Attribution of deaths to multiple causes

17. 09. 2024

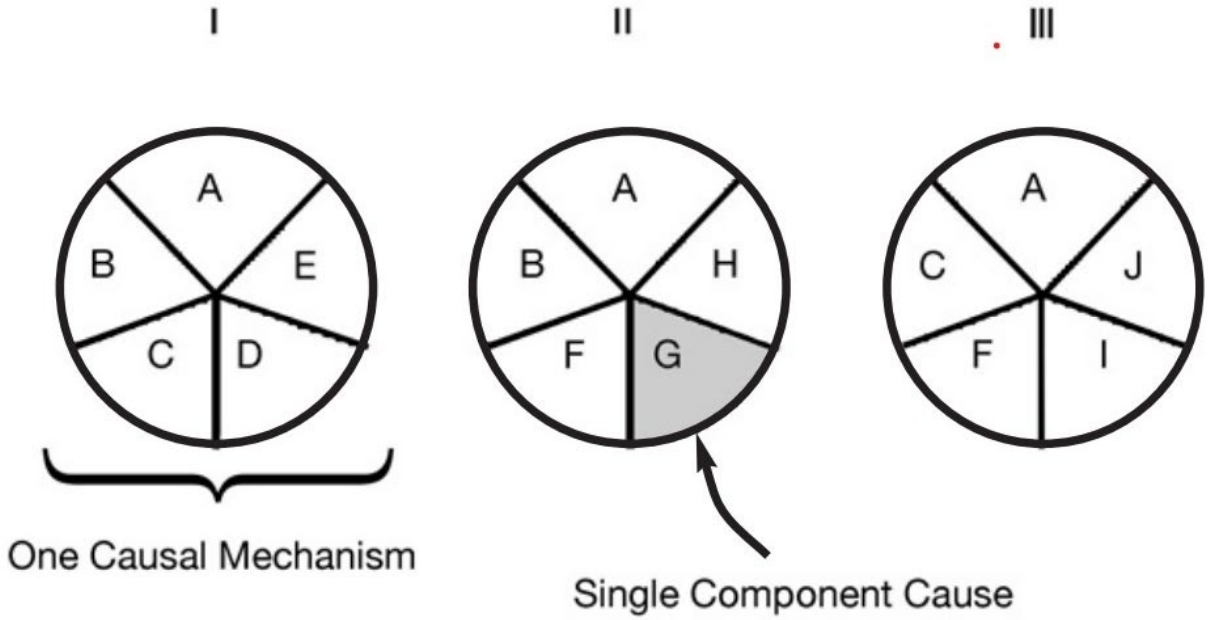
Elizabet Ukolova

The research and this conference paper were supported by the SCOR Foundation for Science through the funding for the "SCOR Chair in Mortality Research.

SCOR FOUNDATION FOR SCIENCE

Why?

Because current methods of attribution of deaths to multiple causes do not care about how much individual causes contribute to the actual death.



All conditions are equally important. → The most important is the condition, that actually contributes most to the death from officially coded underlying cause.

Outline

Causal pie models

- What are they good for?
- How do they work?
- How can they be used in multiple-cause of death analysis?

Causal pie models with mediator

- What are they good for?
- How do they work?
- How can they be used in multiple-cause of death analysis?

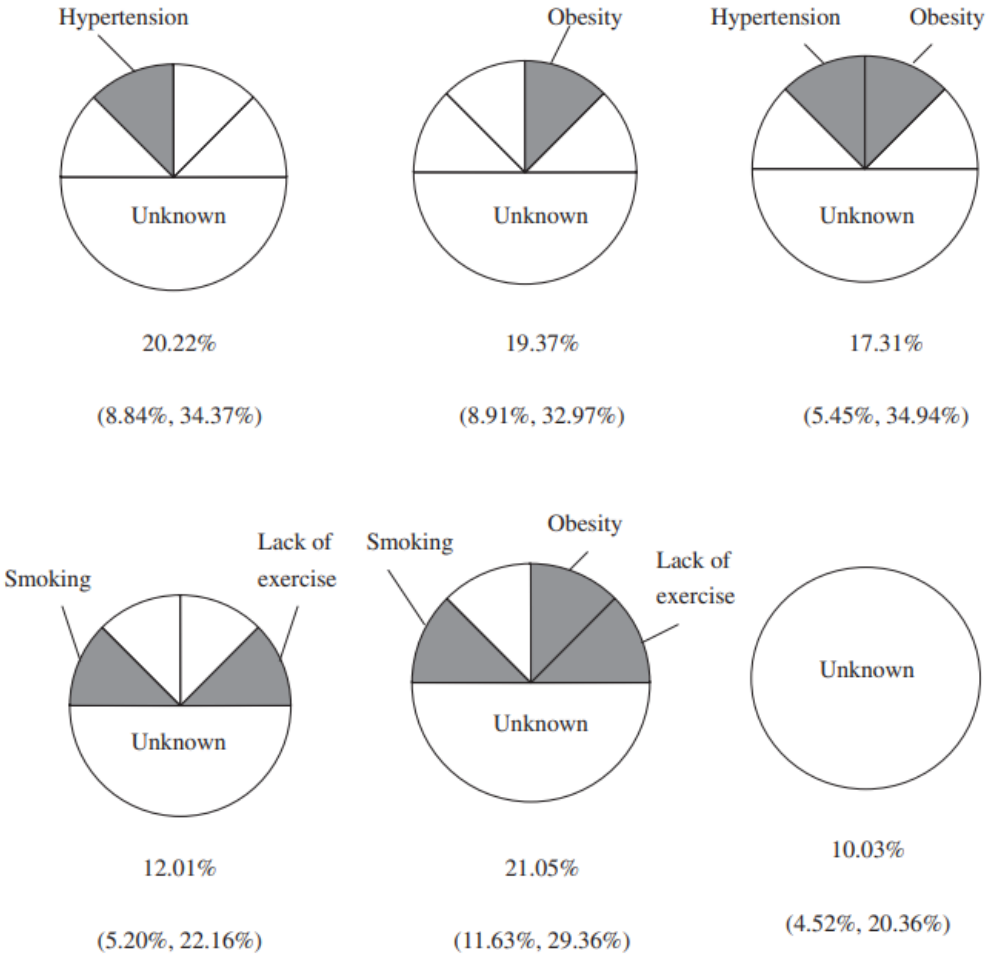
Causal pie models: What are they good for?

„Causal pie model helps to clarify the multifactorial and complex interactive nature in disease causation“ (Liao & Lee, 2010).

Example: smoking, hypertension, obesity and lack of exercise are risk factors for ischemic heart disease.

What is the proportion of subjects who developed the outcome thru a particular causal pie? (quantifying importance of some class of disease causation)

If you reduce some risk factor, what impact does it have on the outcome?



