Indoor dust bacterial and fungal microbiota composition and allergic diseases: a scoping review to construct reusable DAG

Inge Wouters and **Alex Bossers**
Daily supervisors

Javier Mancilla Galindo
MSc Epidemiology student

Lidwien Smit
Examiner

Major Research Project
One Health Epidemiology Microbial Agents
Institute for Risk Assessment Sciences

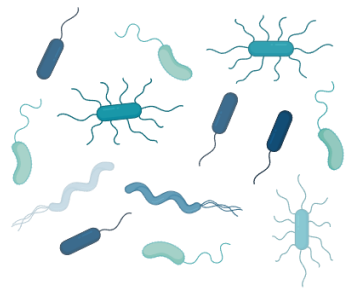
Background

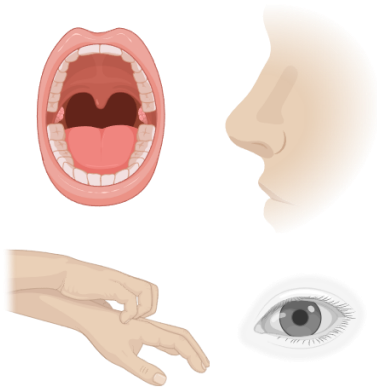
Prevalence Allergic Diseases[1]

Asthma: 3.42% (95%CI: 2.98 – 3.91)
Atopic dermatitis: 1.69% (95%CI: 1.62– 1.76)

Allergy is an exaggerated immune response against predominantly environmental or other exogenous antigens.[2]

Different microbes can promote or protect form allergies as a result of complex immunologic, genetic, and enviromental interactions.[3]





Indoor dust particles are an important source of exposure to microbes since humans spend up to 90% of their time indoors.[4]

1. Ferrari A, et al. *Lancet;* 2024; 403(10440): 2133-2161.
2. Cruz RH, et al. *Allergol Immunopathol;* 2023; 51: 16–21.
3. Kelly MS, et al. *J Allergy Clin Immunol Pract*; 2022; 10: 2206-2217.e1.
4. Shan Y, et al. *Int Microbiol* International Microbiology; 2019; 22: 297–304.

Background

Causal inference is increasingly being applied in microbiome studies

Examples

- RCT microbiota transplantation[5]
- Mendelian randomization[6]
- Mediation analysis[7]

Challenges

- No interference assumption[8]
- Consideration of interactions[9]
- Cross-sectional data[10]

5. Yu EW, et al. *PLOS Medicine.* 2020;17(3):e1003051
6. Li Z, et al. *Gastroenterology.* 2024;166(2):354-355
7. Lin L, et al. *Nature Medicine.* 2023;29(7):1750-1759
8. Hudgens, MG, et al. *J Am Stat Assoc.* 2008;103(482):832–842.
9. VanderWeele TJ. *Epidemiology.* 2014;25(5):749-61.
10. Corander J, et al. *Lancet Microbe.* 2022;3(11):e881-e887

Objectives

1. Characteristics of studies of the indoor dust microbiome
2. Study design, causal statements, and variable selection
3. Directed acyclic graph (DAG)
indoor dust microbiome → allergies

Methods

Search Strategy

- WorldCat.org and SciELO
- Keywords:

  kw:(dust) AND kw:(indoor OR dwelling OR house OR household OR residence OR store OR mall OR hospital OR workplace OR office OR school OR university OR sport OR closed environment) AND kw:(microbiota OR microbiome OR microbial community OR virome)

- initial (08-03-2023), updated: 15-09-2023 and 30-01-2024
- English and Spanish
- 2000-2024

