Multimorbidity patterns at older ages: A study using graphical model and network analysis

Linh Hoang Khanh Dang, Rossella Miglio, Rosella Rettaroli, Giulia Roli

Department of Statistical Sciences, University of Bologna

Abstract

One key factor to construct appropriate healthcare and pension planning for sustainable aging populations is the possibility of identifying precise multimorbidity patterns at older ages and seizing their progression (trajectories) in time. In both developed and developing countries, understanding the structure of multimorbidity, and most ideally across time, is an urging challenge in order for groups who share the same degree of vulnerability and needs to receive assistance and intervention in a timely manner. Compared to traditional approach in literature like factorial and clustering analysis, the combination of probabilistic approach from graphical model and intuitive visualization from network analysis is emerging quickly as powerful tool in recent years to not only efficiently explore the richness of administrative health data, but also to provide an analysis framework whom predictability can be assessed. Applying these methods on longitudinal data of individuals aged 50 and above residing Emilia-Romagna region (northern Italy) in 2011 and followed up to 2019, we study the multimorbidity patterns at older ages, their changes across time, and the profiles of patients within identified patterns.

Introduction

In aging populations, multimorbidity, defined as having two or more chronic diseases, is a medical and social challenge of imminent priority, for both individuals who suffer from it, and for the healthcare and pension system whose sustainability depend on it. For the past decades, literature on the patterns of multimorbidity are concentrated in approach using methods such as factorial analysis and clustering (Prados-Torres et al., 2014). However, factorial analysis often results in diseases being included in several groups at the same time, making interpretation of resulted patterns rather complicated. On the other side, even if cluster analysis allows data partitioning into separate groups, it cannot integrate the strength of relationship between diseases within a cluster (Jones et al., 2022). New approach to model complex structures while maintaining the relationship intensity between research subjects using network analysis has been flourishing in studying diseases network in medical and biological research (Barabási et al., 2011; Hidalgo et al., 2009). This approach also keeps expanding further to other research fronts, including the network of causes of death (Egidi et al., 2018) and of

multimorbidity patterns (Batista et al., 2022). However, the choice of association measures between diseases varies across studies and can bring drastically different results on the same data (Monchka et al., 2022). The usual association measurements (e.g., lift, relative risk, phi) can only serve as description tools and limit the network analysis at its descriptive level. Recently, graphical models have become a more widely-used method to abstract complex systems in various disciplines by providing probabilistic associations between subjects. Of which, its application combined with network analysis in multimorbidity research also began, even only based on one-year data (Alvarez-Galvez and Vegas-Lozano, 2022). Using the longitudinal administrative health data of population aged 50 and above in 2011 from Emilia-Romagna region (northern Italy), and followed up to 2019, we study the multimorbidity patterns, their network structure, their network formation across time, and the profile of individuals in these patterns.

Data

The data sources for this study include census data in 2011, population register, and hospital administrative data from 2011 to 2019 in Emilia-Romagna region, which are linked together using a linkage record process based on unique personal identification. In addition, the census data combined with population register provide us a representative data of the population residing in the region at census date, free of selection bias that could undermine the analysis results. Apart from information about chronic conditions, we also obtain other covariates including sex, date of birth (and date of death, if the individual passed away), civil status, education level, occupation for each individual. Individuals who were present at census date but moved out of the region during follow-up period are right-censored. Individuals who moved into the region after census date are excluded. Only individuals who were diagnosed or recorded with at least two chronic diseases during observation period are included in the study (in total: 1,010,610 individuals, of which 557,601 females and 453,009 males).

Methods

Weighted, undirected pairwise chronic diseases network is constructed using mixed graphical models (Yang et al., 2014). Each node is one chronic disease, coded as a binary variable, where absence of disease is coded "0" and presence of disease is coded "1" for all individuals included in the study. The association analysis is limited to diseases dyads (i.e., mixed graphical models of interaction at order 2). Mixed graphical model estimations are done using package `mgm` (version 1.2-14) in R (Haslbeck and Waldorp, 2020). Network structure visualization and network centrality measures are made using package `qgraph` (version 1.9.5) in R (Epskamp et al., 2012). All analyses are done separately for men and women.