Causal pie models: How do they work?

1. Fit the positive linear odds model.
   - Assumption: Causal pie components can only increase the risk of the outcome and they act additively.
2. Use model coeffitients to calculate odds ratios.
3. Use odds ratios to calculate population attributable fraction.
4. Use population attributable fractions to calculate causal pie weights.

Causal pie models: How do they work?

Step 1

Variables	Regression coefficients	Standard errors	p (one-sided)
Intercept b_0	0.0006	0.0003	0.0223
Hypertension b_1	0.0025	0.0010	0.0064
Obesity b_2	0.0021	0.0008	0.0049
Smoking × lack of exercise b_3	0.0032	0.0014	0.0125
Hypertension × obesity b_4	0.0031	0.0017	0.0303
Smoking × obesity × lack of exercise b_5	0.0094	0.0028	0.0004

$$\frac{(b_0+b_1)/(1-(b_0+b_1))}{b_0/(1-b_0)}$$

$$\frac{(b_0+b_1+b_2+b_4)/(1-(b_0+b_1+b_2+b_4))}{b_0/(1-b_0)}$$

Step 2

	Risk factor profiles/target levels				Number of subjects (%)			
	Smoking	Hypertension	Obesity	Lack of exercise	Cases ($n_1 = 159$)	Controls ($n_0 = 26{,}813$)	ORs	PAFs (%)
No risk factor/eliminating all	0	0	0	0	0 (0.00)	1448 (5.40)	1.00	89.97
One risk factor/eliminating three risk factors	1	0	0	0	0 (0.00)	617 (2.30)	1.00	89.97
	0	1	0	0	4 (2.51)	617 (2.30)	5.17	69.75
	0	0	1	0	0 (0.00)	751 (2.80)	4.50	70.60
	0	0	0	1	4 (2.51)	4424 (16.50)	1.00	89.97
Two risk factors/eliminating two risk factors	1	1	0	0	0 (0.00)	188 (0.70)	5.17	69.75
	1	0	1	0	2 (1.26)	349 (1.30)	4.50	70.60
	1	0	0	1	5 (3.14)	1716 (6.40)	6.33	77.96
	0	1	1	0	5 (3.14)	965 (3.60)	13.83	33.07
	0	1	0	1	5 (3.14)	2520 (9.40)	5.17	69.75
	0	0	1	1	10 (6.29)	3137 (11.70)	4.50	70.60
Three risk factors/eliminating one risk factor	1	1	1	0	3 (1.89)	348 (1.30)	13.83	33.07
	1	1	0	1	7 (4.40)	697 (2.60)	10.50	57.74
	1	0	1	1	20 (12.58)	1501 (5.60)	25.50	37.54
	0	1	1	1	47 (29.56)	5497 (20.50)	13.83	33.07
Four risk factors/eliminating none	1	1	1	1	47 (29.56)	2038 (7.60)	34.83	0.00

Causal pie models: How do they work?

	Risk factor profiles/target levels				Number of subjects (%)			
	Smoking	Hypertension	Obesity	Lack of exercise	Cases ($n_1 = 159$)	Controls ($n_0 = 26{,}813$)	ORs	PAFs (%)
No risk factor/eliminating all	0	0	0	0	0 (0.00)	1448 (5.40)	1.00	89.97
One risk factor/eliminating three risk factors	1	0	0	0	0 (0.00)	617 (2.30)	1.00	89.97
	0	1	0	0	4 (2.51)	617 (2.30)	5.17	69.75
	0	0	1	0	0 (0.00)	751 (2.80)	4.50	70.60
	0	0	0	1	4 (2.51)	4424 (16.50)	1.00	89.97
Two risk factors/eliminating two risk factors	1	1	0	0	0 (0.00)	188 (0.70)	5.17	69.75
	1	0	1	0	2 (1.26)	349 (1.30)	4.50	70.60
	1	0	0	1	5 (3.14)	1716 (6.40)	6.33	77.96
	0	1	1	0	5 (3.14)	965 (3.60)	13.83	33.07
	0	1	0	1	5 (3.14)	2520 (9.40)	5.17	69.75
	0	0	1	1	10 (6.29)	3137 (11.70)	4.50	70.60
Three risk factors/eliminating one risk factor	1	1	1	0	3 (1.89)	348 (1.30)	13.83	33.07
	1	1	0	1	7 (4.40)	697 (2.60)	10.50	57.74
	1	0	1	1	20 (12.58)	1501 (5.60)	25.50	37.54
	0	1	1	1	47 (29.56)	5497 (20.50)	13.83	33.07
Four risk factors/eliminating none	1	1	1	1	47 (29.56)	2038 (7.60)	34.83	0.00

	Smoking	Hypertension	Obesity	Lack of exercise	Cases	Controls	OR	PAFs
	0	0	0	0	0,0000	0,0540	1,00	
	1	0	0	0	0,0000	0,0230	1,00	
	0	1	0	0	0,0251	0,0230	5,17	0,0202
	0	0	1	0	0,0000	0,0280	4,50	
	0	0	0	1	0,0251	0,1650	1,00	
	1	1	0	0	0,0000	0,0070	5,17	0,0000
	1	0	1	0	0,0126	0,0130	4,50	
	1	0	0	1	0,0314	0,0640	6,33	
	0	1	1	0	0,0314	0,0360	13,83	0,0212
	0	1	0	1	0,0314	0,0940	5,17	0,0253
	0	0	1	1	0,0629	0,1170	4,50	
	1	1	1	0	0,0189	0,0130	13,83	0,0128
	1	1	0	1	0,0440	0,0260	10,50	0,0175
Target scenario	**1**	**0**	**1**	**1**	0,1258	0,0560	25,50	
	0	1	1	1	0,2956	0,2050	13,83	0,1994
	1	1	1	1	0,2956	0,0760	34,83	0,0792
							Σ	**0,3756**

Step 3

$$PAF = \sum_{j=1}^{n} n \times \frac{OR_j - OR_j^*}{OR_j}$$

What proportion of cases would be eliminated, if hypertension is eliminated?

Causal pie models: How do they work?