Data extraction

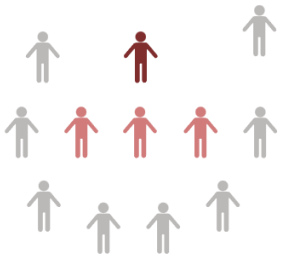
Country

Microbiome type

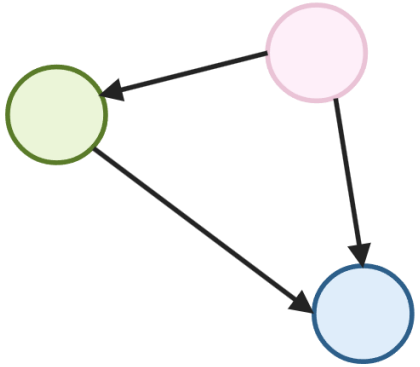
Environmental determinants

Indoor environment

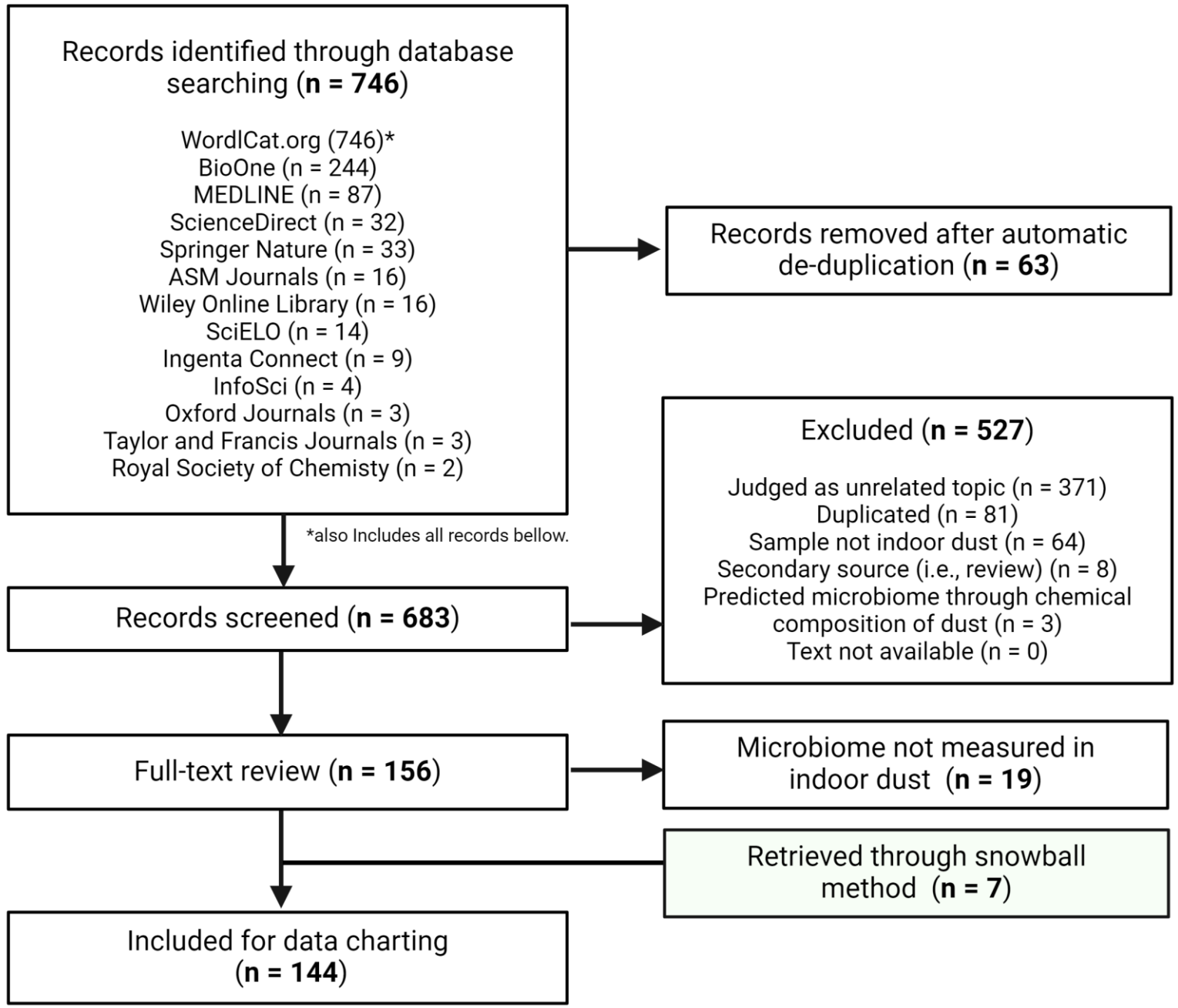
Sample collection method



Study Design

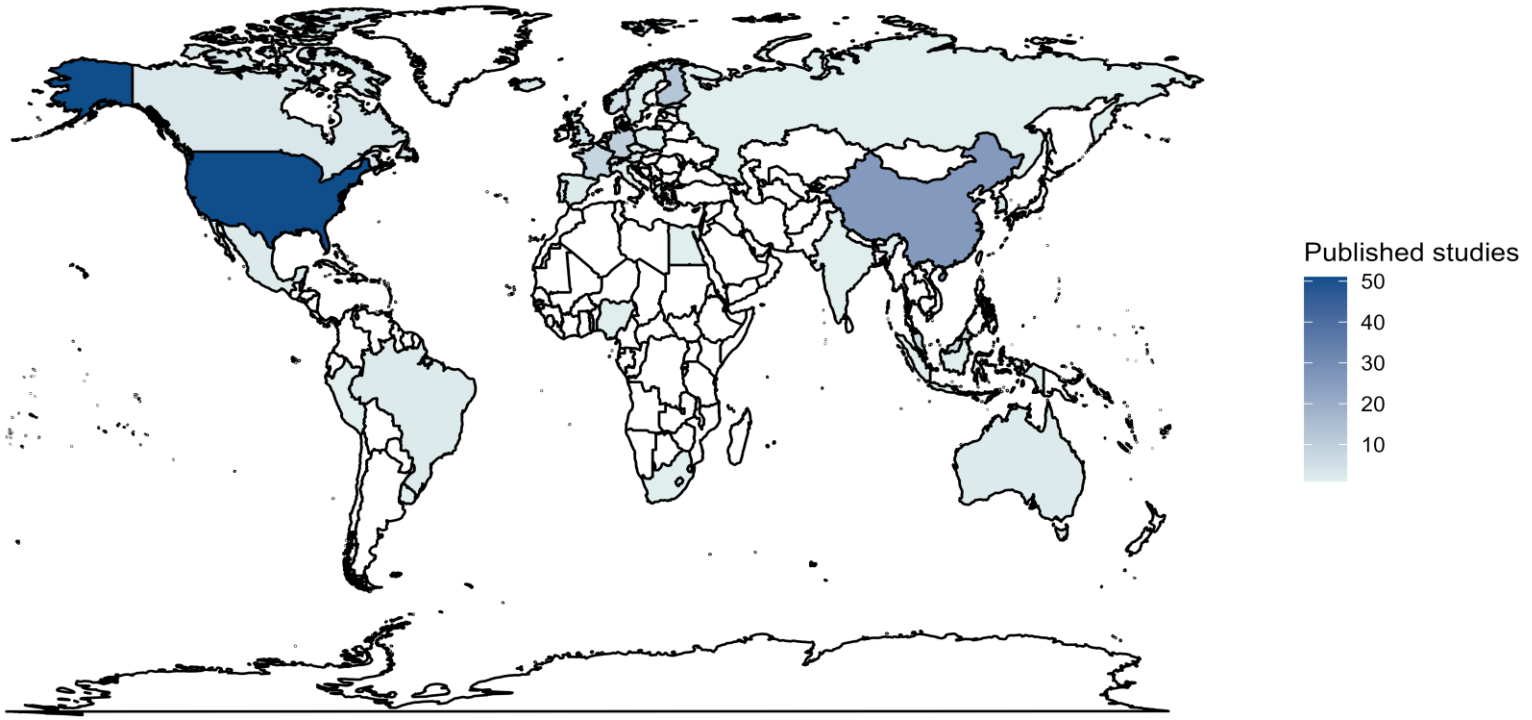


Causal statements and variable selection


Records identified through database searching (n = 746)
WordlCat.org (746)*
BioOne (n = 244)
MEDLINE (n = 87)
ScienceDirect (n = 32)
Springer Nature (n = 33)
ASM Journals (n = 16)
Wiley Online Library (n = 16)
SciELO (n = 14)
Ingenta Connect (n = 9)
InfoSci (n = 4)
Oxford Journals (n = 3)
Taylor and Francis Journals (n = 3)
Royal Society of Chemisty (n = 2)
*also Includes all records bellow.
Records removed after automatic de-duplication (n = 63)
Records screened (n = 683)
Excluded (n = 527)
Judged as unrelated topic (n = 371)
Duplicated (n = 81)
Sample not indoor dust (n = 64)
