

The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia

Michella G Hill, Moira Sim, Brennen Mills

The known: Australians are searching the internet for health care advice with increasing frequency. Symptom checkers are common online tools for obtaining diagnostic and triage advice.

The new: Thirty-six symptom checkers were evaluated with a range of clinical vignettes. The correct diagnosis was listed first in 36% of tests, and included among the first ten results in 58%; triage advice was appropriate in 49% of cases, including about 60% of emergency and urgent cases but only 30–40% of less serious case vignettes.

The implications: Symptom checkers may provide unsuitable or incomplete diagnostic or triage advice for users in Australia, resulting in inappropriate care advice.

Australians generally embrace online technology: in 2016–17, 86% of households had internet access¹ and about 89% of surveyed adults owned smartphones.² An abundance of material once accessible only to subject specialists is now more readily available on websites and via smartphone applications (apps), and widespread internet access has changed health information-seeking behaviour.³ About 80% of Australians report searching the internet for health information, and nearly 40% seek self-treatment advice online.^{4,5} An Australian study⁶ found that 34% of people attending emergency departments had searched the internet for information about their condition.

Symptom checkers (SCs) are algorithm-based programs that provide potential diagnoses and triage advice. Medical students routinely use diagnostic apps to support their education,⁷ and similar products are publicly available. SCs may not function as intended; they can encourage users to seek professional care even when self-care would be suitable, increasing the burden on health care systems.^{8,9} Not all SCs are affiliated with reputable organisations,³ and concerns have been raised about data security, privacy, the credentials of app authors and editors, and the accuracy of health information.^{9,10} The United States Food and Drug Administration recently launched a voluntary software pre-certification program aimed at providing assurance of effectiveness and transparency of product performance and safety.¹¹ In Australia, the Therapeutic Goods Administration oversees the regulation of medical software products and devices, but does not regulate SCs because they are not classified as medical devices.^{12,13}

Semigran and colleagues recently tested 23 free online SCs with standardised patient vignettes; they found that correct diagnoses headed the result list in 34% of evaluations, and appropriate triage advice was provided in 57% of cases.⁸ In a study in a British primary care clinic, a self-assessment tool recommended a more urgent level of care than the general practitioner in 56% of cases and less urgent advice for only 5% of consultations.¹⁴

The aim of our study was to investigate the diagnosis and triage performance of free SCs accessible in Australia.

Abstract

Objectives: To investigate the quality of diagnostic and triage advice provided by free website and mobile application symptom checkers (SCs) accessible in Australia.

Design: 36 SCs providing medical diagnosis or triage advice were tested with 48 medical condition vignettes (1170 diagnosis vignette tests, 688 triage vignette tests).

Main outcome measures: Correct diagnosis advice (provided in first, the top three or top ten diagnosis results); correct triage advice (appropriate triage category recommended).

Results: The 27 diagnostic SCs listed the correct diagnosis first in 421 of 1170 SC vignette tests (36%; 95% CI, 31–42%), among the top three results in 606 tests (52%; 95% CI, 47–59%), and among the top ten results in 681 tests (58%; 95% CI, 53–65%). SCs using artificial intelligence algorithms listed the correct diagnosis first in 46% of tests (95% CI, 40–57%), compared with 32% (95% CI, 26–38%) for other SCs. The mean rate of first correct results for individual SCs ranged between 12% and 61%. The 19 triage SCs provided correct advice for 338 of 688 vignette tests (49%; 95% CI, 44–54%). Appropriate triage advice was more frequent for emergency care (63%; 95% CI, 52–71%) and urgent care vignette tests (56%; 95% CI, 52–75%) than for non-urgent care (30%; 95% CI, 11–39%) and self-care tests (40%; 95% CI, 26–49%).

Conclusion: The quality of diagnostic advice varied between SCs, and triage advice was generally risk-averse, often recommending more urgent care than appropriate.