1 ~ In this causal pie I have this risk factor

1 ~ I eliminate this

smoking	obesity	hyper	exercise
0	1	0	0
0	0	1	0
0	0	0	1
0	0	0	0
0	0	0	0
1	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1
1	1	0	0
1	0	1	0
1	0	0	1
0	1	1	0
0	1	0	1
0	0	1	1
1	1	1	0
1	1	0	1
1	0	1	1
0	1	1	1
1	1	1	1

0	1	0	0	0	1	1	1	0	0	0	1	1	1	0	1		
0	0	1	0	0	1	0	0	1	1	0	1	1	0	1	1		
0	0	0	1	0	0	1	0	1	0	1	1	0	1	1	1		
0	0	0	0	1	0	0	1	0	1	1	0	1	1	1	1		
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0		
1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0		
1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0		
1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0		
1	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0		
1	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0		
1	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0		
1	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0		
1	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0		
1	1	1	1	0	1	1	0	1	0	0	1	0	0	0	0		
1	1	1	0	1	1	0	1	0	1	0	0	1	0	0	0		
1	1	0	1	1	1	0	1	1	0	0	1	0	0	1	0	0	
1	0	1	1	1	0	1	1	1	0	0	0	1	1	1	0	0	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	

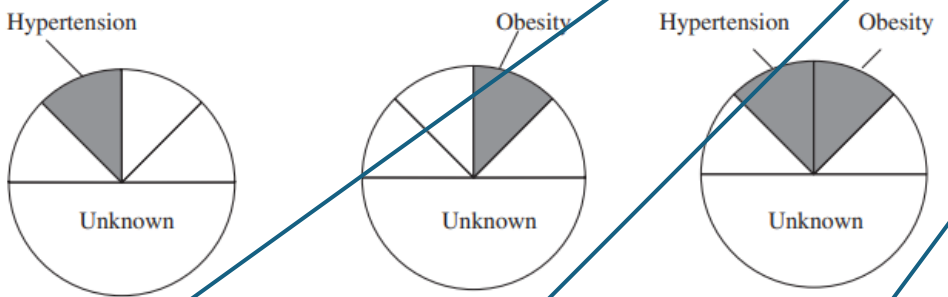
×

Step 4

W

0,1003	0,1003
0,1003	0,0000
0,3025	0,2022
0,2940	0,1937
0,1003	0,0000
0,3025	0,0000
0,2940	0,0000
0,2204	0,1201
0,6693	0,1731
0,3025	0,0000
0,2940	0,0000
0,6693	0,0000
0,4226	0,0000
0,6246	0,2105
0,6693	0,0000
1,0000	0,0001
Σ	1,00000

=



20.22% (8.84%, 34.37%)

19.37% (8.91%, 32.97%)

17.31% (5.45%, 34.94%)

$$W = A^{-1} \times (1 - F)$$

W... Is a vector of causal pie weights for each target level;

A... Is a matrix, those elements equal to 1 if causal pie denoted in the column is affected by the intervention denoted in the row, 0 otherwise.

F... Is a vector of PAFs.

Interpretation: The independent effect due to hypertension is in 20,22%, the independent effect due to obesity is 19,37% and the joint effect of both is in 17,31% of cases.

Causal pie models: How can they be used in MCD analysis?

Example: Death from COVID-19 in USA in 2021 happened on average with 4,3 additional causes of death. Which of them could be the most important one and what proportion of death developed thru the causal pie including this/these particular cause/s?

Causal pie models: How can they be used in MCD analysis?

	Respiratory	Hypertension	Diabetes	Ischemic HD	Dementia	OR	Cases	Controls	Cases %	Controls %	PAF	PAF %
Elim none	0	0	0	0	0	1,00	103468	894476	0,336	0,421	0,000	0,000
Elim 1	0	0	0	0	1	1,00	16287	245111	0,053	0,115	0,003	0,259
	0	0	0	1	0	1,00	11035	217828	0,036	0,102	0,003	0,252
	0	0	1	0	0	1,11	8557	57334	0,028	0,027	0,010	0,994
	0	1	0	0	0	1,00	10948	124307	0,036	0,058	0,004	0,431
	1	0	0	0	0	1,22	89310	175408	0,290	0,082	0,075	7,455
Elim 2	0	0	0	1	1	1,12	1530	26687	0,005	0,013	0,005	0,454
	0	0	1	0	1	1,11	1446	14007	0,005	0,007	0,013	1,291
	0	0	1	1	0	1,11	3026	40821	0,010	0,019	0,013	1,310
	0	1	0	0	1	1,11	3990	49947	0,013	0,023	0,005	0,536
	0	1	0	1	0	1,11	3434	63286	0,011	0,030	0,005	0,540
	0	1	1	0	0	1,12	7188	49339	0,023	0,023	0,014	1,433
	1	0	0	0	1	1,22	5091	18728	0,017	0,009	0,077	7,735
	1	0	0	1	0	1,22	6386	22259	0,021	0,010	0,077	7,733
	1	0	1	0	0	1,27	7402	9728	0,024	0,005	0,086	8,625
	1	1	0	0	0	1,22	8187	17948	0,027	0,008	0,079	7,897
Elim 3	0	0	1	1	1	1,14	331	4906	0,001	0,002	0,016	1,551
	0	1	0	1	1	1,15	868	13375	0,003	0,006	0,005	0,543
	0	1	1	0	1	1,14	1272	14117	0,004	0,007	0,016	1,551
	0	1	1	1	0	1,13	2331	33642	0,008	0,016	0,016	1,551
	1	0	0	1	1	1,28	457	1797	0,001	0,001	0,080	7,959
	1	0	1	0	1	1,28	564	1254	0,002	0,001	0,090	8,959
	1	0	1	1	0	1,28	1819	3756	0,006	0,002	0,090	9,014
	1	1	0	0	1	1,27	1439	3698	0,005	0,002	0,080	8,013
	1	1	0	1	0	1,26	2179	5895	0,007	0,003	0,080	8,009
	1	1	1	0	0	1,30	6084	7145	0,020	0,003	0,092	9,160
Elim 4	0	1	1	1	1	1,20	306	4657	0,001	0,002	0,016	1,551
	1	0	1	1	1	1,34	116	323	0,000	0,000	0,093	9,295
	1	1	0	1	1	1,36	269	884	0,001	0,000	0,080	8,013
	1	1	1	0	1	1,35	525	1104	0,002	0,001	0,093	9,295
	1	1	1	1	0	1,35	1589	3056	0,005	0,001	0,093	9,295
Elim 5	1	1	1	1	1	1,52	137	318	0,000	0,000	0,093	9,295

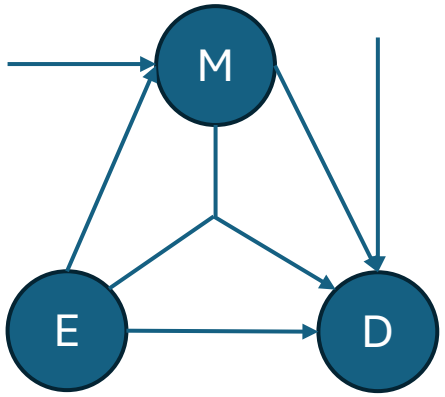
Preventing other respiratory diseases would lead to 7,5% of reduction in COVID-19 deaths. Joint effect of other diseases is negliable, however interactions with diabetes contribute more then 1% as well.

