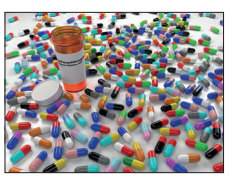
Interpreting big-data analysis of retrospective observational data

Huizinga, T.W.J.; Knevel, R.

8 US Office of Disease Prevention and Health Promotion. Social determinants of health. Aug 18, 2020. https://www.healthypeople.gov/2020/topics-objectives/topic/social-determinants-of-health (accessed Sept 1, 2020).
9 Cho K. Introduction to VA Phenomics Library. May 12, 2020. https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/video_archive.cfm?SessionID=3814 (accessed Sept 1, 2020).
10 US Department of Veterans Affairs. HSR&D Collaborates to Develop VA COVID-19 Social Risks Screening Questions. May 6, 2020. https://www.hsrd.research.va.gov/news/research_news/hausmann-050620.cfm (accessed Sept 1, 2020).

Interpreting big-data analysis of retrospective observational data


Shutterstock

Published Online
August 21, 2020
https://doi.org/10.1016/S2665-9913(20)30289-7

See Articles page e698

Current data technology offers fantastic new opportunities to generate data that can inform us about the safety of drugs. These data will affect the way we use drugs by balancing benefits of specific agents with better and more information on their associated risks. Nowadays, possibilities to use big data to deal with safety concerns are enormous, and it is difficult not to get enthusiastic reading papers that take this approach. An outstanding example is the use of claims data of 160 000 patients with rheumatoid arthritis to assess the risk for lower-tract gastrointestinal perforation associated with tocilizumab and tofacitinib in comparison with other biological drugs.[1]

In *The Lancet Rheumatology*, Jennifer Lane and colleagues present a study using claims data and electronic medical records (mostly of patients with rheumatoid arthritis) to analyse the long-term risks of cardiovascular complications (among other outcomes) in about 1 000 000 users of hydroxychloroquine compared with more than 300 000 users of sulfasalazine.[2] This analysis is relevant because the European League Against Rheumatism (EULAR) guidelines for the treatment of patients with systemic lupus erythematosus (SLE) recommend hydroxychloroquine for all patients with SLE, and in practice the drug is given for decades.[3] Although Lane and colleagues identified no excess risk with short-term use (30 days) of hydroxychloroquine alone, long-term use was associated with increased cardiovascular mortality (calibrated hazard ratio [HR] 1·65 [95% CI 1·12–2·44]). When combined with azithromycin, hydroxychloroquine use was associated with increased risk of 30-day cardiovascular mortality (calibrated HR 2·19 [95% CI 1·22–3·95]), chest pain or angina (1·15 [1·05–1·26]), and heart failure (1·22 [1·02–1·45]).

Most doctors will feel that a study as large as that of Lane and colleagues is most likely relevant, and they will try to weigh the information presented to optimise treatment strategies for their patients. It has been convincingly shown that most published data are false,[4] and the corollary that the hotter a scientific field (with more scientific teams involved), the less likely the research findings are to be true is a relevant consideration given the recent discussions around use of hydroxychloroquine in patients with COVID-19.

So what considerations can be made? Might this be a false positive result? In such a retrospective analysis of observational data, there can of course be confounding by indication. It is important to note that the authors used state-of-the-art methods to deal with the challenges of studying retrospective electronic medical record data; they did a new-user cohort study and a self-controlled case series to avoid the risk of bias in a case-control design, using propensity scores, fitting models with ten-fold cross validation, and negative control outcome analyses. The study thus provides a relevant guide for researchers in the field of electronic medical record analyses. Still, the question remains whether the results should guide our current standards of care.

