

Digital and online symptom checkers and assessment services for urgent care to inform a new digital platform: a systematic review

Duncan Chambers,* Anna Cantrell, Maxine Johnson, Louise Preston, Susan K Baxter, Andrew Booth and Janette Turner

School of Health and Related Research (SchARR), University of Sheffield, Sheffield, UK

*Corresponding author d.chambers@sheffield.ac.uk

Declared competing interests of authors: Andrew Booth is a member of the National Institute for Health Research Complex Review Support Unit Funding Board.

Published August 2019

DOI: 10.3310/hsdr07290

Scientific summary

Digital/online symptom checkers for urgent care

Health Services and Delivery Research 2019; Vol. 7: No. 29

DOI: 10.3310/hsdr07290

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

This systematic review deals with digital and online symptom checkers and assessment services used by patients seeking guidance about an urgent health problem. These services generally provide people with possible diagnoses and/or suggest a course of action based on their reported symptoms. NHS England intends to introduce a digital platform (NHS111 Online) to operate alongside the NHS111 telephone service.

The 'digital 111' service is seen as key to reducing demand for the telephone 111 service, enabling resources to be redirected to support 'integrated urgent and emergency care systems', as outlined in the NHS *Five Year Forward View* (NHS England. *Five Year Forward View*. Leeds: NHS England; 2014) and its 2017 update *Next Steps on the NHS Five Year Forward View* (NHS England. *Next Steps on the NHS Five Year Forward View*. Leeds: NHS England; 2017).

Thus, it is anticipated that a digital 111 platform will help to manage demand and increase efficiency in the urgent and emergency care system, complementing the agenda of locally based Sustainability and Transformation Partnerships. However, there is a risk of increasing demand, duplicating health-care contacts and providing advice that is not safe or clinically appropriate.

In 2017, NHS England carried out pilot evaluations of different digital or online triage systems in four regions of England (NHS England. *NHS111 Online Evaluation*. Leeds: NHS England; 2017). The evaluations aimed to assess whether or not digital/online triage was acceptable to users and connected them to appropriate clinical care. The full report of these evaluations is not yet published at the time of writing this report.

Objectives

The objective of this systematic review was to inform further developments of the proposed digital platform by summarising and critiquing the previous research in this area, from both the UK and overseas.

Methods

Focused searches of seven databases were performed, supplemented by phrase searching for names of symptom checker systems (e.g. webGP, iTriage) and citation searches of key included studies. The bibliographic databases searched were MEDLINE, EMBASE, The Cochrane Library, CINAHL (Cumulative Index to Nursing and Allied Health Literature), HMIC (Health Management Information Consortium), Web of Science and the Association of Computing Machinery (ACM) Digital Library, from inception up to April 2018.

Inclusion and exclusion criteria are listed in the sections below.

Population

Members of the general population, including adults and children, who are looking for information online or digitally to address an urgent health problem, which includes issues arising from both acute illness and long-term chronic illness. Non-urgent problems were excluded.

Intervention

The intervention is made up of online or digital services that are designed to assess symptoms, provide health advice and direct patients to appropriate services. This reflects the role of the NHS111 telephone service. Services that provide only health advice were excluded, as were those that offer treatment (e.g. online cognitive-behavioural therapy services).

Comparator

The 'gold standard' comparator is the current practice of telephone assessment (e.g. NHS111) or face-to-face assessment (e.g. in a general practice, urgent-care centre or emergency department). However, studies with other relevant comparators (e.g. comparative performance in tests or simulations) or with no comparator were included if they addressed the research questions.

Outcomes

The main outcomes of interest were:

- safety (e.g. any evidence of adverse events arising from following or ignoring advice from online/digital services)
- clinical effectiveness (any evidence of clinical outcomes associated with the use of online/digital services)
- cost-effectiveness (including costs and resource use)
- accuracy – this refers to the ability to provide a correct assessment and the ability to distinguish between high- and low-acuity (urgency and level of care needed) problems and, hence, direct patients to appropriate services, avoiding over- or undertriage. Direction to appropriate services is also referred to as patient disposition
- impact on service use/diversion (including possible multiple contacts with health services)
- compliance with advice received
- patient/carer satisfaction
- equity and inclusion (e.g. barriers to access, characteristics of patients using the service compared with the general population).

This list is not exhaustive and other relevant outcomes from included studies were extracted.