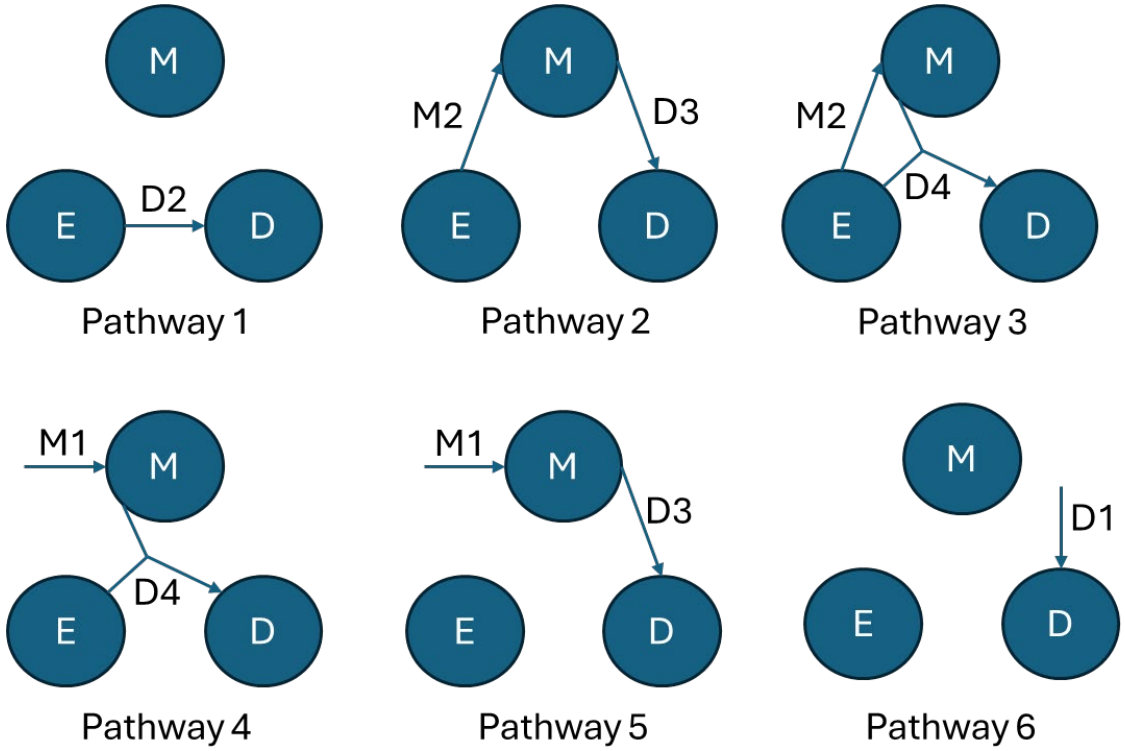
CPW_{00000} is 0,907, which means that 90% of COVID-19 deaths would not be prevented even if eliminating all of these diseases. Most frequent causes associated with diabetes were not that important in contributing to COVID-19 deaths. Among the rest, other respiratory disease and diabetes are the most important multiple causes of death.

Causal pie models with mediator: What are they good for?

Relations between entities (for example exposure, outcome and mediator) are modelled as mechanisms/pathways. Such models are based on structural equation modelling.

The model:

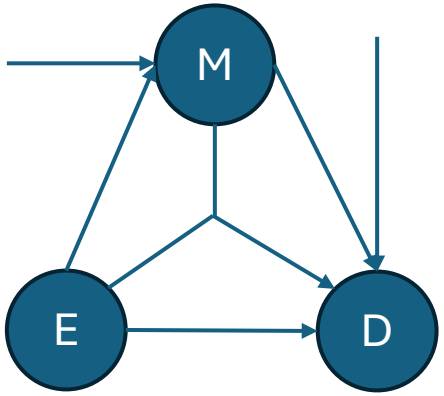




Aim: To calculate, which of these pathways is most likely followed.

Causal pie models with mediator: How do they work?

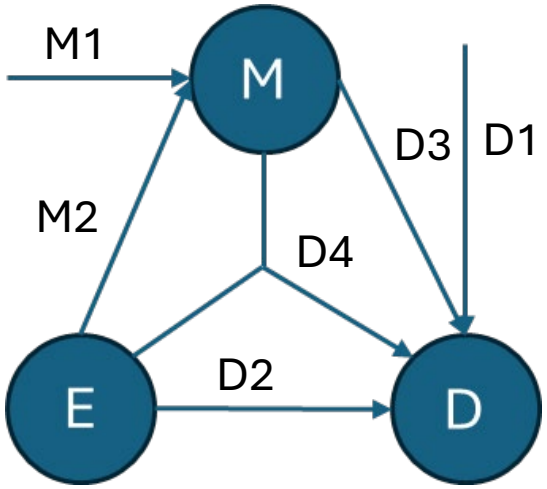
1. Estimation of the transition rates in the direct acyclic graphs (parameters of the model).
2. Using transition rates to calculate attributable fractions (population, or among the exposed etc.).



Mediator status and disease status	Exposure status	
	E=0	E=I
M=0, D=0	m_1	n_1
M=I, D=0	m_2	n_2
M=0, D=I	m_3	n_3
M=I, D=I	m_4	n_4
Total	$m = m_1 + m_2 + m_3 + m_4$	$n = n_1 + n_2 + n_3 + n_4$

Causal pie models with mediator: How do they work?

1. Estimations of the transition rates (parameters) in the direct acyclic graphs.



MLE estimation → maximize following expression:

$$l = \sum_{i=1}^{4} m_i \times \log(p_{0i}) + \sum_{i=1}^{4} n_j \times \log(p_{1j})$$

Where p_{0i} and p_{1j} jsou transition probabilities for either exposed or unexposed to *i*/*j*.
They are equal to the set of equations.
Where λ are initial incidence rates, for example λ_{M1} is equal to proportion of unexposed subject who aquire the mediator.