Secondary source (i.e., review) (n = 8)
Predicted microbiome through chemical composition of dust (n = 3)
Text not available (n = 0)
Full-text review (n = 156)
Microbiome not measured in indoor dust (n = 19)
Retrieved through snowball method (n = 7)
Included for data charting (n = 144)

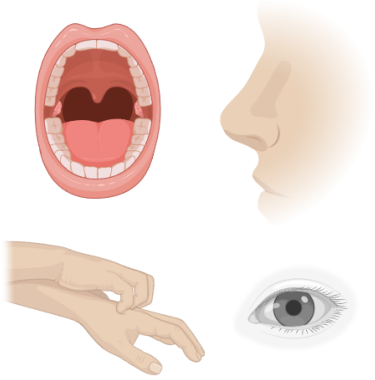
World region	n	%
Europe & Central Asia	70	40.0
North America	54	30.9
East Asia & Pacific	41	23.4
Latin America & Caribbean	6	3.4
Sub-Saharan Africa	2	1.1
Middle East & North Africa	1	0.6
South Asia	1	0.7

Indoor dust microbiome research articles per country



Income	n	%
High	131	74.9
Upper-middle	40	22.9
Lower-middle	4	2.3
Low	0	0%

Topics

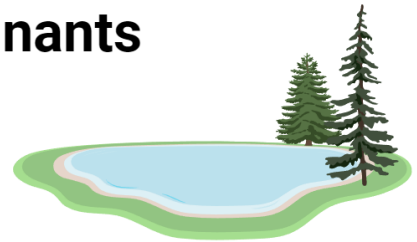


Allergies

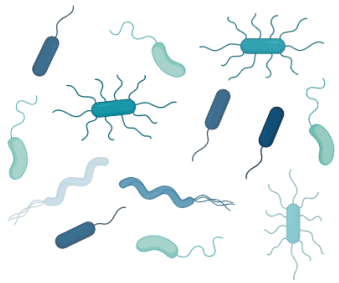
28.5% (n = 41)

Environmental Determinants

80.6% (n = 116)



Microbiome type

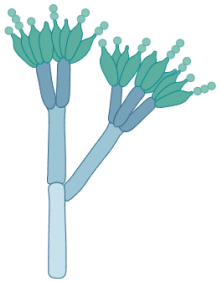


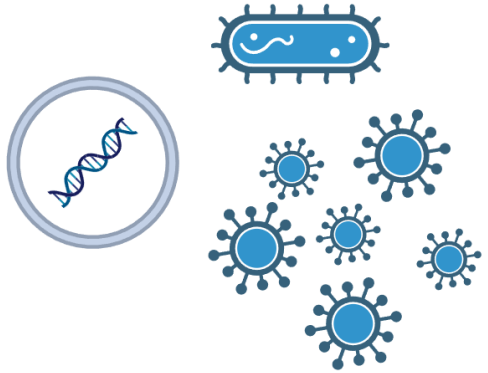
Bacterial

88.9% (n = 128)