Methods

Symptom checkers

To identify the most prominent free online SCs in Australia, we searched for “symptom checker”, “medical diagnosis”, “health symptom diagnosis”, “symptom”, and “cough fever headache” (ie, flu-like symptoms) with five popular search engines (Google, Yahoo, Ask, Search Encrypt, Bing) between November 2018 and January 2019. The first three pages of results for each search engine were filtered to identify SCs that were publicly available, in English, for humans, and not restricted to specific medical conditions. The two major smartphone application distribution services (Google Play [Android] and App Store [Apple]) were also searched. As app store ranking algorithms place the most popular apps at the top of search results,⁹ we included the first fifteen listed apps for each store. SCs were characterised as to whether they requested demographic information, by the number of diagnoses provided in responses, and by whether they claimed to use artificial intelligence (AI) algorithms.

Patient vignettes

To assess the diagnostic and triage performance of the SCs, 30 patient vignettes from the study by Semigran and colleagues⁸ were adapted and supplemented by 18 new symptom-based scenarios, including several reflecting Australia-specific illnesses (Supporting Information, table 1). Clinical information for the

vignettes was drawn from medical resource websites and training material for health care professionals.^{8,15–17} The disposition advice provided by the SCs was allocated to one of four triage categories (emergency, urgent, non-urgent, self-care; **Box 1**). The “correct” diagnoses and triage advice for each vignette were confirmed by two GPs (including author MS) and an emergency department specialist with a combined 87 years’ clinical experience; disagreements related to contextual factors (eg, timing, local access to resources) were resolved by discussion. For our study, the triage category for deep vein thrombosis was deemed “urgent”, not “emergency”,⁸ to reflect the recent benefits of oral anticoagulant therapy. Medical conditions described were characterised as being common (85% of vignettes) or uncommon (15%: similar to the 10% reported for general practice presentations in Australia¹⁸) on the basis of Bettering the Evaluation and Care of Health (BEACH) general practice research data (2006–07 to 2015–16),¹⁹ Australian Institute of Health and Welfare emergency department care statistics (2016–17),¹⁷ and information provided by the Victorian government Better Health Channel website;¹⁶ these assessments were confirmed by the two GPs. For vignettes describing conditions with similar symptoms, several results could be deemed correct (**Supporting Information**, table 1).

SCs require users to enter their symptoms and, in some cases, other information. Accordingly, vignettes were summarised, in lay language, as the chief complaint and the core symptoms, and only this information was entered into the SC. To maintain consistency, one investigator entered the information for each vignette. As some SCs were specific for adult or childhood conditions, not every vignette could be evaluated in all SCs.

Symptom checker performance: diagnosis

Only the first ten results for each vignette were examined. Accurate diagnosis was defined as including the correct diagnosis as the top result, or as being among the top three or top ten potential diagnoses; “incorrect diagnosis” was defined as

the correct condition not being included in the top ten results. A vignette was deemed “unassessable” if the key symptoms could not be entered. Diagnostic performance was based on 1170 assessable vignettes across the 27 SCs providing diagnostic information (**Box 2**).

Symptom checker performance: triage

Appropriate triage advice can be more immediately useful to patients than a probable diagnosis,²⁰ as timely care can improve outcomes, particularly for emergency and urgent conditions. We defined triage accuracy by whether the SC provided triage advice concordant with our assessment. If several triage options were provided, the most urgent level was accepted. All triage advice was analysed, regardless of whether a diagnosis was provided. If triage advice depended on the offered diagnosis, the vignette was deemed “unassessable”. For example, Buoy Health triage advice for the appendicitis vignette ranged from “self-care” for the diagnosis of rotavirus infection to “emergency department” for appendicitis; that is, unambiguous triage advice could not be provided based on the entered symptoms alone. Triage performance was based on 688 assessable vignettes across the 19 SCs providing triage advice (**Box 2**).

Data analysis

Means with 95% confidence intervals (CIs) based on the binomial distribution were calculated in SPSS 25. Diagnostic and triage accuracy by diagnosis frequency (common, uncommon), vignette triage category, and SC characteristics (AI algorithms, demographic questions, maximum number of diagnosis results) were assessed in one-way analyses of variance (ANOVAs) and independent samples *t* tests. Normality of data distribution was assessed with the Shapiro–Wilk test; when non-normally distributed, the data were log transformed. Homogeneity of variance was assessed in Levene tests; when non-homogeneous, Welch ANOVA was used for analysis. A post hoc sensitivity analysis of triage performance excluded SCs that never recommended self-care. Individual SC performance was expressed as separate proportions of correct results for diagnosis and triage.