$$p_{01} = e^{-(\lambda_{M1}+\lambda_{D1})\times T}$$

$$p_{02} = \frac{\lambda_{M1}}{\lambda_{D3}-\lambda_{M1}} \times \left[e^{-(\lambda_{M1}+\lambda_{D1})\times T} - e^{-(\lambda_{D1}+\lambda_{D3})\times T}\right]$$

$$p_{03} = \frac{\lambda_{D1}}{\lambda_{M1}+\lambda_{D1}} \times \left[1 - e^{-(\lambda_{M1}+\lambda_{D1})\times T}\right]$$

$$p_{04} = 1 - p_{01} - p_{02} - p_{03}$$

$$p_{11} = e^{-(\lambda_{M1}+\lambda_{M2}+\lambda_{D1}+\lambda_{D2})\times T}$$

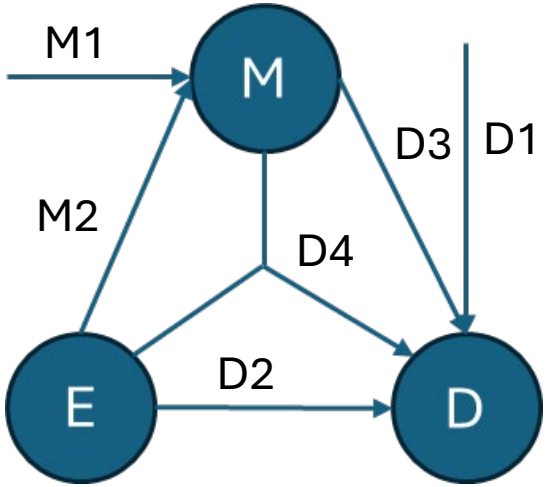
$$p_{12} = \frac{\lambda_{M1}+\lambda_{M2}}{\lambda_{D3}+\lambda_{D4}-(\lambda_{M1}+\lambda_{M2})} \times \left[e^{-(\lambda_{M1}+\lambda_{M2}+\lambda_{D1}+\lambda_{D2})\times T} - e^{-(\lambda_{D1}+\lambda_{D2}+\lambda_{D3}+\lambda_{D4})\times T}\right]$$

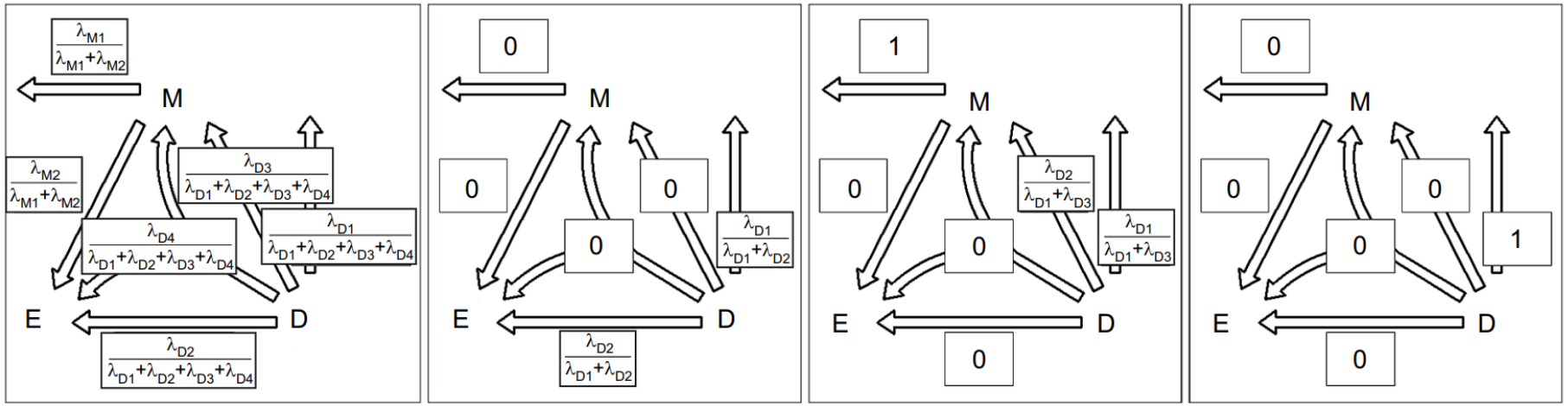
$$p_{13} = \frac{\lambda_{D1}+\lambda_{D2}}{\lambda_{D1}+\lambda_{D2}+\lambda_{M1}+\lambda_{M2}} \times \left[1 - e^{-(\lambda_{M1}+\lambda_{M2}+\lambda_{D1}+\lambda_{D2})\times T}\right]$$

$$p_{14} = 1 - p_{11} - p_{12} - p_{13}$$

Causal pie models with mediator: How do they work?

2. Using transition rates to calculate attributable fractions (population, or among the exposed etc.).





E=1, M=1 | E=1, M=0 | E=0, M=1 | E=0, M=0

Mediator status and disease status	Exposure status	
	E=0	E=1
M=0, D=0	m_1	n_1
M=1, D=0	m_2	n_2
M=0, D=1	m_3	n_3
M=1, D=1	m_4	n_4
Total	$m = m_1 + m_2 + m_3 + m_4$	$n = n_1 + n_2 + n_3 + n_4$

Causal pie models with mediator: How can they be used in MCD analysis?

Example: Death certifiers were „forced“ to record other conditions predominantly into the second part of the death certificate, since the ACS (Automated Coding System) does not select underlying causes of death predominantly from there. What pathway between leading triads of causes of death in the USA in 2021 could be the most probable one?

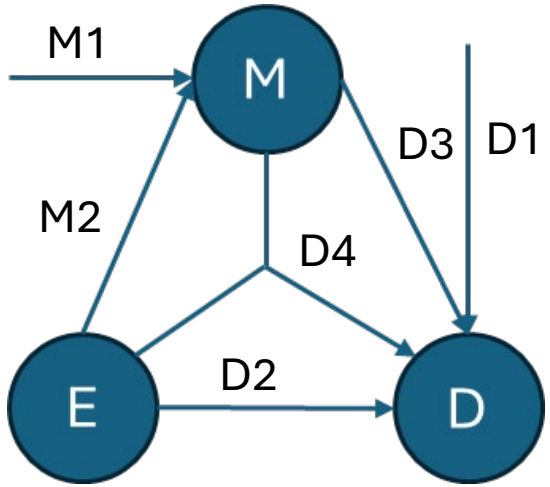
Causal pie models with mediator: How can they be used in MCD analysis?

Each cause of death was tried in each position.