Fungal

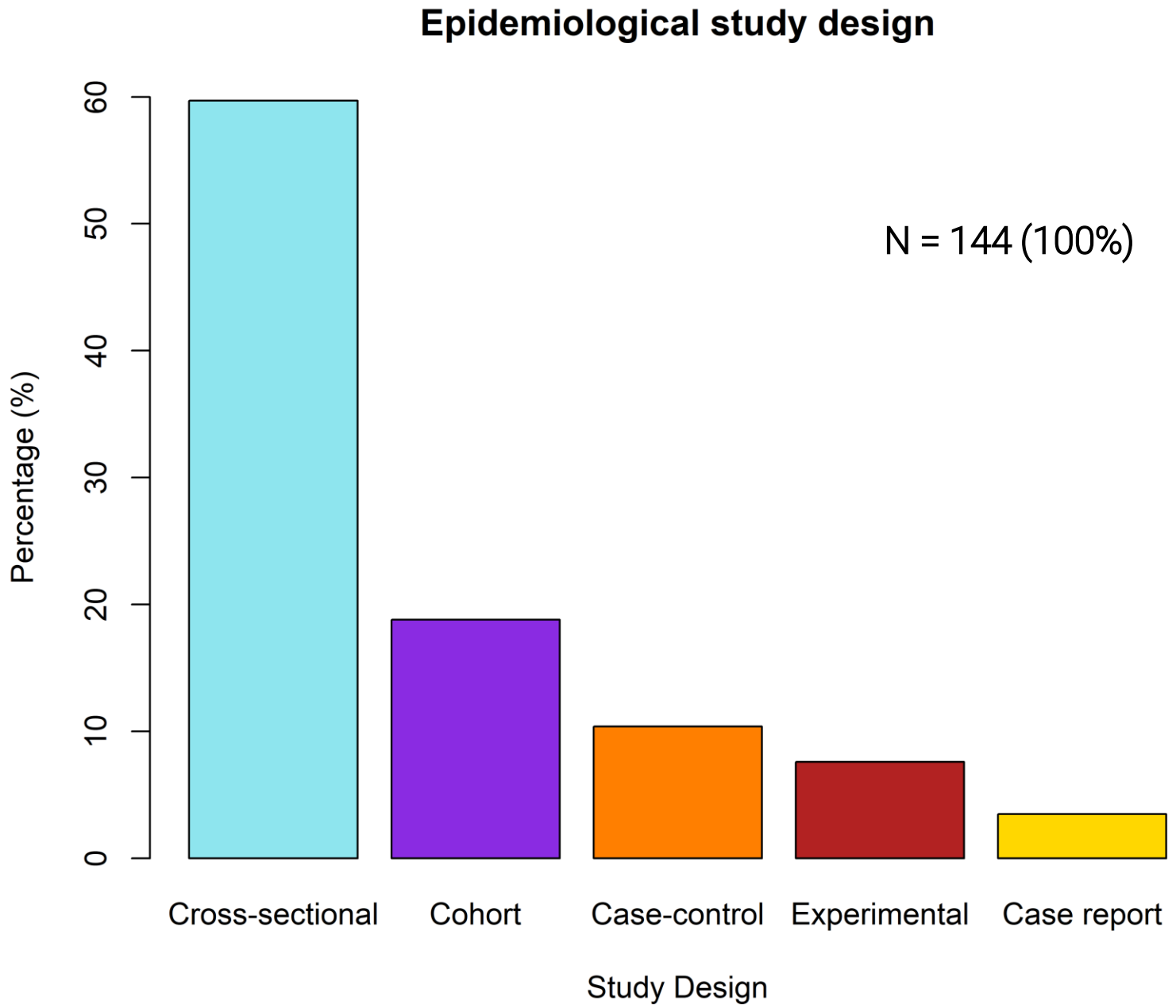
57.6% (n = 83)





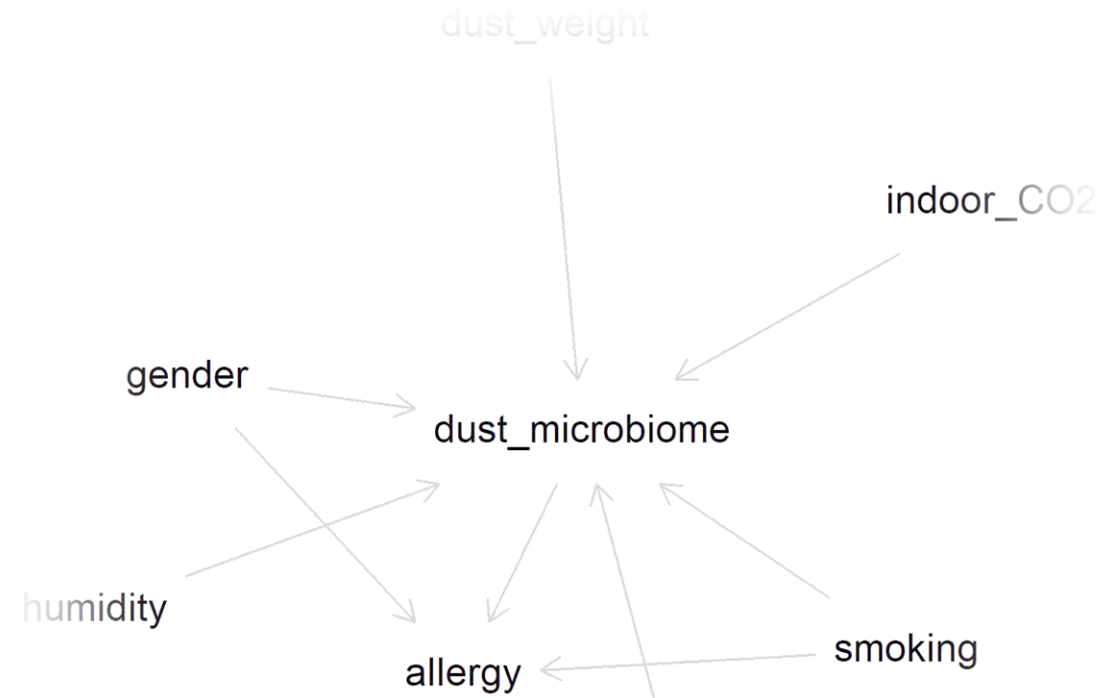
Other

6.9% (n = 10)


Epidemiological study design
Percentage (%)
0
10
20
30
40
50
60
N = 144 (100%)
Cross-sectional
Cohort
Case-control
Experimental
Case report
Study Design