Ethics approval

Our study received ethics approval from the Edith Cowan University Human Research Ethics Committee on 27 March 2018 (reference, 20312).

Results

We identified 36 SCs for evaluation; ten provided both diagnostic and triage advice, 17 diagnostic information only, and nine triage advice only (**Supporting Information**, table 2).

Diagnostic performance

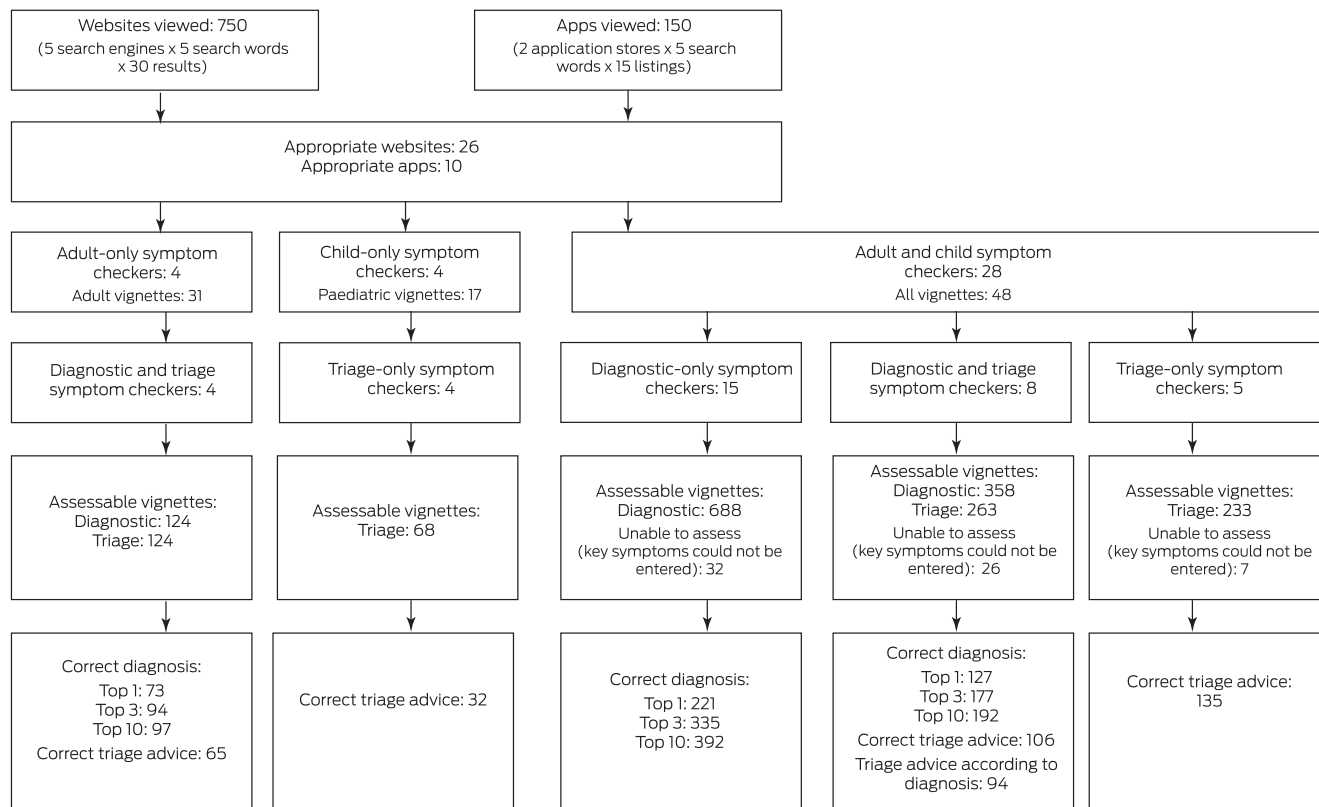
The 27 diagnostic SCs listed the correct diagnosis first in 421 of 1170 SC vignette tests (36%; 95% CI, 31–42%), among the top three results in 606 tests (52%; 95% CI, 47–59%), and among the top ten results in 681 tests (58%; 95% CI, 53–65%). The proportions of correct results listed first differed by patient condition in the vignette: emergency, 27% (95% CI, 21–34%); urgent, 45% (95% CI, 39–51%); non-urgent, 39% (95% CI, 33–47%); and self-care, 32% (95% CI, 25–42%). Correct diagnoses were more frequently listed first for vignettes describing common (42%; 95% CI, 37–50%) than for those describing uncommon conditions (4%; 95% CI, 1–7%) (**Box 3**; **Supporting Information**, table 3).

