What Can Digital Disease Detection Learn from (an External Revision to) Google Flu Trends?

Mauricio Santillana, PhD, MS, D. Wendong Zhang, MA, Benjamin M. Althouse, PhD, ScM, John W. Ayers, PhD, MA

Background: Google Flu Trends (GFT) claimed to generate real-time, valid predictions of population influenza-like illness (ILI) using search queries, heralding acclaim and replication across public health. However, recent studies have questioned the validity of GFT.

Purpose: To propose an alternative methodology that better realizes the potential of GFT, with collateral value for digital disease detection broadly.

Methods: Our alternative method automatically selects specific queries to monitor and autonomously updates the model each week as new information about CDC-reported ILI becomes available, as developed in 2013. Root mean squared errors (RMSEs) and Pearson correlations comparing predicted ILI (proportion of patient visits indicative of ILI) with subsequently observed ILI were used to judge model performance.

Results: During the height of the H1N1 pandemic (August 2 to December 22, 2009) and the 2012–2013 season (September 30, 2012, to April 12, 2013), GFT's predictions had RMSEs of 0.023 and 0.022 (i.e., hypothetically, if GFT predicted 0.061 ILI one week, it is expected to err by 0.023) and correlations of r=0.916 and 0.927. Our alternative method had RMSEs of 0.006 and 0.009, and correlations of r=0.961 and 0.919 for the same periods. Critically, during these important periods, the alternative method yielded more accurate ILI predictions every week, and was typically more accurate during other influenza seasons.

Conclusions: GFT may be inaccurate, but improved methodologic underpinnings can yield accurate predictions. Applying similar methods elsewhere can improve digital disease detection, with broader transparency, improved accuracy, and real-world public health impacts.
(Am J Prev Med 2014;47(3):341–347)

Introduction

The rapid escalation of digital methods is changing public health surveillance.[1–3] By harvesting web data, investigators claim to validly estimate cholera,[4] dengue,[5,6] influenza,[7,8] kidney stones,[9,10] listerosis,[11] methicillin-resistant *Staphylococcus aureus*,[12] mental health,[13] and tobacco control[14] trends, but are they actually valid?

The novelty of digital data has generally remained the central focus in these studies, whereas the methods and disinterested interpretations have been overlooked. Therefore, studies demonstrating modest associations with ground truth outcomes (e.g., R^2=0.15,[14] R^2=0.25,[4] or R^2=0.62[11]) have been presented as accurate, without further model validation. Most notable is Google Flu Trends (GFT),[8] not because it is potentially the most flawed but because it is oft-cited and many subsequent studies modeled their approach after GFT or even used weaker methods.[6,12,15,16]

Concerns about GFT's accuracy came to light via media reports in 2009 when it misrepresented the epidemic curve and required updating that Autumn.[17] Again during 2012–2013, media reports questioned the revised GFT,[18] followed by separate peer-reviewed analyses suggesting GFT was typically inferior to traditional sentinels owing to inaccuracies.[19,20] Most recently, Google again updated their model to improve GFT operation but did not identify their revisions or describe its performance.[21] Many, unfortunately, are unaware of these problems.

From the School of Engineering and Applied Sciences (Santillana, Zhang), Harvard University, Cambridge, Massachusetts; Santa Fe Institute (Althouse), Santa Fe, New Mexico; and the Graduate School of Public Health (Ayers), San Diego State University, San Diego, California

Address correspondence to: John W. Ayers, PhD, MA, 879 Compass Way, San Diego CA 92123. E-mail: ayers.john.w@gmail.com.

0749-3797/$36.00

http://dx.doi.org/10.1016/j.amepre.2014.05.020

The head of the CDC Influenza Surveillance and Outbreak Response Team told *Nature News* that she monitors GFT (and other digital disease detection sentinels) "all the time," likely in the sense that some data are better than no data.[18] Moreover, some investigators are beginning to use GFT as ground truth for epidemiologic studies.[22] However, if GFT (and by extension similar systems for other outcomes) are invalid, should public health officials be paying any attention?