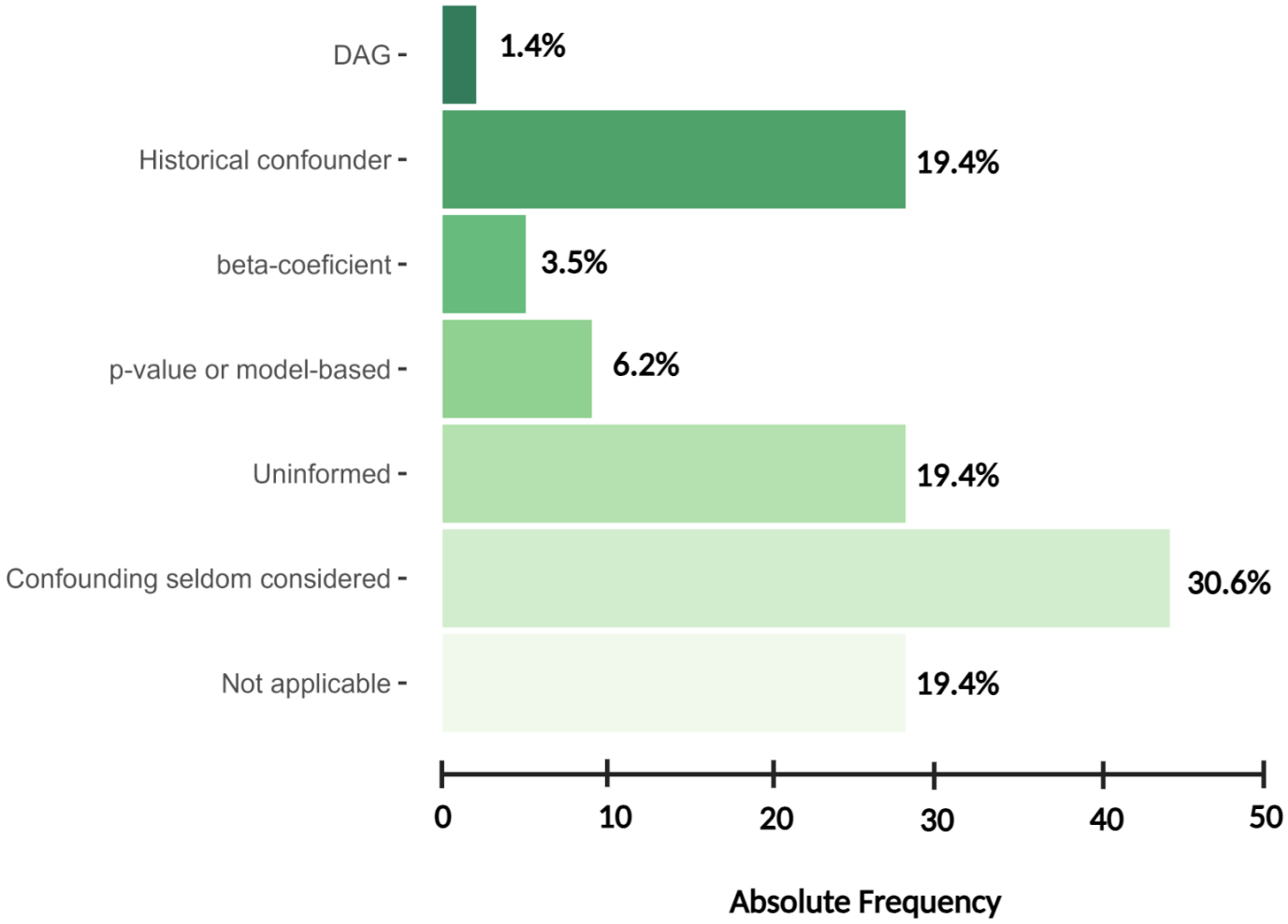
Discussion and Conclusions (Part 1)

- Research efforts have been focused in high and upper-middle income countries of three out of seven world regions.
- Environmental determinants (80.6%) and allergies (28.5%) have been recurrent topics.
- Most studies have assessed the bacterial dust microbiome, followed by fungal microbiome. The indoor dust virome remains understudied.
- Cross-sectional is the most common epidemiological study design (~60%)



Causal statements were made in 81.9% (n = 118) of studies

Method for selecting and accounting for confounding variables



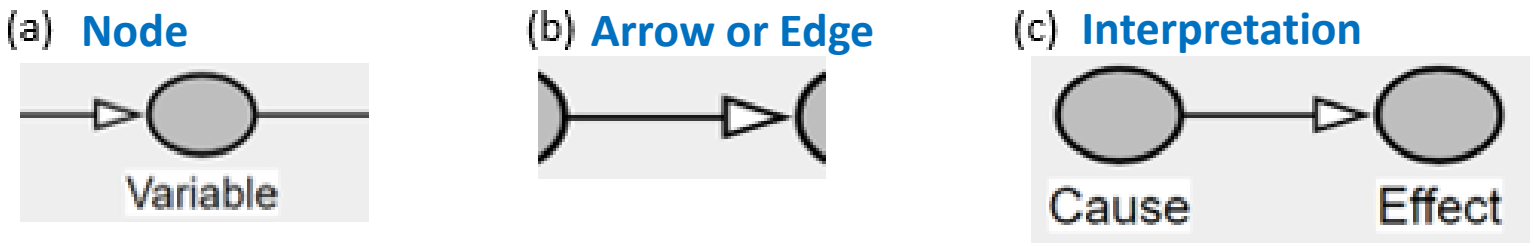
Lederer DJ, et al. *Annals of the American Thoracic Society* 2019;16(1);22-28

Constructing and documenting a reusable Directed Acyclic Graph (DAG)

dust microbiome $\rightarrow$ *allergy*

Directed Acyclic Graphs

Directed acyclic graphs (DAGs) posit causal relationships between variables as arrows between nodes.