Preliminary results

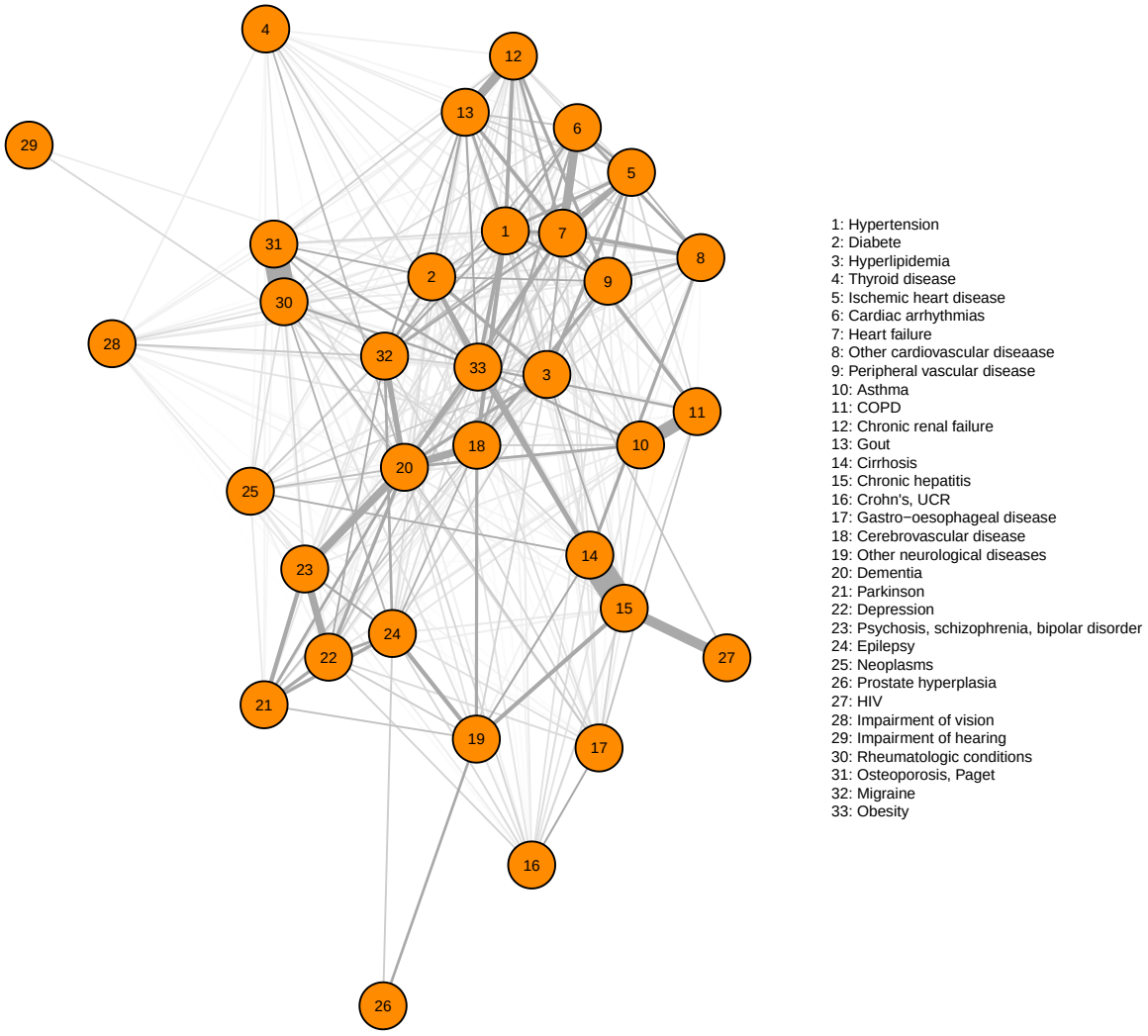
Mapping the multimorbidity network

Figure 1 and 2 below present the chronic diseases networks that are estimated by mixed graphical models, respectively for women and men during the observation period. The non-directional relations between two nodes indicate that two diseases are diagnosed during observation period. The thickness of the edges linking two nodes represent the probability of observing a tie formed between two given chronic diseases. For both females and males, several groups of diseases can be found through their visibly strong linkage. One expected linkage can be observed between arrhythmia (or cardiac arrhythmias), heart failure, ischemic heart disease, peripheral vascular disease, hypertension and other cardiovascular diseases. Another strong linkage includes dementia, psychosis, schizophrenia, bipolar disorder, migraine, depression, Parkinson, epilepsy and other neurological disease. The strong linkage between cirrhosis and chronic hepatitis is found to be linked with the group of obesity-diabetes-hyperlipidemia at one side and with HIV at another side. Three other linkages of strength are between rheumatologic conditions and osteoporosis-Paget's disease, between asthma and COPD, and between chronic renal failure and gout. However, if the group chronic renal failure-gout seems to have weak association with the group rheumatologic conditions - osteoporosis-Paget's disease for females, this relationship is especially strong for males. In addition, other than with mental disorders and diseases, dementia also holds strong association with cerebrovascular disease, which is in turn directly connected with obesity for females, but not for males.

Measuring the centrality of multimorbidity network

Combined with the visual structure of diseases network, the role each chronic disease has in the network can be assessed using centrality indicators. Figure 3 and figure 4 respectively report the values of betweenness centrality and closeness centrality for females and males. In this preliminary analysis, focus can be first given to betweenness centrality that informs us of diseases that serve as connecting nonadjacent diseases. In other words, diseases of great betweenness centrality are diseases only through which interactions between two nonadjacent diseases can be made. These diseases play an important roles as bridges, or "gatekeepers" for patterns of multimorbidity to form. For females, the top 5 gatekeepers are dementia, cerebrovascular disease, cirrhosis, heart failure and obesity. For males, the top 5 gatekeepers are dementia, cirrhosis, cerebrovascular diseases, heart failure and osteoporosis and Paget's disease. These diseases also share relatively higher closeness index values. This latter is computed using the geodesic distance (i.e., the length of the shortest path between a pair of nodes). Hence, the determined gatekeeper-diseases with higher closeness centrality measures are also diseases that are more connected through short path to other diseases. Attention to these diseases might be essential for healthcare planning and related studies.

