

Implied graphs from individual studies

Indoor dust microbiome (exposure) and allergic diseases (outcome)

Javier Mancilla Galindo, MSc student

Supervisors: Inge Wouters and Alex Bossers

Examiner: Lidwien Smit

2024-05-27

Table of contents

Packages and session information	2
Load data	3
Selection of studies	3
Mapping individual studies into implied graphs	3
Num 1 (Fu et al., 2023)	4
Num 9 (Fu et al., 2022)	7
Num 14 (Cox et al., 2022)	7
Num 15 (Du et al., 2022)	7
Num 19 (Fu et al., 2021a)	8
Num 25 (Ta et al., 2021)	8
Num 27 (Hyytiäinen et al., 2021)	8
Num 32 (Vandenborght et al., 2021)	9
Num 33 (Adams et al., 2021)	9
Num 35 (Fu et al., 2021b)	9
Num 37 (Fu et al., 2020)	10
Num 38 (Gangneux et al., 2020)	10
Num 49 (Karvonen et al., 2019)	10
Join edges tables and assess characteristics	11
Package references	12
References studies	13

Packages and session information

```
if (!require("pacman", quietly = TRUE)) {  
  install.packages("pacman")  
}  
  
pacman::p_load(  
  tidyverse,    # Used for basic data handling and visualization.  
  dagitty,      # Used to produce implied graphs and hypothetical DAG, and to  
                # analyze properties of such graphs.  
  DiagrammeR,  # Used to plot example DAG of study Num 1  
  gt,          # Print and save html tables.  
  report       # Used to generate package citations in markdown format.  
)
```

R version 4.4.0 (2024-04-24 ucrt)
Platform: x86_64-w64-mingw32/x64
Running under: Windows 11 x64 (build 22631)

Matrix products: default

locale:

[1] LC_COLLATE=Spanish_Mexico.utf8 LC_CTYPE=Spanish_Mexico.utf8
[3] LC_MONETARY=Spanish_Mexico.utf8 LC_NUMERIC=C
[5] LC_TIME=Spanish_Mexico.utf8

time zone: Europe/Berlin

tzcode source: internal

attached base packages:

[1] stats graphics grDevices utils datasets methods base

other attached packages:

[1] report_0.5.8	gt_0.10.1	DiagrammeR_1.0.11	dagitty_0.3-4
[5] lubridate_1.9.3	forcats_1.0.0	stringr_1.5.1	dplyr_1.1.4
[9] purrr_1.0.2	readr_2.1.5	tidyr_1.3.1	tibble_3.2.1
[13] ggplot2_3.5.1	tidyverse_2.0.0	pacman_0.5.1	

Load data

```
Formal class 'DataFrameCollection' [package ".GlobalEnv"] with 0 slots
Named list()
..$ data          : tibble [144 x 31] (S3: tbl_df/tbl/data.frame)
..$ countries     : tibble [175 x 5] (S3: tbl_df/tbl/data.frame)
..$ collectors    : tibble [184 x 2] (S3: tbl_df/tbl/data.frame)
..$ buildings     : tibble [159 x 2] (S3: tbl_df/tbl/data.frame)
..$ environmental_determinants: tibble [595 x 3] (S3: tbl_df/tbl/data.frame)
..$ references    : 'data.frame': 144 obs. of 90 variables:
```

Selection of studies

I will create a subset of studies named `subset_studies` that includes only those studies for which the *microbiome* is the exposure and *allergy* is the outcome.

Out of the 144 entries in `data`, there are **23 studies** meeting the inclusion criteria for mapping into a directed acyclic graph according to the procedure described by (Ferguson et al., 2020).

Note: Studies where the reverse association (*allergy* -> *microbiome*) was studied could have also been included in the mapping procedure. However, none of the studies included reported having studied such directionality of association.