1 Triage categories and triage advice provided by symptom checkers

Triage category	Advice provided in symptom checkers
Emergency (requiring immediate medical care)	<ul style="list-style-type: none"> Call an ambulance or emergency (Australia: 000) Go to the emergency department Seek medical help immediately; this may be a life-threatening condition
Urgent (requiring medical attention within 24 hours)	<ul style="list-style-type: none"> See a GP immediately See a GP within 4 hours See a GP within 24 hours Go to an urgent care facility Call a nurse immediately
Non-urgent (see a health care provider in the near future)	<ul style="list-style-type: none"> Call a GP or primary health care provider See a GP or primary health care provider Go to a specialist* Go to a retail clinic Have an e-visit Minor injuries unit Walk-in clinic Telemedicine Speak to a nurse about your child
Self-care (assistance people provide themselves)	<ul style="list-style-type: none"> Stay at home Go to a pharmacist

* In Australia, this requires referral by a general practitioner; this advice is therefore not used on Australian website symptom checkers. ♦

2 Flowchart for selection and assessment of symptom checkers providing diagnostic or triage advice



The eight SCs that used AI algorithms according to the suppliers also requested demographic information. These SCs listed the correct diagnosis first (46%; 95% CI, 40–57%) more frequently than the other 19 diagnostic SCs (32%; 95% CI, 26–38%). Diagnostic accuracy was not related to the maximum number of diagnoses provided by an SC (Box 4).

The performance of individual SCs in terms of correct first diagnosis ranged between 12% and 61%, and for correct diagnosis in the top ten results between 30% and 81%. Some SCs were available from different sources (website, App Store, Google

Play), and the performance of the variants differed (Supporting Information, table 4).

Triage performance

The 19 triage SCs provided correct advice for 338 of 688 vignette tests (49%; 95% CI, 44–54%). Appropriate advice was more frequent for emergency care (63%; 95% CI, 52–71%) and urgent care vignette tests (56%; 95% CI, 52–75%) than for non-urgent care (30%; 95% CI, 11–39%) and self-care tests (40%; 95% CI, 26–49%) (Box 5; Supporting Information, table 5). The advice for 40%

3 Accuracy of diagnosis provided by 27 symptom checkers, by severity of patient condition in vignette and by frequency of diagnosis

	Total vignettes	Correct diagnosis					
		Listed first		Listed in top 3		Listed in top 10	
		Number*	Proportion (95% CI)	Number*	Proportion (95% CI)	Number*	Proportion (95% CI)
All vignettes	48	421/1170	36% (31–42%)	606/1170	52% (47–59%)	681/1170	58% (53–65%)
Patient condition							
Emergency	13 (27%)	92/343	27% (21–34%)	139/343	41% (33–50%)	161/343	47% (40–56%)
Urgent	14 (29%)	163/365	45% (39–51%)	231/365	63% (58–70%)	245/365	67% (62–73%)
Non-urgent	11 (23%)	99/252	39% (33–47%)	145/252	58% (51–67%)	168/252	67% (61–75%)
Self-care	10 (21%)	67/210	32% (25–42%)	91/210	43% (36–53%)	107/210	51% (42–60%)
Diagnosis frequency							
Common	41 (85%)	414/986	42% (37–50%)	593/986	60% (54–69%)	666/986	68% (61–76%)
Uncommon	7 (15%)	7/184	4% (1–7%)	13/184	7% (3–12%)	15/184	8% (4–13%)