Study design

We did not restrict inclusion by study design (and included relevant audits or service evaluations in addition to formal research studies) but included studies had to evaluate (quantitatively or qualitatively) some aspect of an online/digital service. Studies were not excluded on grounds of quality. Studies of health-care system from any high-income country were eligible for inclusion.

Excluded

The following types of studies were excluded from the review:

- studies that merely describe services without providing any quantitative or qualitative outcome data
- conceptual papers and projections of possible future developments
- studies conducted in low- or middle-income countries' health-care systems.

Screening studies for inclusion, data extraction and quality assessment (using appropriate tools for different study designs) were carried out by one reviewer with a 10% sample checked for accuracy and consistency. Final decisions on study inclusion were taken by consensus of the review team. To characterise the included digital and online systems as interventions, we identified studies reporting on a particular system and extracted data from all relevant studies using a modification of the Template for Intervention Description and Replication (TIDieR) checklist that we designated Template for Intervention Description for Systems for Triage (TIDieST).

A narrative synthesis of the included studies was performed and structured around the predefined research questions and key outcomes. The overall strength of evidence for each outcome was classified as 'stronger', 'weaker', 'conflicting' or 'insufficient', based on study numbers and design.

Results

We included 29 publications describing 27 studies (nine from the UK). Studies were published between 2006 and 2018 and were diverse in terms of their design and methodology. The overall strength of the evidence base varied between outcomes, but in absolute terms the evidence was weak, being based largely on observational studies and with a substantial component of grey literature.

We were able to extract data on eight systems using the TIDieST checklist. When appropriate, data from multiple studies were combined in one checklist. Some of these systems appear to be no longer in use and we were unable to obtain sufficient data for some systems currently being used and/or evaluated. Four of the included systems were designed to cover a full range of symptoms and four others covered a more limited range (i.e. three for influenza-like illness and one for minor respiratory symptoms). Most systems were accessed through web pages, often linked to health-care providers or government organisations. The 'babylon check' system was the main exception because it was designed for access using a smartphone application (app). Published research studies provided relatively little detail about the systems, possibly reflecting a need for commercial confidentiality (Middleton K, Butt M, Hammerla N, Hamblin S, Mehta K, Parsa A. *Sorting Out Symptoms: Design and Evaluation of the 'Babylon Check' Automated Triage System*. London: Babylon Health; 2016).

We found little evidence to support the hypothesis that digital and online symptom checkers are detrimental to patient safety. However, the studies that reported patient safety outcomes were mostly short term and involved relatively small samples. Some were limited to people with specific types of symptoms (e.g. influenza-like illness or respiratory symptoms) and others recruited from specific population groups (e.g. students), meaning that participants are not representative of all users of urgent-care services.

Study participants generally expressed high levels of satisfaction, albeit in uncontrolled studies. For example, in the NHS England pilot evaluation 70–80% of users were satisfied with their experience at each of the pilot sites. This was based on a sample of > 1500 users, the majority of whom were involved in the London pilot using the 'babylon check' app (NHS England. *NHS111 Online Evaluation*. Leeds: NHS England; 2017).

The evidence from the studies included in the review suggests that digital and online systems have yet to achieve a high level of accuracy in the assessment of specific conditions. This finding applies to both 'general purpose' symptom checkers and those symptom checkers that are limited to particular conditions. However, some of the included studies did not recruit representative populations and others were based on standardised vignettes rather than real-world data. Studies that compared symptom checkers with health professionals tended to use the doctors' clinical diagnosis as the reference standard.

Results for the accuracy of triage were inconsistent between studies. The studies used similar methods to those evaluating diagnostic accuracy. A review of 23 symptom checkers found that triage level was appropriate overall for 57% of patient vignettes, with considerable variation between systems and conditions (Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ* 2015;**351**:h3480). The NHS England evaluation also found differing levels of agreement with clinical expert opinion across the four systems evaluated. In general, algorithm-based triage tended to be more risk averse than the triage of health professionals.

We also found inconsistent evidence on effects on service use, but there was some indication that symptom checkers can influence the pattern of service use. The strongest evidence came from a randomised trial of an intervention that was specifically designed to promote self-care and covered

respiratory symptoms only (Little P, Stuart B, Andreou P, McDermott L, Joseph J, Mullee M, *et al.* Primary care randomised controlled trial of a tailored interactive website for the self-management of respiratory infections (Internet Doctor). *BMJ Open* 2016;**6**:e009769). In this study, the intervention group had fewer contacts with doctors (but more contact with NHS Direct) than the control group, despite having a longer duration and greater severity of illness. The NHS England evaluation found a small shift towards self-care with digital triage compared with telephone triage, and another study reported that the webGP symptom checkers diverted 18% of patients from requesting a general practitioner appointment (Madan A. *WebGP: The Virtual General Practice*. London: Hurley Group; 2014). These are findings from the grey literature and should be interpreted with caution.

