

Univerzita Karlova
Přírodovědecká fakulta

Studijní program: Demografie

Studijní obor: Demografie



Elizaveta Ukolova

Analýza vztahů mezi vícečetnými příčinami smrti aplikací metod analýzy sítí

Diplomová práce

Vedoucí: RNDr. Boris Burcin, Ph.D.

Praha, 2022

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 22. 4. 2022

.....

Elizaveta Ukolova

Poděkování

Děkuji panu RNDr. Borisi Burcinovi, Ph.D. za investovaný čas, konstruktivní a snad i upřímnou zpětnou vazbu, úsilí vynaložené na hledání cest a pro mne velmi hodnotné rady nejen ohledně zpracovávání této práce.

Děkuji RNDr. Markétě Majerové, Ph.D. za poskytnutí materiálů pro výpočet CDAI, dále také odpovědi na devět otázek ohledně vícečetných příčin smrti a rekomandaci čtyřech článků týkajících se této problematiky.

Děkuji RNDr. Kláře Hulíkové Tesárkové, Ph.D. za udělení svobody při zpracovávání úkolů z předmětů Demografické aplikace SAS I a Základy ekonometrie, díky čemuž se mi naskytla velmi cenná příležitost se doptat na některé věci z programování, které je fundamentem této práce.

Děkuji své milé sestřičce Már za přečtení syré verze tohoto textu a poskytnutí některých korektur.

Děkuji svému otci, Sergeji Ukolovi, za jeho neutuchající podporu a za to, že je mi svým pracovním nasazením inspirací.

Abstrakt

Práce je zaměřena na aplikaci metod analýzy sítí při zkoumání vztahů mezi vícečetnými příčinami smrti. V důsledku prodloužení průměrné délky života a rostoucího podílu osob umírajících na chronická onemocnění ve vyspělých zemích může být zkoumání vícečetných příčin smrti relevantním tématem, protože je díky tomuto přístupu možné analyzovat i okolnosti umírání, nejen jedinou příčinu iniciující celý chorobný řetězec vedoucí ke smrti. Smyslem práce je tedy modelování systému příčin smrti, v němž je snahou podchytit komplexitu vztahů mezi příčinami smrti uváděnými na hlášení o úmrtí. Jedna síť je tvořena pro prvou část hlášení o úmrtí, kde je ve vztazích mezi příčinami smrti rozlišeno, která ze dvou příčin tvořících dvojici byla příčinou a která důsledkem v chorobném řetězci vedoucím ke smrti. Druhá síť je konstruována pro druhou část hlášení o smrti, jež je od první sítě odlišná v tom, že v ní není rozlišována orientace vazby. Takto utvořené systémy jsou analyzovány za použití rozličných měr centralit, rozkladu systému na podgrafy a zkoumáním sousedů vybraných chorob. Součástí práce je i explorační analýza dat o vícečetných příčinách smrti, od níž se odvíjí pravidla pro vymezení obou systémů.

Díky aplikaci metod analýzy sítí na data první části hlášení o úmrtí bylo zjištěno, že choroby oběhové, trávicí, dýchací soustavy i některé novotvary se váží na nemoci z těchto kapitol Mezinárodní klasifikace nemocí, jako jsou ony samy. Naopak nemoci endokrinní, výživy a přeměny látek, nervové soustavy a poruchy chování se váží spíše vně kapitol, kam samy náležejí. Dalším poznatkem bylo, že v síti žen je nalézáno více podskupin nemocí než v síti mužů, která se také vyznačovala tím, že v ní uzly centrální měly spíše menší dosah, tj. vyšší lokální centralitu. Síť mužů i žen se více než v 70 % vazeb shodovaly a rozdíly mezi ženami a muži tkvěly hlavně ve vnějších příčinách smrti, některých novotvarech, chorobách krve a krvetvorných orgánů, žlučníku, pohybového aparátu a některých nemocí nervové soustavy. Nejdůležitější uzly v orientovaných sítích se většinou shodovaly s těmi, které se na hlášení o úmrtí vyskytovaly nejčastěji, ale záleželo na tom, která metrika byla pro vyjádření „důležitosti“ uzlu v systému použita.

V sítích tvořených z druhé části hlášení se systémy za ženy vyznačovaly vyššími průměrnými stupni, dále byly zjištěny choroby charakteristické tím, že s jinými nejběžnějšími nesdílely mnoho společných sousedů, a také bylo poukázáno na méně časté příčiny smrti, které se ale shodovaly ve většině chorob, které k nim bývají zapisovány na hlášení o úmrtí.

Klíčová slova: vícečetné příčiny smrti, analýza sítí, závislost, hlášení o úmrtí, základní příčina smrti, přispívající příčiny smrti

Počet znaků: 198 765

Abstract

The aim of the thesis is the application of network analysis methods in the study of relationships between multiple causes of death. As a result of increasing life expectancy and of the growing proportion of people dying from chronic diseases in developed countries, the analysis of multiple causes of death may be a relevant topic, as this approach makes it possible to analyse the circumstances of dying, not just the single cause initiating the entire disease chain leading to death. The purpose of this work is to construct the system of causes of death, in which the complexity of the relationships between the causes of death reported on the death certificate would be captured. First network is created from the first part of the death certificate, and we distinguish between the “causing” condition and the consequent one in a pair of multiple causes of death. The second network is designed for the second part of the death certificate, and in this network, we do not differentiate the direction of the link. Networks are then analysed with various centrality measures and community detection, and we also examine neighbours of selected diseases. Essential part of the thesis is the explanatory analysis of multiple causes of death, because it determines methods later used in network analysis.

It was found that diseases of the circulatory, digestive, respiratory tracts as well as some neoplasms are more closely linked to diseases from the same categories. Conversely, metabolic, and behavioural disorders and the nervous system diseases are more often linked to diseases from distinct chapters of International Classification of Diseases. Another finding was that more communities were detected in networks for females than in network for males, which was also characterised by higher local centrality of the diseases. More than 70 % of links are same in networks for women and men, the differences between them lay mainly in external causes of death, neoplasms, blood and hematopoietic, gallbladder and musculoskeletal diseases and some diseases of the nervous system. The most important nodes in directed networks mostly coincided with those that occurred most frequently on death certificates, but it depended on which metric was used to determine the “importance” of the node in the system.

The networks from second part of death certificate for females were denser than those for males. In these networks diseases with “specific” neighbours were spotted. Diseases, which are not the most frequent ones, but are very alike in terms of links they have, were detected too.

Keywords: multiple causes of death, network analysis, death certificate, underlying cause of death, contributory causes of death

OBSAH

<i>Seznam tabulek</i>	8
<i>Seznam obrázků</i>	9
<i>Seznam zkratek</i>	11
1 Úvod	12
2 Zdroj dat	14
2.1 Hlášení o úmrtí	14
2.2 Analýzy vícečetných příčin smrti	18
3 Základní explorační analýza vícečetných příčin smrti	20
3.1 Průměrné počty zapsaných příčin smrti	20
3.2 Struktura zemřelých podle počtu kódů na hlášení o úmrtí	22
3.3 Místo vyplnění základní příčiny smrti.....	23
3.4 Nejčastěji zapsané kódy	24
3.5 Jiné a nepřesně určené příčiny smrti	27
3.6 Základní příčiny smrti podle průměrného počtu zapsaných příčin smrti	30
4 Použité metody a ukazatele	33
4.1 Vymezení orientované sítě.....	33
4.2 Vymezení neorientované sítě	38
4.3 Síťové ukazatele	42
5 Validace postupů při vymezení systému	48
5.1 Opodstatnění využití pravidel pro vymezení orientované sítě	48
5.2 Základní statistiky komorbidit	51
6 Výzkumné otázky	54
7 Výsledky analýzy orientované sítě příčin smrti	55
7.1 Jaký je systém příčin smrti?	55

7.2	Jaká je modularita grafů a jaké nemoci společně utvářejí komunity?.....	59
7.2.1	Novotvary.....	60
7.2.2	Sepse a další susedé.....	62
7.2.3	Kardiovaskulární choroby.....	64
7.2.4	Nemoci trávicí soustavy.....	67
7.2.5	Respirační nemoci.....	69
7.2.6	Choroby mozku a duševní choroby.....	70
7.2.7	Vnější příčiny smrti.....	72
7.2.8	Ostatní podgrafy.....	76
7.3	Kdo jsou susedé smrti?.....	78
7.4	Jaké role hrají uzly v síti?.....	80
7.5	Kdo jsou autority a hubs?.....	85
7.6	Jaké vazby se nevyskytují u opačného pohlaví?.....	91
8	Výsledky analýzy neorientované sítě.....	94
8.1	Standardní bimodální graf.....	95
8.2	Bimodální graf shodného susedstva.....	97
9	Shrnutí, diskuze a závěr.....	101
9.1	Shrnutí.....	101
9.1.1	Orientovaná síť.....	101
9.1.2	Neorientovaná síť.....	103
9.2	Diskuze a závěr.....	103
	<i>Seznam zdrojů.....</i>	<i>106</i>
	<i>Seznam příloh [USB].....</i>	<i>117</i>

Seznam tabulek

Tab. 1:	Nejčastěji uváděné kódy podle částí hlášení a v proměnné „coded UCD“, ženy, USA, 2018	25
Tab. 2:	Nejčastěji uváděné kódy podle částí hlášení a v proměnné „coded UCD“, muži, USA, 2018.....	26
Tab. 3:	Nejčastější jiné a nepřesně určené UCD, obě pohlaví dohromady, USA, 2018	28
Tab. 4:	Vybrané kódy s nejvyšším průměrným počtem zapsaných příčin smrti, ženy, USA, 2018	31
Tab. 5:	Vybrané kódy s nejvyšším průměrným počtem zapsaných příčin smrti, muži, USA, 2018.....	31
Tab. 6:	Vybrané kódy s nejnižším průměrným počtem zapsaných příčin smrti, ženy, USA, 2018	32
Tab. 7:	Vybrané kódy s nejnižším průměrným počtem zapsaných příčin smrti, muži, USA, 2018.....	32
Tab. 8:	Značení subpopulací podle toho, jaké nesou znaky	35
Tab. 9:	Studie, na něž se bylo orientováno při rozmyslech o vymezení systému příčin smrti z části 1 hlášení o úmrtí	49
Tab. 10:	Množiny vazeb podle signifikance v mírách užitých pro vymezení systému.....	51
Tab. 11:	Sousedé smrti	78
Tab. 12:	Korelační koeficienty mezi použitými metrikami, ženy a muži, USA, 2018	86
Tab. 13:	Uzly s nejvyšším stupněm v síti s vazbami, které nemají ženy	91
Tab. 14:	Uzly s nejvyšším stupněm v síti s vazbami, které nemají muži	92
Tab. 15:	Ilustrace z bimodálního grafu shodného sousedstva	95
Tab. 16:	Průměrné stupně ve standardních bimodálních grafech, ženy a muži, USA, 2018.....	95
Tab. 17:	Choroby s nejvyššími stupni v bipartitních grafech shodného sousedstva	99

Seznam obrázků

Obr. 1:	Část 1 a část 2 hlášení o úmrtí pro zaznamenání příčin smrti.....	15
Obr. 2:	Průměrný počet uváděných příčin smrti na HOÚ v závislosti na věku zemřelého, USA, 2018.....	21
Obr. 3:	Věková struktura zemřelých podle počtu zapsaných příčin smrti, USA, 2018.....	22
Obr. 4:	Zemřelí podle počtu zapsaných příčin smrti, celá populace a populace zemřelá s jinými a nepřesně určenými základními příčinami smrti, USA, 2018.....	29
Obr. 5:	Vývojový diagram pro orientovanou síť příčin smrti, USA, 2018.....	37
Obr. 6:	Schéma bipartitního grafu.....	38
Obr. 7:	Maticový počet na úpravu algoritmu pro produkci dat na bimodální síť.....	41
Obr. 8:	Typologie rolí podle stupně uzlu a mezilehlosti.....	44
Obr. 9:	Hrana s vysokou mezilehlostí.....	46
Obr. 10a:	Základní statistiky komorbidit, orientovaná síť, ženy, USA, 2018.....	52
Obr. 10b:	Základní statistiky komorbidit, orientovaná síť, ženy, USA, 2018.....	53
Obr. 11:	Zastoupení rozmezí kódů podle jejich prvních písmen v síti orientované, USA, 2018.....	58
Obr. 12:	Vztah prevalence uzlů a stupňů uzlů.....	59
Obr. 13:	Podgrafy novotvarů, muži (A), ženy (B a C), USA, 2018.....	60
Obr. 14:	Podgrafy sepse a další susedé, muži (A), ženy (B), USA, 2018.....	62
Obr. 15:	Podgrafy se srdečními chorobami, muži (A), ženy (B a C), USA, 2018.....	64
Obr. 16:	Podgrafy s chorobami trávicího ústrojí, muži (A), ženy (B a C), USA, 2018.....	67
Obr. 17:	Podgrafy s repiračními nemocemi, muži (A), ženy (B), USA, 2018.....	69
Obr. 18:	Podgrafy s demencemi, muži (A), ženy (B), USA, 2018.....	71
Obr. 19:	Podgrafy s vnějšími příčinami smrti, muži (A), ženy (B), USA, 2018.....	73
Obr. 20:	Podgrafy sdružující převážně otravy, muži (A), ženy (B), USA, 2018.....	74
Obr. 21:	Ostatní podgrafy s vnějšími příčinami, muži (A a C), ženy (B a D), USA, 2018.....	76

Obr. 22:	Role uzlů v síti z hlediska stupně uzlu a jeho mezilehlosti, ženy, USA, 2018	80
Obr. 23:	Srovnání rolí uzlů u mužů a žen na úrovni prvních písmen kódů, USA, 2018.....	83
Obr. 24a:	Pořadí chorob z hlediska AS, HS a výše stupně v podgrafech, ženy, USA, 2018....	86
Obr. 24b:	Pořadí chorob z hlediska AS, HS a výše stupně v podgrafech, ženy, USA, 2018....	87
Obr. 25:	Pořadí chorob z hlediska AS, HS a výše stupně v podgrafech, muži, USA, 2018....	89

Seznam zkratk

ACME	Automatic Classification of Medical Entry
AS	Authority score
CB	mezilehlost
CCD	Contributory Cause of Death
CDAI	Cause of Death Association Indicator
CDC	Centers for Disease Control and Prevention
D	stupeň uzlu
DHHS	Department of Health and Human Services
DS	dýchací soustava
EAC	Entity axis codes
G-N	Girvan-Newman algorithm
HCDD	Human Cause of Death Database
HOÚ	hlášení o úmrtí
HS	hub score
inD	stupeň uzlu dovnitř
MCD	multiple cause of death
MKN	mezinárodní klasifikace nemocí
NBER	National Bureau of Economic Research
NCHS	National Center for Health Statistics
OR	odds ratio
OS	oběhová soustava
outD	stupeň uzlu ven
PPK	průměrný počet zapsaných příčin smrti
RAC	Record axis codes
RR	relativní riziko
SNA	metody analýzy sítí
TS	trávicí soustava
UCD	underlying cause of death
VNM	vazby, které nemají muži
VNŽ	vazby, které nemají ženy

Kapitola 1

Úvod

Téma vícečetných příčin smrti začalo v literatuře rezonovat již od poloviny minulého tisíciletí, kdy někteří autoři zastávali názor, že „*statistiky ukazující kombinace příčin smrti jsou blíže pravdě než ony založené na jediné příčině, protože většina úmrtí je výsledkem kombinací příčin smrti*“ (Janssen, 1940, s. 5). Ruku v ruce se změnami v úmrtnosti doprovázejícími epidemiologické přechody, tj. růst intenzit úmrtnosti na chronické a degenerativní nemoci nebo koncentrace zemřelých do spíše vyšších věků, se potřeba analyzovat vícečetné příčiny smrti ještě navýšila (Israel et al., 1986), leč mnozí autoři vyjadřovali skepsi před daty o vícečetných příčinách smrti, neboť vyvstávaly otázky okolo jejich srovnatelnosti napříč časem a prostorem (Guralnick, 1966). Za populaci USA však existují volně dostupná data o zemřelých se záznamy o všech příčinách smrti uvedených na jejich hlášení o úmrtí, která jsou zároveň systematizována podle minimálně měnných principů již od roku 1960, což je také opodstatněním, proč byla tato populace, konkrétně za rok 2018, vybrána pro analýzu. Jiným důvodem, jenž přispěl s prostorovému zasazení práce, je, že data o zemřelých v USA obsahují i další podrobné údaje o způsobu vyplnění hlášení o úmrtí, které byly nezbytné pro provedení analýzy vztahů mezi příčinami smrti.

Tato analýza bude provedena aplikací metod analýzy sociálních sítí. Sítě (grafy) jsou způsobem vyjádření vztahů mezi objekty a jsou tvořeny uzly (nody, entitami), které mohou být společně propojeny hranami (linky, vazbami). Jsou nástrojem, jenž lze používat v nejrůznějších oblastech: od sociologie (sítě vztahů v sociálních skupinách), kde mají také kořeny, přes geografii (sítě dopravní, migračních toků), informatiku (sítě odkazů na weby), kriminologii (sítě organizovaného zločinu), genetiku (sítě zachycující vztahy mezi geny a chorobami) atd. (viz Barabási a Pósfai, 2016). V neposlední řadě mohou nacházet využití i v demografii. Egidi et al. (2018) publikovali studii, jejíž náplní bylo právě zkoumání vztahů mezi vícečetnými příčinami smrti prostřednictvím sítí. Tento nápad má poměrně blízko k mnohem častější sféře aplikace grafů, a to k sítím komorbidit, v rámci nichž autoři zkoumají vztahy mezi nemocemi, kterými trpí živí lidé. Sítě komorbidit a dosud jediná modelovaná síť příčin smrti od Egidi et al. (2018) jsou však analyticky nesrovnatelné, protože úroveň agregace jednotlivých nemocí bývá v komorbiditních sítích zpravidla mnohem podrobnější, než v síti příčin smrti od Egidi et al. (2018), v níž autoři setrvávají na úrovni seskupení příčin smrti podle „European shortlist for causes of death“.

Cílem této práce je analýza asociací mezi příčinami smrti pomocí metod analýzy sítí, a to na oné podrobnější úrovni, analogicky jako v sítích komorbidit. Velký potenciál studií zaměřených na modelování biomedicínských sítí (mezi nimi i grafy komorbidit) může být viděn nejen v možnosti nabytí komplexnějšího pohledu na vztahy mezi nemocemi, ale i v predikci vazeb, které se ve vyvíjejících se systémech utvářejí (Barabási a Pósfai, 2016). Konkrétní využitelnost poznatků, k nimž lze díky SNA dojít i potenciál aplikace metody SNA v demografii, jsou rozvedeny až v závěru práce, neboť pro porozumění tomu, co struktury sítí implikují, je vhodné napřed definovat termíny fundamentální pro SNA.

Z důvodu, že SNA zřejmě nemusí být standardní demografickou metodou, se může jevit být nezbytným se odchýlit od standardní struktury závěrečné práce. Její první část je věnována popisu principů vzniku dat o vícečetných příčinách smrti v USA, což je následováno explorační analýzou těchto dat. Naopak v začátku práce chybí samostatná kapitola s rešerší literatury, a to z důvodu, že dosavadní zjištění jsou analyzována paralelně při srovnání vlastních výsledků a výsledků dosud provedených studií, poněvadž tento přístup vnáší do práce větší přehlednost¹. Možné postupy, přístupy a metody, které taktéž standardně spadají pod kapitolu s rešerší literatury, jsou rozebírány v rámci metodologické části práce, která sama navazuje na kapitolu o datových zdrojích a tu s explorační analýzou. Po metodologické kapitole následují teprve výzkumné otázky, a to protože bez přednostního definování termínů SNA a představení metod mohou formulace těchto otázek postrádat význam. Navíc je nezbytné nejdříve uvést způsob, jakým bude systém příčin smrti vymezen, teprve potom s reflexí způsobu vymezení klást otázky. Zbylé kapitoly jsou zaměřeny již standardně na výsledky provedené SNA a závěr.

¹ Srovnatelnost vlastních výsledků s poznatky z literatury je v daném tématu velmi stěžejní, zároveň není nutné se stále odkazovat na jedinou kapitolu.

Kapitola 2

Zdroj dat

Smyslem druhé kapitoly je popis procesu vzniku a zpracování dat o příčinách smrti v USA. Síť je totiž modelována z párů příčin smrti, jejichž vzájemný vztah je determinován místem, kam jsou na hlášení o úmrtí vyplněny. Proto je v úvodu kapitoly nastíněna metodika vyplnění tohoto hlášení, následně jsou popsány základní aspekty fungování algoritmů, které ze záznamů objevujících se na hlášení o úmrtí skládají datovou matici vícečetných příčin smrti. Přijatý způsob analýzy dat vzešlých z celého procesu zpracování je konkretizován v závěru této kapitoly.

2.1 Hlášení o úmrtí

Segment hlášení o úmrtí (HOÚ) pro zápis příčin smrti je rozdělen do dvou částí (obrázek 1). Prvá obsahuje stavy, jejichž posloupanost vyústila v smrt, druhá pak příčiny pouze k ní přispívající, ale bez přímé prokazatelné vazby na základní příčinu smrti (UCD) (National Center for Health Statistics, 2009a; dále NCHS). Základní příčina smrti je ta, která celý chorobný řetězec iniciovala, a platí, že by měla mít vazbu na bezprostřední příčinu, kterou je rozuměna choroba, či stav stojící na úplném závěru řetězce vedoucího ke smrti, tj. předcházející pouze úmrtí samotnému. Mezi UCD a bezprostřední příčinu smrti může osoba vyplňující HOÚ zapisovat i další, co do příčinnosti na sebe navazující stavy, jež bývají označovány jako předchozí příčiny úmrtí (ÚZIS ČR, 2006).

Bez ohledu na místo, kam se druhy příčin uvádějí, je forma jejich zaznamenávání unifikovaná podle aktuální verze Mezinárodní klasifikace nemocí (MKN). V datech o zemřelých Centra pro kontrolu a prevenci nemocí (CDC) by měla vždy být, a za rok 2018 rovněž i je, zachována podrobnost zápisu aspoň na úroveň třímístných kódů z číselníku, který je dostupný v české verzi např. na MKN-10 (2021), odkud jsou rovněž i citovány celé názvy všech chorob v následujících pasážích textu. Analýza je tudíž provedena nad třímístnými kódy, jelikož se jedná o standardní úroveň, s níž bývá pracováno při konstrukcích sítí komorbidit (viz např. studie citované v tabulce 9).

Obrázek 1: Část 1 a část 2 hlášení o úmrtí pro zaznamenání příčin smrti

CAUSE OF DEATH (See instructions and examples)		Approximate interval: Onset to death
<p>32. PART I. Enter the <u>chain of events</u>--diseases, injuries, or complications--that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary.</p>		
<p>IMMEDIATE CAUSE (Final disease or condition resulting in death) →</p>	<p>a. _____ Due to (or as a consequence of):</p>	<p>_____</p>
<p>Sequentially list conditions, if any, leading to the cause listed on line a. Enter the UNDERLYING CAUSE (disease or injury that initiated the events resulting in death) LAST</p>	<p>b. _____ Due to (or as a consequence of):</p> <p>c. _____ Due to (or as a consequence of):</p> <p>d. _____</p>	<p>_____</p> <p>_____</p> <p>_____</p>
<p>PART II. Enter other <u>significant conditions contributing to death</u> but not resulting in the underlying cause given in PART I</p>		<p>33. WAS AN AUTOPSY PERFORMED? <input type="checkbox"/> Yes <input type="checkbox"/> No</p>
		<p>34. WERE AUTOPSY FINDINGS AVAILABLE TO COMPLETE THE CAUSE OF DEATH? <input type="checkbox"/> Yes <input type="checkbox"/> No</p>

Zdroj: National Center for Health Statistics Instruction manual part 2a, 2009

Maximálně lze na standardní HOÚ v USA uvést až dvacet příčin smrti, přičemž počet v žádné z částí, ani množství příčin na jednom řádku nejsou nikterak omezeny, pokud v součtu nepřekračují uvedený maximální počet. Na který řádek osoba vyplňující HOÚ kód příčiny zapíše, ale hraje roli. Platí totiž, že choroba uvedená v části 1 na spodnějším řádku, by měla zapříčínovat chorobu zapsanou na svrchnějším, což je vyjádřeno pravidlem „due to“ (NCHS, 2009b) na obrázku 1. Z toho vyplývá, že správné místo pro vyplnění UCD, jakožto stavu udavšního se nejdříve ze všech v části 1 HOÚ uvedených, je vždy na řádku, nad nímž již není žádná další choroba. Například, pokud by na obrázku 1 byla na každém z řádků *a* až *d* uvedena právě jedna příčina smrti, pak by v souladu s pravidly měla být základní ona na řádku *d*. Není však výjimkou, že tomu tak není, protože v datech CDC o vícečetných příčinách smrti existují dvě proměnné, v nichž jsou kódovány příčiny smrti, které jsou považované, že by mohly iniciovat chorobné řetězce. Jedna z nich obsahuje „reported UCD“, což jsou pouze přímo příčiny smrti zapsané na HOÚ na nejvyšších řádcích², druhá obsahuje „coded UCD“ a vzniká z „reported UCD“, na níž jsou aplikována pravidla pro výběr UCD (Chamblee, 1986). Těmto pravidlům se podřizuje ACME, software pro automatizované kódování UCD. Používá algoritmus postavený na bohaté bázi rozhodovacích schémat, kterými by se měly řídit i osoby vyplňující HOÚ a pokud by nedělaly chyby, byly by jimi zakódované UCD v 99 % shodné s těmi, které by vybral ACME (NCHS, 2009a). Zmíněná pravidla pro výběr UCD jsou velmi široce popsána v obsažném instrukčním manuálu NCHS (2009a, 2009b, 2009c), avšak pro účel této práce jsou zde uvedena pouhá tři základní a jediné modifikační, neboť se níže vyskytují pasáže textu s odkazem jen na tato pravidla.

Pokud je na HOÚ více než jeden kód, tak je úkolem ACME identifikovat mezi nimi původní předchozí příčinu³. Aby to provedl, zkontroluje zprvu, zda je na každém z řádků zapsaný kód vedený v rozhodovacích schématech jako přijatelná příčina kódu uvedeného pod ním samotným, tedy zda platí pravidlo sekvence (NCHS, 2009a). Dále ACME zkoumá, zda platí obecný princip⁴ (NCHS, 2009a), podle něhož příčina na nejvyšším řádku by měla být vybrána jako UCD, pokud by mohla způsobit všechny uvedené pod ní, tudíž nejen bezprostředně pod ní, jako v prvním pravidlu sekvence. Pokud není aplikovatelný tento obecný princip, uchyluje se ACME k pravidlům pro výběr⁵, která jsou podle NCHS (2009a) tři:

- Podívá se na kód, který je uveden na nejvyšším řádku HOÚ a vybere ze všech ostatních na HOÚ uvedených ten, který by podle do něj implementovaných rozhodovacích tabulek mohl onu příčinu na nejvyšším řádku způsobit.
- Pokud žádný takový kód ACME neidentifikuje, ponechá jako UCD původní příčinu z nejvyššího řádku.
- Pokud kód, který v bodech výše ACME vybral, je v příslušných rozhodovacích tabulkách veden jako očividný přímý důsledek kteréhokoli z ostatních kódů na HOÚ, tak zapíše tento.

² Nejvyšším řádkem je myšlen ten, který je nejvíce spodní na HOÚ. Nejvyšším je označován proto, že v datech je značen nejvyšší cifrou.

³ *Originating antecedent cause*; pozn.: nejedná se o předchozí příčinu uvedenou v ÚZIS ČR (2006)

⁴ *General principle*

⁵ *Selection rules*

Nad to existuje nejen šest dalších modifikačních pravidel pro resekci UCD, ale i převelice detailně popsané kvantum specifických situací. Z modifikačních pravidel je pro kontext této práce postačující pouze prvé, jímž se ACME řídí, pokud v některém z předchozích kroků vybral jako UCD kód jsoucí na seznamu „ill-defined“ (jiných a nepřesně určených) příčin smrti, čímž se označují jiné a nepřesně určené či definované příčiny. V souladu s prvním modifikačním pravidlem totiž musí výběr provést znovu, a to jakoby tato „ill-defined“ příčina smrti na HOÚ uvedena vůbec nebyla. Seznam takových příčin smrti podle NCHS (2009a) obsahuje na úrovni třímístných kódů kódy R00–R99⁶, ale bez R95 („*syndrom náhlé smrti dítěte, kojence [SIDS]*“), z jiných kapitol je na něm taktéž I99 („*jiná a neurčená onemocnění oběhové soustavy*“). Lze najít i jiné seznamy nepřijatelných základních příčin smrti (např. Flagg et al., 2021), ale ACME bere v tomto směru ohled pouze na výše vypsané.

Proces zpracování dat o vícečetných příčinách smrti v USA nekončí ve chvíli, kdy ACME definitivně vybere UCD. Navazuje na něj další zpracování, jehož smyslem je nastolit vnitřní konzistenci zbylých příčin smrti uvedených na HOÚ a očistit je. Tato část zpracování je prováděna softwarem Transax (DHHS, 1986), který překládá kódy mezi dvěma typy, které se v datové matici CDC objevují. Prvým typem jsou „entity axis codes“ (EAC), druhým „record axis codes“ (RAC), přičemž druhé zmíněné vytváří Transax z prvně zmíněných. Proměnné EAC obsahují kódy shodné s těmi, které osoba vyplňující HOÚ na formulář vypsal, tedy i s chybami ve smyslu duplicitních, redundantních nebo protiřečících si kódů (Chamblee, 1986). Po těchto Transax pátrá, chyby opravuje a produkuje tak RAC. Jako jeden z nejtypičtějších příkladů pro úpravu entit uvádí CDC (1997) situace, kdy jsou na HOÚ (nebo rovněž v EAC) uvedeny kódy pro cirhózu jater (K74) a pro závislost na alkoholu (F10). Transax by podle CDC (1997) tyto dva sloučil do kódu K70 vyjadřujícího alkoholickou cirhózu jater, který by byl u zemřelé osoby jako jediný uveden v RAC.

Hlavní předností EAC je, že pouze z nich lze určit místo zapsání každého z kódů na hlášení, což je vlastnost, na které tato práce staví a je odůvodněním, proč je zde využito právě a pouze EAC. Navíc i podle Pechholdové (2021) bylo pro účel této práce vhodnější aplikovat právě EAC. CDC (1997) mezi dalšími důvody pro upřednostnění EAC dat před RAC uvádí, že jsou vhodná k examinaci etiologických vztahů mezi chorobami, testování validity předpokladů pro výběr UCD, nebrání v potaz pravidel Transax, nebo i v neposlední řadě, že jsou EAC v některých ohledech detailnější. Stěžejní klad RAC vyplývá z výše uvedeného, sice, že jsou to data očištěná od nesmyslných kódů, duplicit a „*dalších nedokonalostí*“ (CDC, 1997, s. 26), harmonizovaná a obsahují kódy, které Transax považuje za nejvhodnější (CDC, 1997).

Výsledná datová matice, jsou-li v potaz brány jen proměnné vícečetných příčin smrti (tj. bez sociodemografických charakteristik zemřelého), obsahuje osmdesát sloupců, z nichž 60 je pro EAC a zbylých 20 pro RAC. U RAC je na každou potenciálně vyplnitelnou příčinu smrti vyhrazena jedna proměnná, do níž se zapíše pouze čistý, zpravidla aspoň třímístný MKN kód. Oproti čemuž u EAC jsou ke každé proměnné pro zapsání kódu samotného přidruženy ještě další dvě. Prvá z nich nese informaci o místě zápisu kódu na HOÚ, nabývá hodnot 1–6, přičemž 1–5 znamená, že se příčina vyskytuje v první části HOÚ a hodnota čísla identifikuje řádek zapsání, hodnota 6 vyjadřuje, že kód byl zapsán do druhé části HOÚ.

⁶ Příznaky, znaky a abnormální klinické a laboratorní nálezy nezařazené jinde

Druhá skupina proměnných z těch, které jsou u EAC navíc oproti RAC, je užitečná pro určení toho, zda byla příčina smrti zapsána na řádek samostatně, či kolikátá v pořadí na něm byla uvedena, pokud se jich na jediném vyskytovalo vícero. Tedy, v případě, že se čísla řádku u jednoho zemřelého opakují (prvá proměnná u EAC kromě té se samotným kódem), neměla by se opakovat čísla sekvencí (druhá proměnná u EAC kromě té se samotným kódem). Naopak, pokud se čísla v první proměnné opakují u jednoho člověka, který zemřel s více než jednou příčinou smrti, tak u shodných čísel v první proměnné by se neměla opakovat čísla v druhé proměnné, protože pokud by tomu tak ryze teoreticky bylo, označovalo by to, že jsou kódy na jediném řádku zapsané paralelně.

Tedy, například kdyby se v datech CDC vyskytoval hypoteticky kód 12131415, bylo by možné z něj vyčíst, že zemřelá osoba měla na HOÚ uvedené čtyři příčiny smrti (ke každé proměnné s EAC jsou vždy dvě číselné proměnné konkretizující lokaci zápisu), všechny příčiny byly zaznamenány do první části HOÚ (neobjevuje se cifra 6) a všechny byly na jediném řádku (liché cifry jsou shodné, kvůli čemuž se musí lišit všechny sudé cifry, označující pořadí zapsání EAC na jediném řádku, tj. výše zmíněná čísla sekvencí). Jiným příkladem je ilustračně kód 216162 indikující, že zemřelý měl na HOÚ tři příčiny smrti, jednu v první části na druhém řádku (cifry 2 a 1) a dvě v druhé části napsané samozřejmě na jediném řádku, protože v části 2 HOÚ žádný další ani není (obrázek 1).

Zároveň samozřejmě vyvstává otázka, proč jsou řádky z první části HOÚ označeny čísly 1–5, když jsou celkem pouze čtyři (obrázek 1). Podle DHHS (2003) platí, že pokud osoba vyplňující HOÚ uváží, že smrti předcházelo a v ní vyústilo stavů pět, lze pod ručně zapsaným písmenem *e* do první části hlášení vpravit ještě onu pátou, navazující na předchozí stav.

2.2 Analýzy vícečetných příčin smrti

Mezi studiemi, které se zabývají analýzou úmrtnosti prostřednictvím vícečetných příčin smrti (MCD), existují nestejnosti ve způsobu pojetí MCD. Jednak jde o rozdílnosti ve významech pojmů, jednak také o odlišné způsoby inkorporování příčin smrti z hlediska jejich místa zápisu na HOÚ.

V některých studiích bývají příčiny smrti děleny stejně jako zde na UCD („underlying causes of death“) a MCD („multiple causes of death“), ale pod UCD jsou míněny všechny, které se vyskytly v první části HOÚ (Romon, 2007; Foreman, 2016). Pro pojmenování klasického způsobu analýzy příčin smrti, tzn. jednočetně, je v těchto situacích použit termín „single UCD“ a analyzována v rámci něho je „direct cause of death“ (přímá příčina smrti). Jako MCD pak autoři takových studií rozumí situace, kdy jsou uvažována i data z druhé části hlášení. U nich se tedy analýzy vícečetných příčin smrti zakládají na tom, že jsou jejich součástí jak všechny příčiny z části 1, tak i všechny pouze přispívající příčiny z části 2.

Podle jiného, evidentně většinového přístupu lze o MCD hovořit vždy, když je uvažována více než jen jediná příčina smrti u jednoho člověka (Dessesquelles 2010; Dessesquelles 2012; Gamage, 2021). Přístupy autorů těchto studií se pak liší v tom, zda do svých analýz vpraví i příčiny z druhé části HOÚ, či nikoli. Narazit lze i na studie, které dokonce uvažují pouze UCD a přispívající příčiny smrti.

Argumentují, že ostatní příčiny z první části HOÚ, nejsou ve své podstatě příčinami smrti, nýbrž jenom prostřednickými faktory odvislými od UCD (Rodriguez, 2019; Park, 2014), protože při zkoumání komorbidit nejsou relevantní, narozdíl od kódů zapsaných ve druhé části HOÚ, kam by měly být vepisovány skutečné komorbidity.

Ve světle výše uvedeného je patrné, že lze zvažovat, který ze způsobů použít. Smyslem této práce je analyzování vztahů mezi vícečetnými příčinami smrti a jejich modelování v rámci jediného systému, který ale existuje odděleně pro každou část HOÚ, neboť data po jeho jednotlivých částech jsou typově odlišná, jak vyplynulo z již napsaného. Entity z první části HOÚ by si měly být vzájemně příčinami a důsledky, kdežto entity z druhé mají s ostatními společné jen to, že byly uvedeny na tentýž list. K analýze je tudíž přistoupeno tak, že data z obou částí jsou analyzována zvlášť, kromě UCD, která vstupuje i do analýzy dat z druhé části HOÚ, kam však sama není zapsána.

Na systému vyplňování HOÚ, jenž byl popsán v této kapitole, jsou postavena pravidla pro utváření dvojic z vícečetných příčin smrti, které představují množinu objektů tvořících síť příčin smrti. Tvořená síť za část jedna je orientovaná, tj. uvažuje se směr vazby mezi chorobami, ve smyslu rozlišení příčin a důsledků. Síť za část dva je neorientovaná a tvořená dvěma disjunktivními množinami, kdy první sestává z UCD a druhá z přispívajících chorob části 2 HOÚ. Blíže jsou principy tvorby sítí vyloženy v metodologické kapitole.

Výše uvedené informace jsou tedy poměrně stěžejní pro rozumění logiky pravidel pro tvorbu dvojic formulovaných v metodologické části práce (kapitola 4). Tato pravidla jsou vybudována tak, že předpokládají, jako by všechny kódy uvedené na HOÚ byly správně umístěné v souladu s pravidly pro vyplnění HOÚ a selekci UCD.

Kapitola 3

Základní explorační analýza vícečetných příčin smrti

Cílem této kapitoly je provést explorační analýzu dat o vícečetných příčinách smrti, z nichž celá následná síťová analýza vychází. Analyzována jsou data o zemřelých v USA a jejich závislých územích v roce 2018, a to podle autorů, kteří se problematikou vícečetných příčin smrti zabývají. Tedy, jsou řešeny analogické aspekty jako v člancích Redelings et al. (2007), Wilkins et al. (1997), Pechholdová (2014) nebo Gorina a Lentzner (2008). Tato kapitola je zařazena zde, protože její výsledky jsou směrodatné pro způsob vymezení systému příčin smrti, jenž se nachází v navazující metodologické kapitole.

3.1 Průměrné počty zapsaných příčin smrti

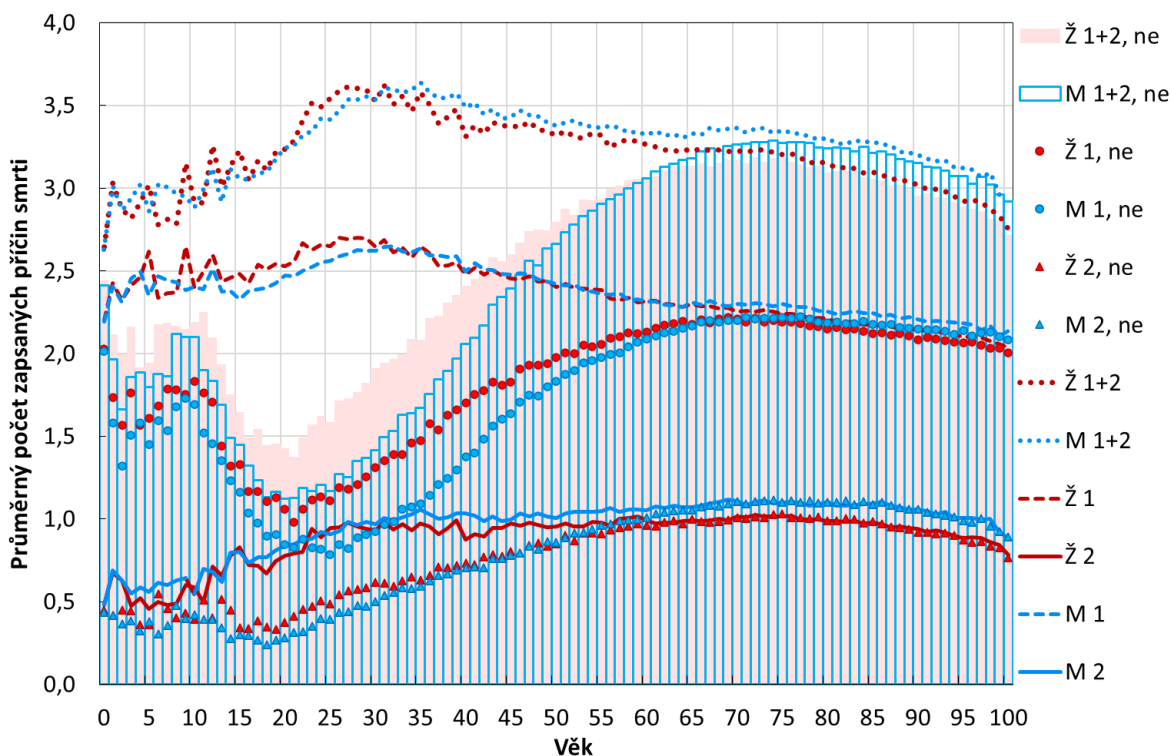
Průměrný počet zapsaných příčin smrti (PPK) je ukazatelem udávajícím informaci o tom, jak moc osoby vyplňující HOÚ využívají prostor vymezený pro zaznamenávání příčin smrti. Jeho výše má implikaci pro síťovou analýzu, neb spíše vyšší PPK předznamenává, že síť příčin smrti bude obsahovat mnoho vazeb, protože průměrně bude na jednom HOÚ vyšší počet příčin smrti, z nichž bude moci být utvořeno více párů chorob.

PPK se lehce mění v závislosti na věku (obrázek 2). PPK počítaný z dat z obou částí HOÚ byl v USA v 2018 nejvyšší v spíše mladších věcích, kdy mezi lety 25 a 35 dosahoval maximum ve výši 3,5 příčin na jedno hlášení. Je zřejmé, že za maximum v těchto věcích stojí příčiny z první části hlášení (obrázek 2). Podle CDC (2021) se jedná o věky, během nichž lidé nejčastěji umírají v důsledku vnějších příčin smrti, u nichž je podle pravidel MKN nutné vyplňovat nejen způsob přivodění vnější příčiny, ale i zranění, v něž vyústila (viz i dál), důsledkem čehož je PPK vyšší u této věkové skupiny v porovnání s jinými. Když se totiž z výpočtu PPK pominou příčiny smrti z rozmezí V01–Y98 („*vnější příčiny nemocnosti a úmrtnosti*“), jsou to právě tyto mladé věkové skupiny, u nichž hodnota PPK poklesne nejmarkantněji (obrázek 2). Konkrétněji se pak opominutí vnějších příčin smrti v rámci analýzy nejdůrazněji promítá do HOÚ mužů a první části.

Spíše vyšší PPK počítané ze všech příčin smrti v mladších věcích v porovnání s věky staršími je rys, jenž je zdrojem odlišnosti mezi USA a jinými vyspělými zeměmi. Například v Itálii či Francii podle Dessesquelles (2012), nebo i v Česku podle Pechholdové (2014), výše popisované maximum v mladších věcích vůbec neexistuje a PPK je maximální až ve spíše starších věcích, konkrétně kolem 60–80 let.

Zároveň ale v těchto věkových skupinách dosahuje týchž hodnot, jako v USA v 2018. Rozdíly v průběhu PPK v závislosti na věku ve Francii s Itálií, nebo i Česku a USA pramení zřejmě z odlišné skladby struktury příčin smrti, kdy v USA má úmrtnost v důsledku vnějších příčin vyšší zastoupení než např. v Česku (World Bank, 2021), nebo z odlišných zvyklostí při vyplňování HOÚ.

Obrázek 2: Průměrný počet uváděných příčin smrti na HOÚ v závislosti na věku zemřelého, USA, 2018



Pozn.: Ž = ženy, M = muži, 1 = část 1 HOÚ, 2 = část 2 HOÚ, 1+2 = obě části HOÚ, ne = bez inkorporace vnějších příčin smrti, tj. kódů V01–Y98

Zdroj dat: NBER (2020), výpočet vlastní

Naopak poznatkem konzistentním napříč dosud provedenými studiemi jmenovanými v úvodu kapitoly je, že v nejvyšších věcích PPK mírně klesá, za čímž podle Pechholdové (2014) může stát jednak hypotéza, že lidé umírající až za hranicí modálního věku smrti mohou být celkově zdravější a tím pádem mít na hlášení uvedeno méně chorob, nebo tento pokles může být manifestací implicitního předpokladu, že ve velmi vysokém věku je „umřít normální“, protože osoby vyplňující HOÚ mohou vykazovat nižší míru snahy dopátrat se po řetězcích chorobných stavů vedoucích ke smrti.

V USA klesá PPK počítané z obou částí HOÚ již od věku zhruba 30 let, avšak ve středním věku je zřetelná stagnace narušená až oním mírným opětovným klesáním od věku 75 let, o němž se zmiňuje Pechholdová (2014). O tom, že v USA je pravděpodobnější v odstavci výše uvedená druhá možnost pro odůvodnění tohoto poklesu ve věcích do zhruba 85 let, svědčí rozklad PPK podle částí hlášení. Totiž zatímco PPK v první části HOÚ klesá, ve druhé části, kam se píše přidružené choroby, se zhruba od věku 50 let zejména u žen drží kolem hodnoty jedna a pokles přichází až po věku 85. Tedy, „průměrná délka chorobného řetězce vedoucího ke smrti“ se až do tohoto věku zkracuje, kdež zatíženost zemřelých přidruženými chorobami zůstává přibližně neměnná. Po věku zhruba 85 let se snižuje PPK počítaný za každou z částí HOÚ.

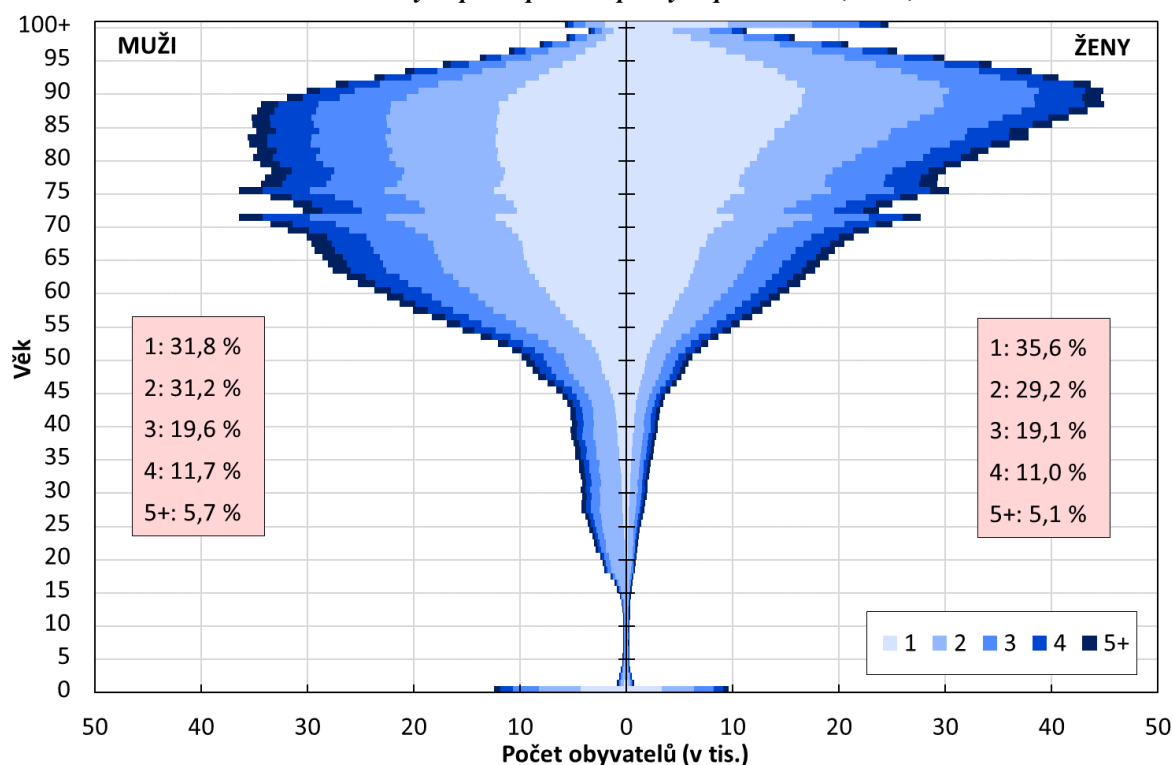
3.2 Struktura zemřelých podle počtu kódů na hlášení o úmrtí

Vedle PPK ovlivňuje síť i struktura zemřelých podle počtu kódů uvedených na HOÚ. Při výpočtech koeficientů asociací mezi příčinami smrti totiž hrají roli počty zemřelých se shodnými dvojicemi kódů, které samozřejmě závisí na tom, kolik osob zemřelo celkem, eventuálně v jakém věku, zejména jedná-li se o příčiny smrti typické pro určité věkové skupiny.