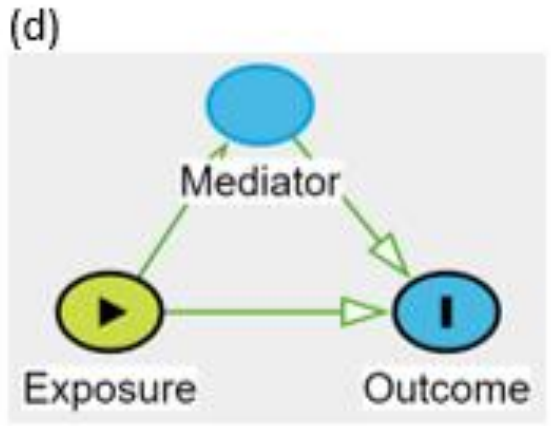


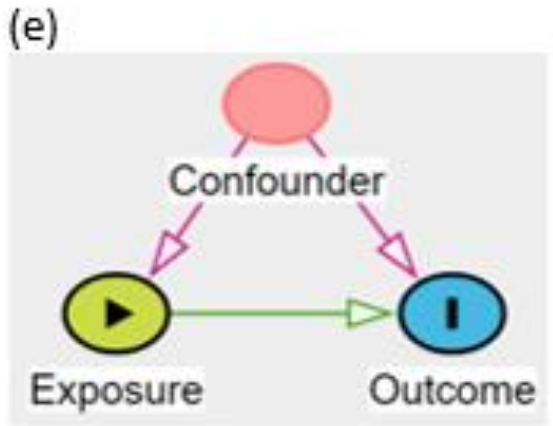
- Directed means that arrows (edges) can only be unidirectional
- Acyclic means that no combination of arrows may constitute a loop

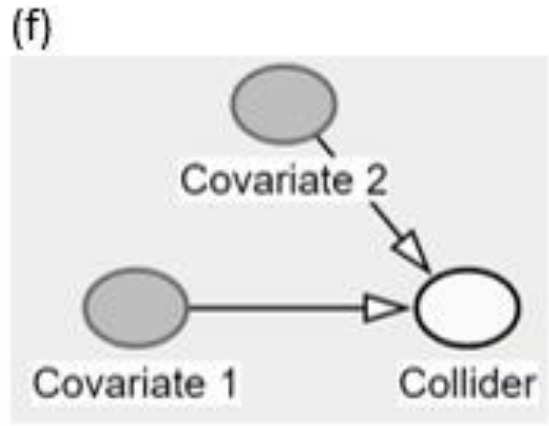
Ferguson KD, et al. *International Journal of Epidemiology* 2020;49;322-29

Directed Acyclic Graphs

DAGs are valuable tools for study design and guiding analysis.







https://dagitty.net/dags.html

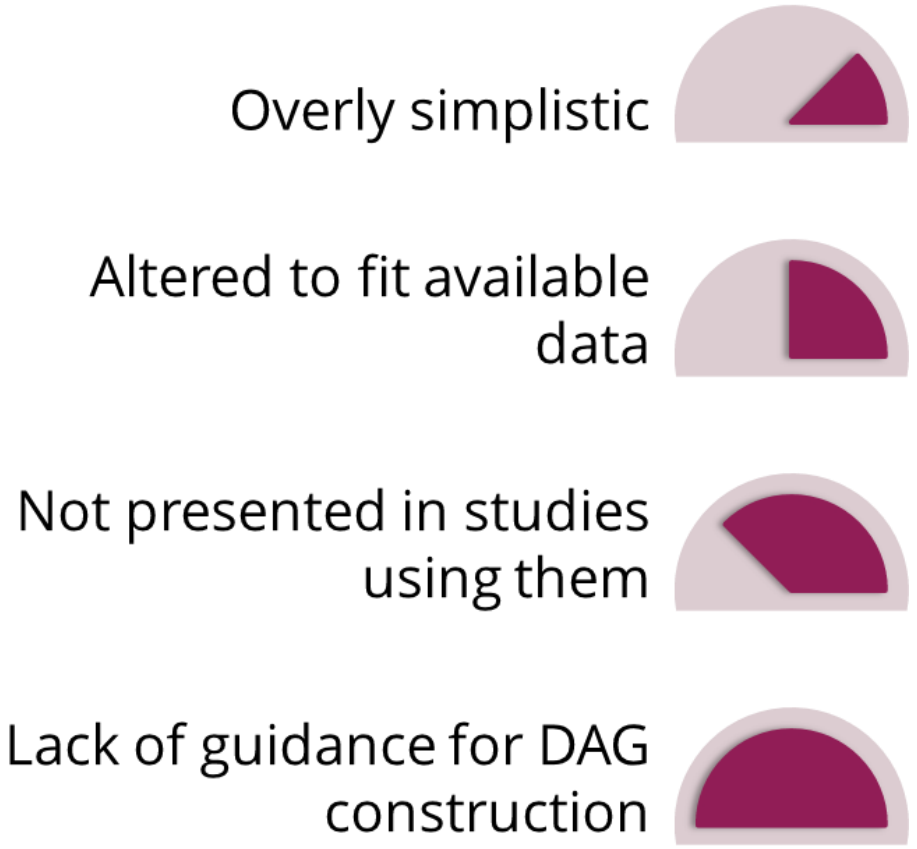
Ferguson KD, et al. *International Journal of Epidemiology* 2020;49;322-29

Directed Acyclic Graphs

- Any variable that influences at least two others should be included.
- Absence of an arrow between nodes indicates no causal effect.
- Researchers should 'work backwards' from a 'saturated' DAG—one in which all variables are inter-connected—and only delete connections that are thought impossible.