Mapping individual studies into implied graphs

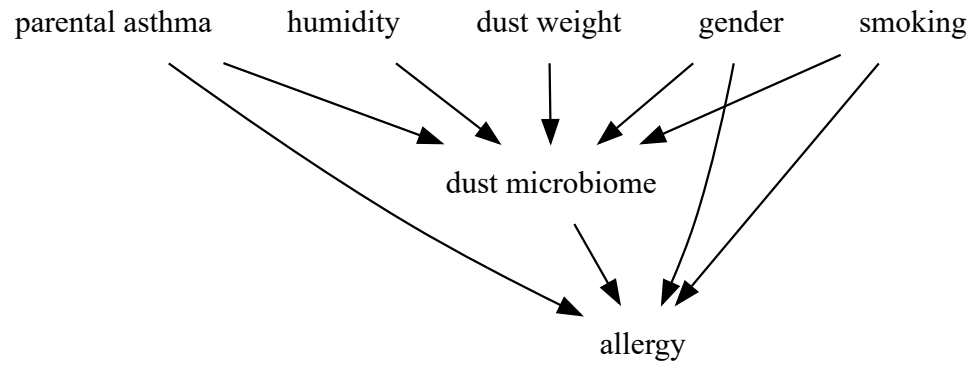
The figures for every graph produced by the subsequent code sourced for every study can be found in the [results folder](#).

Note that the saved figures will have different variable layouts every time they are generated. I have decided not to attach coordinates to variables to simplify the implied graphs' code since the interest at this stage is on the content of the graph rather than visualization.

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.

The implied graph would have the following appearance:



This is an example of the tables that will be produced for each implied graph to be used in the synthesis phase to produce a DAG.

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

Num 9 (Fu et al., 2022)

- Exposure: microbial richness/ taxa absolute quantity
- Outcome: asthma
- Control variables: **gender**, current **smoking**, and **parental asthma**
- Other: Effect of **relative humidity**, **indoor CO2**, and **dust weight** on asthma are potentially *mediated* through the microbiome.

```
source("implied_graphs/9_fu2022a.R")
```

Num 14 (Cox et al., 2022)

The edges in this graph have been extracted as is from Figure 1 in the paper.

```
source("implied_graphs/14_cox2022a.R")
```

By definition, directed acyclic graphs should not contain cycles. Since this DAG is more complex, I will test if it is indeed acyclic.

```
isAcyclic(ig_14_cox2022a)
```

```
[1] TRUE
```

Num 15 (Du et al., 2022)

- Exposure: airborne fungal concentration
- Outcome: asthma / wheeze / allergic rhinitis / eczema
- Control variables: gender, age, urbanicity, and genetic disorders in family
- Other: Concluded that allergy is related to season and humidity, whereas indoor fungal microbiome is correlated with season, humidity, temperature, indoor CO2, and the outdoor microbiome.

```
source("implied_graphs/15_du2022a.R")
```

Num 19 (Fu et al., 2021a)

- Exposure: bacterial richness / bacterial genera
- Outcome: asthma (self-administered questionnaire, supervised by healthcare professionals)
- Control variables: **gender**, current **smoking**, and **parental asthma**
- Other: plants in room, location in building, age of building, and wall surface type were associated with indoor microbiome. These characteristics were not associated with asthma.

```
source("implied_graphs/19_fu2021a.R")
```

Num 25 (Ta et al., 2021)

- Exposure: bacterial relative abundance
- Outcome: early life eczema
- Control variables: None
- Other: None of the conclusions support other variables to include in graph.

```
source("implied_graphs/25_ta2021a.R")
```

Num 27 (Hyytiäinen et al., 2021)

- Exposure: Bacterial richness
- Outcome: Inhalant atopy (sensitization against at least one of the 13 or 8 tested specific inhalant allergens) / allergic rhinitis at 10-1.5 years.
- Control variables: gender, parental atopy (either mother or father reporting asthma, eczema or hay fever, if at least one of the parents were answered the question) and number of older siblings in both cohorts, maternal education, living on a farm and cohort in LUKAS; and parental education (the highest level of either parent) and study center in LISA. season of dust sampling in both cohorts; the number of different pet species indoors in LUKAS; and age of the mother during delivery in LISA.
- Other: None

```
source("implied_graphs/27_hyytiaainen2021a.R")
```


Num 32 (Vandenborgh et al., 2021)

- Exposure: diversity, abundance
- Outcome: severe asthma endotype (T2-high vs T2-low) / asthma exacerbation.
- Control variables: None declared.
- Other: None declared.

```
source("implied_graphs/32_vandenborgh2021a.R")
```

Num 33 (Adams et al., 2021)

- Exposure: bacterial/fungal diversity, richness, individual taxa. Additionally, moisture damage of schools was an exposure of interest.
- Outcome: respiratory symptoms: wheeze, nocturnal dry cough, and rhinitis, on questionnaire administered 2 months **before** exposure assessment.
- Control variables: gender, age, moisture damage in home (note that this is different to moisture damage in school, where samples were taken), educational level
- Other: endotoxin and microbial taxa mediate 30% of the effect of moisture damage in school on respiratory symptoms.

```
source("implied_graphs/33_adams2021a.R")
```