We remain optimistic about the future of GFT and digital disease detection broadly[23–26] because a methodologic problem has a methodologic solution. Herein, a transparent, external evaluation of GFT, as a case study for the scientific status of digital disease detection, is presented. An alternative methodology capable of outperforming GFT is subsequently proposed, with potential application across digital disease detection.

Methods

The methodology behind the original GFT and the 2009 revision (published in 2011) consisted of building a regression for CDC-reported influenza-like illnesses (ILI) with a single explanatory variable. Originally, the single variable was the mean trend for the 45 search terms with the strongest correlation with ILI for September 28, 2003, through March 11, 2007.[8] The revised GFT single variable was the mean trend for the most correlated search terms (approximately 160, the exact number unknown) for September 28, 2003, through September 13, 2009, after removing terms related to influenza complications and general interest in influenza (with the exact terms unknown).[17] Both generated ILI predictions at time (t) using search data from (t) and historic periods, but because CDC-reported data are delayed, these estimates are typically available 1–2 weeks ahead of the CDC's.

The original and updated GFT methodology is problematic for at least three reasons. First, combining multiple queries into a single variable ignores the variability in individual search query tendencies over time and how certain unique queries may be better predictors.[8] Second, the exclusion of search queries in the revised GFT relies on investigator opinion rather than any empirical evidence.[17] Third, the model is static, assuming that queries predicting ILI at time (t) will equally predict ILI at time ($t + x$ years). The language of searches undoubtedly changes over time (e.g., swine flu, H1N1, H1N9) and must be accounted for in any prediction model.

Our alternative approach, inspired by data-assimilation techniques,[27,28] supervised machine learning,[29] and artificial intelligence,[30] expands upon (1) their single explanatory-variable approach, by allowing multiple individual queries to contribute independently to the prediction; (2) their quasi-nonempirical search query selection, by empirically choosing search queries that maximize predictive accuracy in real time; and (3) their use of manual revisions, by dynamically updating how individual queries predict influenza each week to ensure accurate prediction across a changing search and influenza landscape. All improve the transparency of GFT.

These revisions are executed in a multivariable linear model with different coefficients for each specific search query trend. Each query is prescribed a different level of importance based on its coefficient, determined by a LASSO for the best predicting and most parsimonious model.[31] The coefficients change each week based on refitting the model at time (t) to ILI through ($t - 2$), representing the latest available CDC ILI estimates if the system was running in real time. The equation is given by the following simplified notation:

$$\text{logit}[I(t)] = \sum_{i=1}^{n} a_i(t)\text{logit}[Q_i(t)] + \text{e}, \quad (1)$$

where $I(t)$ is the percentage of ILI physician visits, $Q_i(t)$ is the query fraction for term i at time t, $a_i(t)$ is the multiplicative coefficient associated to such term at time t, e is the normally distributed error term, and $\text{logit}(p) = \ln[p/(1-p)]$. The mathematical foundations of this method and other methodological considerations were reported by Zhang.[32]

The criterion is the weekly percentage of confirmed ILI-related physician visits (e.g., fever $>100°$F and cough or sore throat as a percentage of all outpatient healthcare provider visits nationally to $>2{,}900$ reporting clinics) publicly available at (cdc.gov/flu/weekly/). Predictions were made for the entire U.S., following the strategy used in media critiques and Google's revisions.[17,21]

For the explanatory variables, our alternative conceptually relies on the same initial selection strategy used by GFT: time trends for the Google queries most strongly correlated with ILI. Google Correlate (google.org/trend/correlate) returned the weekly z scores of the query fractions spanning January 2004 to May 2013 of the 100 Google terms most correlated with CDC-reported ILI from September 28, 2003, to March 22, 2009, as accessed on October 17, 2013, and still available today. These 100 terms were the fixed inputs in the model, where the query fraction of each term is the total count of a query term in the U.S. aggregated weekly and then scaled by the total count of all queries issued in the same week.[8] This choice simulates the selection process that could occur if this methodology operated in real time and had been implemented since March 22, 2009. Unlike Google, our alternative method did not filter this list of queries based on either the intensity of the correlation or the query content but used all 100 most correlated queries, anticipating the model would then select queries to maximize prediction.