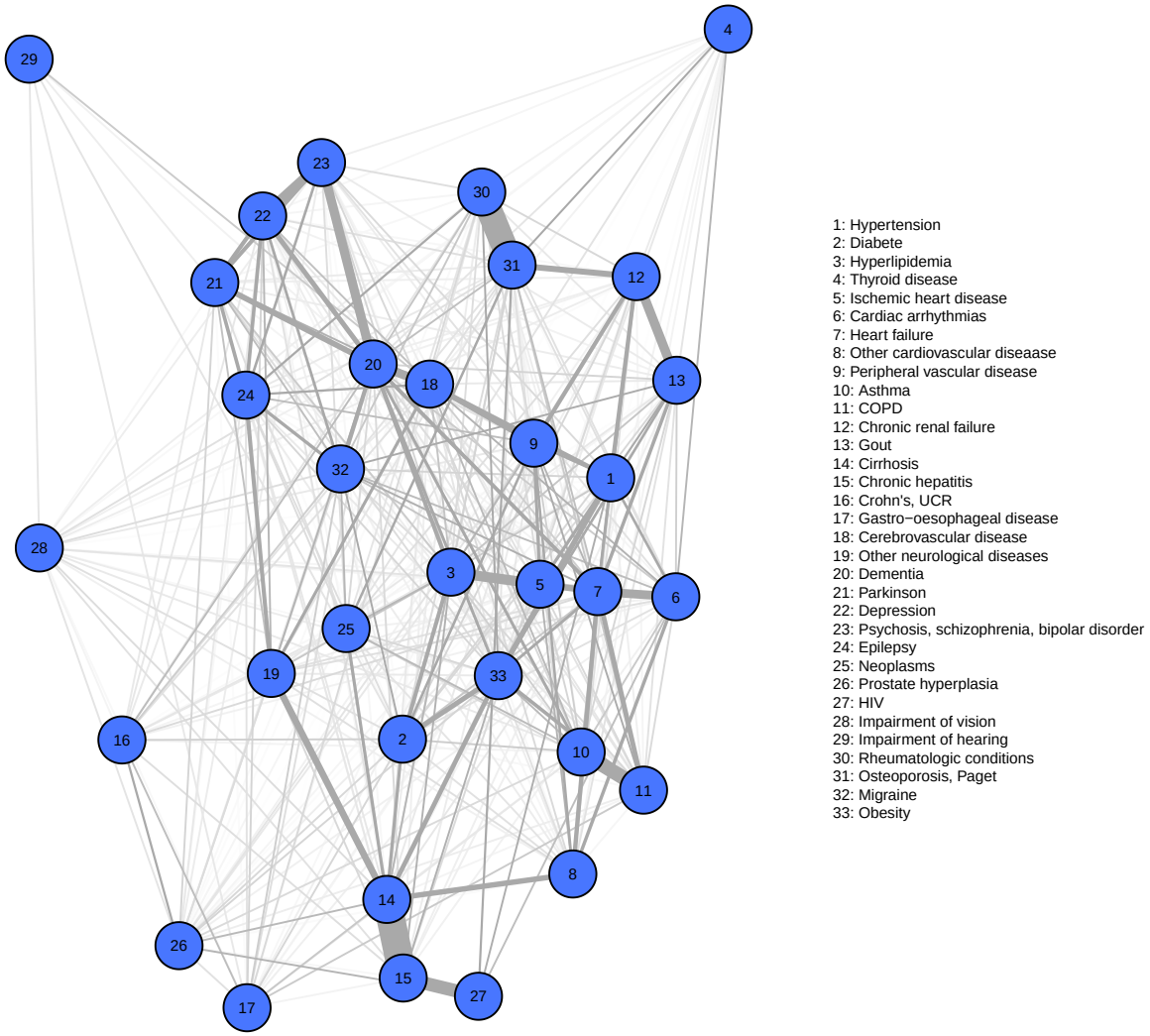
Figure 1: Multimorbidity network using mixed graphical model, females aged 50 and above, observation period 2011-2019



Further developments

Our next steps constitutes firstly of using community detection algorithm to group chronic diseases into clusters within the network structure estimated by graphical models. Community detection is a classic approach to clustering in network analysis, and while there is a good number of algorithms for this task (namely, the walktrap algorithm, the Louvain algorithm, the fast-greedy algorithm, etc.), the choice of a community detection method is not neutral nor easy. Different methods/algorithms might lead to drastically different results. However, throughout the literature, the clarity in the choices that were made is often an issue. The lack of details often limits how one can understand the way results are obtained. In this paper, we will apply the Question-Alignment approach (Smith et al., 2020) to identify the most