Num 35 (Fu et al., 2021b)

- Exposure: microbial richness/abundance
- Outcome: asthma / wheezing, / rhinitis / rhinoconjunctivitis
- Control variables: smoking, gender, parental_allergies
- Other: urbanicity related to dust microbiome and allergy. Effect of dust microbiome could occur through metabolism-related genes and immune-related mechanisms.

```
source("implied_graphs/35_fu2021e.R")
```

Num 37 (Fu et al., 2020)

- Exposure: richness / taxa
- Outcome: history of asthma / asthma
- Control variables: smoking, gender, parental_allergies
- Other: indoor dampness/ visible mold (I will record this as humidity), which did not associate with microbial community variation, was associated with the asthma-associated bacteria. Other environmental characteristics in this study were not significantly associated.

```
source("implied_graphs/37_fu2020a.R")
```

Num 38 (Gangneux et al., 2020)

- Exposure: bacterial and fungal taxa
- Outcome: household with inhabitants with an asthma diagnosis
- Control variables: Not reported
- Other: No additional statements

```
source("implied_graphs/38_j2020a.R")
```

Num 49 (Karvonen et al., 2019)

- Exposure: richness / diversity / taxa
- Outcome: asthma / respiratory symptoms / wheezing phenotypes
- Control variables: living on a farm, maternal history of allergic diseases, sex, number of older siblings, and smoking during pregnancy. Study cohort and time were additional adjustment variables which I will not include in DAG since I believe these are typical variables of control due to the study design.
- Other: increased animal contacts, natural as opposed to mechanical ventilation and timber structures were associated with increase in asthma protection associated microbes and with decrease in the asthma predisposing Lactococcus abundance.

```
source("implied_graphs/49_karvonen2019a.R")
```

Join edges tables and assess characteristics

So far, I have assessed **13** out of the **23 studies** studies included for the DAG mapping.

This procedure produced a total of **140** edges.

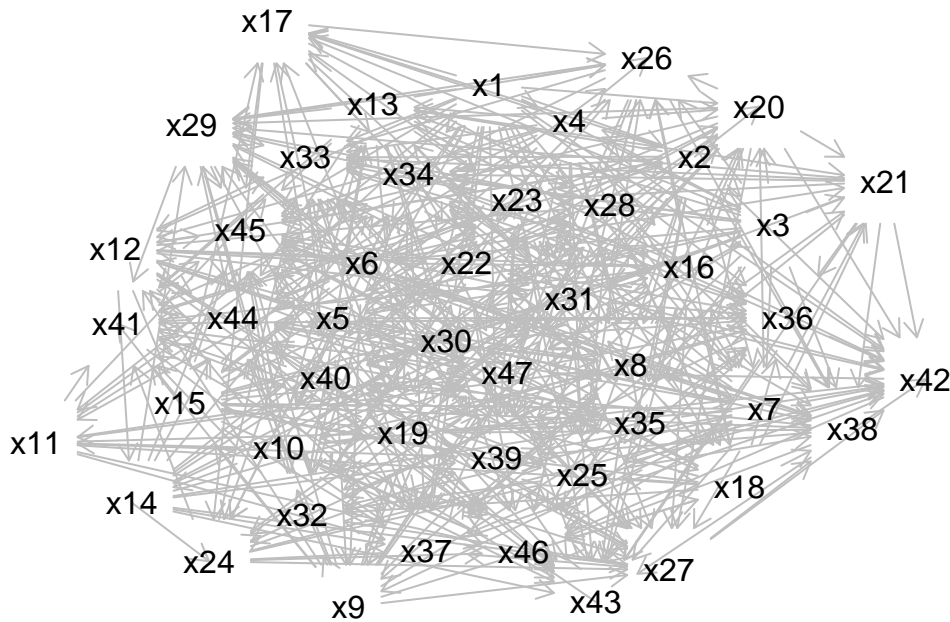
The number of edges after de-duplication is:

n
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 **534** edges that I would need to assess individually.