CI = confidence interval. * Numbers of correct vignette evaluations and applicable evaluations. Some vignettes could not be applied to a given symptom checker (see Methods). ♦

4 Accuracy of diagnosis provided by 27 symptom checkers, by tool characteristics

Characteristics	Number	Correct diagnosis					
		Listed first		Listed in top 3		Listed in top 10	
		Number*	Proportion (95% CI)	Number*	Proportion (95% CI)	Number*	Proportion (95% CI)
All symptom checkers	27	421/1170	36% (31–42%)	606/1170	52% (47–59%)	681/1170	58% (53–65%)
Artificial intelligence algorithms							
Yes	8 (30%)	144/312	46% (40–57%)	201/312	64% (59–74%)	220/312	71% (65–80%)
No	19 (70%)	277/858	32% (26–38%)	405/858	47% (40–54%)	461/858	54% (46–61%)
Demographic questions							
At least age and sex	18 (67%)	293/770	38% (32–47%)	421/770	55% (47–64%)	481/770	62% (55–72%)
No	9 (33%)	128/400	32% (26–38%)	185/400	46% (40–53%)	200/400	50% (44–57%)
Maximum diagnoses listed							
0							
1–5	5 (18%)	83/218	38% (32–43%)	111/218	51% (43–58%)	111/218	52% (44–58%)
6–10	8 (30%)	132/312	42% (32–58%)	178/312	57% (44–76%)	197/312	63% (49–82%)
11 or more	14 (52%)	206/640	32% (23–40%)	317/640	50% (39–59%)	373/640	58% (48–68%)

CI = confidence interval. * Numbers of correct vignette evaluations and applicable evaluations. Some vignettes could not be applied to a given symptom checker (see Methods). ◆

5 Accuracy of triage advice provided by 19 symptom checkers, by severity of patient condition in vignette and by frequency of diagnosis

Characteristics	Total vignettes	Correct triage advice	
		Number*	Proportion (95% CI)
All vignettes	48	338/688	49% (44–54%)
Patient condition			
Emergency	13 (27%)	121/191	63% (52–71%)
Urgent	14 (29%)	115/206	56% (52–75%)
Non-urgent	11 (23%)	46/151	30% (11–39%)
Self-care	10 (21%)	56/140	40% (26–49%)
Diagnosis frequency			
Common	41 (85%)	297/587	51% (46–55%)
Uncommon	7 (15%)	41/101	41% (22–45%)

CI = confidence interval. * Numbers of correct vignette evaluations and applicable evaluations. Some vignettes could not be applied to a given symptom checker (see Methods). ◆

of non-urgent and self-care category vignettes was to seek urgent or emergency care, while for 10% of emergency category vignettes, non-urgent care or self-care was recommended (data not shown). For example, Family Doctor recommended self-care for “acute liver failure” and non-urgent care for “stroke”.

Four SCs (Isabel Healthcare, Everyday Health, Symcat, Doctor Diagnose) did not suggest self-care for any vignette; excluding these tools from the analysis did not markedly increase the proportion of correct responses (275 of 528 tests, 52%; 95% CI, 48–55%). Correct triage advice was more frequent for vignettes describing common conditions (51%; 95% CI, 46–55%) than for those describing uncommon conditions (41%; 95% CI, 22–45%); the accuracy of the five SCs requiring demographic data (53%;

6 Accuracy of triage advice provided by 19 symptom checkers, by tool characteristics

Characteristics	Number	Correct triage advice	
		Number*	Proportion (95% CI)
All symptom checkers	19	338/688	49% (44–54%)
Artificial intelligence algorithms			
Yes	5 (26%)	88/172	52% (50–54%)
No	14 (74%)	250/516	48% (41–54%)
Demographic questions			
At least age and sex	11 (58%)	240/451	53% (49–57%)
No	8 (42%)	98/237	41% (34–51%)
Maximum diagnoses listed			
0	9 (47%)	167/301	55% (48–58%)
1–5	4 (21%)	66/169	39% (11–65%)
6–10	6 (32%)	105/218	49% (42–56%)
11 or more	0	—	—

CI = confidence interval. * Numbers of correct vignette evaluations and applicable evaluations. Some vignettes could not be applied to a given symptom checker (see Methods). ◆

95% CI, 49–57%) was greater than for the 14 that did not (41%; 95% CI, 34–51%) (Box 6). Triage performance varied by SC (Supporting Information, table 6).

