

## COMMENTARY

**‘You should see a doctor’, said the robot: Reflections on a digital diagnostic device in a pandemic age**CHRISTOFFER BJERRE HAASE<sup>1,2</sup> , MARGARET BEARMAN<sup>2</sup>, JOHN BRODERSEN<sup>1,3</sup>, KLAUS HOEYER<sup>1</sup>  & TORSTEN RISOR<sup>1,4</sup><sup>1</sup>Department of Public Health, University of Copenhagen, Denmark, <sup>2</sup>Centre for Research in Assessment and Digital Learning (CRADLE), Deakin University, Australia, <sup>3</sup>Primary Health Care Research Unit, Region Zealand, Denmark, <sup>4</sup>Department of Community Medicine, UiT The Arctic University of Norway, Norway**Abstract**

**Aims:** In three days at the beginning of the COVID-19 pandemic, the Copenhagen Emergency Medical Services developed a digital diagnostic device. The purpose was to assess and triage potential COVID-19 symptoms and to reduce the number of calls to public health-care helplines. The device was used almost 150,000 times in a few weeks and was described by politicians and administrators as a solution and success. However, high usage cannot serve as the sole criterion of success. What might be adequate criteria? And should digital triage for citizens by default be considered low risk? **Methods:** This paper reflects on the uncertain aspects of the performance, risks and issues of accountability pertaining to the digital diagnostic device in order to draw lessons for future improvements. The analysis is based on the principles of evidence-based medicine (EBM), the EU and US regulations of medical devices and the taxonomy of uncertainty in health care by Han et al. **Results:** Lessons for future digital devices are (a) the need for clear criteria of success, (b) the importance of awareness of other severe diseases when triaging, (c) the priority of designing the device to collect data for evaluation and (d) clear allocation of responsibilities. **Conclusions: A device meant to substitute triage for citizens according to its own criteria of success should not by default be considered as low risk. In a pandemic age dependent on digitalisation, it is therefore important not to abandon the ethos of EBM, but instead to prepare the ground for new ways of building evidence of effect.**

**Keywords:** Chatbot, COVID-19, diagnostics, digital diagnostic device, digitalisation, evidence-based medicine, regulations, acute primary health care, uncertainty, symptom checker

**Introduction**

Right at the beginning of the COVID-19 pandemic, the Copenhagen Emergency Medical Services (CEMS) developed a digital diagnostic device to assess symptoms of infection [1]. In just three days, this device was launched in the Capital Region of Denmark. A week later, the device was implemented nationwide in Denmark and was used more than 90,000 times in its first week and almost 150,000 times in the second week. The purpose of the device was presented as twofold [1–4]: (a) to help individual citizens in assessing whether symptoms they

experienced were potentially COVID-19 related and to advise them when and where to seek further medical assistance; and (b) to reduce the number of calls to public health-care helplines. Immediately, politicians and administrators described this digital diagnostic device in press releases as ‘a new digital solution from the Regions’, which ‘has been a great success’ [4,5]. However, frequent use of the device may not equal alleviation of pressure on the helplines; ‘see a doctor’ is indeed likely advice provided by the device. If high usage is not an adequate criterion of success, what would be? And should such a device by default be considered low risk?

Correspondence: Christoffer Bjerre Haase, Section of General Practice, Department of Public Health, University of Copenhagen, Center for Sundhed og Samfund (CSS), Øster Farimagsgade 5, building 24, entrance Q, ground and 1st floor, P.O.B. 2099, 1014 Copenhagen K, Denmark. E-mail: c.haase@sund.ku.dk

Date received 13 August 2020; reviewed 30 September 2020; 30 October 2020; accepted 15 November 2020

© Author(s) 2020

Article reuse guidelines: [sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/1403494820980268

[journals.sagepub.com/home/sjp](https://journals.sagepub.com/home/sjp)