Package references

- Grolemund G, Wickham H (2011). “Dates and Times Made Easy with lubridate.” *Journal of Statistical Software*, 40(3), 1-25. <https://www.jstatsoft.org/v40/i03/>.
- Iannone R, Cheng J, Schloerke B, Hughes E, Lauer A, Seo J (2024). *gt: Easily Create Presentation-Ready Display Tables*. R package version 0.10.1, <https://CRAN.R-project.org/package=gt>.
- Makowski D, Lüdtke D, Patil I, Thériault R, Ben-Shachar M, Wiernik B (2023). “Automated Results Reporting as a Practical Tool to Improve Reproducibility and Methodological Best Practices Adoption.” *CRAN*. <https://easystats.github.io/report/>.
- Müller K, Wickham H (2023). *tibble: Simple Data Frames*. R package version 3.2.1, <https://CRAN.R-project.org/package=tibble>.
- R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rinker TW, Kurkiewicz D (2018). *pacman: Package Management for R*. version 0.5.0, <http://github.com/trinker/pacman>.
- Textor J, van der Zander B, Gilthorpe MS, Liśkiewicz M, Ellison GT (2016). “Robust causal inference using directed acyclic graphs: the R package ‘dagitty’.” *International Journal of Epidemiology*, 45(6), 1887-1894. doi:10.1093/ije/dyw341 <https://doi.org/10.1093/ije/dyw341>.
- Wickham H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.
- Wickham H (2023). *forcats: Tools for Working with Categorical Variables (Factors)*. R package version 1.0.0, <https://CRAN.R-project.org/package=forcats>.
- Wickham H (2023). *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.5.1, <https://CRAN.R-project.org/package=stringr>.
- Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). “Welcome to the tidyverse.” *Journal of Open Source Software*, 4(43), 1686. doi:10.21105/joss.01686 <https://doi.org/10.21105/joss.01686>.
- Wickham H, François R, Henry L, Müller K, Vaughan D (2023). *dplyr: A Grammar of Data Manipulation*. R package version 1.1.4, <https://CRAN.R-project.org/package=dplyr>.
- Wickham H, Henry L (2023). *purrr: Functional Programming Tools*. R package version 1.0.2, <https://CRAN.R-project.org/package=purrr>.
- Wickham H, Hester J, Bryan J (2024). *readr: Read Rectangular Text Data*. R package version 2.1.5, <https://CRAN.R-project.org/package=readr>.
- Wickham H, Vaughan D, Girlich M (2024). *tidyr: Tidy Messy Data*. R package version 1.3.1, <https://CRAN.R-project.org/package=tidyr>.

References studies

- Adams, R.I., Leppänen, H., Karvonen, A.M., Jacobs, J., Borràs-Santos, A., Valkonen, M., Krop, E., Haverinen-Shaughnessy, U., Huttunen, K., Zock, J., Hyvärinen, A., Heederik, D., Pekkanen, J., Täubel, M., 2021. Microbial exposures in moisture-damaged schools and associations with respiratory symptoms in students: A multi-country environmental exposure study. *Indoor Air* 31, 1952–1966. <https://doi.org/10.1111/ina.12865>
- Cox, J., Stone, T., Ryan, P., Burkle, J., Jandarov, R., Mendell, M.J., Niemeier-Walsh, C., Reponen, T., 2022. Residential bacteria and fungi identified by high-throughput sequencing and childhood respiratory health. *Environmental Research* 204, 112377. <https://doi.org/10.1016/j.envres.2021.112377>
- Du, C., Li, B., Yu, W., Yao, R., Cai, J., Li, B., Yao, Y., Wang, Y., Chen, M., Essah, E., 2022. Characteristics of annual mold variations and association with childhood allergic symptoms/diseases via combining surveys and home visit measurements. *Indoor Air* 32, e13113. <https://doi.org/10.1111/ina.13113>
- Ferguson, K.D., McCann, M., Katikireddi, S.V., Thomson, H., Green, M.J., Smith, D.J., Lewsey, J.D., 2020. Evidence synthesis for constructing directed acyclic graphs (ESCDAGs): a novel and systematic method for building directed acyclic graphs. *International Journal of Epidemiology* 49, 322–329. <https://doi.org/10.1093/ije/dyz150>
- Fu, X., Du, B., Meng, Y., Li, Y., Zhu, X., Ou, Z., Zhang, M., Wen, H., Ma'pol, A., Hashim, J.H., Hashim, Z., Wieslander, G., Chen, Q., Jiang, J., Wang, J., Norbäck, D., Xia, Y., Chen, Q., Sun, Y., 2023. Associations between environmental characteristics, high-resolution indoor microbiome, metabolome and allergic and non-allergic rhinitis symptoms for junior high school students. *Environmental Science: Processes & Impacts* 25, 791–804. <https://doi.org/10.1039/D2EM00480A>
- Fu, X., Li, Y., Meng, Y., Yuan, Q., Zhang, Z., Wen, H., Deng, Y., Norbäck, D., Hu, Q., Zhang, X., Sun, Y., 2021a. Derived habitats of indoor microbes are associated with asthma symptoms in Chinese university dormitories. *Environmental Research* 194, 110501. <https://doi.org/10.1016/j.envres.2020.110501>
- Fu, X., Meng, Y., Li, Y., Zhu, X., Yuan, Q., Ma'pol, A., Hashim, J.H., Hashim, Z., Wieslander, G., Zheng, Y.-W., Lai, X.-X., Spangfort, M.D., Wu, J., Mu, P., Wang, J., Norbäck, D., Sun, Y., 2022. Associations between species-level indoor microbiome, environmental characteristics, and asthma in junior high schools of Terengganu, Malaysia. *Air Quality, Atmosphere & Health* 15, 1043–1055. <https://doi.org/10.1007/s11869-021-01080-0>
- Fu, X., Norbäck, D., Yuan, Q., Li, Y., Zhu, X., Hashim, J.H., Hashim, Z., Ali, F., Zheng, Y.-W., Lai, X.-X., Spangfort, M.D., Deng, Y., Sun, Y., 2020. Indoor microbiome, environmental characteristics and asthma among junior high school students in Johor Bahru, Malaysia. *Environment International* 138, 105664. <https://doi.org/10.1016/j.envint.2020.105664>
- Fu, X., Ou, Z., Zhang, M., Meng, Y., Li, Y., Wen, J., Hu, Q., Zhang, X., Norbäck, D., Deng, Y., Zhao, Z., Sun, Y., 2021b. Indoor bacterial, fungal and viral species and functional genes in urban and rural schools in Shanxi Province, China—association with asthma, rhinitis and rhinoconjunctivitis in high school students. *Microbiome* 9, 138. <https://doi.org/10.1186/>

