

NETWORK ANALYSIS: A MULTIVARIATE STATISTICAL APPROACH FOR HEALTH SCIENCE RESEARCH

Análise de redes: uma abordagem de estatística multivariada para pesquisas em ciências da saúde

Daniel Eduardo da Cunha Leme^a , Erika Valeska da Costa Alves^a ,
Vinícius do Carmo Oliveira Lemos^b , André Fattori^a 

ABSTRACT

Network analysis is a graphical statistical technique that allows rapid visualization and interpretation of associations between multiple variables. There are still few theoretical studies on this method, especially in the areas of geriatrics and gerontology research, which cover the study of different social, clinical, or physical and mental health variables. The objectives of this study were to present the main theoretical aspects of network analysis and demonstrate its applicability by analyzing studies that used this technique, offering an accessible language for all levels of knowledge in statistics. The main characteristics of the graphs, basic theoretical concepts, and scientific articles that used networks were demonstrated. This methodological study can help the reader to understand this analytical method, which is still little explored in national research. There is a scarcity of research on this subject in the areas of geriatrics and gerontology; however, technological advances, the availability of statistical programs with new data analysis resources, and the dissemination of information are relevant factors for the expansion of knowledge and the use of network analysis in this context.

KEYWORDS: statistical analysis; biostatistics; epidemiological studies; methods; geriatrics.

RESUMO

A análise de redes é uma técnica estatística gráfica que permite a rápida visualização e interpretação de associações entre múltiplas variáveis. Ainda existem poucos estudos com conteúdo teórico sobre esse método, especialmente nas áreas de pesquisa da geriatria e gerontologia, as quais abrangem o estudo de diferentes variáveis sociais, clínicas ou de saúde física e mental. Os objetivos deste estudo foram apresentar os principais aspectos teóricos e demonstrar a aplicabilidade da análise de rede por meio de estudos que utilizaram essa técnica, oferecendo uma linguagem acessível para todos os níveis de conhecimento em estatística. Foram demonstrados as principais características dos gráficos, conceitos teóricos básicos e artigos científicos que utilizaram redes. O presente estudo metodológico pode auxiliar o leitor na compreensão desse método analítico, ainda pouco explorado no âmbito da pesquisa nacional. Nas áreas de geriatria e gerontologia há escassez de pesquisas que abordam essa temática, entretanto o avanço tecnológico, a disponibilidade de programas estatísticos com novos recursos de análise de dados e a divulgação de informações são fatores relevantes para a expansão do conhecimento e a utilização da análise de redes.

PALAVRAS-CHAVE: análise estatística; bioestatística; estudos epidemiológicos; métodos; geriatria.

^aUniversidade Estadual de Campinas – Campinas (SP), Brazil.

^bUniversidade Federal do Rio de Janeiro – Rio de Janeiro (RJ), Brazil.

Corresponding author

Daniel Eduardo da Cunha Leme – Faculdade de Ciências Médicas, Universidade Estadual de Campinas – Rua Tessália Vieira de Camargo, 126 – Cidade Universitária Zeferino Vaz – CEP: 12950-602 – Campinas (SP), Brazil – E-mail: daniel.eduardo.7@hotmail.com

Received on: 09/28/2019. Accepted on: 12/06/2019

DOI: 10.5327/Z2447-212320201900073



INTRODUCTION

Network science is an approach originated in graph theory, a field of mathematics that studies relationships between objects,^{1,2} which was idealized in the 18th century. Networks also have roots in sociology, especially in the study of relationships between human beings considering groups of family, friends, and work colleagues. Social networks are more relevant in people's behavior and worldview than categories such as sex, social class, or age and demonstrate the importance of each individual through their personal relationships during life.^{1,3}

Modern network science known today evolved methodologically and theoretically in the early 1950s with studies in different areas of knowledge, such as mathematics, sociology, anthropology, and physics.⁴ The evolution of information technology and the availability of new resources in statistical programs, led other areas to adopt this analytical method in their search to explore data, such as marketing, information science, political science and computing.³

In genetics, researchers used network analysis to verify complex relationships between multiple genes in the development of diseases such as cancer.⁵ In health science, an integrative review research demonstrated that network analysis was used in prior studies to check information on supply and demand for medical specialties in order to characterize a profile of increased utilization of primary health services, helping to compress the functioning and management of these services.⁶ In addition, scholars used the network approach to identify risk factors in health and observe correlations between prevalent chronic or communicable diseases to improve the accuracy of diagnosis and treatment.⁷

Health research focuses on the human being, his environment and his social capital. Studies in geriatrics and gerontology address social, clinical, physical, and mental health variables. Thus, univariate analyses may not explain the phenomenon under investigation because they capture the isolated action of each variable regarding the outcome under study.

With network analysis, it is possible to visually explore relationships that occur simultaneously between multiple variables, incorporating advanced tools in statistical analysis, such as bootstrapping techniques and Bayesian inference.^{1,3} In this context, based on the scarcity of studies that address theoretical aspects of networks, especially in the areas of geriatrics and gerontology, this methodological study aimed to present the main theoretical aspects and demonstrate the applicability of network analysis with studies that used this technique, using a language accessible to all levels of knowledge in statistics.