For performance benchmarks, our alternative model was implemented for March 22, 2009, through May 2013, fitting to early trends (since January 1, 2004) as a training period. Pearson correlation coefficients and root mean squared errors (RMSEs) comparing predicted ILI with subsequently observed ILI were used to judge model performance. The latter was added to accurately assess performance when correlation may not (i.e., two trends may have r=1.00 correlation, but differ by a constant factor).

For comparisons to GFT, we relied on the estimates made by GFT as events unfolded. As GFT was updated after the H1N1 season in 2011, the GFT webpage returns those updated results, but the national predictions originally made by GFT were downloaded through December 27, 2009, from Figure 1 in the Google revision.[17] For later periods, GFT predictions were simply downloaded in the summer of 2013 (google.org/flutrends). All analyses were conducted in R, version 2.15.3.

Results

Figure 1 presents GFT's and our alternative model's predictions alongside the subsequently observed ILI trends, where it is readily apparent that the alternative produced more accurate predictions.

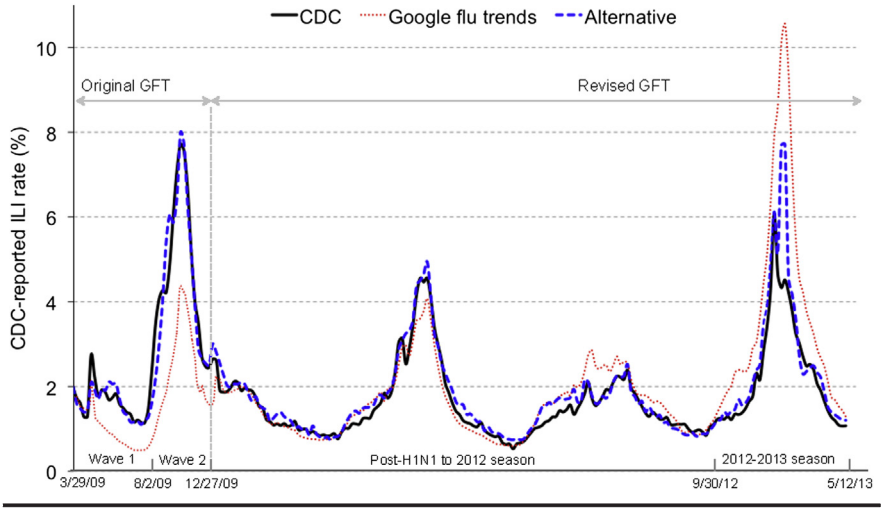

Figure 1. The alternative model outperforms Google Flu Trends

Note: Both GFT (dot) and the proposed alternative model (dash) are shown against the criterion (solid) measure of national CDC-reported ILI (the weekly percentage of confirmed ILI-related physician visits (e.g., fever >100°F and cough or sore throat as a percentage of all outpatient healthcare provider visits nationally to >2,900 reporting clinics). Both GFT and the alternative model generated predictions for ILI at time (*t*) using search data from the same week (*t*) and historic periods, but because ILI is delayed, these estimates were typically available 1–2 weeks earlier than CDC-reported ILI.
GFT, Google Flu Trends; ILI, influenza-like illness

During Wave 1 (March 29 through August 2, 2009) and Wave 2 (August 3 through December 27, 2009) of the H1N1 outbreak, particularly important periods of ILI surveillance, the RMSEs were 0.008 and 0.023 (i.e., if GFT predicted 0.061 ILI, it would be have a usual error of 0.008 or 0.023 each week) with correlations of r=0.290 and r=0.916 for GFT (Table 1). In contrast, our alternative model had RMSEs of 0.002 and 0.006 with correlations of r=0.887 and r=0.961. In practical terms, our alternative model yielded more accurate predictions in 17 of 19 weeks during Wave 1 and every week (21 of 21) during Wave 2. For example, the original GFT predicted an estimated 4.3% peak ILI compared to a 7.7% CDC-reported ILI, an absolute difference of 3.4%, whereas our alternative model predicted an 8.0% peak, an absolute difference of 0.3%.

