BMJ Open

Early detection of COVID-19 in China and the USA: summary of the implementation of a digital decision-support and disease surveillance tool

Yulin Hswen [iD],[1,2,3] John S Brownstein,[2,3] Xiang Xu,[4] Elad Yom-Tov [iD] [5,6]

For numbered affiliations see end of article.

Correspondence to
Dr Yulin Hswen;
yulin.hswen@ucsf.edu

ABSTRACT

Objectives Rapid detection and surveillance of COVID-19 is essential to reducing spread of the virus. Inadequate screening capacity has hampered COVID-19 detection, while traditional infectious disease response has been delayed due to significant demands for healthcare resources, time and personnel. This study investigated whether an online health decision-support tool could supplement COVID-19 surveillance and detection in China and the USA.

Setting Daily website traffic to Thermia was collected from China and the USA, and cross-correlation analyses were used to assess the designated lag time between the daily time series of Thermia sessions and COVID-19 case counts from 22 January to 23 April 2020.

Participants Thermia is a validated health decision-support tool that was modified to include content aimed at educating users about Centers for Disease Control and Prevention recommendations on COVID-19 symptoms. An advertising campaign was released on Microsoft Advertising to refer searches for COVID-19 symptoms to Thermia.

Results The lead times observed for Thermia sessions to COVID-19 case reports was 3 days in China and 19 days in the USA. We found negative cross-correlation between the number of Thermia sessions and rates of influenza A and B, possibly due to the decreasing prevalence of influenza and the lack of specificity of the system for identification of COVID-19.

Conclusion This study suggests that early deployment of an online campaign and modified health decision-support tool may support identification of emerging infectious diseases like COVID-19. Researchers and public health officials should deploy web campaigns as early as possible in an epidemic to detect, identify and engage those potentially at risk to help prevent transmission of the disease.

BACKGROUND

As of 9 May 2020, across 210 countries and territories there have been more than 3.8 million cases of COVID-19. COVID-19 likely began in December 2019, with the first cases resembling pneumonia documented in Wuhan, China.[1 2] Little information was known about the virus and its clinical course,

Strengths and limitations of this study

► Digital detection of COVID-19 in conjunction with the provision of a decision support tool for people who suspected they might be ill.
► Cross-country analysis, USA and China.
► Data could not be linked to actual cases so as to preserve privacy. Therefore, only correlational evidence can be provided to support findings.
► Our data comprised relatively few user interactions. This is partly due to awareness on the side of people and partly to the budget allocated to creating such awareness through ads.

making early detection a challenge.[3] Additionally, successful reporting of those who are infected with COVID-19 relies on healthcare capacity, such as the availability of healthcare workers, and medical resources, which may not always be readily available.[3]

To assist with case detection of infectious diseases, digital surveillance has often been used to supplement traditional epidemiological monitoring approaches.[4] Digital surveillance tools have been used in the past, beginning with tools for surveillance of influenza[5] and later expanded to other conditions.[6 7] However, experience has shown that such tools need careful tuning to successfully track cases of illness.[8] Notwithstanding these limitations, tools based on search queries,[9] advertising[10] and other digital signals[11] have been proposed. Currently, these tools are used for tracking COVID-19,[12] for example, in England[13 14] using search engine data.

In response to the rising case reports of COVID-19 in China, we developed an advertising campaign on Microsoft Advertising in conjunction with Thermia[4], an online health decision-support tool. In late January 2020, before WHO declared COVID-19 as a Public Health Emergency (30 January 2020),[15] we deployed Thermia to identify infections and

advise the public about their potential risks of having COVID-19. Thermia, a decision support to for febrile illness was adapted to include symptoms and human mobility behaviours related to COVID-19 to be able to provide recommendations for treatment of COVID-19. At the beginning of the COVID-19 pandemic differences between the influenza and COVID-19 were not fully understood and travel history was prominent in helping identifying possible COVID-19 infection as defined by the CDC. As well, questions about severity of symptoms were also not included because at the time of the study the CDC had only provided a list of potential symptoms of COVID-19 and did not expand on the severity of the symptoms and their differential diagnosis of the disease. Thus, the modification of Thermia from influenza to COVID-19 composed of adding questions on travel history and a change to the response to users, which was based on CDC guidelines.

Here, we report on the use of the Thermia tool as a method for digital surveillance for supporting the early detection of emerging cases of COVID-19 in China and the USA.