Discussion

In our assessment of SCs, the correct diagnosis was listed first in 36% of tests, similar to the figure of 34% reported by Semigram and colleagues;⁸ the correct diagnosis was included in the top ten results in 58% of our cases. Triage advice was

appropriate in 49% of tests, compared with 57% reported by Semigran and colleagues.⁸ The proportion of correct diagnosis results was larger for programs that use AI algorithms and require demographic information. The questions and diagnosis results of SCs available on multiple platforms were not always consistent, possibly because of differences in AI algorithms or programming.

Triage advice, especially for less serious case vignettes, tended to be risk-averse, although there were also notable instances of the opposite; the former can place unnecessary burdens on health care systems, but the latter can be dangerous. Neither we nor Semigran and colleagues⁸ found that omitting SCs which never suggest self-care from analysis markedly increased the proportion of triage recommendations classified as appropriate.

The diagnostic accuracy of SCs is limited by their programming and how information is presented. For example, the core signs and symptoms in the “heart attack” vignette are chest pain, sweating, and breathlessness. The Health Tools (American Association of Retired Persons) SC found no “possible causes” for this symptom combination, but suggested 51 diagnoses for the lone symptom “chest pain”, with “heart attack” listed second. That is, diagnostic usefulness was reduced by increasing the number of symptoms entered, although the tool prompts users to enter multiple symptoms.

In a recent British National Health System study,²¹ 20% of SC users were directed to the emergency phone number or an emergency department, suggesting that people consult online tools about urgent and emergency conditions. Health and computer literacy could affect how people use SCs,⁴ and this can pose challenges for patient–physician relationships if the SC advice contradicts that of the physician.¹⁰

Some SCs provide a directory of symptoms from which users can select, influencing the diagnostic capability of the program. For example, the Mayo Clinic SC includes 28 chief complaints for adult patients, but the list does not include “fever”. Overseas SCs do not include conditions specific to Australia, such as Ross River virus infection, reducing their local usefulness. The Healthdirect SC, the only Australian SC we assessed, provides triage but not diagnosis advice.

These limitations suggest that triage advice may be the more important function of online SC tools. Each included SC warned that their service was not a substitute for consulting a physician. Diagnosis is not a single assessment, but rather a process requiring knowledge, experience, clinical examination and testing, and the passage of time,¹⁶ impossible to replicate in a single online interaction. Overseas SCs also suggest health care services inappropriate for users in Australia, including self-referral to

specialist care. As reported by other investigators,¹⁰ some SCs did not report specific information about their authors or editors, or about the validity of their advice.

Despite their limitations and the lack of regulation of online health programs, SCs can be useful. Epidemiological data can be tracked on sites such as WebMD, providing timely data to health professionals.²² Users can educate themselves about their own health, potentially improving patient–physician relationships.^{3,5} SCs can also direct people to appropriate care, and some tools are even directly linked with health care services. As the public becomes more educated, active, and informed about health, online tools will be an increasingly popular resource.

Limitations and strengths

Our systematic approach to identifying SCs will have captured the most popular tools, but we may have missed some used by Australians. The vignettes were primarily simple scenarios describing patients without comorbid conditions, and did not reflect the complexity of real patients. The investigator who assessed all SCs was familiar with the clinical vignettes, and user bias was possible. Information about whether programs employed AI algorithms was drawn solely from that provided in the SC. Strengths of our investigation included our use of previously described vignettes and of lay language to test the programs, as well as our incorporating conditions specific to Australia.

Conclusion

The SCs we evaluated provided correct first diagnoses in 36% of cases and appropriate triage advice in 49%. Overseas SCs may not provide advice relevant to users in Australia. The medical terminology used in SCs and its effect on comprehensibility for users and their adherence to the advice provided should be investigated, as should the real life performance of SCs and their impact on health outcomes. To diminish the burden on health care systems, particularly emergency care, it is vital that online tools promptly direct people to appropriate care. Accordingly, online programs should be backed by quality sources, and provide diagnostic or triage advice that is as accurate as possible.