Ferguson KD, et al. *International Journal of Epidemiology* 2020;49;322-29

Problems with DAGs



Ferguson KD, et al. *International Journal of Epidemiology* 2020;49;322-29

ESC-DAG protocol

Evidence Synthesis for Constructing DAGs

Literature Search

- **Ideally** novel systematic review or review of SRs
- Well defined PICO/PECO question

Mapping

- Produce a DAG representing the conclusions of each individual study

Translation

- Assess causal characteristics of each connection

Integration

- Synthesis: Combine DAGs into one
- Recombination: Combine nodes for practical or substantive reasons

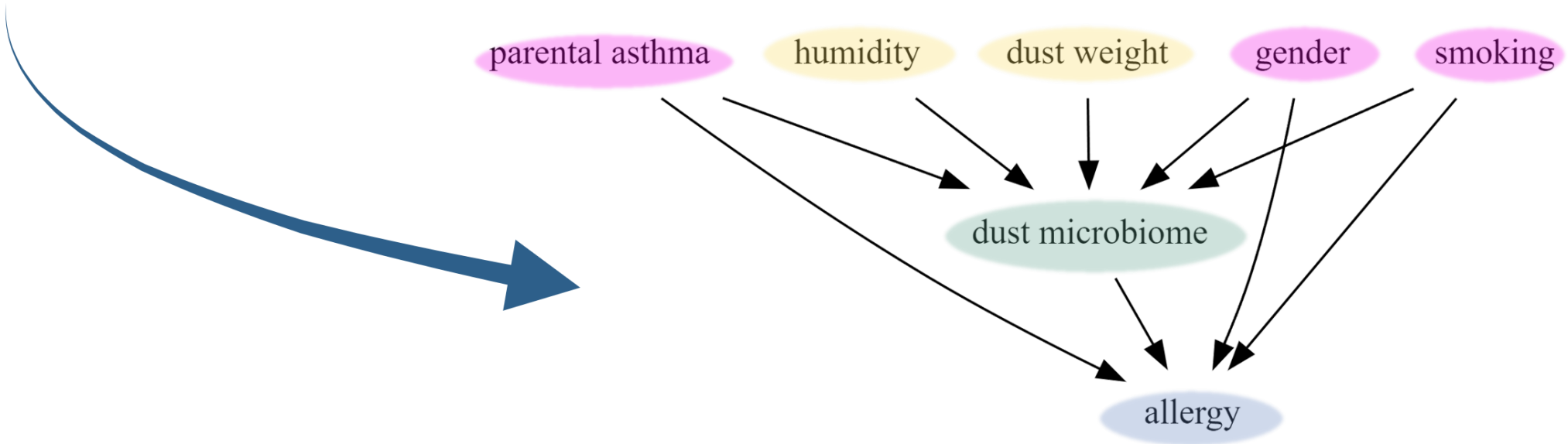
Ferguson KD, et al. *International Journal of Epidemiology* 2020;49;322-29

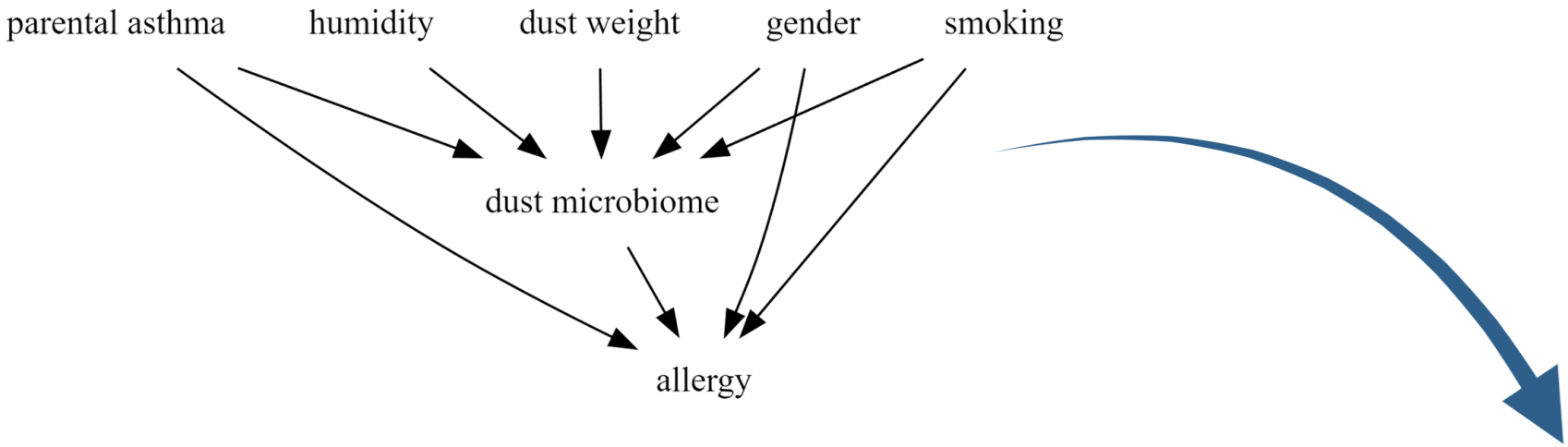
Mapping

23 studies selected
- Exposure: indoor dust microbiome
- Outcome: any allergy

Num 1 (Fu et al., 2023)

- Exposure: microbial richness/concentration
- Outcome: allergic rhinitis / non-allergic rhinitis symptoms
- Control variables: **gender**, current **smoking**, and **parental asthma**
- Other: Effect of **relative humidity** and **dust weight** on rhinitis symptoms is concluded to be *mediated* through the microbiome.



