Reliability of COVID-19 symptom checkers as national triage tools: an international case comparison study

Fatma Mansab, Sohail Bhatti, Daniel Goyal

ABSTRACT

Objectives Triage is a critical component of the pandemic response. It affects morbidity, mortality and how effectively the available healthcare resources are used. In a number of nations the pandemic has sponsored the adoption of novel, online, patient-led triage systems—often referred to as COVID-19 symptom checkers. The current safety and reliability of these new automated triage systems remain unknown.

Methods We tested six symptom checkers currently in use as triage tools at a national level against 52 cases simulating COVID-19 of various severities to determine if the symptom checkers appropriately triage time-critical cases onward to healthcare contact. We further analysed and compared each symptom checker to determine the discretionary aspects of triage decision-making that govern the automated advice generated.

Results Of the 52 clinical presentations, the absolute rate of onward referral to any form of healthcare contact was: Singapore 100%, the USA 67%, Wales 65%, England 62%, Scotland 54% and Northern Ireland 46%. Triage decisions were broadly based on either estimates of ‘risk’ or ‘disease severity’. Risk-based symptom checkers were more reliable, with severity-based symptom checkers often triaging time-critical cases to stay home without clinical contact or follow-up.

Conclusion The COVID-19 symptom checkers analysed here were unable to reliably discriminate between mild and severe COVID-19. Risk-based symptom checkers may hold some promise of contributing to pandemic case management, while severity-based symptom checkers—the CDC and NHS 111 versions—confer too much risk to both public and healthcare services to be deemed a viable option for COVID-19 triage.

INTRODUCTION

Symptom checkers are online platforms where the public can enter details of their illness, answer set questions about their symptoms and then receive advice on what to do next. During the pandemic, many nations have deployed symptom checkers to help identify potential COVID-19 cases and provide advice to the public. Some nations have gone further, using symptom checkers in place of more typical clinical triage systems.

Despite a number of studies highlighting the diagnostic sensitivity of various online COVID-19 symptom checkers, we could find no studies (apart from our previous analysis) examining the safety and reliability of online COVID-19 symptom checkers as a standalone triage tool.

The difference is stark. On the one hand, these accessible web-based questionnaires can direct potential cases toward SARS-CoV-2 testing services—answering the question: should you be tested? In such circumstances, symptom checkers act more as a prompt, conveying the national advice. On the other hand, there is another category of symptom checkers attempting to answer a much more complicated question: do you need medical help?

It is quite an ask, of an automated system. And the stakes are high. There is the unavoidable direct morbidity and mortality impact when triaging acute medical problems. There is also an operational consideration, whereby delaying treatment in...
time-critical conditions leads to a higher overall healthcare burden.\(^3\)\(^-\)\(^6\) Online COVID-19 symptom checkers may carry the potential to offset the inevitable high healthcare demands of pandemic management, allowing valuable resources to be focused on those with real clinical need. This benefit can only be realised if the symptom checker successfully triages COVID-19 pneumonia (and other serious conditions mimicking COVID-19) on to further care early enough for maximal treatment benefits to be achieved. Missing the opportunity to prevent disease progression will invariably lead to higher mortality, delayed recovery (eg, higher rates of long COVID-19)\(^4\)\(^-\)\(^6\) and a more lengthy inpatient stay.\(^4\)\(^-\)\(^6\) There is a very real possibility that symptom checkers, if inappropriately used, can significantly increase the healthcare burden associated with COVID-19—compromising healthcare capacity sooner than is necessary.\(^2\) Both at a patient level and operational level, quality and efficacy studies are essential if clinical activities are to be replaced or even augmented with such patient-led, automated clinical services.

Back in April 2020, we undertook an analysis on national symptom checkers from the UK, USA, Singapore and Japan. At that time, the results of our case-simulation study revealed a low rate of onward referral for both the UK’s ‘111 COVID-19 Symptom Checker’ and the US’s ‘CDC Coronavirus Symptom Checker’—44% and 38%, respectively. It was noted that both symptom checkers triaged simulated cases of severe COVID-19 pneumonia, bacterial pneumonia and sepsis, to stay home with no further healthcare contact.\(^2\) In short, both the US and UK symptom checkers maintained a high threshold to refer patients onward for healthcare contact.