ELEMENTS AND MAIN CHARACTERISTICS OF NETWORKS

Networks are graphical structures composed of nodes, circle-shaped elements that represent variables. Nodes connect to each other through lines called edges. Networks can be classified as unweighted and weighted. In unweighted networks, edges represent only the relationship between nodes, and in weighted networks the magnitude of the relationships is shown. That is, the thicker the connection between nodes, the stronger the relationship between them. In addition, the edges may vary in color depending on the direction of the relationship (positive or negative). The statistical programs by default define that the green or blue color represents a positive relationship, and the red, a negative relationship (Figures 1A and 1B).³

Networks are also classified into directional and non-directional.³ In the former, the edges have arrows at one end, indicating an influence path or a path from one node to another. Some fields of science consider directional nets as attractive models for demonstrating causal structures. For example, the insomnia variable (A) can cause the fatigue variable (B). In this simple model there is a unidirectional path of influence, $A \rightarrow B$. In addition to pathways, directional nets can contain cycles. Thus, in the existence of 2 variables that influence and are influenced in a model, there will be 2 arrows between them, $A \rightleftarrows B$.⁸

It should be noted that causal network structures can be tested in cross-sectional studies, but the lack of temporal information is one of the problems in such studies, because when there is only cross-sectional information in the data, the findings should be interpreted with caution, mainly because of the causal statement. Therefore, networks estimated in cross-sectional studies are preferably non-directional.⁹

In the present study, emphasis was given to non-directional and weighted networks, specifically in relation to the study of correlations or associations between variables in this type of graphic structure commonly used in cross-sectional research. In non-directional networks, edges do not have arrows, i.e., they are lines that do not indicate the direction of relationships between variables.¹⁰

The positioning of the nodes also indicates the magnitude of the relationship. That is, the closer 2 nodes are positioned, the stronger the relationship between them, which is possible because of the positioning algorithms. One of the most used in epidemiological research is the Fruchterman-Reingold (FR) algorithm, available in statistical packages of the R statistical program. The FR is a force-directed method and resembles a system of balls connected by elastic strings. In this example, we can imagine an elastic rope connecting

2 balls in order to join them, while the other balls separate them in different directions. This is what happens in networks, and this dynamics provided by the FR algorithm results in a visually attractive graph, in which the nodes in general do not overlap.¹¹

Network analysis allows the study of data complexity. With this technique, it is possible to verify the simultaneous relations between variables, unlike traditional methods, such as regression analysis and principal component analysis, in which statistical models are obtained that can bring answers to the research hypotheses by adjustments and data reduction. Also, emergent or latent behavior can be observed in networks through dense subgraphs that contain highly correlated variables. This characteristic is similar to the latent variables of structural equations models and the latent factors of confirmatory factor analysis. However, in networks, the variables form a latent behavior characterizing a formative model, and, in structural equations models and confirmatory factor analysis, observable variables are often caused by a latent variable, characterizing a reflective model.^{3,12,13}

NETWORK TYPES

Networks can be estimated by different statistical techniques. Some of them are more common.

Covariance or correlation structures

They are networks based on estimates of correlation or covariance matrices and widely used to verify associations between nodes. However, as a negative point, they present all the associations between the studied variables, making it

difficult to interpret the models, which are characterized by a high amount of data edges, some resulting from spurious correlations. In correlation networks, the coefficients range from -1 to 1, and only interactions with a value equal to 0 are omitted. It should be noted that, in terms of strength between interactions, a correlation of 0.5 equals to a -0.5 correlation. Therefore, both are stronger than a correlation of 0.2 or -0.2.^{3,12}

Partial correlation

In partial correlation networks, edges represent relationships between 2 nodes after conditioning in all other variables of the data set. That is, when there is a connection between 2 variables, there is a correlation between them that cannot be explained by other variables in the system. This condition can be understood as a partial correlation control for all other connections.¹⁴ These networks are also known as concentration graphs or Gaussian graphical models and are part of a class of statistical models called pairwise Markov random fields.¹¹ An example of partial correlation networks can be seen in Figure 2, with 3 nodes and 2 connections. The absence of a connection between 2 nodes indicates that the variables are conditionally independent. Node A interacts strongly and positively with B, and this, negatively with C. Nodes A and C are conditionally independent, given node B. Above all, variable B plays an important role in this structure, and it is possible to interpret these relationships in different ways. First, there is a strong relationship between A and B after controlling C; similarly, C and B are related after controlling A. After controlling B, there is no longer any correlation between A and C; variable B mediates the relationship between A and C. In the correlation