Our alternative predictions were often better than GFT during other influenza cycles. For December 28, 2009, to September 29, 2012, the alternative predictions had smaller error (RMSE=0.002, r=0.978, vs GFT RMSE=0.003, r=0.912), suggesting a relative 33% reduction in error (i.e., [0.003–0.002]/0.003). Again, in practical terms, the alternative model was more accurate in 126 of 176 weeks (72%), with several periods where the alternative model produced predictions that mirrored ILI when GFT was mis-predicting ILI by ±1%.

During the 2012–2013 season, the GFT's RMSE was 0.022 vs 0.009 for our alternative model, with correlations of r=0.927 and r=0.919. In practical terms, our alternative yielded more accurate predictions every single week. For example, GFT predicted a peak ILI of 10.6% vs 6.1% for CDC-reported ILI, an absolute difference of 4.5%, compared to a 7.7% peak estimate from our alternative method, an absolute difference of 1.6%.

The autonomous and dynamic nature of our alternative appears to be a key component to improved predictions (Figure 2). First, the model automatically

Table 1. Predictive accuracy of the alternative model and Google Flu Trends

	H1N1 (Wave 1) (3/29/09–8/2/09)	H1N1 (Wave 2) (8/2/09–12/27/09)	Post-H1N1 to 2012 season (12/27/09–9/30/12)	2012–2013 season (9/30/12–5/12/13)
Correlation				
Alternative	**0.887**	**0.961**	**0.978**	**0.919**
GFT (original)	0.290	0.916	—	—
GFT (updated)	—	—	0.912	0.927
RMSE				
Alternative	**0.002**	**0.006**	**0.002**	**0.009**
GFT (original)	0.008	0.023	—	—
GFT (updated)	—	—	0.003	0.022

Note: Both GFT and the alternative method generated predictions for ILI at time (*t*) using search data from the same week (*t*) and historic periods, but because ILI is delayed, these estimates were typically available 1–2 weeks earlier than CDC-reported ILI.
GFT, Google Flu Trends; ILI, influenza-like illness; RMSE, root mean squared error

excludes or reduces the predictive influence of many non-influenza terms. For example, "chaos tour" is typically given zero or little weight in the predictions without relying on human opinion as in prior revisions to GFT. Second, there is great variability in the coefficient estimates for each term by each week, where the alternative is automatically updated in response to how terms are indicative of ILI or not from the recent past. For example, terms indicative of influenza concern are given diminishing predictive value over time and those indicative of treatment/complications are given more predictive value (i.e., "is flu contagious?" versus "expectorant").

Most importantly, our alternative does not guarantee accurate predictions, but when predictions do go astray, the alternative learns from its mistake in just 2 weeks, rather than waiting 2 years for a manual revision.[17] At the height of the 2012–2013 season, the alternative mis-predicted the peak ILI proportion by a nontrivial amount, but the model self-corrected, changing the queries included/excluded just 2 weeks later (when the model was first aware of the error because of the delay in CDC-reported ILI). In response to the model error after the aforementioned peak, for example, "flu incubation" queries were negatively weighed to lower the prediction of ILI. Clear shifts in coefficient values for dozens of other queries were observed in this time period as well. These can be seen in the latter portion of Figure 2.

Discussion

Our alternative methodology is capable of producing more accurate predictions of influenza activity than GFT, and does so autonomously with dynamic updating of the model each week. With 3–5 million infected and 250,000–500,000 killed by influenza worldwide each year,[33] influenza surveillance is of tremendous importance, providing necessary intelligence for hospitals facing staffing decisions, physicians facing active and accurate diagnoses, employers with workers at risk for infection, and public health officials making recommendations for protecting unvaccinated individuals. Yet, these results have even greater implications as a case study for digital disease detection broadly.

