Review of: "A Multi-factor Model of COVID-19 Epidemic in California"

Matthew Merski[1]

1 University of Warsaw

Potential competing interests: No potential competing interests to declare.

In the paper “A Multi-factor Model of COVID-19 Epidemic in California” by Lewis, the author examines the cumulative case count of COVID-19 cases for California. The author looks at the cumulative case count and case duration for 58 counties in California, While there is a good correlation between county population and the number of cases (correlation coeff. = 0.99), the author did not find good correlations for the other factors examined (correlation coeff. < 0.8). This was, however, a laudable attempt to relate the role of socioeconomic factors with COVID spread and outcomes and is worth following up on. However, the small number of references suggests the author is somewhat unfamiliar with other work that has been done in predicting COVID-19 case numbers. This also becomes a problem in that some of the terminology which has been defined in the field of epidemiology is used inaccurately.

Minor point: The cases and duration fit equations have significant typographical errors, switching between decimal point and comma notation and the duration equation seems to lack a variable at all.

Major point: The author claims that: “We provide a novel prediction model based on correlation coefficients”. However, no actual predictions are done here. The data at one point in time (June 2023) are fit to a model equation. Predictions would be fitting the data at one or more points in time and then predicting the result (either prospectively or retrospectively) and then determining how accurate those predictions were. For example, how accurately would a prediction made from data in October 2021 be at estimating the total number of cumulative cases in California in July 2023? An actual analysis in this vein is needed if any assertions of accuracy are to be made.