During this previous analysis, we compared and contrasted the four symptom checkers, looking specifically for points of divergence. The most notable difference was whether known COVID-19 disease ‘risk factors’ or an estimated ‘disease severity’ were used to calculate triage disposition. Both Singapore’s and Japan’s symptom checkers focused specifically on risk factors—age, duration of symptoms, the presence of breathlessness and comorbidities—and made no attempt to quantify disease severity. We have termed these, ‘risk-based symptom checkers’. Both the US and UK symptom checkers, in April 2020, relied more on qualitative questions designed to estimate how severe a case was, and made no account of the most consistent risk factors—age, duration of symptoms or the presence of breathlessness. In addition, only moderately severe comorbidities affected the CDC symptom checker triage decision and only ‘shielding category’ comorbidities affected triage dispositions in the NHS 111 symptom checker.\(^2\) We have termed these, ‘severity-based symptom checkers’.

At the time of our initial study, the impact of the CDC symptom checker on actual case presentations to US hospitals was unknown. However, the ‘NHS 111 COVID-19 Symptom Checker’ was a known gatekeeper for UK patients with suspected or confirmed COVID-19.\(^8\)\(^-\)\(^11\) As such, the symptom checker’s ‘decision’ to triage cases simulating time critical urgent medical problems—severe COVID-19, bacterial pneumonia and sepsis—to remain at home was considered by the authors as unsafe. No data existed on internal or external quality assurance studies, further compounding the concerns regarding the use of the NHS 111 COVID-19 symptom checker and patient safety. These concerns were raised with NHS Digital—the body responsible for the NHS 111 symptom checker—both prior to publication and following. NHS Digital considered the concerns to be historic and not representative of the improved version of the symptom checker.\(^12\)

We, therefore, undertook a repeat analysis in June 2021.

METHODS
During the first week in June 2021, we undertook a follow-up analysis on national symptom checkers. In summary, we generated four distinct patient scenarios relating to COVID-19. The four scenarios included were fever with cough, comorbidity with fever and cough, immunosuppression with fever and cough, and shortness of breath with fever. We varied patient age, duration of symptoms and the severity of symptoms. In total, this generated 52 separate case simulations, including mild, moderate, severe and critical COVID-19, and COVID mimickers such as bacterial pneumonia and sepsis. Each case was applied by a single investigator (DG) to each symptom checker and the triage decision was recorded. We then calculated the total referral ratio (ie, proportion of all 52 cases that were referred for clinical contact, regardless of the level of designation—call centre, primary care provider, or emergency department). A percentage ratio was generated of total referrals made by each symptom checker. We also noted specific features of each symptom checker for comparison.

In addition to the methods as described previously, we also completed analysis of symptom checkers from Scotland,\(^13\) Wales\(^14\) and Northern Ireland,\(^15\) and thus differentiated the previously analysed NHS England symptom checker from the other three nations.

A more detailed methodology is explained by Mansab et al.\(^2\)

RESULTS
Of the four nations included in the initial analysis of April 2020, Singapore, the USA and all four nations within the UK continue to use symptom checkers as part of the national response to COVID-19. Japan was no longer using an accessible symptom checker, it being replaced by a flow chart. As such, it was excluded from further analysis.

Triage dispositions were slightly different for each symptom checker, but generally followed: stay home, or contact service provider/General Practitioner (GP)/111, or go straight to emergency department/999.
The current analysis was undertaken during a relatively low prevalence time, when COVID-19 inpatient burden was relatively low (table 1). Referral ratio

The rates of onward referral to any further healthcare contact for each national symptom checker were: Singapore 100%, the USA 67%, Wales 65%, England 62%, Scotland 54% and Northern Ireland 46% (figure 1). Previous referral rates in April 2020 were: Singapore 88%, the USA 38% and England 44%. For the triage disposition of individual case simulations, see online supplemental data.