Top triads: COVID-19 + diabetes + respiratory disease

COVID-19 + ischemic heart disease + respiratory disease

	COVID-19	Respiratory	Diabetes/HD
Version 1	E	D	M
Version 2	M	D	E
Version 3	D	E	M
Version 4	M	E	D
Version 5	D	M	E
Version 6	E	M	C

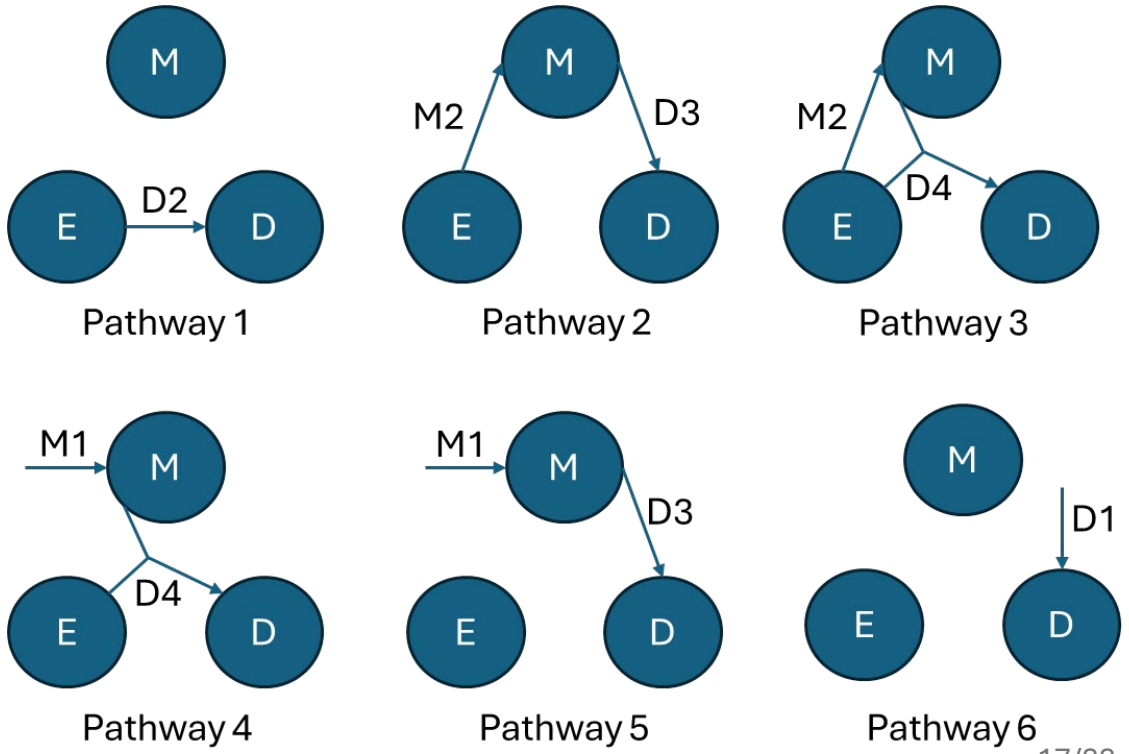


Causal pie models with mediator: How can they be used in MCD analysis?

	I	II	III	IV	V	VI	max	sex	age	Triad	verze	max
F_1_999_40_88_1	0.30	0.01	0.14	0.35	0.03	0.16	IV	F	65-84	COVID-19 + diabetes + respiratory disease	1	0.35
F_1_999_40_88_2	0.00	0.08	0.00	0.02	0.76	0.14	V	F			2	0.76
F_1_999_40_88_3	0.20	0.00	0.00	0.60	0.11	0.09	IV	F			3	0.60
F_1_999_40_88_4	0.01	0.00	0.00	0.40	0.51	0.08	V	F			4	0.51
F_1_999_40_88_5	0.00	0.41	0.00	0.00	0.19	0.41	II	F			5	0.41
F_1_999_40_88_6	0.00	0.10	0.27	0.13	0.05	0.45	VI	F			6	0.45
F_1_999_65_88_1	0.30	0.00	0.04	0.45	0.05	0.16	IV	F		COVID-19 + IHD + respiratory disease	1	0.45
F_1_999_65_88_2	0.00	0.00	0.00	0.04	0.82	0.14	V	F			2	0.82
F_1_999_65_88_3	0.25	0.00	0.00	0.55	0.09	0.12	IV	F			3	0.55
F_1_999_65_88_4	0.00	0.00	0.00	0.24	0.66	0.10	V	F			4	0.66
F_1_999_65_88_5	0.00	0.33	0.00	0.00	0.15	0.52	VI	F			5	0.52
F_1_999_65_88_6	0.00	0.16	0.14	0.07	0.08	0.55	VI	F			6	0.55
F_2_999_40_88_1	0.31	0.02	0.14	0.37	0.05	0.11	IV	F	85+	COVID-19 + diabetes + respiratory disease	1	0.37
F_2_999_40_88_2	0.00	0.10	0.02	0.12	0.66	0.10	V	F			2	0.66
F_2_999_40_88_3	0.24	0.00	0.02	0.57	0.09	0.08	IV	F			3	0.57
F_2_999_40_88_4	0.00	0.00	0.00	0.34	0.58	0.07	V	F			4	0.58
F_2_999_40_88_5	0.00	0.37	0.09	0.03	0.12	0.39	VI	F			5	0.39
F_2_999_40_88_6	0.00	0.20	0.24	0.08	0.07	0.40	VI	F			6	0.40
F_2_999_65_88_1	0.33	0.01	0.05	0.43	0.06	0.12	IV	F		COVID-19 + IHD + respiratory disease	1	0.43
F_2_999_65_88_2	0.00	0.00	0.00	0.05	0.84	0.11	V	F			2	0.84
F_2_999_65_88_3	0.29	0.00	0.00	0.53	0.08	0.09	IV	F			3	0.53
F_2_999_65_88_4	0.00	0.00	0.00	0.19	0.72	0.09	V	F			4	0.72
F_2_999_65_88_5	0.00	0.36	0.01	0.00	0.11	0.51	VI	F			5	0.51
F_2_999_65_88_6	0.00	0.24	0.11	0.04	0.08	0.52	VI	F			6	0.52

Mostly, the most probable pathway to death is Version 2, causal pie V.

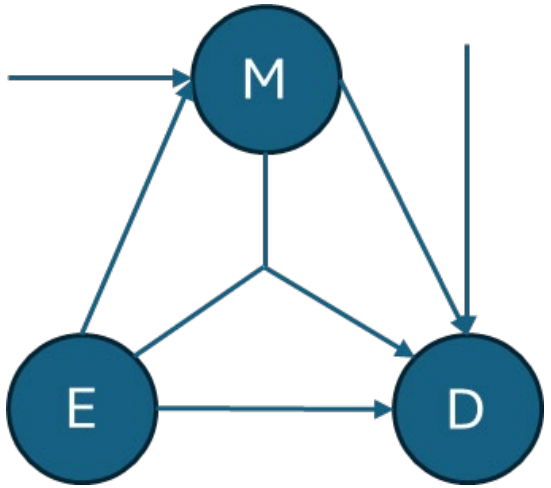
	COVID-19	Respiratory	Diabetes/HD
Version 2	M	D	E
Version 3	D	E	M
Version 4	M	E	D



Causal pie models with mediator: How can they be used in MCD analysis?

Example: In Czechia, 90% of death are due to combinations of causes, most of them are cardiovascular diseases. What role plays contributory cause of death in the relationship between underlying cause and its consequence?