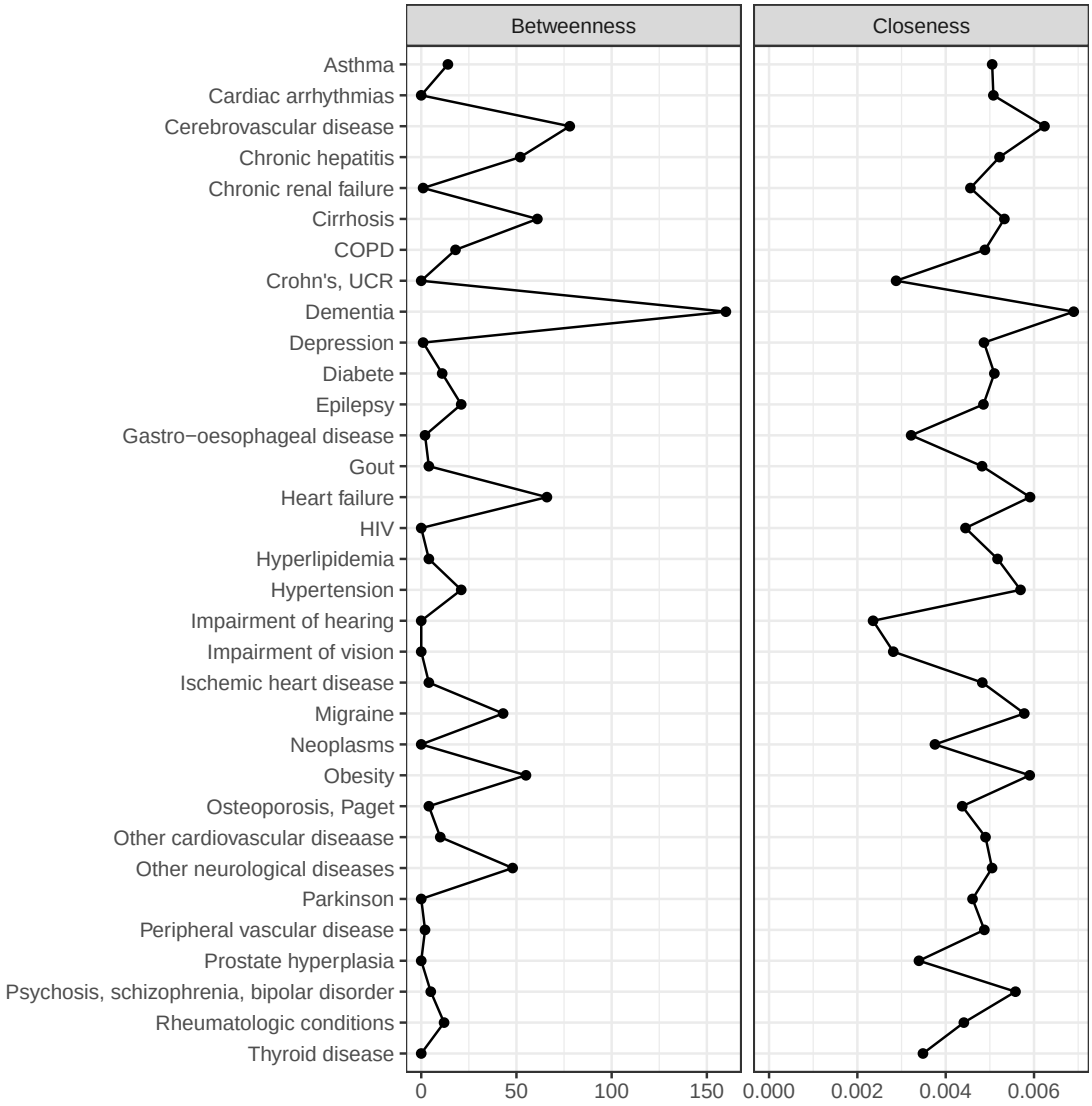
Figure 2: Multimorbidity network using mixed graphical model, males aged 50 and above, observation period 2011-2019



suitable community detection algorithm for our research question and data type, making our classification process transparent and fully documented. The set of most common algorithms to choose from is available in R package `igraph` (Csardi and Nepusz, 2006). The resulted groups of chronic diseases will give us the global view of the multimorbidity patterns existing in the cohort population of individuals aged 50 in 2011 and followed up until 2019.

