Multiple health conditions affect a person from birth to death. The sequential order and the frequency of diseases during the lifecourse of an individual and of entire populations vary because of differences in the environment, society, lifestyle, genetics, and interactions between these factors. Understanding the chronological order of disease occurrence across different subgroups is important to deliver precise intervention at the individual level, to improve the estimate of medical-resource allocation, and to help develop health policies for populations.

An Article by Valerie Kuan and colleagues published in this issue of *The Lancet Digital Health* describes the analysis of electronic health records (EHRs) of nearly 4 million people in the UK National Health Service between 2010 and 2015. Kuan and colleagues developed the first chronological map of health conditions during the human lifecourse to help health-care professionals and researchers understand which parts of the population are susceptible to which health conditions and at which ages.

Previous studies have investigated the global distribution and burden of diseases. These studies focused on the accumulated health challenges across time, countries, and gender rather than the trajectory of diseases at each life stage. Large-scale investigations of the disease spectrum across the lifecourse at the individual level were few, largely because of the tremendous effort needed to longitudinally monitor the participants and precisely document adverse events.

EHRs, encompassing a wide range of data sources including data at the point of care, provide an opportunity to draw a comprehensive and dynamic picture of diseases through the passage of life. Compared with traditional research settings, such as randomised clinical trials and research-driven observational studies, EHR data have the advantages of representing an unselected population, of offering optimal generalisability (especially for specific populations that are usually excluded from clinical trials, such as older participants or patients with kidney-function impairment), and of being able to use very large sample sizes by virtue of the data collection method.

However, the quality-control process in the generation and analysis of EHRs—which might lack the rigour of clinical studies—could be insufficient, a limitation that should be considered when interpreting results based on such data.

Kuan and colleagues developed phenotyping algorithms and codelists for more than 300 diseases and estimated the cumulative incidence, the period prevalence, and the median age at diagnosis for these diseases. They ranked and mapped the 50 most common diseases in each decade of life and provided enhanced visualisation of the data to allow non-specialists to identify patterns within the data. Kuan and colleagues found major types of diseases varied during an individual’s lifecourse, largely evolving from atopic conditions and infections to chronic diseases, which is consistent with previous studies from various age groups or specialties. In addition, they observed variations of the spectrum of diseases according to gender and ethnicity.

Strengths of this study include a large sample size of the general population and harmonisation of different coding systems across several electronic health-record sources by manual curation. Furthermore, the visualisation of results not only enables researchers to quickly identify trends and relationships, but also provides policy makers and other stakeholders with a quick, clear understanding of the information generated from multiple complex statistical estimates. The results from Kuan and colleagues’ study could help to quantitively estimate the medical impacts of aging, and provide evidence to help shape medical education at national levels. More precise interventions to target specific diseases at the population level could be made on the basis of Kuan and colleagues’ results. For instance, they revealed that nutritional anaemia developed earlier in women than in men; hence, relevant monitoring and intervention should be prioritised for women. Their data could indicate that interventions targeting unhealthy lifestyles such as physical inactivity and high salt intake should be initiated before the typical age of the onset of cardiovascular diseases, which is around 40 years according to Kuan and colleagues’ results. Additionally,
the chronological map presented by Kuan and colleagues could be used as prior probability by physicians and employed as a complement to other medical guidelines when evaluating the possible diagnoses for a specific patient. Finally, the phenotyping algorithms and codelists generated by the authors have been made available through an open-source data repository, so they can be reused by other researchers potentially facilitating the definition of diseases from various data sources, a process that can be a bottleneck when using EHRs because of the lack of standards.

Dr Kuan and colleagues’ study was based on EHR data in England, so their results might not be generalisable to other countries (especially those without a centralised health-care system). Furthermore, the combination of data in death registries, Clinical Practice Research Datalink data, and Hospital Episode Statistics Admitted Patient Care data, could be used to investigate the indices related to death, including mortality and years of life lost because of premature death, which are important markers of disease burden that are not included in Kuan and colleagues’ study. Secular trends of common health conditions and their determinants could also be investigated after long-term data accumulation. By undertaking such an investigation, researchers could extend the question of who gets which conditions when and why.

**Ruogu Meng, *Luxia Zhang**

National Institute of Health Data Science and Center for Data Science in Health and Medicine, Peking University, Beijing 100191, China (RM, LZ); and Renal Division, Department of Medicine, Peking University First Hospital, Peking University Institute of Nephrology, Beijing, China (LZ).

zhanglx@bjmu.edu.cn

This study was supported by Grants from the National Natural Science Foundation of China (81771939, 91846101, B301296), Peking University (BMU2018RMX020, PKU2017LC0X5), the National Key Technology R&D Program of the Ministry of Science and Technology of the People’s Republic of China (2016YFC1305400), and the University of Michigan Health System–Peking University Health Science Center Joint Institute for Translational and Clinical Research (BMU20160466, BMU2018J012, BMU2019J005). We declare no competing interests.

Copyright © 2019 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

1 Kuan V, Denaxas S, Gonzalez-Izquierdo A, et al. A chronological map of 308 physical and mental health conditions from 4 million individuals in the National Health Service. *Lancet Digital Health* 2019; 1: e63–77.

2 GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018; 392: 1789–858.

3 Barnett K, Mercer S, Norbury M, Watt G, Wyke S, Guthrie B. The epidemiology of multimorbidity in a large cross-sectional dataset: implications for health care, research and medical education. *Lancet* 2012; 380: 37–43.

4 Network for Excellence in Health Innovation. Real world evidence: a new era for health care innovation. 2015. http://www.nehi.net/writable/publication_files/file/rwe_issue_brief_final.pdf (accessed May 8, 2019).

5 Cohen AT, Schreiber K, Goto S, Torp-Pedersen C. Why do we need observational studies of everyday patients in the real-life setting? *Eur Heart J Suppl* 2015; 17: D2–8.

6 Booth CM, Tannock IF. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. *Br J Cancer* 2016; 115: 551–55.

7 Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-world evidence—what is it and what can it tell us? *N Engl J Med* 2016; 375: 2293–97.