METHODS
Digital Health Surveillance for COVID-19
Thermia[4] is a digital decision-support tool, that was developed by researchers at Harvard Medical School, to provide clinical advice on how to treat febrile illness based on an evidence-based algorithm. Users are referred to the Thermia platform based on their symptoms queries related to febrile illness on a web search engine. At Thermia, users are directed through a series of questions about their temperature, symptoms and biometric characteristics (see figure 1) and are given recommendations on how to further proceed with medical care based on their provided answer. Thermia has also been used for digital surveillance, and was previously validated for early detection of influenza in China.[4]

We modified our validated Thermia platform to provide advice about COVID-19 symptoms based on the recommendations from the US Centers for Disease Control and Prevention (CDC). These recommendations urged individuals to carefully monitor their symptoms and to stay at home and inform their doctor about their recent travel history and symptoms before seeking in-person care.[16] In the third week of January, we began an advertising campaign on Microsoft Advertising in conjunction with Thermia to provide recommended practices for COVID-19 to users from the USA who experienced symptoms consistent with COVID-19 and who had travelled (first to China, and later more generally) in the past 14 days. Persons who queried for symptoms related to COVID-19 were presented with shown ads that asked users whether they had COVID-19-related symptoms and whether they had travelled to China or outside the United States in the past 2 weeks. Users who clicked on the ads were referred to the Thermia website.[17] Additionally, users could find the Thermia website by searching for symptoms using general web search and arrive at the site directly.

The volume of daily web traffic to Thermia from the USA and China enabled us to conduct cross-correlation analysis of traffic volume with case counts. We restricted this study to investigating visits to the Thermia website in China and the USA because the campaign was placed in both English and Mandarin and the ads were shown in the USA and related to travel from China and outside the USA. We also observed the largest data volume in sessions from those two countries. Therefore, although web traffic was seen from Canada and the UK, the amount of daily web traffic was not sufficient to conduct our temporal analyses.

Our hypothesis was that web traffic to Thermia, from the online advertisements or direct visits to the website, would serve as a proxy of COVID-19 cases. Specifically, we expected that Thermia sessions would provide an earlier signal of COVID-19 cases in both China and the USA.

Data
Daily COVID-19 cases from China and the USA were collected through The Center for Systems Science and Engineering at John Hopkins University. Sessions to Thermia were aggregated on a daily basis, stratified to those from China and from USA. The daily time series for COVID-19 cases and Thermia sessions were plotted for China and the USA.

Cross-correlation analysis
We sought to examine the relationship between COVID-19 cases and Thermia sessions, and to identify if past day lags of Thermia sessions are predictive of COVID-19 case counts in China and the USA. Cross-correlations were used to analyse the relationship between two signals and calculate the correlation coefficient at designated lags (displacements) between the two series. The maximum correlation coefficient is the time point that the two series correlate most closely and the coefficient reveals how much two series correlate with one another.[18]

The cross-correlation between two discrete sequences x and y at lag k is defined as:
$$(x * y)\,[k] \triangleq \sum_t x\,[t + k] * y * [t], \text{ where } y * [t] \text{ denotes}$$
the conjugate of y.

Therefore, we used the cross-correlation function (CCF) between the Thermia sessions X_t and the COVID-19 cases Y_t at varying lags to evaluate the lag that may be useful for predicting COVID-19, Y_t. We defined the set of correlations between x_{t+k} and y_t whereby k = 0, 1±, 2±, 3± etc

A negative value for k is a correlation between Thermia sessions at a time (in days) before t and COVID-19 cases at time t.

For Thermia sessions and COVID-19 case time series in China and the USA, we set $k = 30$ (days) to allow the CCF to explore the correlation between thirty-day lags of Thermia sessions and COVID-19 case time series. The CCF analysis includes the date range in which both time

Figure 1 Thermia platform.

series overlap for China and the USA, which is from 22 January 2020 to 23 April 2020.

In the USA, influenza season occurs in the fall and winter and peaks in December and February. Weekly influenza data for the 2019–2020 season is provided from the CDC.

To explore the possibility that Thermia was capturing influenza, and to further validate our cross-correlation results between Thermia sessions and COVID-19 cases, we evaluated the cross-correlation with the CDC influenza time series. Influenza A and Influenza B circulated

in the 2019–2020 season and so we conducted the cross-correlation on the weekly time series of the percent of Influenza A and percent of Influenza B to the weekly time series of Thermia sessions (we could only explore the influenza trends in the USA because of availability of data from the CDC). All analyses were conducted using the R V.3.6.