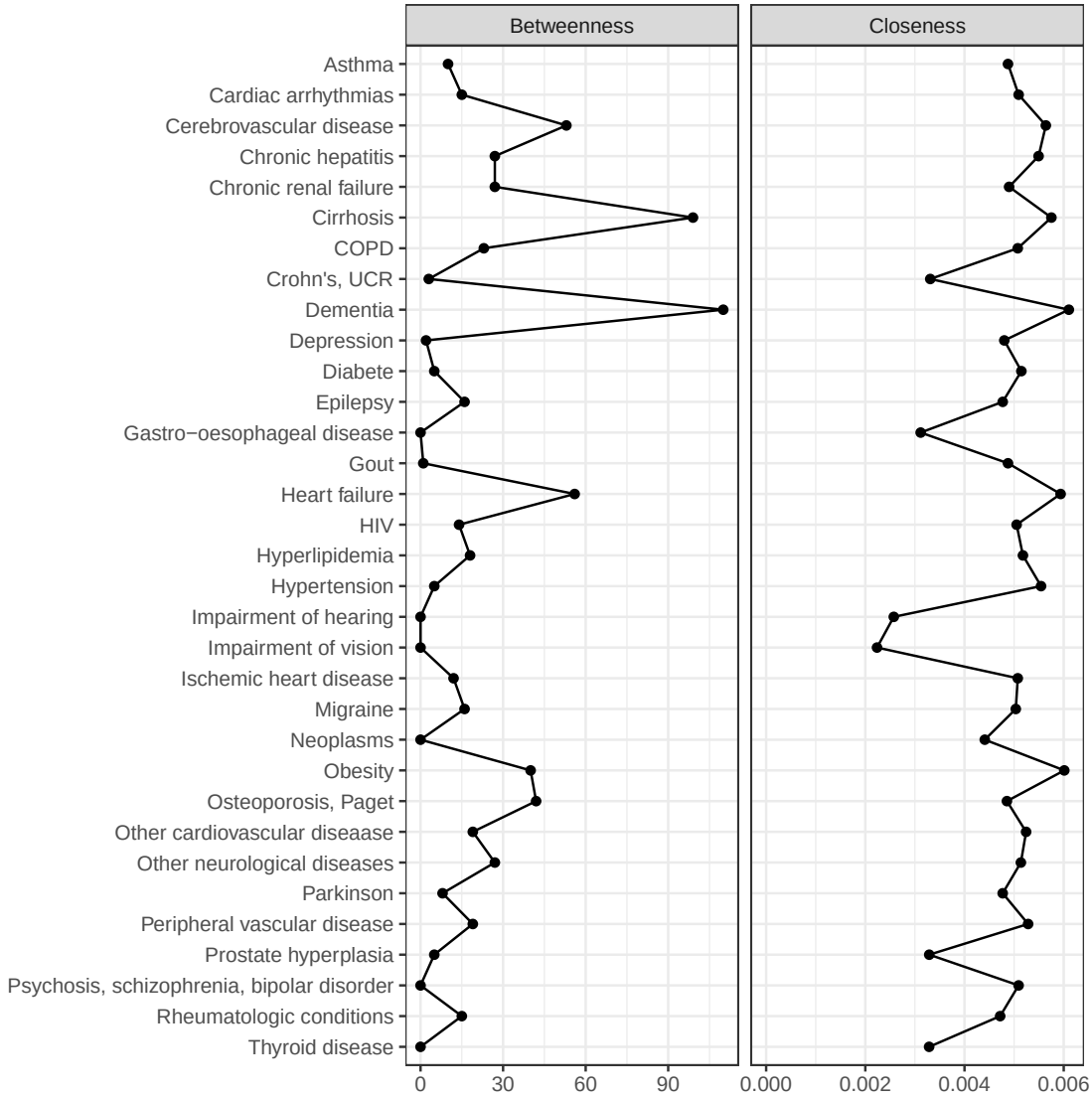
A longitudinal data setting does not only give us enough follow-up time to observe chronic diseases that might take time to manifest and be diagnosed, and thus allows us to depict as close to reality as possible the main multimorbidity patterns of a population aged 50 and above, repeated measures on the same individuals also provide us with information of the changes in chronic diseases each year. Hence, we continue to expand our model to dynamic network mod-

Figure 3: Centrality indexes for females aged 50 and above, observation period 2011-2019



eling, namely the stochastic actor-oriented model (SAOM) by Snijders et al. (2010), to study the ties formation of the chronic diseases within population with multimorbidity, separately for female and male population included in our study. Most of previous studies using network analysis on multimorbidity adopt rather a cross-sectional approach, and the dynamic component that lies in the nature of network analysis still has more to offer. It is worthy to note however that we do not interpret our study in the causal sense of multimorbidity formation. Our objective is to peel the first layer of the process, which is to model the ties formation between chronic diseases in time based on our exhaustive and reliable individual data across different time points (herein, 9 time points). Another approach to the longitudinal multimorbidity data is to use sequence analysis to cluster patterns of chronic diseases at older ages, which could be useful to compare with longitudinal network analysis. This study of the patterns of multimorbidity and its network formation can bring valuable insights for related research, such as the study of multiple causes of death (Dobson et al., 2023; Désesquelles and Meslé, 2004), as well as their competing structure and impact on mortality (Berry et al., 2010), thereby jointly provide

Figure 4: Centrality indexes for males aged 50 and above, observation period 2011-2019



an integrated understanding of the morbidity and mortality process in aging populations.