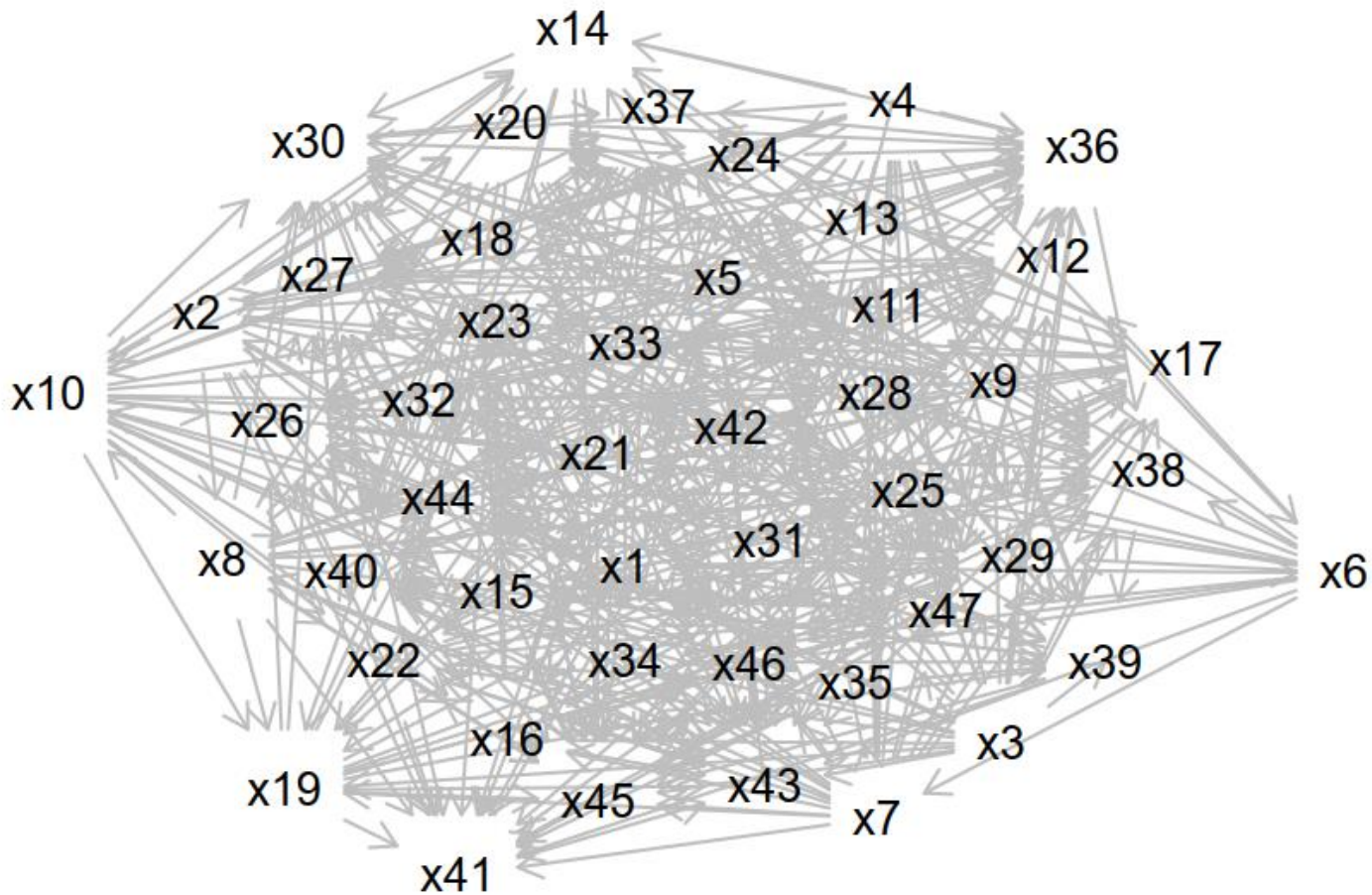


Num	Edge	Explanation
1	dust_microbiome -> allergy	main
1	dust_weight -> dust_microbiome	microbiome mediates the effect of dust weight on allergies
1	gender -> allergy	adjusted for gender
1	gender -> dust_microbiome	adjusted for gender
1	humidity -> dust_microbiome	microbiome mediates the effect of humidity on allergies
1	parental_asthma -> allergy	adjusted for parental_asthma
1	parental_asthma -> dust_microbiome	adjusted for parental_asthma
1	smoking -> allergy	adjusted for smoking
1	smoking -> dust_microbiome	adjusted for smoking

Current Status

- So far, I have assessed 13 out of the 23 studies included for the mapping stage of the ESC-DAG.
- This produced a total of 140 edges. The number of edges after de-duplication is 80.
- The number of variables (nodes) present is 47.
- If I were to only assess relationships between every variable with the exposure and outcome, ignoring all relationships between covariates, I would need to assess 90 edges.

How many edges would I need to assess to produce a saturated DAG?



This hypothetical DAG with the same number of variables contains **547** edges that I would need to assess individually.

Discussion and Conclusions (Part 2)

- Evidence synthesis construction of DAGs remains challenging and with technical limitations.
- Papers citing this method are using it in multiple deviating ways. Only few exceptional compliances, including NASA.[11]
- Authors of ESC-DAG attempted to construct a DAG with their own method and reported multiple logistical and technical limitations; dagitty website crashes with >30 nodes (variables).[12]
- Documentation and openness of ESC-DAGs has not been good. In my approach I am investing additional efforts on it.

11. Antonsen E, et al. npj Microgravity. 2024;19;10(1):32.
12. Campbell T, et al. 2021 PsyArXiv. osf.io/preprints/psyarxiv/sfp4x

Discussion and Conclusions (Part 2)

- ESC-DAGs challenge traditional epidemiologists' conceptions of DAGs.
- Current utility limited to one research question. Naive to expect researchers to follow method for every study
- Indoor dust microbiome DAGs will remain limited if there is a lag in flow of basic knowledge.
- Causal questions of *mediation* in the indoor dust microbiome field anticipated to be of special interest. Documented, reusable, and adaptable DAGs could inform the design and analysis of such studies.

Future work and perspectives

- For my own DAG, I must decide how to approach number of edges to assess in the *synthesis* phase
- Develop a resource to integrate good documentation with easiness of use. Again, documentation and open practices of ESC-DAG development need to improve.
- Future uses of robust and reproducible DAGs could resemble applications in commerce (i.e., Amazon) or NASA risk prevention

Utrecht University

DISCLAIMER

The information in this presentation has been compiled with the utmost care,
but no rights can be derived from its contents.