Atherosclerosis 358 (2022) 55–56

Contents lists available at ScienceDirect

Atherosclerosis

journal homepage: www.elsevier.com/locate/atherosclerosis

Editorial

What electronic health records can and cannot tell us in the era of *big data*

ARTICLE INFO

Keywords
Big data
Electronic health record
Epidemiology
Diabetes
CKD

The increasing availability of data drawn from diverse sources, along with new communication technologies, is influencing the way several social and economic sectors operate, including the healthcare system [1]. Information potentially useful in defining health needs, disease trajectories, or enabling epidemiological assessments is nowadays drawn from administrative as well as medical, pharmaceutical or ancillary claims [1]. In this issue of the journal, Wang and colleagues published an interesting analysis derived from a query of the National Health Insurance registry in Taiwan. They aimed to define the incidence and prevalence of frailty among subjects with diabetes mellitus (DM) and to assess the impact of various comorbidities on the risk of developing frailty [2]. Although the registry has an administrative purpose, it can also be used to define the conceptual framework of a health problem. Frailty is associated with frequent hospitalizations and adverse outcomes and a significant social and economic burden [3,4]. An accurate prediction of which patient may develop malnutrition, sarcopenia, gait instability, decline in exercise or cardiopulmonary tolerance could have repercussions on the allocation of resources that could be devoted to the personalized management of a patient's condition. The main finding of the paper is that chronic kidney disease (CKD) presages a significant risk of frailty irrespective of the time of its development (before or after the diagnosis of diabetes mellitus) [2].

Although the focus of the paper is on the relative contribution of CKD on the risk of developing frailty, chronic obstructive pulmonary disease (COPD), liver and cardiovascular diseases (CVD) also exhibited a significant association with the risk of frailty in patients with diabetes, suggesting that multimorbidity rather than a single ailment may be a better predictor of the risk of developing frailty [2]. According to the Academy of Medical Sciences multiple long-term health conditions (multimorbidity) are defined as the coexistence of two or more long-term physical non-communicable (for example CKD, DM, cardiovascular disease, chronic obstructive pulmonary disease, etc.) or mental health conditions (for example dementia) [5]. The prevalence of multimorbidity increases with age and it is estimated that about one in two (54%) individuals over 65 years of age has a multimorbid condition [6,7]. However, defining the disease trajectory of each subject represents a major challenge for health care professionals. Insurance databases and health related claims, as a source of information, may be of aid in solving the puzzle. Large databases may allow fairly precise estimates of the extent of a problem of interest. For instance, contrary to what is commonly believed [8], Wang et al. [2] described a prevalence of CKD of one in six patients diagnosed with DM, instead of one in three as previously reported [8].

The use of a large and unselected source of information such as the National Health Insurance registry of Taiwan, as opposed to highly selected populations recruited in clinical trials, likely justifies this discrepancy. However, although health record databases (HRD) are large, they cannot overcome the limitations of observational studies and cannot establish a cause-and-effect link that may allow us to understand whether DM or CKD (or other comorbidities) are the key determinants of frailty risk. Additionally, HRD do not possess all the qualities that "big data" are believed to have (the 5-V model defines big data: volume, velocity, variety, veracity, value) [1]. In fact, the lack of variety and velocity of data acquisition and analysis limits the ability of HRD to answer the research question Wang et al. [2] posed of predicting which subject with DM will develop frailty, while predicting any changes in the subject's disease trajectory. Far more importantly, the mono-dimensionality of administrative datasets does not allow the investigations of important aspects such as social and psychological status that may influence the trajectory of disease, the clinical outcome and the use of economic resources [9,10]. Indeed, although Wang et al. analyzed data of more than 800,000 patients [2], integration of other sources of data would be required for a multidimensional analysis and allow for an holistic approach to determine and accurately predict patients' needs

DOI of original article: https://doi.org/10.1016/j.atherosclerosis.2022.06.1010.

https://doi.org/10.1016/j.atherosclerosis.2022.08.002
Received 25 July 2022; Accepted 3 August 2022
Available online 18 August 2022
0021-9150/© 2022 Elsevier B.V. All rights reserved.

[1].