Specific features

The Singapore symptom checker currently refers all cases for a same day assessment at one of the nation’s public health clinics. It continues to refer all patients with any degree of breathing problems directly to the emergency department (scenario 4, online supplemental material). The US (CDC) symptom checker referred twice as many cases on to clinical care than it did the year before. Notably, the advice for those referred had changed from ‘contact medical provider within 24 hours’ to ‘contact medical provider as soon as possible’. Age was also now a considered risk factor for disease severity with all patients over the age of 65 years with suspected COVID-19 being advised to contact their medical provider regardless of disease severity or other comorbidity. The CDC symptom checker continued to triage those under 65 years of age with mild to moderate shortness of breath to stay home with no further clinical contact (scenario 4, online supplemental material).

The UK symptom checkers do not account for age in the triage decision in the case simulations undertaken, except for NHS Wales. NHS Wales ‘111’ symptom checker triaged all cases over the age of 70 years onward to call ‘111’. Scotland, Northern Ireland and England continued to, for example, triage a 72 years old with cough and fever for 7 days to stay home with no healthcare contact or follow-up (scenario 1, online supplemental material).

In comparison to the year previously, the NHS England symptom checker now triaged any case with the subjective sense of shortness of breath onward to further healthcare contact (‘call 111’ for cases with self-rated mild to moderate shortness of breath, and the emergency department for severe shortness of breath). If, though, shortness of breath is a secondary symptom (ie, feeling flu-like with shortness of breath), then patients are still advised to stay home, unless self-identified as severe (table 2).

DISCUSSION

The WHO guidelines on triage recommend that all patients with suspected or confirmed COVID-19 are clinically triaged. In many countries, this proves challenging due to national policy, healthcare accessibility, healthcare resources and the level of community transmission of SARS-CoV-2. Most patients with COVID-19 do not develop complicated illness and will resolve the infection without clinical intervention. Successful clinical pathways are able to pick out the COVID-19 pneumonia reliably and, crucially, early enough for basic medical care to prevent progression of disease.

There have been noticeable improvements in the symptom checkers that were first analysed in April 2020. Notably, the overall referral rate onward to healthcare contact has increased in the USA, UK and Singapore. Other features of the symptom checkers have also been improved. Singapore has included a question on immunosuppression, although it still does not confer such patients to more urgent care. The CDC coronavirus symptom checker has now included age as a defining risk factor. Most symptom checkers have expanded

<table>
<thead>
<tr>
<th>Table 1</th>
<th>National inpatient healthcare burden and key population statistics. Source: World Bank and WHO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Singapore</td>
</tr>
<tr>
<td>Population data</td>
<td></td>
</tr>
<tr>
<td>Patients currently admitted to hospital per 10 000 inhabitants (rate as per April 2020)</td>
<td>0.38 (2.23)</td>
</tr>
<tr>
<td>Mean national age (years)</td>
<td>44.2</td>
</tr>
<tr>
<td>Gross Domestic Product per capita (thousands of US dollars)</td>
<td>59.8</td>
</tr>
<tr>
<td>Physicians per 10 000 head of capita</td>
<td>24</td>
</tr>
<tr>
<td>Total case fatality rate (%)</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Figure 1 Percentage ratio of absolute onward referrals of each national symptom checker. Black represents the total percentage of cases triaged onward to further healthcare contact. Grey represents the percentage of cases triaged to remain at home with no further planned healthcare follow-up.
the list of comorbidities affecting triage decisions. And while the seriousness of breathlessness is given variable attention by each national symptom checker, all seem to now give it more weight in triage decisions. While these changes move towards re-establishing the usual standards of care for patients with COVID-19, our results indicate that COVID-19 symptom checkers are too unreliable in discriminating mild from severe COVID-19, and, as such, are likely to confer too great a risk to public, and, if used in place of clinical triage, will only weaken the healthcare response to the pandemic.20

Specifics

Whether the CDC coronavirus symptom checker has in fact worsened outcomes for the USA or not remains unknown. It has not been formally integrated into the US healthcare system or COVID-19 clinical pathways. The fact it functions as a stand-alone patient-led triage service and is freely accessible to the public is concerning. While not formally or preferentially directed to the symptom checker, the US public can seek clinical advice from the automated service, and thereby, potentially—in the patient’s mind—negate the need to undertake actual clinical triage. Given it cannot reliably differentiate between mild and severe COVID-19, any reassurance provided is neither dependable nor evidence based. For example, the CDC coronavirus symptom checker still advises the 65 years old with a 7-day history of persistent fever and worsening cough to remain at home without contacting their healthcare service provider. In the absence of any discoverable quality and safety studies supporting its use, it would seem prudent to discontinue the use of the CDC symptom checker until the CDC can prove its efficacy and safety, as would be the case for any diagnostic test.