Podíly lidí podle celkového počtu zapsaných příčin smrti se v USA dlouhodobě výrazněji neproměňovaly⁷. Za rok 2018 je rozložení zemřelých podle počtu zapsaných příčin smrti uvedeno v tabulkách v rámci obrázku 3. Více než třetina osob měla na HOÚ příčinu jedinou, přičemž u žen je tento podíl mírně výše podílu za muže. Ostatní, tj. u žen asi 65 %, u mužů 70 %, zemřeli s aspoň dvěma příčinami smrti. Do tady modelované sítě se promítá lvím podílem skladba příčin smrti právě této subpopulace, přičemž platí, že s rostoucím počtem uvedených příčin narůstá zpravidla i počet dvojic, které je z nich možné vytvářet, což vyúsťuje v to, že v rámci sítě může být „nadhodnocena“ část struktury příčin smrti, která je typická pro osoby, které umírají s jejich vyšším počtem. Již bylo zmíněno, že takovými by vždy měli být zemřelí v důsledku vnějších příčin smrti, a i proto je tato oblast úmrtnosti v síti reprezentována ne poskrovnu (více v kapitole s výsledky) a tkví v ní rovněž vysvětlení, proč muži lehce zaostávají za ženami v podílu zemřelých s jedinou příčinou, a to o zhruba 4 p. b. (obrázek 3).

Ve věcích od 15 do 50 let je absolutní počet zemřelých mužů vyšší, přičemž většinou neumírají s příčinou pouze jednou. Platí, že u mužů až do věku 57 let převyšuje počet zemřelých s více než jednou příčinou počet zemřelých s právě jednou, v porovnání s čímž u žen je tento převis jednak menší a jednak také končí již kolem věku 40 let.

Obrázek 3: Věková struktura zemřelých podle počtu zapsaných příčin smrti, USA, 2018



Zdroj dat: NBER (2020)

⁷ Tvrzeno na základě výsledků vlastní analýzy.

3.3 Místo vyplnění základní příčiny smrti

V předchozí kapitole bylo uvedeno, že v datech existují dvě proměnné s UCD, jedna z nich je s reportovaným kódem, druhá obsahuje UCD vybranou softwary a na jejím základě jsou rovněž publikovány oficiální statistiky příčin smrti USA. Již v nejstarších studiích zabývajících se vícečetnými příčinami smrti (Israel et al., 1986) je za relevantní údaj považován podíl případů, kdy oficiální UCD vybraná (nebo vytvořená) softwarem ACME či Transax není osobou vyplňující HOÚ zapsána na nejvyšší řádek. Jinými slovy je zkoumáno, jak často se liší „coded UCD“ a „reported UCD“, což je informace, která je taktéž směrodatná pro navazující metodologickou část, neboť se od ní odvíjí rozhodnutí o tom, kterou z proměnných s UCD při rekonstrukcích chorobných řetězců využít.

Podíl případů, kdy UCD není na nejvyšším řádku (neboli „reported UCD“ a „coded UCD“ se neshodují) se u obou pohlaví za rok 2018 pohybuje kolem 30 %. Tyto případy podle některých autorů poukazují na chyby ve vyplňování HOÚ a jsou podle nich rovněž důvodem, proč je lépe provádět analýzu úmrtnosti nad RAC nikoli nad EAC (NCHS, 2009b nebo Redelings et al., 2006). Pod vlivem těchto názorů byl také i zde při přípravě dat pro síť tvořen řetězec chorobných stavů odvíjející se od oficiální „coded UCD“, nikoli od příčiny zapsané na nejvyšším řádku, ač ve výsledku by ona druhá cesta, kdy by bylo pracováno s „reported UCD“, mohla přivést v některých aspektech i k „smysluplnějším“ výsledkům. Mezi oněmi „mylnými“ 30 % se totiž mimo jiné vyskytují i případy, kdy je na nejvyšším řádku skutečně vyplněn kód, který do řetězce chorobných stavů uvedeného na HOÚ zapadá jedině na pozici iniciátora, zároveň je však kódem, který nesmí být dle mezinárodních pravidel uváděn jako UCD, proto by ACME mělo mít snahu jej při výběru UCD z různých důvodů obejít, např. protože se jedná o blíže nespecifikovanou chorobu, což automaticky zcela nevylučuje že i případně nejdříve nastoupivší ze všech na HOÚ uvedených, čímž zároveň také mající základní vlastnost UCD, tj. iniciující chorobný řetězec ústící v smrt. Seznam nepřijatelných základních příčin smrti, z nichž mnohé má rovněž ACME podle NCHS (2009a) za úkol snažit se při výběru UCD míjet, poskytují Flagg et al. (2021). Rozdělují je na neznámé a chybně zakódované (zejména R kódy), dále bezprostřední a předchozí (např. nespecifikované infekce nebo selhání orgánů) a blíže nespecifikované příčiny smrti.

Souhrnně, mezi oněmi 30 % případů, kdy jsou „coded UCD“ a „reported UCD“ rozdílné, a které jsou tudíž označovány za „chybné“, jsou například i takové, které mají na nejvyšším řádku uvedený symptom nemoci, a až níže nemoc samotnou, což není v rozporu s pravděpodobnou možnou návazností kódů, pouze s pravidly pro kódování UCD. Však stává se, že ACME na pozici iniciátora chorobného řetězce vedoucího ke smrti ponechává UCD explicitně formulovanou jako nepřijatelnou. Ve většině takových případů se však jedná zejména o zemřelé, u nichž je UCD nedohledatelná, protože je u nich na HOÚ jen kód R99 („jiné nepřesně určené a neurčené příčiny smrti“). Více k tomuto v oddílu věnovaném „ill-defined causes of death“, jenž je součástí této kapitoly.

Při zkoumání převládajících způsobů vyplnění HOÚ lze jít i hlouběji. Podle metodologie systému na automatické kódování UCD (Chamblee, 1986), má tento systém nejen pravomoc UCD vybírat, ale rovněž je z těch, které jsou na HOÚ uvedeny, skládat, nebo rovnou vpravovat příčiny zcela nové.

Případů, kdy oficiálně vybraná UCD vůbec není v první části HOÚ, bylo v roce 2018 u mužů i žen kolem sedmi procent. Těchto sedm procent případů není tvořeno jen situacemi, kdy UCD na hlášení vůbec není, ale i situacemi, kdy UCD je pouze mimo první část, tj. je možné, že může být v druhé.

3.4 Nejčastěji zapsané kódy

Dalším dílčím cílem ve studiích citovaných v úvodu kapitoly bývá obvykle nalezení kódů nejhojněji se vyskytujících v každé z částí HOÚ. Toto je relevantní i v kontextu sítí, neboť lze očekávat, že správně modelovaný systém bude v první řadě obsahovat kódy, které se na HOÚ vyskytují nejčastěji. Navíc díky podílům zemřelých na nejčastější příčiny je získána zásadní informace o rozložení zastoupení jednotlivých příčin smrti v populaci.

V tabulkách 1 a 2 je uvedeno prvních patnáct nejčastějších příčin smrti mezi kódy z části jedna, dva a mezi základními příčinami smrti za každé pohlaví. Nejčastějšími kódy z první části jsou myšleny pouze ony neshodné s UCD⁸. Zhruba polovina lidí zemřela s těmito prvními patnácti nejčastějšími příčinami v rámci každé z částí HOÚ, přičemž celkový počet rozdílných třímístných kódů v datasetu zemřelých přesahuje u obou pohlaví tisíc. Tudíž nerovnost v zastoupení příčin smrti⁹ na úrovni třímístných kódů je velmi vysoká a rovněž platí, že se nejčastější diagnózy v různých částech hlášení i v poli pro UCD v některých případech neodlišují (některé kódy se opakují ve víceru sloupcích tabulek 1 a 2).

U obou pohlaví jsou choroby na prvních místech nerozdílné. U mužů nejčastější UCD, ischemická choroba srdeční (I25), je dokonce zaznamenána až v 10 % případů. Na prvním místě ve druhé části hlášení se vyskytuje duševní porucha způsobená užíváním tabáku (F17), která je u žen psána v pětině, u mužů ve čtvrtině případů, kdy je do druhé části hlášení vyplněn aspoň jeden kód. Takto vysoký podíl u mužů může souviset s tím, že druhou a třetí nejčastější UCD jsou u nich choroby postihující dýchací soustavu (DS). U žen je v tomto ohledu druhou nejčastější Alzheimerova nemoc (G30). Mezi nejčastějšími UCD jsou dále u mužů i žen některé novotvary, srdeční choroby, diabetes, další demence a otravy (u žen krve, u mužů narkotiky).

Nepočítaje UCD jsou do první části HOÚ nejčastěji psány kódy, které jsou na seznamu nepřijatelných příčin smrti od Flagg et al. (2021). Vymykají se v tomto směru jenom jiná chronická obstrukční plicní nemoc (J44), primární hypertenze (I10) a otevřená rána hlavy (S01), přičemž prvé uvedené onemocnění je zároveň také jedním z nejvíce se vyskytujících UCD a po něm zmíněná I10 je druhou nejčastější chorobou psanou do druhé části hlášení. V ní také již nepřevládají pouze choroby DS a kardiovaskulární, ale téměř třetinovým dílem jsou zastoupeny choroby metabolické (E), a to u obou pohlaví. Mezi nejčastějšími kódy ve druhé části HOÚ jsou taktéž vnější příčiny smrti a jejich následky.

⁸ Frekvenční tabulky za první část HOÚ byly tvořeny pouze z příčin smrti, které se lišily od UCD v individuálních záznamech zemřelých (není to tedy myšleno tak, že by v tabulkách 1 a 2 nesměly být shodné příčiny smrti v 1. a 3. sloupci).

⁹ Existuje malý počet vysoce častých příčin smrti a velký počet málo častých.

Tabulka 1: Nejčastěji uváděné kódy podle části hlášení a v proměnné „coded UCD“, ženy, USA, 2018

Příčina v části 1 HOÚ	Podíl (v %)	Základní příčina smrti	Podíl (v %)	Příčina v části 2 HOÚ	Podíl (v %)
Srdeční zástava (I46)	7,9	Chronická ischemická choroba srdeční (I25)	7,5	Poruchy způsobené tabákem (F17)	21,2
Respirační selhání nezařazené jinde (J96)	7,0	Alzheimerova nemoc (G30)	6,1	Esenciální (primární) hypertenze (I10)	6,4
Esenciální (primární) hypertenze (I10)	6,8	Jiná chronická obstrukční plicní nemoc (J44)	5,6	Neurčená demence (F03)	3,8
Selhání srdce (I50)	6,5	Neurčená demence (F03)	4,9	Chronické onemocnění ledvin (N18)	3,3
Pneumonie, původce NS (J18)	3,5	Zhoubný novotvar průdušky (bronchu) a plicí (C34)	4,8	Fibrilace a flutter síní (I48)	3,1
Jiná sepse (A41)	3,4	Selhání srdce (I50)	3,3	Neurčený diabetes mellitus (E14)	2,8
Nedostatek předp. norm. fyziol. vývoje (R62) (1)	2,9	Akutní infarkt myokardu (I21)	3,2	Selhání srdce (I50)	2,7
Chronická ischemická choroba srdeční (I25)	2,7	Zhoubný novotvar prsu (C50)	3,1	Jiná chronická obstrukční plicní nemoc (J44)	2,3
Jiné příznaky a znaky týkající se OS a DS (R09) (3)	2,1	Cévní mozková příhoda (mrtvice) (I64) (4)	2,2	Chronická ischemická choroba srdeční (I25)	2,3
Chronické onemocnění ledvin (N18)	2,0	Postižení srdce při hypertenzi (I11)	1,9	Poruchy metabolismu (E78) (2)	1,8
Sekundární zhoubný novotvar DS a TS (C78) (5)	1,9	Pneumonie, původce NS (J18)	1,6	Diabetes mellitus 2. typu (E11)	1,7
Sekundární zhoubný novotvar neurč. lok. (C79) (6)	1,6	Neurčený diabetes mellitus (E14)	1,6	Alzheimerova nemoc (G30)	1,4
Neurčená demence (F03)	1,5	Zhoubný novotvar slinivky břišní (C25)	1,6	Vystavení neurčeným faktorům (X59)	1,3
Fibrilace a flutter síní (I48)	1,4	Jiná sepse (A41)	1,5	Otrava diuretiky a jinými léčivými (T50) (7)	1,2
Jiná chronická obstrukční plicní nemoc (J44)	1,4	Zhoubný novotvar tlustého střeva (C18)	1,4	Obezita – otylost (E66)	1,1
Ostatní	47,5	Ostatní	49,8	Ostatní	43,6

Pozn.: (1) Nedostatek předpokládaného normálního fyziologického vývoje, (2) Poruchy metabolismu lipoproteinů a jiné lipidemie, (3) Jiné příznaky a znaky týkající se oběhové a dýchací soustavy, (4) Cévní mozková příhoda (mrtvice) neurčená jako krvácení nebo infarkt, (5) Sekundární zhoubný novotvar dýchací a trávicí soustavy, (6) Sekundární zhoubný novotvar jiných a neurčených lokalizací, (7) Otrava diuretiky a jinými a neurčenými léky, léčivými, návykovými a biologickými látkami

Zdroj dat: NBER (2020), výpočet vlastní

Tabulka 2: Nejčastěji uváděné kódy podle části hlášení a v proměnné „coded UCD“, muži, USA, 2018

Příčina v části 1 HOÚ	Podíl (v %)	Základní příčina smrti	Podíl (v %)	Příčina v části 2 HOÚ	Podíl (v %)
Srdeční zástava (I46)	7,2	Chronická ischemická choroba srdeční (I25)	10,2	Poruchy způsobené tabákem (F17)	26,9
Respirační selhání nezařazené jinde (J96)	5,9	Zhoubný novotvar průdušky (bronchu) a plíce (C34)	5,2	Esenciální (primární) hypertenze (I10)	4,4
Selhání srdce (I50)	5,9	Jiná chronická obstrukční plicní nemoc (J44)	4,8	Poranění neurčené části těla (T14)	3,4
Esenciální (primární) hypertenze (I10)	5,5	Akutní infarkt myokardu (I21)	4,4	Neurčený diabetes mellitus (E14)	2,9
Chronická ischemická choroba srdeční (I25)	3,8	Selhání srdce (I50)	2,6	Chronická ischemická choroba srdeční (I25)	2,9
Pneumonie, původce NS (J18)	3,3	Alzheimerova nemoc (G30)	2,6	Chronické onemocnění ledvin (N18)	2,9
Jiná sepse (A41)	2,8	Neurčená demence (F03)	2,3	Fibrilace a flutter síní (I48)	2,3
Otrava narkotiky (T40) (1)	2,3	Zhoubný novotvar předstojné žlázy–prostaty (C61)	2,2	Poruchy způsobené více drogami (F19)	2,2
Chronické onemocnění ledvin (N18)	2,0	Neurčený diabetes mellitus (E14)	1,9	Jiná chronická obstrukční plicní nemoc (J44)	2,1
Otevřená rána hlavy (S01)	2,0	Postižení srdce při hypertenzi (I11)	1,8	Selhání srdce (I50)	2,0
Postižení srdce při hypertenzi (I11)	1,9	Zhoubný novotvar slinivky břišní (C25)	1,6	Neurčená demence (F03)	1,9
Jiné příznaky a znaky týkající se OS a DS (R09) (2)	1,7	Cévní mozková příhoda (mrtvice) (I64) (6)	1,5	Otrava diuretiky (T50) (8)	1,6
Sekundární zhoubný novotvar DS a TS (C78) (3)	1,6	Zhoubný novotvar tlustého střeva (C18)	1,5	Diabetes mellitus 2. typu (E11)	1,6
Sekundární zhoubný novotvar neurč. (C79) (4)	1,6	Náhodná otrava narkotiky (X42) (7)	1,5	Poruchy metabolismu lipoproteinů (E78) (9)	1,5
Nedostatek před. norm. fyziol. vývoje (R62) (5)	1,3	Pneumonie, původce NS (J18)	1,5	Poruchy způsobené alkoholem (F10)	1,3
Ostatní	51,2	Ostatní	54,6	Ostatní	40,0

Pozn.: (1) Otrava narkotiky a psychodysleptiky (halucinogeny), (2) Jiné příznaky a znaky týkající se oběhové a dýchací soustavy, (3) Sekundární zhoubný novotvar dýchací a trávicí soustavy, (4) Sekundární zhoubný novotvar jiných a neurčených lokalizací, (5) Nedostatek předpokládaného normálního fyziologického vývoje, (6) Cévní mozková příhoda (mrtvice) neurčená jako krvácení nebo infarkt, (7) Náhodná otrava narkotiky a psychodysleptiky a expozice, (8) Otrava diuretiky a jinými a neurčenými léčivými, návykovými a biologickými látkami, (9) Poruchy metabolismu lipoproteinů a jiné lipidemie

Zdroj dat: NBER (2020), výpočet vlastní

3.5 Jiné a nepřesně určené příčiny smrti

Výše bylo uvedeno, že ač by některé příčiny smrti, zejména ony označované „ill-defined causes of death“ (jiné a nepřesně určené příčiny smrti), neměly být kódovány jako UCD, lze na takové v datech narazit. Je důležité zjistit možné příčiny a jak často k tomuto dochází, aby výsledky síťových analýz týkající se jiných a nepřesně určených příčin smrti nebyly nekorektně interpretovány. Mohlo by k tomu docházet v orientované síti, kde jsou rozlišeny kódy na pozici příčin a důsledků na základě toho, jak byly uvedeny na HOÚ (viz pravidla pro tvorbu příčin a důsledků). Pokud by se kód ze seznamu jiných a nepřesně určených příčin smrti objevil na pozici příčiny, znamenalo by to, že na HOÚ musel být uváděn aspoň ve trojici, pokud by neexistovaly situace, kdy jsou jako UCD ponechávány jiné a nepřesně určené příčiny smrti. Za účelem zjištění, zda je interpretace, že „ill-defined“ příčina smrti na pozici příčiny ve dvojici chorob v orientované síti indikuje nutnost přítomnosti aspoň tří kódů na HOÚ, je korektní, i navzdory tomu, že jiné a nepřesně určené příčiny smrti figurují v UCD, obsahuje práce tuto podkapitolu.

Proměnná s oficiální „coded UCD“ obsahuje sedmkrát méně kódů, které jsou na seznamu jiných a nepřesně určených příčin smrti, který je včleněn do ACME, než proměnná s „reported UCD“, ale stále je obsahuje. Přibližně 1 % osob, které v USA v 2018 zemřely, měly uvedenou jako UCD příčinu, která by podle pravidel ACME měla být při selekci neuvažována. Více než třetinovým dílem je mezi jinými a nepřesně určenými základními příčinami zastoupeno R99, které se v průměru vyskytuje na HOÚ spíše samostatně (tabulka 3). Zároveň však lze zjistit, že za R99 následující nejčastější jiné a nepřesně určené UCD mívají průměrně okolo 1,5 kódů zapsaných na HOÚ v první části celkem. Vysvětlení, proč ač PPK převyšuje u některých UCD hodnotu jedna a i za těchto podmínek zůstávají vedeny jako UCD, lze podávat v kontextu ostatních příčin, které jsou k nim nejčastěji zapisovány. Jedná se totiž rovněž o jiné a nepřesně určené příčiny smrti, nebo v menší míře o typické bezprostřední příčiny smrti, konkrétně I46 („srdeční zástava“) nebo J96 („respirační selhání nezařazené jinde“).

Jak již bylo v prvním odstavci v této podkapitole nastíněno, v části s výsledky je opakovaně skloňováno, že „ill-defined“ příčiny smrti na pozici příčin v párech kódů (viz postup tvorby příčin a důsledků) mají v síti výsadní postavení. Toto tvrzení je korektní, neboť i když se v UCD objevují, tak nefrekventovaně, jak je doloženo v tabulce 3. Podíl zakódování nejčastějších jiných a nepřesně určených příčin do UCD je téměř u všech těchto kódů vzácnější deseti procent, výjimkou jsou R99, R54 a R58 (tabulka 3), jejichž podíly zapisování do UCD sice tuto hranici přesahují, avšak stále je i z těchto příčin smrti většina kódována do polí, kam by kvůli příslušnosti k „ill-defined“ příčinám smrti také býti kódována měla (nejpravější sloupec tabulky 3).

Proč je relevantní toto řešit, když bylo uvedeno, že jiné a nepřesně určené UCD mělo v 2018 jen 1 % zemřelých v USA? Důvodem je, že se do sítí dostávají i v absolutním zastoupení velmi málo časté dvojice chorob, třeba i v řádu pouhé desítky. Sice lze nacházet souvislost mezi intenzitou úmrtnosti na tu kterou chorobu a jejím postavením v síti (viz dále), rozhodně ale analýza intenzity procesu není smyslem užití SNA. Ten spočívá v podchycení komplexity vztahů, jak uvádí Barabási a Pósfai (2016). A komplexita indikuje i přítomnost chorob, které nejsou časté a možná pro analýzu intenzity úmrtnosti relevantní. Díky informacím v tabulce 3 je tedy možné v síti objevit vyskytující se trojice chorob, jejichž součástí jsou jiné a nepřesně určené příčiny smrti, nejen dvojice.

Tabulka 3: Nejčastější jiné a nepřesně určené UCD, obě pohlaví dohromady, USA, 2018

Kód	Název UCD	Počet	Podíl (v%)	PPK	Podíl vUCD(%)
R99	Jiné nepřesně určené a neurčené příčiny smrti	12 255	38,7	1,08	32,1
R54	Stáří (senilita)	4 274	13,5	1,48	21,8
R62	Nedostatek před. norm. fyz. vývoje	3 373	10,6	1,41	6,9
R09	Jiné příznaky a znaky týkající se OS a DS	2 613	8,2	1,43	6,5
R68	Jiné celkové příznaky a znaky	1 739	5,5	1,67	7,5
R57	Šok nezařazený jinde	1 628	5,1	1,72	7,2
R56	Křeče nezařazené jinde	1 037	3,3	1,48	8,6
R06	Nepravidelnosti dýchání	685	2,2	1,97	9,7
R58	Krvácení nezařazené jinde	621	2,0	1,61	17,7
R53	Nevolnost a únava	512	1,6	1,71	4,1

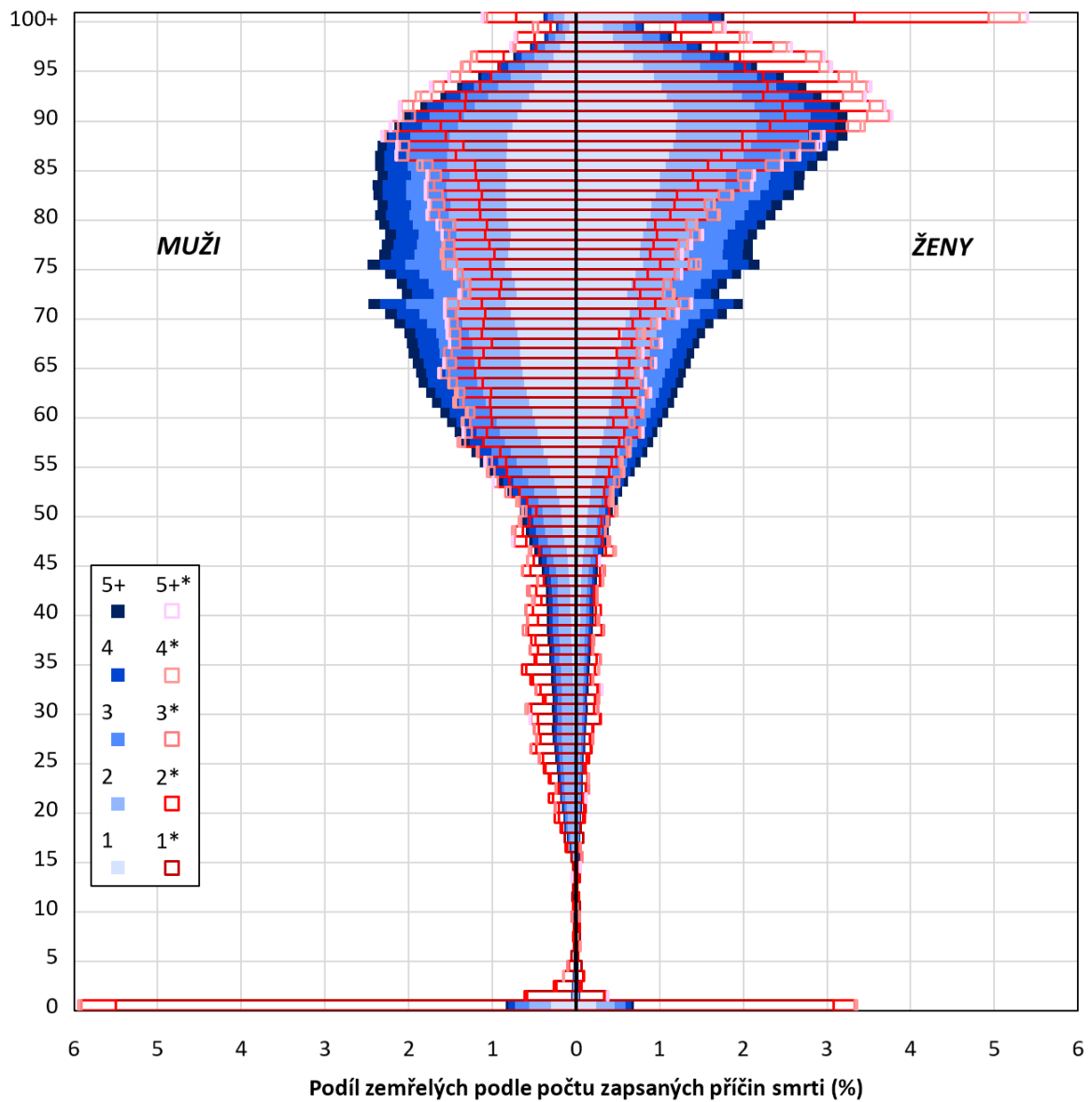
Pozn.: Počet = počet zemřelých na uvedenou „ill-defined“ UCD; Podíl = relativní zastoupení dané „ill-defined“ UCD na všech „ill-defined“ UCD; PPK = průměrný počet zapsaných příčin smrti u osob zemřelých s „ill-defined“ UCD; Podíl v UCD = podíl zemřelých, kteří měli uvedenou „ill-defined“ UCD na všech, kteří ji měli uvedenou na HOÚ

Zdroj dat: NBER (2020)

Pohlavně věková struktura subpopulace zemřelé s jinými a nepřesně určenými UCD se lehce neshoduje s většinovou, která byla prezentována v obrázku 3 výše. Převaha žen nad muži je vyšší, konkrétně zastupují 53 % všech zemřelých s „ill-defined“ UCD, a také je modální věk při smrti v této subpopulaci vyšší, i když jsou v ní nezanedbatelným podílem zastoupeni i kojenci. Na obrázku 4 je srovnání relativních věkových struktur většinové populace zemřelých a zemřelých s jinou a nepřesnou UCD. V případě žen nejzřetelněji více často v relativním vyjádření s „ill-defined“ UCD zemřely osoby starší 90 let a novorozenci. Podobný vzorec je viděn i u mužů, však s rozdílem, že nejvyšší převis se nalézá právě v kojeneckém věku (obrázek 4) a spíše nežli v těch úplně nejstarších, umírají s jinou a nepřesně určenou UCD častěji muži ve věku 20 až 45 let. Ve věcích, kdy muži umírají nejčastěji, naopak subpopulace zemřelých s jinou a nepřesně určenou UCD v zastoupení nápadně z hlediska podílu zaostává.

S již výše zmíněným ohledně průměrných počtů příčin zapsaných na HOÚ u lidí s „ill-defined“ UCD souvisí i to, že zejména v mladých věcích převládají zemřelí s jedinou příčinou. Osob s více než jednou přibývá naopak až od věků vyšších, což je vlastnost inverzní k té, jíž je charakteristické složení zemřelých podle PPK v celé populaci.

Obrázek 4: Zemřelí podle počtu zapsaných příčin smrti, celá populace a populace zemřelá pouze s jinými a nepřesně určenými základními příčinami smrti, USA, 2018



Pozn.: * = za populaci zemřelou s jinými a nepřesnými základními příčinami smrti

Zdroj dat: NBER (2020), zpracování vlastní

3.6 Základní příčiny smrti podle průměrného počtu zapsaných příčin smrti

Před přistoupením k vymezení systému příčin smrti v následující kapitole je klíčové získání údaje chorobách, které nebývají na HOÚ osamělé. Analogicky jako v začátku explorační analýzy bylo PPK počítané podle věku zemřelých osob, tak je nyní počítáno podle základní příčiny smrti. Poznatky, které jsou tímto dílčím krokem získány, nesou informace o tom, jaké choroby by neměly absentovat zejména v sítích neorientovaných a na výsledky prezentované v tabulkách 4–7 je proto opakovaně odkazováno v kapitole 8.

Základní příčiny smrti podle průměrného počtu kódů, které se k nim zapisují, jsou zaznamenány v tabulkách 4–7. Obsahem tabulek jsou jenom UCD vyskytující se častěji než jedním promile a uvedeno je vždy 15 příčin s nejvyšším PPK a následně 15 s nejnižším. Průměr je vážen počtem lidí, kteří na danou UCD zemřeli¹⁰. Na začátek je ke zmínění, že jsou-li brány choroby v populaci nejčastěji se vyskytující, bývají i nejnižší příčky tabulek vyplněny povětšinou kódy, u nichž PPK je výše jedné. Tedy, pokud by věková struktura z obrázku 3 byla vztažena pouze k osobám, které zemřely na nejčastější příčiny smrti, byl by podíl osob zemřelých s jedinou příčinou ještě nižší.

U obou pohlaví bývá nejvíce dalších chorob na HOÚ uváděno v souvislosti s úmrtími v důsledku vnějších příčin, na srdeční onemocnění a nemoci endokrinní, výživy a přeměny látek (začínají písmenem E). U mužů bývají v průměru více než čtyři další choroby psány i při UCD na infekční, bakteriální, nebo virová onemocnění. Není bez povšimnutí, že se na HOÚ v první části taktéž často vyskytuje s vyšším počtem dalších chorob u obou pohlaví C97. Podle české verze instrukční příručky MKN-10 se jedná o kód, který by neměl být zapisován jako základní, je-li uvedeno na HOÚ více příčin smrti (MKN-10 Instrukční příručka, 2020). Užit jej jako UCD lze, pokud se na HOÚ objevují dva nezávislé (např. anatomicky jinde lokalizované) novotvary, z nichž jeden neměl souvislost s druhým. V manuálu NCHS (2009a) popisujícím pravidla ACME není k dohledání, že by C97 nemělo být psáno jako UCD, uváděna jsou pouze dílčí pravidla, kdy zapsáno být má. Z nich plyne analogické, co z MKN-10 a na jejich základě může být vyvozeno, že vyšší průměrný počet zapsaných příčin „implikuje“ zapsání C97 jako UCD, neboť čím více je na HOÚ uvedeno dalších kódů stojících pro novotvary, tím může být pravděpodobnější, že alespoň jeden z nich bude nezávislý (anatomicky jinde lokalizovaný) na ostatních.

Z chvostů seznamů často se vyskytujících chorob podle průměrného počtu k nim připsaných dalších kódů (tabulky 6 a 7) je patrné, že naopak nejméně zapisovaných dalších příčin se váže především na novotvary. Po vícerech zejména u žen se vyskytují i některé R choroby.

Dále jsou v tabulkách 6 a 7 uvedeny ještě některé nemoci nervové soustavy a srdeční, přesněji ony patřící na seznam typických bezprostředních příčin smrti, jmenně srdeční zástava nebo selhání srdce, taktéž i respirační onemocnění téhož druhu, příkladem respirační selhání.

V návaznosti na jednu z předchozích podkapitol týkající se explorační „ill-defined“ příčin smrti by mohlo být patřičným zdůraznit, že právě tyto choroby se často zapisují na HOÚ společně s R kódy a při takovýchto kombinacích je ACME neupřednostňuje při výběru UCD, tudíž jsou to ty, které jsou skryty za vyšším PPK u výjimečných případů, kdy je oficiálně vybraná UCD z kategorie „ill-defined“.

¹⁰ Byl spočítán jako součet součinu počtu zemřelých podle UCD a celkového počtu zapsaných příčin smrti vydělený celkovým počtem zemřelých podle UCD. Cílem je učít větší váhu tomu počtu příčin smrti, s nímž zemřelo více osob.

Tabulka 4: Vybrané kódy s nejvyšším průměrným počtem zapsaných příčin smrti, ženy, USA, 2018

Kód	Základní příčina smrti	PPK	‰
X44	Náhod. otrava j. a neurč. lék., léč., návyk. a biol. lát. a expoz. jej. působení	5,9	6,5
X64	Úmyslné sebeotr. j. a neurč. léky, léč., náv. a biol. lát. a exp. jejího působení	5,5	1,3
I13	Hypertenzní nemoc srdce a ledvin	5,4	3,3
C97	ZN mnohočetných samostatných (primárních) lokalizací	5,3	1,6
Y83	Operace a j. oper. výk. jako příč. abn. reakce nebo pozd. kompl. (bez neh.)	5,2	1,0
X59	Vystavení neurčeným faktorům	5,1	2,3
I12	Postižení ledvin při hypertenzi	5,0	5,0
W80	Ucpání dýchacích cest vdechnutím nebo polknutím j. předmětů	4,9	1,1
X41	Náh. otrava antiép., sed-hypn., antipark. a psych. lék. a exp. jej. půs., NJ	4,7	1,4
E78	Poruchy metabolismu lipoproteinů a jiné lipidemie	4,6	4,4
E11	Diabetes mellitus 2. typu	4,6	10,1
E87	Jiné poruchy tekutin, elektrolytů a acidobasické rovnováhy	4,6	2,3
X42	Náhodná otrava narkot. a psychodyslep. a expoz., NJ	4,5	5,3
I34	Nerevmatická onemocnění dvojčipé chlopně [valvulae mitralis]	4,5	1,2
N28	Jiné poruchy ledviny a ureteru nezařazené jinde	4,4	1,8

Tabulka 5: Vybrané kódy s nejvyšším průměrným počtem zapsaných příčin smrti, muži, USA, 2018

Kód	Základní příčina smrti	PPK	‰
X44	Náhod. otrava j. a neurč. lék., léč., návyk. a biol. lát. a expoz. jej. působ.	5,7	10,2
C97	ZN mnohočetných samostatných (primárních) lokalizací	5,4	2,2
I13	Hypertenzní nemoc srdce a ledvin	5,4	2,8
Y83	Operace a j. oper. výk. jako příč. abn. reakce nebo pozd. kompl.	5,3	1,1
W80	Ucpání dýchacích cest vdechnutím nebo polknutím j. předmětů	5,0	1,2
I12	Postižení ledvin při hypertenzi	5,0	4,1
X59	Vystavení neurčeným faktorům	5,0	2,0
E87	Jiné poruchy tekutin, elektrolytů a acidobasické rovnováhy	4,8	1,8
E11	Diabetes mellitus 2. typu	4,7	11,6
E78	Poruchy metabolismu lipoproteinů a jiné lipidemie	4,7	5,1
X42	Náhodná otrava narkot. a psychodyslep. a expoz., NJ	4,6	14,5
A04	Jiné bakteriální střevní infekce	4,6	1,5
J10	Chřipka způsobená identifikovaným sezónním chřipkovým virem	4,6	1,8
B18	Chronická virová hepatitida	4,6	1,9
N28	Jiné poruchy ledviny a ureteru nezařazené jinde	4,6	1,7

Pozn.: ‰ = prevalence základní příčiny smrti na 1 000 obyvatel, tato poznámka platí i pro tabulky 6 a 7

Zdroj dat: NBER (2020), výpočet vlastní

Tabulka 6: Vybrané kódy s nejnižším průměrným počtem zapsaných příčin smrti, ženy, USA, 2018

Kód	Základní příčina smrti	PPK	%
R99	Jiné nepřesně určené a neurčené příčiny smrti	1,1	3,8
I46	Srdeční zástava	1,3	5,9
R09	Jiné příznaky a znaky týkající se oběhové a dýchací soustavy	1,4	1,1
R62	Nedostatek předpokládaného normálního fyziologického vývoje	1,4	1,7
P07	Poruchy v souvislosti se zkráceným trváním těh. a nízkou por. hmot. NJ	1,4	1,1
R54	Stáří (senilita)	1,5	2,1
C71	Zhoubný novotvar mozku	1,9	5,2
G12	Mišní svalová atrofie a příbuzné syndromy	2,1	2,3
C25	Zhoubný novotvar slinivky břišní	2,2	15,7
J96	Respirační selhání nezařazené jinde	2,2	4,9
C56	Zhoubný novotvar vaječníku	2,2	10,0
C23	Zhoubný novotvar žlučníku	2,3	1,1
I50	Selhání srdce	2,3	32,7
G31	Jiné degenerativní nemoci nervové soustavy nezařazené jinde	2,3	12,3
C16	Zhoubný novotvar žaludku	2,4	3,2

Tabulka 7: Vybrané kódy s nejnižším průměrným počtem zapsaných příčin smrti, muži, USA, 2018

Kód	Základní příčina smrti	PPK	%
R99	Jiné nepřesně určené a neurčené příčiny smrti	1,1	4,8
I46	Srdeční zástava	1,4	6,4
P07	Poruchy v souvislosti se zkráceným trváním těh. a nízkou por. hmot. NJ	1,4	1,5
C71	Zhoubný novotvar mozku	2,0	6,5
G12	Mišní svalová atrofie a příbuzné syndromy	2,2	2,6
J96	Respirační selhání nezařazené jinde	2,3	4,0
C25	Zhoubný novotvar slinivky břišní	2,3	15,9
I50	Selhání srdce	2,4	26,3
C79	Sekundární zhoubný novotvar jiných a neurčených lokalizací	2,4	1,8
I51	Komplikace a nepřesně určené a popsané nemoci srdce	2,4	7,3
C45	Mezoteliom [mesothelioma]	2,5	1,3
I70	Ateroskleróza	2,5	1,5
G31	Jiné degenerativní nemoci nervové soustavy nezařazené jinde	2,5	7,3
C20	Zhoubný novotvar konečníku	2,5	3,1
C18	Zhoubný novotvar tlustého střeva	2,6	14,7

Zdroj dat: NBER (2020), výpočet vlastní

K povšimnutí je, že i v tabulkách 6 a 7 přesahuje u řady příčin ukazatel PPK hodnotu dvě. Byly-li by v tabulkách uvedeny choroby bez omezení prevalence, byl by pak minimální PPK nižší. Tato charakteristika předznamenává, že prevalence uzlu do jisté míry souvisí s jeho PPK, a to ve smyslu přímé úměry. Tato vlastnost vyplývá na povrch ještě důrazněji při analýze vztahů mezi příčinami smrti za pomoci sítí a je opakovaně skloňována v různých kontextech dále v práci.

Kapitola 4

Použité metody a ukazatele

K vymezení sítí, v nichž aktéry představují choroby, a jejich analýze, může být přistoupeno řadou způsobů, z nichž výčet oněch nejčastěji uplatňovaných lze nalézt například v Fotouhi et al. (2018). Výběr metody se odvíjí nejen od povahy podkladových dat (počet aktérů, vazeb, menšinová dominance uzlů...), ale i od druhu konstruované sítě. V případě, že je cílem zkoumat vztahy mezi příčinami úmrtí, jež byly zapsány na HOÚ, hraje roli při volbě druhu dvojaká povaha uzlů pramenící z diferencí mezi částmi hlášení. I proto jsou v práci analyzovány příčiny z rozdílných částí hlášení odděleně. Jak bylo avizováno již v úvodu práce, z kódů uvedených v první části hlášení je tvořena orientovaná síť, v níž je rozlišován směr vazeb, z kódů v druhé části hlášení a ze základní příčiny smrti je tvořena neorientovaná síť, v níž vazby nemají směry. Bližší popis uplatněných postupů pro tvorbu a analýzu každé z těchto sítí je středobodem stávající kapitoly.

4.1 Vymezení orientované sítě

Kódy v první části hlášení jsou charakteristické předpokladem, že by měly mít mezi sebou vztah, z něhož lze rozlišovat „tu předchozí“ a „tu následnou“ příčinu smrti. Na tomto základě z nich byly utvářeny uspořádané dvojice podle logiky vyplňování HOÚ, která byla rozebírána v kapitole se zdroji dat. Platí, že za předpokladu, že by byly na HOÚ všechny příčiny smrti zaznamenány na správnou pozici, jsou níže formulovaná pravidla způsobem, jak rekonstruovat chorobné řetězce vedoucí ke smrti. Znějí následujícíce:

1. Důsledkem nemůže být příčina oficiálně stanovená jako základní.
 - A. Pokud zemřelý měl příčinu smrti shodnou s UCD zapsanou kdekoli v první části hlášení, byla vyjmuta z chorobného řetězce a zapsána do nového řádku s číslem vyšším než pět (čímž se dostala na vyšší řádek, než je nejvyšší možný v původních datech u každé osoby, čímž bylo ošetřeno, aby k tomuto kódu bylo možné dohledávat jen důsledky).
 - B. Pokud zemřelý neměl příčinu shodnou s UCD zapsanou kdekoli v první části hlášení, tak nebylo nic vyjímáno z chorobného řetězce a do nového nejvyššího řádku byla zapsána rovnou příčina shodná s oficiálně stanovenou UCD.

2. Pokud jsou dva kódy na po sobě jdoucích řádcích: kód na řádku s o jedno vyšším číslem evidujeme jako příčinu kódu na řádku s číslem o jedno menším, než je on sám (zpracování pravidla „due to“ na obrázku 1).
3. Pokud jsou kódy na stejném řádku, nemohou mít mezi sebou „příčinný vztah“, proto uijíme logiku:
 - A. Oba vedou k témuž kódu.
 - B. K oběma je vedeno od téhož kódu.
4. Pokud byly kódy zapsány na různých řádcích, z nichž jeden nebyl bezprostředně nad tím druhým, můžou být také analogicky utvořeny dvojice, ale pouze pokud mezi nimi nebyl zapsán žádný jiný kód (tímto bodem se řešily mezery vzniklé přestěhováním základní příčiny smrti, tj. následky aplikace bodu 1 těchto pravidel, a situace, kdy byly kódy zapsány s vynecháním řádků, např. jeden na 1. a druhý na 3.).
5. Smrt je důsledkem kódu uvedeného na nejnižším řádku. Každému zemřelému byl přidělen kód *smrt*, aby:
 - A. byla prokázána snaha nevynechat kompletně úmrtnost lidí, kteří měli uvedenou jen jedinou chorobu.
 - B. byl pokus aspoň částečně řešit možnou nekonzistentnost pramenící z utváření pouze dvojic, místo skupin vyšších řádů (např. je uvažována choroba B, s níž, když je zapříčiněna chorobou A, člověk nejčastěji umírá, když je ale B zapříčiněno chorobou C, není progresse ke smrti z B tak častá, nýbrž člověk umírá nejčastěji např. na chorobu, kterou zapříčiňuje právě B, tedy bez vazby na smrt by se B jevilo být jen tranzientním stavem, tj. nebylo by zřejmé, že s chorobou B lidé rovněž často umírají).

Tedy, pokud např. byly na hlášení čtyři kódy v první části hlášení a vyskytovaly se na 1., 2., 2., a 3. řádku, vzniklé dvojice by byly: 12, 12, 23, 23. Takto a podle ostatních kroků v bodech 1–5 byly utvořeny dvojice ze všech třímístných kódů, které se objevily vždy u jednoho člověka. Na této etapě bylo dospěno ke vztahům, které by se vyskytovaly v syré síti. Fotouhi et al. (2018) takto označuje systémy, v nichž nejsou kladeny žádné nároky na významnost vazeb. Biomedicínské sítě, mezi něž Toor a Chana (2020) řadí i sítě komorbidit, jsou však typické svou obšírností, počty vazeb v syré síti tohoto typu mnohdy přesahují desetitisíce. Za účelem vyzdvihnutí nejrelevantnějších topologických vlastností a architektury rozsáhlých sítí bývá běžnou praxí uplatňování sparsifikačních metod, představení kterých lze nalézt v Hamman et al. (2016). Při výběru nejvhodnějšího postupu v tomto směru je nutné opět vycházet z povahy konkrétních síťových dat. Na to, aby byla v daném případě vazba uznána jako signifikantní a ponechána v systému, byla zavedena kombinace pravidel¹¹:

1. ϕ koeficient je signifikantní na hladině spolehlivosti 1 % (podle Hidalgo et al., 2009)
2. poměr šancí mít vůči nemít důsledek, když má zemřelý člověk danou příčinu, signifikantně převyšuje pět (podle Kim et al., 2016)

¹¹ Vysvětlení, proč zrovna taková kombinace zrovna takových pravidel je podáno v navazující kapitole s validací postupů.

3. jedná se o dvojici, kde důsledkem je „smrt“ a zároveň je příčinou uzel s prevalencí nad 1 %, přičemž podíl osob přesouvajících se z této příčiny do smrti je vyšší než podíl osob přecházející z ní do kteréhokoli jiného chorobného stavu
4. absolutní četnost počtu zemřelých s oběma chorobami náleží k 25 % nejčastějších dvojic (částečně podle Kim et al., 2016)

Tyto podmínky vymezují modelovaný systém, přičemž mezi body 1–3 platí logická spojka „nebo“ (stačí, aby byly vazby aspoň v jednom hledisku z těchto tří vyhovující). Bod čtyři byl dodatečně aplikován na vazby, které splňovaly aspoň jednu z prvních tří podmínek. Opodstatnění použití tohoto způsobu vymezení, v rámci čehož i odpověď na otázku, proč je vhodné metodiky kombinovat, je uvedeno v kapitole 5. V stávající kapitole jsou níže popsány jen způsoby výpočtů koeficientů, statistik a měr uvedených v bodech 1–4.

Hidalgo et al. (2009) používají pro výpočet síly vazby ϕ koeficient a relativní riziko (RR). Píší, že prvně zmíněná míra dobře podchycuje asociace pouze mezi chorobami s podobnou prevalencí, zatímco RR podhodnocuje asociace mezi vysoce prevalentními chorobami, ale nadhodnocuje asociace mezi vzácnými. Vyvozují, že společné použití měr přináší částečné řešení zkruslení obsažené v každé z nich. Konstruuji tedy dvě sítě s tím, že „obě zobrazují signifikantní vazby, jen každá na trochu odlišném měřítku prevalence“ (Hidalgo et al., 2009, s. 5). Cílem zde je však postihnout signifikantní vazby v rámci pouze jediného systému. Proto je podle této studie spočítáno pouze ϕ , a to za použití vzorce:

$$\phi_{ij} = \frac{N_{ij}N_{-i-j} - N_{i-j}N_{-ij}}{\sqrt{N_i N_{-i} N_j N_{-j}}}$$

Kde indexovaná N znamenají počty příslušející polím čtyřpolní tabulky:

Tabulka 8: Značení subpopulací podle toho, jaké nesou znaky

	Má důsledek j	Nemá důsledek j	Celkem
Má příčinu i	N_{ij}	N_{i-j}	N_i
Nemá příčinu i	N_{-ij}	N_{-i-j}	N_{-i}
Celkem	N_j	N_{-j}	N

Zdroj: Hidalgo et al. (2009)

Signifikance koeficientu byla spočtena t-testem, kdy pro výpočet t-statistiky je podle Hidalgo et al. (2009) uplatňován vzorec:

$$t = \frac{\phi\sqrt{n-2}}{\sqrt{1-\phi^2}}$$

Kde n je vyšší hodnota z N_i a N_j , která je podle Hidalgo et al. (2009) volena místo obvyklého N , protože uvažování všech zemřelých by podle něj vedlo k většímu počtu vazeb jevících se jako významné, ač by nemusely být silnými prediktory. Signifikantní jsou ty hodnoty t-statistiky, které přesahují, nebo se aspoň rovnají 2,56 (1 % spolehlivosti).

Přičemž je testováno, zda jsou koeficienty odlišné od nuly, což je běžným postupem v sítích komorbidit, neboť výše koeficientů je ovlivněna frekvencí výskytu choroby i a j tvořící pár, více podrobněji v kapitole 5.