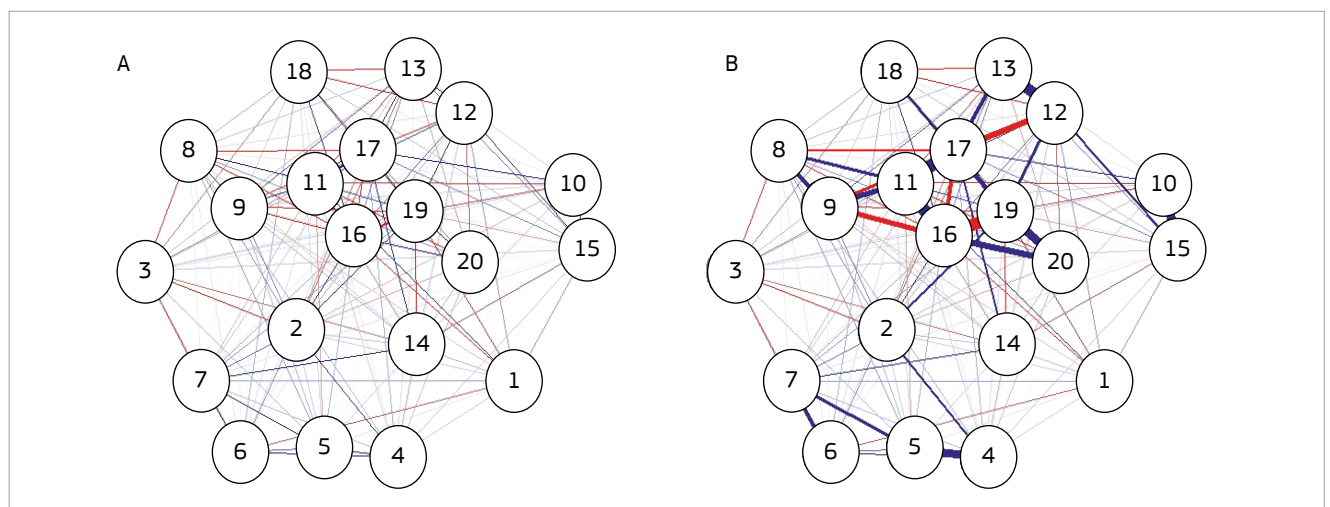


Figure 1 Illustration of network structures (A) unweighted and (B) weighted, performed in the RStudio statistical program. Red connections represent negative relationships, and blue connections, positive relationships.

and partial correlation network, polychoric correlation is used to estimate the relationship between 2 ordinal variables; polyserial, when one variable is ordinal and the other continuous; and Pearson's correlation, when the 2 variables are numeric (discrete or continuous). In the correlation and partial correlation networks, an assumption is the normality of the variable distributions and, in the absence of normal variables; the procedure for transforming these data must be performed. In the programs R or RStudio, this function is called nonparanormal transformation.^{9,15}

Mixed graphic models

It is a structure estimation approach based on generalized covariance matrices. This method is considered a non-parametric extension of the aforementioned statistical methods and the state of the art when one wants to estimate networks with dichotomous, ordinal and numeric variables, discrete or continuous.^{16,17} In health sciences, different variables are commonly analyzed simultaneously, such as age in years (discrete variable); sex (dichotomous); dose of medication (continuous) and height in centimeters (continuous). In the R and RStudio programs, this technique is available in the MGM statistical package. When using it, one must select the MGM function and then assign specific functions for each variable and its distribution. For example, the g function corresponds to the Gaussian distribution for continuous numeric variables; p, the Poisson distribution for discrete numeric variables; and c, the Bernoulli and multinomial distribution for discrete and ordinal variables, respectively.¹⁶

Least absolute shrinkage and selection operator penalty

To control spurious relations and obtain easily interpretable networks, the least absolute shrinkage and selection operator

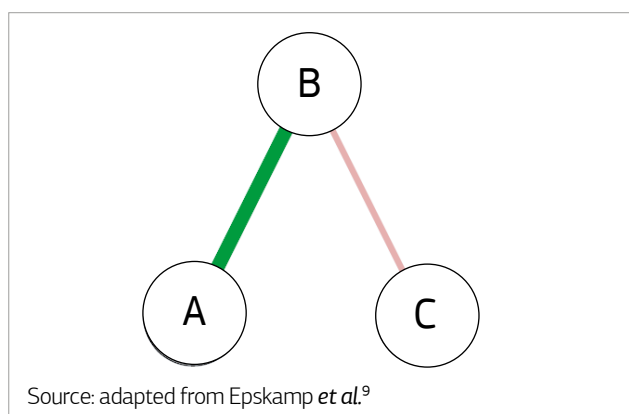


Figure 2 Partial correlation between nodes A and B, and C and B, in a simple network structure.

(LASSO) estimator is used, a method of variable selection and regularization analysis that improves the accuracy of the statistical model.¹⁸ In networks, this tool is increasingly popular to obtain a structure with as few connections as possible to explain the covariance between variables. The LASSO estimator has several variants, one of which is available in the R or RStudio statistical programs; being described as graphical Lasso (gLasso). GLasso selects relevant partial correlation coefficients, and small values are estimated as 0, resulting in parsimonious networks.^{3,9,19} As seen in Figure 3, when comparing the networks presented in items A and B, we found that, after the selection of the gLasso function in the RStudio statistical program, network B had fewer connections in relation to network A, with only partial correlation (Figure 3).