Patient and public involvement

Patients or the public were not included in the design, or conduct, or reporting, or dissemination plans of our research.

RESULTS
Descriptive results

In China, the average number of daily Thermia sessions was 27.1 (SD=23.4, Min=4, Max=123) and the average COVID-19 number of daily cases was 839.6 (SD=2885.2, Min=3, Max=15 136). In the USA, the average number

of Thermia sessions was 17.7 (SD=7.0, Min=4, Max=42) and the average number of COVID-19 cases was 9346.0 (SD=12 881.1, Min=0, Max=34 126). Figure 2A shows a map of the Thermia sessions from China (at the city level) and figure 2B shows Thermia sessions from the United States (at the state level). Figure 3A displays daily Thermia sessions and COVID-19 cases in China and figure 3B shows the same in the USA. Thermia sessions increased and exhibited a first large spike at the end of January 2020. A second spike was observed at the beginning of February 2020. COVID-19 cases began in early February and showed a surge in late February. Descriptive results of Thermia session and COVID-19 cases are presented in table 1.

Cross-correlation between Thermia sessions and COVID-19 case time series

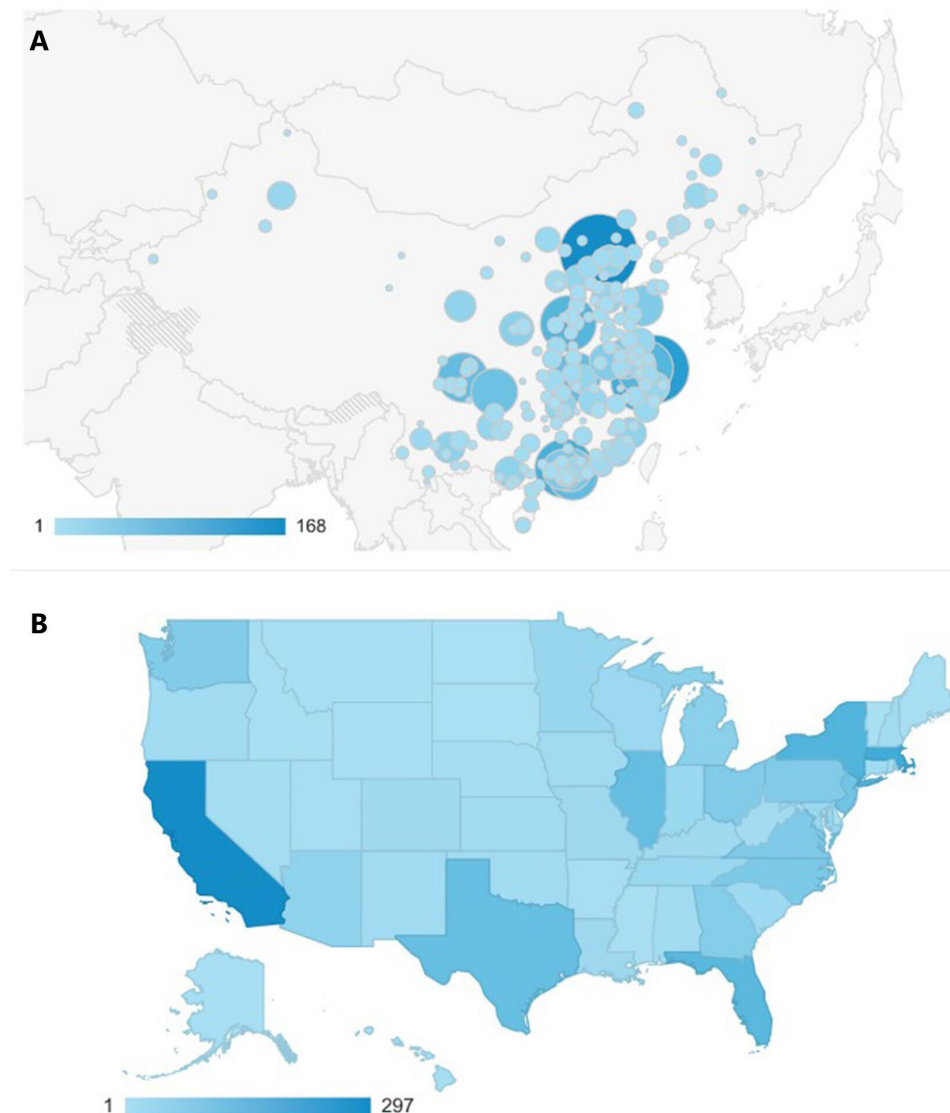Cross-correlation analyses confirmed a significant positive cross-correlation between Thermia sessions and