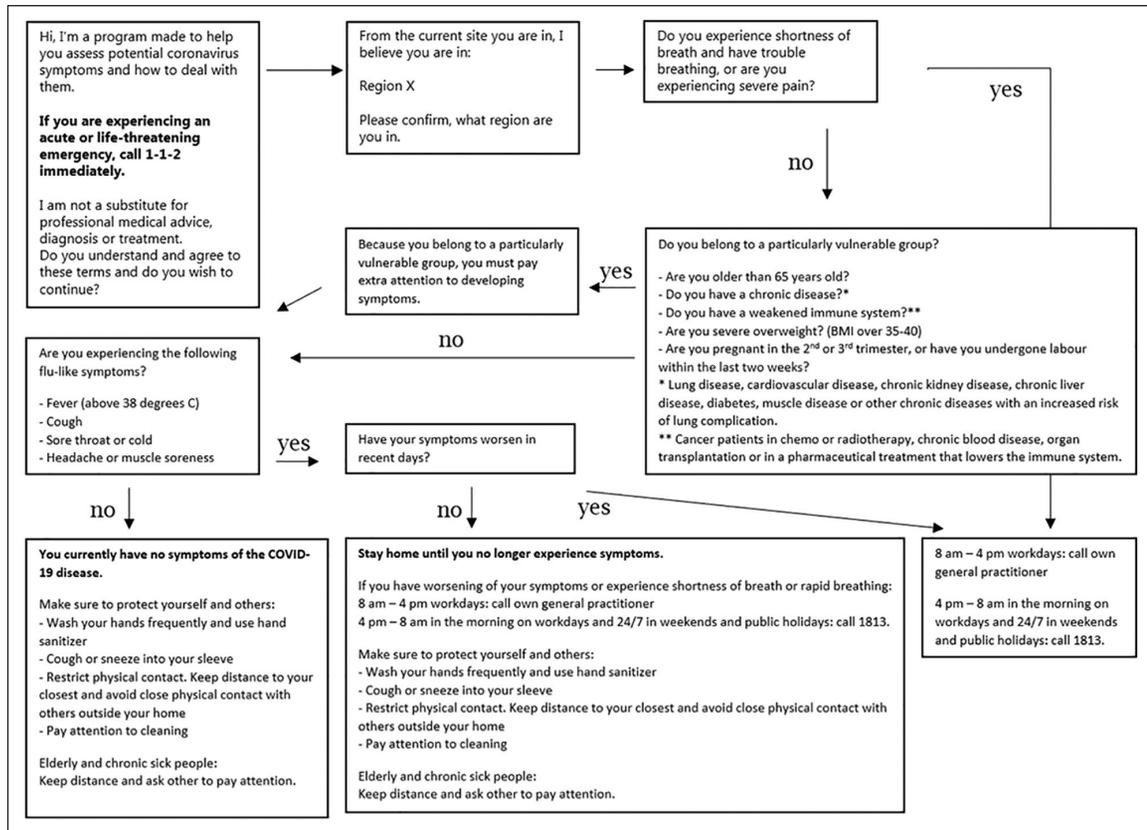


Figure 1. The decision tree of the Danish COVID-19 'chatbot', by the Copenhagen Emergency Medical Services (CEMS). The figure is based on the first version of the launched device and the decision tree published by CEMS elsewhere [1].

With a rapid turn to digital solutions in a pandemic age, there is a need to move beyond the hype [6–8] and base application of possibilities on clear criteria for evidence of effect. This paper reflects on the uncertain aspects of the performance, safety risks and accountability of the digital diagnostic device in order to draw such lessons for the future.

### Digital devices in health care and the absence of evidence

CEMS initiated the development of the device on 12 March 2020, and policy support and research grants have allowed for an ongoing upgrade of the device [9]. Therefore, its concrete content elements have changed over time. With support of their software provider – Microsoft – CEMS designed the device as a simple decision tree (Figure 1) inspired by the decision trees normally used in their telephone assessments of symptoms [1]. The helplines are the entry point into the primary and secondary health-care services in Denmark outside of regular GP office hours.

The initiative represents a digital response to the challenges of diagnostics and triage. When patients facing serious illness decide which action to take, it can have the gravest implications. They must make the right choice. However, for digital devices, clinical

standards as known from evidence-based medicine (EBM) are not in place [10]. Unlike drugs, software must continuously be developed and prove its efficiency through use. In response to our inquiries, we have been informed that this particular device was not subject to EU regulations of medical devices, as it was declared not to be 'intended by its manufacturer to be used specifically for diagnostic or therapeutic purposes' [11]. The Food and Drug Administration (FDA) responded that they would categorise the device as 'lower risk' for which the FDA does not intend to enforce requirements under the FDA&C Act at the moment [12].

The device was also presented as a 'chatbot', although it does not conduct an actual written or oral conversation with the user, which usually defines chatbots [13,14]. The device is more accurately classified as a computerised diagnostic decision support, widely known as a symptom checker [15]. These are typically available online or as apps [16]. Symptom checkers have been used in other countries to manage COVID-19 [17]. So far, however, uncertainty prevails about their effects [16,18]. A systematic review found variation between symptoms checkers but relatively strong evidence that they are inferior in triage compared to health professionals [18]. They are mostly more cautious, which should work against

the goal of minimising pressure on helplines. The review found little evidence of cost-effectiveness and patient compliance [18].