Implications for the Next Google Flu Trends

In a brief working paper, Google recently described the need to revise GFT.[21] They acknowledged that a multivariable approach (one of the improvements implemented here) would enhance the accuracy of GFT, but that paper also shared many of the weaknesses inherent in the original and first revision to GFT that our alternative may overcome. First, the methods lacked transparency, as the working paper did not identify the model they were implementing.[21]

Second, the predictive validity of the revised GFT remains unknown, because Google.org only included 5 weeks of data in the paper estimating the predictive accuracy. Using this small sample of data, however, our alternative method appears to be a better predictor. Last, their revision still relied on investigator opinion to select/omit some queries and failed to incorporate automatic updating, as we added herein.

Therefore, our alternative method may serve as the foundation for another revision to GFT. Specifically, because much of the alternative is automated, it can be scaled up (e.g., Google could apply it to thousands of strongly correlated search terms instead of just 100 as herein). Our study is just an initial step toward improving GFT, as the structure around our model can be further refined to yield even greater accuracy. Moreover, by making the inner workings of GFT and the data behind GFT more public, such improvements may be more quickly realized by other external teams.[19,20]

Implications for Digital Disease Detection

It is important to appraise how the leading system in the field, GFT, produced questionable predictions, while investigators mimicked the methods behind GFT[6,12,15,16] and few levied significant criticisms,[19,20,34,35] and how this omission can serve as a call for action within the field.

One European study found that GFT predictions before and during the H1N1 pandemic were only crudely associated with influenza (ρ=0.39 and 0.52, respectively), but concluded that GFT provided accurate detection.[36] Analyses from Australia,[37] China,[38] Japan,[39] New Zealand,[11] and the U.S.[40] suggest that GFT predicted larger incidences and missed the timing of outbreaks. Yet, these studies concluded that GFT was valid, with one calling the accuracy of GFT "remarkable,"[37] as also quoted in a Google-led publication.[17] This disparity in results and praise is indicative of a larger problem in digital disease detection.

Many of the current studies are the first of their kind and deserve praise as such, but to move this potentially important field forward, investigators and public health leaders need to exercise caution and become discerning scientific consumers. Claims need to be carefully critiqued, as nearly all of the literature on digital disease detection[4,6,9–14,41] relies on weak methodologic approaches or patterns of association similar to GFT, with rare exception.[5,42–45] Thus, investigators should turn to more sophisticated approaches for the development and evaluation of digital systems. The stakes are high in public health surveillance and the methodologic bar must be raised higher accordingly.