Figure 2 Thermia sessions in China and the United States. (A) Thermia sessions in China. (B) Thermia sessions in the USA.
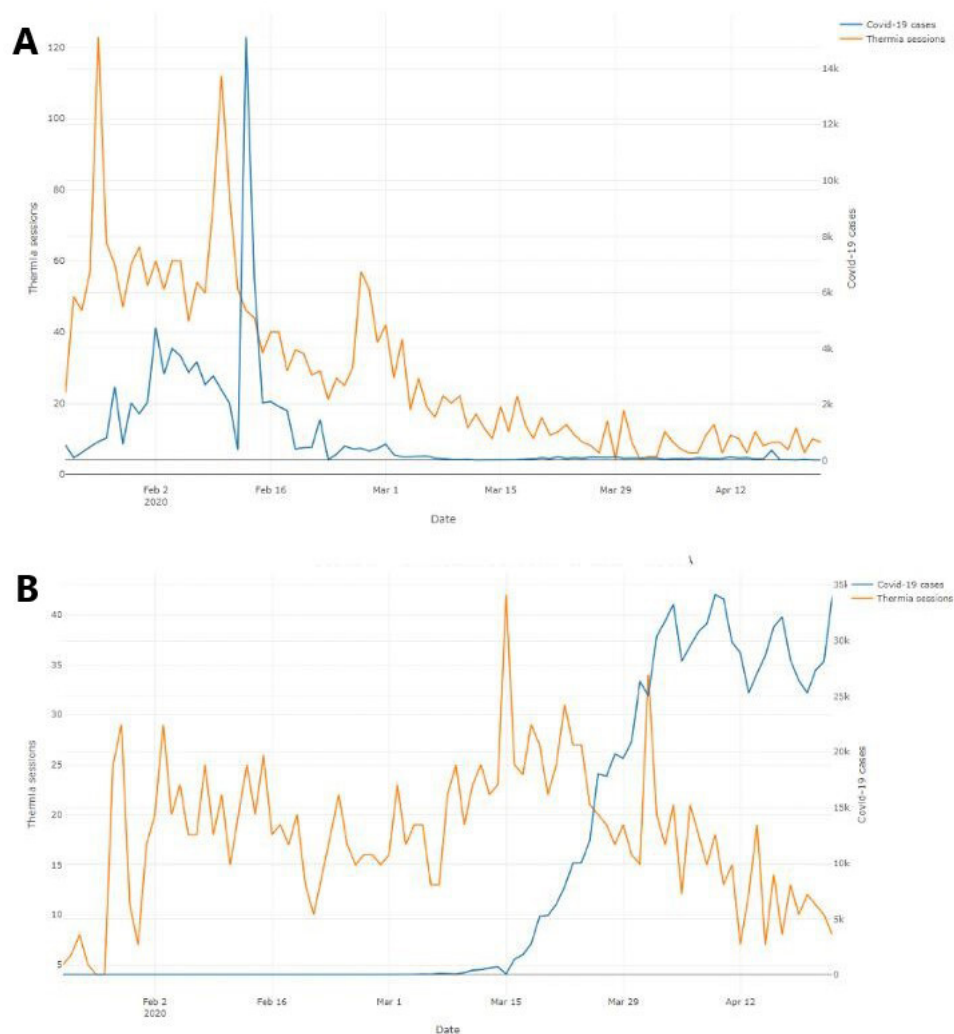
Figure 3 Time series of Thermia sessions and COVID-19 cases. (A) Time series of Thermia sessions and COVID-19 cases in China. (B) Time series of Thermia sessions and COVID-19 cases in the USA.

Table 1	Descriptive statistics for Thermia users	
Daily	**Thermia sessions**	**COVID-19 cases**
China		
Mean	27.12	839.56
SD	23.4466	1885.223
Min	4	3
25% quantile	9	46.25
50% quantile	18	98.50
75% quantile	42.25	555
Max	123	15 136
USA		
Mean	17.96	9346
SD	6.96	12 881.10
Min	4	0
25% quantile	13	0
50% quantile	18	68
75% quantile	22	25 070
Max	42	34 126

COVID-19 cases in China, with the highest CCF at $k = -3$ (days), $(r = 0.68, p < 0.05)$ and a significant positive correlation between Thermia sessions and COVID-19 cases in the USA at $k = -19$ (days) $(r = 0.41, p < 0.05)$. The correlation lagged days for Thermia sessions and COVID-19 cases in China and the USA are presented in the online supplemental appendix.