As the authors state in their discussion, the cohort included patients who were new users of hydroxychloroquine or sulfasalazine with a diagnosis of rheumatoid arthritis, without medication use in the previous 365 days, and with at least 365 days of continuous observation time before the index event. In general, one expects hydroxychloroquine to be used in patients with more comorbidities and, from clinical reasoning alone, there is high potential for differences in the cohorts. This is a limitation of the study, as the authors correctly emphasise. Thoroughly constructed propensity scoring was used to adjust for confounders, but this approach cannot control for all differences and could accidently include intermediary variables.[5]

It is also useful to look at the absolute numbers; the numbers of events for cardiovascular-related mortality was 4·39 per 1000 person-years for patients taking hydroxychloroquine compared with 2·00 per 1000 person-years for patients on sulfasalazine. Given these very low absolute numbers, one needs to consider that if bias between the groups exists, then the differences between 4 per 1000 and 2 per 1000 years of observation might also be caused by bias. Although the self-controlled case series analysis overcome many of these possible biases, the indication for hydroxychloroquine use could still be a confounder. Another unfortunate fact is that normal indicators of causality such as dose–response were missing from the study because of apparent lack of variation in dose of hydroxychloroquine or the inability to obtain data on the association between the duration of hydroxychloroquine use and cardiovascular event rate.

The study by Lane and colleagues also lacks controls to show that the database yields what it should. Maculopathy is a well-known adverse effect of hydroxychloroquine, but the authors were not able to observe an association between hydroxychloroquine use and maculopathy in their databases. This might have been caused by positive control surveillance bias, but the absence of a positive control decreases the convincingness of the data.

Finally, the key question for long-term hydroxychloroquine prescription for patients with SLE is how the benefits balance the risk. The current study did not (and did not intend to) address this question. So although we feel that the study by Lane and colleagues is extremely interesting with regard to methodology, and we foresee the rapid growth of studies linking of electronic health record data and claims data, it is difficult to weigh the current data in the context of daily care of patients with SLE, in which so much convincing evidence exists for the positive effects of hydroxychloroquine as recommended by EULAR.

We declare no competing interests.

*Tom W J Huizinga, Rachel Knevel
t.w.j.huizinga@lumc.nl

Department or Rheumatology, Leiden University Medical Center, Leiden, Netherlands (TWJH, RK), Brigham And Women's Hospital, Department of Medicine, Division of Rheumatology, Boston, MA, USA (RK)

1 Xie F, Yun H, Bernatsky S, Curtis JR, et al. Risk for gastrointestinal perforation among rheumatoid arthritis patients receiving tofacitinib, tocilizumab, or other biologics. *Arthritis Rheumatol* 2016; **68:** 2612–17.
2 Lane JCE, Weaver J, Kostka K, et al. Risk of hydroxychloroquine alone and in combination with azithromycin in the treatment of rheumatoid arthritis: a multinational, retrospective study. *Lancet Rheumatol* 2020; published online Aug 21. https://doi.org/10.1016/S2665-9913(20)30276-9.
3 Fanouriakis A Kostopoulou M, Alunno A, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. *Ann Rheum Dis* 2019; **78:** 736–45.
4 Ioannidis JPA. Why most published research findings are false. *PLoS Med* 2005; **2:** e124.
5 Månsson R, Joffe MM, Sun W, Hennessy S. On the estimation and use of propensity scores in case-control and case-cohort studies. *Am J Epidemiol* 2007; **166:** 332–39.

Accumulating evidence suggests anti-TNF therapy needs to be given trial priority in COVID-19 treatment

Published **Online** September 4, 2020 https://doi.org/10.1016/S2665-9913(20)30309-X

The COVID-19 pandemic continues to wreak havoc on global health-care systems and to claim an increasing number of lives. Although some treatments have shown promise, including dexamethasone and remdesivir, problems remain with access to medication and high mortality despite treatment. Patient selection also appears to be critical, with some patient groups benefitting from treatment, but not others. One potential treatment that deserves higher priority in COVID-19 trials, based on the documented evidence of its effects, is the biological agent anti-TNF.

Feldmann and colleagues[1] described the rationale for trialling anti-TNF therapies in COVID-19. These therapies neutralise TNF, a major component of the cytokine response that is part of the damaging excess inflammatory phase of COVID-19, which is termed hyperinflammation or cytokine release syndrome. This hyperinflammatory response in COVID-19 is characterised by elevated concentrations of serum TNF, interleukin (IL)-6, and IL-8, but relatively little IL-1.[2] However, IL-1 has a short serum half-life, and mononuclear transcriptome data show that genes and pathways upregulated by TNF, IL-1β, and type I interferon predominate.[3] A major component of deteriorating lung function in patients with COVID-19 is capillary leak, a result of inflammation driven by key inflammatory cytokines: TNF, IL-1, IL-6, and vascular