Cover Letter: How Not to Forecast the Flu

Prithwish Chakraborty1,2\* , Bryan Lewis3, Stephen Eubank3, John S. Brownstein4,5,

Madhav Marathe2,3, and Naren Ramakrishnan1,2

1Discovery Analytics Center, Virginia Tech, VA, USA

2Dept. of Computer Science, Virginia Tech, VA, USA

3Network Dynamics and Simulation Science Laboratory, Biocomplexity Institute, Virginia Tech, VA, USA

4 Children’s Hospital Informatics Program, Boston Children’s Hospital, MA, USA

5Dept. of Pediatrics, Harvard Medical School, MA, USA

\* prithwi@vt.edu

Dear Editor,

In recent years, flu forecasting competitions have attracted a lot of interest in the scientific community. We were part of two such efforts (organized by CDC and by IARPA) and were declared the winning team in the IARPA forecasting tournament. As a participant in these competitions, our team has come to appreciate several key facets of flu forecasting that pose unique challenges and opportunities for computational, statistical, and modeling researchers. Exiting research often makes simplifying assumptions that are easily violated. For example, most studies on flu forecasting are retrospective/or do not instabilities associated with flu surveillance – rendering such models impractical for a real-time forecasting system.

In this paper, we have gathered together some of our "lessons learnt" in the form of a "what not to do" manifesto and presented evidences to support our lessons using surveillance data from several Latin American countries (where surveillance is more unstable) and the United States (where surveillance is relative more stable). We hope that these insights help communicate the underlying complexity of the problem to a broader audience and will help lower the bar for new flu modelers. Relevant prior work from our group and others' are cited in the article and provided in the supplementary information.

We look forward to your review.

Sincerely,

The authors