Exploratory cross-correlation between Thermia sessions and influenza A/B in the United States

The CDC weekly influenza data for the 2019–2020 season showed that influenza A began to decline mid-February while influenza B declined in early January and both had close to zero cases mid-March. Conversely, Thermia sessions begin to rise in mid-Feb and peak mid-March. These results highlight that Thermia was likely capturing cases of COVID-19 and not influenza A or B since there were no cases of influenza shortly after mid-March (figure 4). In the USA, the cross-correlation analyses confirmed a significant negative cross-correlation between Thermia sessions and the CDC per cent Influenza A, with the highest CCF at $k = -4$ (weeks), $(r = -0.51, p < 0.05)$ and a significant negative correlation
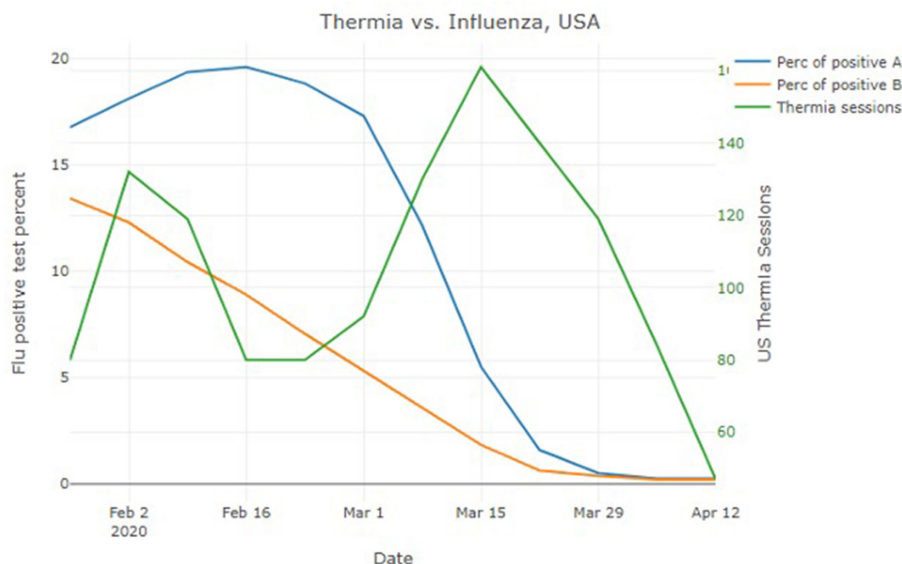
Figure 4 Time series of Thermia sessions and influenza positive a/b percentage.

between Thermia sessions and COVID-19 cases in the USA at k = −2 (weeks) (r = −0.75, p < 0.05). The correlation lagged days for Thermia sessions and percent influenza A and influenza B in the USA are presented in the online supplemental appendix.

DISCUSSION

Most epidemiological monitoring tools, especially those dependent on online interactions (eg, search based), rely on a combination of factors for their success. The first of these is that there is a significant lag between the appearance of symptoms and the first time that people visit the medical system or that only a small part of the infected population visits the medical system. When symptoms are severe enough to warrant an urgent visit to a hospital and when most infected people visit a medical provider, data from the medical system will be superior to that of search-based infodemiological systems. When no symptoms exist, people will not query about them, making the monitoring tools ineffective. Second, good ground-truth data are needed to calibrate these systems. In the case of influenza, for example, researchers often use past seasons to tune the models.