While Wang et al. [2] should be commended for their effort, a standard approach to data analysis does not allow for a shift from a view based on the presence/absence of a certain disease to an approach based on defining the risk of groups of individuals [1]. In this regard, big data analytics and artificial intelligence (AI) can support people-centered and integrated health services [1]. Certainly AI represents a frontier of medical research on which great expectations are placed [11]. Machine learning technology, among others, is now widely used for analyzing medical dataset and developing learning algorithms to accurately classify and predict, which patients will be more prone to develop the outcome of interest with reduced subjectivity compared to conventional approaches [11]. Although AI may not overcome the mono-dimensionality of administrative databases, it could provide a view of how the combination of different morbid conditions interact to determine the risk of frailty in DM, and the disease trajectory of different subgroups of patients. Nonetheless, while AI represents the next epidemiology research frontier, a number of limitations and barriers must be overcome before the full potential of this approach can be benefited from. For instance, a noticeable obstacle to the full implementation of AI is the ability to link and integrate different data sources while respecting the privacy of the individual [1].

In summary, the study by Wang et al. [2] has the undoubted merit of having used an administrative database that is large and, likely, representative of the population of Taiwan. It therefore represents a step forward for epidemiological and medical research and a call for the use of various and diverse data sources to develop people-centered integrated health services.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Timo Schulte, Sabine Bohnet-Joschko, How can big data analytics support people-centred and integrated health services, A Scoping ReviewInt J Integr Care 22 (2) (2022 Jun 16) 23, 10.5334/ijic.5543. eCollection 2022 Apr-Jun.

[2] Lee Wang, Chao, et al., The frailty risk trajectory associated with kidney and cardiovascular morbidities among patients with incident diabetes: a population-based study, Atherosclerosis S0021–9150 (22) (2022 Jun 14), 01297-7.

[3] J. Walston, E.C. Hadley, L. Ferrucci, et al., Research agenda for frailty in older adults: toward a better understanding of physiology and etiology: summary from the American geriatrics society/national institute on aging research conference on frailty in older adults, J. Am. Geriatr. Soc. 54 (2006) 991–1001.

[4] A. Ekram, R.L. Woods, C. Britt, S. Espinoza, M.E. Ernst, J. Ryan, The association between frailty and all-cause mortality in community-dwelling older individuals: an umbrella review, J. Frailty Aging 10 (2021) 320–326.

[5] The Academy of Medical Sciences, Multimorbidity: a priority for global health research. The Academy of Medical Sciences, URL: https://acmedsci.ac.uk/file-download/99630838. (Accessed 7 September 2021).

[6] C. Violan, Q. Foguet-Boreu, G. Flores-Mateo, C. Salisbury, J. Blom, M. Freitag, et al., Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies, PLoS One 9 (7) (2014 Jul), e102149, https://doi.org/10.1371/journal.pone.0102149 [FREE Full text] [Medline: 25048354].

[7] A. Cassell, D. Edwards, A. Harshfield, K. Rhodes, J. Brimicombe, R. Payne, et al., The epidemiology of multimorbidity in primary care: a retrospective cohort study, Br. J. Gen. Pract. 68 (669) (2018 Apr) e245–e251, https://doi.org/10.3399/bjgp18X695465 [FREE Full text] Medline: 29530918.

[8] H. Kramer, Screening for kidney disease in adults with diabetes and prediabetes, Curr. Opin. Nephrol. Hypertens. 14 (2005) 249–252.

[9] A.E.M. Liljas, F. Brattström, B. Burström, P. Schön, J. Agerholm, Impact of integrated care on patient-related outcomes among older people - a systematic review, Int. J. Integrated Care 19 (3) (2019 Jul) 6, https://doi.org/10.5334/ijic.4632 [FREE Full text] [Medline: 31367205].

[10] S. Baxter, M. Johnson, D. Chambers, A. Sutton, E. Goyder, A. Booth, The effects of integrated care: a systematic review of UK and international evidence, BMC Health Serv. Res. 18 (1) (2018 May) 350, https://doi.org/10.1186/s12913-018-3161-3 [FREE Full text] [Medline: 29747651].

[11] Hamdan O. Alanazi, Abdul Hanan Abdullah, Kashif Naseer Qureshi, A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care, J. Med. Syst. 41 (4) (2017 Apr) 69, https://doi.org/10.1007/s10916-017-0715-6. Epub 2017 Mar 11.

Antonio Bellasi*
Department of Medicine, Division of Nephrology, Ente Ospedaliero Cantonale, 6900, Lugano, Switzerland

Paolo Raggi
Division of Cardiology, Department of Medicine, University of Alberta, Edmonton, AB, Canada

* Corresponding author. Department of Medicine, Division of Nephrology, Ente Ospedaliero Cantonale, 6978, Lugano, Switzerland.
E-mail address: antonio.bellasi@eoc.ch (A. Bellasi).