NUMBER OF NODES, ESTIMATED PARAMETERS AND SAMPLE SIZE

The greater the number of nodes, the greater the number of edges and estimated parameters. A correlation, partial correlation, or mixed graphic models network with 10 nodes, contains, in general, 55 estimated total parameters ($10 + 10 \times 9/2 = 45$), where 10 limit parameters are added to 45 paired association parameters. With 20 nodes ($20 + 20 \times 19/2 = 190$), there are 210 parameters. There must be an empirical basis for including nodes in a network analysis. In addition, to estimate a safe number of parameters, the observations in the sample must exceed the number of variables analyzed. For example, a sample with 100 observations is sufficient for 20 nodes, but a sample with 30 observations and 30 nodes may present unreliable results.¹⁵

The sample size is a relevant aspect in network estimates. As the sample increases, the parameters are estimated more accurately and there is greater statistical power. The literature refers to a number of 500 or more observations for a partial correlation network analysis, but this value may be lower when using the Lasso penalty, because with this estimator, a reduced number of parameters is sufficient to estimate the covariance of the data. However, in small samples, the correlation results must be interpreted with greater caution.^{15,18}

CENTRALITY MEASURES

In network science, centrality measures demonstrate the importance of nodes in a system. When evaluating a set of variables, not all of them have the same relevance in a model, and there are centrality indices that help in the detection of important variables and the role they play in the network

structure, within the scope of their relationships. Centrality indices are also used to model or predict various network processes and provide the researcher with an intervention guide.⁹

Three measures of centrality in networks are best known:

- betweenness (degree of connectivity): represents the number of times that a node is part of the shortest path between all pairs of connected nodes in the network;
- closeness: a measure obtained by the inverse of the distance of a node with all the others in the network, that is, the more central the node is positioned in the network, the closer it is to all other nodes;
- strength: derives from the sum of all the paths that connect one node to the others and is based on the mean of the correlation weights, with the centrality index being the most important (Figure 4).²⁰

SOFTWARE AND STATISTICAL PACKAGES

The following describes different statistical programs or packages that provide resources for creating and viewing networks:

- Igraph: is a free statistical package available in several formats for download, mainly as a package of the R and Python™ statistical programs. With this tool, it is possible to estimate directional and non-directional networks. This package can execute 2- and 3-dimensional networks, in addition to estimating node centrality measures;²¹
- qgraph: free statistical package available for download of the R and RStudio statistical programs. It was developed due to the need to study networks in psychometrics and psychopathological approaches. With this tool, it is possible to estimate and visualize the correlation and partial correlation networks and those with Lasso penalty. It is also used to visualize outputs of structural equation models and exploratory and confirmatory factor analysis, elaborated with the *sem* and *lavaan* functions. The qgraph differs from other statistical packages, as it aims to present non-sparse and weighted graphs, as well as more advanced visualization features of networks with multiple nodes, maintaining their quality in the output. Another relevant feature is that the qgraph was designed for intermediate and advanced users in R statistical language. It is also possible to visualize analyses of centrality measurement and specific tests of stability and accuracy in networks with resampling techniques (bootstrap);²²
- bootnet: free statistical package available for download, function also present in the statistical programs R and RStudio. This method allows the estimation and visualization of networks, as well as the development of stability and accuracy tests in networks. It is also possible to create non-directional and directional networks with bootnet;²³
- pcalg: free package available in the R program that only enables the construction of directional networks for the study of the inductive causal network. After the elaboration of the analyses in this package, the visualization of the directional networks occurs through the qgraph package;^{3,24}