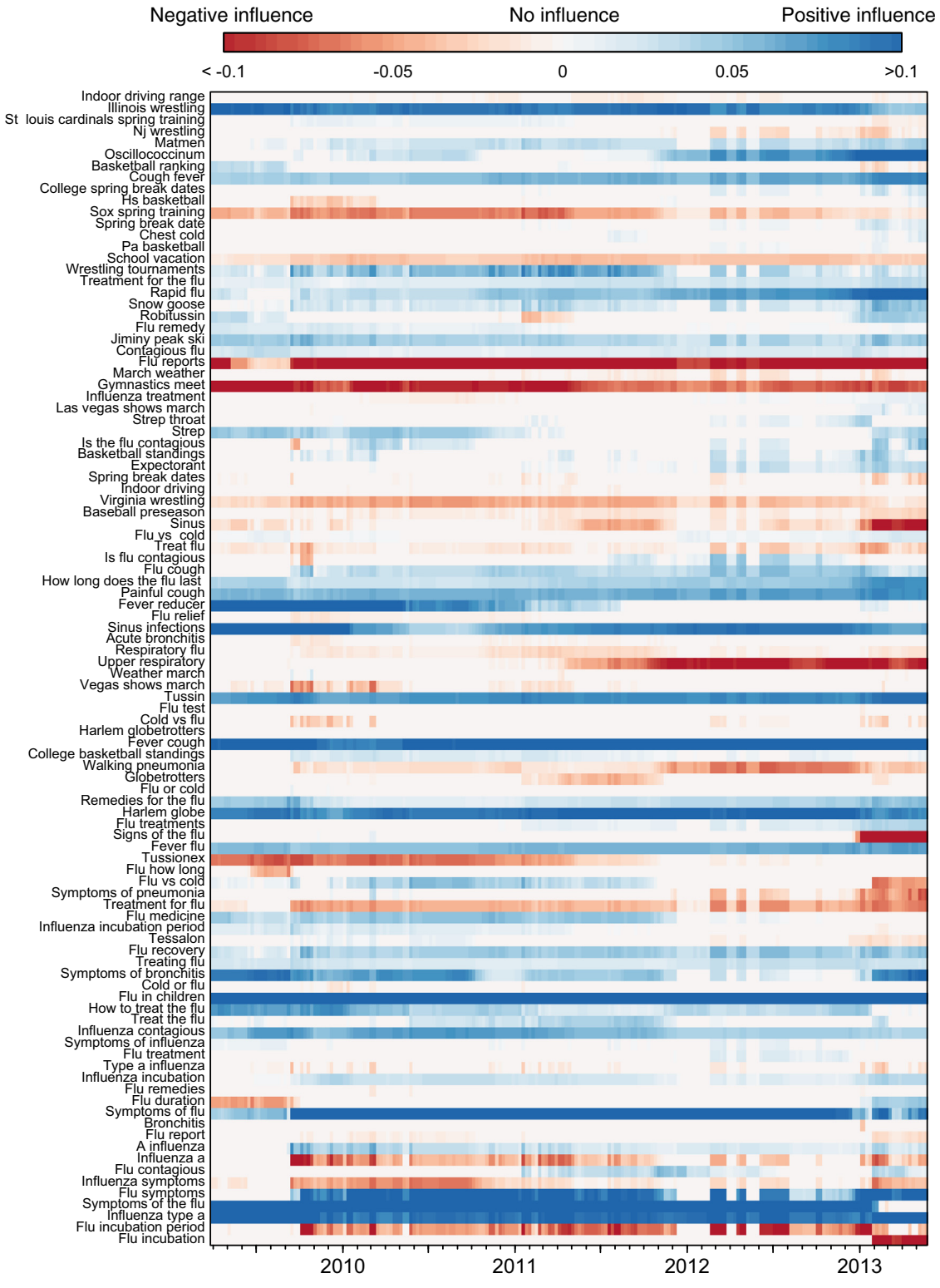

Figure 2. Search term inclusion and dynamic updating over time in the alternative model

Note: Each line is an indicator of the estimated coefficient for a specific search term (*y* axis) updated each week (*x* axis). Colors indicate the relative importance of the search term (positive in blue, negative in red) on prediction over time. For example "flu in children" has a uniformly positive prediction and "St. Louis Cardinals spring training" has a nearly uniform zero effect. This figure shows how the alternative model is autonomously selecting queries and updating the relative value assigned to these queries each week

Yet, we remain optimistic about digital disease detection, because as with all adolescent fields there needs to be periodic methodologic critique and revision. Building on this study and the work of others,[19,20] digital disease detection may better realize its aims.

Limitations

Though advancing the underpinnings of GFT and digital disease detection, the alternative method is not without limitations. As usual, the alternative method was only evaluated at one geographic resolution, and model fits may vary in other geographies, especially those for which there is no ground truth to compare the model predictions against. In addition, CDC-reported ILI estimates (the ground truth) sometimes are later revised as the influenza season progresses,[46] thus introducing uncertainties in the (or any) prediction methodology.

In this regard, the dynamic alternative approach is capable of incorporating information as it becomes available, and thus will automatically produce an improved model every time CDC ILI information is either initially released or revised. Finally, Google queries are but one source of information and a multi-sourced model would be preferred, but through appropriate modeling, searches can be leveraged to achieve accurate predictions that better realize the implications of GFT.[8]

Conclusions

The methods behind digital disease detection are wanting, but by outlining a correction that could improve the accuracy of influenza detection, we hope public health officials find value in digital disease detection and investigators refine our approach to achieve more transparent and accurate surveillance that can have real-world public health impacts.

This work was improved by comments from presentations at Harvard Medical School, the New York City Department of Heath and Mental Hygiene, and Stanford Medical School, with special appreciation for John S. Brownstein, Mark Dredze, John Ioannidis, Donald Olson, Keith Schnakenber, Diana Z. Li, Zhenbu Zhang, and the eight anonymous *American Journal of Preventive Medicine* reviewers. The authors agree to make their data and code available to other investigators wishing to replicate this study.