Podmínka pro bod dvě uvedená v pravidlech pro sparsifikaci, je použita jako ve studii Kim et al. (2016). Autoři, poté co stanoví signifikanci vazeb pomocí χ_2 (mezi kterýmžto a ϕ existuje vztah¹², viz také Chmiel et al., 2014), uvažují jako nejdůležitější vazby ty, v nichž poměr šancí (OR) přesahuje hodnotu pět. Užívají vzorec:

$$OR = \frac{\frac{N_{i,j}}{N_{\neg ij}}}{\frac{N_{i\neg j}}{N_{\neg i\neg j}}}$$

Pozn.: Symboly ve vzorcích jsou definovány v tabulce 8.

Interval spolehlivosti byl spočten podle Katz et al. (1978):

$$CI = e^{(\ln(OR) \pm 2,56 \times \sqrt{\frac{1}{N_{ij}} + \frac{1}{N_{\neg ij}} + \frac{1}{N_{i\neg j}} + \frac{1}{N_{\neg i\neg j}}})}$$

Dvě uvedená pravidla sama o sobě nestačila, protože použití ϕ a OR řeší pouze částečně problém v párech, kdy je jeden uzel velice častý a druhý málo, neboť vztahy mezi takovými uzly nejsou vyřešeny komplementárním užitím ϕ a OR (nebo RR, důvod opět v kapitole 5), což plyne z tezí Fotouhi et al. (2018) nebo z Hidalgo et al. (2009) a citovaných i zde na předchozích stranách. Toto se nejzřetelněji promítá do párů, v nichž důsledkem je *smrt*, a to kvůli existenci prázdného pole v kontingenční tabulce v buňce s počtem osob nemajících tento důsledek. Pravidlo tři bylo tedy použito pouze na dvojice příčiny a kódu pro *smrt*, protože do výpočtu podmíněné pravděpodobnosti nevstupuje $N_{\neg j}$, ani $N_{i\neg j}$, ani $N_{\neg i\neg j}$, proto nebude ve všech případech vycházet nulová, jako by tomu bylo při výpočtu ϕ , kde počty osob nemajících důsledek do výpočtu vstupují.

Princip vpravení *smrti* spočíval v identifikaci uzlů, z nichž zemřelé osoby přechází do umělé příčiny smrti *smrt* častěji než do kterékoli jiné choroby, a to z choroby, kterou mají na HOÚ uvedenou na bezprostředně vyšším řádku. Podle vzoru Chmiel et al. (2014), byla užitá podmíněná pravděpodobnost, nikoli však s podmínkou j , ale s i (v Chmiel et al., 2014 není konstruována orientovaná síť, proto není směřodonné uvažovat pořadí výskytu jevů):

$$p_{ij} = \frac{N_{i,j}}{N_i}$$

Když p_{ij} , kde j stojí pro *smrt*, bylo vyšší než součet p_{ij} přes všechna j nerovná *smrti*, byl vztah mezi *smrtí* a chorobou uznán jako relevantní, neboť, stejně jako již bylo uvedeno výše, toto znamenalo, že vyšší podíl osob přechází z chorobného stavu právě do *smrti*, nežli kamkoli jinam. Takových příčin bylo ale velmi mnoho (viz obrázek 5), přičemž valná většina z nich měla velmi nízkou individuální prevalenci a byla bez signifikantní vazby na kterýkoli další kód. Zároveň však počty výskytu dvojic mezi *smrtí* a příčinou tohoto charakteru hrály poměrně významnou roli v rozložení N_{ij} , takže by nebyly odfiltrovány v bodě 4 pro vymezení systému,

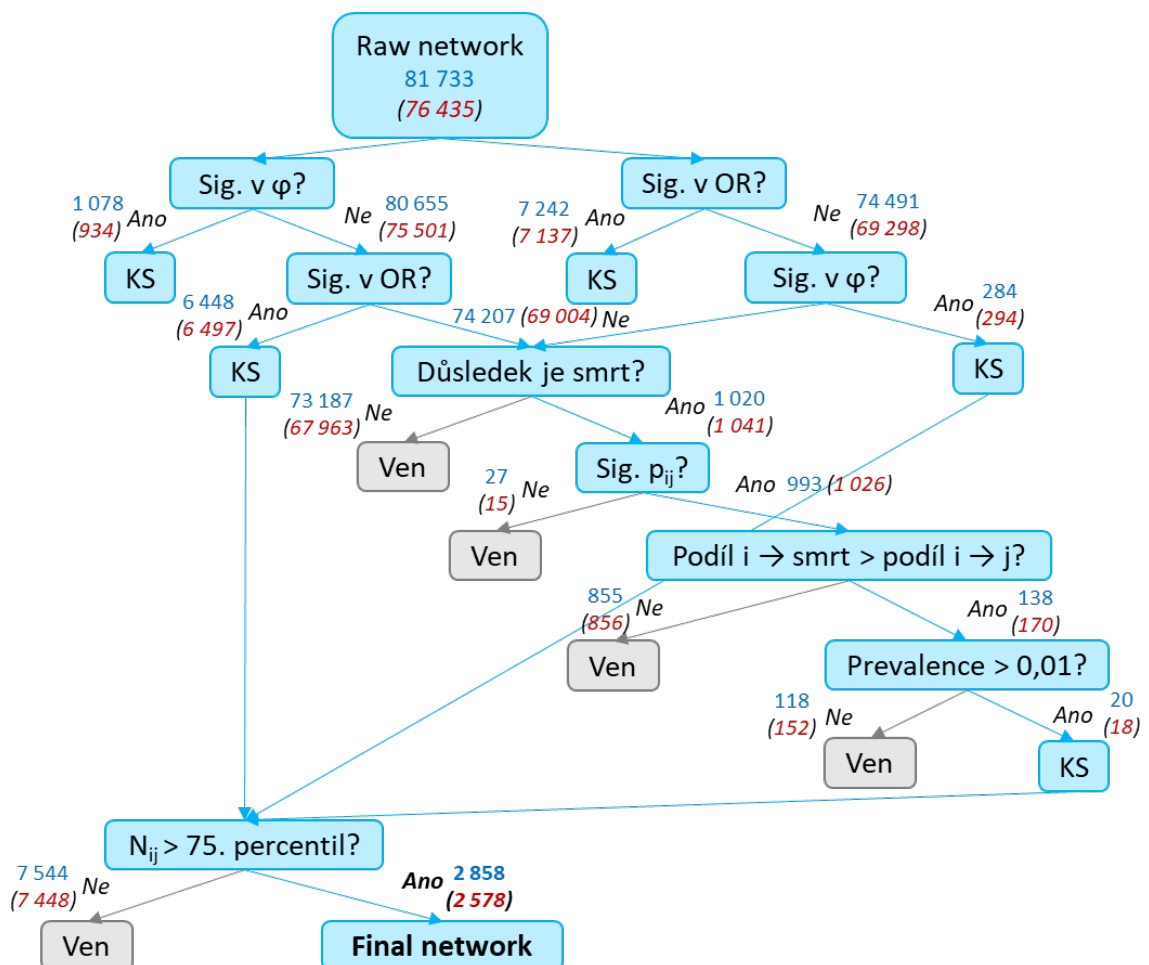
¹² $\chi_2 = N \times \phi^2$

přítom však nenesou informace prvořadě relevantní pro daný kontext, protože dvojice se *smrtí* nejsou ve své podstatě dvojicemi příčin smrti. Aby tedy sítí výrazně nedominovaly vazby mezi méně častými kódy a uměle implementovanou *smrtí*, byl vznesen požadavek na minimální prevalenci příčiny jsoucí ve dvojici se *smrtí*, sice vyšší než 1 %.

Počet vazeb splňujících podmínky 1–3 se u obou pohlaví pohyboval kolem osmi tisíc, což je zhruba desetina vazeb figurujících v syré síti. Z těchto osmi tisíc bylo ponecháno pouze 25 % nejčastěji se vyskytujících podle předlohy již citované studie Kim et al. (2016), kde však autor filtruje podle intenzity vazeb, nikoli jejich frekvence. Zde je však pracováno s neváženou sítí (vazba buď splňuje definované podmínky a tím pádem existuje, nebo je nesplňuje a neexistuje), což pramení z multiplexní povahy sítě. Ta totiž nevyklučuje situace, kdy by byla vazba přítomná pouze např. v φ vrstvě, ale již ne v OR, kvůli čemuž by nebylo možné stanovit jednotnou míru determinující sílu vztahu. Navíc, zjednodušení v ohledu ignorace váhy hrany je řešením spíše korektnějším, nežli že by takovým nebylo, protože existuje zkreslenost v síle asociací (znovu kapitola 5).

Na obrázku 5 je rekapitulován výše uvedený postup. Implikace použitého způsobu vymezení systému jsou rozebírány v navazující části práce. Odpověď na otázku, „*jak by se systém změnil, kdyby byly bývaly použity jiné podmínky pro jeho vymezení*“ je hledána v kapitole věnované validaci zvoleného postupu, na níž již bylo nesčetněkrát odkazováno výše.

Obrázek 5: Vývojový diagram pro orientovanou síť příčin smrti, USA, 2018



Pozn.: KS = kompletní síť, modrá čísla jsou počty vazeb u mužů, červená jsou počty vazeb u žen

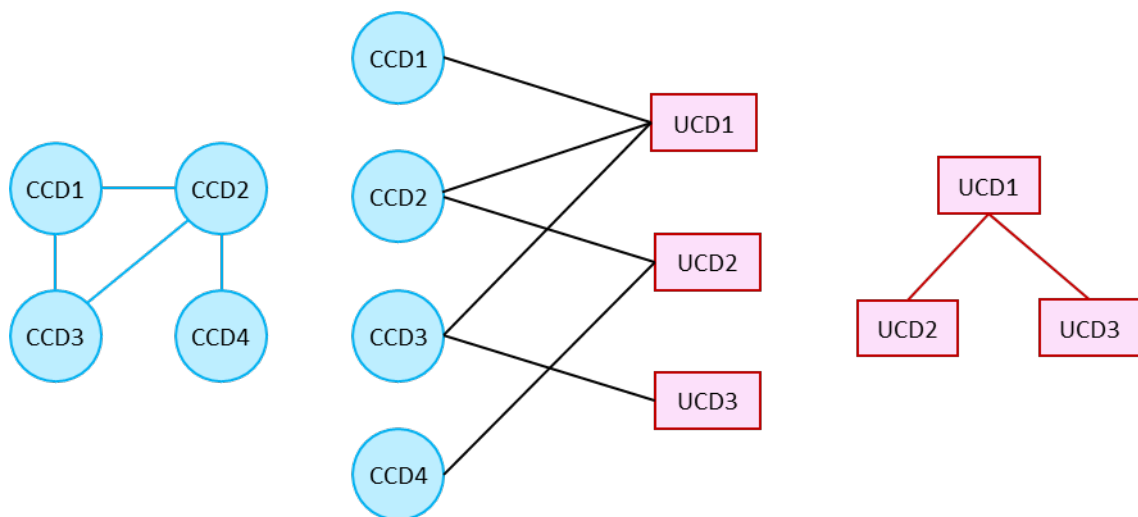
Zdroj: zpracování vlastní

4.2 Vymezení neorientované sítě

Část dvě HOÚ je vyhrazena pro stavy, které se u zemřelé osoby objevovaly současně se základní příčinou smrti, ale nijak se na ni co do přičinnosti nevázály. Ve světle tohoto se může nabízet myšlenka, že užití síťové analýzy analogicky jako na první část hlášení nedává pro data z druhé smysl, pokud je cílem zkoumat nejen vzájemný vztah mezi oněmi přispívajícími příčinami, ale mezi přispívajícími a základními, jako v této práci. Jediné, co tyto přispívající a základní příčiny smrti spojuje, je, že se vyskytovaly u jednoho člověka zároveň, nikoli v návaznosti na sebe. I proto je zde pro analýzu vztahů mezi nimi použita síť neorientovaná a problém plynoucí z významové odlišnosti základní příčiny a přispívajících bude řešen užitím sítě bimodální.

Bimodální (bipartitní) síť je taková, „jejíž uzly lze rozdělit na dvě nespojité množiny U a V tak, že každá vazba spojuje U -uzel s V -uzlem“ (Pavlopoulos et al., 2018, s. 2). Jinými slovy, uzly z množiny U se na sebe váží, pokud mají vztah ke stejným uzlům z množiny V a totéž platí pro V uzly, tj. tyto se na sebe váží, pokud se jejich sousedé z množiny U neodlišují. Zde prvou množinu tvoří základní příčiny smrti, druhou přispívající choroby¹³. Výstupem jsou tedy dva grafy pro každé pohlaví, z nichž jeden zachycuje vazby mezi přispívajícími příčinami, které existují, pokud se tyto váží na stejnou základní příčinu smrti, druhý pak obsahuje vazby mezi UCD, které vyjadřují, že sousedi v této síti mívají často ve druhé části hlášení stejné přispívající choroby (obrázek 6). Toto umožňuje obejít situace, kdy by se na sebe vázaly choroby A a B, aniž by bylo rozlišitelné, zda je to kvůli tomu, že je lidé často mívají pospolu ve druhé části, nebo zda je mívají spolu kvůli tomu, že jedna je základní a druhá k ní přispívající.

Obrázek 6: Schéma bipartitního grafu



Pozn.: CCD jako přispívající příčina smrti, UCD jako základní příčina smrti; Smyslem tohoto obrázku je grafické vyjádření významu vztahů v bipartitní síti. Např. modrá hrana mezi CCD1 a CCD2 existuje, protože existuje černá hrana mezi CCD1 a UCD1 a zároveň mezi CCD2 a UCD1. Vazba mezi CCD1 a CCD2 tedy vyjadřuje, že tyto dvě CCD mají vazbu ke shodné příčině smrti do množiny UCD.

Zdroj: podle Goh a Choi (2012), zjednodušeno

¹³ Když je tedy dále v textu používáno slovní spojení „z opačných množin“, je tímto pro množinu UCD myšlena množina CCD, pro CCD je opačnou množinou naopak UCD.

Výše v této podkapitole byl představen obecný princip bimodálních sítí. Následující pasáž textu je věnována již konkrétně postupům užitým pro tvorbu neorientované sítě příčin smrti. Jak bylo zjištěno v jiných studiích, nebo i zde v explorační analýze, do druhé části HOÚ se v průměru píše méně kódů než do první, navíc jsou pro vytvoření neorientované sítě páry příčin smrti tvořeny pouze napříč množinami CCD a UCD, nikoli i v rámci pouze každé z nich samostatně. Proto je také celkový počet dvojic, které je možno z CCD a UCD vytvořit, menší, pročež formulace pravidel pro vymezení systému a sparsifikační postupy jsou prostší.

Pro výpočet síly vztahu je užit ukazatel asociace mezi příčinami smrti¹⁴ (CDAI). Jsou ponechány vazby signifikantní mezi dvojicemi kódů, které se vyskytly dostatečně často, tedy, analogicky jako u orientované sítě, patří k 10 % nebo 25 % nejčastějších dvojic (konkrétněji dále). Vzorec pro výpočet CDAI je převzat z Human Cause of Death Database (HCDD, 2017):

$$CDAI_{u,c} = \frac{\sum_x \frac{d_x^{u,c}}{\bar{d}_x^u} \times \bar{d}_x}{\sum_x \frac{d_x^c}{\bar{d}_x} \times \bar{d}_x}$$

Kde $d_x^{u,c}$ je počet zemřelých lidí ve věku x se základní příčinou u a přispívající c .

Kde \bar{d}_x^u je počet zemřelých ve věku x s příčinou u jako základní.

Kde d_x^c je počet zemřelých lidí ve věku x s příčinou c jako přispívající (bez ohledu na základní).

Kde \bar{d}_x je celkový počet zemřelých ve věku x (bez ohledu na základní příčinu).

Kde \bar{d}_x je standard (zde v práci je použita standardní věková struktura zemřelých pro evropskou populaci).

Standardizací se odstraňuje efekt rozdílného věkového rozložení v subpopulacích zemřelých podle základní příčiny smrti (Dessesquelles et al., 2010). Ve své nestandardizované podobě CDAI odpovídá poměru pozorovaných a za předpokladu nezávislosti očekávaných četností ve čtyřpolní tabulce, jako byla uvedena u vzorce pro ϕ (podle značení v ní lze vzorec pro CDAI bez standardu zjednodušit na $(N_{ij}/(N_i \times N_j)) \times N$). Proto, když CDAI vychází nesignifikantně odlišen od jedné (sta), lze docházet k interpretaci, že pravděpodobnost zemřít s oběma chorobami je táž jako pravděpodobnost zemřít s jednou, či druhou. Dessesquelles et al. (2010) konstatují, že „CDAI je poměr standardizované pozorované míry prevalence ku standardizované očekávané míře prevalence“ (Dessesquelles et al., 2010, s. 12).

Zda ve vztahu mezi chorobami existuje signifikantní odlišnost od jedné (sta), bývá podle HCDD dopočítáváno pomocí intervalů spolehlivosti pro CDAI, které HCDD (2017, s. 2) definuje jako:

$$CI_{95\%}(CDAI) = \exp(\log(CDAI) \pm 2,56 \times SE(\log(CDAI)))$$

Kde $SE(\log(CDAI))$ je směrodatná odchylka, která se dopočítává podle HCDD (2017, s. 2) vzorcem:

$$SE(\log(CDAI)) = \sqrt{\frac{\sum_x \left(\frac{\bar{d}_x}{\sum_x \bar{d}_x} \right)^2 \times \frac{d_x^{u,c}}{(\bar{d}_x^u)^2}}{\left(\sum_x \frac{d_x^{u,c}}{\bar{d}_x^u} \times \frac{\bar{d}_x}{\sum_x \bar{d}_x} \right)^2}}$$

¹⁴ Cause of death association indicator

Na základě výsledků explorační analýzy lze konstatovat, že zhruba každá druhá zapsaná přispívající příčina smrti náleží k oněm uvedeným v tabulkách s nejčastějšími kódy (tabulky 1 a 2). Toto má za důsledek, že bimodální grafy konstruované standardními postupy popsány výše bývají silně zahlceny velkým množstvím vazeb, a to i když (nebo rovněž „protože“) je uvažováno pouze 10 % nejčastěji se vyskytujících párů základní a přispívající příčiny. Silně nerovnoměrné rozložení prevalencí v obou těchto typech množin uzlů, z nichž bimodální graf vzniká, totiž předznamenává, že se uzly z každé z množin budou vázat na spíše menší počet různých uzlů z množiny opačné, čímž bude narůstat podíl uzlů, které mají aspoň jednoho souseda společného, v důsledku čehož se stupně takové bimodální sítě budou pohybovat ve výrazně vyšších řádech, než tomu bylo např. u zde konstruované orientované sítě. Tudíž bimodální graf utvářený standardními postupy implementovanými v programu *Gephi 0.9.2*, v prostředí kteréhož byly všechny sítě vytvářeny, generuje z dvojic mezi přispívajícími a základními příčinami smrti síť s velmi vysokou hustotou. Konkrétně například síť přispívajících příčin smrti u mužů by obsahovala o něco více než 10 % všech možných vazeb, které v neorientované síti mohou existovat (za předpokladu uvažování všech signifikantních vazeb z hlediska CDAI a omezení z hlediska častosti dvojice). Při počtu uzlů přibližně 500 toto činí více než 20 000 vazeb mezi přispívajícími chorobami. U žen i množin základních příčin nastává podobné zahlcení.

Proto se nabízí postup tvorby bipartitního grafu modifikovat, a to v první řadě nastavením minimálního počtu společných sousedů z protilehlé množiny, kterým musí příčiny z téže množiny disponovat, aby mohly společně sousedit. Nastavení prahu v tomto smyslu však vede ke grafu, jemuž dominují dvojice mezi chorobami, kde jedna má sousedů mnoho a druhá málo. Disparita v počtu sousedů zkrátka generuje prostor pro shodu v „aspoň nějakých vazbách“. Za účelem minimalizace tohoto efektu lze nastavovat minimální počet společných sousedů relativně vzhledem k celkovému počtu a uvažovat pouze příčiny, které mají tento celkový počet sousedů v protilehlé množině vyšší než určitý stanovený práh. Uvedené řešení však není součástí *Gephi 0.9.2*, pročež je v následujícím odstavci shrnut postup, kterým bylo dojito k dopočtu datové matice, na jejímž základě bylo možné vytvořit bipartitní graf aplikací standardního algoritmu, který v *Gephi 0.9.2* generuje síť unimodální.

Zprvu byla datová matice obsahující vypočtený ukazatel CDAI a komponenty, které jsou pro jeho výpočet potřebné, redukována na pouze signifikantní vazby, načež byla přeskládána na matici, jejíž řádky obsahovaly kódy pro přispívající příčiny smrti (CCD) a sloupce kódy pro základní příčiny smrti (UCD). Buňky této matice byly vyplněny počty osob zemřelých s danou dvojicí kódů. Tyto počty byly přepsány na 1, pokud počet lidí zemřelých s onou dvojicí přesahoval 75. percentil, na nulu jinak. U tohoto „modifikovaného“ bipartitního grafu je uvažováno jen 25 % nejčastějších dvojic mezi CCD a UCD, oněch výše zmíněných 10 % platí pro standardní verzi bipartitního grafu.

Matice v této podobě (na obrázku 7 označena A) byla posléze transponována a dále obě matice společně vynásobeny (obrázek 7). Výsledkem součinu byla matice, v jejíchž řádcích i sloupcích byly uvedeny přispívající příčiny smrti (na obrázku 7 označena AA^T). Diagonála matice AA^T nesla informaci o celkovém počtu základních příčin smrti, na něž se dané přispívající váží. Pole mimo diagonálu AA^T obsahovala počet shodných sousedů z množiny UCD pro přispívající příčinu smrti i (CCD _{i}) a přispívající příčinu smrti j (CCD _{j}).

V matici AA^T byly následně uplatněny parametry pro shodu v sousedstvu (aspoň polovina sousedů¹⁵ musí být u příčin z téže množiny nerozdílná) i pro minimální počet sousedů (musel ležet nad mediánem). Celý postup byl následně zopakován z pohledu druhé množiny, tj. původní matice A neobsahovala řádky s přispívajícími příčinami, ale se základními. Sloupce v ní byly tudíž sestaveny z oněch přispívajících, tj. matice AA^T sestávala z vektorů základních příčin smrti, nikoli přispívajících.

Obrázek 7: Maticový počet na úpravu algoritmu pro produkci dat na bimodální síť

4.3 Síťové ukazatele

V předcházejících podkapitolách byly popsány způsoby vymezení každé z modelovaných sítí. Poslední podkapitola v této části textu obsahuje popis metod, které jsou aplikovány při analýze vymezených systémů. Níže popisované metody SNA jsou použitelné jak pro orientovanou, tak pro neorientovanou bimodální síť, avšak zde jsou jimi analyzovány jen první zmíněné, protože existují limity z hlediska rozsahu textu. Tedy, analýza bipartitních grafů je pojata v této práci spíše velmi okrajově.

Sítě lze analyzovat na globální nebo komunitní úrovni, dále na úrovni uzlů a vazeb (Borgatti et al., 2009). Jako ve kterékoli jiné statistické metodě, i zde existuje pestrá škála ukazatelů, jimiž lze systémy na každé úrovni popisovat (viz např. Barabási a Pósfai 2016, Khokhar 2015). Specifikum síťových ukazatelů však tkví v jejich relativně nepřímocharé interpretaci. Ta plyne v první řadě ze skutečnosti, že ač řada metrik vyjadřujících sílu rozličných vztahů v síti vychází po vzoru klasických korelačních koeficientů v intervalu $(-1, 1)$, tak nula neznačí, že má síť náhodný charakter (Blondel et al., 2008). Dalším specifikem je neaplikovatelnost klasických statistických testů, která vyplývá z porušeného fundamentálního předpokladu nezávislosti pozorování. Neméně významný je i mnohdy chybějící kontext pro interpretaci ukazatelů, protože, jak je tomu i v této práci, další systém vymezený podle stejných podmínek neexistuje a síť je tvořena pouze za jedinou populaci v jediný rok. Právě tato skutečnost byla stěžejní při výběru ukazatelů i přístupů k analýze vůbec. Je také důvodem, proč není toliko kladen zřetel na základní deskriptivní síťové metriky, jejichž výčet je k dispozici například v Barabási a Pósfai (2016) nebo v Ding (2011).

V části s výsledky je nejprve analyzována síť na úrovni uzlů a je dvěma úhly pohledu zkoumáno, jaké hrají uzly role. Jsou použity čtyři metriky, sice stupeň (*degree*) a mezilehlost uzlu (*betweenness*) (prvý úhel pohledu), skór (váha) autority (*authority score*) a hub-skór (váha) (*hub score*) (druhý úhel pohledu). Všechny svým způsobem vypovídají o důležitosti jednotlivých aktérů v síti.

Stupeň (D) značí „počet hran, které jsou incidentní na daném uzlu“ (Khokhar, 2015, s. 118). V orientovaných sítích bývá rozlišován stupeň dovnitř (inD), jenž je v konkrétním případě roven počtu signifikantních příčin vedoucích k danému uzlu, který je jim tudíž důsledkem, a stupeň ven (outD), který odpovídá počtu uzlů, jimž je daný uzel jako signifikantní příčina. Všechny druhy stupňů mohou být užity i k výpočtu globálního ukazatele, klasicky průměrného stupně sítě (dovnitř i ven).

Vedle stupňů uzlů (dále po vzoru literatury pouze „stupňů“) je jiným základním ukazatelem centrality i již zmíněná mezilehlost. Ta je ukazatelem důležitosti proto, že vysoký skór mezilehlosti (CB) uzlu indikuje vysoký počet nejkratších cest procházejících tímto uzlem. Ač tedy uzel s vysokou mezilehlostí nemusí mít vysoký stupeň, jeho odebráním by síť mohla být rozbita na menší, nepropojené komponenty.

Nerghes a Lee (2015) definují CB jako „součet podílů nejkratších cest v nichž daný uzel leží (...) na počtu všech pozorovaných nejkratších cest pro každý pár uzlů“ (Nerghes a Lee, 2015, s. 7). Výpočet se provádí podle algoritmu Brandes (2001). Tento zprvu vytvoří všechny možné dvojice uzlů v síti, následně nalezne nejkratší cestu mezi těmito dvěma uzly, potom zaznamená uzly ležící v této cestě, načech spočítá v kolika cestách leží uzel, pro který je skór počítán. Standardní vzorec pro výpočet CB je proveden podle Freeman (1979) v Brandes (2001):

$$CB(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

Kde $\sigma_{st}(v)$ je počet nejkratších cest z uzlu s do uzlu t , v nichž leží uzel v .

Kde σ_{st} je počet nejkratších cest z uzlu s do uzlu t .

Nabízí se otázka, co je míněno pojmem nejkratší cesta. Jedná se o hojně používanou metriku v sítích sociálních nebo geografických, k jejímuž odvození opět není přístupováno striktně univerzálně. Pro účel rozumění algoritmu na výpočet CB se může zdát být dostačující informace, že nejkratší cestou mezi dvěma uzly je taková, v rámci níž je počet skoků mezi uzly, které obsahuje, nejmenší, přičemž v konkrétním případě se neopomíjí orientace hrany mezi sousedními uzly.

Míru mezilehlosti lze počítat nejen pro uzly, ale také pro hrany v grafu. Mezilehlosti hran je užito i níže při detekci komunit podle Girvan a Newman (2002) a jedná se o zcela analogický ukazatel, i počítaný stejným principem, jako CB, jen do procesu výpočtu vstupuje místo množiny s uzly množina s hranami.

Společně tvoří D a CB základ pro explorační roli aktérů v síti podle Nerghes a Lee (2015). Tito definují čtyři typy rolí vymezených metrikami D a CB (obrázek 8). Prvním druhem jsou kódy mající vysoké D i CB, jež jsou označeny jako globálně centrální (*globally central*), protože nejenže pohlcují mnoho vazeb v síti, rovněž jimi prochází i řada nejkratších cest, k čemuž potřebují mít poměrně vyrovnané inD a outD, jinak by totiž stály pouze na počátku, resp. konci nejkratších cest, jejichž počet by se odvíjel pouze právě od inD, resp. outD.

Druhé jsou uzly lokálně centrální (*locally central*), které jsou charakteristické vysokým D, ale nízkým CB, což indikuje, že jsou významnými aktéry jen v určité části grafu. V orientovaných sítích jsou to mnohdy uzly, které mají vysoký stupeň dovnitř a nízký ven, což v síti přičiní úmrtí znamená, že jsou častými důsledky a důsledky mohou být proto, že jsou častými bezprostředními příčinami smrti, nebo jsou to i nemoci samy o sobě vysoce smrtelné. Mohou to být i kódy s nízkým stupněm dovnitř a vysokým ven, což o nich vypovídá, že bývají častými spouštěči chorobných řetězců vedoucích ke smrti, nebo jsou to například kódy, které musí osoba vyplňující hlášení o smrti na certifikát zapsat z různých „formálních důvodů“ (např. C97, viz část s výsledky analýzy).

Třetím druhem jsou uzly v roli vrátných (*gatekeepers*), které mají nízký D a vysokou CB. Jsou-li aktéři lidé, pak je vrátným taková osoba, která je strategicky umístěna tak, že má sice málo kontaktů, ale bez ní by byly narušeny toky z různých oblastí sítě. Vrátní v síti přičiní úmrtí jsou poměrně vzácní. Jejich status analogicky značí, že se jedná o příčiny přemostující odlišné klastry, které se mají v síti přičin smrti zřejmě tendenci tvořit. Přičemž, aby mohli přemostovat, musejí být jednomu klastru jako příčina a druhému jako důsledek (viz část s výsledky).

Posledním typem jsou uzly marginální (*marginal*), ty nabývající nízkých hodnot D i CB. V reálných sítích činí významný podíl všech entit, což souvisí s povahou pravděpodobnostního rozložení stupňů uzlů, podle kterého se v reálných sítích mají tendence utvářet menší počty uzlů s vysokým stupněm (D) a velké počty uzlů s nízkými stupni (Barabási a Pósfai, 2016). Jinými slovy, i v kontextu sítí lze uplatňovat Paretovo pravidlo, podle něhož 20 % uzlů stojí za 80 % vazbami systému. Tato teze je v analýze biomedicínských sítí dosti stěžejní, důvod čehož je blíže rozebrán v závěrečné kapitole práce (kapitola 9). Nyní, při rozboru rolí, je podstatné pouze to, že rozložení stupňů uzlů i skóre mezilehlosti je tímto významně sešikmené v prospěch marginálních uzlů, což vede autory Nergheš a Lee (2015) k tomu, že celou analýzu rolí staví na pořadí uzlů podle D a CB, nikoli na jejich hodnotách. Tento postup byl přejat také zde a uzly byly seřazeny podle výše stupně i mezilehlosti sestupně (měly-li stejný skóre, měly i stejné pořadí). Pořadí bylo normalizováno do intervalu (0, 100).

Obrázek 8: Typologie rolí podle stupně uzlu a mezilehlosti



Zdroj: Nergheš a Lee (2015), převráceno z hlediska řazení uzlů podle C a CB

Vzhled výsledného strukturálního prostoru vyobrazeného na obrázku 8 navádí na otázku ohledně způsobu vymezení hranic mezi kvadranty. Nergheš a Lee (2015) se jejich rigidně strohému vymezení vyhýbají dopočtem měř vyjadřujících, jak moc daný uzel leží, či neleží v daném kvadrantu. U každého uzlu dopočítávají čtyři hodnoty podle vzorců:

$$\begin{aligned}
 gc &= (100 - \text{norm}D^{\text{rank}}) \times (100 - \text{norm}CB^{\text{rank}}) \\
 lc &= (100 - \text{norm}D^{\text{rank}}) \times \text{norm}CB^{\text{rank}} \\
 gk &= \text{norm}D^{\text{rank}} \times (100 - \text{norm}CB^{\text{rank}}) \\
 m &= \text{norm}CB^{\text{rank}} \times \text{norm}D^{\text{rank}}
 \end{aligned}$$

Kde gc je globálně centrální, lc je lokálně centrální, gk je vrátný a m je marginální.

Kde norm znamená, že pořadí je normalizováno tak, aby maximální hodnota byla 100.

Kde D^{rank} a CB^{rank} jsou pořadí daného uzlu z hlediska těchto dvou metrik.

Lze rozumět, že maximální možná výše, které může uzel jak v každé z těchto měř zvlášť, tak i součtem přes všechny čtyři druhy dohromady, dosáhnout, je 100^2 . Čím vyšší hodnota míry, tím více neoddiskutovatelně uzel v daném kvadrantu leží a jedná se tak podle autorů o způsob obejít subjektivního názoru při vytyčování hranic v pořadí D a CB.

Již výše bylo uvedeno, že stupně a mezilehlost tvoří jednu z možných optik, pomocí kterých je možné odvodit role a poukázat tak na nejlivnější uzly v grafu. Druhým možným úhlem pohledu jsou skóry autorit (AS) a hub-skóry (HS). Čím jsou tyto vyšší, tím „důležitější“ uzel je, protože uzly, které na něj ukazují (v případě AS), nebo na které on sám ukazuje (v případě HS), jsou rovněž „důležitější“ (Khokar, 2015). Roli tudíž nehraje tolik stupeň uzlu samotného, ale spíše atributy jeho sousedů, díky čemuž je možné detekovat neduhy stupňových charakteristik uzlů. Ty spočívají ve skutečnosti, že uzel může mít i nejvyšší D v síti, ale to například pouze díky vazbám s marginálními uzly. Reálně tedy takový aktér nemusí být dobře ukotven v systému a pozbyvat tak na autoritě, což se právě skrze metriky AS a HS projeví. Konfrontace D a AS tedy umožňuje odlišovat uzly univerzálně populární od mocných, tedy těch s „důležitými sousedy“, nejen s jejich vysokým počtem.

Souhrnně, silná autorita může, ale rovněž nemusí mít vysoký stupeň, ale aby byla autoritou, musí se na ni vázat uzly globálně centrální. Hubs mají oproti tomu vazby k více uzlům s vysokou autoritou (Khokar, 2015). Výpočet AS a HS bývá prováděn algoritmem, jehož autorem je Kleinberg (1998), který uvádí pseudokód:

„Iterate(G, k)

G : a collection of n linked pages

k : a natural number

Let z denote the vector $(1, 1, 1 \dots 1) \in \mathbb{R}^n$

Set $x_0 := z$.

Set $y_0 := z$.

For $i = 1, 2, \dots, k$

Apply the I operation to (x_{i-1}, y_{i-1}) , obtaining new x -weights x'_i .

Apply the O operation to (x_{i-1}, y_{i-1}) , obtaining new y -weights y'_i .

Normalize x'_i , obtaining x_i .

Normalize y'_i , obtaining y_i .

End

Return (x_k, y_k) .“ (Kleinberg, 1998, s. 9–10)

Jedná se o iterativní algoritmus grafu G , který se bude opakovat k krát. Na začátku procesu je definován vektor z s počtem prvků n , kde n odpovídá počtu uzlů v systému. Je položeno x_0 a y_0 rovno z . Vektory x_0 a y_0 představují prvopočáteční AS a HS. Tedy, každému uzlu je dána shodná váha z hlediska autority a hub, která je rovna jedné. Následně je k krát opakována operace I, čímž jsou získávány nové a nové AS, a operace O, čímž jsou získávány též nové a nové HS. Operaci I, již probíhá update x_i vah (neboli AS „meziskóru“) Kleinberg (1998) definuje:

$$x^{(p)} \leftarrow \sum_{q:(q,p) \in E} y^{(q)}$$

Vzorcem je rozuměno, že AS uzlu p (levá strana) je součtem HS každého uzlu z těch, které na daný uzel p ukazují (pravá strana). Operaci O probíhá obdobně update y_i vah:

$$y^{(p)} \leftarrow \sum_{q:(p,q) \in E} x^{(q)}$$

Zápis vyjadřuje, že HS uzlu p (levá strana) je součtem AS všech uzlů, na které daný p ukazuje (pravá strana). Na konci každé iterace je provedena normalizace na obou druzích vah (skórů), a to vydělením každé váhy součtem skórů všech uzlů v síti.

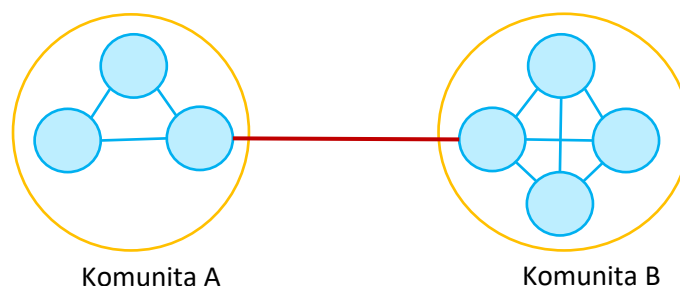
Z výše uvedeného vyplývá, že v první iteraci jsou AS každého uzlu rovné stupni dovnitř a HS každého uzlu jsou rovny stupni ven. Oba druhy skórů ve své normalizované podobě slouží jako vstup pro další iteraci. Kleinberg (1998) konstatuje, že ve většině sítí při dostatečně velkém k dojde ke stabilizaci AS i HS u všech uzlů. Uzly, hodnoty jejichž vah se stabilizují až při vyšších k , mají potenciál být autoritami nebo hubs. V práci je prostřednictvím těchto metrik poukázáno na jádro sítě a na její periferie. Navíc, jak bude diskutováno v kapitole s výsledky, lze s pomocí AS a HS odhalit skupiny příčin smrti, které vznikají právě tím, že se váží na stejné významné uzly.

Ostatně detekce skupin (komunit) je tématem dalších odstavců. Komunity jsou definovány jako skupiny uzlů, uvnitř nichž je výskyt vazeb pravděpodobnější než výskyt vazeb mezi nimi a uzly z jiných částí sítě (Borgatti et al., 2009). Ne ve všech reálných sítích mají entity tendence komunity utvářet a ukazatel vyjadřující míru, do níž lze síť dekomponovat na menší podgrafy je modularita. Jedná se o koeficient nabývající hodnot z intervalu $(-1, 1)$ a čím vyšší je, tím jednoznačnější rozklad sítě na menší grafy lze provést.

Postupům využitelným pro detekci komunit se věnuje Fortunato (2010). Mimo jiné zmiňuje, že se v principu značně podobají metodám shlukové analýzy. I pro výpočet modularity lze volit mezi divizivními metodami (kdy se jediná množina uzlů rozkládá postupně na menší) a aglomerativními (kdy se na počátku samostatné uzly postupně spájí se sobě blízkými). Jedním z prvních algoritmů pro detekci komunit byla, a i dnes je v mnohých programech s balíky na analýzu sítí implementovaná, divizivní metoda od Girvan a Newman (2002), která je aplikována i v této práci.

Girvan a Newman (2002) (G-N) píší, že svou metodu staví na nalézání periferií potenciálních komunit. Periferie podle nich tvoří hrany s vysokou mezilehlostí, neboť vysokou CB budou mít hrany, přes něž vede mnoho nejkratších cest, což ilustruje obrázek 9 vyjadřující, že při přechodu z kteréhokoli uzlu v shluku A do kteréhokoli z B je nutné vždy putovat přes červenou hranu.

Obrázek 9: Hrana s vysokou mezilehlostí



Pozn.: Pro spojení kteréhokoli uzlu z komunity A s kterýmkoli uzlem z komunity B je nutné inkorporovat do cesty červenou hranu. Podle G-N algoritmu bude mít červená hrana vysokou mezilehlost a tudíž bude ležet na okrajích různých podskupin.

Zdroj: zpracování vlastní

Algoritmus podle Girvan a Newman (2002) běží ve čtyřech krocích:

1. Spočítá mezilehlost pro všechny hrany v síti.
2. Odstraňuje hranu s nejvyšší mezilehlostí.
3. Přepočítá mezilehlost pro všechny hrany ovlivněné odebráním.
4. Zopakuj od kroku 2 až nezbydou žádné hrany.

To, které hrany jsou nakonec v grafu ponechány, se rozhoduje podle výše modularity, jež je simultánně dopočítávána. Iterace algoritmu, během níž je modularita grafu nejvyšší, je považována za nejlepší. Standardní vzorec pro výpočet modularity definují Blondel et al. (2008) jako:

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j)$$

Kde A_{ij} je váha hrany mezi i a j .

Kde m je rovné $\frac{1}{2} \sum_{i,j} A_{ij}$.

Kde k je součet vah hran napojených na i , resp. j . (inD, outD).

Kde c je označení komunity, k níž je uzel i , resp. j připojen (jedná se o funkci, jejímž úkolem je vrátit hodnotu jedna, pokud jsou už uzly ve stejné komunitě, neboli c_i je rovno c_j).

Khokar (2015), vysvětluje, že ve vzorci nad je srovnávána pravděpodobnost, že vazba mezi uzly existuje v rámci jednoho modulu (komunity), a pravděpodobnost, že by i v náhodném grafu vazba spadla přímo do téhož modulu.

Souhrnně psáno, celý postup na detekci komunit odhaluje hrany mezi uzly, které mají nejmenší pravděpodobnost skončit spolu v podgrafu a to podle jeho autorů „s vysokým stupněm úspěchu“ (Girvan a Newman, 2002, s. 7). Pravdou je, že výše uvedený vzorec pro výpočet modularity byl původně navržen pro aplikaci na síť vážené a neorientované. Nicméně Fortunato (2010) píše, že jej lze použít i na síť nevážené a rovněž argumentuje, že rozlišování orientace hrany ve výpočtech modularity nedává smysl. Takové stanovisko zastává, protože podle něj lze o komunitách hovořit již v momentě, kdy mezi uzly vazby pouze existují, bez ohledu na to, kdo je vysílajícím a kdo přijímajícím aktérem, protože mezi nimi nelze poukázat na toho, který je v jejich vztahu důležitější.

Kapitola 5

Validace postupů při vymezení systému

V této kapitole je zodpovězena otázka, proč bylo přistoupeno ke kombinacím postupů autorů při konstrukci orientovaného grafu a nebyla jen ryze následovaná některá z předlohových studií. Jelikož tedy u neorientované sítě nebyly kombinovány koeficienty asociace a byl užit jen CDAI, není této síti v stávající kapitole ani věnována pozornost.

Struktura kapitoly je taková, že po opodstatnění využití pravidel pro vymezení orientované sítě je vložena pasáž věnující se základním komorbiditním statistikám, protože v ní jsou doložena na datech tvrzení, která jsou důvody, proč člověk ten který krok při vymezení systému dělal.

5.1 Opodstatnění využití pravidel pro vymezení orientované sítě

V tabulce 9 jsou uvedeny dosud provedené studie vztahů mezi chorobami, pro něž je společný aspoň jejich počáteční krok, tj. podchycení vztahů mezi většinou třímístnými MKN kódy. U každé studie jsou v tabulce 9 vždy zapsány míry asociace použité při vymezení systému, resp. metoda v případě sítě příčin smrti podle Egidi et al. (2018). V tomto směru mohou být studie, které figurovaly jako vzory pro tuto práci, rozřazeny do třech skupin. V první použili pro nalezení asociací autoři společně k základní míře ϕ některou další, nejčastěji RR, v druhé aplikovali pouze RR, v třetí některý z jiných koeficientů, či metod. Hned se nabízí otázka, proč zde v práci nebyl sledován buď jen postup autorů užívajících ϕ a další komplementární míru z těch, které bývají užívány s ním v kombinaci, nebo ryze postup Kim et al. (2016) s pouze OR, aby bylo dosaženo asi žádoucího nekombinování přístupů.

Tabulka 9: Studie, na něž se bylo orientováno při rozmyslech o vymezení systému příčin smrti z části 1 hlášení o úmrtí

Autor	Rok	Užitý koeficient (metoda)	Typ
Chmiel et al.	2014	ϕ s úpravou, p_{ij}	multiplex („A“)
Hidalgo et al.	2009	ϕ , RR	2 networks
Kim et al.	2016	OR	uniplex
Folino et al.	2010	ϕ , RR	2 networks
Egidi et al.	2018	IPFP procedura	uniplex
Jensen et al.	2014	RR	uniplex
Jeong et al.	2017	RR s úpravou	uniplex
Kalgotra et al.	2017	Salton cosine index	uniplex

Pozn.: multiplex = výstupem je jediná síť z více vrstev, uniplex = výstupem je jediná síť z jedné vrstvy

Zdroj: zpracování vlastní

Pouhé použití OR vede k výsledkům, v nichž chybí vazby mezi velice častými příčinami smrti (např. chronická ischemická choroba srdeční (I25) a selhání srdce, chronická obstrukční plicní nemoc a respirační selhání, nebo I25 a srdeční zástava a řada dalších párů). Jedná se o dvojice, kde příčina může být signifikantním rizikovým faktorem pro důsledek ($OR > 1$), ale do výsledné sítě by se nedostala, protože podle Kim et al. (2016) jsou uvažovány pouze dvojice, jejichž $OR > 5$, kam jde obtížněji se dostat, jsou-li N_i a N_j velmi časté v populaci (viz následující oddíl). Snížení minimální hodnoty OR z 5 na 1 by nebylo řešením, protože by potom bylo stejně nutné se uchýlit k dalším sparsifikačním krokům, poněvadž by tímto značně vzrostl počet významných asociací v síti (v daném případě až třináásobně). A vzhledem k tomu, že cílem práce je analýza vztahů mezi chorobami, je snahou se průměrným stupněm sítě blížit např. výsledkům Hidalgo et al. (2009).

Protože tedy pouhé použití OR a nutnost nastavení jeho minimální hodnoty na 5 ústí v síť chudou na vazby mezi příčinami smrti s vysokou častostí, bylo společně s OR použito ϕ , které podchycuje velmi dobře vztahy mezi podobně prevalentními chorobami, tedy i mezi těmi velmi často se vyskytujícími (také doloženo v následujícím oddílu). Ponechány byly vazby signifikantní v aspoň jednom z těchto dvou koeficientů. V případě použití dvou měř asociace však podle předlohových studií v tabulce 9 bývá zvykem buď konstruovat síť dvě na odlišných škálách prevalence (Hidalgo et al., 2009; Folino et al., 2010), nebo jedinou, ale signifikantní v obou mírách zároveň (Chmiel et al., 2014). Cílem práce je konstrukce pouze jediné sítě, proto se nebylo uchýleno k postupům podle Hidalgo et al. (2009), ani Folino et al. (2010). Současně by však pravidlo „A“¹⁷ jako u Chmiel et al. (2014) v daném případě postrádalo smysl, neboť ϕ a OR jsou společně použita právě za tím účelem, že jedno kompenzuje nedostatky druhého a vzájemná konstelace množin signifikantních vazeb (viz dál) z hlediska každé z těchto měř indikuje, že pravidlo „A“ je krokem zpět.

Nabízí se otázka, proč se tedy již v počátku nerozhodnout pro míry jako u Chmiel et al. (2014) a tvořit multiplexní síť ryze podle nich. Evidentně by totiž na kombinaci měř ϕ a p_{ij} mohlo být použito pravidlo „A“, čili by nebylo nutné kombinovat přístupy.

¹⁷ Tímto se rozumí, že je tvořen průnik dvou vrstev sítě. Vrstvu sítě představují vazby signifikantní z hlediska jednoho koeficientu asociace.

Důvodem je, že součástí konstrukce sítě podle těchto autorů je provedení nezbytného sparsifikačního kroku, jehož princip má však přesah až do problematiky týkající se dynamiky sítí, což je oblast síťových analýz, která si žádá hlubší vhled a je nad rámec této práce. Užití pravidla „NEBO“ zůstalo tedy jediným smysl nepostrádajícím a realistickým postupem při vymezení systému.

Nicméně stále není vysvětleno, proč se tedy již v samotném úvodu nevydat cestou jako autoři, kteří užívají ϕ a RR, když by tímto mohlo být dosaženo asi žádoucího nekombinování vícero různých metodik a zároveň by mohl být kompenzován neduh každé z těchto měř, protože RR má podobné přednosti jako OR, jak konstatují Hidalgo et al. (2009). Důvody jsou dva.

Za prvé, hodnoty OR a RR se odlišují zejména v případech, kdy v kontingenčních tabulkách neexistují výraznější asymetrie (Viera, 2008). Pokud je jedna z měř rovna jedné, bude rovna jedné i druhá. Když je $RR > 1$, bude za předpokladu neexistence výraznějších asymetrií OR vycházet výše nežli RR (Zhang et al., 1998). To je vlastnost, která může navádět na myšlenku, že OR je v dané situaci, kdy je minimální hodnota stanovena na 5, výhodnější než RR.

