Check for updates

Editorial

COVID-19 and Statistical Challenges

Soleiman Kheiri[1*] ID

[1]Modeling in Health Research Center, Shahrekord University of Medical Sciences, Shahrekord, Iran

*****Corresponding Author:** Soleiman Kheiri, Tel: 03833346712, Email: kheiri@skums.ac.ir

Since the outbreak of COVID-19 in December 2019, there has been an explosion of statistics and information about the disease, the number of cases, the number of deaths, and the number of recoveries. During this period several statistical and mathematical models have been developed and used to predict the disease. Much of this information has been helpful and paved the way for disease control; however, inaccurate or ambiguous information has been published in some cases, which can briefly be divided into three main categories.

The first category is related to the publication of official statistics by governmental centers in countries, which has faced many errors. Although some of these errors are unintentional due to the definition of the disease based on definitive polymerase chain reaction (PCR) testing, death within 28 days of infection, or due to the similarity of the disease outcomes with other diseases,[1] in many cases, the statistics regarding the disease and its consequences have been presented by governments with a manipulation, mainly on the small number of patients with COVID-19.[2] This issue caused a great deal of controversy among academic centers, and many of them attempted to explore the differences between the official statistics provided by countries and other types of data based on other sources. Fewer reports on the actual number of positive cases of the disease may lead to false optimism and thus negligence in dealing with COVID-19 and an increase in the number of patients.

The second category contributes to incorrect or incomplete use of statistical indexes. During this period, some information has occasionally been published by some public media in which statistical indexes have not been used correctly and appropriately. For example, mentioning the percentage of increase in positive cases over a specified period of time without mentioning the base number or the percentage of deaths without mentioning the number, can provide incorrect information to the audience. The presentation of incomplete statistics and information is considered among the examples of lying with statistics, indicating that academic researchers need to tackle and challenge this disagreeable phenomenon.[3]

The third category belongs to the challenges of epidemiological modeling in COVID-19. With the advent of COVID-19, many models emerged to predict its incidence and consequences. Although many of the predictions were true, some of them were incorrect or inaccurate. The main reasons for the inaccuracy of these predictions were the consideration of incorrect or weak assumptions, the existence of incomplete data, the short-term view, the use of point estimates instead of interval estimates, and the lack of a multidimensional view of the problem. Accordingly, considering the above-mentioned points and findings from the observed problems in predicting the course of the disease and the resulting mortality during the COVID-19 pandemic, one can hope for the ability to model the prediction of similar diseases in the future.[4]

Conflict of Interest Disclosures
None declared.

Ethical Approval
Not applicable.

References
1. Oliver D. David Oliver: mistruths and misunderstandings about COVID-19 death numbers. BMJ. 2021;372:n352. doi: 10.1136/bmj.n352.
2. Adiguzel FS, Cansunar A, Corekcioglu G. Truth or dare? detecting systematic manipulation of COVID-19 statistics. Journal of Political Institutions and Political Economy. 2020;1(4):543-57. doi: 10.1561/113.00000021.
3. Huff D. How to Lie with Statistics. WW Norton & Company; 1993.
4. Ioannidis JPA, Cripps S, Tanner MA. Forecasting for COVID-19 has failed. Int J Forecast. 2020. doi: 10.1016/j.ijforecast.2020.08.004.