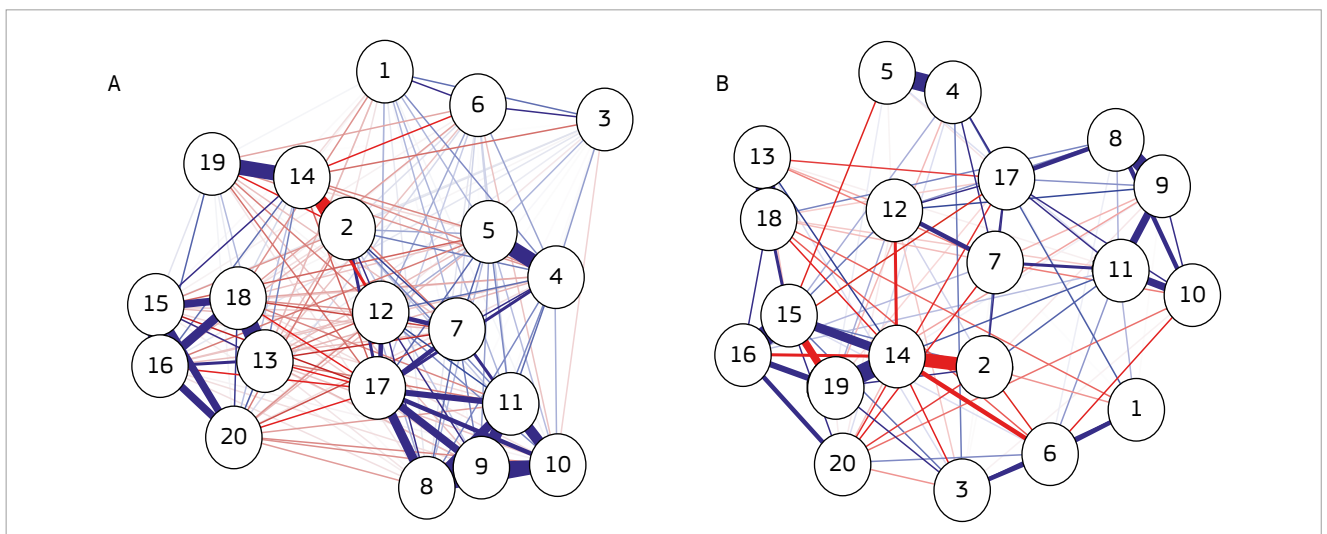


Figure 3 Network structures estimated via (A) partial correlation and (B) gLasso estimator, with fewer connections developed with the RStudio statistical program.

- JASP: free, downloadable, and easy to use graphic statistical program, specially designed for researchers used to statistical programs such as Statistical Package for the Social Sciences (SPSS®) and STATA®. With JASP, it is possible to develop networks of correlation and partial correlation and with the gLasso method. It is important to emphasize that this statistical program is also used to perform factor analysis and models of structural equations, besides producing tables and plots with appropriate formatting for publication of scientific articles or books;^{25,26}
- NodeXL Basic®: free statistical package available for Microsoft Excel®. In 2015, a non-free version of this tool was developed, NodeXL Pro®. With both packages it is possible to create and visualize networks, however the paid version of the application has advanced functions in social networks, such as access to importers of social media network data, advanced network metrics, and automation.^{14,27}

NETWORKS IN HEALTH SCIENCES

In the health sciences, especially in epidemiological and public health research, research topics commonly studies include behavioral and relational situations inherent to the process of disease transmission, associations between chronic

diseases, dissemination of information and innovations in health, and influence of groups in relation to health risk behaviors. In this context, network analysis is an appropriate technique to explore and describe multiple conditions considering a complex data structure.¹

With network analysis, it is possible to study biological phenomena seen as complex and multifactorial. In fact, when studying cancer, it is known that mutations occur in dozens of genes in the body and that there is no specific gene for this condition. Another example can be observed when studying human consciousness, a phenomenon that results from the evolution of the nervous system, which cannot be explained by a single neuron, involving billions of synapses.²⁸

The more we know about the functioning of individual genes or neurons, the less we understand the system as a whole and, for this reason, studies are currently oriented towards understanding the complexity of the phenomena. According to contemporary science, we have reached the limits of reductionism, and, at this point, a new concept of data treatment has emerged amidst advances in statistics, contrasting this paradigm.^{3,28,29}

Thus, networks were applied in studies on health problems prevalent in the population, specifically demonstrating important associations between multiple genes and diseases in the cardiovascular, endocrine, gastrointestinal, and immune