Stále existuje však onen druhý (a stěžejnější) důvod, proč je zde vhodné OR místo RR. Podstatné je uvědomění si, že studie od autorů předlohových studií v tabulce 9 spojuje fakt, že ti používající RR pracují s daty longitudinálními, tj. se záznamy zemřelých pacientů, u nichž mohou z dat zjistit okamžik diagnostiky chorob. Oproti čemuž ve studii Kim et al. (2016) pracují autoři se vzorkem pojištěnců, o nichž je známo pouze jakými chorobami k analyzovanému roku trpěli, analogicky jako ve zde užitém datovém souboru zemřelých za rok 2018. Ranganathan et al. (2015) vyzdvihává, že užití relativního rizika není zcela korektní v nekohortních datech, neboť není známa opravdová exponovaná populace, která do výpočtu RR vstupuje. Odlišná povaha dat je tedy oním druhým důvodem upřednostnění OR místo RR zde v práci, nehledě na skutečnost, že by se podle předlohových studií nabízelo právě spíše RR.

Pravidla 1–3 pro vymezení systému, která byla v uplynulých odstavcích obhajována, platila pro všechny dvojice kromě párů, kde důsledkem byla smrt. Myšlenka, že by implementace umělého kódu pro smrt mohla být dalším možným způsobem obohacení síťové analýzy, pochází od Fotouhi et al. (2014). Píší, že „*dalším zajímavým směrem vpřed by mohlo být inkorporování smrti. Člověk by mohl přidat smrt jako novou chorobu. Očividně by stupeň ven byl roven nule, ale mohlo by být poučné studovat stupně dovnitř a pozici smrti v síti...*“ (Fotouhi et al., 2014, s. 33). Dosud však smrt inkorporována nebyla, důvodem čehož může být, že není možné s ní zacházet stejně, jako s ostatními kódy (neaplikovatelnost koeficientů asociace, viz metodická část). V předchozí kapitole je popsán způsob, jak bylo snahou *smrt* vpravit do sítě zde. Je podstatné napřímo uvést, že tento způsob oporu v literatuře nenachází. Však se díky němu v síti rýsují frekventované choroby, s nimiž většina lidí přímo umírá, aniž by měla na HOÚ zapsanou nějakou další chorobu. Může se jevit být ne očekávaným, že takových chorob není mnoho, jak mohlo již vyplynout z počtu vazeb v obrázku 5, nebo i jak bude komentováno v části s výsledky.

5.2 Základní statistiky komorbidit

V předchozí podkapitole byla užitá tvrzení ohledně vztahů mezi výší užitých měř asociací a frekvencemi výskytu chorob tvořících vícečetné příčiny smrti, i další tvrzení ohledně konstelace množin signifikantních vazeb OR a ϕ . Tato tvrzení jsou v nynější podkapitole podepřena základními statistickými charakteristikami sítí. Záměrem je „ilustrovat“ že pro konkrétní data platí výroky, které ovlivňovaly rozhodnutí o způsobu vymezení systému.

V tabulce 10 jsou uvedeny počty vazeb podle kategorií specifikovaných v prvních dvou sloupcích nalevo této tabulky a podle toho, zda se v síti vyskytují, či nikoliv. U obou pohlaví vstupuje do výsledné sítě pouhých necelých 4 %¹⁸ všech možných vazeb, které šlo z datasetu zemřelých poskládat podle v úvodu definovaných pravidel pro utváření dvojice příčin a důsledků. Většina vazeb, které se ve zde konstruované síti objevují, spadají do skupiny tři (tabulka 10), v níž jsou vazby jsoucí signifikantní z hlediska OR, ale nesignifikantní z ϕ . V této kategorii jsou také obsaženy dvojice chorob, které se v síti nevyskytují kvůli N_{ij} nižšímu než stanovený práh častosti. Vazby, které jsou významné v obou mírách se v síti objevují všechny, obdobně jako vazby, které jsou významné pouze v ϕ , kterých je zároveň co do počtu nejméně (tabulka 10). Jedná se o ty vazby, které by byly ztraceny, kdyby nebylo užito kombinací přístupů vícera autorů. Zároveň, kdyby bylo použito pravidla „A“, skončily by v síti povětšinou pouze vazby, které jsou signifikantní aspoň v ϕ , jichž je opět pomálu a lze zvědět, že nejsou těmi, které jsou společné sítím od autorů uvedených v tabulce 9.

Tabulka 10: Množiny vazeb podle signifikance v mírách užitých pro vymezení systému

Je vazba signifikantní?		Označení skupiny	Vyskytuje se vazba v síti?			
			Muži		Ženy	
ϕ	OR		Ano	Ne	Ano	Ne
sig.	sig.	Skupina 1	794 (0,97)	0 (0,00)	640 (0,84)	0 (0,00)
sig.	nesig.	Skupina 2	284 (0,35)	0 (0,00)	294 (0,38)	0 (0,00)
nesig.	sig.	Skupina 3	1 762 (2,16)	4 686 (5,73)	1 627 (2,13)	4 870 (6,37)
nesig.	nesig.	Skupina 4	18 (0,02)	74 189 (90,77)	17 (0,02)	68 987 (90,26)

Pozn.: V závorkách šikmě jsou procenta. Součty ve sloupcích „ano“ za každé pohlaví jsou rovný celkovému počtu vazeb ve finální síti, který byl již uveden na nejnižším patru vývojového diagramu v obrázku 5. Hlavní poznatek plynoucí z této tabulky je ten, že nejvíce vazeb se v síti vyskytuje právě díky sjednocení množin sig. vazeb z hlediska ϕ a OR a že vazeb v ϕ vrstvě by bylo pomálu.

Zdroj dat: NBER (2020), výpočet vlastní

Figury A, B, C a D na obrázku 10a jsou vytvořeny podle Hidalgo et al. (2009). Účelem je na nich a E a F (obrázek 10b) ilustrovat některá tvrzení užitá v předchozích odstavcích. Obrázky 10a a 10b jsou prezentovány pouze za ženy. Výsledky za muže v tomto směru se samozřejmě neodlišují.

Horní figury A s B vyobrazují rozložení ϕ , resp. OR podle toho, zda se vazba, sílu asociace kteréžto vyjadřují, v síti nachází, či nikoli. Lze vidět, že existují dvojice chorob, které mají mezi sebou stejně silnou asociaci, ale zatímco některé se v síti objevují, jiné ne. Jsou to choroby, jejichž OR se nachází bezprostředně nad hodnotou jedna.

¹⁸ Což je zřejmé i z obrázku 5: ve finální síti se u mužů vyskytuje 2 858 vazeb, u žen 2 578. Tato čísla jsou přibližně necelými čtyřmi procenty všech vazeb v syré síti, které jsou v obrázku 5 uvedeny na nejsvrchnějších patrech diagramu (za muže 81 733, za ženy 76 435).

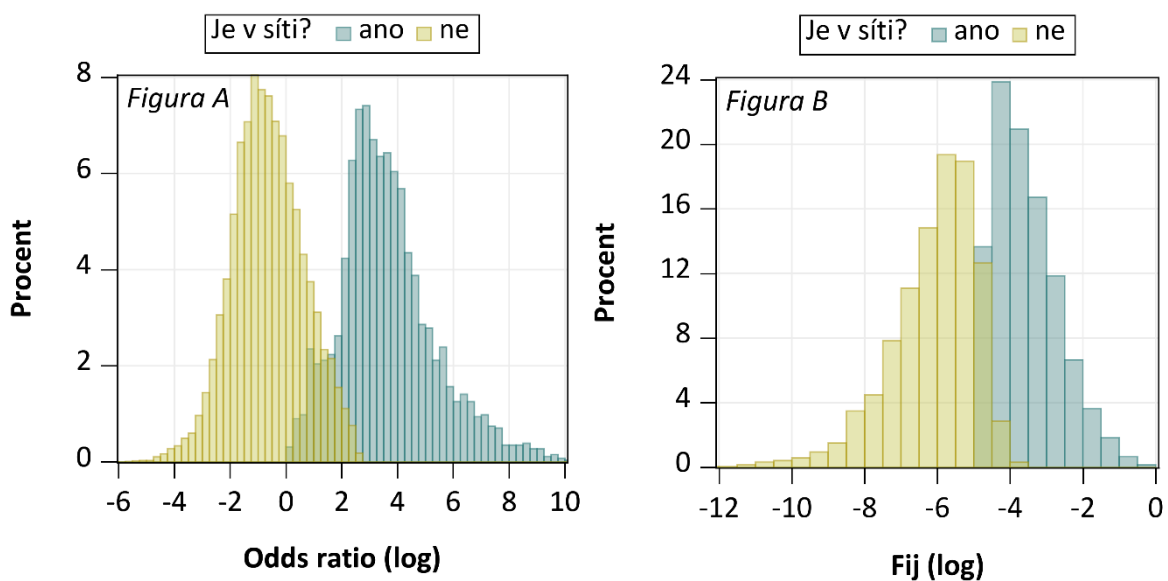
Mezi těmito dvojicemi jsou i ony vazby mezi velmi častými nemocemi, které se do sítě promítají právě pouze díky komplementárnímu použití dvou měr, ale také jsou to dvojice, které se sice neliší ve výši OR, ale ve své prevalenci (nejsou z hlediska N_{ij} v horním kvartilu).

Dále je v C po vzoru Hidalgo et al. (2009) ilustrováno rozložení prevalencí, avšak pouze těch chorob, které se v síti vyskytují. I v logaritmické transformaci je zřejmé, že nejvíce je chorob s nízkou prevalencí a čím je vyšší, tím je chorob méně. Což dokládá povýše i poníže mnohokrát opakovaná tvrzení odvolávající se na Paretovo pravidlo.

Ve figuře D je vztah mezi ϕ , OR a prevalencí výskytu komorbidit. Zatímco málo často se vyskytujícími dvojicemi by mohla být proložena regresní přímka s nevysokými reziduálními hodnotami, tak u dvojic s vyšší prevalencí jsou výsledovatelné větší diskrepance mezi OR a ϕ . Tyto diskrepance se stávají být většími s tím, jak se prevalence dvojice zvyšuje. Lze vidět, že dvojice, které bývají nejčastější, mívají poměrně nízká OR, což opět podporuje adekvátnost použitého postupu pro vymezení systému a podkládá některá výše užitá tvrzení.

Figura E má za cíl doložit zmiňovaná zkrácení v ϕ . Jsou proti sobě vynesena pořadí v prevalenci chorob i a j tvořících dvojici, datové body za tyto dvojice jsou podbarveny podle výše ϕ . Nejvyšší hodnoty ϕ jsou koncentrovány v pomyslné elipse kolem diagonály grafu jdoucí z origa do protilehlého rohu, což ilustruje skutečnost, že ϕ je nejvyšší u dvojic tvořených chorobami s podobným pořadím z hlediska jejich prevalence. Naopak nejnižší hodnoty ϕ tvoří ten obsáhlejší, kvůli čemuž také nápadnější, červený shluk, který se formuje kolem pomyslné přímky spojující poslední pořadí u proměnné N_i i N_j . Tedy, mezi dvojicemi tvořenými chorobami, z nichž jedna má vysoké pořadí z hlediska prevalence a druhá nízké, bývají zpravidla i nízká ϕ . Což ovšem neplatí pro míru OR (figura F). U ní existuje odlišná zákonitost, a to že vyšších hodnot OR dosahují zpravidla dvojice mezi méně častými chorobami. Stále však platí, že nejvyšší hodnoty koeficientů asociace nabývají páry chorob nejméně častých. Tento problém tedy nebyl vyřešen, ale na druhou stranu váhy hran dané hodnotami měr asociace nikde v práci nefigurují a jejich zkrácení tedy není v této situaci překážkou.

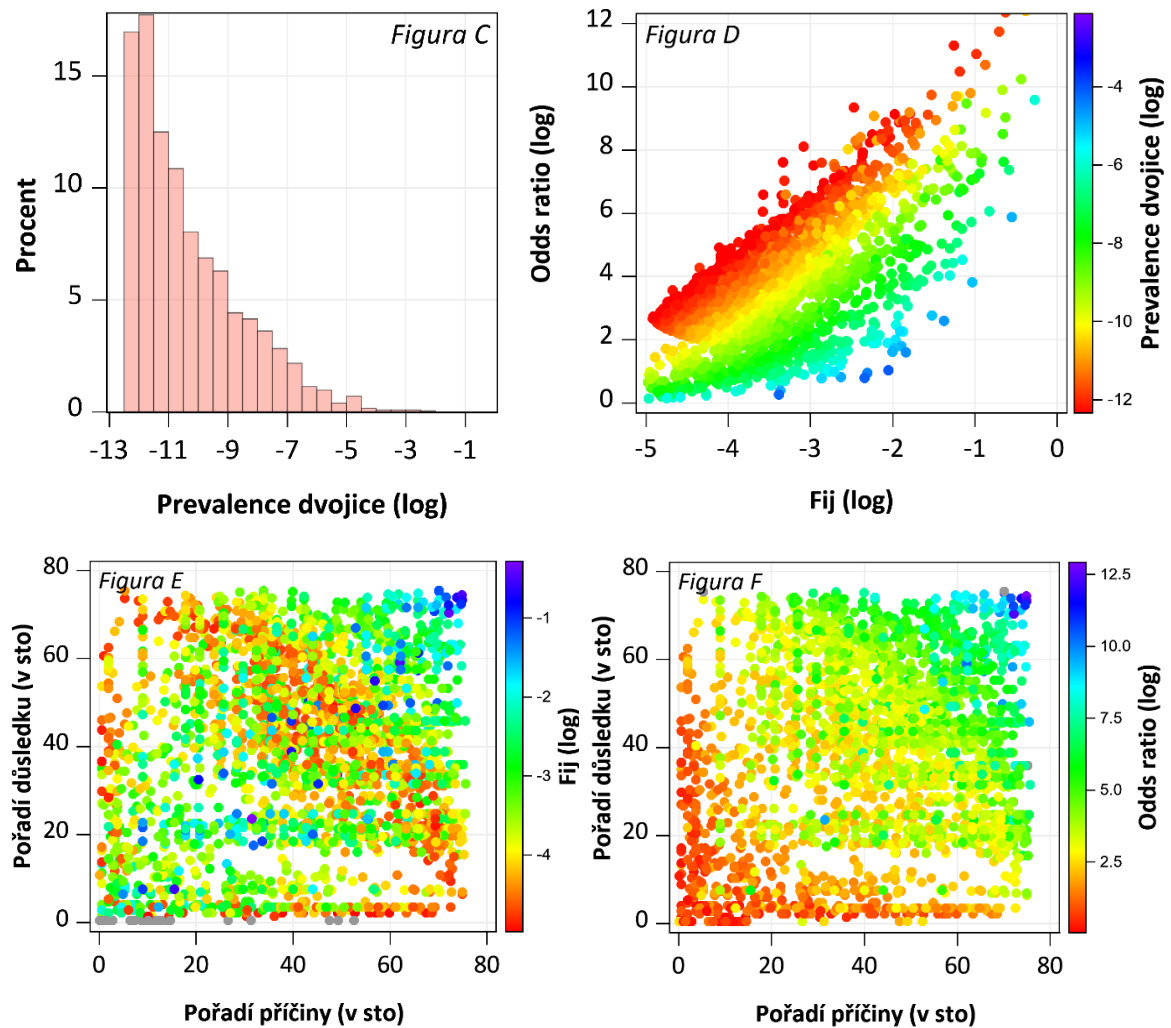
Obrázek 10a: Základní statistiky komorbidit, orientovaná síť, ženy, USA, 2018



Pozn.: viz obr. 10b

Zdroj dat: NBER (2020), zpracování vlastní

Obrázek 10b: Základní statistiky komorbidit, orientovaná síť, ženy, USA, 2018



Pozn.: Figura A = rozložení OR podle toho, zda se vazba v finální síti vyskytuje, či nikoli;
 Figura B = rozložení ϕ podle tohoto, zda se vazba v síti vyskytuje, či nikoli; Figura C = rozložení dvojic chorob podle prevalence; Figura D = vztah mezi OR a ϕ ovlivněný prevalencí dvojice; Figura E = vztah mezi rozdílem v pořadí ve frekvenci výskytu choroby na pozici příčiny a tímž pořadím choroby na pozici důsledku a výši koeficientu ϕ ; Figura F = vztah mezi rozdílem v pořadí ve frekvenci výskytu choroby na pozici příčiny a tímž pořadím choroby na pozici důsledku a výši koeficientu OR; $F_{ij} = \phi$

Zdroj dat: NBER (2020), výpočet vlastní

Kapitola 6

Výzkumné otázky

Ve studii Egidi et al. (2018) si autoři kladou za cíl zkoumat postavení uzlů v síti příčin smrti skrze různé míry centrality, mezi nimi i pomocí stupňů uzlů a jejich mezilehlosti. Píší, že je to právě SNA, jež umožňuje „identifikovat příčiny smrti, které hrají důležité role v procesu umírání a také vzorce úmrtnosti zahrnující více vzájemně se ovlivňujících příčin“ (Egidi et al., 2018, s. 23). Výzkumné otázky 5 a 6 (uvedeny níže) jsou tedy formulovány na základě této studie. Téma rolí má přesah i do otázek 1 a 4 a obou otázek týkajících se bipartitního grafu (7 a 8). V těchto však nejsou role zkoumány systematicky pomocí výše prezentovaných ukazatelů. Jsou formulovány obecněji a jejich smyslem je, jak také z jejich znění vyplývá, zjistit „jaký je systém“. Takové znění otázky totiž umožňuje interpretovat výsledky tak, jako to dělají autoři předlohových studií (např. Kim et al., 2016; Chmiel et al., 2014; Kalgotra et al., 2017) a objevovat shodné rysy zde konstruovaného systému a jiných existujících. Jedna výzkumná otázka míří na *smrt*. Uzel, který ji symbolizuje v síti, nemůže být brán jako „rovný“ ostatním, neboť pro páry, v nichž figurovala, platila odlišná pravidla vstupu do sítě. Proto je položena samostatná otázka 3. Otázka číslo 2 je docela stěžejní pro tuto práci. Analýza podskupin, které se poměrně jasně v orientované síti vytvářejí, umožňuje nahlédnout do architektury sítě, neboť systém jako komplet čítá přes 2 000 chorob a do takto komplexního systému může být „snadnější proniknout“, když je zkoumán po částech, nikoli jako celek. Navíc je analýza podskupin často řešeným aspektem v sítích komorbidit (např. Jones et al., 2021; Hidalgo et al., 2009) a jako dílčí téma je zařazeno i do studie Egidi et al. (2018).

Výzkumné otázky pro orientovanou síť jsou tedy formulovány:

1. Jaký je systém příčin smrti z hlediska složení?
2. Jaká je modularita grafů a jaké nemoci společně tvoří komunity?
3. Kdo jsou sousedé smrti?
4. Jaké vazby se nevyskytují v sítích opačného pohlaví?
5. Kdo jsou autority a hubs?
6. Jaké role hrají uzle v síti s z hlediska stupňů uzlů a jejich mezilehlosti?

Pro druhořadou neorientovanou bimodální síť bude snahou zodpovědět otázky 7 a 8:

7. Jaký se standardní bimodální graf?
8. Jaký je bimodální graf shodného sousedstva?

Kapitola 7

Výsledky analýzy orientované sítě příčin smrti

Finální orientovaná síť příčin smrti za populaci USA v roce 2018, odděleně podle pohlaví, je v příloze 1 a 2. Barvy uzlů jsou přiděleny na základě prvního písmena jejich kódu. Velikost uzlů je proporcionální k jejich stupni. Barvy hran se odvíjí od barev uzlů, na které se váží. Začínají-li oba sousední uzly tímž písmenem má hrana mezi nimi shodnou barvu s jejich, je-li každý uzel z jiné MKN kapitoly, je barva hrany mixem barev takových uzlů. Tady je důležité podotknout, že grafické zobrazení systému nemá žádný dopad na jeho metriky ani vlastnosti. I tak však nese význam. Algoritmy pro zobrazování sítě se totiž zpravidla řídí několika společnými pravidly: uzly, které se na sebe navzájem váží, se snaží lokalizovat co nejbližší k sobě a dobře v systému ukotvené uzly (centrální) se snaží umístit co nejhluběji do prostředu grafu a uzly marginální naopak na okraj.

Pomocí finálních grafů byly vytvořeny ještě další dvě sítě. Prvá z nich zobrazuje vazby, které mají ženy, ale nemají muži (příloha 4) a druhá naopak vazby pouze mužské, ale v grafu žen chybějící (příloha 3). Tyto grafy obsahují i dvojice uzlů s obrácenou přičinností mezi pohlavími, tj. pokud síť jednoho pohlaví obsahovala dvojici chorob AB a druhého jen BA, byly i tyto choroby uvažovány jako u opačného pohlaví chybějící. Všechny komentáře výše k vykreslení a rozložení platí pro všechna vyobrazení sítě v přílohách, tedy i pro síť s rozdíly mezi muži a ženami.

7.1 Jaký je systém příčin smrti?

Jak bylo zřejmé již z obrázku 5, síť žen obsahuje 791 uzlů a 2 578 hran, síť mužů 828 uzlů a 2 858 hran. Podíl těchto dvou čísel vyjadřuje průměrný stupeň sítě, který je u mužů lehce vyšší než u žen (3,45, resp. 3,25). Na první pohled je zřejmé, že součástí hlavně sítě žen jsou shluky uzlů, které jsou s jádrem sítě (největším klastrem) spojeny relativně slabě. V těchto slaběji přivázaných komponentách se shlukují příčiny typické pro úmrtnost specifických subpopulací. Konkrétně u obou pohlaví je zjevný klastr kojenecké úmrtnosti (P kódy) s nímž jsou provázány vrozené vady (Q). U žen se objevuje i homogenní skupina kódů z MKN kapitoly s kódy O sdružující příčiny smrti spjaté s těhotenstvím, porodem a šestineděním, které jsou na zbytek sítě navázány pouze díky dvěma O chorobám. Dalším svěbytným shlukem je ten s V, S, T a X kódy (zachycuje vnější příčiny smrti spjaté s nehodami, úrazy...), který je u mužů zřetelně mohutnější, a to jednak protože čítá více kódů, a dále také proto, že je spojen s tou částí vnějších příčin, která je u žen

oddělená, sice s menším vyčnívajícím hlukem tvořeným převážně T a X kódy, v němž jsou sdruženy příčiny zejména z rozmezí T36–T50, které jsou podle MKN označeny jako „otrava léky, léčivý, návykovými a biologickými látkami“. Spolu s těmito uzly sousedí „náhodné otravy“ (X40–X49) a „úmyslná sebepoškození“ (X60–X84). Vysvětlení, proč jsou tyto dva klastry u mužů spojeny a u žen odděleny, tkví zřejmě ve větší provázanosti W kódů („jiné vnější příčiny náhodných poranění“), které se u mužů váží nejen na S kódy, jako u žen, ale i na kódy související s předávkováním drogami. U žen se řada W pojí pouze na T75 („účinky jiných vnějších příčin“). Zkrátka, mužské klastry s vnějšími příčinami jsou mnohem spleťtější. Ostatně v příloze 3 s vazbami, které nemají ženy, se rýsuje nemalý podgraf s kódy V, S, T a X.

Ani o jádru sítě však nelze říci, že by nebyly zřetelné tendence určitých typů nemocí se párovat společně. Síť příčin smrti je v mnohém podobná sítím komorbidit, v první řadě je jejich společným rysem poznatek Kim et al. (2016), že „uzly vyjadřující choroby ze stejných kategorií jsou obecně více blízce spojeny“ (Kim et al., 2016, s. 3). Tato teze je v analyzovaném případě uplatnitelná zejména pro choroby postihující konkrétní orgány, nebo mající stejnou etiologii vzniku. Například nemoci oběhové (I kódy), dýchací (J kódy), trávicí (K kódy), vylučovací soustavy (N kódy) mívají často sousedy téhož druhu, jako jsou ony samy. Nápadná je rovněž pospolitost novotvarů (kódy C), jakožto nemocí, které sice nepostihují orgány jediné soustavy, ale vznikají podle jednotných biologických principů. Novotvary jsou povětšinou napojeny na C97, což je kód sám o sobě neoznačující novotvar žádného orgánu, nýbrž je definován jako „zhoubné novotvary mnohočetných samostatných (primárních) lokalizací“. Bývá do hlášení vyplněn, pokud osoba trpěla více než jedním novotvarem zároveň a podle metodických pravidel MKN by neměl být kódován jako UCD. Podle výsledků zde však nejčastěji bývá prostřednictvím ACME vybrán právě jako iniciátor chorobných řetězců, čemuž potenciální vysvětlení bylo nastíněno v explorační analýze. Není však pravda, že C97 jediné „drží“ pohromadě klastr s novotvary, neboť ty se poměrně často také váží mezi sebou navzájem, a to zejména u žen.

Oproti tomu uzly označující především nemoci endokrinní, výživy a přeměny látek (kódy E), poruchy duševní a poruchy chování (kódy F), nebo nemoci nervové soustavy (kódy G) mají tendence se na HOÚ vyskytovat spíše společně s kódy z jiných kapitol MKN. V E nebo F kapitolách nejsou nemoci zařazeny na základě specifické anatomické lokalizace (Kim et al., 2016), nýbrž se jedná o choroby, které mohou ovlivnit celý organismus, a proto se také zde v síti v první řadě váží na od nich samých odlišné choroby. Pro nemoci G uvedené neplatí zcela, protože tato kapitola sdružuje nemoci určité, konkrétně nervové, soustavy. G nemoci mohou být charakterizované jako „multifaktoriální se složitou patofyziologií“ (Piscopo et al., 2021) a problémy s nervovou soustavou se většinou manifestují skrze další zdravotní problémy.

Při standardních vícečetných analýzách příčin smrti se obvykle neuvažují choroby z kapitol R, ani některé další, které označují například komplikace zdravotní péče, či vnější příčiny nezjištěného charakteru (některé W, Y kódy). V sítích komorbidit je někdy volen analogický postup (např. Chmiel et al., 2014, Kim et al. 2016...), jindy autoři pracují s kompletními údaji, tedy včetně R, Y a dalších ne blíže definovaných stavů (např. Hidalgo et al., 2009), což bylo následováno i zde, kdy z hlášení o smrti nebyly vyjímány kódy žádného druhu a lze vidět, že například diagnózy v rámci skupiny R mají v síti nemalé zastoupení. Může se jevit být logickým, že jak R, tak Y se rovněž mají tendence vázat spíše vně své kapitoly.

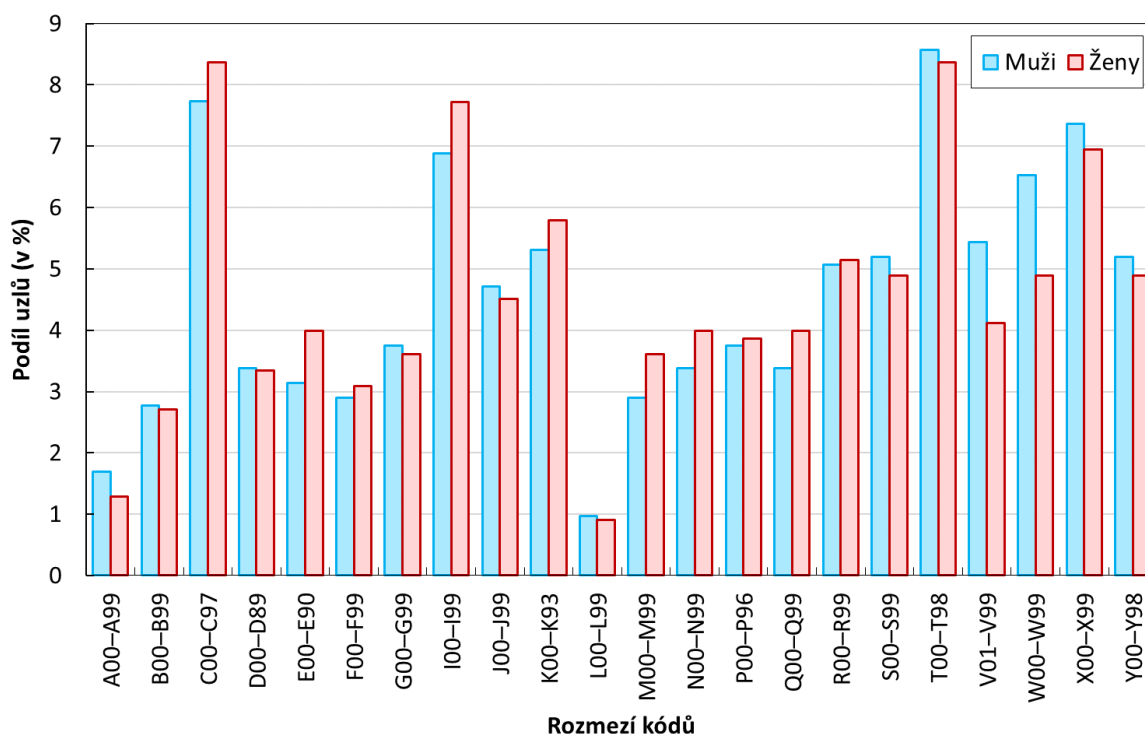
Při srovnání výsledné sítě s dosud provedenými analýzami velkých sítí nemocí lze narážet na nápadný rozdíl. Zatímco v síti příčin úmrtí utvářených z první části HOÚ zcela chybí choroby očí a uší (H00–H95), v sítích komorbidit bývají jedněmi z nejčastěji se vyskytujících. Tedy, lidé často chodí do zdravotnických zařízení kvůli H00–H95¹⁹, ale tyto choroby osoby vyplňující hlášení o smrti nepovažují za podstatné součásti řetězců chorobných stavů vedoucích ke smrti. Na druhou stranu v bimodálním grafu, jenž je zkoumán níže, figurují i tyto choroby, ač vzácně. Podobně jako H například i nemoci L nebo M mají v sítích komorbidit významně vyšší zastoupení než zde v příčinách smrti.

V obrázku 11 jsou uvedeny podíly zastoupení rozmezí kódů na celkovém počtu uzlů. Je očividné, že se nejedná o standardní členění MKN do jejích kapitol, nýbrž o rozdělení ryze podle prvního písmena kódu. Důvod je ryze praktický – proces třídění dat, explorační i síťové analýzy, a nakonec i vizualizace byly mnohem pohodlněji proveditelné podle prvních písmen kódů. Opřít se lze i o další studie (např. Strauss et al., 2021), kde autoři také nerozlišují kódy podle příslušnosti ke kapitolám MKN, nýbrž stejně jako zde. V poznámce pod obrázkem 11 jsou rozepsány celoslovné názvy jednotlivých rozmezí. Jsou-li tedy v dalších pasážích textu zmiňovány například uzly z kapitoly D, jsou jimi myšleny kódy z rozmezí D00–D89, tudíž ležící ve dvou kapitolách MKN-10, sice II a III (MKN-10 Instrukční příručka, 2020).

Nejčastějšími chorobami, které splňují výše zavedené podmínky pro vymezení systému, jsou u obou pohlaví novotvary, srdeční choroby a poranění, otravy a některé jiné následky vnějších příčin nebo vnější příčiny z rozmezí X00–X99, kde jsou zařazeny expozice živlům, úmyslná sebepoškození či též napadení. Zatímco mužská síť je bohatší na kódy týkající se vnějších příčin, v ženské je vyšší zastoupení uzlů symbolizujících novotvary, srdeční choroby, choroby trávicí soustavy, nebo endokrinní.

Překvapujícím se může jevit být zastoupení chorob skupiny R, které je srovnatelné například s respiračními nemocemi, jedněmi z nejčastějších příčin smrti v USA vůbec (CDC, 2021). Však zároveň nelze opominout, že určité druhy kódů, mezi nimi i R, mají vyšší šanci být psány ne samostatně na HOÚ (jedná se o např. symptomy nemocí), protože sytí síť více, analogicky jako kódy stojící pro vnější příčiny smrti, jak již bylo uvedeno např. v kapitole s explorační analýzou. Taktéž je k podotknutí, že ne všechna rozmezí kódů obsahují v MKN stejný počet nemocí, což se samozřejmě taktéž promítá do zastoupení rozmezí uvedených v obrázku 11.

¹⁹ Síť komorbidit jsou zpravidla postaveny na datech o návštěvách lékařů, která poskytují zdravotní pojišťovny.

Obrázek 11: Zastoupení rozmezí kódů podle jejich prvních písmen v síti orientované, USA, 2018

Pozn.: Při srovnání mužů a žen jsou vynechány O kódy (Těhotenství, porod a šestinedělí), ostatní rozmezí označují:

A00–A99 Některé infekční a parazitární nemoci

B00–B99 Některé infekční a parazitární nemoci

C00–C97 Zhoubné novotvary

D00–D89 Nezhoubné, nezjištěné, in situ novotvary a nemoci krve, krvetvorných orgánů a některé poruchy týkající se mechanismu imunity

E00–E90 Nemoci endokrinní, výživy a přeměny látek

F00–F99 Poruchy duševní a poruchy chování

G00–G99 Nemoci nervové soustavy

I00–I99 Nemoci oběhové soustavy

J00–J99 Nemoci dýchací soustavy

K00–K93 Nemoci trávicí soustavy

L00–L99 Nemoci kůže a podkožního vaziva

M00–M99 Nemoci svalové a kosterní soustavy a pojivové tkáně

N00–N99 Nemoci močové a pohlavní soustavy

P00–P96 Některé stavy vzniklé v perinatálním období

Q00–Q99 Vrozené vady, deformace a chromozomální abnormality

R00–R99 Příznaky, znaky a abnormální klinické a laboratorní nálezy nezařazené jinde

S00–S99 Poranění

T00–T98 Poranění, otravy a některé jiné následky vnějších příčin

V01–V99 Dopravní nehody

W00–W99 Jiné vnější příčiny

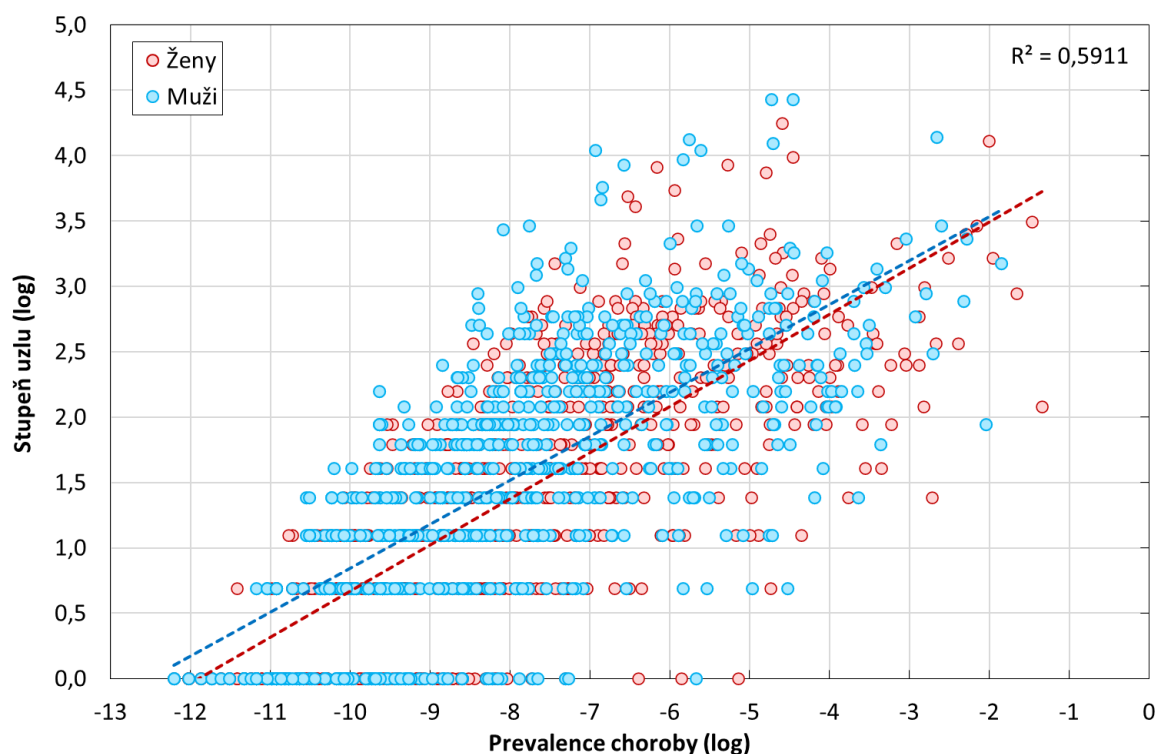
X00–X99 Jiné vnější příčiny

Y00–Y98 Jiné vnější příčiny

Zdroj dat: NBER (2020), výpočet vlastní

Další nezbytnou součástí odpovědi na otázku „jaký je systém“, je vlastnost opakovaně pronikající do mnohých interpretací dále v práci. Je jí zákonitost, že s rostoucí prevalencí chorob se zvyšuje stupeň uzlu (obrázek 12), jak již bylo předzvěstováno výše při analýze kódů podle průměrného počtu zapsaných příčin smrti (kapitola 3). Lineárním regresním modelem by se u obou pohlaví dalo vysvětlit až 60 % variability dat.

Obrázek 12: Vztah prevalence uzlů a stupňů uzlů



Pozn.: R^2 je za obě pohlaví dohromady, neodpovídá ani jedné z vyobrazených přímek.

Zdroj dat: NBER (2020), výpočet vlastní

7.2 Jaká je modularita grafů a jaké nemoci společně utvářejí komunity?

Pro detekci komunit (podskupin, podgrafů) byl použit G-N algoritmus, který nalézá hrany ležící na okrajích potenciálních shluků, tudíž mající vyšší mezilehlost. Maximální nalezená modularita při rozkladu kompletní sítě byla u mužů i žen 0,71. Čím vyšší modularita, tím snáze je síť rozložitelná na podgrafy, přičemž dosáhnout může maximálně hodnoty jedna. V literatuře lze narazit i na interpretaci, že např. modularita s hodnotou 0,70 indikuje, že 70 % hran je lokalizováno v grafu v rozporu s náhodnou sítí (Diviák et al., 2021).

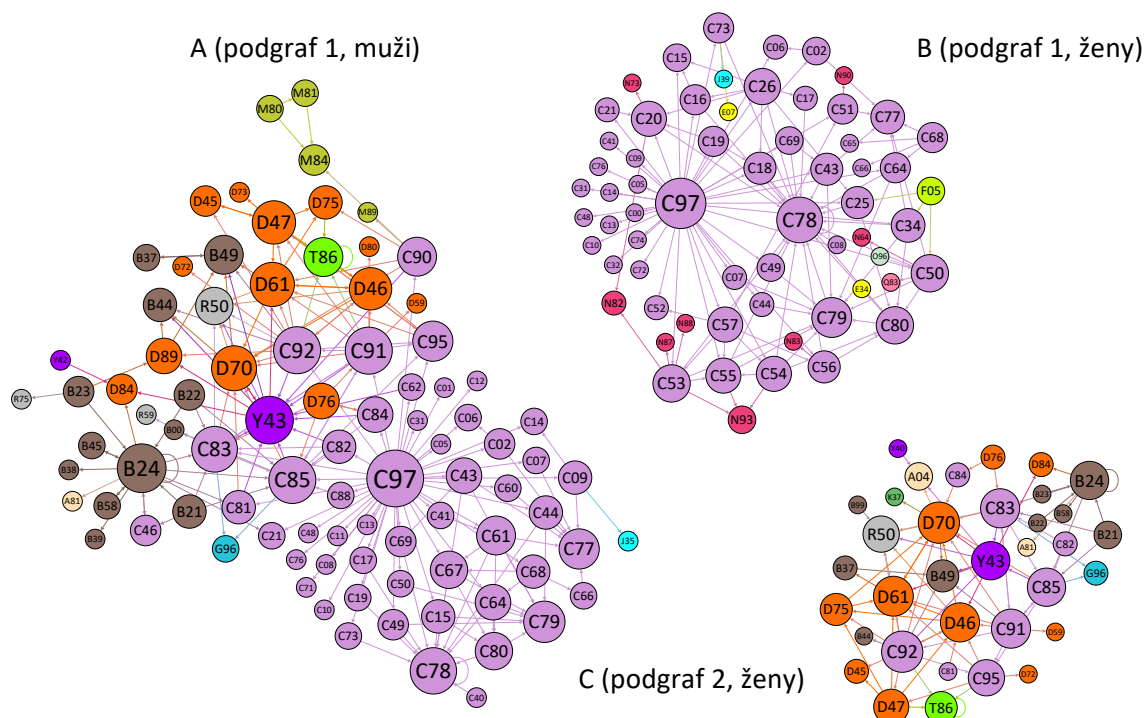
U mužů bylo detekováno 20 podskupin, u žen 28. Tyto podgrafy jsou nejen včleněny do textu níže, ale i v příloze 5. V tabulkách v rámci této přílohy jsou shrnuty základní metriky podgrafů a jejich stručná charakteristika. Ze základních metrik je v tabulkách mimo jiné i modularita podgrafu. To, že je ve všech případech nižší než v celém grafu (což je samozřejmé, jinak by byl ten podgraf dekomponován hlouběji, viz popis fungování G-N algoritmu v metodické kapitole), znamená, že další rozklad by již nebyl tak jednoznačný, jako v rámci kompletního systému. Např. v případě klastru s kojeneckou úmrtností by se dalo tvrdit, že nalézání dalších komunit nižších řádů by znamenalo vytvoření téměř náhodných podskupin, oproti čemuž třeba u mužů v klastru se sepsí je modularita sice nižší než v celém grafu, ale stále dosahující hodnoty 0,5.

Níže je popisován obsah vzniklých klastrů s důrazem na rozdíly mezi pohlavími, přičemž komentovány (a na obrázcích ilustrovány) jsou pouze podgrafy obsahující více než 2 % uzlů z celé sítě.

7.2.1 Novotvary

Již výše bylo naznačeno, že v kompletním grafu je i bez výpočtu modularity u obou pohlaví poměrně nápadná tendence novotvarů se vázat na sebe navzájem. Podle výsledku G-N algoritmu patří zejména u mužů převaha C kódů do jediného klastru. U žen se tento rozpadá na dva, kdy první je tvořen zejména C81–96 („zhoubné novotvary, podle zjištění nebo předpokladu primární, mízní, krevetvorné a příbuzné tkáně“), D50–89 („nemoci krve a krevetvorných orgánů a některé poruchy imunity“) a chorobami souvisejícími s HIV/AIDS (B21–24). Druhý se skládá z většiny ostatních novotvarů (obrázek 13).

Obrázek 13: Podgrafy novotvarů, muži (A), ženy (B a C), USA, 2018



Zdroj: vlastní zpracování v Gephi 0.9.2

Nabízí se otázka, proč nebyl algoritmem u mužů klastr s novotvary rozložen analogicky jako u žen. Výsledek je ovlivněn větší hustotou vazeb v komunitě C na obr. 13 u žen (nebo v příloze 5, podgraf 2), než je tomu u mužů mezi týmiž kódy, což dokládá i síť vazeb vyskytujících se pouze v ženské populaci (množství uzlů z komunity 2 a spíše centrální umístění těchto uzlů v grafu, viz příloha 4). Za vyšší hustotou vazeb mezi chorobami v podgrafu 2 u žen stojí zejména nemoci spojené s anemií (D61), či dále agranulocytóza (D70), což jsou choroby, jimiž trpí ženy výrazně častěji (Regitz_Zagrosek, 2012).

Dalším faktorem přispívajícím k odlišnému rozložení klastrů s novotvary u žen a mužů je i menší počet vazeb mezi uzly z klastru 2 a 1 u žen. Lze spatřit uzly povětšinou marginálnější povahy, které se pouze u mužů váží jak na centrální uzel v hypotetickém podgrafu 1 (C97), tak na uzly, které by obsahoval jejich hypotetický podgraf 2, kdyby rozklad proběhl analogicky jako u žen. Příkladem jsou kódy B21 („onemocnění virem lidské imunodeficiency [HIV], mající za následek zhoubný novotvar“), C88 („zhoubné imunoproliferativní nemoci“) a C90 („mnohočetný myelom a plazmocytární novotvary“).

Poslední skutečností přispívající k tomu, že jen u žen existují oddělené 1 a 2 je i relativně vyšší provázanost uvnitř hlavního klastru s novotvory (B, obr. 13), která existuje kvůli C50 („*zhoubný novotvar prsu*“) a C51–C58 („*zhoubné novotvary ženských pohlavních orgánů*“). U mužů se samozřejmě také vyskytují pohlavně specifické novotvary, ale jednak je jich méně, dále také bývají mnohé z nich vzácnějšími (s rostoucí prevalencí se zvyšuje stupeň uzlu, viz obr. 12) a také se méně často objevují tyto nemoci u jednoho člověka najednou nebo obecně jako komorbidity.

V podgrafech, kde dominují novotvary jsou v malém zastoupení i kódy z jiných kapitol. U žen jsou to choroby z N²⁰ sousedící s C51–C58, ale i poměrně pestré sousedstvo C50 (z kapitol O, Q, F; celé pojmenování viz pozn. u obr. 11, kapitola 3, nebo v přílohách), což může být zřejmě opět odrazem skutečnosti, že novotvar prsu je v USA vůdčím novotvarem co do incidence (National Cancer Institute, 2021, dále NCI).

U mužů neexistuje velký počet vazeb mezi novotvory a N, ale na rozdíl od žen jsou u nich více ukotveny kódy související s HIV/AIDS, které sousedí kromě C81–C96 i s některými novotvory trávicího ústrojí (TS). Vyšší riziko osob s HIV mít tyto choroby je uváděno v řadě studií (Brunnberg et al., 2017). Shluk kódů související s touto nemocí rovněž reflektuje i fakt, že v USA je častější výskyt HIV u mužů než u žen, a to mimo jiné i protože ony v síti nemají R75 („*laboratorní průkaz viru lidské imunodeficiency [HIV]*“), které nevyhovělo podmínce čtyři ve vymezení systému.

Lze si všimnout, že ženy na rozdíl od mužů nemají v klastru číslo 1 ani 2 (obr. 13) kódy M80–M85 („*poruchy hustoty a struktury kosti*“) ani M89 („*jiná onemocnění kosti*“). Tyto choroby u žen jsou zřejmě specifickou oblastí úmrtnosti, protože spolu s některými R²¹ a E²² formují samostatný drobný klastr nazvaný jako „*pohybová soustava*“ v příloze 5, avšak čítající méně nežli 2 % entit. U mužů mají M80–M85 spíše charakter marginálních přidružených chorob k např. C90.

Závěrem ke klastru novotvarů k všimnutí, že do něj u mužů nespádl druhý v USA nejfrekvencovaněji se vyskytující, a to C34 („*zhoubný novotvar průdušky (bronchu) a plíce*“) (NCI, 2021). Ten se nachází v komunitě, již dominují respirační nemoci, poněvadž jednak v nich má mnoho sousedů a jednak se váže i na další novotvary, které pouze u mužů sousedí s J kódy a spadají tudíž do podgrafu s respiračními nemocemi, sice C74 („*zhoubný novotvar nadledviny*“), C32 („*zhoubný novotvar hrtanu*“) a C39 („*zhoubný novotvar jiných a nepřesně určených lokalizací v dýchací soustavě a nitrohručních orgánech*“). Opět, častější výskyt, který do značné míry dokáže predikovat stupeň uzlu, chorob C32 a C39 u mužů byl prokázán v literatuře (SEER, 2021 a Kokoska, 1995).

²⁰ Nemoci močové a pohlavní soustavy

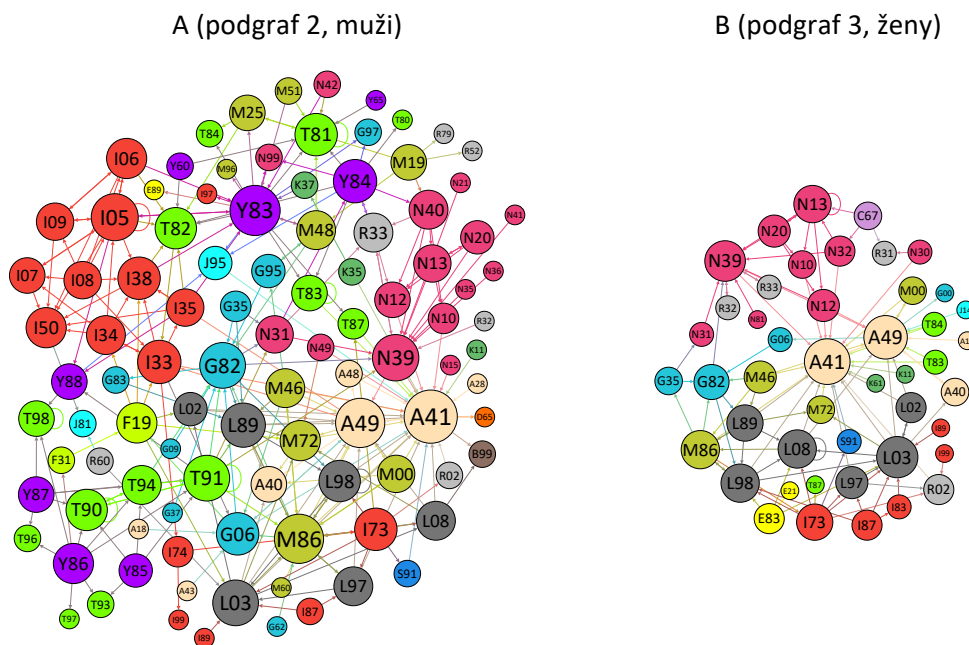
²¹ Příznaky, znaky a abnormální klinické a laboratorní nálezy nezařazené jinde

²² Nemoci endokrinní, výživy a přeměny látek

7.2.2 Sepse a další susedé

V obrázku 14 jsou uvedeny podgrafy, jimž dominuje sepse. V příloze 5 se za ženy jedná o podgraf 3, za muže o podgraf 2.