JWA and BMA share an equity stake in Directing Medicine LLC, a consultancy that helps clinician-scientists implement methods in digital disease detection, some embodied in this work. JWA also holds an equity stake in HealthWatcher Inc., an entity that develops online data processing software, and reports being a paid speaker or advisor to several organizations in the past year. No financial disclosures were reported by the other authors of this paper.

References

1. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection—harnessing the Web for public health surveillance. N Engl J Med 2009;360(21):2153–5, 2157.
2. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med 2011;40(5S2):S154–S158.
3. Ayers JW, Althouse BM, Dredze M. Could behavioral medicine lead the web data revolution? JAMA 2014;311(14):1399–400.
4. Chunara R, Andrews JR, Brownstein JS. Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. Am J Trop Med Hyg 2012;86(1):39–45.
5. Althouse BM, Ng YY, Cummings DA. Prediction of dengue incidence using search query surveillance. PLoS Negl Trop Dis 2011;5(8):e1258.
6. Chan EH, Sahai V, Conrad C, Brownstein JS. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. PLoS Negl Trop Dis 2011;5(5):e1206.
7. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. AMIA Annu Symp Proc 2006;2006:244–8.
8. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009;457(7232):1012–4.
9. Breyer BN, Sen S, Aaronson DS, Stoller ML, Erickson BA, Eisenberg ML. Use of Google Insights for Search to track seasonal and geographic kidney stone incidence in the U.S. Urology 2011;78(2):267–71.
10. Willard SD, Nguyen MM. Internet search trends analysis tools can provide real-time data on kidney stone disease in the U.S. Urology 2011;81(1):37–42.
11. Wilson K, Brownstein JS. Early detection of disease outbreaks using the Internet. CMAJ 2009;180(8):829–31.
12. Dukic VM, David MZ, Lauderdale DS. Internet queries and methicillin-resistant *Staphylococcus aureus* surveillance. Emerg Infect Dis 2011;17(6):1068–70.
13. Yang AC, Tsai SJ, Huang NE, Peng CK. Association of Internet search trends with suicide death in Taipei City, Taiwan, 2004–2009. J Affect Disord 2011;132(1–2):179–84.
14. Cavazos-Rehg PA, Krauss MJ, Spitznagel EL, et al. Monitoring of non-cigarette tobacco use using Google Trends. Tob Control 2014: In press.
15. Yuan Q, Nsoesie EO, Lv B, Peng G, Chunara R, Brownstein JS. Monitoring influenza epidemics in China with search query from Baidu. PLoS One 2013;8(5):e64323.
16. Ocampo AJ, Chunara R, Brownstein JS. Using search queries for malaria surveillance, Thailand. Malar J 2013;12:390.
17. Cook S, Conrad C, Fowlkes AL, Mohebbi MH. Assessing Google flu trends performance in the U.S. during the 2009 influenza virus A (H1N1) pandemic. PLoS One 2011;6(8):e23610.
18. Butler D. When Google got flu wrong. nature.com/news/when-google-got-flu-wrong.
19. Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. PLoS Comput Biol 2013;9(10):e1003256.
20. Lazer D, Kennedy R, King G, Vespignani A. The parable of Google Flu: traps in big data analysis. Science 2014;343(6176):1203–5.
21. Copeland P, Romano R, Zhang T, Hecht G, Zigmond D, Stefansen C. Google disease trends: an update. Menlo Park CA: Google.org, 2013. patrickcopeland.org/papers/isntd.pdf.
22. Nsoesie EO, Buckeridge DL, Brownstein JS. Guess who's not coming to dinner? Evaluating online restaurant reservations for disease surveillance. J Med Internet Res 2014;16(1):e22.
23. Ayers JW, Ribisl KM, Brownstein JS. Tracking the rise in popularity of electronic nicotine delivery systems (electronic cigarettes) using search query surveillance. Am J Prev Med 2011;40(4):448–53.