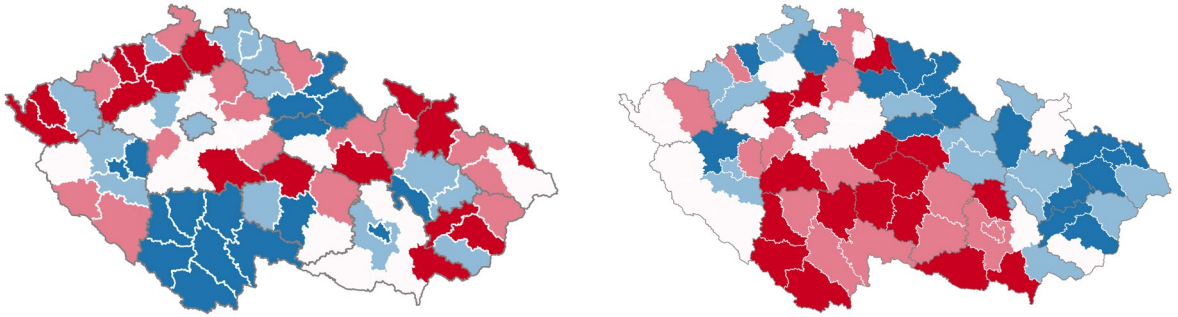
Contributory cause of death



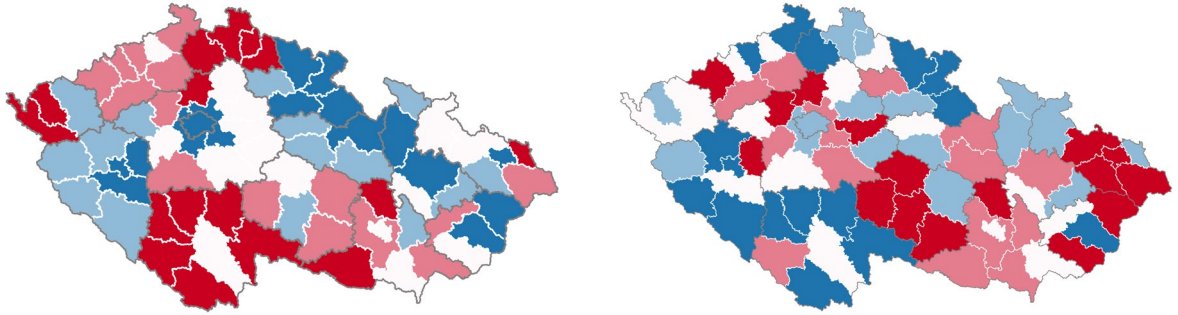
Underlying cause of death

Its consequence recorded on the death certificate

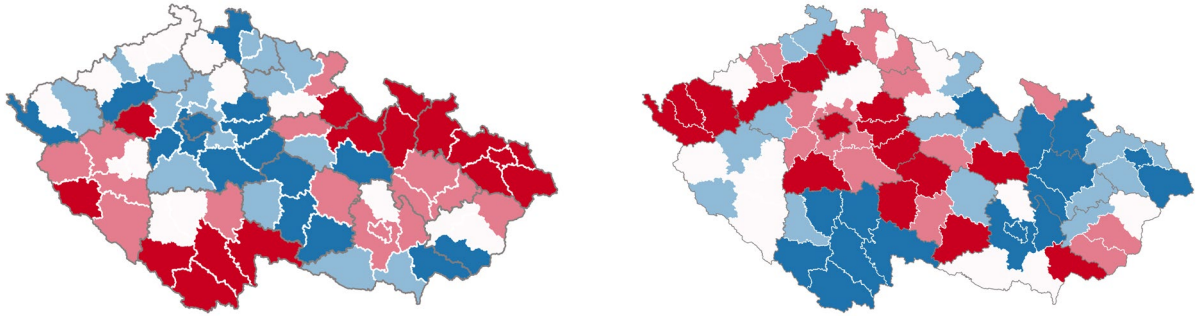
I25: Chronic ischaemic heart disease



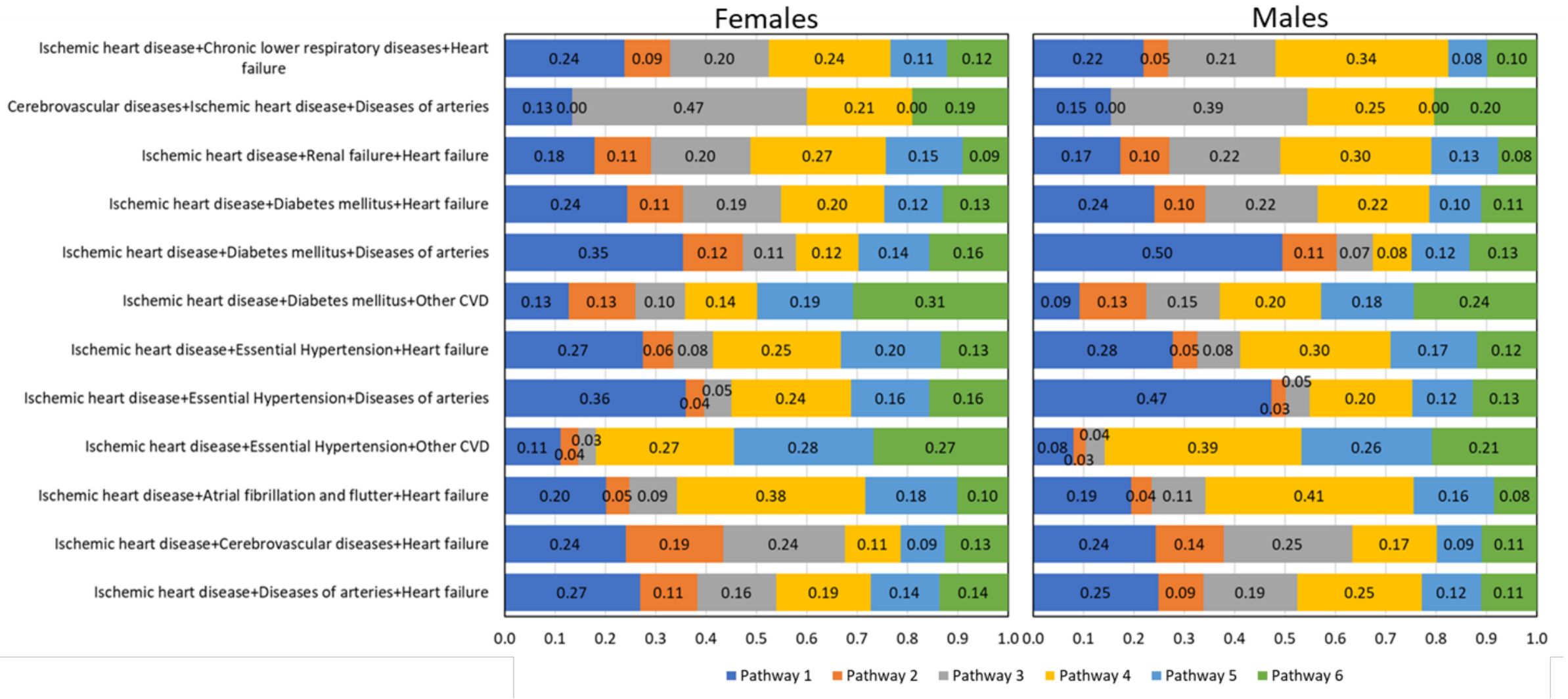
I21: Acute myocardial infarction



I50: Heart failure

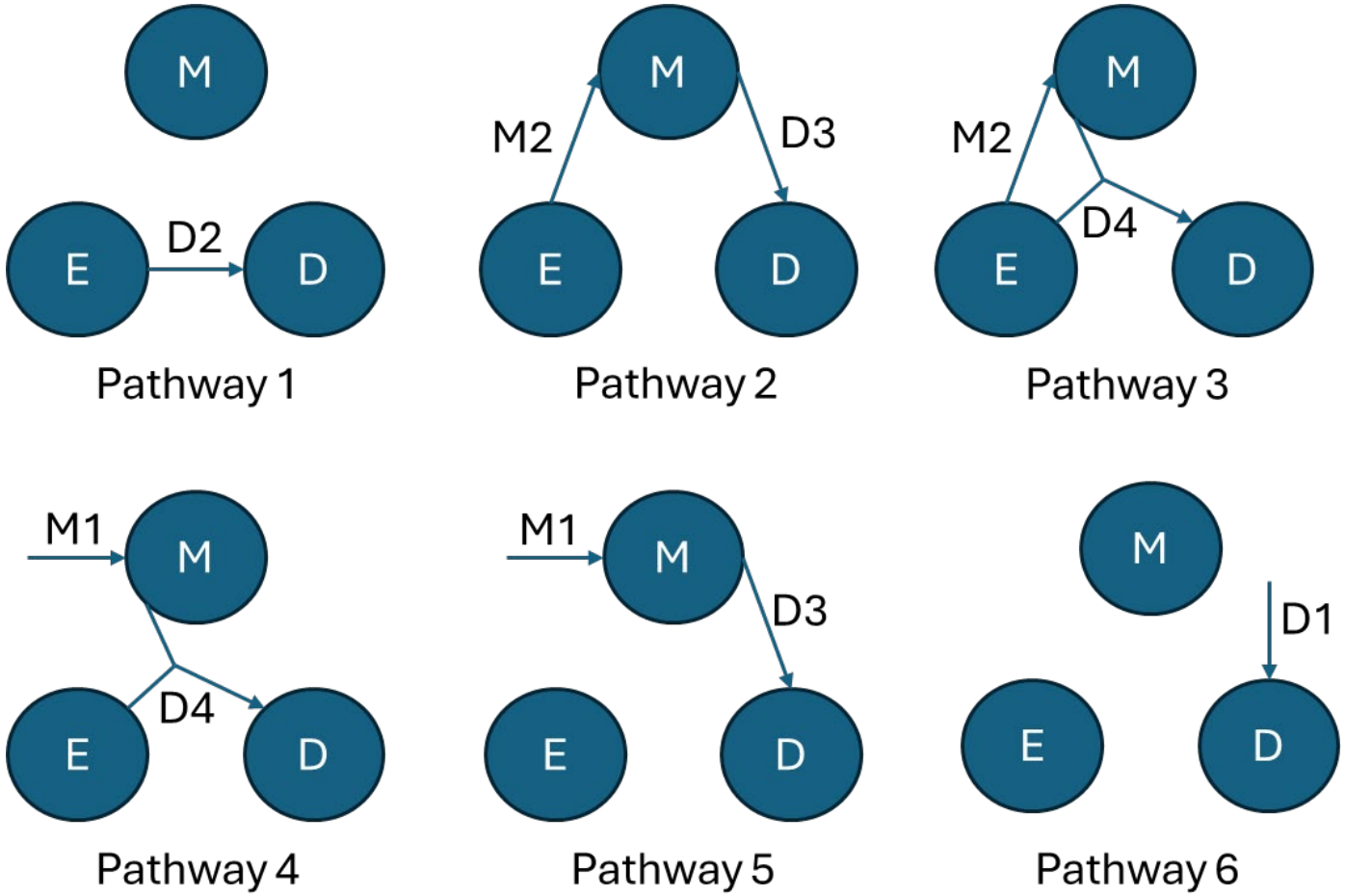


Causal pie models with mediator: How can they be used in MCD analysis?



We classified pathways to death based on 3 aspects: (i) whether the relationship between the underlying and immediate cause is influenced by the contributory cause, (ii) whether the immediate cause results from the underlying cause, and (iii) whether there is an interaction between contributory and underlying causes in their influence on the immediate cause.

Causal pie models with mediator: How can they be used in MCD analysis?



We classified pathways to death based on 3 aspects: (i) whether the relationship between the underlying and immediate cause is influenced by the contributory cause, (ii) whether the immediate cause results from the underlying cause, and (iii) whether there is an interaction between contributory and underlying causes in their influence on the immediate cause.

Limitations

Is it „OK“ to force the regression coefficients to be strictly positive?

Is it „OK“ to include only aditivity?

Is it „OK“ to omit other factors (age)? (this can be solved easily, of course)

Is it „OK“ to consider that **all** cause of death records could really contribute to death?

Is it „OK“ not to have the true „Exposure“ in the causal pies? (~ people who are actually alive)

Is it „OK“ to omit certain edges in directed acyclic graphs?

And many others...

Limitations

Is it „OK“ to force the regression coefficients to be strictly positive?

Is it „OK“ to include only aditivity?

Is it „OK“ to omit other factors (age)? (this can be solved easily, of course)

Is it „OK“ to consider that **all** cause of death records could really contribute to death?

Is it „OK“ not to have the true „Exposure“ in the causal pies? (~ people who are actually alive)

Is it „OK“ to omit certain edges in directed acyclic graphs?

And many others...

Conclusion (and „why shoud we care window“)

We tried to show, how could deaths be attributed to multiple causes based on how important they are in the lethal process. Next, we tried to use models to compute probabilities of different pathways to death.

no unified rules for selection of immediate and contributory conditions → no official statistics with them → understand the pathway and interaction → step towards such statistics → better understanding of lethal process

Thank you