Singapore remains consistent in its approach to the clinical management of COVID-19.21 All suspected cases of COVID-19 are clinically assessed by a physician and followed-up by primary care.22 This is consistent with the WHO technical guidelines and clinical recommendations for triage and management of COVID-19.19 During our previous analysis in April 2020, Singapore was suffering a surge of COVID-19 infections. The symptom checker was then set to advise the young, non-breathless patient with no comorbidities and a short duration of illness to self-isolate and contact the public health clinic if symptoms worsened or had not improved by day 4.4 This remains a compromise to normal clinical care but may have achieved a low-risk reduction in healthcare burden. Now, with SARS-CoV-2 prevalence less than in April 2020, all suspected cases are clinically assessed.

Unlike the CDC and Singapore COVID-19 symptom checkers, the NHS ‘111’ symptom checkers have a clear and critical role in the national clinical response to COVID-19.8–11 Current advice in the UK to the public

<table>
<thead>
<tr>
<th>Triage criteria</th>
<th>Risk-based symptom checker</th>
<th>Severity-based symptom checker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of symptoms</td>
<td>Duration of symptoms affects triage outcomes. Patients with symptoms over 4 days are always triaged to further care.</td>
<td>Duration of symptoms does not alter triage outcomes. No length of illness leads to triage in to further healthcare contact.</td>
</tr>
<tr>
<td>Age</td>
<td>Age affects triage outcomes. Cases over the age of 65 years are always triaged on for further healthcare contact. During times of low SARS-CoV-2 healthcare burden, age is removed as a restriction to further clinical assessment.</td>
<td>Age affects triage outcomes. Cases over the age of 65 years are now triaged on to further healthcare contact.</td>
</tr>
<tr>
<td>Comorbidity</td>
<td>Comorbidity affects triage outcomes. Cases with any comorbidity are triaged on to further care.</td>
<td>Comorbidity affects triage outcomes. Cases with moderately severe comorbidities are triaged on for further healthcare contact.</td>
</tr>
<tr>
<td>Shortness of breath</td>
<td>Any degree of shortness of breath is triaged straight to the emergency department.</td>
<td>Patients with severe breathlessness are triaged to the emergency department. Patients with mild to moderate shortness of breath are advised to stay home with no clinical follow-up.</td>
</tr>
</tbody>
</table>

*NHS Scotland only triages cases with comorbidities that are on the shielding category list for further care. NHS Northern Ireland only triages cases with immunosuppression or conditions that have become more difficult to control since symptoms began. NHS Wales relies on shielding categories and also includes diabetes and pregnancy with heart conditions. NHS England use shielding categories and also consider immunosuppression, diabetes, heart disease, respiratory disease, kidney failure, liver disease or neurological disease.
diversion of patients away from actual clinical triage, the NHS 111 symptom checkers are likely to be contributing to the UK’s poor pandemic response, including the high morbidity and mortality. Also of growing concern is the impact the NHS 111 symptom checkers are likely to have on the resilience of society to tolerate background levels of SARS-CoV-2 and post-pneumonia complications (eg, long COVID-19) by delaying presentation of COVID-19 pneumonia to timely, appropriate medical care.

Given the NHS ‘111’ symptom checkers have ventured into diagnosis and, arguably, clinical management, are currently gatekeepers to further healthcare access, fail to reliably triage severe COVID-19 on to further care, fail to account for age as a risk factor (except NHS Wales) and are likely to miss COVID-19 mimickers such as bacterial pneumonia, considerable improvements are needed to render the current NHS 111 COVID-19 symptom checkers fit for purpose.