24. Ayers JW, Althouse BM, Allem JP, et al. Novel surveillance of psychological distress during the great recession. J Affect Disord 2012;142:323–30.
25. Ayers JW, Althouse BM, Noar SM, Cohen JE. Do celebrity cancer diagnoses promote primary cancer prevention? Prev Med 2014;58:81–4.
26. Ayers JW, Althouse BM, Johnson M, Cohen JE. Circaseptan (weekly) rhythms in smoking cessation considerations. JAMA Intern Med 2014;174(1):146–8.
27. Ghil M, Malanotte-Rizzoli P. Data assimilation in meteorology and oceanography. Adv Geophys 1991;33:141–266.
28. Wang B, Zou X, Zhu J. Data assimilation and its applications. Proc Natl Acad Sci U S A 2000;97(21):11143–4.
29. Hastie T, Tibshirani R, Friedman J, Franklin J. The elements of statistical learning: data mining, inference and prediction. New York: Springer, 2001.
30. Russell SJ, Norvig P, Canny JF, Malik JM, Edwards DD. Artificial intelligence: a modern approach. Englewood Cliffs NJ: Prentice Hall, 1995.
31. Tibshirani R. Regression shrinkage and selection via the lasso. J Roy Stat Soc B 1996;58(1):267–88.
32. Zhang, Wendong. Development of a Real-Time Estimate of Flu Activity in the United States Using Dyamically Updated Lasso Regressions and Google Search Queries. Harvard senior honors theses from the Harvard School of Engineering and Applied Sciences, Accession 19083, Box 4, 2013.
33. WHO. Influenza fact sheet, 2009. http://who.int/mediacentre/factsheets/fs211/en/.
34. Patwardhan A, Bilkovski R. Comparison: flu prescription sales data from a retail pharmacy in the U.S. with Google Flu Trends and U.S. ILINet (CDC) data as flu activity indicator. PLoS One 2012;7(8):e43611.
35. Liu F, Lv B, Peng G, Li X. Influenza epidemics detection based on Google search queries. In: Goal FL, ed. Recent progress in data engineering and internet technology. New York: Springer, 2012:371–6.
36. Valdivia A, Monge-Corella S. Diseases tracked by using Google trends, Spain. Emerg Infect Dis 2010;16(1):168.
37. Kelly H, Grant K. Interim analysis of pandemic influenza (H1N1) 2009 in Australia: surveillance trends, age of infection and effectiveness of seasonal vaccination. Eur Surveill 2009;14(31):2.
38. Kang M, Zhong H, He J, Rutherford S, Yang F. Using Google trends for influenza surveillance in South China. PLoS One 2013;8(1):e55205.
39. Wada K, Ohta H, Aizawa Y. Correlation of "Google Flu Trends" with sentinel surveillance data for influenza in 2009 in Japan. Open Public Health J 2011;4:17–20.
40. Ortiz JR, Zhou H, Shay DK, Neuzil KM, Fowlkes AL, Goss CH. Monitoring influenza activity in the U.S.: a comparison of traditional surveillance systems with Google Flu Trends. PLoS One 2011;6(4):e18687.
41. Dugas AF, Hsieh YH, Levin SR, et al. Google Flu Trends: correlation with emergency department influenza rates and crowding metrics. Clin Infect Dis 2012;54(4):463–9.
42. Paul MJ, Dredze M. You are what you tweet: analyzing Twitter for public health. Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011); 2011 Jul 17–21; Barcelona, Spain.
43. Althouse BM, Allem J-P, Childers MA, Dredze M, Ayers JW. Population health concerns during the U.S.' Great Recession. Am J Prev Med 2014;46(2):166–70.
44. Ayers JW, Althouse BM, Allem JP, Rosenquist JN, Ford DE. Seasonality in seeking mental health information on Google. Am J Prev Med 2013;44(5):520–5.
45. Ayers JW, Althouse BM, Ribisl KM, Emery S. Digital detection for tobacco control: online reactions to the U.S.' 2009 cigarette excise tax increase. Nicotine Tob Res 2014;16(5):576–83.
46. Shaman J, Karspeck A. Forecasting seasonal outbreaks of influenza. Proc Natl Acad Sci U S A 2012;109(50):20425–30.