There was very limited evidence on patients' reactions to online triage advice and whether the patients follow the advice or seek further help or information. Preliminary findings from the NHS England evaluation suggest that patients may be more likely to seek further advice for more urgent conditions, but further confirmation is required.

Over half of the included studies considered equity and inclusion either directly or by comparing users with non-users of digital triage systems. Not surprisingly, there was a clear consensus that younger and more highly educated people are more likely to use these digital/online services, whereas older and less educated patients are more likely to prefer telephone or face-to-face contact with health-care practitioners. This could have implications for health equity if urgent-care pathways prioritise (or appear to prioritise) requests originating from digital sources.

Limitations

Findings from symptom checker systems for specific conditions may not be applicable to more general systems and vice versa. We have also included studies of symptom checkers as part of electronic consultation systems in general practice, which is, again, a slightly different setting from a general digital 111 service. Most studies were screened by one reviewer.

Conclusions

The current evidence base covers diverse interventions, study designs and outcomes. Major uncertainties surround the probable impact of digital 111 services on most of the important outcomes, but precedent suggests that once introduced their use will increase rapidly. It will be important to monitor and evaluate these services using all available data sources and by commissioning high-quality research.

The studies included in the review suggest that there is a high level of uncertainty about the impact of digital 111 on the urgent-care system and the wider health-care system. The health service may need to respond to short-term increases (or decreases) in demand and/or shifts from one part of the system to another. This may increase pressure on the system, at least in the short term. In the longer term, if the use of the 111 telephone service decreases as planned, there may be opportunities to redeploy staff to fill other roles in the urgent and emergency care system.

Priorities for research include comparisons of different systems, rigorous economic evaluations based on real-world data, investigations of the pathways followed by patients using the new service and investigations of the role of behaviour change theory in the development and implementation of symptom checkers.

Study registration

The study is registered as PROSPERO CRD42018093564.

Funding

Funding for this study was provided by the Health Services and Delivery Research programme of the National Institute for Health Research.

Health Services and Delivery Research

ISSN 2050-4349 (Print)

ISSN 2050-4357 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HS&DR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hsdr. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Services and Delivery Research* journal

Reports are published in *Health Services and Delivery Research* (HS&DR) if (1) they have resulted from work for the HS&DR programme or programmes which preceded the HS&DR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HS&DR programme

The Health Services and Delivery Research (HS&DR) programme, part of the National Institute for Health Research (NIHR), was established to fund a broad range of research. It combines the strengths and contributions of two previous NIHR research programmes: the Health Services Research (HSR) programme and the Service Delivery and Organisation (SDO) programme, which were merged in January 2012.

The HS&DR programme aims to produce rigorous and relevant evidence on the quality, access and organisation of health services including costs and outcomes, as well as research on implementation. The programme will enhance the strategic focus on research that matters to the NHS and is keen to support ambitious evaluative research to improve health services.

For more information about the HS&DR programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/hsdr>

This report

The research reported here is the product of an HS&DR Evidence Synthesis Centre, contracted to provide rapid evidence syntheses on issues of relevance to the health service, and to inform future HS&DR calls for new research around identified gaps in evidence. Other reviews by the Evidence Synthesis Centres are also available in the HS&DR journal. The research reported in this issue of the journal was funded by the HS&DR programme or one of its preceding programmes as project number 16/47/17. The contractual start date was in April 2018. The final report began editorial review in June 2018 and was accepted for publication in October 2018. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HS&DR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Chambers *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

NIHR Journals Library Editor-in-Chief

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

NIHR Journals Library Editors

Professor John Powell Chair of HTA and EME Editorial Board and Editor-in-Chief of HTA and EME journals. Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK, and Honorary Professor, University of Manchester, and Senior Clinical Researcher and Associate Professor, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals) and Editor-in-Chief of HS&DR, PGfAR, PHR journals

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Consultant Advisor, Wessex Institute, University of Southampton, UK

Ms Tara Lamont Director, NIHR Dissemination Centre, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Wellbeing Research, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk