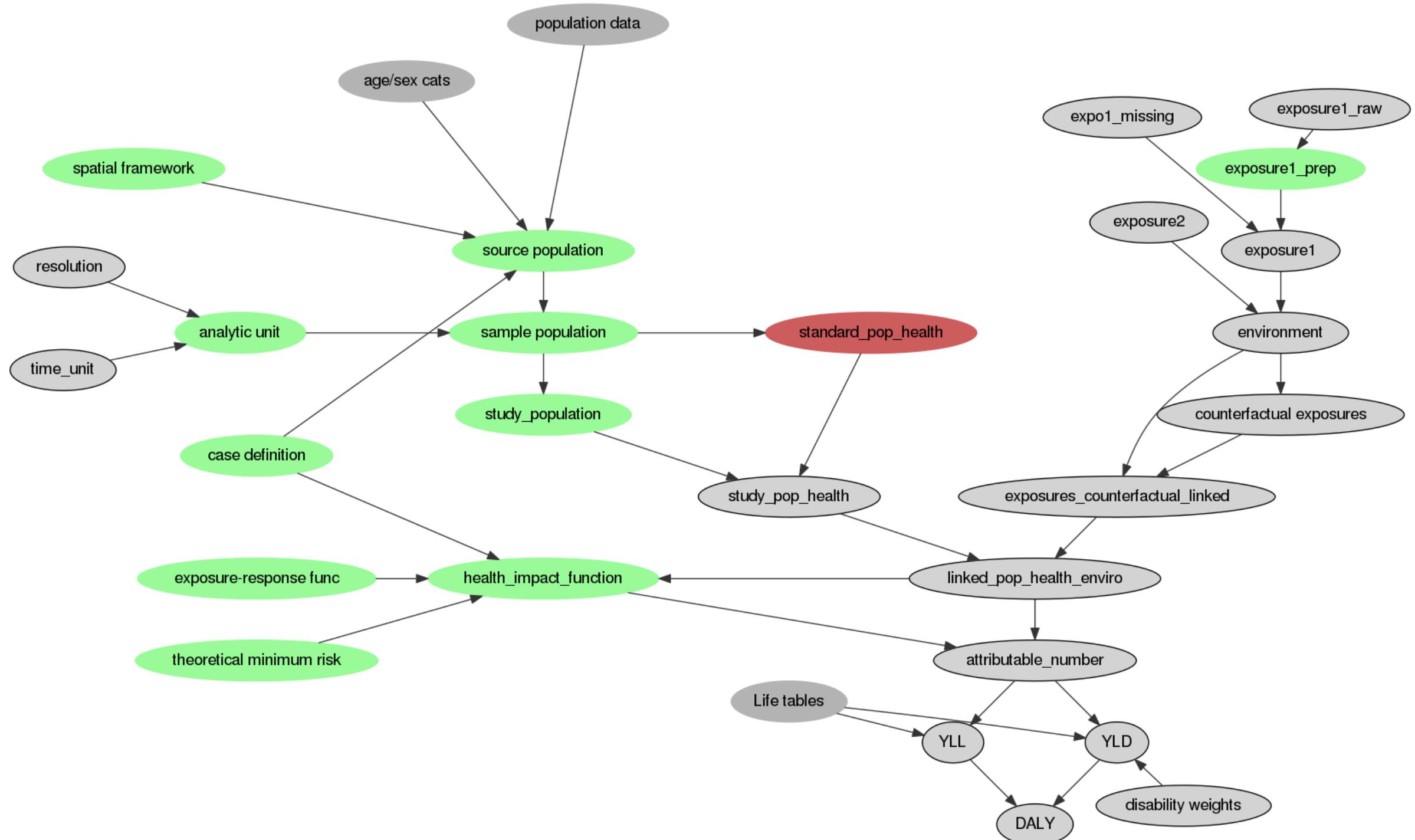


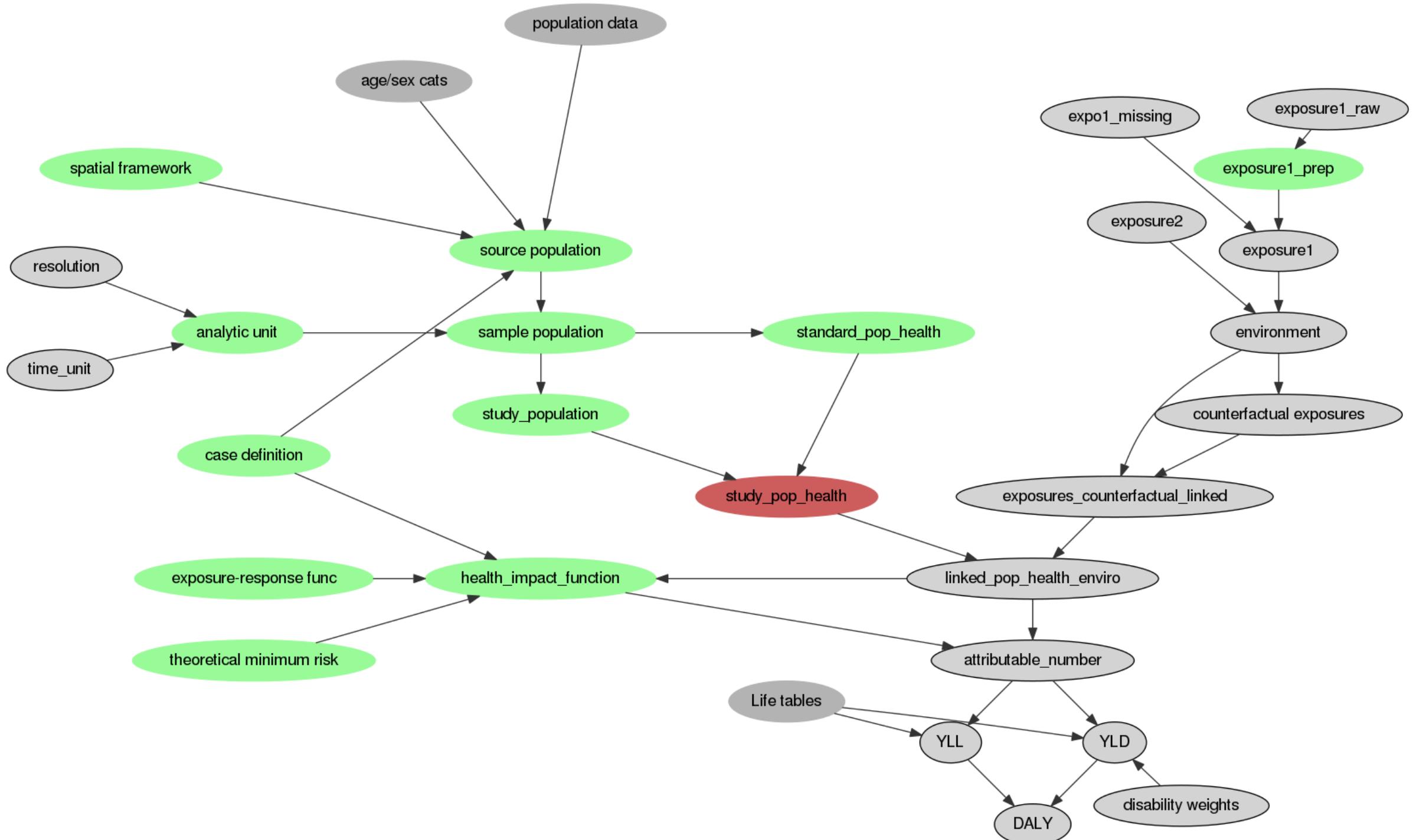
Sample population:  
Married,  
registered  
nurses who  
were aged 30 to  
55 in 1976, who  
lived in the 11  
most populous  
states and  
whose nursing  
boards agreed  
to supply the  
study with their  
members'  
names and  
addresses.

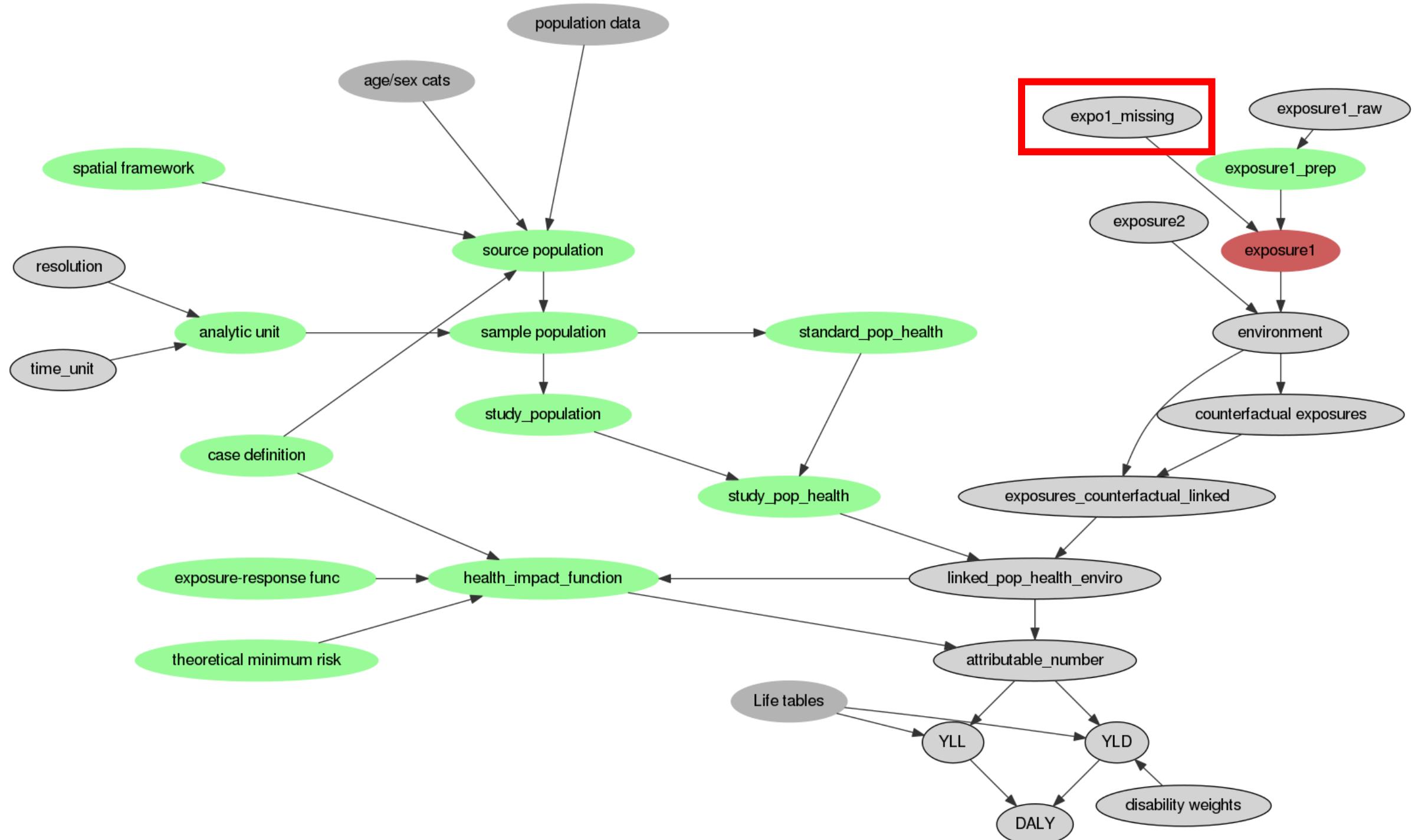
(Source: Madhukar Pai)

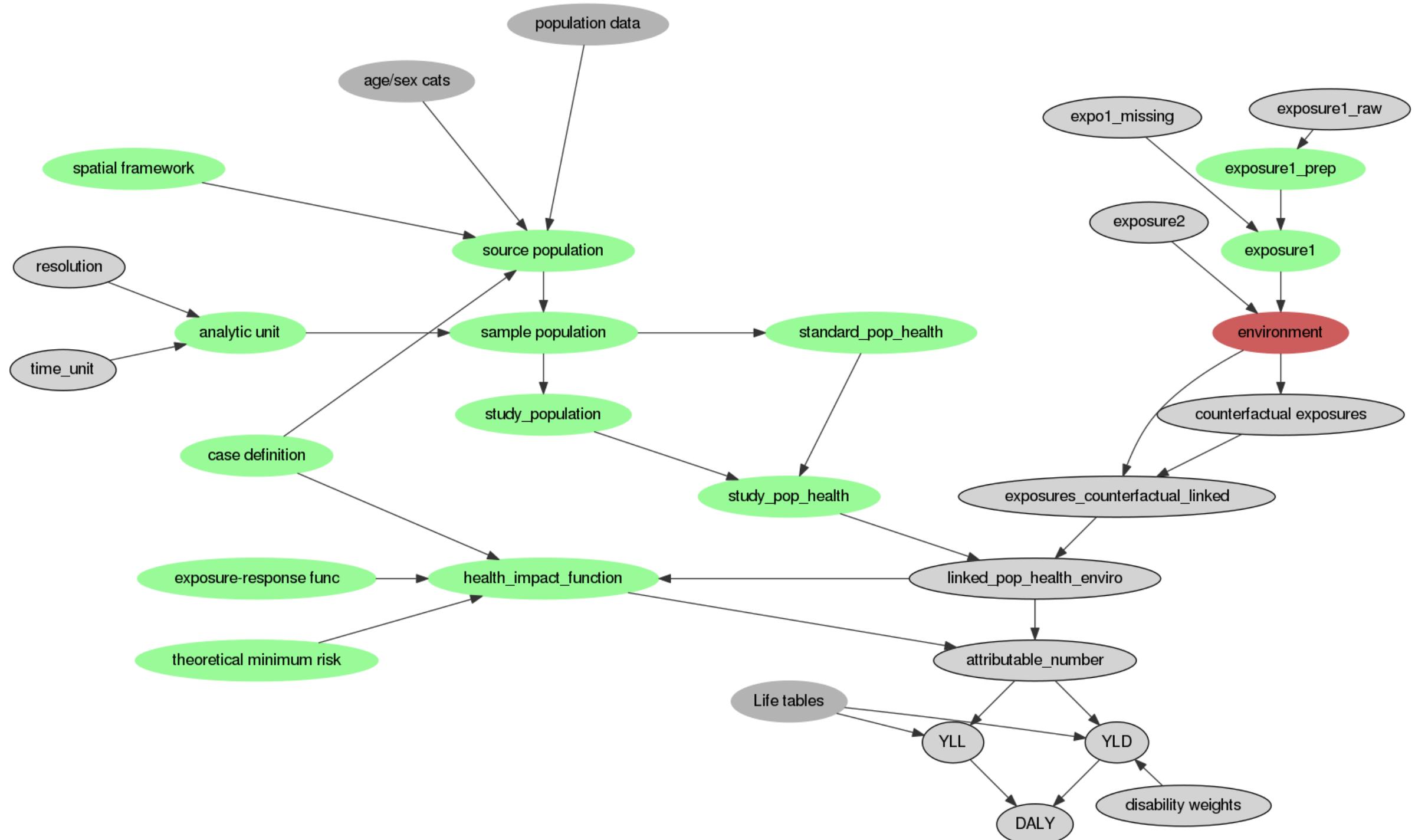


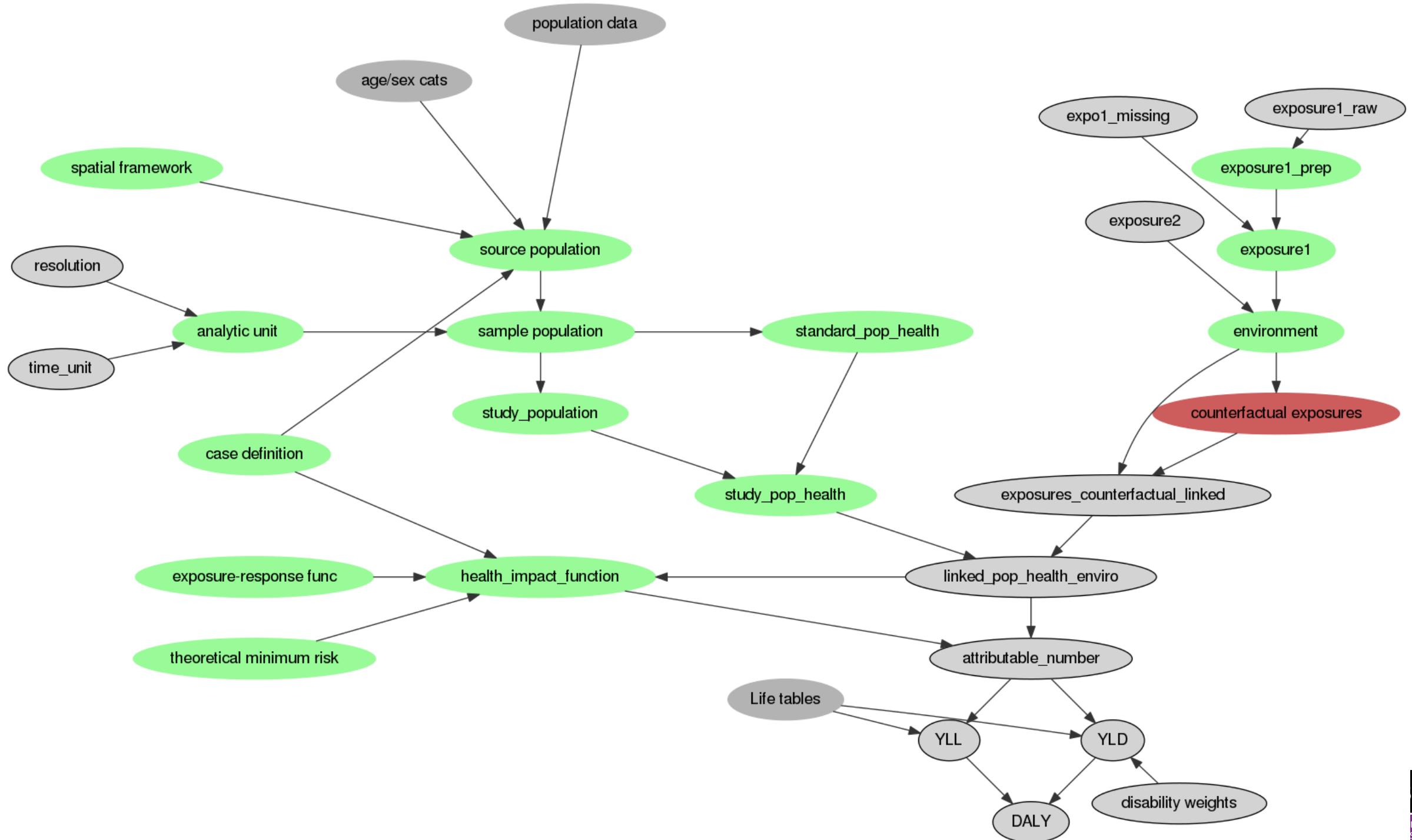












# The counterfactual exposure scenario

- “What If”?
  - Could PM2.5 in all the areas across a region be reduced to the minimum observed anywhere in the region?
  - Assuming the minimum observed PM2.5 is achievable everywhere then that could be the counterfactual



Curtin University

WHO COLLABORATING CENTRE FOR CLIMATE  
CHANGE AND HEALTH IMPACT ASSESSMENT

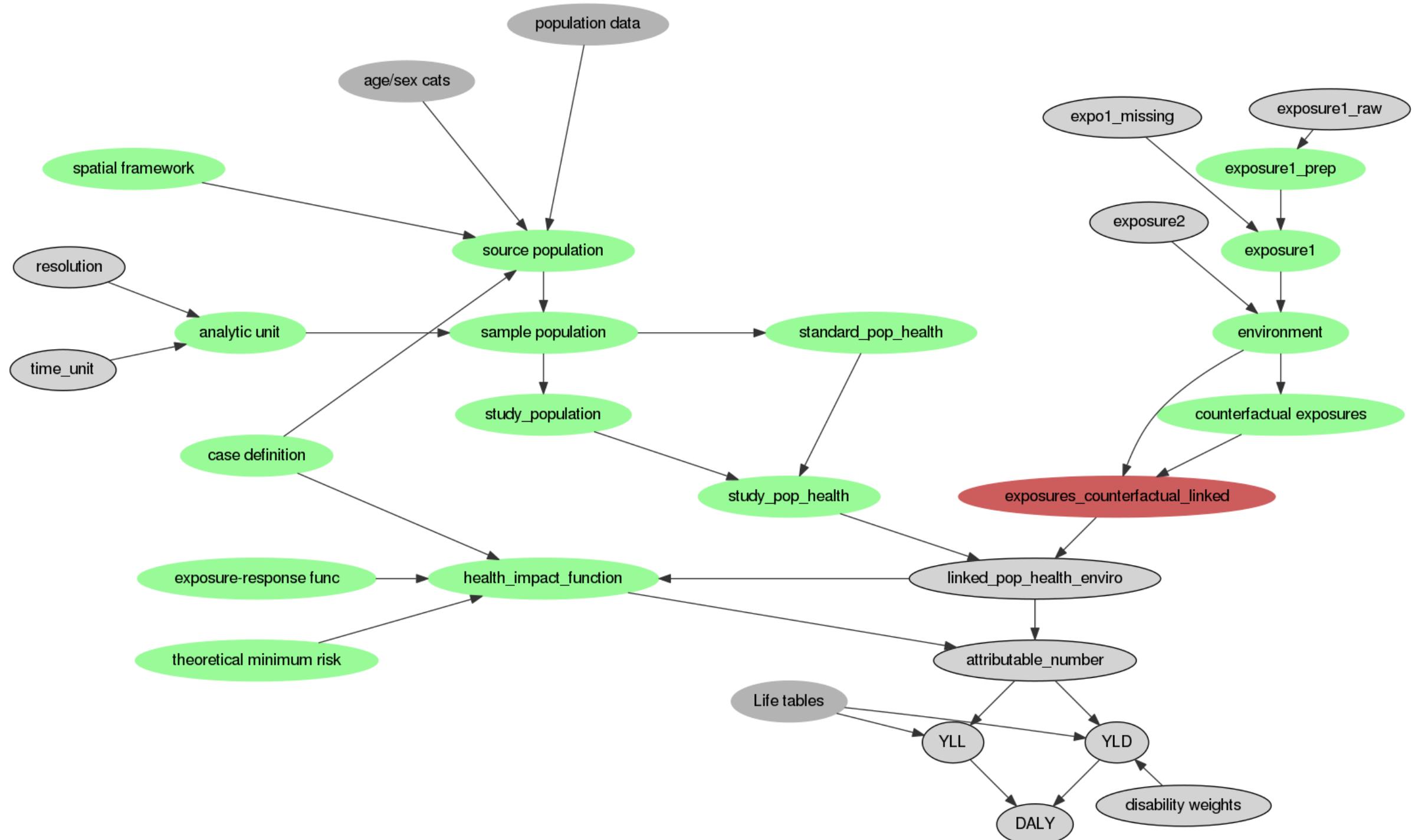
# Terminology Issue!

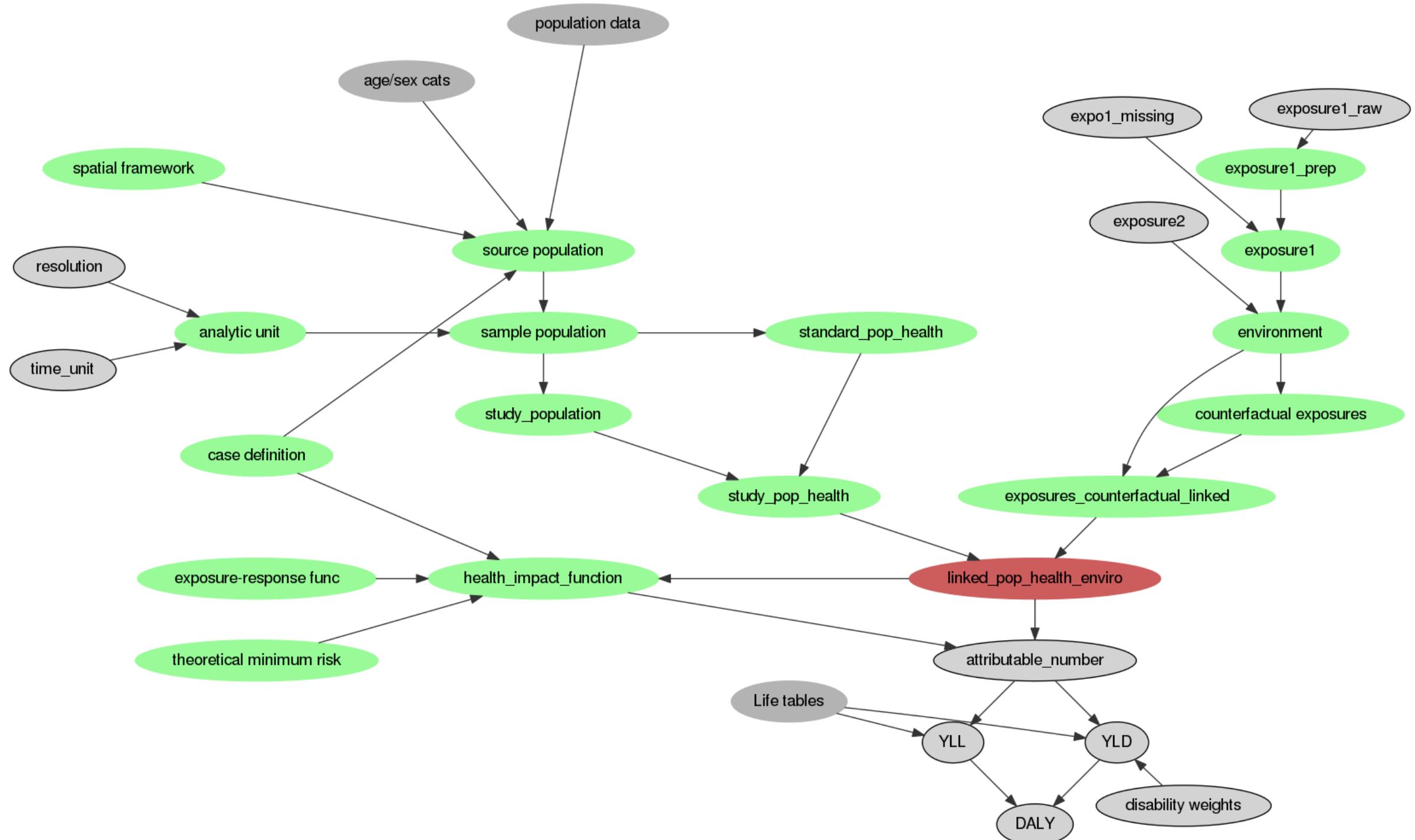
- Burden:
  - What is the current attributable number in population X
  - Due to the accumulated exposure over life course until now
- Impact:
  - What would be in a counterfactual (future) population X'
  - Due to the changed exposure over life course from now on
  - Accounting for demographic change (aging)
  - And (potentially) population ‘churn’ or ‘turnover’

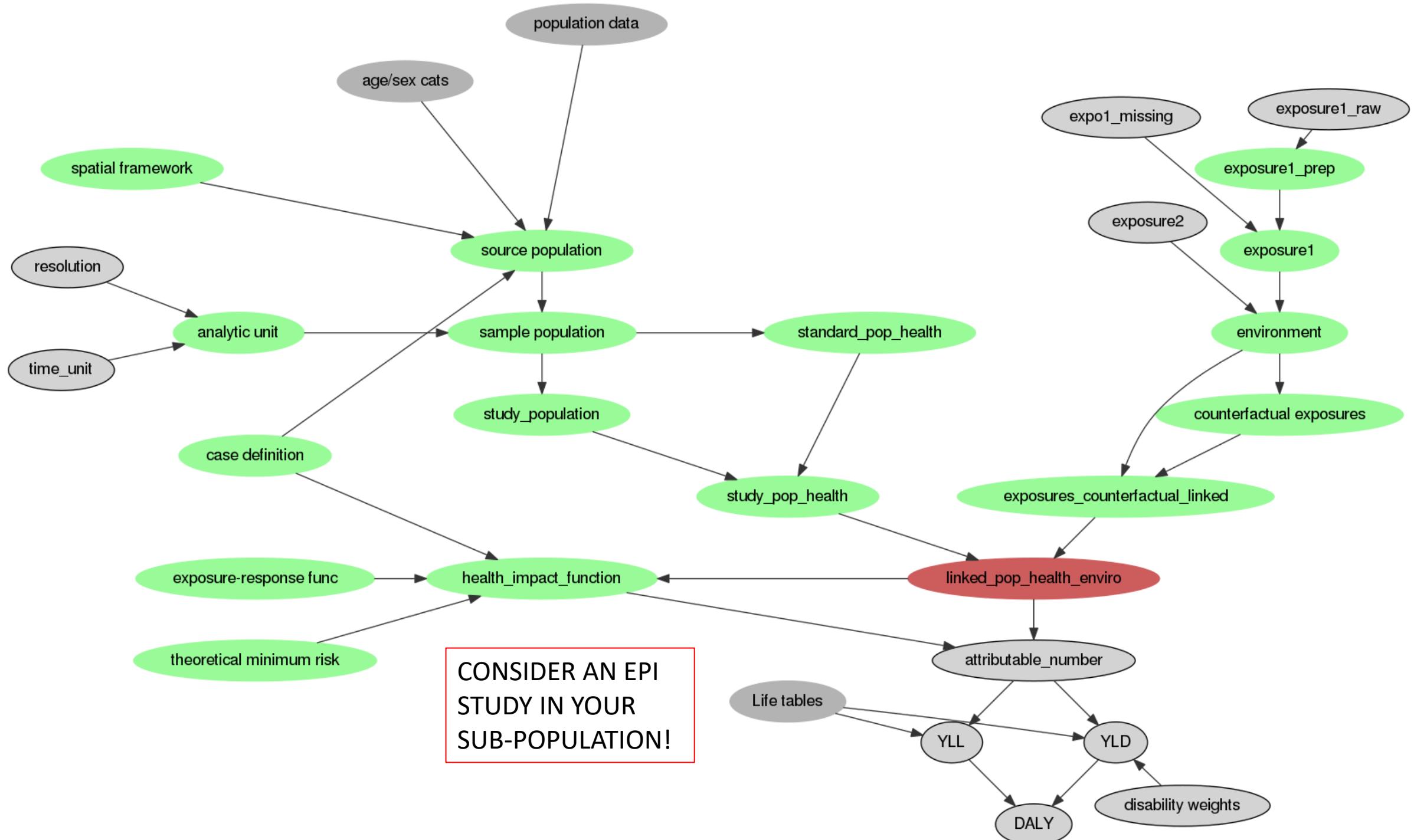


Curtin University

WHO COLLABORATING CENTRE FOR CLIMATE  
CHANGE AND HEALTH IMPACT ASSESSMENT







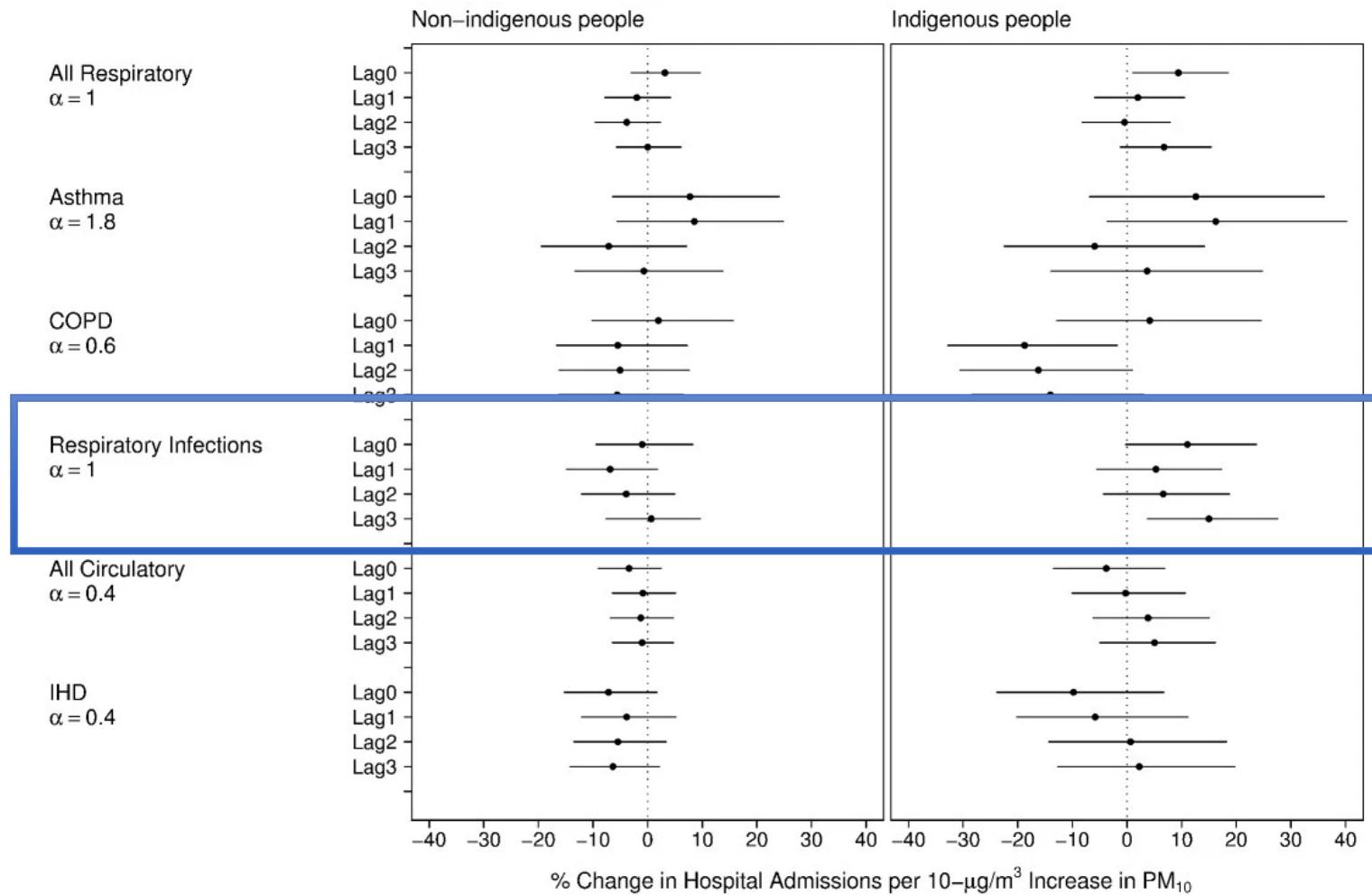
# The principles of HIA (from enHealth 2017)

1. Democracy
2. Equity
3. Sustainable Development
4. Ethical use of evidence
5. Comprehensive approach to health

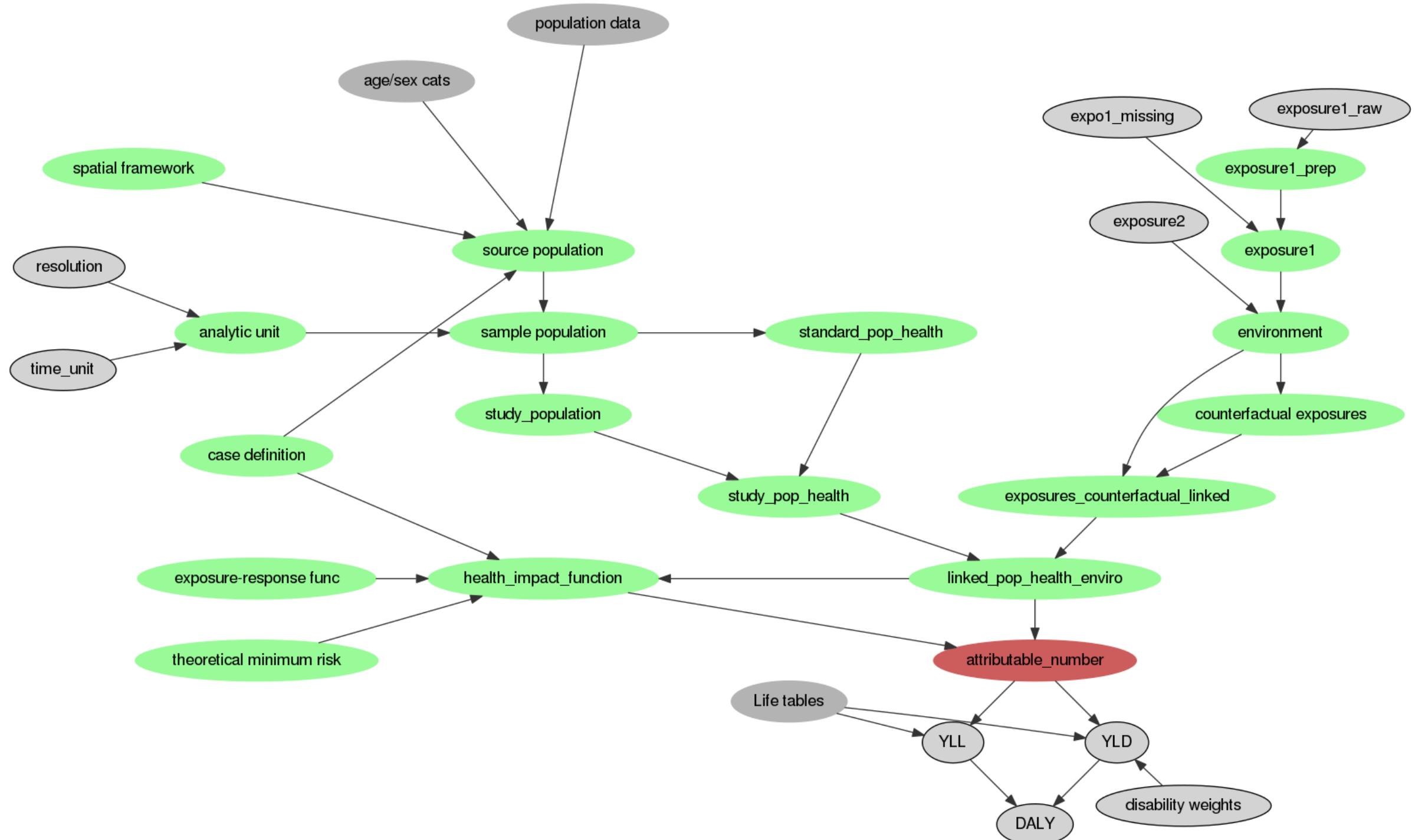
# Differential effect estimates

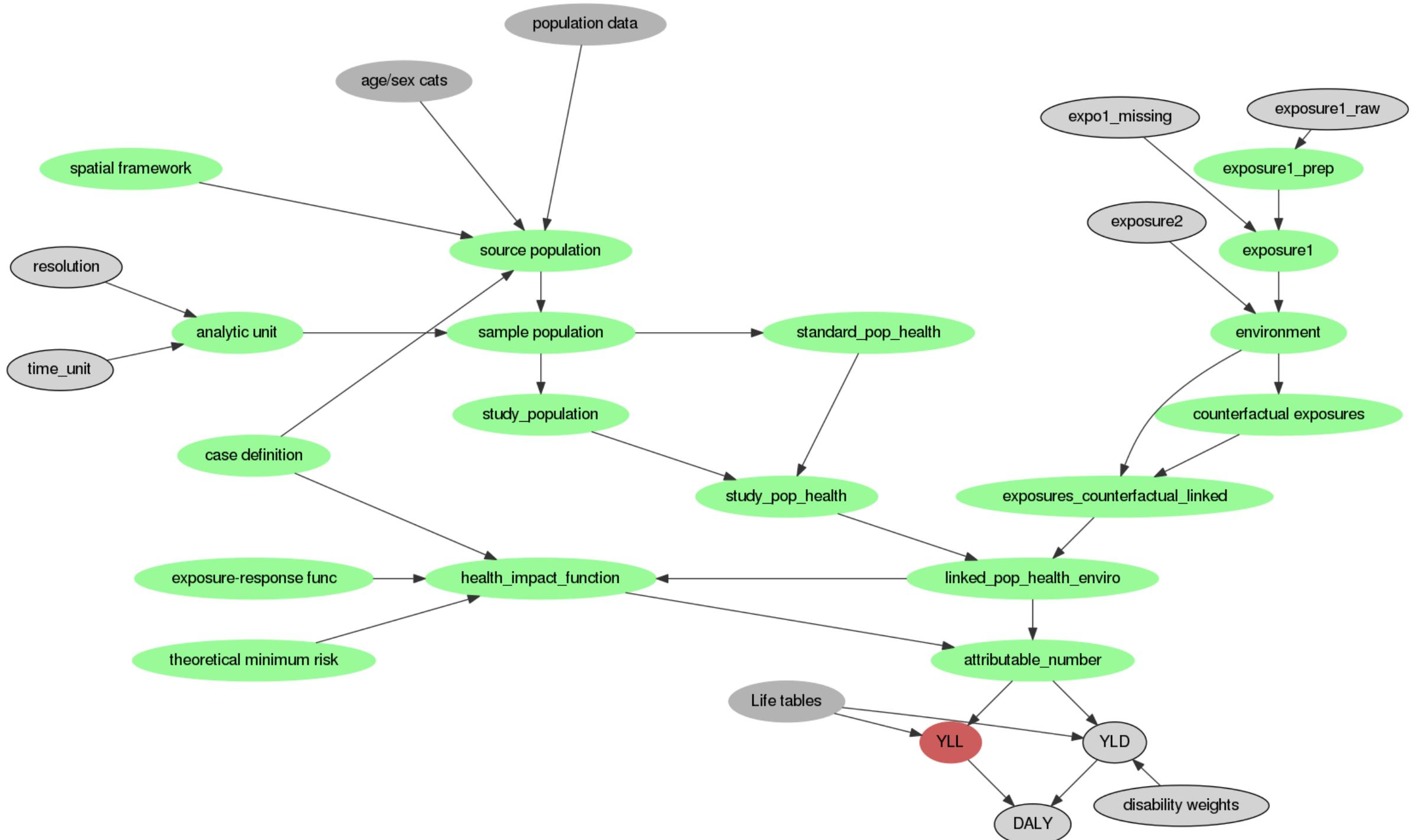
- Exposure-response estimates often assumed to be the same for all areas/subpopulations
  - Using effect-estimates specific to subpopulations would likely show an exacerbation in socio-economically disadvantaged groups
  - Local and international studies have found differential effects between subgroups
- Krewski, D. (2009). *Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality.*
  - Less than high school 1.13 (0.9-1.42) for  $10\mu\text{g}/\text{m}^3 \text{PM}_{2.5}$
  - High school or more 0.97 (0.9-1.05) for  $10\mu\text{g}/\text{m}^3 \text{PM}_{2.5}$

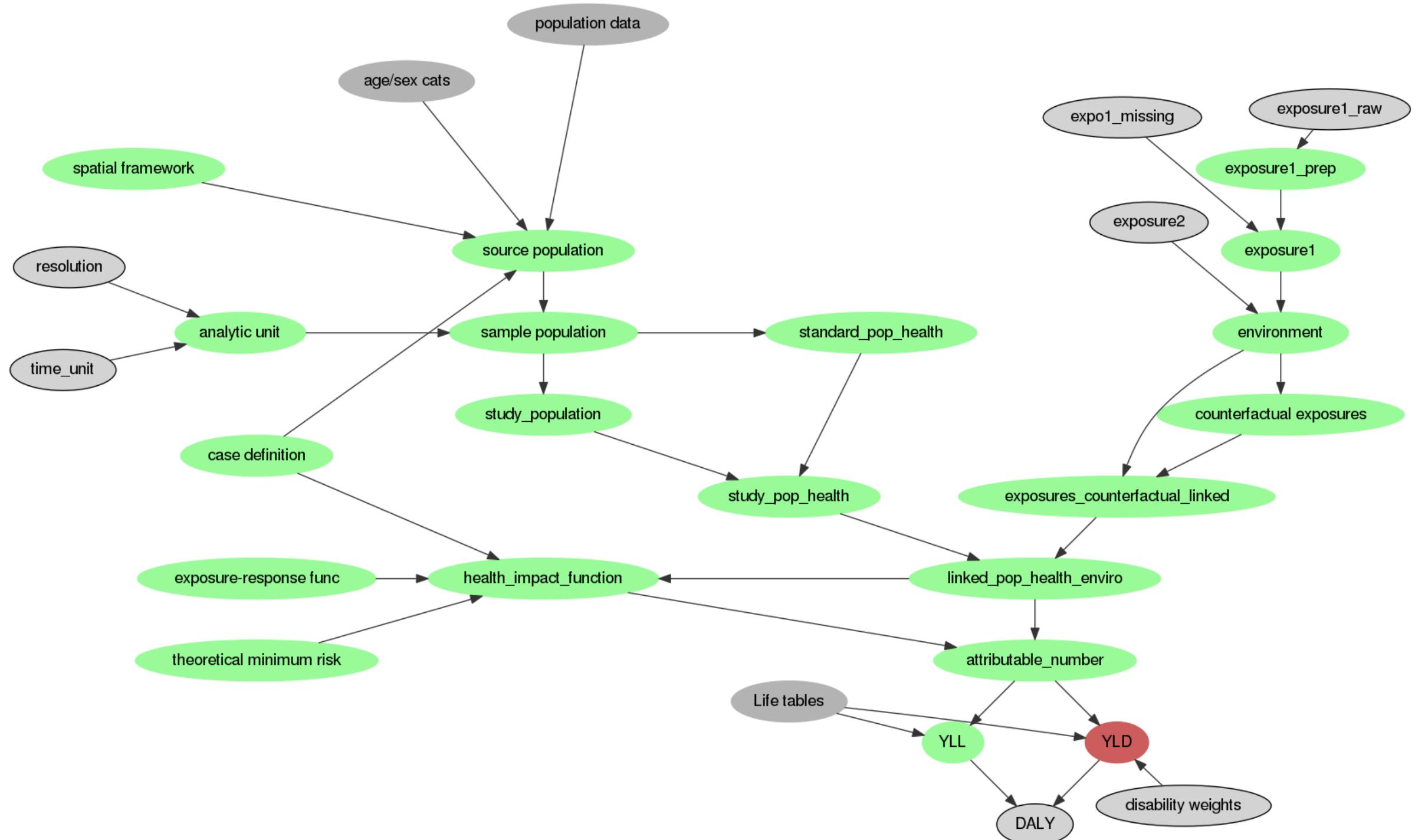
# Disadvantage and differential effect estimates

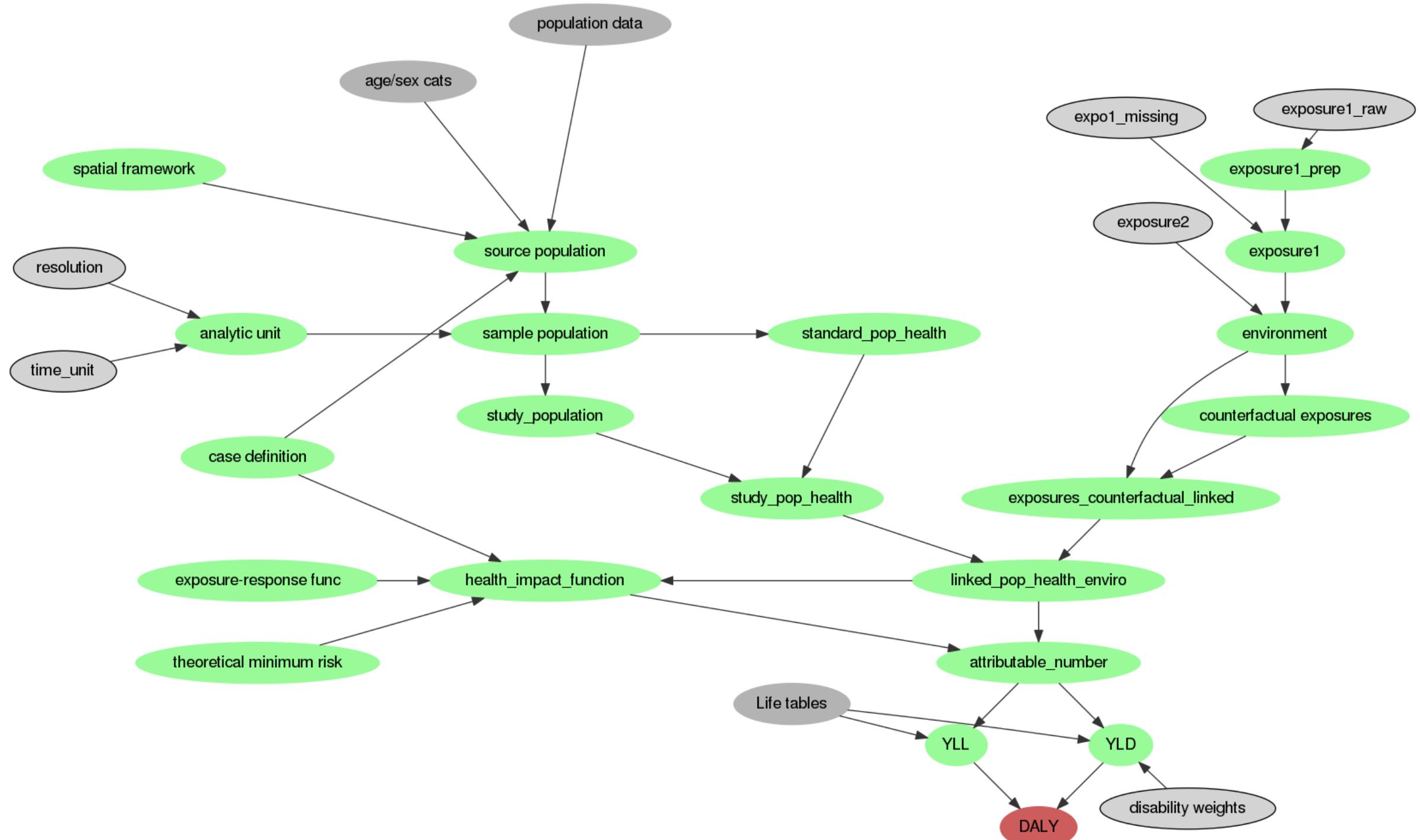


Hanigan, I.C., Johnston, F.H. & Morgan, G.G. Vegetation fire smoke, indigenous status and cardio-respiratory hospital admissions in Darwin, Australia, 1996–2005: a time-series study. *Environ Health* **7**, 42 (2008). <https://doi.org/10.1186/1476-069X-7-42>









# Part two: Impact

- Now we move on to the more desirable question
  - ‘What-If?’
- But we need to be careful about the interpretation
  - Burden is useful for many questions but the limits are clear and caveats required
  - Impact is based on scenarios, assumptions and calculations that move beyond the quantification of the epidemiological relations, into modelling of system dynamics.



Curtin University

WHO COLLABORATING CENTRE FOR CLIMATE  
CHANGE AND HEALTH IMPACT ASSESSMENT

File Edit View Layout Model Options Windows Help

SyntheSim

A  
B  
C  
D  
E  
F  
G  
H  
I  
J  
K  
L  
M  
N  
O  
P  
Q  
R  
S  
T  
U  
V  
W  
X  
Y  
ZLock  
Sketch  
Variable  
Level  
Arrow  
Rate  
Shadow Variable  
IO Object Comment  
Delete  
Equations  
Relative ModeTree  
User Tree  
Loops

Document

Document All

Causes Strip

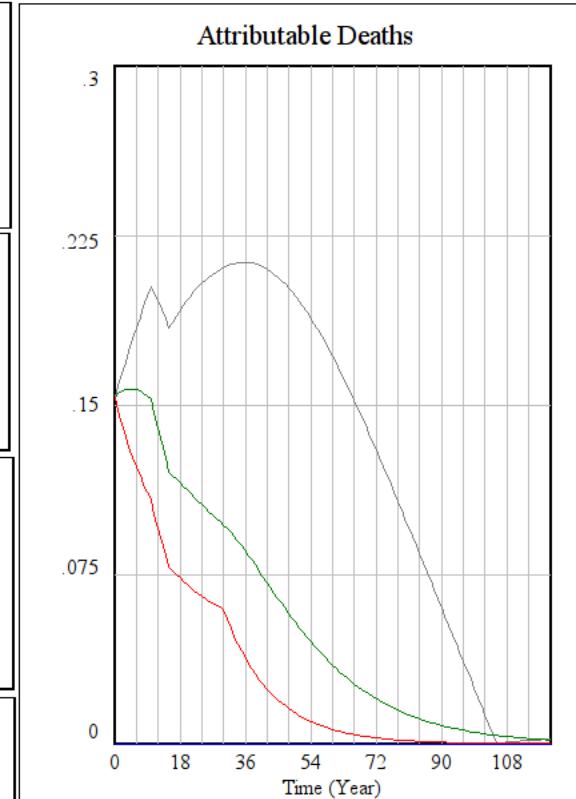
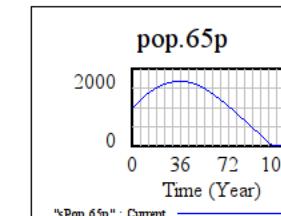
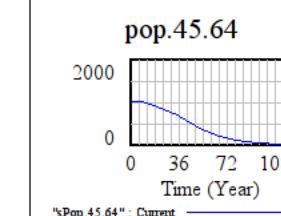
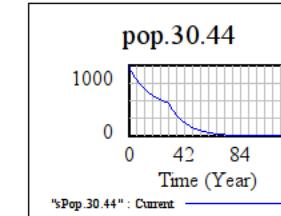
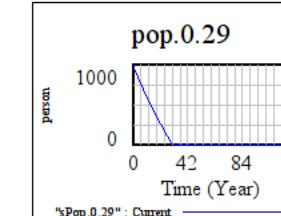
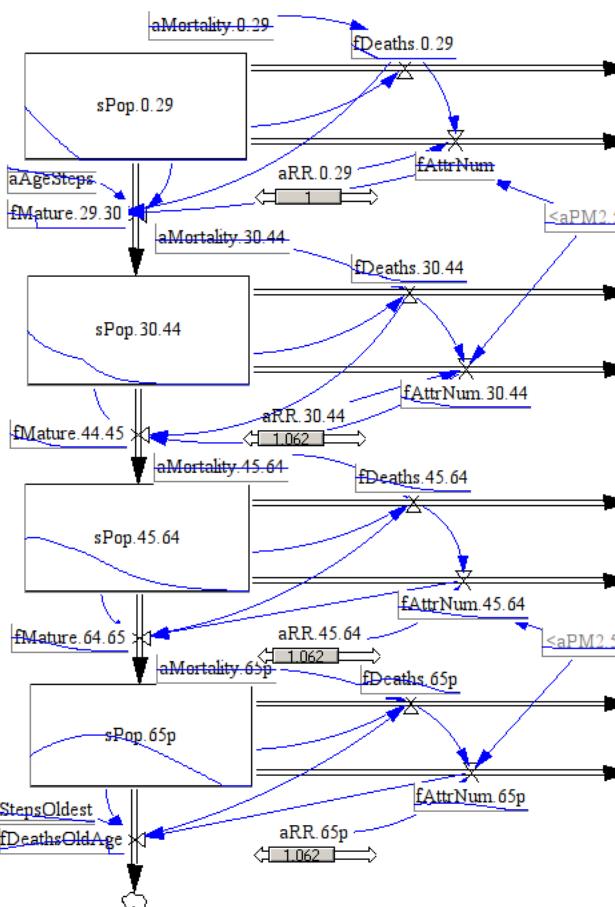
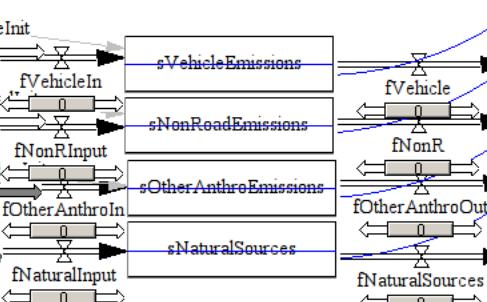
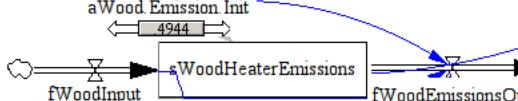
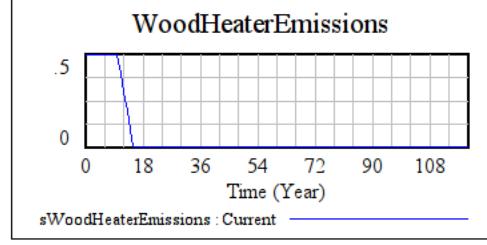
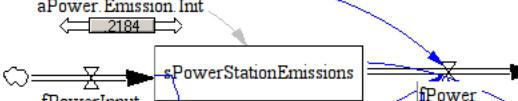
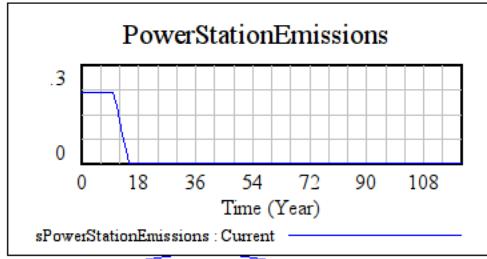
Graph