Acknowledgements

Longitudinal data used for this study are prepared in collaboration with Nicola Caranci and Daniela Fortuna from Department of Innovation in Health and Social Care, Direction of Health and Welfare, Emilia-Romagna region (Bologna, Italy). This research was funded by the 2020 PRIN-Research Project of National Relevance "SOcial and health Frailty as determinants of Inequality in Aging (SOFIA).

References

Alvarez-Galvez, J. and E. Vegas-Lozano (2022). Discovery and classification of complex multimorbidity patterns: unravelling chronicity networks and their social profiles. *Nature: Sci-*

entific Reports 12(20004).

Barabási, A., N. Gulbahce, and J. Loscalzo (2011). Network medicine: a network-based approach to human disease. *Nature Reviews Genetics 12*, 56–68.

Batista, S. R., A. L. L. Sousa, B. P. Nunes, R. R. Silva, P. C. B. V. Jardim, and for the Brazilian Group of Studies on Multimorbidity (GBEM) (2022). Identifying multimorbidity clusters among brazilian older adults using network analysis: Findings and perspectives. *PLOS ONE 17*(7), 1–16.

Berry, S. D., L. Ngo, E. J. Samelson, and D. P. Kiel (2010). Competing risk of death: an important consideration in studies of older adults. *Journal of the American Geriatrics Society 58*(4), 783–787.

Csardi, G. and T. Nepusz (2006). The igraph software package for complex network research. *InterJournal Complex Systems*, 1695.

Dobson, A., P. McElwee, M. Baneshi, J. Eynstone-Hinkins, L. Moran, and M. Waller (2023). A new data driven method for summarising multiple cause of death data. *BMC Medical Research Methodology 23*(83).

Désesquelles, A. and F. Meslé (2004). Intérêt de l'analyse des causes multiples dans l'étude de la mortalité aux grands âges: l'exemple français. *Cahiers québécois de démographie 33*, 83–116.

Egidi, V., G. Rivellini, M. A. Salvatore, and S. D'Angelo (2018). A network approach to studying cause-of-death interrelations. *Demographic Research 38*(16), 373–400.

Epskamp, S., A. O. J. Cramer, L. J. Waldorp, V. D. Schmittmann, and D. Borsboom (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software 48*(4), 1–18.

Haslbeck, J. M. B. and L. J. Waldorp (2020). mgm: Estimating time-varying mixed graphical models in high-dimensional data. *Journal of Statistical Software 93*(8), 1–46.

Hidalgo, C. A., N. Blumm, A.-L. Barabási, and N. A. Christakis (2009, 04). A dynamic network approach for the study of human phenotypes. *PLOS Computational Biology 5*(4), 1–11.

Jones, I., F. Cocker, M. Jose, M. Charleston, and A. L. Neil (2022). Methods of analysing patterns of multimorbidity using network analysis: a scoping review. *Journal of Public Health* (31), 1217–1223.

Monchka, B., C. Leung, N. Nickel, and L. L.M (2022). The effect of disease co-occurrence measurement on multimorbidity networks: a population-based study. *BMC Medical Research Methodology 22*(165).

Prados-Torres, A., A. Calderón-Larrañaga, J. Hancco-Saavedra, B. Poblador-Plou, and M. van den Akker (2014). Multimorbidity patterns: a systematic review. *Journal of Clinical Epidemiology 67*(3), 254–266.

Smith, N. R., P. N. Zivich, L. M. Frerichs, J. Moody, and A. E. Aiello (2020). A guide for choosing community detection algorithms in social network studies: The question alignment approach. *American Journal of Preventive Medicine 59*(4), 597–605.

Snijders, T. A., G. G. van de Bunt, and C. E. Steglich (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks 32*(1), 44–60. Dynamics of Social Networks.

Yang, E., Y. Baker, P. Ravikumar, G. Allen, and Z. Liu (2014). Mixed Graphical Models via Exponential Families. In S. Kaski and J. Corander (Eds.), *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*, Volume 33 of *Proceedings of Machine Learning Research*, Reykjavik, Iceland, pp. 1042–1050. PMLR.