Obrázek 14: Podgrafy sepse a další susedé, muži (A), ženy (B), USA, 2018



Zdroj: vlastní zpracování v Gephi 0.9.2

Uzlem s jedním z nejvyšších stupňů, přičemž zejména inD, v síti je u obou pohlaví A41 („jiná sepse“). Podgraf, v němž tato choroba leží, je u mužů daleko větší než u žen, a to z důvodu, že se v něm nachází hned dva centrální uzly, sice kromě zmíněné sepse i Y83 („operace a jiné operační výkony jako příčina abnormální reakce nebo pozdější komplikace, bez zmínky o nehodě při výkonu“). I proto by tento klastr u mužů byl poměrně dobře dále rozložitelný na menší (modularita 0,5).

Řada susedů A41 je však jak u mužů, tak u žen táž. V první řadě bývá příčinou sepse kód N39 („jiná onemocnění močové soustavy“), který je důsledkem rozličných N chorob. Dále bývá sepse důsledkem kódů z rozmezí L00–L08 („infekce kůže a podkožního vaziva“) a L80–L99 („jiné nemoci kůže a podkožního vaziva“). L choroby mají v síti procentuálně nejmenší zastoupení, zároveň jsou ale často důsledky chorob z různých kapitol, tudíž mají i spíše vyšší stupně v síti. V těchto ohledech se poměrně podobají samotné sepsi. Tedy, nejen A41 je silně podceňována při jednočetných analýzách úmrtnosti (Pechholdová, 2017), ale i některá L, jež bývají často předstupni samotné sepse a zároveň důsledky rozličných stavů. To, že se např. L02–L03 a L89 do jednočetných analýz příčin smrti nepromítají, pramení ze skutečnosti, že se nachází na některých seznamech nepřipustných základních příčin smrti, stejně jako určitý druh sepse z A41 (Flagg et al., 2021).

Početně vícero v tomto klastru je u obou pohlaví i příčin M susedících s A41 a A49 („bakteriální infekce neurčené lokalizace“), nebo G, z nichž se vazby mezi mnohými sbíhají v G82 („paraplegie a tetraplegie“), jež samo se váže např. na L89, L98, M86, M46 nebo N39.

Klastr za muže je objemnější zejména díky Y85–Y89 („*následky vnějších příčin nemocnosti a úmrtnosti*“) a T90–T98 („*následky poranění, otravy a jiných následků vnějších příčin*“). Tyto diagnózy se váží nejen mezi sebou navzájem, ale mají i řadu sousedů v mužském klastru se sepsí (třeba již zmiňované L89, M96, N39...). Mnohé z těchto vazeb se u žen v síti vůbec nevyskytují, a proto u nich výše uvedená Y a T opět tvoří samostatný drobný klastr (příloha 5 ženy, nečíslovaný podgraf vnějších příčin čítající méně než 2 % uzlů) spojený se zbytkem sítě pouze sousedy T91 („*následky poranění krku a trupu*“).

Ze studií je známo, že vnější příčiny jsou u mužů jednak častějšími a jednak jedná-li se o zranění, bývají i závažnějšími (Westergaard, 2019). Tyto poznatky by mohly souviset s rozdíly mezi pohlavími z hlediska vazeb některých Y85–Y89 a T90–T98 (u mužů jsou tyto kódy v síti více ukotvené), však ne u Y88 („*následky při zdravotní péči jako vnější příčina*“). Pouze u mužů je Y88 důsledkem některých srdečních nemocí postihujících chlopně (I35, I05) a J95 („*stavy dýchací soustavy po výkonech nezařazené jinde*“). Není pravda, že by se tytéž choroby u žen na komplikace nevázaly vůbec, váží se však pouze na Y83 („*operace a jiné operační výkony jako příčina abnormální reakce nebo pozdější komplikace, bez zmínky o nehodě při výkonu*“). Rozdíl mezi Y88 a Y83 spočívá v tom, že Y88 se uvádí, pokud se komplikace vyskytne teprve více než rok od zákroku (ICD-10, 2019). Uvedené rozdíly ve vazbách se jeví nebýt možným interpretovat v kontextu jiných podobných studií, protože v sítích komorbidit se T–Z kódy vynechávají. Nicméně existují studie, podle nichž je konkrétně po operacích kvůli I35 („*nerevmatická onemocnění aortální chlopně*“) u žen vyšší „in-hospital mortality“ (Chaker, 2017), což by mohlo souviset například s neexistencí vazby mezi I35 a Y88, nýbrž jen mezi I35 a Y83 u žen.

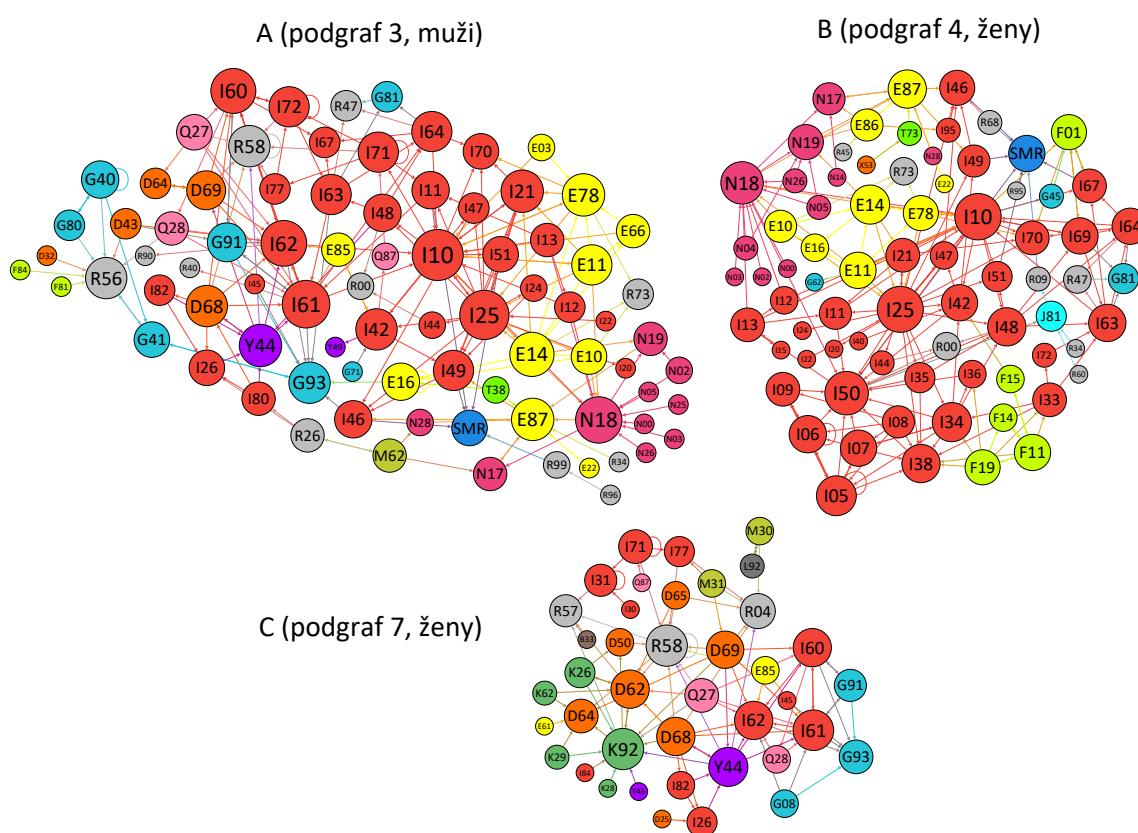
Přítomnost pouze pro muže specifických vztahů mezi komplikacemi a srdečními chorobami (zejména, ale nejen, I05–I09, „*chronické revmatické choroby srdeční*“) je jedním z důvodů, proč jsou mnohá I jen u mužů v klastru se sepsí. Tyto kódy bývají navíc příčinami I50 („*selhání srdce*“), na něž se váží některé marginální srdeční choroby, které se tudíž vyskytují pouze u mužů rovněž v tomtéž podgrafu na obr. 14 (nebo v příloze 5, muži, podgraf 2). Postavení srdečních chorob je ostatně další rys odlišující síť mužů a žen (viz další pododdíl).

Závěrem může vyvstávat otázka, kam u žen spadá Y83. To, navzdory tomu, že sousedy mezi uzly v ženském klastru se sepsí má, spadá do podgrafu, jemuž dominují choroby trávicí soustavy a jiné Y kódy (příloha 5, ženy, podgraf 5). Síť mužů a žen se tedy odlišuje i tím, že u mužů se komplikace péče váží častěji na srdeční choroby, u žen na choroby trávicí soustavy.

7.2.3 Kardiovaskulární choroby

U mužů mají kardiovaskulární choroby převahu jak v podgrafu 2 rozebíraném v předchozím oddíle, tak dominantně i v 3, u žen v 4 a v 7, přičemž by se dalo říci, že kódy obsažené v mužském 3 jsou u žen právě v oněch dvou podgrafech 4 a 7. Muži ale mají v absolutním vyjádření více chorob z kapitoly I²³ než ženy v celé síti. Podgraf 3 za muže a 4 za ženy jsou tvořeny mimo srdečních chorob také R, N a E. U mužů dále ještě ve větším zastoupení lze vidět G²⁴ a D²⁵, která u žen spadají s některými druhy I a K do podgrafu 7 (obr. 15, C), nebo jsou zcela mimo v malém podgrafu (příloha 5, ženy, nečíslovaný podgraf s méně než 2 % uzlů označený „křeče“) s centrálním kódem R56 (viz dál). Podgrafy, jimž je věnována stěžejní pozornost v tomto oddíle jsou na obrázku 15.

Obrázek 15: Podgrafy se srdečními chorobami, muži (A), ženy (B a C), USA, 2018



Zdroj: vlastní zpracování v Gephi 0.9.2

Obě pohlaví mají tedy klastr s převahou I, R, N a E, v němž je i umělý kód pro smrt (obr. 15, A a B), protože právě zde má nejvíce sousedů, sice I46 („srdeční zástava“), I49 („jiné srdeční arytmie“) a I51 („komplikace a nepřesně určené a popsané nemoci srdce“), u žen pak ještě I50 a I67 („jiná cévní onemocnění mozku“), pouze u mužů I11 („postižení srdce při hypertenzi“).

²³ Nemoci oběhové soustavy

²⁴ Nemoci nervové soustavy

²⁵ Některé novotvary a nemoci krve, krevtovorných orgánů a některé poruchy týkající se mechanismu imunity

Všechny tyto choroby splňují podmínky, aby mohly být sousedy smrti, tedy jsou jedněmi z nejčastěji se vyskytujících příčin smrti v USA v 2018 a zároveň většina osob, jimž jsou zapsány na HOÚ, již nemívají žádnou „reálnou chorobu“ jako jejich důsledek, a proto ze stavu s těmito chorobami přechází lidé většinou rovnou do smrti. Podrobnější analýza sousedů smrti je rozvedena až níže.

Již bylo uvedeno, že v klastru jsou významně zastoupeny nemoci močové a pohlavní soustavy (N). Většina z nich u obou pohlaví sousedí s N18 („*chronické onemocnění ledvin*“). Samo N18 je jednou z nejčastějších komorbidit u lidí s I10 („*esenciální (primární) hypertenze*“) i s dalšími srdečními chorobami (Pugh, 2019), což se zde promítá do jejich společného sousedství i sdílení téhož klastru. Další choroby jsou příčinami mnohých I a N18 jsou E10–E14 („*diabetes mellitus*“). Analýzou hlášení o smrti se zabývali Fedeli et al. (2015) a rovněž došli ke zjištění, že „*pravděpodobnost reportování diabetu je vyšší se spoluexistujícími oběhovými chorobami a selháním ledvin*“ (Fedeli et al., 2015, s. 1) než s jinými chorobami. E10–E14 mají v síti i další vazby. Uvnitř stávajícího klastru jsou to ještě další E (E78: „*poruchy metabolismu lipoproteinů a jiné lipidemie*“, E16: „*jiné poruchy vnitřní sekrece slinivky břišní*“, E66 „*obezita – otylost*“) u mužů nebo G62 („*jiné polyneuropatie*“) u žen. „*Diabetická neuropatie je častou neurologickou komplikací diabetes mellitus 1. a 2. typu*“ (Mazanec et al., 2009, s. 1), proto také existují kódy pro společný zápis těchto dvou chorob (specifický čtyřmístný E nebo G63.2 indikující právě diabetickou neuropatii). Může se nabízet myšlenka, že vazba mezi E11 a G62 u žen poukazuje spíše na chyby ve vyplňování hlášení o smrti, neboť podle pravidel MKN by se při komorbiditě diabetu a neuropatie měla uvádět ona konkrétní pod kódem G63.2 nebo ta, jejíž kód začíná na E (MKN-10, 2021).

Také další choroby G se v klastru vyskytují, ale sousedí spíše s I kódy. Konkrétně je to u obou pohlaví G81 („*hemiplegie*“), jejíž příčinou bývá často mrtvice (Siddique et al., 2009), která spadá do I60–I69 („*cévní nemoci mozku*“). Muži mají tyto choroby v hlavním klastru s I (podgraf 4), ženy v 7 spolu s převážně D a K (viz dál).

Další podobnosti mezi pohlavími lze spatřovat v R, která jsou v klastrech s I, N, E a G z nejrůznějších důvodů. Jednak jsou tam komplikace či symptomy vázané na tyto choroby (R00–R09: „*příznaky a znaky týkající se oběhové a dýchací soustavy*“, R73: „*zvýšená hladina krevní glukózy*“, R47: „*poruchy řeči nezařazené jinde*“), dále také R68 („*jiné celkové příznaky a znaky*“) a u žen R95 („*syndrom náhlé smrti dítěte (kojence) [SIDS]*“), protože u nich R95 nemá jiného souseda než smrt.

Nabízí se otázka, proč u žen existuje specifický klaster 7. G-N algoritmus rozkládá graf na komunity detekováním hran s vysokou mezilehlostí. Když existují vyšší hustoty vazeb mezi kódy některých druhů, nabývá hrana spojující tyto hustější podskupiny uzlů vyšších hodnot CB. Za tímto lze nalézat vysvětlení existence klastru 7 u žen, neboť mají v porovnání s muži více vazeb mezi I v klastru 4 (což neznamená, že mají I z tohoto klastru u žen více vazeb obecně). Zároveň jsou u žen I z 4 méně hustě propojena s I z 7 a tato mají naopak více vazeb k rozličným G (G08: jisté zánětlivé onemocnění NS, G93: „*jiné poruchy mozku*“), které muži nemají. V 7 je u žen většina I kódů z I60–I69. Jsou v něm i mnohé D65–D69 („*poruchy koagulace, purpura a jiné krvácivé stavy*“) a K kódy.

U mužů jsou tyto součásti jednak hlavního klastru se srdečními chorobami, nebo jsou k nalezení v klastru s převahou K (podgraf 4, muži), či novotvarů (podgraf 1, muži). U žen jsou D65–D69 v podgrafu 7 kvůli častějším vazbám s I (např. D69 a I60), nebo častějším vazbám D mezi sebou, či na choroby K.

Ze studií věnujících se rozdílům mezi pohlavími v epidemiologii chorob zmíněných v odstavci výše vyplývá, že rizikovými faktory cerebrovaskulárních nemocí (I60–69) jsou mnohá onemocnění srdce (Dufek et al., 2003). Zároveň v nižších věcích jsou těmito chorobami více zasaženi muži, což by s ohledem na zjištěný vztah prevalence a stupně uzlů mohlo částečně vysvětlovat větší propojenost cerebrovaskulárních nemocí a dalších I chorob u mužů. Zároveň však ve starších věcích (85+) jsou I60–69 častěji postiženy ženy a podle některých studií mají také závažnější následky než muži (Roy-O'Reilly, 2015). V kontextu čehož je k zopakování, že ženy mají více vazeb mezi I60–69 a G, konkrétně třeba s G93 („jiné poruchy mozku“).

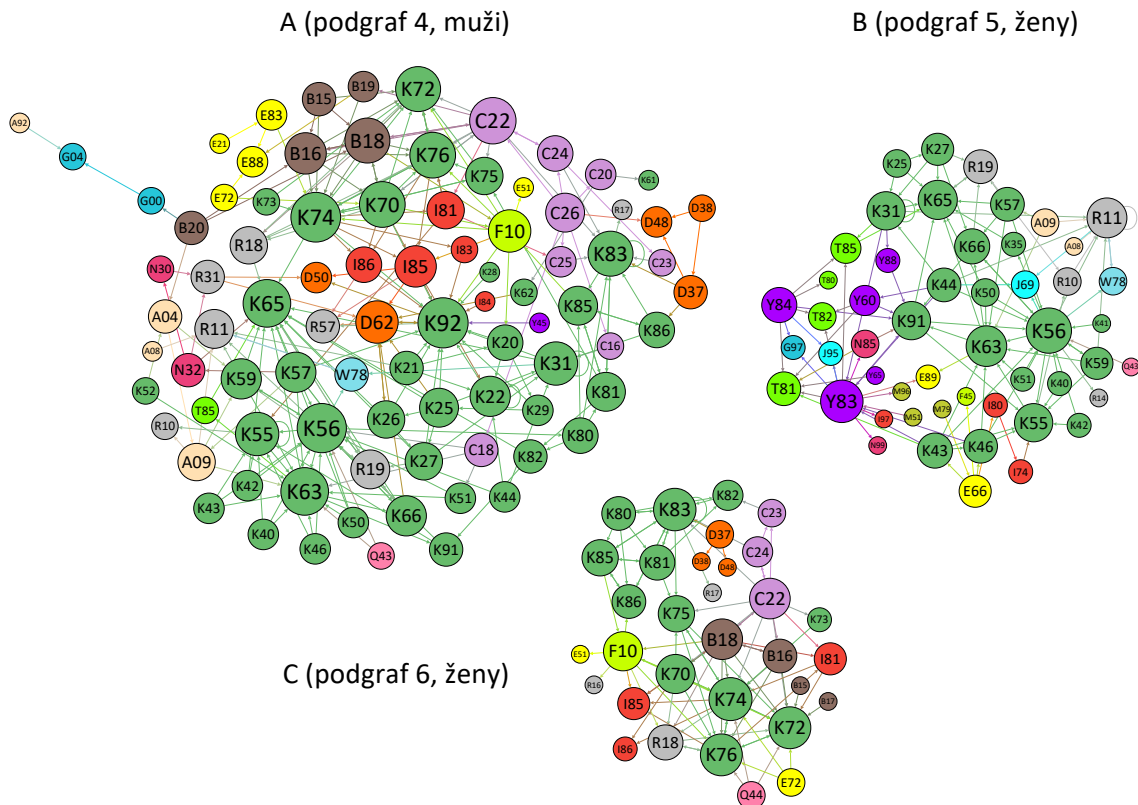
Proč však mají v ženské síti uzly D65–D69 jiné postavení než v mužské, výše uvedené nevysvětluje. Existují však studie, které zkoumají faktory ovlivňující „pohlavní dimorfismus incidence mrtvice a jejích následků“ (Haast et al., 2012, s. 2). Autoři studií dochází k závěru, že jedněmi z nich mohou být i rozdíly mezi muži a ženami v etiologii chorob souvisejících se srážlivostí krve, poněvadž ty jsou podle citované studie jak častou příčinou, tak i důsledkem cerebrovaskulárních nemocí (Samai, 2015). Empirické důkazy o tom, že I a D choroby, které v podgrafu 7 spolu sousedí jen u žen, jsou výrazně ovlivněny pohlavně specifickými faktory (působení estrogenu a chromozomy), a proto také spolu mnohdy souvisí, existuje (viz citované studie v tomto odstavci).

Leč jsou podgrafy uvedené v obrázku 15 srovnávány navzájem, tak by ženský klastr 4 a 7 dohromady nedával přesně klastr 3 u mužů. Významným rozdílem je totiž absence R56 („křeče nezařazené jinde“) a jeho sousedů v ženských 4 a 7. Ten je totiž centrálním kódem v jiném menším podgrafu (jak bylo uvedeno v úvodu oddílu, jedná se o podgraf „křeče“ v příloze 5, ženy) a vyskytuje se odděleně, protože má mezi sousedy řadu Q, kdežto u mužů R56 s chorobami tohoto typu nesousedí, poněvadž analogická Q jsou napojena spíše rovnou na G choroby (G40, G41: choroby související s epilepsií, G80: „mozková obrna“). U žen napojení mezi Q (Q02, Q85, Q90) a G existuje právě skrze R56. Vnitřní konzistentnost vazeb mezi těmito uzly (váží se na sebe, jen zřídka na jiné kódy) je důvodem, proč existuje u žen jedinečný izolovaný drobný podgraf s centrálním uzlem R56.

7.2.4 Nemoci trávicí soustavy

Zatímco příčiny smrti z kapitoly K utvářejí u mužů převážně jeden klastr (příloha 5, muži, podgraf 4), který je co do počtu uzlů čtvrtým největším, u žen jsou K dominantními druhy příčin smrti ve dvou menších (příloha 5, ženy, podgrafy 5 a 6) a významné zastoupení mají i v již výše rozebíraném podgrafu 7, a to díky K92 („jiné nemoci trávicí soustavy“) a jeho sousedům. Kromě příloh jsou podgrafy s převahou chorob TS uvedeny i v obrázku 16.

Obrázek 16: Podgrafy s chorobami trávicího ústrojí, muži (A), ženy (B a C), USA, 2018



Zdroj: vlastní zpracování v *Gephi 0.9.2*

Na rozdíl od předchozích případů je zde docela evidentní, proč se u žen většina chorob trávicí soustavy nachází ve dvou klastrech. V onom menším z nich (podgraf 6, ženy) jsou totiž K70–K77 („nemoci jater“), K80–K87 („nemoci žlučníku, žlučových cest a slinivky břišní“), novotvary těchto orgánů, potom také B15–B19 („virová hepatitida“), a příčinou mnohých chorob z tohoto klastru je F10 („poruchy duševní a poruchy chování způsobené užíváním alkoholu“). U mužů se analogické vazby vyskytují taktéž a lze soudit, že modularita o výši 0,50 nasvědčuje o poměrně jasné existenci dalších komunit v podgrafu 4 u mužů. Výše uvedené kódy u nich však nejsou oddělené i protože muži mají menší hustotu vazeb mezi nemocemi K80–87. U žen jsou choroby žlučníku častější, protože většinou bývají důsledkem K80 („žlučové kameny [cholelithiasis]“), což je nemoc, kterou ženy trpí až dvakrát častěji než muži (Figueiredo et al., 2017).

Obecně ze studií vyplývá, že muži mají vyšší incidenci i úmrtnost na novotvary trávicí soustavy (American Cancer Society, 2021), také častěji mívají onemocnění jater kvůli vyšší spotřebě alkoholu (Buzzetti et al., 2017). Zároveň ale ženy mívají horší progresi některých nemocí jater (např. K70: „alkoholické onemocnění jater“) (Durazzo, 2014) a také bývají méně často sledovány na choroby TS.

Westergaard et al. (2019) dokonce upozorňují na to, že z jimi provedené analýzy nemocničních dat vychází najevo, že zatímco muži bývají nejprve diagnostikováni chorobou TS a až potom případně novotvarem, u žen, právě kvůli nižší častosti monitoringu přichází diagnóza TS až po novotvarech. V síti je v tomto kontextu k uvedení například opačný směr vazby mezi C22 („*zhoubný novotvar jater a intrahepatálních žlučových cest*“) a K70 mezi jednotlivými pohlavími. U mužů diagnóza K70 předchází časově²⁶ C22, u žen naopak bývá nejdříve diagnostikovaná C22, načež teprve K70.

Dále se u žen v podgrafu 5 (obr. 16, B) vyskytuje značné množství Y kódů, zatímco muži mají v trávicím klastru pouze jeden jediný, sice Y45 („*analgetika, antipyretika a protizánětlivá léčiva*“). Na tato Y se váží některá z T80–T88, což jsou sousedství, která se u mužů vyskytují v již povýše rozebíraném podgrafu 2 s dominantní sepsí (obr. 14, A). Jak již bylo uvedeno, ženy mívají komplikace péče spíše v souvislosti s trávicí soustavou, nikoli s I nebo jinými kódy v mužském podgrafu 2. To nemusí nutně reflektovat horší postoperační stavy u žen s chorobami TS, ale podle některých studií i to, že mají tendenci komplikace spojené s TS hlásit častěji než muži (Myles et al., 2004).

Na trávicí choroby se váží i mnohá R (R10–R19: „*příznaky a znaky týkající se trávicí soustavy a břicha*“), I80–89 („*nemoci žil, mízních cév a mízních uzlin nezařazené jinde*“) i E, z nichž u žen je v tomto klastru i E66 („*obezita*“), která u mužů spadá do podgrafu 3 s převahou srdečních chorob (obr. 15, A). U žen se E66 na srdeční choroby i E11–E14 (které mají centrální postavení v obrázku 15) váže samozřejmě také, zároveň má ale rozličné sousedstvo v podgrafu 5 u žen, kromě K kódů i M79 („*jiná onemocnění měkké tkáně nezařazené jinde*“), I80 (druh zánětu žil) nebo F45 („*somatoformní poruchy*“). V USA trpí ženy obezitou více než muži, kteří ale mívají častější spojení obezity s kardiovaskulárními nemocemi (Cooper et al., 2021). Podle některých studií ženy považují obezitu za výrazněji snižující kvalitu jejich života než muži, o čemž svědčí i přidružené psychologické nemoci k obezitě častější u žen (Tronieri et al., 2017), zároveň jsou ale duševní choroby obecně více zastoupeny v populaci žen (Kalgotra et al., 2017). V síti příčin smrti pouze u žen sousedí E66 se zmíněným F45. A odlišné sousedstvo E66 rozebírané v tomto odstavci je důvodem, proč je obezita v odlišném podgrafu u žen než u mužů, konkrétně v podgrafu s převahou nemocí trávicí soustavy (obr. 16).

²⁶ Korektněji bývá na HOÚ psána nad K70.

V podgrafech s respiračními chorobami se vyskytují M choroby z rozmezí M05–M06 (druhy artritid) a M30–M36 („*systemová onemocnění pojivové tkáně*“). „*Přibližně 30 % až 40 % pacientů s artritidou vykazuje buď radiografické nebo plicní funkční abnormality svědčící o intersticiální fibróze nebo restriktivním plicním onemocněním*“ (Roschmann et al., 1987, s.1). Toto může být v síti vyjádřeno tím, že M05–M06 jsou u obou pohlaví příčinami J84 („*jiné intersticiální plicní nemoci*“). Rovněž existují studie, podle kterých více než polovina pacientů s některou z diagnóz M30–M36 po dvou až čtyřech letech získá intersticiální plicní nemoc (Végh et al., 2005). Všechny choroby počínající písmenem M z klastru 4 u žen i 40 u mužů (obr. 17) jsou příčinami J84, které samo je důsledkem pouze dvou J. Jedná se tedy o chorobu s výrazně vychýleným poměrem stupně dovnitř ku ven ve prospěch prvého.

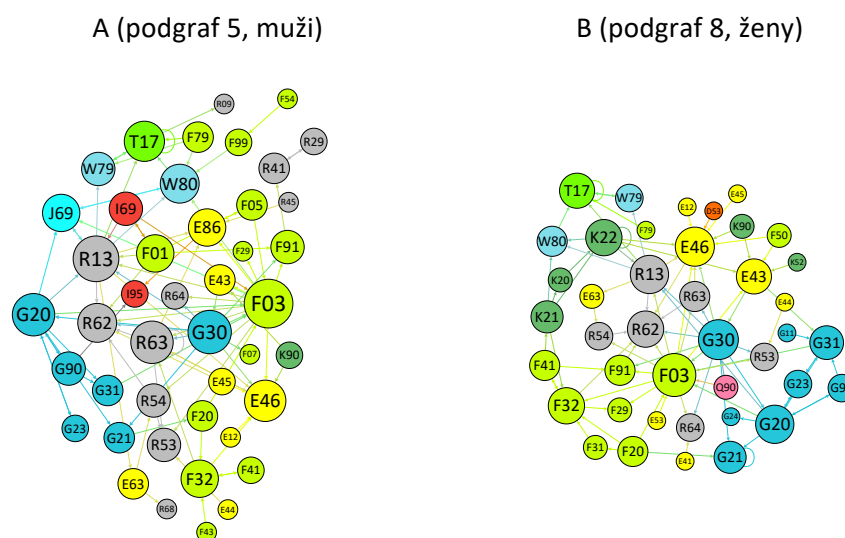
Zejména u žen jsou v klastru přítomny i mnohé infekční a parazitární nemoci a choroby G. Tyto označují vzácnější onemocnění týkající se funkce svalů (G12: „*míšní svalová atrofie a příbuzné syndromy*“, G70), jež bývají příčinami J96 („*respirační selhání nezařazené jinde*“). Dalším G je G47 („*poruchy spánku*“), které je příčinou J44 („*jiná chronická obstrukční plicní nemoc*“). „*Porucha nespavosti je velmi častá u pacientů s chronickou obstrukční plicní nemocí*“ (Budhiraja et al., 2015, s. 3). Avšak podle citované studie bývá G47 spíše důsledkem J44, ač zde v síti příčinnost mezi těmito nemocemi vychází v opačném směru. Může se tedy jednat o důsledek aplikace pravidel pro tvorbu příčin a důsledků odvíjejících se od oficiální UCD.

Příčinnost neodpovídající výsledkům studií se vyskytuje i u F17 („*poruchy duševní a poruchy chování způsobené užíváním tabáku*“), kdy u obou pohlaví je tato choroba jak na pozici příčiny, tak i překvapivě důsledku chorob respiračního traktu, konkrétně je u mužů i žen zde v síti F17 důsledkem J44 a 43 („*rozedma plic*“), přičemž je ve skutečnosti vůdčí příčinou těchto nemocí (Snider, 1989). Podle pravidel pro vyplňování hlášení o smrti z 2014 by F17 neměla být uváděna jako základní příčina smrti, pokud je znám bližší fyziologický stav (ICD-10 Mortality Manual, 2014). Rovněž je však F17 častým rizikovým faktorem nejen mnoha respiračních chorob, proto také může mít přímou vazbu na stavy z první části hlášení. Pokud je do ní zapsána, podle pravidel by neměla být spouštěčem, zároveň ale podle epidemiologického konsenzu je pravděpodobnější příčinou respiračních nemocí než jejich důsledkem. Nabízí se myšlenka, že obojaká vazba mezi F17 a dalšími kódy může částečně pramenit z možné neexistence úplně jednoznačně vhodného místa na hlášení o smrti, kam by mohla být zapsána, když je častým a v literatuře hojně uváděným původcem některých respiračních chorob, zároveň by ale neměla být chorobou iniciující řetězce vedoucí ke smrti.

7.2.6 Choroby mozku a duševní choroby

Ačkoli řada F i G kódů se nachází též v klastrech s převahou chorob zasahujících jiné orgány, největší koncentrace těchto nemocí se nalézá v podgrafu 5 u mužů a v podgrafu 8 u žen. Tyto podgrafy jsou k vidění v obrázku 18. Je vidno, že tyto podgrafy, podobně jako ony s respiračními chorobami (však možná ještě důrazněji) se významně u mužů a žen neodlišují ani co do struktury, natož velikosti a hustoty. U obou pohlaví by stávající uskupení chorob mohlo být výstižně nazváno jako „úmrtnost kvůli rozličným demencím“.

Obrázek 18: Podgrafy s demencemi, muži (A), ženy (B), USA, 2018



Zdroj: vlastní zpracování v Gephi 0.9.2

Klastry na obrázku 18 se u obou pohlaví skládají z centrálního F03 („neurčená demence“) a jejich sousedů z téže kapitoly, dále kódů G20–G26 („extrapyramidové a pohybové poruchy“, sem patří třeba Parkinsonova nemoc) a G30–G32 („jiné degenerativní nemoci nervové soustavy“ pod toto spadá Alzheimerova nemoc). Tyto choroby jsou si jednak samy sobě příčinami a důsledky a jednak sousedí s marginálnějšími uzly, z nichž většina spadá do kapitoly R a E, u žen i K.

„Závažná demence často způsobuje komplikace jako je imobilita, poruchy polykání a podvýživa“ (Alzheimer’s Association, 2022), což jsou rovněž choroby ležící v klastru s demencemi i zde v síti (R13, R63, R64, viz dále). Zároveň se v souvislosti s demencemi často píše o přidružených chorobách, nejčastěji srdečních a oběhových, metabolických, pohybového aparátu, trávicí soustavy nebo duševní (Santiago et al., 2021 a Duthie et al., 2011). Podgrafy číslo 5 (muži) a 8 (ženy) v síti vytvořené z dat v první části hlášení evidentně zachycují stavy, které lidé s demencemi mívají z příčiny gradace jejich základní příčiny smrti, tj. demence. Ony komorbidity (tj. výše zmíněné I, E nemoci aj.) bývají systematicky psány spíše do druhé části hlášení, jak bude vidět v neorientované síti. Rovněž jednou z nejčastějších komorbidit je u lidí s demencí F32 („depresivní fáze“) a její přítomnost v podgrafech zde v síti do jisté míry odporuje tezí výše ohledně částí hlášení. Existují ale starší články, podle nichž F32 může být rizikovým faktorem pro vznik demence (Jorm et al., 2001). Na druhou stranu se podle jiných studií ukazuje, že F32 často vzniká v důsledku G30 a závažnější F32 vede k horšímu průběhu G30 (Moretti et al., 2002). Další studie poukazují na to, že počáteční symptomy G30 se do jisté míry pouze překrývají se symptomy F32 (Alzheimer’s Association, 2022), aniž by se obě diagnózy ve vyšší míře vyskytovaly u jednoho člověka zároveň. Tedy, mechanismus za asociací mezi F32 a G30 není podle studií jednoznačný, i proto vazba mezi těmito chorobami existuje jak v orientované síti, tak v neorientované, čemuž takto není u ostatních vazeb s demencemi.

Podgrafy s demencemi jsou především u mužů bohaté na mnohá R. Jedná se o stavy a symptomy plynoucí z již zmíněných komplikací demencí, u obou pohlaví např. R13 („*dysfagie*“, to je potíží polykání), R53 („*nevolnost a únava*“), R54 („*stáří (senilita)*“), R63 („*příznaky a znaky týkající se příjmu potravy a tekutin*“), R64 („*kachexie*“, to je nedostatek prospívání u dospělého). Dále u mužů i žen je důsledkem těchto R, demencí samotných i metabolických chorob R62 („*nedostatek předpokládaného normálního fyziologického vývoje*“). Jedná se o kód, který by měl indikovat odchylky od standardního vývoje v dětském věku. Hypotézy, proč se často objevuje na hlášeních o smrti u lidí s demencemi se samozřejmě nabízí³¹, ale oporu v literatuře nenachází, poněvadž R kódy se obecně považují za neúčinné při řešení otázek týkající se úmrtnosti. Podle Flagg et al. (2021) jsou, když už, tak vhodné spíše k indikaci kvality dat o příčinách smrti – čím méně R, tím kvalitnější data.

Pouze u mužů jsou v klastru i další R, například některá R40–R46 („*příznaky a znaky týkající se vědomí, vnímání, emočního stavu a chování*“), a to mimo jiné i protože ženy nemají v podgrafu 8 kód E86 („*snížení objemu plazmy nebo extracelulární tekutiny*“, třeba dehydratace), poněvadž u nich tato choroba spadá spolu s dalšími z této kapitoly do podgrafu se srdečními chorobami a diabetem (4, ženy), neboť se oproti mužskému E86 váže na více N nemocí. Jinak se choroby počínající písmenem E v podgrafu demencí u mužů a žen shodují, neboť u obou jich nejvíce spadá do rozmezí E40–E46 („*podvýživa*“).

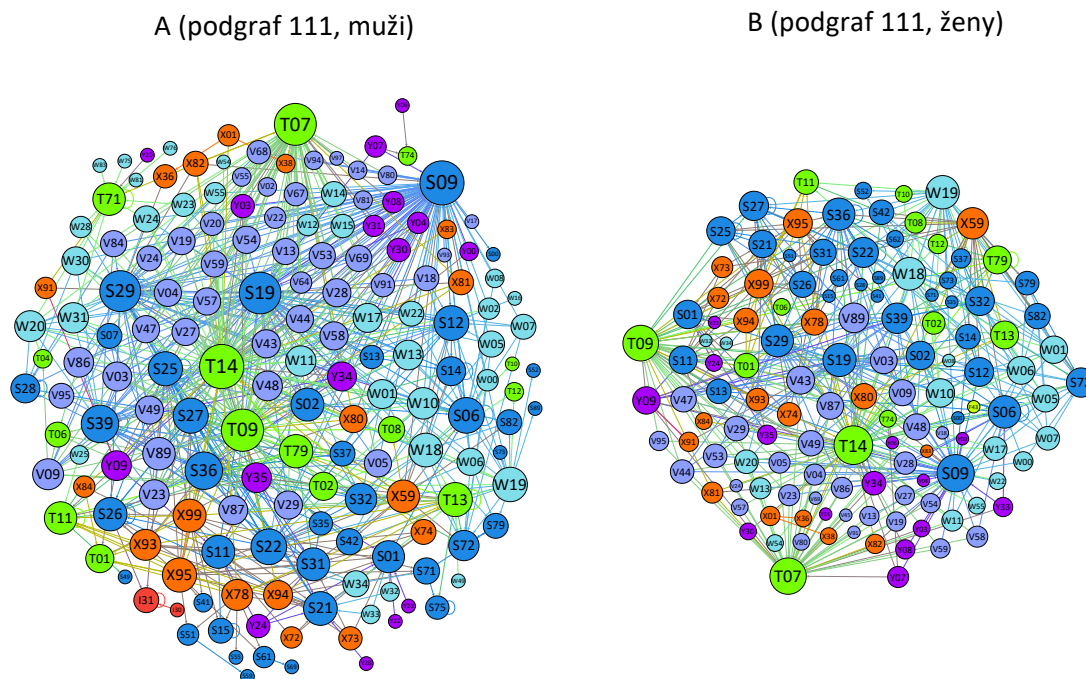
Pouze u žen se ještě objevují některé nemoci TS, které se váží na metabolické choroby, nebo na F32 a F41 („*jiné anxiózní poruchy*“). Dále při analýze přidružených chorob může být relevantním, že podle Brunnström a Englund (2009) bývá častou bezprostřední příčinou smrti u lidí s demencí pneumonie. V mužském podgrafu se J69 („*pneumonie způsobená pevnými a tekutými látkami*“) skutečně objevuje, ale v ženském nikoli. Však jiné nemoci z kategorií F a G, zejména F03, se i u žen na některé respirační choroby skutečně váží, ale tyto leží mimo podgraf 8, nebo u mužů 5.

7.2.7 Vnější příčiny smrti

Již bylo zmíněno, že nemalý podíl uzlů jsou v síti odkazem na vnější příčiny smrti. Ani u žen, ani u mužů však nejsou všechny slity do jediného klastru, nýbrž do několika, z nichž každý je samostatný, protože zachycuje odlišný typ vnějších příčin smrti než ostatní. Největší klastr nejen mezi těmi sdružujícími vnější příčiny, ale také v celé síti, je tvořen uzly vyjadřujícími úrazy zapříčiněné nehodami, a to nejčastěji v dopravě (příloha 5, podgraf 111 u žen i mužů). Dalším je klastr s příčinami smrti v důsledku otrav (příloha 5 podgraf 9 muži, 12 ženy). V síti jsou ještě další dva relativně větší podgrafy vnějších příčin, první z nich sdružuje úmrtnost kvůli kontaktu s ohněm (dýmem, hořlavinou...) (táž příloha 5, podgraf 10 muži, 13 ženy), druhý obsahuje příčiny smrti v důsledku vzdorování vodě (tonutí) nebo dalším živlům (vítr) (příloha 5, podgraf 11 muži, 14 ženy). V obrázku 19 jsou podgrafy 111 za každé pohlaví.

³¹ Např. u mužů je R62 důsledkem E45 („*opožděný vývoj následující po protein-energetické podvýživě*“, v tom zakrnění, malý vzrůst). Může se v něm tedy skrývat skutečnost, že bývá psán dospělým lidem, kteří se v důsledku výše citovaných komplikací spjatých s Alzheimerem zmenší. Pod R62.8, který je podle NCHS (2006) mezi čtyřmístnými R nejméně frekventovanější, se skutečně skrývá „nedostatečný vzrůst“, ale ne u dospělých. Vazby R62 mohou tedy opět poukazovat spíše na systematické chyby při vyplňování hlášení.

Obrázek 19: Podgrafy s vnějšími příčinami smrti, muži (A), ženy (B), USA, 2018



Zdroj: vlastní zpracování v *Gephi 0.9.2*

Podgraf 111 u žen obsahuje přes 15 % uzlů a 20 % hran celé sítě, u mužů do téhož podgrafu spadá přes 20 % uzlů a 25 % hran. Důvod, proč kódy spjaté s úmrtími na vnější příčiny zabírají podstatnou část všech entit v síti byl již uveden v části s explorační analýzou. Mladí lidé mívají nejčastěji na hlášeních o smrti aspoň dva kódy, protože jsou skupinou, kde lidé nejvíce umírají právě v souvislosti s vnějšími příčinami a ty by podle pravidel pro kódování příčin z kapitol S, T, V a W neměly být nikdy psány samostatně na hlášení o smrti. Požadovanou informaci je totiž nejen druh vnější příčiny (pád, dopravní nehoda, otrava alkoholem...), ale i její následky, tj. např. lokalizace úrazu na těle a jeho povaha, nebo srdeční potíže kvůli otravě atd. I proto je hustota podgrafů 111 u obou pohlaví o poznání vyšší v porovnání s kompletní sítí. Existence hlubší a spletné vnitřní organizace v tomto podgrafu se odráží i na spíše nižší modularitě tohoto klastru a také na skórech autorit a hub, jak bude rozvedeno dále.

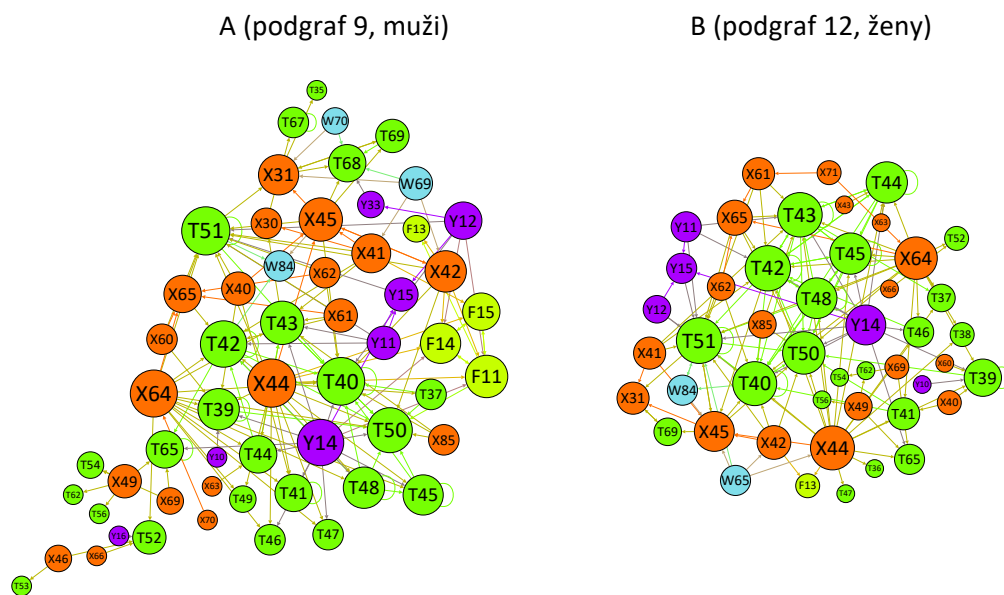
Kromě typů dopravních nehod (V01–V99), rozličných jiných vnějších příčin z rozmezí W00–X59 (v tom ony pády, udeření, pořezání, výbuchy...), lokalizací poranění (S00–S99) a jejich druhů (T00–T78) jsou v podgrafech 111 ve větším výskytu i některé X a Y kódy. Z X jsou to nejčastěji příčiny spadající do X60–X84 („úmyslné sebepoškození“), z Y do Y10–Y34 („případ (událost) nezjištěného úmyslu“), ale i některá napadení nebo kódy stojící pro špatná zacházení, zanedbání (Y06, Y07).

Vlivem rozdílné velikosti sítí samotných i podgrafů 111 u mužů a žen existuje řada vazeb mezi vnějšími příčinami, které mají pouze muži, ale i některé vazby, které mají jen ženy (přílohy 3 a 4). Na první pohled je vidět, že způsoby smrti při dopravních nehodách i typy smrtelných zranění jsou u žen mnohem méně pestré než u mužů. Naopak u žen existuje více vazeb u kódů W18, W19 (pády na rovině a neurčené), za čímž stojí i to, že se více než u mužů váží na příčiny nejsoucí vnějšími, sice choroby pohybového aparátu (osteoporózy), jimiž ženy trpívají více než muži.

Westergaard et al. (2019) analyzovali dvojice chorob, které měly rozdílný příčinný vztah u žen a mužů. Mimo jiné identifikovali i několik dvojic zranění (vnějších příčin) a došli k závěru, že rozdíly mezi muži a ženami z hlediska příčinnosti zranění a typů kódů mohou reflektovat genderové rozdíly v domácím násilí. Zde v síti je potenciálně v návaznosti na toto zjištění znatelné, že ve specificky ženských vazbách mezi vnějšími příčinami se často vyskytují vágní kódy, např. X59 („vystavení neurčeným faktorům“, v tom např. zlomeniny z nejasné příčiny), ale i oni sousedé W18, jmenně třeba T79 („některé časné komplikace úrazů“), dále i Y33 („jiné určené případy (události), nezjištěného úmyslu“), Y07 („jiné špatné zacházení“). Tyto v síti s vazbami ryze mužskými mají menší stupně, marginálnější postavení, nebo zcela chybí. Tyto rozdíly mezi muži a ženami však nemusí být odůvodněny poznatky, o nichž píše autoři Westergaard et al. (2019).

Vnější příčiny smrti se sdružují i v dalších, velikostně však drobnějších komunitách. V podgrafu (9 muži, 12 ženy), které se nachází v obrázku 20, se se nachází většinou otravy.

Obrázek 20: Podgrafy sdružující převážně otravy, muži (A), ženy (B), USA, 2018



Zdroj: vlastní zpracování v Gephi 0.9.2

Podgrafy z obrázku 20 utvářejí u obou pohlaví X40–X49 („náhodná otrava škodlivými látkami a expozice jejich působení“) a X60–X84, přičemž z tohoto oddílu se nejčastěji jedná o X60–X69 (úmyslné sebeotravy), z X zde leží i X70, X71 (způsoby sebevražd) jež bývají důsledky předávkování. X v tomto klastru tedy většinou sdružují činy úmyslné, oproti T kódům, které stojí pro náhodné příčiny.

Z nich jsou k nalezení v podgrafech 9 u mužů a 12 u žen nejvíce příčiny smrti z rozmezí T36–T50 („otrava léky, léčivý, návykovými a biologickými látkami“) nebo T51–T65 („toxické účinky látek ze zdrojů převážně mimo lékařství“), zejména u mužů i další, které jsou zpravidla důsledky otrav, sice např. T66–T78 („jiné a neurčené účinky vnějších příčin“). Kódy Y v tomto podgrafu stojí pro skutky, jejichž záměrnost nebyla zjištěna. Nejčastěji leží v rozmezí Y10–Y16 (otravy s nezjištěným záměrem). U mužů se v tomto klastru vyskytuje i Y33 („jiné určené případy (události) nezjištěného úmyslu“), poněvadž je důsledkem Y12 či T68, ale neváže se na zranění na těle, jako u žen. V podgrafu se nalézají v menší četnosti i příčiny smrti z kapitoly W, a to zejména u mužů. Jedná se o nehody, u nichž osoba vyplňující hlášení o smrti vyhodnotila, že se staly v důsledku nadužití psychotropních látek, ilustračně třeba vztah mezi W69 (utnutí) a X42 (otrava halucinogeny).