Future Direction

The use of symptom checkers as part of the national clinical care pathway for COVID-19 (and future pandemics) requires considerably more research and validation.29 Currently, none of the symptom checkers pose a viable option in replacing clinical triage, and as such, effort should focus on resourcing clinical triage services.

Data relating to the use of symptom checkers and the effects these have on future healthcare burden have not yet been analysed. At an operational level, the possibility of severity-based symptom checkers leading to an increase in healthcare burden, including an increase in high-dependency admissions, should sponsor caution and an urgent review of any care pathways depending on such forms of patient-led triage. Our analysis suggests, severity-based COVID-19 symptom checkers (such as the NHS 111 or CDC versions) are likely to increase the healthcare burden associated with the pandemic (in comparison to clinically led, remote triage).

There is the equally challenging obstacle of national versus local triage to overcome. The current NHS 111 symptom checker triages nationally using the same referral thresholds. This may have contributed to the disproportionate healthcare activity across the UK.27 Where a national symptom checker is ‘set’ to respond to critical demand in, for example, London, those using the symptom checker in an area of low demand, for example, the Lake District, will also be held to the same, compromised and rationed access to healthcare. This goes against the principles of triage, in that triage decisions must be responsive to resource availability. It is not justifiable—orlogistically savvy—to ration access to healthcare preemptively or without a definitive need to.

In the short term—pending further safety studies—a ‘risk-based symptom checker’ may provide a possible low-risk solution to signposting potential COVID-19 cases, under pandemic conditions. The usual standard remains an actual clinical assessment, but where healthcare resources are insufficient for such a standard of
care, an untested symptom checker that can be adjusted in response to risk and demand would be preferable to an untested symptom checker attempting to determine clinical severity from an automated algorithm. A national symptom checker may still have a role in risk stratifying, and with the benefit of postcode localisation, there may be an ability to adjust the ‘risk necessary to take’ more accurately and based on local demands. Such a national service providing local risk stratification must be dynamic and responsive to demands, be under constant data collection and review, and be viewed as a considerable compromise to usual standards of care.

Whatever future version of COVID-19 symptom checkers manifest, they must be designed with the intention of detecting progressive COVID-19 or those at risk of severe disease, not designed with the intention of preventing healthcare contact. Triage itself is not resource saving. But the effort invested in the triage process yields high returns when cases of progressive COVID-19 pneumonia are detected early enough to avoid costly, protracted and complicated admissions. Triage systems must be viewed for what they are: an opportunity to maximise the use of available resources to prevent death, avoid disability and improve healthcare resilience.

CONCLUSION
The use of symptom checkers to triage patients during a pandemic or major incident is novel and untested. Our case simulation study provides little reassurance for their ongoing use. Even during a period of low healthcare burden, the symptom checkers deployed by both the USA and UK maintained a high threshold for onward referral. Neither symptom checker reliably triaged treatable, time-critical cases in to healthcare contact or follow-up and were unable to consistently differentiate mild from severe COVID-19. Of further concern, age is not factored in the triage decisions of the NHS 111 symptom checkers (except NHS Wales)—an unusual practice in clinical triage and well-below national and international standards of care.

Beyond the patient safety concerns, there is no evidence that COVID-19 symptom checkers reduce the healthcare burden associated with the pandemic. Our results suggest, by delaying the presentation of time-critical cases to medical care, it is quite likely the NHS 111 symptom checkers increase the healthcare burden associated with the SARS-CoV-2 pandemic in the UK.

In the absence of any safety, efficacy or quality assurance studies to support the use of symptom checkers as triage tools, our results necessitate a recommendation for the NHS 111 symptom checker and CDC coronavirus symptom checker to be subject to further analysis prior to their ongoing use in COVID-19 clinical care pathways. The stakes of patient triage are simply too high, and the reliability of symptom checkers is simply too poor, to justify their ongoing use.

Contributors All authors contributed to the conception, methodology and analysis and did final review and edit. DG undertook the majority of the write-up. FM undertook edits and contributed to the write-up and revision of the manuscript. DG is the guarantor of the study.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

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.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID id
Daniel Goyal http://orcid.org/0000-0003-0418-8859

REFERENCES