### Lessons for future digital devices

With little evidence and high levels of urgency, it is important to help developers to be better prepared for digital responses to pandemic threats. Since diagnostic uncertainty is at the heart of triage function, we draw on the taxonomy of uncertainty in health care by Han et al. [19] and combine it with the principles of EBM and the EU and US regulations of medical devices.

Based on the available information also graciously offered by the developers about the purposes, design and success of the device, the following elements contain lessons for the future:

(1) *Clear criteria of success*

The stated criterion of success is 'frequent use', but this is insufficient when a device intervenes in the diagnostic process. Criteria should be based on assessments of risks and benefits that includes long-term consequences – also for other affected actors. In this case, the appropriateness of recommendations should be a criterion of success.

(2) *What if the users don't have COVID-19 but some other serious condition?*

The scope of the device is limited to detection of COVID-19 (Figure 1). Nonetheless, users may be experiencing symptoms of other potentially acute, fatal and relatively easily treatable diseases, such as meningitis or acute coronary syndrome, which are not appropriately assessed in the device. During a pandemic, it is important not to expose users to the risk of overlooking symptoms from severe diseases not related to COVID-19. Thus, interventions such as this device should not by default be considered low risk.

(3) *Lack of possibility to evaluate the impact of the device*

The device collects minimum feedback from users in a form of comments and answers about their demographic. Since software should prove its effect through use, it is important to design it with a plan for appropriate data collection to document achievement of the stated criteria of success. This device did not collect sufficient data to evaluate its effect. Therefore, the impact of the device should not be considered low risk.

(4) *No accountability for consequences of use of the device*

Clear allocation of responsibilities could stimulate additional critical thoughts during the development phase and provide users with appropriate

Table I. Suggested approach before implementation of a non-regulated digital diagnostic device.

- 
- (1) **Information:** Be transparent and explicit by providing easily understandable and easily accessible information to the users and other relevant actors about aspects of the device in following order:
  - (2) **Aim:** State all aims of the device, including clearly defined criteria of achieving those aims.
  - (3) **Safety risks and implications for others:** State any individual or societal safety risks of using the device, as well as consequences for any actors who may be directly or indirectly affected by the device.
  - (4) **Before implementation:** Describe how aims, safety risks and implications for others have been investigated and evaluated before implementation. Make it clear if any aspects of aims, safety or implications for others have not been investigated and evaluated.
  - (5) **Monitoring and re-evaluation:** Describe the plan of how to measure/monitor and re-evaluate each of the stated aspects of aims, safety and implications for others. Make it clear if an aspect is planned not to be measured/monitored and re-evaluated.
  - (6) **Contact:** Provide contact information for further questions and suggestions for improvements.
  - (7) **Accountability:** Clarify which institution/company/organisation is responsible for steps 1–6.
- 

The information provided by the approach outlined above should be publically accessible along with the device to inform users and other affected actors. The approach is based on principles of evidence-based medicine, the EU and US regulations of medical devices and the taxonomy of uncertainty in health care by Han et al. [10,12,19].

contact points in case of concerns or challenges. Accountability may prompt politicians, administrators and health-care workers to ensure the quality control of unregulated medical devices. Finally, accountability has a symbolic value with respect to maintaining public trust in pandemic situations (and clinical matters in general).

Despite these concerns, it is remarkable how quickly the device was developed and implemented, especially at the beginning of the outbreak in Denmark, when every institution and workplace were in a state of emergency. Further, CEMS cannot be accountable for the many uncertain aspects of a new disease, including symptoms, prognosis and treatment. Finally, if CEMS is forced to take action on overloaded helplines, they cannot be required to meet the same demands that commercial devices do or guidelines that takes years to produce and formulate. The device should be seen in this specific context, which was – and is – extraordinary.

However, based on the mentioned lessons, a device meant to substitute triage for citizens should not by default be considered low risk. In a pandemic age dependent on digitalisation, developers, initiators and health authorities could take advantage of the overall principles of EBM and medical device regulations (Table I). The suggested step-wise approach does not

require many resources. Sometimes you have to build the boat while sailing [20,21], but it only sustains the need for gathering evidence along the way too, while taking into account how privacy concerns and user agreements may significantly hinder or bias the collection of the aimed evidence in practice. We therefore need to update existing scientific and regulatory principles and facilitate their easy use in order to ensure that future devices have the intended performance and safety.

### Acknowledgements

We thank the developers from the Copenhagen Emergency Medical Services for sharing insights into the process and reasoning of making the device during a busy and challenging period of the COVID-19 pandemic. We also thank Markéta Jurovská for her support and helpful discussions on digital strategies and solutions, Rola Ajjawy for her constructive feedback on an earlier version of this manuscript, and two referees.

### Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

### Funding

The authors disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement Number 682110).

### ORCID iDs

Christoffer Bjerre Haase  <https://orcid.org/0000-0003-3895-518X>

Klaus Hoeyer  <https://orcid.org/0000-0002-2780-4784>

### References

- [1] Jensen TW, Holgersen MG, Jespersen MS, et al. Strategies to handle increased demand in the COVID-19 crisis a corona telephone hotline and a web-based self-triage system. *Prehosp Emerg Care*. Epub ahead of print 9 October 2020. DOI: 10.1080/10903127.2020.1817212.
- [2] Bach D. How international health care organizations are using bots to help fight COVID-19, <https://news.microsoft.com/transform/how-international-health-care-organizations-are-using-bots-to-help-fight-covid-19/> (accessed 28 July 2020).
- [3] Region Hovedstaden had a test chatbot developed for self assessment of corona symptoms, <https://en.proactive.dk/kunder/offentlig/region-hovedstaden> (accessed 28 July 2020).
- [4] Ny digital løsning fra regionerne skal aflaste læger i hele landet, <https://www.regioner.dk/services/nyheder/2020/marts/ny-digital-loesning-fra-regionerne-skal-aflaste-laeger-i-hele-landet> (accessed 28 July 2020).
- [5] Næsten 150.000 danskere har brugt chatbot til at vurdere symptomer, <https://www.regioner.dk/services/nyheder/2020/april/naesten-150000-danskere-har-brugt-chatbot-til-at-vurdere-symptomer> (accessed 28 July 2020).
- [6] Cassell EJ. The sorcerer's broom: medicine's rampant technology. *Hastings Cent Rep* 1993;23:32.
- [7] Hoeyer K. Data as promise: reconfiguring Danish public health through personalized medicine. *Soc Stud Sci* 2019;49:531–55.
- [8] Wachter R. *The digital doctor: hope, hype, and harm at the dawn of medicine's computer age*. New York: McGraw-Hill, 2015.
- [9] Novo Nordisk Fonden supports 11 new, important projects in connection with the coronavirus epidemic, <https://novonordiskfonden.dk/en/news/novo-nordisk-fonden-supports-11-new-important-projects-in-connection-with-the-coronavirus-epidemic/> (accessed 28 July 2020).
- [10] Guyatt GH, Oxman AD, Schünemann HJ, et al. GRADE guidelines: a new series of articles in the *Journal of Clinical Epidemiology*. *J Clin Epidemiol* 2011;64:380–2.
- [11] European Commission. Guidance document medical devices – scope, field of application, definition – qualification and classification of stand alone software – MEDDEV 2.1/6, <https://ec.europa.eu/docsroom/documents/17921/attachments/1/translations> (accessed 4 December 2020).
- [12] Food and Drug Administration. Policy for device software functions and mobile medical applications, <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/policy-device-software-functions-and-mobile-medical-applications> (accessed 4 December 2020).
- [13] Nadarzynski T, Miles O, Cowie A, et al. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. *Digit Health* 2019;5:205520761987180.
- [14] Oxford Advanced Learner's Dictionary. Chatbot noun – definition, pictures, pronunciation and usage, <https://www.oxfordlearnersdictionaries.com/definition/english/chatbot?q=chatbot> (accessed 28 July 2020).
- [15] Fraser H, Coiera E and Wong D. Safety of patient-facing digital symptom checkers. *Lancet* 2018;392:2263–4.
- [16] Semigran HL, Linder JA, Gidengil C, et al. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ* 2015;351.
- [17] Gasser U, Ienca M, Scheibner J, et al. Digital tools against COVID-19: taxonomy, ethical challenges, and navigation aid. *Lancet Digit Health* 2020;2:e425–34.
- [18] Chambers D, Cantrell AJ, Johnson M, et al. Digital and online symptom checkers and health assessment/triage services for urgent health problems: systematic review. *BMJ Open* 2019;9:1–13.
- [19] Han PKJ, Klein WMP and Arora NK. Varieties of uncertainty in health care: a conceptual taxonomy. *Med Decis Mak* 2011;31:828–38.
- [20] Young A. Ethnography and bioethics: boat repair at sea. *Med Health Care Philos* 2002;5:91–3.
- [21] Green S, Carusi A and Hoeyer K. Plastic diagnostics: the remaking of disease and evidence in personalized medicine. *Soc Sci Med*. Epub ahead of print 18 May 2019. DOI: 10.1016/j.socscimed.2019.05.023.