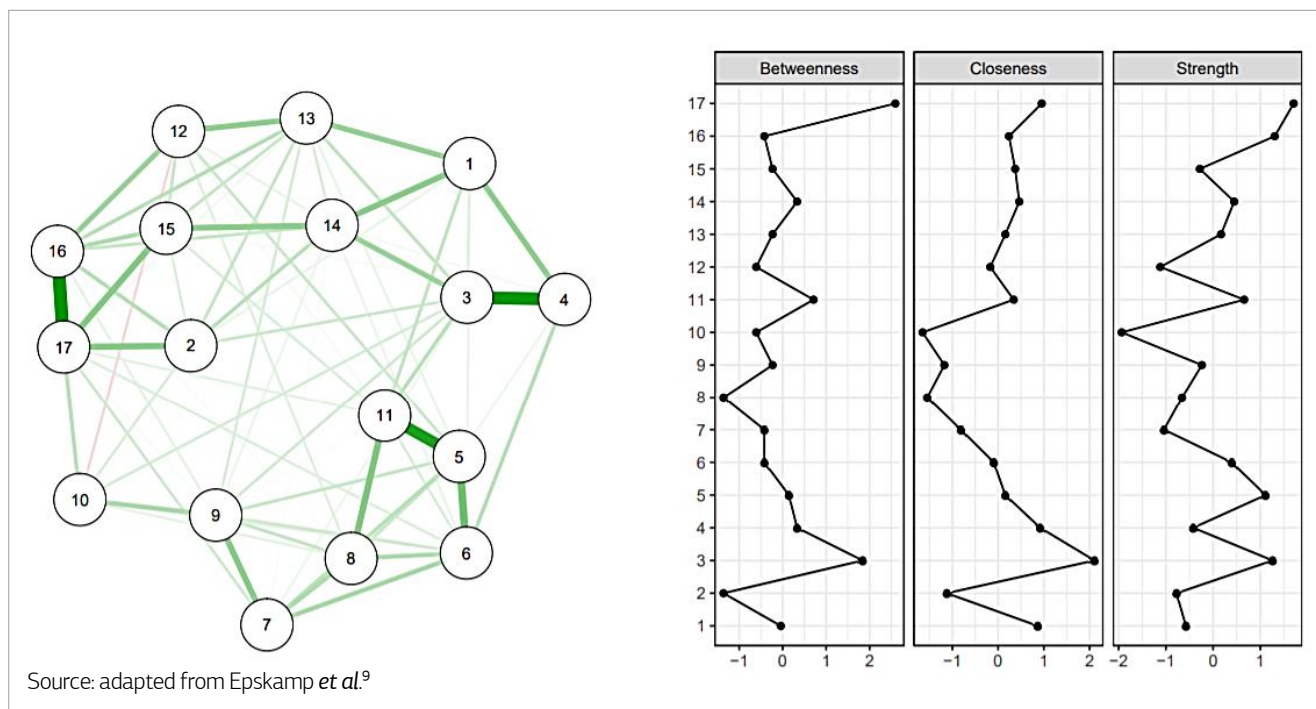


Figure 4 Simulated partial correlation network structure in the study by Epskamp et al.⁹ Considering standardized z-score values, nodes 3, 11, and 17 have high indexes, being relevant in the model.

systems and cancer, as well as comparing differences in the population's multimorbidity pattern.³⁰

Nair et al.³¹ demonstrated the application of network analysis to identify the main inflammatory genes in coronary artery disease (CAD). The researchers studied 124 genes associated with inflammation and CAD. Based on the analysis of centrality measures, 5 genes with a key role in the network were observed, which were more central and obtained greater numbers of connections. They are: genes for interleukin 6 (IL-6), vascular endothelial growth factor A (VEGFA), interleukin-1 beta (IL-1 β), tumor necrosis factor (TNF), and prostaglandin-endoperoxide synthase-2 (PTGS2). According to the authors, the network approach in the study of genes related to cardiovascular disease allows the observation of complex relationships, considering all possible variables in a model and individual components that are more prevalent in the analyses and, thus, allows a greater understanding of the disease evolutionary process and identification of target genes for a specific treatment.

Also regarding the study of chronic diseases, Kalgotra et al.³² studied electronic medical records of 21.9 million patients evaluated during the period from 2000 to 2016 in 662 American hospitals, in order to verify the correlations between the prevalent chronic diseases and demonstrate a multimorbidity profile in the analyzed population. For the estimation of the networks, each node represented a specific group of multiple diseases in different organic systems (infectious, cardiometabolic, neurological, endocrine, psychological, respiratory, dermatological and musculoskeletal diseases). A difference in the characteristics of multimorbidity networks between sexes was observed. Women were more prone to multimorbidity and obtained networks with stronger connections compared to men. Still, among women, multimorbidity was strongly related to psychological illnesses. In men, cardiometabolic diseases were prevalent.

Network analysis is widely used in psychopathology studies to help understand mental disorders. In the theory of psychopathological networks, psychological conditions arise from complex direct interactions between the symptoms represented by nodes, which are interconnected through biological, psychological, and social mechanisms. Associations and predictions using this statistical technique allow implications for the understanding of the diagnosis and treatment of mental disorders.³³

Van Wanrooij et al.³⁴, in a study conducted with 3526 older adults aged between 70 and 78 years, verified the relation of the items 'helplessness' and 'worthlessness' from the Geriatric Depression Scale (GDS) with functional loss, and the item 'memory problems' with dementia. For the authors,

the network analysis approach allowed the study of the relations between specific depressive symptoms with functional losses and dementia and also demonstrated how individual items of the scale can contribute differently to each association, which was not observed in previous studies.

Solmi et al.³⁵ analyzed the complex relationship between different demographic and nutritional variables, depressive symptoms, quality of life, multimorbidity, number of drugs in use, alcohol consumption, and level of physical activity in a population of 3552 aged and middle-aged adults. Using the partial correlation networks, the main findings demonstrated that the depressive symptoms variable was strongly and negatively related to quality of life relevant to physical and mental health, income, and education, and maintained an indirect relationship with multimorbidity and polypharmacy in quality of life concerning physical health. The authors stressed that these findings are relevant in the development of multidimensional health strategies, aimed at physical and, above all, psychological vulnerabilities in the population.