U mužů se v podgrafu 9 nachází ještě F10–15 („poruchy duševní a poruchy chování způsobené užíváním psychoaktivních látek“). Spíše nežli to, že se u mužů nachází, může být otázkou, proč se u žen nenachází. Důvodem je častější vazba mezi těmito kódy a I chorobami u žen která zapříčiňuje to, že se F10–F15 nacházejí v jejich podgrafu se srdečními nemocemi (obrázek 15, nebo příloha 5, ženy, podgraf 4). Podle některých zdrojů mají muži větší pravděpodobnost závislosti na drogách (úmrtnost spjatá s tímto je u nich více zastoupená než u žen), ale zároveň vyšší riziko progresu od zneužívání návykových látek k závislosti na zneužívání návykových látek mají ženy (Becker a Hu, 2008). Podle Fernandez-Montalvo et al. (2014, s. 1–2) jsou „ženy více náchylné k strádání na vedlejší účinky návykových látek a předávkování“. Tuto tezi autoři uplatňují v souvislosti s faktem, že ženy-alkoholičky trpívají více na jaterní choroby a ženy-uživatelky drog na kardiovaskulární nemoci v porovnání s muži.

USA je v rámci zemí vyspělého světa v čele z hlediska dostupnosti opioidů i v intenzitě úmrtnosti spjaté s těmito látkami³² (OECD, 2019), která také podle CDC (2022) stojí v USA za více než 70 % všech úmrtí z důvodu předávkování drogami a obnáší v nejvyšším zastoupení smrt v důsledku heroínu a opioidních léků. Věkově standardizovaná míra úmrtnosti na předávkování drogami byla v USA v 2018 21,8 na 100 000 obyvatelů (CDC WONDER, 2021), což je přibližně stejná hodnota jako na E10–E14 („Diabetes mellitus“) (21,4 na 100 000) (CDC WONDER, 2021). Exponenciální nárůst úrovně úmrtnosti důsledku užívání drog v USA od 1980 vedl k tomu, že úmrtí na předávkování opioidy se začala považovat za epidemii, jak píše Volkow (2014).

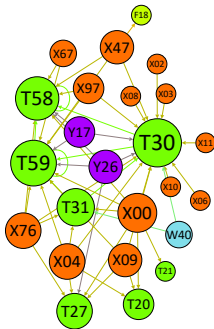
Podgrafy 12 a 9 (obr. 20) se podle pohlaví odlišují, což by mohlo být interpretováno v kontextu zjištěných rozdílů mezi ženami a muži v „opioid overdose epidemic“ (dále epidemie předávkování se drogami). V první řadě je nápadné především to, že existuje více jen ženských vazeb, ač podgraf 12 u žen je menší než 9 u mužů. Síť žen totiž obsahuje více vazeb mezi T kódy s náhodnými otravami zejména z T36–T50. „Nežádoucí lékařské, psychiatrické a funkční důsledky spojené s poruchami užívání návykových látek jsou u žen často závažnější“ (McHugh, 2020), což Zucker a Prendergast (2020) nebo Clayton et al. (2014) zdůvodňuje skutečností, že i přestože ženy a muži reagují na léky jinak, mnoho předepisovaných léčiv nebylo testováno adekvátně na ženách.

³² Opioid-related mortality

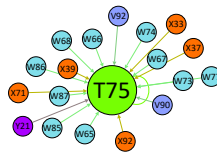
Poslední dva drobné klastry vnějších příčin obsahující aspoň 2 % uzlů kompletní sítě sdružují kódy vázající se v drtivé většině na T30 („popálenina a poleptání, část těla neurčena“) (podgraf 13 ženy, 10 muži) a T75 („účinky jiných vnějších příčin“, třeba tonutí, zásah elektrickým proudem...) (podgraf 14 ženy, 11 muži). V těchto podgrafech se nalézají opět jak činy náhodné, úmyslné, nezjištěného úmyslu i násilné, které většinou zapříčinily T30, resp. T75. Tyto komunity se nachází na obrázku 21.

Obrázek 21: Ostatní podgrafy s vnějšími příčinami, muži (A a C), ženy (B a D), USA, 2018

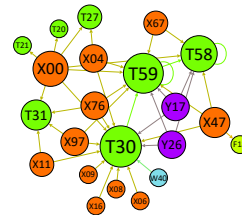
A (podgraf 10, muži)



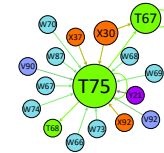
C (podgraf 11, muži)



B (podgraf 13, ženy)



D (podgraf 14, ženy)



Zdroj: vlastní zpracování v Gephi 0.9.2

7.2.8 Ostatní podgrafy

Výše již bylo uvedeno, že u žen se vyskytují ještě další drobné klastry s méně než 2 % uzlů, které nejsou zcela izolované od hlavního grafu, jako dvojice některých chorob v kompletní síti (příloha 1). Prvním takovým je podgraf sdružující M nemoci (příloha 5, ženy podgraf „pohybová soustava“), zejména z M80–85 a M15–M19. Tyto choroby jsou u mužů v klastru s novotvory, nebo sepsí, neboť M nemoci mají u mužů povahu marginálních uzlů, ale u žen se váží hustě nejen samy na sebe, ale i na některá R nebo E. Ostatně v síti s vazbami, které nemají muži, patří M choroby k uzlům s nejvyšším stupněm.

V literatuře věnované rozdílům v nemocnosti mezi muži a ženami bývá vznesena hypotéza, že ač choroby spojené s osteoporózou, která patří do M80–M85, se více vyskytují u žen, rozdíl mezi pohlavími nemusí být tak velký, nýbrž jsou muži v tomto směru pouze méně často diagnostikováni a pokud už, tak často až v důsledku zlomeniny (Westergaard et al., 2019). Baggio (2013, s. 9) píše, že „testovány na tuto nemoc bývají až čtyřikrát častěji ženy než muži“. Tato fakta mohou být důvodem, proč tvoří pouze u žen osteoporózy samostatný klaster, kdež u mužů jsou částečně v podgrafu 1 (obrázek 13). O tom, že u žen mají nejen kódy související s osteoporózou, ale i obecně patřící do kapitoly M více vazeb, vypovídá i vyšší stupeň u těchto kódů v síti s vazbami nevyskytujícími se u mužů (příloha 4).

Výše bylo psáno, že další specificky ženský podgraf vzniká okolo R56 („*křeče nezařazené jinde*“), poněvadž se více váže na vrozené nemoci než u mužů, a proto neleží v podgrafu s kardiovaskulárními chorobami a E10–14 jako u nich (obr. 15, A, nebo podgraf 3, muži). Jen u žen existuje i malinký podgraf sdružující Y85–87 (následky nehod a úmyslného sebepoškození) a některé důsledky T90–98 (následky poranění a otrav). Tento je oddělen kvůli menší hustotě vazeb mezi těmito kódy a zbytkem grafu. Více než 1 % uzlů u žen obsahuje ještě podgraf kolem T71 („*zadušení*“) a také O³³ kódů. Klastř s příčinami smrti spjatými s těhotenstvím se na zbytek sítě váže pouze přes některé srdeční choroby, duševní choroby a M32 („*systémový lupus erythematodes*“), který „*je jednou z nejběžnějších autoimunních nemocí zasahujících ženy během jejich reprodukční doby a zvyšuje riziko rozličných komplikací spjatých s těhotenstvím*“ (Ruiz-Irastorza et al., 2008, s. 1).

³³ Těhotenství, porod a šestinedělí

7.3 Kdo jsou sousedé smrti?

Za stávajících podmínek pro vymezení systému se na smrt mohou vázat pouze příčiny, z nichž největší podíl lidí odchází přímo do ní, nikoli do jiných stavů, a zároveň se jedná o příčiny vyskytující se častěji než jedním procentem. V tabulce 11 jsou uvedeny kódy toto splňující. U žen jsou nejvíce zastoupeny novotvary, choroby srdeční a příznaky z kapitoly R. Obdobně je tomu u mužů, kteří však mají spíše nežli novotvary častější poranění, otravy a následky vnějších příčin.

Tabulka 11: Sousedé smrti

	Kapitola	Kód	inD/outD
Ženy	Některé infekční a parazitární nemoci	A41	6,6
	Novotvary	C25, C56, C78, C79	0,6; 0,4; 8,3; 5,5
	Nemoci nervové soustavy	G30, G31	0,2; 1,3
	Nemoci oběhové soustavy	I46, I49, I50, I51, I67	3,0; 1,5; 5,3; 3,0; 1,3
	Nemoci dýchací soustavy	J96	32,0
	Příznaky, znaky a abnormální klinické a laboratorní nálezy nezařazené jinde	R09, R57, R62, R68, R95	3,0; 3,5; 4,5; 1,5; 0,0
	Poranění, otravy a některé jiné následky vnějších příčin	S01, S21 T01, T07, T40	6,0; 2,0 7,0; 47,0; 4,0
Muži	Novotvary	C25, C78, C79	1,0; 12,0; 10,0
	Nemoci nervové soustavy	G30	0,2
	Nemoci oběhové soustavy	I11, I46, I49, I50, I51	1,0; 2,5; 3,0; 5,0; 1,5
	Nemoci dýchací soustavy	J96	28,0
	Nemoci trávicí soustavy	K72	6,5
	Příznaky, znaky a abnormální klinické a laboratorní nálezy nezařazené jinde	R09, R57, R62, R68, R95, R99	3,0; 3,0; 7,0; 3,0; 0,0; 1,0
	Poranění, otravy a některé jiné následky vnějších příčin	S01, S09, S21 T01, T07, T40, T50, T71	7,0; 13,0; 2,8 8,0; 59,0; 5,0; 13,0; 8,5

Zdroj dat: NBER (2020), výpočet vlastní

V posledním sloupci tabulky 11 je poměr stupně dovnitř ku stupně ven. Hodnoty vyšší než jedna značí, že příčina ve dvojici se smrtí bývá často důsledkem jiných kódů, což může indikovat, že se jedná o časté bezprostřední příčiny smrti, nebo v případě S a T kódů spíše smrtelné důsledky nehod a úrazů. Nejvíce nevyrovnaný poměr ve prospěch inD mají u žen i mužů sekundární zhoubné novotvary (C78 a C79), I50 („selhání srdce“), J96, S01 a T01 (otevřené rány), nebo T07 („neurčená mnohočetná poranění“). Pouze u žen je to A41, pouze u mužů lze k výčtu řadit i T40 a T50 (otravy narkotiky nebo léčiv), K72 („selhání jater nezařazené jinde“), ještě jeden kód pro rány hlavy (S09) a další T kódy stojící pro úrazy, nebo zadušení (T71).

Příčiny s poměrem inD ku outD menší než jedna oproti tomu bývají často na počátcích chorobných řetězců, přičemž se zároveň signifikantně vážou na smrt. Lze spekulovat, že se jedná o choroby hojně se vyskytující samostatně na hlášení o smrti, což může takto být například i pro jejich smrtelnost, nebo v případě G30 („Alzheimerova nemoc“) i kvůli tomu,

že se jedná o příčinu smrti typičtější pro spíše vysoký věk, kdy se celkově průměrný počet příčin zapsaných na hlášení snižuje v porovnání s věky nižšími. Kromě G30 jsou v tabulce 11 pro obě pohlaví společné i další choroby, které mají nižší, nebo vyrovnaný poměr mezi inD a outD, jedná se o C25 („zhoubný novotvar slinivky břišní“), a R95 („syndrom náhlé smrti dítěte (kojence) [SIDS]“), jen u žen ještě C56 („zhoubný novotvar vaječníku“), jen u mužů má například I11 („postižení srdce při hypertenzi“) vyrovnaný poměr mezi inD a outD.

Může vyvstávat otázka, proč se sepse, která je v literatuře u obou pohlaví často zmiňována jako poměrně významná vícečetná příčina smrti, jež však bývá v klasických jednočetných analýzách opomíjená kvůli tomu, že nebývá psána jako základní (Pechholdová, 2017), neváže též u mužů na smrt, jako je tomu u žen. Jednak je záhodno zmínit, že podíl mužů, nepřecházejících z A41 do smrti je 0,51 (u žen činí zhruba 0,47), tudíž jejich podíl je poměrně vyrovnaný s těmi, kteří z A41 do smrti jdou. Zároveň pouze u mužů bývá sepse signifikantní příčinou respiračního selhání (J96), které, protože je velmi častým kódem, významně snižuje podíl lidí jdoucích ze sepse kamkoli jinam, tedy i do smrti.

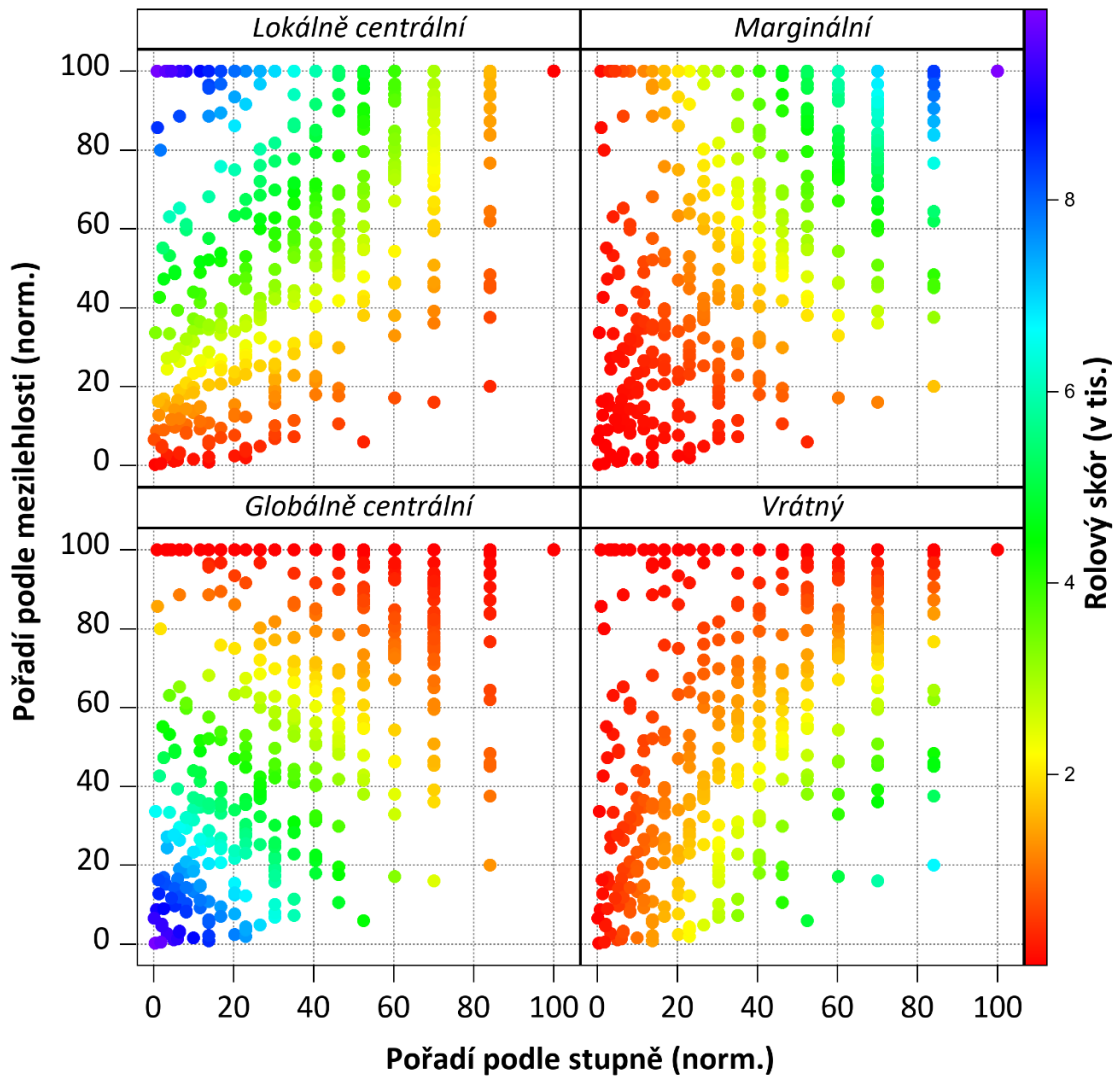
Které další příčiny, ač splňují podmínku dostatečné prevalence, se na smrt neváží kvůli vyššímu podílu lidí jdoucích do jiných stavů, nikoli smrti? Mezi nimi lze nalézt v první řadě chronická onemocnění, z nichž pro muže i ženy jsou společná F03 („neurčená demence“), I25 („chronická ischemická choroba srdeční“), I10 („primární hypertenze“), J44 („jiná chronická obstrukční plicní nemoc“), N18 („chronické onemocnění ledvin“) nebo G93 („jiná onemocnění mozku“), u mužů i E14 a E11 (typy diabetu), K70 („alkoholické onemocnění jater“) a K74 („cirhóza jater“), G20 („Parkinsonova nemoc“) a F10 (poruchy chování kvůli závislosti na alkoholu). K příčinám smrti uvedeným na hlášení by se měl vyplňovat i přibližný časový interval mezi nástupem daného stavu a smrtí a tyto intervaly by měly být kratší směrem ke svrchnějším řádkům části I HOÚ (viz obrázek 1) (ICD-10 Mortality Manual, 2014). Pokud tedy výše uvedené příčiny v tomto odstavci nejsou na hlášení v části I samy (a to nemusejí být, protože musí splňovat podmínku vysoké častosti, což předznamenává vyšší stupeň, viz obrázek 12), nebývají uváděny na nižší řádky, protože se jedná o nemoci spíše dlouhodobějšího charakteru.

Na smrt se však neváží ani mnohé vysoce prevalentní kódy, které by se asi nedaly nazvat jako chronické, nebo „dlouhodobějšího charakteru“, sice N17 („akutní selhání ledvin“), N19 („neurčené selhání ledvin“), J69 a J18 (druhy pneumonií), I21 („akutní infarkt myokardu“), u mužů i C18 („zhoubný novotvar tlustého střeva“). U N kódů lze vysvětlení hledat v jednom konkrétním důsledku, na který se tyto váží, sice E87 („jiné poruchy tekutin, elektrolytů a acidobasické rovnováhy“). Je to stav, kdy jsou v těle v nerovnováze některé minerální látky, což právě bývá běžně asociováno se selháním ledvin (Dhondup et al., 2017). Podobně i pneumonie často vedou k dalším respiračním nemocím, které pojmenovávají možná méně „obecné“ stavy, konkrétně vedou třeba k J12 („virový zánět plic“), nebo k J96 („respirační selhání“). Trajektorie pneumonií jsou i takové, že vedou k infekčním, bakteriálním nemocím, jež se potom opět váží na některé J kódy. Rovněž odchody z pneumonií směřují do R, jmenovitě R06 („nepravidelnosti dýchání“). Tedy, když bývají do hlášení vyplňovány nejčastější akutní stavy, bývají spíše mnohdy příčinami konkrétnějších, méně častých chorobných stavů nežli bezprostředními příčinami smrti, chorobami vázajícími se přímo na smrti samotnou.

7.4 Jaké role hrají uzly v síti?

Na obrázku 22 jsou proti sobě vynesena normalizovaná pořadí z hlediska mezilehlosti uzlu a jeho stupně. Graf je rozdělen do čtyř panelů podle v metodologické části definovaných typů rolí uzlů. Rozložení datových bodů v každém panelu je totožné jako u tří zbylých, avšak body symbolizující choroby se odlišují barvou. Tato se odvíjí od výše skóru, který byl za danou chorobu spočten podle metodiky v kapitole 4. Čím „chladnějším“ odstínem je bod podbarven, tím více jasně by o něm mohlo být tvrzeno, že má charakter toho kterého rolového typu, neboť v něm nabyt vysokého skóru. Protože se jedná o pouhé schematické zobrazení (nejde o to, která choroba je přesně která, nýbrž pouze o ilustraci, že např. vrátných je málo, naopak globálně centrálních uzlů i marginálních, zejména oněch ryze marginálních, zase nejvíce) jsou v obrázku 22 prezentována jenom data za ženy. Obrázek za muže se výrazně z hlediska umístění datových bodů neodlišuje.

Obrázek 22: Role uzlů v síti z hlediska stupně uzlu a jeho mezilehlosti, ženy, USA, 2018



Zdroj dat: NBER (2020), výpočet vlastní

Síť příčin smrti téměř neobsahuje choroby, které hrají role ryze vrátných. Uzly s vysokou mezilehlostí mívají také zpravidla vyšší stupně. Tedy, příčiny smrti, které obvykle nejsou vyhraněnými spouštěči chorobných řetězců, ani nestojí vyloženě na jejich koncích (to by potom měly nízké CB), bývají chorobami do úmrtních listů často psanými v kombinacích s odlišnými dalšími kódy, protože mají relativně vyšší stupně. Nejvíce neplatné je toto tvrzení lze pro několik málo kódů, které jsou v prvním panelu v obrázku 22 znázorněny zeleně. U žen se jedná v první řadě o B99 („jiné a neurčené infekční nemoci“). U obou pohlaví se jedná o chorobu, která bývá jako jedna z mála často psána jako důsledek k A41, což je také důvodem vysokých hodnot mezilehlosti B99. U mužů má tato choroba vyšší stupeň, zároveň také sama o sobě není žádná další chorobě příčinou, proto nemá oproti ostatním uzlům význačně vyšší mezilehlost, ani nízký stupeň, aby byla ryzím vrátným.

U obou pohlaví by jako vrátný mohlo být označeno i P27 („chronické respirační nemoci vzniklé v perinatálním období“). Důvod byl již částečně naznačen v metodologické části, a to že vysokou mezilehlost mívají také uzly figurující jako „mosty“ mezi jinak nepropojenými částmi sítě. Přesně takové je i P27, které je jedním z mála uzlů, které propojují klastr kojenecké úmrtnosti se zbytkem sítě, konkrétně je příčinou k I27 („jiné kardiopulmonální nemoci“). Pravidla pro společné zapisování kódů pro perinatální úmrtnost a ostatních popisuje WHO (2016). Podle ní bez ohledu na to, zda je úmrtnost kojenců zaznamenávána na zcela specifické HOÚ, nebo na standardní, ale kromě segmentu pro zápis příčin smrti vyobrazený na obrázku 1 i do sekce vyhrazené pro údaje o matce, sledují se nejen příčiny smrti novorozence, ale rovněž i nemoci nebo obecněji zdravotní stavy matky. Vazby mezi P a některými srdečními nebo respiračními nemocemi tedy vyjadřují, že k základní příčině smrti novorozence (P kód) přispěla choroba matky (I, J kódy), jak popisuje WHO (2016).

Podobné vysvětlení se nabízí i pro poměrně vysoký skór vrátného u choroby I95 („hypotenze“). Bývá totiž u žen příčinou k P20 („nitroděložní hypoxie“), u mužů k P22 („dechová tíseň novorozence“) a zároveň je i jedním z mála důsledků sepse, protože přes tuto chorobu má potenciál vést velmi mnoho nejkratších cest. Z čehož plyne, že postavení kódu I95 v mužské síti je ovlivněno i morbiditou žen, protože bývá zapisován matkám chlapců zemřelých v kojeneckém věku.

Závěrem má smysl k vrátným uvést, že mezilehlost nemocí souvisí s mezilehlostí jejich sousedů, příkladem čehož je u žen M41 („skolióza“). Tato je jedním z mála důsledků výše uvedeného vrátného I27 a příčinou J98 („jiné poruchy dýchací soustavy“), což jí staví do pozice choroby, která se váže na nemoc nevzdálenou od perinatální úmrtnosti, taktéž sama má vztah ke kardiovaskulárním nemocem, nebo k chorobám dýchací soustavy.

Uzlů s charakterem globálně centrálních je již o poznání více. S ohledem na vztah mezi prevalencí a stupněm uzlu lze čekat, že těmto uzlům dominují uzly nejčastěji se objevující na HOÚ. U obou pohlaví se jedná především o A41, I10 a J44, u žen ještě F03, u mužů s lehkým odstupem C34 a N18. Ač platí, že se většinou jedná o choroby se spíše vyšší prevalencí, existují výjimky, tudíž jsou mezi globálně centrálními nemocemi přítomny i kódy nečasté. Pro ně je společné, že jsou většinou chorobami centrálními pouze ve svých klastrech, ale zároveň mají díky svým sousedům spíše vyšší mezilehlost, příkladem ono P22 u mužů a P20 u žen, tj. centrální příčiny smrti v klastru s kojeneckou úmrtností a také spojky s hlavní částí sítě.

Dále jsou takovými především u mužů i některá zranění (S36: nitrobrěšních orgánů, S22: žeber), u žen třeba D68 („jiné poruchy koagulace“) nebo i vzácné Q24 („jiné vrozené vady srdce“), které je centrem mezi Q kódy i příčinou I27.

Vysokých hodnot globální centrality (ne však zcela nejvyšších) dosahují i některé choroby nacházející se zcela na konci sestupně řazeného žebříku prevalence, u žen je tokovou nejdůležitější F11 („poruchy duševní a poruchy chování způsobené užíváním opioidů“). Tento kód bývá důsledkem otrav a M nemocí, příčinou srdečních chorob a R52 („bolest nezařazená jinde“) (souvisí s M chorobami) a také má reciproký vztah s dalšími duševními onemocněními. I když tedy F11 náleží do klastru se srdečními chorobami, figuruje rovněž jako most mezi řadou dalších oblastí příčin smrti. U mužů mají podobnou vlastnost některá méně běžná zranění a Q choroby.

Lokálně centrální uzly bývají významné jen v určitých oblastech sítí a jsou ukotveny tak, že se nevyznačují velkým dosahem. Nejvyššího skóru lokální centrality dosahuje smrt, C97 a mnohé vnější příčiny, spadající především do klastru s úrazy a zraněními, z čehož je možno soudit, že vysokého skóre v této charakteristice nabývají choroby s nevyrovnaným poměrem stupně dovnitř a ven, nebo ležící v podgrafových, které obsahují příčiny smrti vázající se na „specifickou“ oblast úmrtnosti, která je v síti zároveň zastoupena hojným počtem chorob. Lokálně centrální charakter mají tím pádem i jedny z nejčastějších příčin smrti, sice u obou pohlaví J96 („respirační selhání nezařazené jinde“) a I21 („akutní infarkt myokardu“), u žen N19 („neurčené selhání ledvin“), nebo T07 („neurčená mnohočetná poranění“), u mužů C78 („sekundární zhoubný novotvar dýchací a trávicí soustavy“), T40 („otrava narkotiky a psychodysleptiky (halucinogeny)“), T71 („(za)dušení“) nebo S01 („otevřená rána hlavy“). To, že vysokého skóru lokální centrality dosahuje velké množství uzlů z klastru vnějších příčin smrti, má dopad na skóry autorit a hub, je-li síť brána jako celek. Na tento výrok je navázáno při odpovídání na výzkumnou otázku v následujícím oddíle textu.

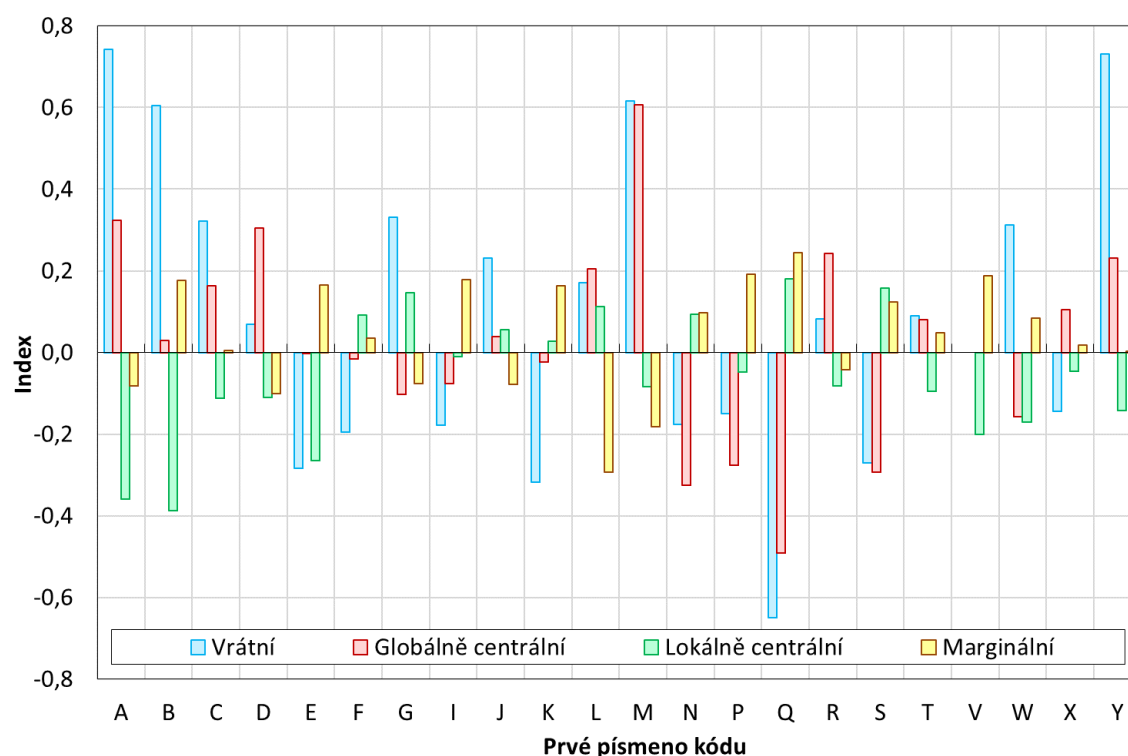
Posledním typem role jsou uzly marginální, kterých je u mužů i žen zcela nejvíce (datové body na obrázcích se mohou překrývat). Pro marginální uzly platí analogie téhož, co pro globálně centrální, tedy, že většinou nejvyššího marginálního skóru dosahují choroby s nízkou prevalencí. Výjimkami jsou pak u obou pohlaví R09 („jiné příznaky a znaky týkající se oběhové a dýchací soustavy“) a I51 („komplikace a nepřesně určené a popsání nemoci srdce“). V obou případech se jedná o „předstupně smrti“, kdy první zmíněný bývá součástí řetězce J choroba→R09→smrt, druhý I choroba→I51→smrt. Kvůli striktně jednosměrným vazbám nemají R09 a I51 prostor pro nabytí vyšší mezilehlosti, zároveň ale ani nesousedí s větším počtem uzlů, protože jsou marginální, nikoli však zároveň málo časté, jako jejich druzi.

U žen je dále jedním z nejprevalentnějších marginálních uzlů I49 („jiné srdeční arytmie“), který má podobné postavení jako I51 a dále jsou to mnohé R kódy, mezi nimiž nejčastěji R54 („stáří (senilita)“) a R06 („nepravidelnosti dýchání“). To, že jsou R časté, ale paradoxně marginální, může být spíše odrazem přijatého způsobu tvoření příčin a důsledků, než že by skutečně v síti neměly „důležité“ místo. Jako iniciátor chorobného řetězce vedoucího ke smrti byl totiž uvažován kód oficiálně stanovený jako základní příčina smrti, čímž byly choroby R, které by jako UCD zapsány být neměly a v 99 % ani nejsou (viz kapitola 3), znevýhodněny v rolích příčin, neboť aby mohly být příčinou, musel existovat dostatečný počet osob, který měl na hlášení podobný řetězec o alespoň 3 chorobách, z nichž aspoň jednu R.

U mužů jsou marginálními uzly se spíše vyšší prevalencí rovněž R choroby, ale na rozdíl od žen ještě N28 („jiné poruchy ledviny a ureteru nezařazené jinde“) a G31 („jiné degenerativní nemoci nervové soustavy nezařazené jinde“). V případě žen není marginální povaha těchto chorob uplatnitelným pojmem a vlastně ani u mužů, je-li N28 zmiňováno v kontextu s bimodálním grafem (viz dále).

Závěrem k typům rolí je na obrázku 23 porovnání mužů a žen z hlediska odlišnosti v rolích na úrovni prvních písmen kódů. Byl spočítán průměrný skóre za každou ze skupin uzlů začínajících stejným písmenem, který byl následně relativizován sumou skóre přes všechny kapitoly, to celé za každé pohlaví zvlášť. Tyto podíly jednotlivých skupin kódů začínajících stejným písmenem na celkovém součtu skóre v každé z čtyř typů rolí byly dány do poměru žen vůči mužům a od výsledku byla odečtena jednička pro centralizaci na nulu. Výsledné hodnoty byly zaneseny do obrázku 23 a vyjadřují, jakou podobu má rozdíl mezi ženami a muži z hlediska míry, do níž má daná skupina tendenci ležet na každé ze čtyřech vymezených rolových typů. Pokud je rozdíl mezi pohlavími kladný, mají kódy ze sítě žen v této kapitole tendenci ležet na dané roli více, než je tomu u mužů. Záporný rozdíl značí, že v dané kapitole a roli jsou naopak kódy častější spíše u mužů.

Obrázek 23: Srovnání rolí uzlů u mužů a žen na úrovni prvních písmen kódů, USA, 2018



Zdroj dat: NBER (2020), výpočet vlastní

V roli vrátých mívají ženy vyšší skóre zejména v kapitolách A, Y, B a M. Naopak v prospěch mužů je rozdíl nejzřetelnější v Q chorobách. Choroby A a B mívají muži častěji v rolích lokálně centrálních, což může souviset s tím, že podle některých studií mívají muži tyto choroby závažnější (viz interpretace podskupin) a také s tím, že velký podíl B chorob je spjatý s HIV/AIDS, která právě u nich také formuje poměrně svébytnou skupinu uzlů. Rozdíly mezi muži a ženami v M chorobách byly již rozebírány dříve. Zmíněné se odráží i na skórech zde.

Výše rozdílu u chorob Y reflektuje spíše skutečnost, že Y choroby nabývají ve skórech vrátných u obou pohlaví velmi nízkých hodnot, na úrovni kterých je dvojnásobného rozdílu mezi muži a ženami dosáhnout snazší (absolutně se tyto skóry liší málo). Ve skutečnosti bývají u obou pohlaví Y choroby nejčastěji marginální.

Muži mají vyšší lokální centralitu nejen v A a B, ale i v E chorobách nebo ve vnějších příčinách. Absolutně mívají muži více vyhraněně lokálně centrálních uzlů než ženy, čemuž tak není v žádném dalším typu role. To může být jednak z důvodu větší mohutnosti podgrafu 111, jednak se může jednat o souvislost s poznatky o kterých píše Kalgotra (2017). Dochází ke zjištění, že propojenost rozličných částí celé mužské sítě je i podle výsledků jeho SNA nižší. Toto staví do souvislosti s tím, že muži navštěvují lékaře v průměru méně často, a proto může být historie jejich zdravotního stavu méně zaznamenaná. Svou studií tedy upozorňuje na existenci „health-survival paradox“ (Kalgorta, 2017, s. 1), tedy, že muži žijí kratší dobu, ale podle výsledků analýz sítí komorbidit jsou „zdravější“³⁴ ve smyslu toho, že jsou to spíše ženy, kdo trpívá častěji na kombinace rozličných nemocí. Zde v práci sice disponují vyšším průměrným stupněm orientované sítě muži, to je však způsobeno inkorporací úmrtnosti na vnější příčiny. Jak bude komentováno u bimodální sítě, provázanost UCD a příčin smrti z druhé části HOÚ, kam se podle některých autorů teprve zapisují komorbidit v pravém slova smyslu, je u žen dalece hlubší.

Shrnuto, vyšší lokální centralita u mužů může být buď odrazem větší spletnosti podgrafu 111 (poněvadž je to velmi z hlediska množství kódů rozsáhlá specifická oblast úmrtnosti se zbytkem sítě propojená spíše slabě), nebo menší propojenosti rozličných částí sítě, je-li uvažována jako celek. Pro druhé svědčí nejen citace výše, ale i výsledek G-N algoritmu, který mužskou síť rozkládá na menší počet větších podgrafů, i když v některých případech s vyšší modularitou.

Z hlediska globálně centrálních uzlů převazují skóry za ženy nad skóry za muže nejvíce v M, naopak vyšší u mužů jsou v N, Q, P a S. U chorob N je toto důsledkem skutečnosti, že ženy mívají více různých a méně častých N nemocí do párů s nemocemi z odlišných kapitol (C, K nemocí). Proto také mívají ženy častěji N nemoci jako marginální. Obrázek 23 ovšem vypovídá o opaku, avšak to jen proto, že celkově mají ženy vyšší marginální skóry přes všechny kapitoly než muži. Jinak řečeno, vyšší globální centralita N chorob u mužů nevypovídá o tom, že by tyto choroby byly „významnější“ v síti mužů, nýbrž o tom, že v síti žen tvoří tyto choroby nejenom významnou samostatnou oblast úmrtnosti, ale i jsou častými méně významnými komorbiditami k dalším nemocem, čemuž tak není u mužů, což u nich zapříčiňuje nižší marginálnost N (v součtu přes všechny rolové skóry musí vycházet 10^4).

V marginálním skóru je nejnevyrovnanější poměr v kapitolách L a M, a to ve prospěch mužů, kterým jsou tyto choroby častěji méně centrální nemoci v síti, jak vyšlo také najevo při analýze klastrů.

³⁴ Termín je v uvozovkách, neboť podstata paradoxu nespočívá v tom, že jsou muži skutečně zdravější, ale že se takovými jeví z dat být, nehledě na příčinu (např. důvodem může být samozřejmě méně kontaktů s institucemi sbírajícími data o chorobách).

7.5 Kdo jsou autority a hubs?

Skór autority (AS) choroby je součet hub skórů (HS) uzlů, které na ni ukazují, HS choroby je součet AS uzlů na ni ukazujících. V předcházejícím oddíle textu u komentáře k lokálně centrálním uzlům bylo naznačeno, že existují oblasti grafu, kde jsou nahuštěny lokálně centrální uzly (podgrafy 111), což je důvodem, proč jsou zde v práci metriky AS a HS generovány izolovaně v rámci každého z podgrafů, nikoli za celou síť jedním rázem. Počítání AS a HS za kompletní graf totiž ústí v téměř nulové hodnoty těchto metrik u většiny uzlů neležících v podgrafech 111, což souvisí s tím, že kompletní graf je poměrně dobře rozložitelný na menší části (modularita přes 0,7), které zároveň, jak bylo zjištěno dříve, si mnohdy nejsou sobě navzájem podobny architekturou, konkrétně podgrafy 111 obsahují u žen asi pětinu, u mužů asi čtvrtinu všech vazeb v síti, a to mezi řadou lokálně centrálních uzlů. Výše modularity indikuje, že podgrafy, v tom i 111, nesoucí až dvojnásobně více vazeb než po něm následující největší, mají spíše slabší vazebnost ke zbytku systému. Z definice AS a HS plyne, že velikost relativně izolované komponenty grafu se odráží na výši hodnot těchto ukazatelů. V tomto konkrétním případě tak, že v podgrafech 111 budou AS a HS vycházet významně vyšší než kdekoli jinde v síti, protože v ostatních částech systému se na sebe vzájemně nikde toliko hustě neváže tak velký počet uzlů, navíc řada z nich se spíše vyšší centralitou. A protože je AS a HS mírou, která se vztahuje k určité hladině, která je dána globálními maximy AS a HS, nemají ostatní uzly šanci se s oněmi z podgrafu 111 svými hodnotami AS a HS rovnat.

Mimo podgraf 111 se však také vyskytuje řada uzlů s nejvyššími stupni. To, že by jejich AS a HS počítané za celý graf dosahovaly spíše nízkých hodnot v porovnání s uzly z podgrafu 111, nasvědčuje skutečnosti, že v částech sítě mimo oblasti se „specifickou“ úmrtností choroby s vysokými stupni nenasazují vzájemně mezi sebou vztahy, natož reciproké. Kontrastní k tomuto zjištění je, že v sítích z druhé části hlášení a UCD panuje přesně opačná zákonitost, sice, že se na sebe navzájem velmi časté nemoci sklony navazovat mají.

Příloha 6 obsahuje totožný rozklad systému na komunity jako příloha 5, jen grafika je nastavena podle odlišných parametrů. Barva uzlů se odvíjí od výše HS – čím vyšší hodnoty skóru choroba v rámci komunity nabývá, tím studeněji je podbarvena. Velikost uzlů je proporcionální k hodnotě AS – čím vyšší je, tím větší poloměr uzlu má.

AS a HS jsou zde v práci použity pro rozlišení uzlů univerzálně populárních a skutečně „vlivných“ a nakonec pro identifikaci „mocných“ uzlů, které dosahují vysokých AS a zároveň HS. Než však je k samotné analýze přistoupeno, může se nabízet otázka, zda vůbec má smysl AS a HS rozebírat. Nemělo by, pokud by tyto metriky byly vysoce korelované s ostatními dosud použitými. Hodnoty korelačních koeficientů v tabulce 12 signalizují, že tomu tak v některých případech skutečně je. Podle očekávání jsou zřejmě společně závislé stupeň dovnitř a AS a stupeň ven a HS, i když slaběji. To se také promítá do zřejmě existujícího vztahu mezi stupněm (celkovým) a AS s HS. To, že korelační koeficient HS zejména s prevalencí vychází poměrně slab, znamená neplatnost tvrzení, že choroby, které jsou na pozici příčiny k důsledku, jenž je autoritou, bývají zároveň také chorobami častými. V rámci klastrů tedy zapříčiňují chorobné řetězce k autoritám spíše kódy méně frekventované.

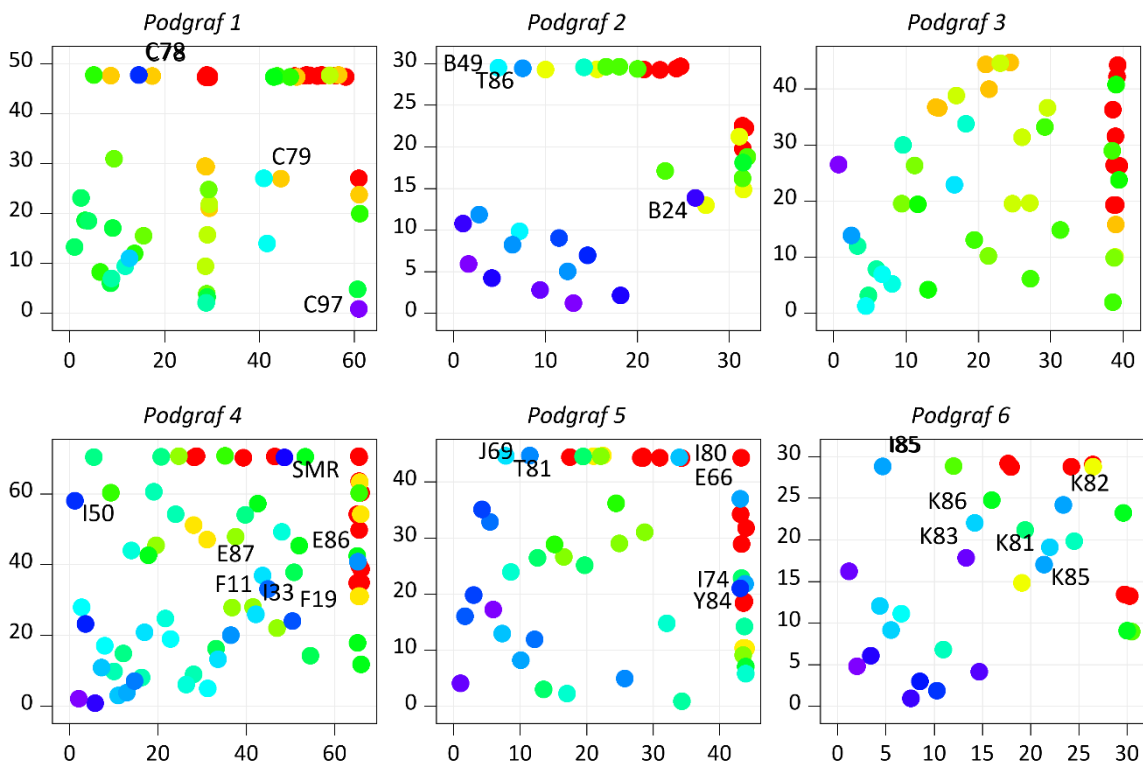
Tabulka 12: Korelační koeficienty mezi použitými metrikami, ženy a muži, USA, 2018

	Stupeň dovnitř	Stupeň ven	Skór autority	Skór hub	Prevalence
Stupeň	0,72 (0,66)	0,77 (0,76)	0,52 (0,49)	0,36 (0,34)	0,79 (0,78)
Stupeň dovnitř		0,22 (0,13)	0,82 (0,81)	-0,06 (-0,14)	0,67 (0,63)
Stupeň ven			0,09 (0,05)	0,64 (0,64)	0,55 (0,51)
Skór autority				-0,11 (-0,09)	0,52 (0,52)
Skór hub					0,21 (0,18)

Pozn.: korelováno pořadí

Zdroj dat: NBER (2020), výpočet vlastní

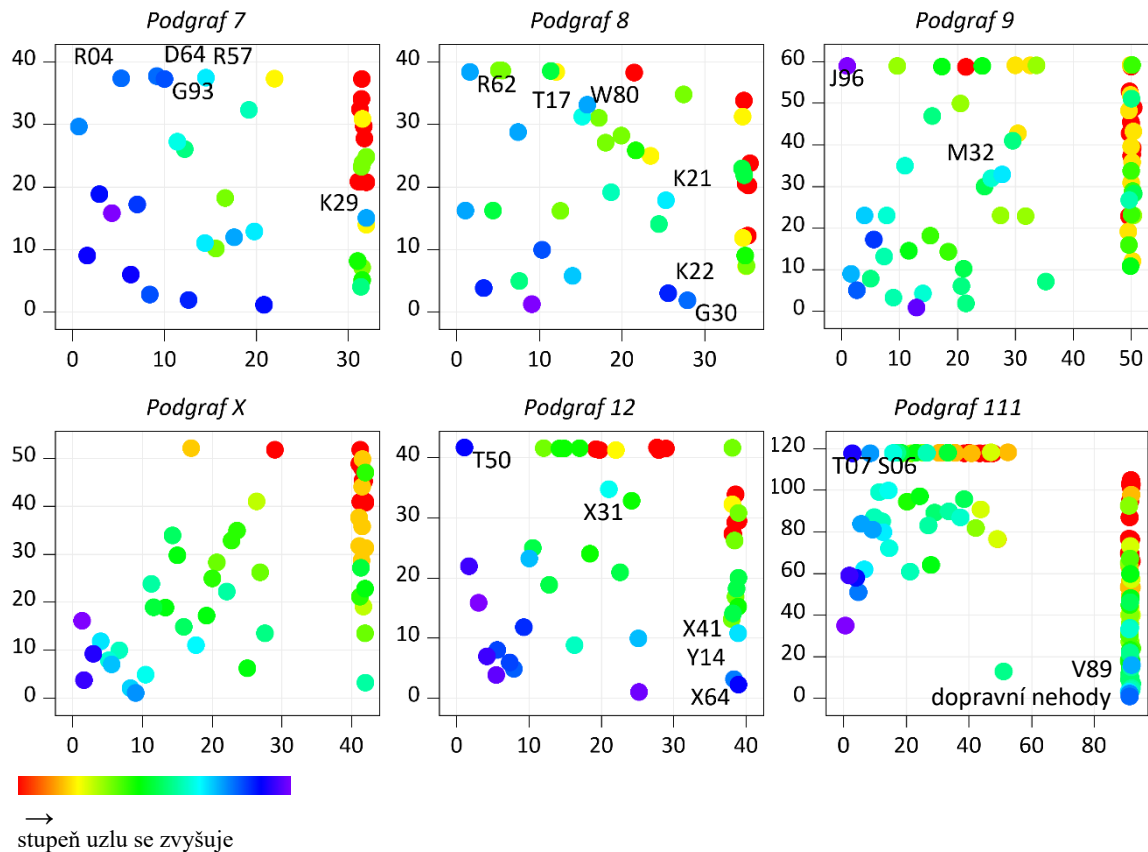
Zřejmě však stále existují choroby, pro něž je typická diskrepance mezi stupněm a AS s HS, tedy uzly univerzálně populární. Ty, které by se daly takto označit, disponují spíše vyššími stupni v porovnání s nižšími AS a HS. V obrázcích 24a a 24b za ženy a 25 za muže mají povahu univerzálně populárních choroby, které jsou představeny datovými body podbarvenými „studeneč“ a zároveň lokalizovanými spíše v horních pravých rozích panelů. Samozřejmě takto vágní vymezení univerzálně populárních uzlů umožňuje nalézat nanejvýše choroby, které by mohly mít rysy takových uzlů, nikoli ty, které to přesně jsou a které nikoli.