Table

Table Time

Units Clock

Res Compare



# “Coached” practice time

- For those of you who have downloaded the tools from Github
  - [https://github.com/cardat/DatSciTrain\\_HIA\\_Health\\_Impact\\_Functions\\_explained](https://github.com/cardat/DatSciTrain_HIA_Health_Impact_Functions_explained)
- We will start by looking at the 02\_Lifetables\_with\_spreadsheet
  - We will not have time to look at the other tools today
  - Feel free to play around with these later, and send any feedback via Github issues or pull requests.



Curtin University

WHO COLLABORATING CENTRE FOR CLIMATE  
CHANGE AND HEALTH IMPACT ASSESSMENT

Our approach:  $YLL = \sum \text{Attributable N due to exposure} \times \text{Life Expectancy w/o exposure}$

### Scenario A with no change in exposure

	RR per 10	1.060	beta	0.006	deltaX	0	pop	death	AN	Expected	$M_x$	$e_x$	YLL
x	n	a <sub>x</sub>											
<1	0	1	0.1				2,533	20		20	0.007896	71.99	0.00
1-4	1	4	0.5				11,130	1		1	0.000090	71.55	0.00
5-9	5	5	0.5				15,519	2		2	0.000129	67.58	0.00
10-14	10	5	0.5				16,409	4		4	0.000244	62.62	0.00
15-19	15	5	0.5				16,133	9		9	0.000558	57.70	0.00
20-24	20	5	0.5				21,482	10		10	0.000466	52.85	0.00
25-29	25	5	0.5				15,997	22		22	0.001375	47.97	0.00
30-34	30	5	0.5				16,026	35	0	35	0.002184	43.28	0.00
35-39	35	5	0.5				19,800	34	0	34	0.001717	38.73	0.00
40-44	40	5	0.5				16,076	39	0	39	0.002426	34.04	0.00
45-49	45	5	0.5				13,404	59	0	59	0.004402	29.43	0.00
50-54	50	5	0.5				13,027	108	0	108	0.008290	25.02	0.00
55-59	55	5	0.5				10,051	136	0	136	0.013531	20.98	0.00
60-64	60	5	0.5				10,220	176	0	176	0.017221	17.27	0.00
65-69	65	5	0.5				9,190	320	0	320	0.034820	13.60	0.00
70-74	70	5	0.5				7,427	445	0	445	0.059917	10.72	0.00
75-79	75	5	0.5				5,231	414	0	414	0.079144	8.61	0.00
80-84	80	5	0.5				2,884	355	0	355	0.123093	6.63	0.00
85+	85	11	0.5				1,840	347	0	347	0.188587	5.30	0.00
												0.00	TOTAL YLL

### Scenario B with reduction of 3 $\mu\text{g}/\text{m}^3$ PM<sub>2.5</sub>

	RR per 10	1.060	beta	0.006	deltaX	3	pop	death	AN	Expected	$M_x$	$e_x$	YLL
x	n	a <sub>x</sub>											
<1	0	1	0.1				2,533	20		20	0.007896	72.18	0.00
1-4	1	4	0.5				11,130	1		1	0.000090	71.75	0.00
5-9	5	5	0.5				15,519	2		2	0.000129	67.77	0.00
10-14	10	5	0.5				16,409	4		4	0.000244	62.81	0.00
15-19	15	5	0.5				16,133	9		9	0.000558	57.89	0.00
20-24	20	5	0.5				21,482	10		10	0.000466	53.04	0.00
25-29	25	5	0.5				15,997	22		22	0.001375	48.16	0.00
30-34	30	5	0.5				16,026	35	1	34	0.002146	43.47	26.37
35-39	35	5	0.5				19,800	34	0	33	0.001687	38.92	22.93
40-44	40	5	0.5				16,076	39	1	38	0.002384	34.22	23.13
45-49	45	5	0.5				13,404	59	1	58	0.004325	29.60	30.27
50-54	50	5	0.5				13,027	108	2	106	0.008147	25.20	47.16
55-59	55	5	0.5				10,051	136	2	134	0.013297	21.14	49.82
60-64	60	5	0.5				10,220	176	3	173	0.016923	17.42	53.14
65-69	65	5	0.5				9,190	320	6	314	0.034217	13.74	76.20
70-74	70	5	0.5				7,427	445	8	437	0.058878	10.84	83.62
75-79	75	5	0.5				5,231	414	7	407	0.077772	8.72	62.59
80-84	80	5	0.5				2,884	355	6	349	0.120960	6.73	41.40
85+	85	11	0.5				1,840	347	6	341	0.185319	5.40	32.45
												549.06	TOTAL YLL

Adjust the death rate ( $M_x$ ) by the attributable number given RR and exposure change (deltaX)

Difference in life expectancy ( $e_x$ ) at birth = 72.18 - 71.99, about 69 days are lost in scenario A

This is the potential years of life expectancy gained in scenario B

# Summary, Conclusions and Feedback

- Demonstrated an end-to-end example of health burden analysis, discussed moving from burden to impact
- Discussed key points for interpretation
- Access a worked example for hands-on “Coached” practice
- Now: please provide feedback using the survey at Slido.com with #3370 651