This statistical technique was used to analyze the quality of life between different groups, as shown in the research by Kossakowski et al.³⁶, which compared the health-related quality of life measured by the short form health survey 36 (SF-36) in 1742 healthy individuals, with a mean age of 57 years, and 485 cancer patients, with a mean age of 46 years. The authors demonstrated the correlations between the nodes that represented items on the SF-36 scale and also observed similarities in the network structures. The 2 networks obtained strong correlations between the items limitation to walk more than 1 kilometer and limitation to walk 1 block, of the function / activity dimension; reduction of time devoted to work or other activity that one likes and performed less tasks that one would like, of the dimension of emotional aspects; body pain during the last 4 weeks and body pain that interfered with normal work in the last 4 weeks, of the pain dimension; and time feeling depressed without anything capable of making one cheerful and time feeling discouraged or dejected, of the mental health dimension. The authors concluded that network analysis is an alternative to classical statistical techniques to analyze data on quality of life in different populations and suggested that the overall structure of SF-36 is dominant in both networks analyzed, supporting the validity of the quiz subscales.

FINAL CONSIDERATIONS

The present study presented methodological concepts of network analysis and discussed their applicability based on specific scientific articles on the subject. Although the

use of network analysis is increasingly common in health science research,^{1,31,32,35} there are still few studies that use this method and, above all, there is a lack of methodological research on networks. Health research, especially in the fields of geriatrics and gerontology, commonly encompasses the study of complex biological, psychological and social conditions that often require advanced statistical tools to be analyzed.

With technological advances, availability of statistical programs with new data analysis resources, and dissemination of theoretical information on network science, new studies should be published. We emphasize that this is the first research on this theme in Brazil, and it was not possible to discuss all the theory that encompasses network analysis. Some concepts about the technique are still evolving, and, therefore, the search for knowledge and updating must be constant on the part of researchers.

For health professionals who do not act as researchers, we encourage the study of this method for a better interpretation of network data. It should be pointed out that this study aimed to disseminate the knowledge of networks to all readers of health sciences, especially those who are professionals

or researchers working in the area of aging. To assist in the study of this statistical technique, with regard to the interpretation of results and methodological aspects in networks, we also present a supplementary material, which demonstrates a population survey with older adults.

The R and RStudio programs served as a theoretical basis for this study, but there are other programs available for researchers with an advanced, intermediate, or basic level in statistics. It is important to point out that, to estimate networks in R or RStudio, it is necessary to understand the statistical programming language, from basic commands, such as selection of variables, to advanced ones, such as the development of networks and performance of resampling techniques or bootstrap statistics. These programs are free, and the commands in statistics are available in articles, which were mentioned in the present study.

SUPPLEMENTARY MATERIAL

Link: https://s3-sa-east-1.amazonaws.com/publisher.gn1.com.br/ggaging.com/pdf/GGA_1900073_IN_material+suplementar.pdf

REFERENCES

1. Luke DA, Harris JK. Network analysis in public health: history, methods, and applications. *Annu Rev Public Health*. 2007;28:69-93. <https://doi.org/10.1146/annurev.publhealth.28.021406.144132>
2. Braga MJC, Gomes LFAM, Ruediger MA. Mundos pequenos, produção acadêmica e grafos de colaboração: um estudo de caso dos Enanpads. *Ver Adm Pública*. 2008;42(1):133-54. <http://dx.doi.org/10.1590/S0034-76122008000100007>
3. Machado W, Vissoci J, Epskamp S. Análise de rede aplicada à Psicometria e à Avaliação Psicológica. In: Hutz CS, Bandeira DR, Trentini CM, editores. *Psicometria*. Porto Alegre: Artmed; 2015. p. 125-46.
4. Prell C. *Social network analysis: History, theory and methodology*. Los Angeles: Sage; 2012.
5. Zheng F, Wei L, Zhao L, Ni F. Pathway network analysis of complex diseases based on multiple biological networks. *BioMed Res Int*. 2018;2018:5670210. <https://doi.org/10.1155/2018/5670210>
6. David HMSL, Faria MGA, Dias JAA, Silva TF, Souza VMD, Dias RS. Social network analysis in primary health care: an integrative review. *Acta Paul Enferm*. 2018;31(1):108-15. <http://dx.doi.org/10.1590/1982-0194201800016>
7. Jin J, Zhou S, Xu Q, An J. Identification of risk factors in epidemiologic study based on ROC curve and network. *Sci Rep*. 2017;7:46655. <https://doi.org/10.1038/srep46655>
8. Pearl J. *Causality: models, reasoning and inference*. Cambridge: Cambridge University Press; 2000. v. 29.
9. Epskamp S, Maris GK, Waldorp LJ, Borsboom D. Network psychometrics. *arXiv preprint arXiv:160902818*. 2016.
10. Epskamp S, Waldorp LJ, Möttus R, Borsboom D. The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behav Res*. 2018;53(4):453-80. <https://doi.org/10.1080/00273171.2018.1454823>
11. Jones PJ, Mair P, McNally R. Visualizing psychological networks: a tutorial in R. *Front Psychol*. 2018;9:1742. <https://dx.doi.org/10.3389%2Fpsyg.2018.01742>
12. Neves ASBM. Evidências de Validade da PSS-10 e PSS-14: Estudo com Análise Fatorial e de Rede [dissertação]. Campinas: Pontifícia Universidade Católica de Campinas; 2018.
13. Epskamp S, Rhemtulla M, Borsboom D. Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*. 2017;82:904-927. <http://doi.org/10.1007/s11336-017-9557-x>
14. Hansen D, Shneiderman B, Smith MA. *Analyzing social media networks with NodeXL: Insights from a connected world*. Oxfordshire: Morgan Kaufmann; 2010.
15. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*. 2018;50:195-212. <https://doi.org/10.3758/s13428-017-0862-1>
16. Haslbeck J, Waldorp LJ. Structure estimation for mixed graphical models in high-dimensional data. *arXiv preprint arXiv:151005677*. 2015.
17. Chen S, Witten DM, Shojaie A. Selection and estimation for mixed graphical models. *Biometrika*. 2015;102(1):47-64. <https://dx.doi.org/10.1093/biomet/asu051>
18. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Statist Soc B (Methodological)*. 1996;58(1):267-288.
19. Van Borkulo CD, Borsboom D, Epskamp S, Blanken TF, Boschloo L, Schoevers RA, et al. A new method for constructing networks from binary data. *Sci Rep*. 2014;4:5918. <https://doi.org/10.1038/srep05918>
20. Newman M. *Networks*. Oxford: Oxford University Press; 2018.
21. Salter-Townshend M, White A, Gollini I, Murphy TB. Review of statistical network analysis: models, algorithms, and software. *Stat Anal Data Min*. 2012;5(4):243-64. <https://doi.org/10.1002/sam.11146>
22. Epskamp S, Cramer AO, Waldorp LJ, Schmittmann VD, Borsboom D. qgraph: Network visualizations of relationships in psychometric data. *J Stat Softw*. 2012;48(4):1-18. <https://doi.org/10.18637/jss.v048.i04>
23. Epskamp S. A preview of new features in bootnet 1.1 [Internet]. 2018 [acessado em 17 set. 2019]. Disponível em: http://psychosystems.org/bootnet_1.1