Obrázek 24a: Pořadí chorob z hlediska AS, HS a výše stupně v podgrafech, ženy, USA, 2018

→
stupeň uzlu se zvyšuje

Pozn.: Podgraf 1 = novotvary; Podgraf 2 = novotvary a nemoci krve, krevetvorných orgánů; Podgraf 3 = sepse a další sousedé; Podgraf 4 = srdeční choroby (choroby oběhové soustavy); Podgraf 5 = choroby trávicí soustavy a komplikace; Podgraf 6 = choroby jater a žlučnku; Na ose x je pořadí podle skóru autority, na ose y je pořadí podle skóru hub.

Zdroj dat: NBER (2020), výpočet vlastní

Obrázek 24b: Pořadí chorob z hlediska AS, HS a výše stupně v podgrafech, ženy, USA, 2018

Pozn.: Podgraf 7 = nemoci cévní, trávicí soustavy a krve; Podgraf 8 = demence, Podgraf 9 = respirační nemoci; Podgraf X = kojenecká úmrtnost; Podgraf 12 = otravy; Podgraf 111 = vnější příčiny smrti.

Na ose *x* je pořadí podle skóru autority, na ose *y* je pořadí podle skóru hub.

Zdroj dat: NBER (2020), výpočet vlastní

U žen (obr. 24a, b) v podgrafech se sepsí (3³⁵), kojeneckou úmrtností (X) nebo vnějšími příčinami (111) takřka neexistují uzly, které by disponovaly relativně vyššími stupni, ale nižšími AS a HS. V těchto komunitách více než v ostatních spíše platí, že čím nižší je stupeň uzlu, tím nižší jsou rovněž i metriky AS a HS³⁶ což indikuje silnější vzájemnou provázanost uzlů, na něž se váží i marginální choroby. Tyto dominantní uzly jsou v příloze 6 vyobrazeny s větším poloměrem, nebo jsou podbarveny „studeně“.

Oproti tomu zejména v podgrafech s převahou srdečních nemocí (4) nebo nemocí TS (5, 6) u žen mohou existovat některé uzly s rysy univerzálně populárních. V prvním zmíněném klastru se jedná o „fiktivní“ uzlu pro *smrt*, jež byla při tvorbě dvojic příčin a důsledků zapsána všem zemřelým. Poměrně nevysokými AS a HS v porovnání se stupněm disponují v podgrafu srdečních nemocí také choroby E86 a E87 (poruchy přeměny látek), což může být odrazem u otázky „jaký je systém“ zmiňovaného poznatku ohledně tendence těchto chorob mít spíše pestré sousedy, a to mimo klastr do něhož samy spadají (kvůli relativně nižším AS i HS). Také poruchy způsobené užíváním drog (F11, F19) nebo i některé srdeční nemoci by se výší prevalence mohly rovnat s uzly s nejvyššími AS a HS v rámci podgrafu 4, avšak samy právě v metrikách AS a HS zaostávají. Opět lze připomenout, že zejména uvedené F choroby se u žen vyznačovaly pestrým sousedstvem.

³⁵ Čísla v závorkách označují čísla podgrafů.

³⁶ Což neznamená, že zároveň choroby se spíše vyšším stupněm mají zároveň také spíše vyšší AS i HS.

V podgrafech 5 a 6 mohou mít povahu univerzálně populárních uzlů např. I80 (druh zánětu žil), E66 („*obezita*“). V okolnostech, za jakých je E66 zapisována do části 1 HOÚ a do části 2 lze tedy spatřovat rozdíl. U výsledků bimodální sítě je totiž evidentní, že do části 2 bývá E66 zapisována společně s nejčastějšími UCD, oproti čemuž v části 1 sousedí jen s některými kódy z E10–E14, I10 („*esenciální (primární) hypertenze*“) a jinak s chorobami marginálnější povahy, s kterýmižto také sdílí pro jejich početnost klastr. V podgrafu 6 jsou ještě univerzálně populárním typem choroby žlučníku.

Navzdory tomu, že mají podobné stupně jako choroby jater, s nimiž obývají klastr, tak jsou to právě spíše nemoci K70–K77 („*nemoci jater*“), které dosahují vyšších AS i HS a to zřejmě díky tomu, že choroby žlučníku nemají přímou vazbu na B18 („*chronická virová hepatitida*“), které je tím pádem příčinou pouze mnohým nemocím jater.

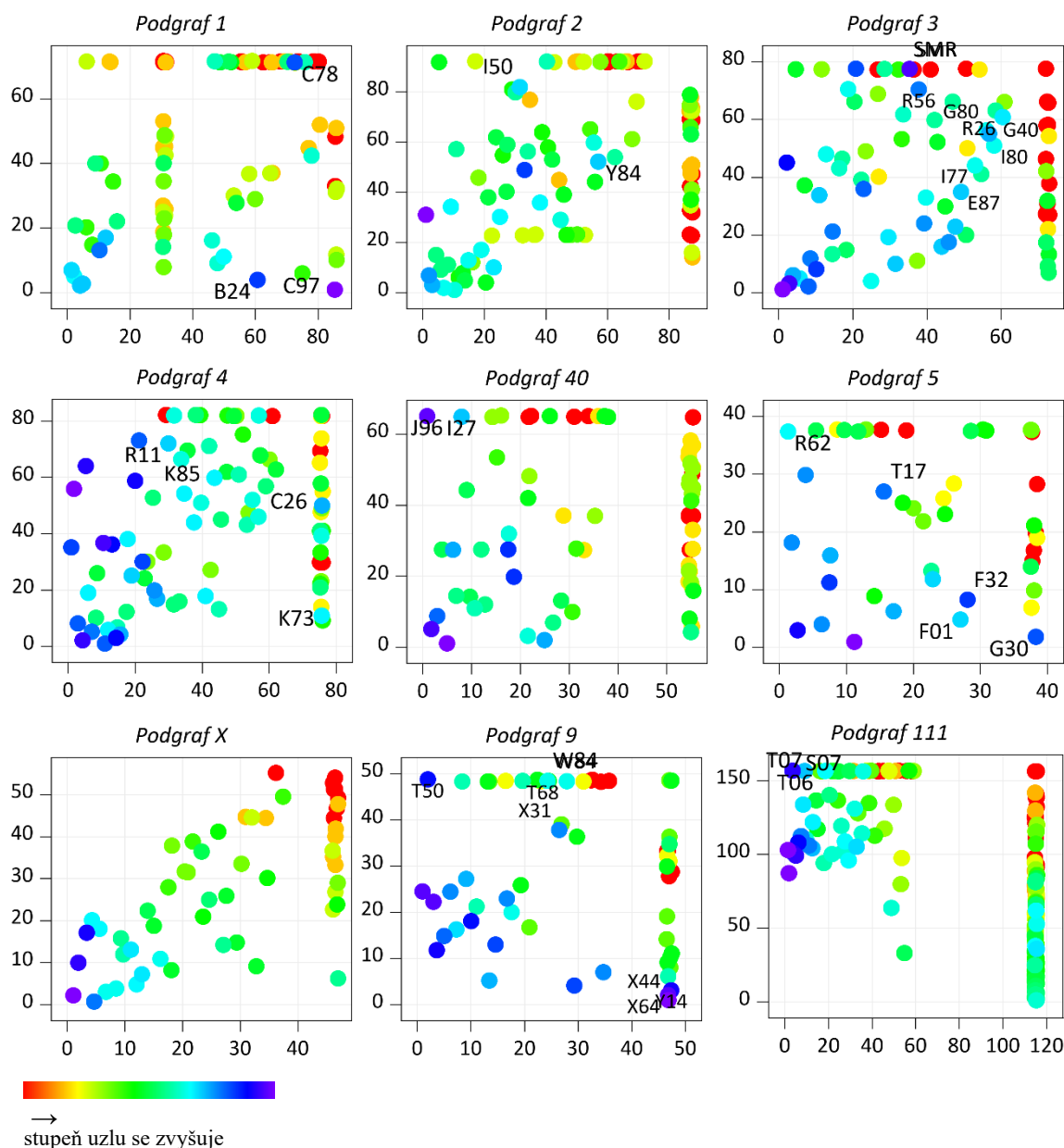
I v některých dalších ženských podgrafech se rýsují uzly inklinující svými charakteristikami spíše k univerzálně populárním. V klastru sdružujícím převážně nemoci cév, krve a některých chorob TS (podgraf 7) je takovým třeba R57 („*šok nezařazený jinde*“), dále v klastru s demencemi (podgraf 8) by to mohly být kódy související s udušením nebo vdechnutím předmětů (W80, T17), v klastru s otravami (podgraf 12) kód X31 („*vystavení nadměrnému přírodnímu chladu*“). Evidentně se tedy jedná o choroby, které bývají součástí spíše „specifických“ chorobných řetězců, nikoli však vyloženě vzácných, jako jiné choroby z jejich komunit s podobnými výšemi AS či HS.

Na obrázku 25 jsou analogické grafy za populaci mužů. Stejně jako u žen i jejich podgrafy kojenecké úmrtnosti (X) a vnějších příčin (111) jsou univerzálně populárních uzlů spíše prosté. Podobný těmto dvěma klastrům je v tomto směru i ten s respiračními nemocemi (40).

Naopak zejména podgrafy s chorobami trávicí soustavy (4) a srdečními (2) mohou obsahovat uzly s nízkými AS a HS a vysokými stupni. V prvně jmenované komunitě může takovým být C26 („*zhoubný novotvar jiných a nepřesně určených trávicích orgánů*“), s odstupem taktéž i choroby z rozmezí K80–K87 („*nemoci žlučníku, žlučových cest a slinivky břišní*“), které ovšem v porovnání se ženami nejsou v populaci mužů toliko častými, jak bylo citováno výše. V podgrafu se srdečními nemocemi mají povahu univerzálně populárních uzlů některá R (R56: „*křeče nezařazené jinde*“, R26: „*abnormality chůze a pohyblivosti*“), dále nemoci nervové soustavy, I80 (druh zánětu žil) a *smrt*, která však dosahuje v AS poměrně ne úplně nízkých hodnot v porovnání s ostatními uzly v jejím podgrafu.

Na rozdíl od žen se u mužů v podgrafu s novotvary (1) nachází uzel výrazněji univerzálně populární, jmenovitě C78 („*sekundární zhoubný novotvar dýchací a trávicí soustavy*“). U žen totiž nabývá vyšších hodnot v HS kvůli sousedství s kódy z rozmezí C51–C58 („*zhoubné novotvary ženských pohlavních orgánů*“).

Obrázek 25: Pořadí chorob z hlediska AS, HS a výše stupně v podgrafech, muži, USA, 2018



Pozn.: Podgraf 1 = novotvary; Podgraf 2 = sepse a souseďé; Podgraf 3 = srdeční choroby (nemoci oběhové soustavy); Podgraf 4 = choroby trávicí soustavy; Podgraf 40 = respirační nemoci; Podgraf 5 = demence; Podgraf X = kojenecká úmrtnost; Podgraf 9 = otravy; Podgraf 111 = vnější příčiny. Na ose x je pořadí podle skóru autority, na ose y je pořadí podle skóru hub.

Zdroj dat: NBER (2020), výpočet vlastní

Vedle uzlů s vlastnostmi univerzálně populárních lze v obrázcích 24 a 25 vidět, že ve většině podgrafů se nachází alfa uzly, které se vyznačují vyššími hodnotami AS i HS a zároveň vysokými stupni. Jedná se tedy o choroby, které jsou zapříčiněny, nebo zapříčiňují nemoci s vyšší centralitou. Alfa uzly se u obou pohlaví spíše nenachází v podgrafech 111. Jak již bylo výše uvedeno, jedná se o klastr obsahující rozličné cesty ke smrti sbíhající se většinou v několika málu důsledcích, které jsou v příloze 6 vyznačeny vyššími poloměry. Podobnou strukturu má u žen i podgraf se sepsí a jejími sousedy (3). U mužů má v tomtéž podgrafu (3) kód A41 postavení výraznějšího alfa uzlu, neboť sama zapříčiňuje některá další bakteriální a virová onemocnění, což není platné pro ženy.

Které další uzly, ač mají vyšší stupně, nemají povahu alfa uzlů z důvodu zaostávání v aspoň některé z metrik AS nebo HS? U žen i mužů v podgrafech s převahou novotvarů (podgrafy 1) je takovým C97 („*zhoubné novotvary mnohočetných samostatných (primárních) lokalizací*“), jako nod objevující se v síti pouze v pozici příčiny a zároveň v klastru s novotvary disponující jedním z nejvyšších stupňů. Podobným příkladem je u mužů i žen v klastrech s respiračními nemocemi (podgraf 4 ženy, 40 muži) uzel J96 („*respirační selhání nezařazené jinde*“), jenž je častým důsledkem, navíc zřejmě i ne marginálních chorob.

Další uzly, které kvůli nízkým hodnotám v některé z metrik AS nebo HS nemohou být označeny jako alfa uzly, jsou k nalezení v podgrafech, kde jsou koncentrovány otravy (ženy 12, muži 9). V těchto klastrech je vysokou autoritou T50 („*otrava diuretiky a jinými a neurčenými léky, léčivy, návykovými a biologickými látkami*“) a naopak významným hubem jsou otravy léky, ať už náhodné (X44), úmyslné (X64), nebo s úmyslem nezjištěným (Y14). Řadu uzlů s podobnými konstelacemi měr AS, HS a D lze nalézt i v hlavním grafu s vnějšími příčinami, sice 111 u obou pohlaví. Jak již psáno bylo, jedná se o uzly s relativně větším poloměrem v příloze 6 (dominantní autority), nebo drobné a podbarvené modře (dominantní hubs). V některých případech nejsou alfa uzly kvůli jedné z měr AS nebo HS ještě R choroby. U mužů i žen se takové nacházejí v podgrafech s demencemi (ženy 8, muži 5), konkrétně již výše rozebíraný R62 („*nedostatek předpokládaného normálního fyziologického vývoje*“). V tomtéž klastru má i G30 („*Alzheimerova nemoc*“) vysoké HS, avšak nízké AS.

Ostatní choroby, které byly dosud rozebírány jako centrální v rámci analýzy podskupin, mají povahu spíše alfa uzlů. Platí tedy, že na úrovni podgrafů mají spíše vyšší AS a HS také choroby s vyššími stupni. Je k repetici, že prezentované výsledky jsou ovlivněny výpočtem AS a HS na úrovni komunit. Uzly, které mají mnoho marginálních sousedů a zároveň se také váží na některé centrální, s nimiž ale nesdílí klastr, měly při tomto přístupu podhodnocené AS i HS.

7.6 Jaké vazby se nevyskytují u opačného pohlaví?

V analýze podskupin bylo diskutováno, že se v síti vyskytují vazby, které se objevují pouze u jednoho či druhého pohlaví. Kromě specifíků plynoucích z anatomických rozdílů se jednalo například i o některé nemoci krve a krvetvorných orgánů nebo kostí, jež jsou i podle studií „ženskými chorobami“, nebo nemoci kardiovaskulární a vnější příčiny, jež bývají nazývány naopak „mužskými“ (Regitz-Zagrosek a Seeland, 2013).

Muži nemají ve své síti asi 25 % vazeb, které mají ženy, ženská síť neobsahuje asi třetinu vazeb, které obsahuje mužská. Větší část vazeb je tedy pro obě pohlaví společná. Zároveň síť s vazbami, které nemají ženy obsahuje trojnásobek nejvyššího stupně uzlu sítě vazeb, jež nemají muži, sice S09 má stupeň 31 (a existuje řada dalších nodů s podobným stupněm), oproti maximu T14 se stupněm 13 (a jedná se o jediný uzel se stupněm vyšším než 10 v síti vazeb, které nemají muži). Tedy, mnoho z hran specifických pouze pro muže, tj. těch ze sítě vazeb, které nemají ženy, je napojeno na malý počet kódů se spíše vyšším stupněm. Centrální choroby jak sítě vazeb, které nemají muži (dál VNM), tak sítě vazeb, které nemají ženy (dál VNŽ) jsou v tabulkách 13 a 14. Samotné sítě rozdílů ve vazbách mezi muži a ženami jsou v přílohách 3 a 4.

Tabulka 13: Uzly s nejvyšším stupněm v síti s vazbami, které nemají ženy

Kód	Celý název nemoci	Stupeň uzlu
S09	Jiná a neurčená poranění hlavy	31
S19	Jiná a neurčená poranění krku	30
S29	Jiná a neurčená poranění hrudníku	30
T14	Poranění neurčené části těla	27
T09	Jiná poranění páteře a trupu, úroveň neurčena	26
S39	Jiná a neurčená poranění břicha, dolní části zad a pánve	19
T07	Neurčená mnohočetná poranění	17
S36	Poranění nitrobřišních orgánů	15
T13	Jiná poranění dolní končetiny, úroveň neurčena	14
B24	Neurčené onemocnění virem lidské imunodeficiency [HIV]	13
C97	ZN mnohočetných samostatných (primárních) lokalizací	12
M86	Zánět kostní dřeně (osteomyelitida)	12
Y43	Nežádoucí účinky léčby – prostředky působící přímo na celý organismus	12
A41	Jiná sepe	11
J84	Jiné intersticiální plicní nemoci	11
N40	Zbytnění prostaty	11
S25	Poranění krevních cév hrudníku	11
S27	Poranění jiných a neurčených nitrohručních orgánů	11
T11	Jiná poranění horní končetiny, úroveň neurčena	11
W31	Kontakt s jiným a neurčeným strojem	11
P36	Bakteriální sepe novorozence	10
T30	Popálenina a poleptání, část těla neurčena	10

Zdroj dat: NBER (2020), výpočet vlastní

Nejvyšší stupně v VNŽ mají z velké části kódy stojící pro poranění. To, že ženy mají méně pestré způsoby, jak se k poraněním dostávají, bylo psáno již povýš. Až v druhé části tabulky 13 jsou choroby nejsoucí důsledky vnějších příčin smrti, nejvyšší stupeň z nich má B24, a to díky tomu, že pouze u mužů má více sousedů mezi dalšími infekčními nemocemi, novotvary TS, nebo meningitidou.

Ač bylo výše uvedeno, že graf za muže je na sousedy chorob M chudší, v tabulce 13 se objevuje M86 („*zánět kostní dřene (osteomyelitida)*“), která byla již v uplynulých pasážích textu vícekrát uváděna jako příklad nodu po různých stránkách odlišného svým postavením v sítích diferencovaných podle pohlaví. Jedná se o jednu z mála M nemocí, která výrazně častěji postihuje muže (Mohsien et al., 2014) a zejména u nich je častou komorbiditou k diabetu, a i proto její výskyt v populaci během posledních dekád zejména u mužů narůstal (Kremers et al., 2015). Podobně jako u M86, i A41, 14. uzel podle výše stupně v síti VNŽ, je příčinou smrti, na níž muži umírají výrazněji častěji než ženy (Nasir et al., 2015). Podle některých studií (např. Angele et al., 2014) ženy na sepsi umírají méně často, protože některé ryze ženské hormony působí protektivně. Na druhou stranu v jiných studiích vědci považují roli hormonů za přeceněnou, např. Ghuman et al. (2013) zkoumali rozdíly mezi chlapci a dívkami do osmi let (v tom i P36, které je rovněž v tabulce 13). Autoři dospěli k závěru, že „*protože hladiny testosteronu a dalších pohlavních hormonů jsou v tomto věku extrémně nízké, usuzujeme, že na převaze mužů v těžké sepsi se podílejí i jiné faktory než testosteron*“ (Ghuman et al., 2013, s. 17).

Tabulka 13 obsahuje i kód J84. Pouze u mužů je totiž často důsledkem chorob z rozmezí J60–J70 („*nemoci plic způsobené zevními činiteli*“). Tyto kódy se zapisují lidem, kteří získali chorobu plic expozicí rizikovým environmentálním faktorům, tj. např. při práci v těžebním průmyslu, s azbestem apod.

Na nižších příčkách tabulky 13 jsou některé další vnější příčiny nebo nemoci orgánů, které ženy nemají. V tabulce 14 jsou uzly s nejvíce vazbami, které naopak nemají muži.

Tabulka 14: Uzly s nejvyšším stupněm v síti s vazbami, které nemají muži

Kód	Celý název nemoci	Stupeň uzlu
T14	Poranění neurčené části těla	13
C53	Zhoubný novotvar hrdla děložního [cervicis uteri]	10
M19	Jiná artróza	10
M32	Systémový lupus erythematodes	10
O99	Jiné nemoci matky ZJ, ale kompl. těhotenství, porod a šestinedělí	10
T43	Otrava psychotropními léčivými nezařazenými jinde	10
W19	Neurčený pád	10
A41	Jiná sepse	9
C50	Zhoubný novotvar prsu	9
C97	ZN mnohočetných samostatných (primárních) lokalizací	9
M81	Osteoporóza bez patologické zlomeniny	9
T42	Otrava antiepileptiky, sedativy-hypnotiky a antiparkinsoniky	9
C55	Zhoubný novotvar dělohy, část NS	8
C78	Sekundární zhoubný novotvar dýchací a trávicí soustavy	8
D64	Jiné anemie	8
D69	Purpura a jiné krvácivé stavy	8
F32	Depresivní fáze	8
J96	Respirační selhání nezařazené jinde	8
R52	Bolest nezařazená jinde	8
T44	Otrava léčivými působícími primárně na autonomní nervovou soustavu	8
T51	Toxický účinek alkoholu	8
W18	Jiný pád na rovině	8

Zdroj dat: NBER (2020), výpočet vlastní

Lze vidět (tab. 14), že nejvyšší stupeň je mnohem nižší, než v VNŽ a také uzly s nejvyššími stupni v VNM tvoří převážně poranění. Více je naopak chorob, přesněji novotvarů orgánů, které nemají muži, nebo nemocí jež jsou považovány za spíše „ženské“, konkrétně choroby pohybového aparátu, závažné pády související s vyšším věkem a M, choroby krve, ale i kód pro bolest. Tyto byly již zmiňovány při analýze podskupin.

V tabulce 14 je však uveden i F32, který bývá rovněž diagnostikován častěji ženám (Wilhelm 2001). U něj však platí, že i v síti VNŽ má stupeň rovný osmi. Deprese je tedy u mužů i žen v síti stejně centrální (ač bývá podle výše citovaných zdrojů považována za nemoc „ženskou“), její sousedé se však u pohlaví většinou odlišují. Prvním rozdílem je více sousedů z řad úmyslných sebepoškození u mužů, kdy pouze u nich se F32 váže i na X72, X73 (střelnou zbraň). V USA muži páchají sebevraždy častěji než ženy a pouze u nich je užití střelné zbraně za tímto účelem nejčastější volbou (Curtin a Martinez, 2020). S tímto může souviset, že pouze u mužů bývá deprese příčinou S01 („otevřená rána hlavy“).

Dále existuje obousměrná vazba mezi F32 a R52. Kód pro bolest se v tabulce 14 vyskytuje také, neboť u mužů má spíše marginální povahu, protože sousedí mimo depresi jenom s M19 („jiná artróza“). Vazby mezi trojicí chorob M19, F32 a R52 poukazují na další rozdíl mezi muži a ženami. Protože u mužů se vazby vyskytují v posloupnosti $M19 \rightarrow R52 \leftrightarrow F32$, lze možná rozumět³⁷, že u mužů, aby mohla M19 být příčinou F32, musejí mít zapsanou R52. U žen existuje mezi M19 a F32 přímá vazba, což je další rozdíl z hlediska sousedstva F32 mezi pohlavími. Posledním podstatným rozdílem v sousedstvu F32 je i to, že u žen se častěji váže na další choroby duševní, kromě F03, jsou to pouze u nich i F29 („neurčené neorganické psychózy“) a F31 („bipolární afektivní porucha“). Podle Weissman a Klerman (1977) mají ženy nejen F32 častěji, ale rovněž k ní častěji mívají i jiné komorbidity související s duševním zdravím, vyjma chorob indikujících zneužívání alkoholu a drog, což jsou choroby duše, které se na F32 naopak váží častěji u mužů.

³⁷ To, že muselo existovat signifikantní množství mužů, kteří zemřeli s touto trojicí, nikoli jen s některou z dvojic z těchto tří kódů, je zřejmé z toho, že R52 je i příčinou F32, zároveň ale nemohly být tyto dva kódy uvedeny na hlášení samostatně, protože by pak R52 muselo být základní příčinou smrti (což nebývá v 99 %, kdy je zapisováno na HOÚ) neboť byly brány ony oficiálně kódované a mezi nimi se R vyskytují pomálu (vyjma R95), viz část s explorační analýzou „ill-defined“ příčin smrti. Před R52 tedy musel být často kódován ještě nějaký stav, v tom i M19.

Kapitola 8

Výsledky analýzy neorientované sítě

Neorientované sítě jsou konstruovány dvě. Prvá zobrazuje vztahy mezi příčinami, které mají alespoň n sousedů z protilehlé množiny společných. Podmnožiny takto vytvořeného standardního bimodálního grafu jsou ilustrovány v přílohách 7a a 7b za ženy, 8a a 8b za muže. Druhá bimodální síť zobrazuje vztahy mezi uzly, které se shodují většinou svých sousedů, jejichž celkový počet náleží k 50 % nejvyšších s těch, co se v datech objevují. Podmnožiny bimodálního grafu shodných sousedů jsou vyobrazeny v přílohách 9a a 9b za ženy a 10a a 10b za muže. Všechny tyto přílohy mohou být posléze základem pro projekci kompletního bimodálního grafu.

Než bude přistoupeno k samotným výsledkům, je znova rekapitulována podstata interpretace toho, co přílohy s nestandardními bipartitními grafy obsahují. Ilustračně, například u žen v síti společného sousedstva u typu UCD (příloha 9a) společně sousedí F32 a M81³⁸. Jejich sousedstvo z druhé množiny, tj. z množiny s CCD, je uvedeno v tabulce 15. Každá z dvojic mezi F32 a jejími CCD, stejně tak jako mezi M81 a jejími CCD, náleží k 25 % nejčastějších dvojic mezi UCD a CCD, neboli obrázku 7 v matici A by v buňkách k nim příslušejícím byla uvedena hodnota 1. CCD, které jsou vyznačeny zeleně, jsou stejné u F32 i M81. Shod je v sumě šest, celkem různých sousedů má každá z příčin devět, většina je tedy pro obě z těchto příčin stejná, protože mezi nimi v grafu shodného sousedstva existuje vztah. Je-li se opět navraceno k obrázku 7, lze uvést, že hodnota šest by byla uvedena mimo diagonálu v matici AA^T a hodnota devět na ní v téže matici.

Analogicky by mohly být interpretovány i vazby v síti přispívajících příčin smrti. Konkrétně u mužů v síti shodného sousedstva je jednou z dvojic i F39 a E53 (příloha 10b). Tyto přispívající příčiny se často zapisují do druhé části HOÚ lidem zemřelým na UCD v tabulce 15. Spolu v síti sousedí, protože aspoň polovina základních příčin smrti, k nimž přispívají, je očividně pro obě příčiny stejná.

³⁸ Výběr ilustračních párů proveden tak, aby se nejednalo o příčiny s vysokým počtem sousedících.

Tabulka 15: Ilustrace z bimodálního grafu shodného sousedstva

Vybraný příklad		Vazby do opačné množiny
Pár UCD	F32	D64, E03, E78, F03, F32, F41, I10, M19, N18
	M81	E03, E78, F03, I10, I48, M19, M81, N18, X59
Pár CCD	F39	E14, I10, G20, F10, F03, G30, X70, X74
	E53	G30, F03, G20, E11, E53, I10

Pozn.: zeleně společní sousedi

Zdroj: zpracování vlastní

8.1 Standardní bimodální graf

Při odpovědích na výzkumnou otázku mířenou na autority a hubs byla citována teze Kalgorta (2017) ohledně „health-survival paradox“. Průměrné stupně standardních bimodálních sítí jsou u žen skutečně vyšší než u mužů, a to bez ohledu na n (tabulka 16). Rozdíly v tomto duchu naznačují, že ženy mohou mít více protkány základní příčiny s přispívajícími než muži, což je interpretovatelné v souladu s podstatou „health-survival paradox“ v kontextu vyšší lokální centrality mužské sítě (návaznost na kapitolu 7, podotázku ohledně rolí uzlů). Předpoklad jeho platnosti se umožňuje naklonit spíše k vysvětlení, že lehce vyšší průměrný stupeň v orientované síti u mužů může být zapříčiněn oněmi vazbami vzniklými v segmentu sítě, v němž se koncentrují vnější příčiny smrti. Na druhou stranu jsou to právě muži, kteří mívají ve vyšších věcích uváděno do části 2 HOÚ v průměru o něco více kódů (obrázek 2), zároveň ale jsou to také oni, kteří mají pravděpodobnost dožití se těchto věků nižší.

Tabulka 16: Průměrné stupně ve standardních bimodálních grafech, ženy a muži, USA, 2018

n	Ženy		Muži	
	Množina UCD	Množina CCD	Množina UCD	Množina CCD
1	111,6	110,3	81,9	75,4
5	59,7	54,9	53,9	51,3
10	30,3	31,8	27,8	25,9
15	22,2	20,2	16,9	16,7
20	15,5	13,0	11,4	11,8

Zdroj dat: NBER (2020), výpočet vlastní

V sítích s vazbami mezi základními příčinami smrti mají mnohé nejčastější UCD (tabulka 1 a 2) rovněž vysoké stupně. Sdílejí tedy mnoho společných přispívajících příčin smrti, avšak v tomto směru zároveň řadu z nich předčívají choroby jiné, které však již nepatří k nejčastějším UCD. Jedná se zejména o řadu I kódů (např. I12: „postížení ledvin při hypertenzi“, I35: „nerevmatická onemocnění aortální chlopně“), dále J (J10: „chřipka způsobená identifikovaným sezónním chřipkovým virem“), E (E87, E78: poruchy přeměny látek), nebo W (W18, W19: druhy pádů). U E chorob lze soudit, že je tato vlastnost vysvětlitelná v kontextu jejich postavení v sítích komorbidit obecně (pestré sousedstvo, tendence vázat se k odlišným příčinám, přičemž podle Horras et al. (2015) zejména chronického charakteru, dále menší počet kódů z téže MKN kapitoly). Rovněž spojitost mezi smrtí na J10 a vysoce častými komorbiditami bývá ve studiích zkoumána, i prokazována (CDC, 2021b).

Některé z nemocí vedených příkladem v začátku odstavce (E87, E78, J10) zároveň figurují v tabulkách podle průměrného počtu zapsaných příčin (PPK) (tabulky 4–7) mezi těmi nemocemi, které jich mají i v první části hlášení průměrně nejvíce. Není to však pravidlem. Choroby jako C97 nebo mnohé z kapitol s vnějšími příčinami smrti (X44, X64: druhy otrav konkrétními preparáty, Y83: „*operace a jiné operační výkony jako příčina abnormální reakce nebo pozdější komplikace, bez zmínky o nehodě při výkonu*“) nedosahují v bipartitním grafu natolik vysokých stupňů jako nejcentrálnější uzly v těchto druzích sítí. Byť tedy v první části HOÚ bývají součástmi těch nejdelsích chorobných řetězců, vazeb do druhé části HOÚ mají citelně méně, což bylo ostatně indikováno již v části textu s explorační analýzou, kdy byla v obrázku 2 zachycena změna v PPK po umazání vnějších příčin smrti.

Některé UCD, jež jsou uvedeny v tabulce 1 a 2, nemají vysoké stupně v množině základních příčin standardního bimodálního grafu. Jedná se u obou pohlaví zejména o novotvary, jejichž shluk se s narůstajícím n stává být stále menším, až téměř schází (příloha 7a a 8a). Samozřejmě je nutné uvažovat, že v přílohách 7a a 8a mají větší potenciál s růstem n setrvat choroby s vyšší prevalencí, avšak i téměř nejčastější novotvar, jmenovitě C34, je u žen naposledy spatřen při $n = 10$, u mužů při $n = 15$. Tato choroba tudíž bývá buď zapisována s menším počtem přispívajících příčin smrti, či s příčinami rozdílnými od většiny jiných nejčastějších UCD.

Podobné platí i pro F03, zejména u žen také pro G30. Charakter sousedstva právě těchto chorob se v porovnání s ostatními co do prevalence podobnými příčinami smrti v obou množinách značně vyděluje. Při srovnání s výsledky v sítích s přispívajícími příčinami totiž lze upozorovat, že když bývají tyto příčiny (u mužů spíše pouze F03) zapisovány do druhé části HOÚ, tak zejména k UCD, které jsou společně řadě ostatních nejčastějších přispívajících příčin smrti. Když se však zapíše do první části HOÚ jako UCD, tak se spíše specifickými přispívajícími příčinami, zároveň ale v sítích s růstem n na rozdíl od novotvarů setrvávají, proto množství sousedů by měly mít srovnatelné, jako jiné jim z hlediska prevalence blízké choroby.

V části bipartitního grafu s přispívajícími příčinami smrti jsou dominantní uzly shodně s těmi nejčastěji uváděnými v tabulce 1 a 2, s výjimkou vůdčí CCD co do prevalence, sice F17. Tato příčina smrti má v sítích zhruba o polovinu méně sousedů než ostatní CCD, které se řadí mezi nejčastější. Tudíž bývá zapisována do HOÚ osobám, které zemrou na spíše specifické UCD, čemuž tak není u ostatních nejčastějších CCD, třeba I48 („*fibrilace a flutter síní*“), I10 („*esenciální (primární) hypertenze*“) nebo N18 („*chronické onemocnění ledvin*“). Povaha F17 může být zřejmě projekcí sousedstva C34 (tyto choroby společně souvisí), které je rovněž co do počtu nižší, jak bylo výše uvedeno.

Často zapisovanou chorobou do části 2 HOÚ je E66 a ve vazbě na výše zjištěné ohledně jejích sousedů a metrik v orientované síti (viz analýza AS a HS), lze rozumět, že v bipartitním grafu má vlastnosti opačné, neboť do části 2 bývá naopak zapisována osobám, které zemřely s nejběžnějšími UCD. Toto neplatí pro F11 či F19 (poruchy chování v důsledku užívání drog rozlišených dle druhu), které s E66 byly povýš interpretovány společně kvůli podobnostem jevícím se z části 1 HOÚ. Tyto kódy mají ryze specifické sousedstvo v množině UCD, což je z pochopitelných důvodů zpodobňuje s F17.

Závěrečným poutavým rysem v podmnožinách standardního bimodálního grafu jsou seznamy podle výše stupně „proložené“ méně častými chorobami, které jsou ale zároveň klasifikovány v týchž oddílech MKN jako ty nejběžnější příčiny. Příkladem může jít trojice chorob N17, N18, N19, nebo dvojice W18 a W19, či K74 a K76. Tyto, i další, bývají v seznamu podle výše stupně poměrně vysoko, ač některé choroby zaostávají z hlediska prevalence. Jiným slovem, pokud se jedná o chorobu, která sdílí oddíl s jinou vysoce prevalentní, může existovat tendence ji zapisovat na HOÚ s týmiž komorbiditami a také ve větším rozsahu, než by tomu bylo u jiných příčin smrti nenáležejících k nejčastějším. Toto se může dít samozřejmě nejen v důsledku „podobnosti“ s onou nejčastější chorobou, ale také kvůli typu choroby samotné (vyšší PPK).

Při zaměření pozornosti na síť omezenou $n = 20$, nelze minout, že muži a ženy se ani v jedné z množin výrazněji ve skladbě tohoto grafu neodlišují. Když už, tak lze nalézt u mužů v CCD množině méně F nemocí, nebo méně výraznou dominanci N chorob v množině UCD.

Nakonec jsou k uvedení choroby jsoucí při nižších n ne výrazně centrálními, ale setrvávající v síti až do $n = 20$. Nejnápadněji se rýsují v ženské množině CCD, konkrétně se jedná o některé choroby E a F, nejzřetelněji jsou takovými E03, E43, F41 a F32. Nemoci s těmito vlastnostmi tedy nemají do množiny UCD vazby s „leckerou“ základní příčinou smrti (jako nejběžnější I nebo J nemoci), nýbrž s určitou, však zároveň rovněž nemalou skupinou, poněvadž v síti setrvávají až do $n = 20$.

8.2 Bimodální graf shodného sousedstva

V přílohách 9a–10b naopak nevyčnívají tak okatě (či dokonce i chybí) vztahy, jimiž disponují nejfrekventovanější příčiny smrti. Uzly se v nich vážou, pokud je většina sousedů společná. V první řadě je zřejmé, že v důsledku toho jsou sítě v kontrastu se standardními podmnožinami bimodálního grafu méně husté, jejich průměrné stupně se pohybují mezi téměř třemi (u množiny CCD mužů) a téměř šesti (u množiny UCD žen) (přílohy 9a–10b) a celkové počty uzlů v sítích nedosahují tří set. O poznání méně chorob tedy sdílí vyloženě většinu společných sousedů z opačné množiny.

Rovněž lze vidět, že zejména sítě s CCD jsou do značné míry fragmentované. Roztříštěnost znamená, že existují menší skupiny vzájemně si velmi podobných přispívajících příčin smrti z hlediska vazeb na základní příčiny. U některých skupin je toto předpokladatelné, třeba v komponentách se stejno písmennými kódy, konkrétně X, V, W, C nebo P. I v rámci větších skupin se však rýsuje tendence kódů ze stejných kapitol, např. J, K nebo I společně sdílet vazby, jako tomu bylo i v orientované síti. V případě neorientované sítě by se k těmto chorobám daly zařadit ještě zejména v CCD síti u mužů nemoci F, které společně s převahou chorob E, R a G formují komponentu relativně slabě propojenou s ostatními částmi sítě. Z dat lze zjistit, že se jedná o příčiny smrti, které shodně sousedí s nemocemi G30, F03, G20 a E11. Choroby obsažené v tomto klastru tudíž mohou být oněmi, které stojí za tím, že jedny z nejčastějších příčin smrti (G30 a F03) nemají ve standardních bimodálních grafech stupně srovnatelné s jinými, stejně častými UCD, jak bylo rozvedeno výše. U obou pohlaví tvoří v sítích shodných sousedů UCD G30 a F03 izolované klastry (u mužů jsou jejich součástí i další choroby jmenované výše), jež jsou projekcí oněch vazeb klastru F, E, R a G chorob v sítích CCD.

V tabulce 17 jsou zapsány nejvíce centrální uzly z hlediska výše jejich stupně. Je evidentní, že se ve většině případů zcela nejedná o choroby častěji se vyskytující v populaci. Podchytit tyto nebylo v bimodálních grafech shodného sousedstva vůbec ani cílem. V těchto sítích je vůči orientovanému grafu paradoxní, že u obou grafů žen a pouze u grafu základních příčin smrti mužů sdílejí většinou uzly s nejvyššími stupni vazby vzájemně. Jedná se o prvek, který odlišuje tyto sítě od oné orientované z dat části 1 HOÚ, jež se vyznačovala tím, že se centrální nody vázaly naopak spíše na choroby o stupních ne toliko vysokých, jako jejich vlastní. S čímž zároveň souvisí existence regionů v síti, kde jsou nahromaděny spíše „tučné“ vazby, které, jak již psáno bylo, symbolizují, že choroby, mezi nimiž se tyto vazby nacházejí, sdílejí vysoký počet společných sousedů, nejen jejich podíl. Nápadný je tento region zejména v UCD množině žen, ale i v CCD množině mužů. Např. N18 a I50 společně sdílejí mírně více než 110 UCD ze zhruba 170, s nimiž každá z příčin má signifikantní vazbu. V podobných řádech se pohybuje i sousedstvo ostatních nejčastějších CCD, ale i jiných uzlů, které s nimi sousedí, jmenně ono E87, jež již bylo zmiňováno výše v kontextu standardního bimodálního grafu, kterému ostatně řada chorob uskupená v těchto oblastech dominovala a také v nich setrvala s rostoucím n .

U mužů v množině UCD však není region s „tučnými“ vazbami nápadný, tudíž u nich ty nejčastější příčiny smrti společnou většinu přispívajících nesdílejí, nebo společných sousedů nemají výrazně mnoho. I proto mají zejména u nich nejčastější příčiny smrti (konkrétně I25, I50, I48, N18 a J96) povětšinou výrazně marginálnější postavení, či některé i zcela chybějí, např. I25 nejen proto, že se u mužů váže na více rozdílných základních příčin smrti než u žen, ale řada z jeho sousedů jsou i odlišní od třeba sousedů I50. Většinová shoda v přispívajících chorobách naopak platí u I48 a E11, které tvoří s několika dalšími uzly samostatnou izolovanou komponentu v grafu v příloze 10a. V sítích existují i další spíše okrajové uzly vyznačující se svou vysokou prevalencí (např. K choroby, nebo již zmiňované F03 a G30, nebo u žen novotvary C54, C53).

Tabulka 17: Choroby s nejvyššími stupni v bipartitních grafech shodného sousedstva

	Množina přispívajících příčin smrti	Množina základních příčin smrti
Ženy	<p>R68 – jiné celkové příznaky a znaky D68 – jiné poruchy koagulace R57 – šok nezařazený jinde K92 – jiné nemoci trávicí soustavy D69 – purpura a jiné krvácivé stavy E88 – jiné poruchy metabolismu G93 – jiné poruchy mozku A49 – bakteriální infekce neurčené lokalizace I42 – kardiomyopatie J96 – respirační selhání nezařazené jinde N28 – jiné poruchy ledviny a ureteru nezařazené jinde</p>	<p>A48 – jiné bakteriální nemoci nezařazené jinde A08 – střevní infekce viry a jinými určenými mikroorganismy J38 – nemoci hlasivek a hrtanu nezařazené jinde I87 – jiná onemocnění žil N30 – zánět močového měchýře A86 – neurčená virová encefalitida Q27 – jiné vrozené vady periferní cévní soustavy D41 – novotvar nejistého nebo neznámého chování močových orgánů D58 – jiné dědičné hemolytické anemie G92 – toxická encefalopatie K42 – pupeční kýla</p>
Muži	<p>F09 – neurčená organická nebo symptomatická duševní porucha E41 – nutriční marasmus F22 – poruchy s trvalými bludy J12 – virový zánět plic (pneumonie nezařazený jinde) J39 – jiné nemoci horních dýchacích cest A49 – bakteriální infekce neurčené lokalizace E88 – jiné poruchy metabolismu F29 – neurčené neorganické psychózy J60 – pneumokonióza uhlokopů R57 – šok nezařazený jinde A41 – jiná sepse D69 – purpura a jiné krvácivé stavy</p>	<p>J82 – plicní eozinofilie nezařazená jinde Q27 – jiné vrozené vady periferní cévní soustavy I87 – jiná onemocnění žil D50 – anemie z nedostatku železa A49 – bakteriální infekce neurčené lokalizace J42 – neurčená chronická bronchitida K82 – jiné nemoci žlučníku N30 – zánět močového měchýře (cystitida) D59 – získané hemolytické anemie G03 – meningitida způsobená jinými a neurčenými příčinami J06 – akutní infekce horních dýchacích cest na více místech a neurčených lokalizací K52 – jiná neinfekční gastroenteritida a kolitida</p>

Zdroj: zpracování vlastní

Nejčastěji jsou v množinách UCD zastoupeny nemoci trávicí a oběhové soustavy, dále vnější příčiny, což odpovídá nejčastěji se objevujícím rozmezím kódů i na obrázku 11 u orientované sítě, ovšem s výjimkou novotvarů a poranění, otrav a jiných následků vnějších příčin (T), které mají v sítích shodného sousedstva menší zastoupení. Pro příčiny T však platí, že společně s R kódy a nemocemi oběhové soustavy dominují množinám CCD u obou pohlaví. V této množině se také poprvé objevují choroby očí a uší (H), jichž je srovnatelně jako infekčních a parazitárních nemocí, které ale jsou zároveň častěji uzly centrálními (tabulka 17).

Některé uzly vystupující jako centrální, mají spíše nižší počet sdílených sousedů, oproti mnohým jiným chorobám, které jsou svým postavením marginálnější a jejichž příklady byly uvedeny výše (např. ono E11 a I48). V porovnání s jinými uzly mají málo sousedů v protilehlých množinách například choroby sdružené u mužů okolo J82 („*plicní eozinofilie nezařazená jinde*“) v UCD síti, kolem J chorob a E41 („*nutriční marasmus*“) v CCD množině a u žen kolem R68 („*jiné celkové příznaky a znaky*“) v CCD množině. Tyto choroby disponují relativně nevysokým celkovým počtem sousedů, však stále náležejícím k 50 % chorob s jejich nejvyšším počtem.

Relativně nevysoký počet sousedů však není příznačný pro uzly z kapitoly A, zejména A49 nebo A41. Konkrétně například mužské sousedství CCD A49 a E88 je založeno na tom, že se váží na více než 50 shodných základních příčin smrti z celkového počtu asi 60 u E88 a zhruba 75 u A49. Podobná situace panuje nejen u ostatních A nemocí, ale i obecně v oblasti sítě, kde se nachází dvojice těchto nemocí. Souhrnně, vysoký počet shodných sousedů nesdílejí pouze příčiny zcela nejčastější, ale i právě takové jako A49, R choroby nebo některé D nemoci.

Kapitola 9

Shrnutí, diskuze a závěr

Záměrem poslední kapitoly je nejdříve shrnutí zjištěného, posléze je pozornost zaměřena na rekapitulaci nejzřejmějších nedostatků, které celý text doprovázely. Uzavřena je práce nastíněním využitelnosti poznatků, k nimž lze díky SNA dojít i potenciál aplikace metod analýzy sítí v demografii.

9.1 Shrnutí

Hlavní poznatky, k nimž bylo dospěno díky aplikací metod analýzy sítí při analýze dat z hlášení o úmrtí, by se daly postihnout v bodech, které následují v dalších dvou oddílech textu.

9.1.1 Orientovaná síť

- Použitá pravidla pro tvorbu příčin a důsledků a pro vymezení systému přivedly k tomu, že ačkoli asi polovina osob umírá s nízkým počtem nejčastějších příčin smrti, tak finální síť obsahovala u každého pohlaví asi 800 uzlů a více než 2,5 tisíce hran.
- Choroby postihující „konkrétní“ orgány oběhové, trávicí a dýchací soustavy měly tendence v síti navazovat vztahy se sobě „podobnými“ nemocemi, kdež nemoci endokrinní, výživy a přeměny látek, dále choroby nervové soustavy a poruchy duševní a poruchy chování, stejně jako příznaky, znaky a abnormální klinické a laboratorní nálezy nezařazené jinde se vázaly spíše k sousedům z odlišných kapitol MKN, než jsou ony samy.
- Specifická pravidla pro vyplňování HOÚ osobám zemřelým v důsledku vnějších příčin se otiskla do velikosti sítě (shluk, kterým dominovaly tyto příčiny smrti, byl nejmohutnější) i její architektury (v podgrafu s vnějšími příčinami se u obou pohlaví nacházelo více než 20 % všech vazeb i uzly s nejvyššími hodnotami skóru autorit a hub).
- Vedle vnějších příčin byly nejhojněji zastoupeny novotvary, nemoci oběhové soustavy, s odstupem i choroby trávicí soustavy, naopak nejméně bylo v orientované síti nemocí kůže a podkožního vaziva a infekčních a parazitárních nemocí.
- Nemoci, které se na HOÚ objevovaly častěji, měly v síti i vyšší stupně (jedná se zejména o choroby v tabulkách 1 a 2).