- Gangneux, J.-P., Sassi, M., Lemire, P., Le Cann, P., 2020. Metagenomic Characterization of Indoor Dust Bacterial and Fungal Microbiota in Homes of Asthma and Non-asthma Patients Using Next Generation Sequencing. *Frontiers in Microbiology* 11, 1671. <https://doi.org/10.3389/fmicb.2020.01671>
- Hyytiäinen, H., Kirjavainen, P.V., Täubel, M., Tuoresmäki, P., Casas, L., Heinrich, J., Herberth, G., Standl, M., Renz, H., Piippo-Savolainen, E., Hyvärinen, A., Pekkanen, J., Karvonen, A.M., 2021. Microbial diversity in homes and the risk of allergic rhinitis and inhalant atopy in two European birth cohorts. *Environmental Research* 196, 110835. <https://doi.org/10.1016/j.envres.2021.110835>
- Karvonen, A.M., Kirjavainen, P.V., Täubel, M., Jayaprakash, B., Adams, R.I., Sordillo, J.E., Gold, D.R., Hyvärinen, A., Remes, S., Mutius, E. von, Pekkanen, J., 2019. Indoor bacterial microbiota and development of asthma by 10.5 years of age. *Journal of Allergy and Clinical Immunology* 144, 1402–1410. <https://doi.org/10.1016/j.jaci.2019.07.035>
- Ta, L.D.H., Tay, C.J.X., Lay, C., Sessions, P.F. de, Tan, C.P.T., Tay, M.J.Y., Lau, H.X., Zulkifli, A.B., Yap, G.C., Tham, E.H., Ho, E.X.P., Goh, A.E.N., Godfrey, K.M., Eriksson, J.G., Knol, J., Gluckman, P.D., Chong, Y.S., Chan, J.K.Y., Tan, K.H., Chong, K.W., Goh, S.H., Cheng, Z.R., Lee, B.W., Shek, L.P., Loo, E.X.L., 2021. Household environmental microbiota influences early-life eczema development. *Environmental Microbiology* 23, 7710–7722. <https://doi.org/10.1111/1462-2920.15684>
- Vandenborght, L.-E., Enaud, R., Urien, C., Coron, N., Girodet, P.-O., Ferreira, S., Berger, P., Delhaes, L., 2021. Type 2–high asthma is associated with a specific indoor mycobiome and microbiome. *Journal of Allergy and Clinical Immunology* 147, 1296–1305.e6. <https://doi.org/10.1016/j.jaci.2020.08.035>