24. Kalisch M, Hauser A, Maechler M, Colombo D, Entner D, Hoyer P, et al. Package 'pcalg'. 2019.
25. Wagenmakers E-J, Love J, Marsman M, Jamil T, Ly A, Verhagen J, et al. Bayesian inference for psychology. Part II: Example applications with JASP. *Psychon Bull Rev*. 2018;25(1):58-76. <https://doi.org/10.3758/s13423-017-1323-7>
26. Love J, Selker R, Verhagen J, Marsman M, Gronau QF, Jamil T, et al. Software to sharpen your stats. *APS Observer*. 2015;28(3):27-9.
27. Himelboim I, Smith MA. NodeXL. *Int Encyclopedia Communication Res Methods*. 2017:1-3. <https://doi.org/10.1002/9781118901731.iecrm0167>
28. Barabasi A-L. The network takeover. *Nature Physics*. 2011;8(1):14-6. <https://doi.org/10.1038/nphys2188>
29. Mazzocchi F. Complexity in biology. *EMBO Rep*. 2008;9(1):10-4. <https://dx.doi.org/10.1038/sj.embor.7401147>
30. Goh K-I, Cusick ME, Valle D, Childs B, Vidal M, Barabási A-L. The human disease network. *Proc Natl Acad Sci U S A*. 2007;104(21):8685-90. <https://doi.org/10.1073/pnas.0701361104>
31. Nair J, Ghatge M, Kakkar VV, Shanker J. Network analysis of inflammatory genes and their transcriptional regulators in coronary artery disease. *PLoS One*. 2014;9(4):e94328. <https://doi.org/10.1371/journal.pone.0094328>
32. Kalgotra P, Sharda R, Croff JM. Examining health disparities by gender: a multimorbidity network analysis of electronic medical record. *Int J Med Inform*. 2017;108:22-8. <https://doi.org/10.1016/j.ijmedinf.2017.09.014>
33. Borsboom D. A network theory of mental disorders. *World Psychiatry*. 2017;16(1):5-13. <https://doi.org/10.1002/wps.20375>
34. van Wanrooij LL, Borsboom D, van Charante EPM, Richard E, van Gool WA. A network approach on the relation between apathy and depression symptoms with dementia and functional disability. *Int Psychogeriatr*. 2019:1655-63. <https://doi.org/10.1017/S1041610218002387>
35. Solmi M, Konayagi A, Thompson T, Fornaro M, Correll CU, Veronese N. Network analysis of the relationship between depressive symptoms, demographics, nutrition, quality of life and medical condition factors in the Osteoarthritis Initiative database cohort of elderly North-American adults with or at risk for osteoarthritis. *Epidemiol Psychiatr Sci*. 2019;29:e14. <https://doi.org/10.1017/S204579601800077X>
36. Kossakowski JJ, Epskamp S, Kieffer JM, van Borkulo CD, Rhemtulla M, Borsboom D. The application of a network approach to Health-Related Quality of Life (HRQoL): introducing a new method for assessing HRQoL in healthy adults and cancer patients. *Qual Life Res*. 2016;25(4):781-92. <https://doi.org/10.1007/s11136-015-1127-z>