- Graf byl dobře rozložitelný na komunity, přičemž u žen jich G-N algoritmus našel více než u mužů, zřejmě kvůli vyšší provázanosti, nebo obecněji jiná vazebná struktura mezi chorobami krve a krevtovných orgánů, jater a žlučníku, pohybové soustavy a některých méně běžných chorob nervové soustavy nebo novotvarů.
- Naopak velmi podobnými u mužů i žen byly podgrafy sdružující respirační onemocnění, demence a kojeneckou úmrtnost.
- Sousedy smrti byly frekventované bezprostřední příčiny smrti, ale i choroby označitelné spíše jako samy o sobě smrtelné. Naopak přímo se na smrt nevázaly některé akutní stavy, či častá selhání orgánů, která bývají na HOÚ následována zápisem méně běžných, ale „blíže specifikovanějších“ stavů.
- V síti je málo vrátných, nejvíce by se takto daly označit choroby zprostředkující vztahy mezi klastrem kojenecké úmrtnosti na hlavním těle sítě, tj. choroby, kterými trpěly matky dětí zemřelých v kojeneckém věku.
- Globálně centrálními kódy byly nejen nejčastější příčiny smrti, ale i ty pouze s pestrým sousedstvím (př. poruchy chování v důsledku abúzu návykových látek, nebo kód stojící pro bolest).
- Lokálně centrálními uzly byly časté bezprostřední příčiny smrti, kódy mající danou pozici, kam by měly být na HOÚ zapsány (C97) a centrální uzly v oblastech se „specifickou“ oblastí úmrtnosti.
- Marginálními uzly byly vzácnější choroby, ale i běžnější, např. některé R kódy.
- Graf za muže nabývá vyšší lokální centrality než graf za ženy, u nichž je tělo sítě zřejmě více provázané, což se odráží i na odlišném rozkladu sítě na podgrafy.
- Dominantní choroby vzájemné vztahy spíše nenavazují, výjimkou jsou centrální uzly v podgrafech s vnějšími příčinami smrti.
- Univerzálně populárními v rámci svých podgrafů jsou metabolické nemoci, choroby žlučníku, *smrt* a některé nemoci nervové soustavy.
- Alfa uzly jsou vysoce frekventované nemoci, ale některé takové nemoci nejsou alfa uzly, přesněji se jedná o nemoci s častým vychýleným poměrem stupně dovnitř ku ven, nebo kódy popisující konkrétní „události“ (př. umrznutí).
- Asi čtvrtina vazeb v ženské síti je ryze ženská a asi třetina vazeb mužské sítě je ryze mužská, tudíž většina vazeb je u mužů i žen shodná.
- Mnoho vazeb, které nemají ženy, ale jen muži, existuje díky malé skupince uzlů s vysokým stupněm a jedná se většinou o zranění.
- Jsou k nalezení sice stejně centrální uzly u mužů i žen, za to kompletně se odlišující svým sousedstvím (např. deprese).
- Bylo poukázáno na vazby, které by mohly být odrazem systematických chyb ve vyplňování hlášení o úmrtí (např. linka mezi nedostatkem předpokládaného normálního fyziologického vývoje a demencí).

9.1.2 Neorientovaná síť

- I když se nejčastější uzly z množiny základních příčin smrti váží na nejčastější uzly z množiny přispívajících chorob, některé choroby disponují spíše „specifičtějšími“ sousedy z protilehlých množin, nebo jich mají méně (př. demence, novotvary a poruchy chování kvůli užívání tabáku)
- Jádra bimodálních sítí jsou tvořena přibližně šedesáti nemocemi, mezi nimiž převažují choroby dýchací, oběhové, trávicí soustavy, dále choroby metabolické nebo pohlavní a močové soustavy (konkrétně viz příloha 7 a 8), v sítích UCD jsou takovými ještě vnější příčiny smrti.
- Standardní bimodální sítě žen jsou hustější.
- Choroby očí a uší se ve zde modelovaných systémech objevují pouze v bimodálních sítích.
- Dvojic chorob, které disponují aspoň polovinou shodných vazeb do protilehlých množin je méně a sítím zaměřeným na podchycení vztahů za takto formulovaných podmínek dominují spíše choroby vzácnější, leč se i zde objevují ony nejčastější UCD, resp. CCD.
- Vynořily se choroby svými vazbami do protilehlých množin většinou podobné, a zároveň mající těchto shodných vazeb bohatě (především nemoci duševní, nervové soustavy a bakteriální a virové).

9.2 Diskuze a závěr

Na metodickou část práce navazovala kapitola věnovaná pokusům o validaci aplikovaných postupů. Ačkoli byly poskytnuty argumenty, které by možná mohly být označeny jako podporující jejich adekvátnost, je práce zatížena řadou nedostatků. Pouze z hlediska obsahového jen ty nejproblematictější aspekty by se daly shrnout do následujících bodů (a při té příležitosti rovněž i dovysvětlit):

- Relevance modelování sítě příčin smrti až na úrovni *třímístných* MKN kódů je diskutabilní. Ač se jedná o standardní postup při tvorbě sítí komorbidit, lze asi dospět k závěru, že agregace na některou z vyšších úrovní (ne třeba až tak povrchní jako v případě Egidi et al. (2018)) by do práce vnesla nejen větší prostor systematictější analýzy, ale i zvýšila utříděnost a přehlednost. Eventuálně se ve světle napsaného může zdát rozumnějším konstruovat sítě za vybrané choroby (např. řada studií zkoumá komorbiditní sítě diabetu, demencí nebo covidu-19). Shrnutí, práce je označitelná za přílišně široce zabranou, a to v důsledku slepého následování dosud provedených studií zapříčiněného nezkušeností s daným tématem. Na druhou stranu přínos a smysl analýzy sítí je právě nad daty velikými.
- Analyzována byla data z hlášení o úmrtí, jež jsou zatížena řadou zkreslení, která ústí v to, že počet zapsaných příčin smrti může být predikován řadou dalších proměnných z datasetu, od kapitoly, do níž náležela základní příčina smrti, po to, zda byla zemřelá osoba pitvána, či nikoli. Vztahy, které byly a nebyly zjištěny, jsou ovlivněny tím, kdo, jak a kde zemřel, i kdo, jak a kde HOÚ vyplňoval. Navíc byla pravidla pro utváření dvojic příčin a důsledků i UCD a CCD postavena na nerealistickém předpokladu správného vyplňování HOÚ zcela v souladu s pravidly ACME.

- Je nutné brát v úvahu značné zjednodušení celé komplexity systémů komorbidit pramenící z utváření pouhých dvojic chorob, nikoli vícetic.
- Protože s růstem počtu vyplněných kódů na HOÚ roste rovněž počet vytvořitelných dvojic, je v síti příčin smrti výrazně nadhodnocena úmrtnost lidí s dvěma a více kódy. A takovou nemusí být reprezentativní populace USA, už jen kvůli tomu, že počet zapsaných kódů je odvislý od řady dalších, nejen sociodemografických proměnných (např. typ UCD, místo smrti, zda byla zemřelá osoba pitvána).
- Data nebyla skutečně analyzována po kapitolách MKN, nýbrž podle prvních písmen kódů, ačkoli interpretována byla mnohdy právě jako ony kapitoly MKN.
- V kapitole s analýzou podgrafů není pozornost věnována podgrafům s kojeneckou úmrtností, ačkoli jako všechny ostatní prezentované obsahují více než 2 % uzlů celé sítě.
- Chybí vysvětlení důvodu, proč, když si člověk dal práci se standardizací koeficientu asociace mezi UCD a CCD k témuž nedospěl při tvorbě orientované sítě. Však hájit se lze jednak tím, že v předlohových studiích o sítích komorbidit se nestejná věková rozložení nemocných podle chorob nijak nekomentují, jednak tím, že lze zvědět, že CDAI se od OR významněji odlišuje když už, tak pouze ve věcích, kdy zemřelo velmi málo osob (zjištěno z vlastních výpočtů).
- Testování signifikance asociací bylo plně přejato z předlohových studií, ač v nich ale autoři pracují zpravidla s daty v některých hlediscích jinými (nejedná se o vyčerpávající záznamy, ani nebývají transversální).
- Představení změn, které by se staly, kdyby byly použity jiné metody pro vymezení systému, končí u pokusů o ukázání, že by vznikla síť diametrálně odlišná, kdyby byla pravidla pro vymezení systému nakombinována jinak. Pro nedostatek prostoru není věnována dostatečná pozornost ukázání těch změn, k nimž by došlo, kdyby byla bývala použita úplně jiná metoda. Lze se však pro jednou opět opřít o Fotouhi et al. (2018), zabývajícími se dopady využití metod postavených na zcela odlišných principech na podobu výsledné sítě. Konstatují, že jimi zkoumané „*metody* (v tom i aplikace koeficientů asociace) *podchycují rozdílné aspekty komorbiditních vzorců, člověk si musí všimnout jejich vlastností a volit mezi nimi s ohledem na oblast svého zájmu*“ (Fotouhi et al., 2018, s. 18). Vlastnosti použitých koeficientů asociace byly podrobeny kritice výše a volba metody je adekvátní pro zde vytyčený cíl, neboť jak Fotouhi et al. (2018) píší, jsou metody běžných statistických koeficientů vhodné a dostačující, když je cílem pouze nalezení signifikantních vazeb mezi chorobami, nikoli nad to např. i uvažování přesné časové souslednosti, v níž se u člověka objevily.
- Na vazby analyzované v podgrafech bylo nahlíženo jako na seriózní, pokud pro ně existovala opora v epidemiologické literatuře. Lze se setkat s tvrzeními typu „*epidemiologie je někdy považována za nezdravou vědu kvůli často dokumentovaným nekonzistencím v poznatcích a neúspěších při replikacích*“³⁹ (Dorak et al., 2012, s. 4). Ve světle čehož by mohlo být vhodné upřímně uvést, že při dohledávání opory vazbám v literatuře nebyl vůbec brán zřetel na závěry studií jsoucích v nesouladu s tím, co bylo snahou literaturou podepřít.

³⁹ „*Epidemiology is sometimes considered to be an unhealthy science because of frequently documented inconsistencies in findings and replication failures.*“

Od posledního zmíněného bodu se odvozuje otázka ohledně relevance zkoumané problematiky. Jednak podle Kim et al. (2014) mohou grafy sloužit také jako zdroj nových pohledů na komorbiditu, neboť nezřídka se v nich rýsují i nekonfirmované vazby. Avšak i pokud je na ně nahlíženo spíše jako na potenciální produkt nevhodně zvolených, či nakombinovaných kroků ve vymezení systému, či chyb v datech, stále nemusí sítě příčin smrti pozbývat na své relevanci. Na s. 48 (kap. 5) je psáno, že jsou „v tabulce 9 uvedeny dosud provedené studie vztahů mezi chorobami, pro něž je společný jejich aspoň počáteční krok...“. Což nastiňuje, že vymodelování a případně analýza systému nebývá stěžejní náplní studií, které figurovaly jako inspirace pro tuto práci. Stěžejním cílem je ve většině z nich zkoumání dynamiky systémů komorbidit, čímž je rozuměn jejich vývoj v čase. V části s explorační analýzou bylo nastíněno, že šikmé rozložení stupňů uzlů je charakteristika pro grafy velmi klíčová. Před lety byl totiž objeven zákon, kterým se řídí toto rozložení stupňů uzlů ve většině reálných sítí (Barabási a Pósfai, 2016), což otevřelo dveře k využívání sítí pro predikci vzniku vazeb mezi entitami. Vzhledem k tomu, že většina studií o sítích komorbidit je zaměřena na predikci linků (ne jen na pouhé vymodelování systému, eventuálně jeho analýzu), lze až v tomto spatřovat hlavní potenciál celého snažení, které analýzu medicínských sítí obklopuje. Ovšem studie zabývající se predikcí vazeb v systémech příčin smrti neexistují, ač by zřejmě mohly, vzhledem k tomu, že mají se sítěmi komorbidit řadu shodných vlastností, na což bylo opakovaně naráženo zde v práci. Tudíž zkoumání dynamiky sítí příčin smrti by mohlo být směrem, jímž by se možná mohlo zdát být smysluplným se pokoušet směřovat případný další výzkum této problematiky. Podle Chmiel et al. (2014) totiž lze ze sítí komorbidit pro vymezené věkové intervaly usuzovat na vztahy, které budou existovat v sítích za tutéž populaci, až zestárne.

Avšak toto není jediným možným směrem. Potenciál lze spatřovat i při setrvání u statických sítí, ty např. mohou sloužit jako nástroj pro nalézání regionálních rozdílů ve zvyklostech vyplňování HOÚ, odhalování oblastí rozkolů mezi ACME a osobami vyplňujícími HOÚ, nebo již výše jmenované zkoumání rozdílů v sítích za různě vymezené subpopulace (podle rasy, vzdělání, místa smrti, UCD...). Zajímavé, dokonce možná i užitečné, by mohlo být zkoumat, jak sítě vypadly dříve (existují časové řady podkladových dat od 1960), nebo jak se na nich projevovaly revize klasifikací MKN.

Rovněž nelze přejít i možnosti skrývající se za bimodálními sítěmi. Podle literatury lze bipartitní grafy používat například při analýzách vztahů mezi množinou symptomů a množinou nemocí. Tento přístup umožňuje vidět nemoc jako komunitu přímo spojených symptomů (Cramer et al., 2010) a uplatnění nalézá prvořadě při analýze nemocí chronických a duševních, kdy zejména prvně zmíněné jsou v současnosti jedněmi z vůdčích příčin úmrtí v zemích vyspělého světa.

I když by tedy mohl vznikat dojem, že celá práce je jen způsobem, jak komplikovaně udělat základní deskriptivní analýzu hlášení o úmrtí, nemusí tomu tak se světle všech možných cest, které se díky sítím nabízejí, vůbec být.

Seznam zdrojů

- Alzheimer's Association. (2022). Alzheimer's Disease Facts and Figures. Special report. *Alzheimers Dement* 2022;18. Dostupné online: <https://www.alz.org/media/documents/alzheimers-facts-and-figures.pdf>.
- American Cancer Society. Cancer Facts & Figures 2021. (2021). Atlanta: *American Cancer Society*; 2021. Dostupné online: <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2021/cancer-facts-and-figures-2021.pdf>.
- Ashtiani, M., Salehzadeh-Yazdi, A., Razaghi-Moghadam, Z., Hennig, H., Wolkenhauer, O., Mirzaie, M., & Jafari, M. (2018). A systematic survey of centrality measures for protein-protein interaction networks. *BMC systems biology*, 12(1), 1-17. Dostupné online: <https://bmcsystbiol.biomedcentral.com/articles/10.1186/s12918-018-0598-2>.
- Baggio, G., Corsini, A., Floreani, A., Giannini, S., & Zagonel, V. (2013). Gender medicine: a task for the third millennium. *Clinical chemistry and laboratory medicine*, 51(4), 713-727. Dostupné online: <https://www.degruyter.com/document/doi/10.1515/cclm-2012-0849/html>.
- Barabási, A.-L., Pósfai, M. (2016). *Network science*. Cambridge: Cambridge University Press. ISBN: 9781107076266 1107076269. Dostupné online: <http://networksciencebook.com/>.
- Bazhenova, L., Newton, P., Mason, J., Bethel, K., Nieva, J., & Kuhn, P. (2014). Adrenal metastases in lung cancer: clinical implications of a mathematical model. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*, 9(4), 442–446. Dostupné online: <https://doi.org/10.1097/JTO.000000000000133>.
- Becker, J. B., & Hu, M. (2008). Sex differences in drug abuse. *Frontiers in neuroendocrinology*, 29(1), 36–47. Dostupné online: <https://doi.org/10.1016/j.yfrne.2007.07.003>
- Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10), P10008. Dostupné online: https://iopscience.iop.org/article/10.1088/1742-5468/2008/10/P10008/meta?casa_token=xfYPUuuUTZAAAAAA:XVTeXjCYj_K5ARcSMFW3AJIM59lJQoAHSQMmlsRN660dX7m8G9M4Kpv3wm-DNshOJmPW4BtXcu8.

- Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, 323(5916), 892-895. Dostupné online: https://www.science.org/doi/full/10.1126/science.1165821?casa_token=sXGwUcV7Z80AAAAA%3AVe9ce1tdm8N1ECcvOo11O0rMftu3BXmGms21bTfwIPbgXCbphXqjhfLxwWDrEgrATuValA1L6YgINU.
- Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of mathematical sociology*, 25(2), 163-177. Dostupné online: https://www.tandfonline.com/doi/abs/10.1080/0022250x.2001.9990249?casa_token=cFmOTEIAopUAAAAA:ZYhvhBCTdgEp0KI5Vu6Ycgyi7EWW_8HxcjPUMqIrOY1KWswg-7uDMaCvF4clXXULJ2H0gF6Y6scKjg.
- Brunnberg, U., Hentrich, M., Hoffmann, C., Wolf, T., & Huebel, K. (2017). HIV-associated malignant lymphoma. *Oncology Research and Treatment*, 40(3), 82-87. Dostupné online: <https://www.karger.com/Article/Abstract/456036>.
- Brunnström, H. R., & Englund, E. M. (2009). Cause of death in patients with dementia disorders. *European journal of neurology*, 16(4), 488-492. Dostupné online: <https://doi.org/10.1111/j.1468-1331.2008.02503.x>.
- Budhiraja R, Siddiqi TA, Quan SF. (2015). Sleep disorders in chronic obstructive pulmonary disease: etiology, impact, and management. *Journal of Clinical Sleep Medicine*, 2015;11(3):259–270. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/25700872/>.
- Buzzetti, E., Parikh, P. M., Gerussi, A., & Tsochatzis, E. (2017). Gender differences in liver disease and the drug-dose gender gap. *Pharmacological Research*, 120, 97-108. Dostupné online: <https://doi.org/10.1016/j.phrs.2017.03.014>.
- CDC WONDER. (2021). Causes of death. Dostupné online: <https://wonder.cdc.gov/controller/datarequest/D158;jsessionid=D4B646C686B99A97378577C2BD9A>.
- CDC. (2021). Summary Health statistics Tables. Dostupné online: <https://www.cdc.gov/nchs/fastats/injury.htm>.
- CDC. (2021b). People at Higher Risk of Flu Complications. Dostupné online: <https://www.cdc.gov/flu/highrisk/index.htm>.
- CDC. (2022). Drug Overdose Deaths. Dostupné online: <https://www.cdc.gov/drugoverdose/deaths/index.html>.
- CDC. (1997). Multiple Cause Mortality. Documentation of the mortality tape file for 1997 data. Dostupné online: https://wonder.cdc.gov/wonder/sci_data/mort/mcmort/type_txt/mcmort97.asp.
- Clayton, J. A., & Collins, F. S. (2014). Policy: NIH to balance sex in cell and animal studies. *Nature*, 509(7500), 282-283. Dostupné online: <https://www.nature.com/articles/509282a>.
- Cooper, A. J., Gupta, S. R., Moustafa, A. F., & Chao, A. M. (2021). Sex/Gender Differences in Obesity Prevalence, Comorbidities, and Treatment. *Current obesity reports*, 1-9. Dostupné online: <https://link.springer.com/article/10.1007/s13679-021-00453-x>.
- Cramer, A. O., Waldorp, L. J., Van Der Maas, H. L., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and brain sciences*, 33(2-3), 137-150. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/20584369/>.

- Department of Health and Human Services. (2003). Physicians Handbook on Medical Certification of Death. National Center for Health Statistics. Maryland. DHHS Publication No. 2003-1108. Dostupné online: https://www.cdc.gov/nchs/data/misc/hb_cod.pdf.
- Désesquelles, A. F., Salvatore, M. A., Pappagallo, M., Frova, L., Pace, M., Meslé, F., & Egidi, V. (2012). Analysing multiple causes of death: which methods for which data? An application to the cancer-related mortality in France and Italy. *European Journal of Population/Revue Européenne de Démographie*, 28(4), 467-498. Dostupné online: <https://link.springer.com/article/10.1007/s10680-012-9272-3>.
- Désesquelles, A., Salvatore, M. A., Frova, L., Pace, M., Pappagallo, M., Meslé, F., & Egidi, V. (2010). Revisiting the mortality of France and Italy with the multiple-cause-of-death approach. *Demographic Research*, 23, 771-806. Dostupné online: <https://www.demographic-research.org/volumes/vol23/28/default.htm>.
- Dhondup, T., & Qian, Q. (2017). Electrolyte and acid-base disorders in chronic kidney disease and end-stage kidney failure. *Blood purification*, 43(1-3), 179-188. Dostupné online: <https://doi.org/10.1159/000452725>.
- Ding, Y. (2011). Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks. *Journal of informetrics*, 5(1), 187-203. Dostupné online: <https://doi.org/10.1016/j.joi.2010.10.008>.
- Diviák, T., Lehotský, L., Ocelík, P. (2021). Seminář k analýze sociálních sítí. Praha: Akademie věd ČR. Květen, 2021.
- Dorak, M. T., & Karpuzoglu, E. (2012). Gender differences in cancer susceptibility: an inadequately addressed issue. *Frontiers in genetics*, 3, 268. <https://doi.org/10.3389/fgene.2012.00268>.
- Dufek, M. (2003). Cerebrovaskulární onemocnění ve stáří. *Neurologie pro praxi*, 1, 2003. Dostupné online: <https://www.solen.cz/pdfs/neu/2003/01/05.pdf>.
- Durazzo, M., Belci, P., Collo, A., Prandi, V., Pistone, E., Martorana, M., ... & Bo, S. (2014). Gender specific medicine in liver diseases: a point of view. *World Journal of Gastroenterology: WJG*, 20(9), 2127. Dostupné online: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3942817/>.
- Duthie A., Chew D., Soiza R.L. (2011). Non-psychiatric comorbidity associated with Alzheimer's disease, *QJM: An International Journal of Medicine*, Volume 104, Issue 11, November 2011, Pages 913–920. Dostupné online: <https://doi.org/10.1093/qjmed/hcr118>.
- Egidi, V., Salvatore, M. A., Rivellini, G., & D'Angelo, S. (2018). A network approach to studying cause-of-death interrelations. *Demographic Research*, 38, 373-400. Dostupné online: <https://www.demographic-research.org/volumes/vol38/16/default.htm>.
- Fedeli, U., Zoppini, G., Goldoni, C. A., Avossa, F., Mastrangelo, G., & Saugo, M. (2015). Multiple causes of death analysis of chronic diseases: the example of diabetes. *Population health metrics*, 13(1), 1-10. Dostupné online: <https://link.springer.com/article/10.1186/s12963-015-0056-y>.

- Fernandez-Montalvo, J., Lopez-Goñi, J. J., Azanza, P., & Cacho, R. (2014). Gender differences in drug-addicted patients in a clinical treatment center of Spain. *The American Journal on Addictions*, 23(4), 399-406. Dostupné online: <https://doi.org/10.1111/j.1521-0391.2013.12117.x>
- Figueiredo, J. C., Haiman, C., Porcel, J., Buxbaum, J., Stram, D., Tambe, N., ... & Setiawan, V. W. (2017). Sex and ethnic/racial-specific risk factors for gallbladder disease. *BMC gastroenterology*, 17(1), 1-12. Dostupné online: <https://bmcgastroenterol.biomedcentral.com/articles/10.1186/s12876-017-0678-6>.
- Flagg, L. A., & Anderson, R. N. (2021). Unsuitable underlying causes of death for assessing the quality of cause-of-death reporting. *National Vital Statistics Reporting*, 2021. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/33541519/>.
- Folino, F., Pizzuti, C., & Ventura, M. (2010, September). A comorbidity network approach to predict disease risk. In *International Conference on Information Technology in Bio-and Medical Informatics* (pp. 102-109). Springer, Berlin, Heidelberg. Dostupné online: https://link.springer.com/chapter/10.1007/978-3-642-15020-3_10.
- Foreman, K. J., Naghavi, M., & Ezzati, M. (2016). Improving the usefulness of US mortality data: new methods for reclassification of underlying cause of death. *Population health metrics*, 14(1), 1-9. Dostupné online: <https://pophealthmetrics.biomedcentral.com/articles/10.1186/s12963-016-0082-4>.
- Fortunato, S. (2010). Community detection in graphs. *Physics reports*, 486(3-5), 75-174. Dostupné online: <https://doi.org/10.1016/j.physrep.2009.11.002>.
- Fotouhi, B., Momeni, N., Riolo, M. A., & Buckeridge, D. L. (2018). Statistical methods for constructing disease comorbidity networks from longitudinal inpatient data. *Applied Network Science*, 3(1), 1-34. Dostupné online: <https://link.springer.com/article/10.1007/s41109-018-0101-4>.
- Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks*, 1: 215-239. Dostupné online: <https://www.bebr.ufl.edu/sites/default/files/Centrality%20in%20Social%20Networks.pdf>.
- Gamage, U. S. H., Adair, T., Mikkelsen, L., Mahesh, P. K. B., Hart, J., Chowdhury, H., ... & Lopez, A. D. (2021). The impact of errors in medical certification on the accuracy of the underlying cause of death. *PloS one*, 16(11), e0259667. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/34748575/>.
- Gephi 0.9.2. Bastian M., Heymann S., Jacomy M. (2009). *Gephi: an open source software for exploring and manipulating networks*. International AAAI Conference on Weblogs and Social Media. Dostupné online: <https://gephi.org/>.
- Ghuman, A. K., Newth, C. J., & Khemani, R. G. (2013). Impact of gender on sepsis mortality and severity of illness for prepubertal and postpubertal children. *The Journal of pediatrics*, 163(3), 835-840. Dostupné online: <https://doi.org/10.1016/j.jpeds.2013.04.018>.
- Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12), 7821-7826. Dostupné online: <https://doi.org/10.1073/pnas.122653799>.

- Goh, K. I., & Choi, I. G. (2012). Exploring the human diseasome: the human disease network. *Briefings in functional genomics*, 11(6), 533-542. Dostupné online: <https://academic.oup.com/bfg/article/11/6/533/236949?login=false>.
- Gorina Y, Lentzner H. (2008). Multiple causes of death in old age. *Aging Trends*. 2008 Feb;(9):1-9. PMID: 19174856. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/19174856/>.
- Guralnick, L. (1966). Some problems in the use of multiple causes of death. *Journal of Chronic Diseases*, 19(9), 979-990. Dostupné online: [https://doi.org/10.1016/0021-9681\(66\)90031-2](https://doi.org/10.1016/0021-9681(66)90031-2).
- Haast, R. A., Gustafson, D. R., & Kiliaan, A. J. (2012). Sex differences in stroke. *Journal of Cerebral Blood Flow & Metabolism*, 32(12), 2100-2107. Dostupné online: <https://doi.org/10.1038/jcbfm.2012.141>.
- Hamann, M., Lindner, G., Meyerhenke, H., Staudt, C. L., & Wagner, D. (2016). Structure-preserving sparsification methods for social networks. *Social Network Analysis and Mining*, 6(1), 1-22. Dostupné online: <https://link.springer.com/article/10.1007/s13278-016-0332-2>.
- Hidalgo, C. A., Blumm, N., Barabási, A. L., & Christakis, N. A. (2009). A dynamic network approach for the study of human phenotypes. *PLoS computational biology*, 5(4), e1000353. Dostupné online: <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1000353>.
- Horras S., Bepko J., Longstreet N. (2015) Fluid, Electrolyte, and Acid–Base Disorders. In: Paulman P., Taylor R. (eds) *Family Medicine*. Springer, Cham. https://doi.org/10.1007/978-1-4939-0779-3_103-1.
- Human Cause-of-Death Database. (2017). French Institute for Demographic Studies (France) and Max Planck Institute for Demographic Research (Germany). Methods Protocol. Dostupné online: www.causeofdeath.org.
- Chaker Z, Badhwar V, Alqahtani F, Aljohani S, Zack CJ, Holmes DR, Rihal CS, Alkhouli M. (2017). Sex Differences in the Utilization and Outcomes of Surgical Aortic Valve Replacement for Severe Aortic Stenosis. *Journal of American Heart Association*. 2017 Sep 21;6(9):e006370. doi: 10.1161/JAHA.117.006370. PMID: 28935681; PMCID: PMC5634288. Dostupné online: <https://www.ahajournals.org/doi/10.1161/JAHA.120.020095>.
- Chamblee, C. (1986). Title : TRANSAX, the NCHS system for producing multiple cause-of-death statistics, 1968-78 Personal Author(s) :, Ronald F.;Evans, Marshall C.; Corporate Authors(s) : *National Center for Health Statistics* (U.S.) Published Date : June 1986 Series : DHHS publication ; no. (PHS) 86-1322;Vital and health statistics. no. 20 Series 1 Programs and collection procedures; URL : <https://stacks.cdc.gov/view/cdc/13008>.
- Chmiel, A., Klimek, P., & Thurner, S. (2014). Spreading of diseases through comorbidity networks across life and gender. *New Journal of Physics*, 16(11), 115013. Dostupné online: <https://iopscience.iop.org/article/10.1088/1367-2630/16/11/115013/meta>.
- ICD-10 classification of mental and behavioural disorders. (2019). World Health Organization. Dostupné online: https://icd.who.int/browse10/2019/en#.

- ICD-10-Mortality Manual. (2014). Section I – Instructions for classifying the underlying cause of death, 2014. Dostupné online: https://www.cdc.gov/nchs/data/dvs/2a_2014.pdf.
- Israel, R. A., Rosenberg, H. M., & Curtin, L. R. (1986). Analytical potential for multiple cause-of-death data. *American journal of epidemiology*, 124(2), 161-179. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/3524199/>.
- Janssen, T. A. (1940). Importance of tabulating multiple causes of death. *American Journal of Public Health and the Nations Health*, 30(8), 871-879. Dostupné online: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1531179/>.
- Jensen, A. B., Moseley, P. L., Oprea, T. I., Ellesøe, S. G., Eriksson, R., Schmock, H., ... & Brunak, S. (2014). Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. *Nature communications*, 5(1), 1-10. Dostupné online: <https://www.nature.com/articles/ncomms5022>.
- Jeong, E., Ko, K., Oh, S., & Han, H. W. (2017). Network-based analysis of diagnosis progression patterns using claims data. *Scientific reports*, 7(1), 1-12. Dostupné online: <https://www.nature.com/articles/s41598-017-15647-4>.
- Jones, P. J., Ma, R., & McNally, R. J. (2021). Bridge centrality: A network approach to understanding comorbidity. *Multivariate behavioral research*, 56(2), 353-367. Dostupné online: <https://doi.org/10.1080/00273171.2019.1614898>.
- Jorm, A. F. (2001). History of depression as a risk factor for dementia: an updated review. *Australian & New Zealand Journal of Psychiatry*, 35(6), 776-781. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/11990888/>.
- Kalgotra, P., Sharda, R., & Croff, J. M. (2017). Examining health disparities by gender: A multimorbidity network analysis of electronic medical record. *International journal of medical informatics*, 108, 22-28. Dostupné online: <https://doi.org/10.1016/j.ijmedinf.2017.09.014>.
- Katz, D. J. S. M., Baptista, J., Azen, S. P., & Pike, M. C. (1978). Obtaining confidence intervals for the risk ratio in cohort studies. *Biometrics*, 469-474. Dostupné online: <https://www.jstor.org/stable/2530610>.
- Khokhar, D. (2015). *Gephi cookbook*. Packt Publishing Ltd. ISBN: 1783987405. Dostupné online: <https://oiiipdf.com/gephi-cookbook>.
- Kim, J. H., Son, K. Y., Shin, D. W., Kim, S. H., Yun, J. W., Shin, J. H., ... & Yun, J. M. (2016). Network analysis of human diseases using Korean nationwide claims data. *Journal of Biomedical Informatics*, 61, 276-282. Dostupné online: <https://www.sciencedirect.com/science/article/pii/S1532046416300326>.
- Kleinberg, J. M. (1998, January). Authoritative sources in a hyperlinked environment. In *SODA* (Vol. 98, pp. 668-677). Dostupné online: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.585.7341&rep=rep1&type=pdf>.
- Kokoska, M. S., Piccirillo, J. F., & Haughey, B. H. (1995). Gender differences in cancer of the larynx. *Annals of Otology, Rhinology & Laryngology*, 104(6), 419-424. Dostupné online: <https://doi.org/10.1177/000348949510400601>.

- Kremers, H. M., Nwojo, M. E., Ransom, J. E., Wood-Wentz, C. M., Melton III, L. J., & Huddleston III, P. M. (2015). Trends in the epidemiology of osteomyelitis: a population-based study, 1969 to 2009. *The Journal of bone and joint surgery. American volume*, 97(10), 837. Dostupné online: [10.2106/JBJS.N.01350](https://doi.org/10.2106/JBJS.N.01350).
- Mazanec, M. R. (2009) et al. Diabetická neuropatie z pohledu neurologa. *Neurologie pro praxi*. 2009 10(6), 378-383 Dostupné online: <https://www.neurologiepropraxi.cz/>.
- McHugh, R. K. (2020). The importance of studying sex and gender differences in opioid misuse. *JAMA Network Open*, 3(12), e2030676-e2030676. Dostupné online: doi:[10.1001/jamanetworkopen.2020.30676](https://doi.org/10.1001/jamanetworkopen.2020.30676).
- Mezinárodní klasifikace nemocí: Mezinárodní statistická klasifikace nemocí a přidružených zdravotních problémů ve znění desáté decenální revize MKN-10. (2021). Praha: Ústav zdravotnických informací a statistiky České republiky. Dostupné online: <https://mkn10.uzis.cz/o-mkn>.
- MKN-10 Instrukční příručka. (2020). Dostupné online: <https://www.uzis.cz/res/f/008277/mkn-10-instrukcni-prirucka-20200101.pdf>
- Mohsien, R. A., Al Mohammedawi, M. C., & Yahya, A. G. (2014). Virulence factors enhancing microbial infection in chronic osteomyelitis and antibiotic susceptibility pattern. *American Journal of Medical Case Reports*, 2(6), 126-132. Dostupné online: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.949.354&rep=rep1&type=pdf>
- Moretti, R., Torre, P., Antonello, R. M., Manganaro, D., Vilotti, C., & Pizzolato, G. (2008). Risk factors for vascular dementia: hypotension as a key point. *Vascular health and risk management*, 4(2), 395. Dostupné online: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2496988/>.
- Myles, P. S., Hunt, J. O., & Moloney, J. T. (2004). Postoperative 'minor' complications: comparison between men and women. *Anaesthesia*, 52(4), 300-306. Dostupné online: <https://doi.org/10.1111/j.1365-2044.1997.89-az0091.x>.
- Nasir, N., Jamil, B., Siddiqui, S., Talat, N., Khan, F. A., & Hussain, R. (2015). Mortality in sepsis and its relationship with gender. *Pakistan journal of medical sciences*, 31(5), 1201. Dostupné online: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4641283/>.
- National Bureau of Economic research. (2020). Mortality Data – Vital Statistics NCHS Multiple Cause of Death Data. Dostupné online: <https://data.nber.org/mortality/2018/mort2018.sas7bdat.zip>.
- National Cancer Institute. (2021). Cancer Fast Stats. Dostupné online: <https://seer.cancer.gov/statfacts/>.
- National Center for Health Statistics. (2009a). Vital statistics, instructions for classifying the underlying causes of death: *NCHS instruction manual; part 2a*. Dostupné online: <https://www.cdc.gov/nchs/nvss/instruction-manuals.htm>. Accessibility verified 20.07.2021.
- National Center for Health Statistics. (2009b). Vital statistics, instructions for classifying the underlying causes of death: *NCHS instruction manual; part 2b*. Dostupné online: <https://www.cdc.gov/nchs/nvss/instruction-manuals.htm>. Accessibility verified 20.07.2021.

- National Center for Health Statistics. (2009c). Vital statistics, instructions for classifying multiple causes of death. *NCHS instruction manual; part 2c*. Hyattsville, Maryland: Public Health Service. Published annually. <https://www.cdc.gov/nchs/nvss/instruction-manuals.htm>. Accessibility verified 20.07.2021.
- Nerghes, A., Lee, J. S., Groenewegen, P., & Hellsten, I. (2015). Mapping discursive dynamics of the financial crisis: a structural perspective of concept roles in semantic networks. *Computational Social Networks*, 2(1), 1-29. Dostupné online: <https://computationsocialnetworks.springeropen.com/articles/10.1186/s40649-015-0021-8>.
- NCHS. (2006). Deaths From Each Cause by Month, Race, and Sex: United States, 1999-2006. Dostupné online: https://www.cdc.gov/nchs/data/dvs/MortFinal2006_WorktableIV_part2.pdf.
- OECD. (2019). Addressing Problematic Opioid Use in OECD Countries, *OECD Health Policy Studies*, OECD Publishing, Paris. Dostupné online: <https://doi.org/10.1787/a18286f0-en>.
- Park, J., & Peters, P. A. (2014). Mortality from diabetes mellitus, 2004 to 2008: A multiple-cause-of-death analysis. *Health reports*, 25(3), 12-17. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/24648135/>.
- Pavlopoulos GA, Kontou PI, Pavlopoulou A, Bouyioukos C, Markou E, Bagos PG. (2020). Bipartite graphs in systems biology and medicine: a survey of methods and applications. *Gigascience*. 2018 Apr 1;7(4):1-31. doi: 10.1093/gigascience/giy014. Erratum in: *Gigascience*. 2020 Jan 1;9(1): PMID: 29648623; PMCID: PMC6333914. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/29648623/>.
- Pechholdová, M. (2014). Multiple cause-of-death data in the Czech Republic: an exploratory analysis. *Demografie*, 56, 335-346. Dostupné online: <https://www.czso.cz/csu/czso/demografie-review-for-population-research-no-42014-2nzp8b9ndl>.
- Pechholdová, M. (2017). Sepsis-related mortality in the Czech Republic: multiple causes of death analysis. *Epidemiology Mikrobiology Imunology*, 66(2), 73-79. Dostupné online: https://www.researchgate.net/profile/Marketa-Pechholdova/publication/318710281_Sepsis-related_mortality_in_the_Czech_Republic_Multiple_causes_of_death_analysis/links/5dcba284585156b3510331d/Sepsis-related-mortality-in-the-Czech-Republic-Multiple-causes-of-death-analysis.pdf.
- Pechholdová, Markéta. (2021). Odborná asistentka katedry demografie Vysoké školy ekonomické v Praze [písemné sdělení]. Praha, 17.6.2021.
- Piscopo, P., Bellenghi, M., Manzini, V., Crestini, A., Pontecorvi, G., Corbo, M., ... & Confaloni, A. (2021). A sex perspective in neurodegenerative diseases: microRNAs as possible peripheral biomarkers. *International Journal of Molecular Sciences*, 22(9), 4423. Dostupné online: <https://www.mdpi.com/1422-0067/22/9/4423>.
- Pugh, D., Gallacher, P.J. & Dhaun, N. (2019). Management of Hypertension in Chronic Kidney Disease. *Drugs* 79, 365–379. (2019). <https://doi.org/10.1007/s40265-019-1064-1>.

- Ranganathan, P., Aggarwal, R., & Pramesh, C. S. (2015). Common pitfalls in statistical analysis: Odds versus risk. *Perspectives in clinical research*, 6(4), 222. Dostupné online: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4640017/>.
- Redelings Matthew D., Matthew Wise, Frank Sorvillo. (2007). Using Multiple Cause-of-Death Data to Investigate Associations and Causality between Conditions Listed on the Death Certificate, *American Journal of Epidemiology*, Volume 166, Issue 1, 1 July 2007, Pages 104–108, <https://doi.org/10.1093/aje/kwm037>.
- Regitz-Zagrosek, V., & Seeland, U. (2012). Sex and gender differences in clinical medicine. *Sex and Gender Differences in Pharmacology*, 3-22. Dostupné online: https://link.springer.com/chapter/10.1007/978-3-642-30726-3_1.
- Rodriguez, F., Blum, M. R., Falasinnu, T., Hastings, K. G., Hu, J., Cullen, M. R., & Palaniappan, L. P. (2019). Diabetes-attributable mortality in the United States from 2003 to 2016 using a multiple-cause-of-death approach. *Diabetes research and clinical practice*, 148, 169-178. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/30641162/>.
- Romon, I., Jouglu, E., Balkau, B., & Fagot-Campagna, A. (2008). The burden of diabetes-related mortality in France in 2002: an analysis using both underlying and multiple causes of death. *European journal of epidemiology*, 23(5), 327-334. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/18386133/>.
- Roschmann, R. A., & Rothenberg, R. J. (1987, February). Pulmonary fibrosis in rheumatoid arthritis: a review of clinical features and therapy. In *Seminars in arthritis and rheumatism* (Vol. 16, No. 3, pp. 174-185). WB Saunders. Dostupné online: [https://doi.org/10.1016/0049-0172\(87\)90020-5](https://doi.org/10.1016/0049-0172(87)90020-5).
- Roy-O'Reilly, M., & McCullough, L. D. (2014). Sex differences in stroke: the contribution of coagulation. *Experimental neurology*, 259, 16-27. Dostupné online: <https://doi.org/10.1016/j.expneurol.2014.02.011>.
- Ruiz-Irastorza, G., & Khamashta, M. (2008). Lupus and pregnancy: ten questions and some answers. *Lupus*, 17(5), 416–420. <https://doi.org/10.1177/0961203308090027>.
- Sally C. Curtin a Pedro Martinez. (2020). QuickStats: Age-Adjusted Suicide Rates, by Sex and Three Most Common Methods — United States, 2000–2018. *MMWR Morbidity and Mortality Weekly Report*, 2020;69:249. DOI: <http://dx.doi.org/10.15585/mmwr.mm6909a7>.
- Samai, A. A., & Martin-Schild, S. (2015). Sex differences in predictors of ischemic stroke: current perspectives. *Vascular health and risk management*, 11, 427. Dostupné online: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4524454/>.
- Santiago JA, Potashkin JA. (2021) The Impact of Disease Comorbidities in Alzheimer's Disease. *Front Aging Neuroscience*. 2021 Feb 12;13:631770. doi: 10.3389/fnagi.2021.631770. PMID: 33643025; PMCID: PMC7906983. Dostupné online: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7906983/>.
- SEER*Explorer: An interactive website for SEER cancer statistics [Internet]. Surveillance Research Program, *National Cancer Institute*. [Cited 2021 September 27]. Dostupné online: <https://seer.cancer.gov/explorer/>.

- Siddique, Md & Nur, Zannatun & Mahbub, Md. Shahriar & Alam, Md & Miah, Md. (2009). Clinical Presentation and Epidemiology of Stroke :A Study of 100 Cases. *Journal of Medicine*. 10. 86. Dostupné online: https://www.researchgate.net/publication/206644682_Clinical_Presentation_and_Epidemiology_of_Stroke_A_Study_of_100_Cases.
- Snider, G. L. (1989). Chronic obstructive pulmonary disease: risk factors, pathophysiology and pathogenesis. *Annual review of medicine*, 40(1), 411-429. Dostupné online: https://www.annualreviews.org/doi/abs/10.1146/annurev.me.40.020189.002211?casa_token=7kWgNPVOp6gAAAAA%3AzzD_P6_eYpQFcr_qvJcNw_QbumlEqEJ25CwCJLRUn--FEFOuQsBjrW-yOXIOEhtOrrgsWX_cGsn-g&journalCode=med.
- Strauss, M. J., Niederkrotenthaler, T., Thurner, S., Kautzky-Willer, A., & Klimek, P. (2021). Data-driven identification of complex disease phenotypes. *Journal of the Royal Society Interface*, 18(180), 20201040. Dostupné online: <https://royalsocietypublishing.org/doi/full/10.1098/rsif.2020.1040>.
- Tan, L., Ye, Y., Xiao, K. et al. (2019). A clinicopathological analysis of adrenal tumors in patients with history of extra-adrenal cancers. *BMC Cancer* 19, 838 (2019). <https://doi.org/10.1186/s12885-019-6061-y>.
- Toor, R., & Chana, I. (2021). Network Analysis as a Computational Technique and Its Benefaction for Predictive Analysis of Healthcare Data: A Systematic Review. *Archives of Computational Methods in Engineering*, 28(3), 1689-1711. Dostupné online: <https://link.springer.com/article/10.1007/s11831-020-09435-z>.
- TRANSAX, the NCHS system for producing multiple cause-of-death statistics. (1986). 1968-78 Personal Author(s) : Ronald F.;Evans, Marshall C.; Corporate Authors(s) : *National Center for Health Statistics* (U.S.) Published Date : June 1986 Series : DHHS publication ; no. (PHS) 86-1322;Vital and health statistics. no. 20 Series 1 Programs and collection procedures; Dostupné online: <https://stacks.cdc.gov/view/cdc/13008>
- Tronieri, J. S., Wurst, C. M., Pearl, R. L., & Allison, K. C. (2017). Sex differences in obesity and mental health. *Current psychiatry reports*, 19(6), 1-11. Dostupné online: <https://link.springer.com/article/10.1007/s11920-017-0784-8>.
- Ústav zdravotnických informací a statistiky ČR. (2006). Jak správně postupovat při vyplňování Listu o prohlídce mrtvého (LPM). Stručný průvodce. ISBN: 80-7280-637-8. Dostupné online: https://www.uzis.cz/sites/default/files/knihovna/lpm_pruvodce_2006.pdf.
- Végh, J., Szilasi, M., Soós, G., Dévényi, K., Dezso, B., Soltész, P., ... & Bodolay, E. (2005). Interstitial lung disease in mixed connective tissue disease. *Orvosi Hetilap*, 146(48), 2435-2443. Dostupné online: <https://europepmc.org/article/med/16408383>.
- Viera AJ. (2008). Odds ratios and risk ratios: what's the difference and why does it matter? *South Medical Association*, J. 2008 Jul;101(7):730-4. doi: 10.1097/SMJ.0b013e31817a7ee4. PMID: 18580722. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/18580722/>.
- Volkow, N. D. (2014). America's addiction to opioids: Heroin and prescription drug abuse. *Senate Caucus on International Narcotics Control*, 14, 1-16. Dostupné online: <https://spdsafeplace.com/documents/departments/cityAttorney/opioidLitigation/FN54-Testimony-Dr.NoraVolkow-05-14-14.pdf>.

- Weissman, M. M., & Klerman, G. L. (1977). Sex differences and the epidemiology of depression. *Archives of general psychiatry*, 34(1), 98-111. doi:10.1001/archpsyc.1977.01770130100011. Dostupné online: <https://pubmed.ncbi.nlm.nih.gov/319772/>.
- Westergaard, D., Moseley, P., Sørup, F. K. H., Baldi, P., & Brunak, S. (2019). Population-wide analysis of differences in disease progression patterns in men and women. *Nature communications*, 10(1), 1-14. Dostupné online: <https://www.nature.com/articles/s41467-019-08475-9>.
- Wilhelm, K. A. (2001). Men and depression. *Australian family physician*, 30(1), 102-5. Dostupné online: <https://search.informit.org/doi/abs/10.3316/meditext.351163644136192>.
- Wilkins, K., Wysocki, M., Morin, C., & Wood, P. (1997). Multiple causes of death: Some conditions and diseases contribute to deaths much more often than they are the underlying cause. *Health Reports-Statistics Canada*, 9, 19-30. Dostupné online: <https://www150.statcan.gc.ca/n1/en/pub/82-003-x/1997002/article/3235-eng.pdf?st=e58qcQ3c>.
- World Bank. (2021). Cause of death by injury. Dostupné online: <https://data.worldbank.org/indicator/SH.DTH.INJR.ZS>.
- World Health Organization. (2016). The WHO application of ICD-10 to deaths during the perinatal period: ICD-PM. Dostupné online: <https://apps.who.int/iris/bitstream/handle/10665/249515/9789241549752-eng.pdf>.
- Zhang J, Yu KF. (1998). What's the Relative Risk? A Method of Correcting the Odds Ratio in Cohort Studies of Common Outcomes. *JAMA*, 1998;280(19):1690–1691. doi:10.1001/jama.280.19.1690. Dostupné online: <https://jamanetwork.com/journals/jama/fullarticle/188182>.
- Zucker, I., & Prendergast, B. J. (2020). Sex differences in pharmacokinetics predict adverse drug reactions in women. *Biology of sex differences*, 11(1), 32. Dostupné online: <https://doi.org/10.1186/s13293-020-00308-5>.

Seznam příloh [USB]

- Příloha 1: Finální síť, ženy [USB]
Příloha 2: Finální síť, muži [USB]
Příloha 3: Ryze mužské vazby (vazby, které nemají ženy) [USB]
Příloha 4: Ryze ženské vazby (vazby, které nemají muži) [USB]
Příloha 5: Podgrafy v síti žen a podgrafy v síti mužů [USB]
Příloha 6: Skóry autorit a hub v podgrafech sítě žen a v podgrafech sítě mužů [USB]
Příloha 7: Standardní bimodální graf, ženy, množina základních příčin smrti (7a) a množina přispívajících příčin smrti (7b) [USB]
Příloha 8: Standardní bimodální graf, muži, množina základních příčin smrti (8a) a množina přispívajících příčin smrti (8b) [USB]
Příloha 9: Bimodální graf shodného sousedstva, ženy, množina základních příčin smrti (9a) a množina přispívajících příčin smrti (9b) [USB]
Příloha 10: Bimodální graf shodného sousedstva, muži, množina základních příčin smrti (10a) a množina přispívajících příčin smrti (10b) [USB]

Přílohy 1 a 2 by měly být rovněž dostupné na:

https://ukolovae.github.io/see184/network_females/ (finální síť žen)

https://ukolovae.github.io/see184/network_males/ (finální síť mužů)

Důvodem existence příloh tohoto typu je možnost dohledávání sousedů jednotlivých chorob. To může být uskutečněno zadáním příslušného třímístného MKN-kódu do pole „*Search by name*“ vlevo. Po provedení tohoto by se vpravo měl objevit seznam sousedů hledané choroby, její stupňové charakteristiky a název podgrafu, do něhož náleží.