COVID-19 has the first set of attributes (eg, lag between symptoms and visit to the medical system,[13] but, especially at the beginning of the epidemic, there was a limited understanding of the symptoms and although we modified Thermia to provide information to people on their condition, there was insufficient ground truth to tune a symptom-search model. For instance, most recent evidence has shown that symptoms such as loss of taste and smell are highly indicative of a COVID-19 infection. Stated differently, multiple questions were needed to ascertain the severity of symptoms, and these were obtained through the use of Thermia. Thus, the combination of ads and questionnaire allowed us to go beyond simple searches. Later in the pandemic, as

more ground-truth data became available, tuning was made possible, as shown in Lampos et al.[14]

Our findings indicate that website traffic to Thermia (ie, Thermia sessions) had a lead time of 3 days for COVID-19 cases in China and a lead time of 19 days in the USA. This provides evidence that an advertising campaign coupled with a digital health decision-support tool may be effective at identifying early signals of a novel respiratory pathogen like COVID-19. Previous evidence using Thermia as a digital surveillance tool has been validated for early influenza detection in China;[4] however, Thermia has not been used as a tool to detect the emergence of new infectious diseases like COVID-19. Patients often search the Internet for information about their health prior to meeting with a provider to make decisions about how to treat themselves, or whether or not they should see a provider.[19 20] Health-seeking behaviour in the form of queries to online search engines often precedes provider visits,[5] thus, search queries related to COVID-19 symptoms on the web may have also played a role in generating predictive signal by visiting Thermia. Because of the deployment of our advertisement campaign, the 3-day lead time of Thermia website traffic to COVID-19 cases in China may have been a result of patients seeking information about their COVID-19 symptoms on the internet before they were tested for COVID-19. Thermia sessions in China had peaks around mid-January and mid-February whereas the rise in COVID-19 cases in China occurred at the start of February and showed a peak in the early weeks in February.

Interestingly, a 19-day lead time of Thermia sessions to COVID-19 cases was seen in the USA. It has been documented that the scarcity of supplies, limited access to screening and problems with test kits, have all hampered the ability to effectively detect and monitor COVID-19 cases in many parts of the country.[21 22] Thus, these differences in lead time may be due to faulty test kits and the long delay in large-scale testing that occurred in the United States compared with China,

where testing was initiated earlier and more widely.[23 24] China, by contrast, sought to mobilise large scale testing capacity to test all inhabitants in high-risk areas of the country. As of 29 June 2020, China has carried out one test for every 15 people, compared with 1 in 11 in the USA.[25] These number were more divergent in the early phase of the pandemic because of the limited tested that was provided by the USA.[26] This is also why we opted to use COVID-19 cases instead of ratio of deaths to cases, as information on deaths at the beginning of the epidemic was noisy due to the dearth of tests and would likely have been noisier than comparison to case numbers. Furthermore, the occurrence of cryptic cases of COVID-19 from updated evidence show the virus was circulating in the USA in early February.[27] These data indicate a sustained community transmission had started before the detection of the first US cases. Our results, showing a greater amount of sessions in California, may be due to COVID-19 cases that were not originally detected. The findings here suggest that the 19-day lead time of Thermia session in the USA is a signal for early COVID-19 cases that were not captured by traditional public health monitoring.

Furthermore, models of the transmission of COVID-19 have implied that the virus may have initially spread undetected and suggest that the first infections occurred much earlier than reported.[28 29] For instance, it was assumed that the first US fatality due to COVID-19 occurred in Seattle on 28 February 2020, but postmortem testing on deaths from 6 February to 17 February 2020 have confirmed that COVID-19 was spreading in the San Francisco Bay area weeks earlier than previously documented.[29] This new fatality data suggest that the virus had been spreading for at least a nearly 3-week period in Santa Clara prior to early February, which was largely believed to be because of limited capacity for testing.[29 30] Our results align with this current set of information of an almost 3-week lead time (19 days) of Thermia sessions to COVID-19 cases in the USA.

A limitation of this study is that we were not able to confirm whether visitors to the Thermia website were COVID-19 positive cases, as users were not followed up with and tested for COVID-19. However, we were able to validate that our results from Thermia sessions in the USA were more likely to be reporting COVID-19 cases than influenza because a negative cross-correlation was seen for Influenza A and B, implying that as Thermia sessions increased, cases of Influenza A and B declined. This could also be interpreted that as COVID-19 cases began to rise in the USA, Influenza A and B cases waned. It is unclear if Thermia would be valid to differentiate COVID-19 to influenza if the prevalence of these disease were parallel over the same time period. Although our validation is based on the different trajectories of the prevalence of influenza and COVID-19, the advertising campaign for Thermia was adapted to include for symptoms of COVID-19 and travel related to the primary areas of emergence of COVID-19. Therefore, there is less of a likelihood of user visiting Thermia who were cases of influenza and our web campaign advertisement for Thermia may have supported detection of cases of COVID-19 from persons who had recently travelled outside the USA. Future studies should test Thermia's ability to differentiate influenza from COVID-19 during the seasonal influenza season.

Furthermore, the cross-correlation of =0.41 in the USA is slightly lower than traditionally research that evaluates the use of online digital tools for earlier detection. However, these correlations between Google Flu Trends and influenza case counts have ranged between (=0.42) and (=0.88).[31] Therefore, the fact that we detect a significant correlation of (=0.41) in the USA with Thermia, a specific surveillance tool adapted for COVID-19, a novel disease, is meaningful. This lower correlation and the greater lead time in the USA of 19 days is seen compared with a 3-day lead China this may be a result of greater testing capacity in China at the beginning of the pandemic.[25] However, the different lag time between countries could also change in time as user interactions with Thermia develop. This presents a challenge for the ability of health authorities to make practical use of the data from this system. Another limitation of our study is the relatively few user interactions we were able to obtain. This is partly due to awareness on the side of people and partly to the budget allocated to creating such awareness through ads. Thus, analyses are based on very few user interactions that were collected with a minimum of four per day and low average of interaction in both USA and China.

Finally, we are unable to isolate the unique users because of limitations in privacy as the data is provided in an aggregated format. Thus, it is possible that the same user will have more than one Thermia visit. However, since Thermia is decision-support tool that provided recommendations for COVID-19, it is unlike a user would revisit unless it was to retrieve information about another user whom they were using the Thermia to get information for.

CONCLUSION

Early deployment of critical information about a novel disease like COVID-19 using a web-based campaign and health decision-support tool may be able to predict the emergence of the disease and help increase public awareness. Considering people often turn to the Internet to find out information early in an outbreak, directing people to a validated and evidence-based web platform may have the ability to generate predictive warning signals.

Here, we demonstrated the ability to rapidly respond to a novel disease outbreak by quickly creating a system which provided people with decision support and, through the data collected by it, provide surveillance information that could be used by health authorities. The use of both Thermia and advertising allowed us to go gain several advantages, including directly approaching people who may yet be invisible to the health system, obtain information beyond searches themselves (because of the use of Thermia), and create awareness through a targeted advertising campaign.

In the future, it would be important for public health researchers, and policy-makers to work with industry leaders in the field of technology to deploy web campaigns as early on as possible in an epidemic to detect, identify and engage

those at risk to more effectively identify the transmission of infectious diseases such as COVID-19.

Author affiliations
[1]Epidemiology and Biostatistics, Bakar Computational Health Institute, University of California San Francisco, San Francisco, California, USA
[2]Computational Epidemiology Lab, Harvard Medical School, Boston, Massachusetts, USA
[3]Innovation Program, Boston Children's Hospital, Boston, Massachusetts, USA
[4]Department of Statistics, Boston University, Boston, Massachusetts, USA
[5]Microsoft Research, Herzeliya, Israel
[6]Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Haifa, Israel

Contributors YH retrieved funding, conceived the idea for the study, designed the platform and methods, collected the data and wrote the manuscript. JSB retrieved funding for the project, helped design the platform, supervised the findings of this work. XX cleaned data, assisted with analysis of the data and translation of text. EY-T encouraged YH to investigate the topic, supervised the findings of the work, verified the analytical methods and helped write the manuscript. All authors contributed to the final version of the manuscript.

Disclaimer The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Map disclaimer The depiction of boundaries on the map(s) in this article does not imply the expression of any opinion whatsoever on the part of BMJ (or any member of its group) concerning the legal status of any country, territory, jurisdiction or area or of its authorities. The map(s) are provided without any warranty of any kind, either express or implied.

Competing interests EY-T is an employee at Microsoft, owner of Bing. All work described in the paper was done as part of his salaried employment.

Patient consent for publication Not required.

Ethics approval Institutional Review Board approval was granted by Boston Children's Hospital for this study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Data collection was approved by the Boston Childrens Hospital Institutional Review Board and cannot be shared.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs
Yulin Hswen http://orcid.org/0000-0003-3203-1322
Elad Yom-Tov http://orcid.org/0000-0002-2380-4584

REFERENCES

1 Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020;20:533–4.
2 World Health Organization. Coronavirus disease (COVID-19) pandemic World Health organization, 2020. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019?gclid=CjwKCAiA7939BRBMEiwA-hX5JwLIsa-X_b8NDiR_H-H69rP6s7m0f4dQ-G-rPi31X8xLmtGoKiPGERoCtQEQAvD_BwE
3 Rajgor DD, Lee MH, Archuleta S, *et al*. The many estimates of the COVID-19 case fatality rate. *Lancet Infect Dis* 2020;20:776–7.
4 Hswen Y, Brownstein JS, Liu J, *et al*. Use of a digital health application for influenza surveillance in China. *Am J Public Health* 2017;107:1130–6.
5 Ginsberg J, Mohebbi MH, Patel RS, *et al*. Detecting influenza epidemics using search engine query data. *Nature* 2009;457:1012–4.
6 Oren E, Frere J, Yom-Tov E, *et al*. Respiratory syncytial virus tracking using Internet search engine data. *BMC Public Health* 2018;18:445.
7 Patrick Copeland and Raquel Romano and Tom Zhang and Greg Hecht and Dan Zigmond and Christian Stefansen. *Google disease trends: an update*, 2013.
8 Lazer D, Kennedy R, King G, *et al*. Google flu trends still appears sick: an evaluation of the 2013-2014 flu season. *SSRN Journal* 2014.
9 Wagner M, Lampos V, Yom-Tov E, *et al*. Estimating the population impact of a new pediatric influenza vaccination program in England using social media content. *J Med Internet Res* 2017;19:e416.
10 Eysenbach G. *Infodemiology: tracking flu-related searches on the web for syndromic surveillance*, 2006.
11 Butler P, Ramakrishnan N, Nsoesie EO, *et al*. Satellite imagery analysis: what can Hospital parking lots tell us about a disease outbreak? *Computer* 2014;47:94–7.
12 Ting DSW, Carin L, Dzau V, *et al*. Digital technology and COVID-19. *Nat Med* 2020;26:459–61.
13 Yom-Tov E, Lampos V, Cox IJ, *et al*. Providing early indication of regional anomalies in COVID19 case counts in England using search engine queries. *arXiv* 2020.
14 Lampos V*et al*. Tracking COVID-19 using online search. *arXiv* 2020.
15 Rolling updates on coronavirus disease (COVID-19). *World Health organization web site*, 2020.
16 Coronavirus (Covid-19). Center for Disease Control. *Center for disease control and prevention web site*, 2020.
17 Thermia. Available: https://thermia.io/ [Accessed Dec 4 2020, 2020].
18 Podobnik B, Horvatic D, Petersen AM, *et al*. Cross-correlations between volume change and price change. *Proc Natl Acad Sci U S A* 2009;106:22079–84.
19 Nagler RH, Gray SW, Romantan A, *et al*. Differences in information seeking among breast, prostate, and colorectal cancer patients: results from a population-based survey. *Patient Educ Couns* 2010;81 Suppl:S54–62.
20 Ofran Y, Paltiel O, Pelleg D, *et al*. Patterns of information-seeking for cancer on the Internet: an analysis of real world data. *PLoS One* 2012;7:e45921.
21 Goodnough A, Thomas K, Kaplan S. *Testing falls woefully short as Trump seeks an end to Stay-at-Home orders*, 2020.
22 Cohen J, Kupferschmidt K. Countries test tactics in 'war' against COVID-19. *Science* 2020;367:1287–8.
23 Ioannidis JPA. *A fiasco in the making? as the coronavirus pandemic takes hold, we are making decisions without reliable data*, 2020.
24 Niehus R, De Salazar PM, Taylor A, *et al*. *Quantifying bias of COVID-19 prevalence and severity estimates in Wuhan, China that depend on reported cases in international travelers*, 2020. Available: http://medrxiv.org/lookup/doi/10.1101/2020.02.13.20022707
25 BBC. *Coronavirus: has China or the US tested more?* 2020.
26 Schuchat A. *Public health response to the initiation and spread of pandemic COVID-19 in the United States, February 24–April 21, 2020*, 2020.
27 CDC COVID-19 Response Team, Jorden MA, Rudman SL, *et al*. Evidence for limited early spread of COVID-19 within the United States, January-February 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:680–4.
28 Chinazzi M, Davis JT, Ajelli M, *et al*. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 2020;368:395–400.
29 Chavez N, Karimi F, Levenson E. *Coronavirus spread 'under the radar' in US major cities since January, researchers say*, 2020.
30 Chappell B. *First known U.S. COVID-19 death was weeks earlier than previously though*, 2020.
31 Dugas AF, Hsieh Y-H, Levin SR, *et al*. Google flu trends: correlation with emergency department influenza rates and crowding metrics. *Clin Infect Dis* 2012;54:463–9.