Acknowledgements: Our investigation was supported by an Australian Government Research Training Program Scholarship to Michella Hill. We acknowledge E. Khong (School of Medical and Health Sciences, Edith Cowan University) and G. Arendts (University of Western Australia Medical School) for their assistance with defining “correct” responses to the vignettes.

Competing interests: No relevant disclosures. ■

Received 29 August 2019, accepted 10 December 2019

© 2020 AMPCo Pty Ltd

- 1 Australian Bureau of Statistics. 8146.0. Household use of information technology, Australia, 2016–17. Mar 2018. <http://www.abs.gov.au/ausstats/abs@.nsf/mf/8146.0> (viewed Apr 2018).
- 2 Deloitte Australia. Mobile consumer survey 2018: the Australian cut. http://images.content.deloitte.com.au/Web/DELOITTEAUSTRALIA/%7B51598b8c-1c72-4ceb-8272-3a052f322992%7D_20181103-tel-inbound-mobile-consumer-survey-2018-report-2.pdf (viewed Apr 2019).
- 3 Lupton D, Jutel A. “It’s like having a physician in your pocket!” A critical analysis of self-diagnosis

- smartphone apps. *Soc Sci Med* 2015; 133: 128–135.
- 4 Cheng C, Dunn M. Health literacy and the Internet: a study on the readability of Australian online health information. *Aust N Z J Public Health* 2015; 39: 309–314.
- 5 Healthdirect Australia. About healthdirect symptom checker. Updated Apr 2016. <https://www.healthdirect.gov.au/about-healthdirect-symptom-checker> (viewed Apr 2018).
- 6 Cocco AM, Zordan R, Taylor DM, et al. Dr Google in the ED: searching for online health information by adult emergency department patients. *Med J Aust*. 2018; 209: 342–347.

<https://www.mja.com.au/journal/2018/209/8/dr-google-ed-searching-online-health-information-adult-emergency-department>.

- 7 Mueller J, Jay C, Harper S, et al. Web use for symptom appraisal of physical health conditions: a systematic review. *J Med Internet Res* 2017; 19: e202.
- 8 Semigran HL, Linder JA, Gidengil C, et al. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ* 2015; 351: h3480.
- 9 Boudreaux ED, Waring ME, Hayes RB, et al. Evaluating and selecting mobile health apps: strategies for healthcare providers and

- healthcare organizations. *Transl Behav Med* 2014; 4: 363–371.
- 10 Jutel A, Lupton D. Digitizing diagnosis: a review of mobile applications in the diagnostic process. *Diagnosis* 2015; 2: 89–96.
 - 11 United States Food and Drug Administration. Developing a software precertification program: a working model. Jan 2019. <https://www.fda.gov/media/119722/download> (viewed May 2019).
 - 12 Australian Department of Health, Therapeutic Goods Administration. Regulation of software as a medical device. Updated Mar 2020. <https://www.tga.gov.au/regulation-medical-software-and-mobile-medical-apps> (viewed Mar 2020).
 - 13 Australian Department of Health, Therapeutic Goods Administration. Australian regulatory guidelines for medical devices (ARGMD). Updated Oct 2019. <https://www.tga.gov.au/publication/australian-regulatory-guidelines-medical-devices-argmd> (viewed Mar 2020).
 - 14 Poote A, French D, Dale J, et al. A study of automated self-assessment in a primary care student health centre setting. *J Telemed Telecare* 2014; 20: 123–127.
 - 15 Murtagh J, Rosenblatt J. Murtagh's general practice. Sixth edition. Sydney: McGraw-Hill Education, 2015. <https://murtagh.mhmedical.com/book.aspx?bookid=1522> (viewed Mar 2020).
 - 16 Department of Health and Human Services (Victoria). Better Health Channel: conditions and treatments [website]. <https://www.betterhealth.vic.gov.au/health/ConditionsAndTreatments> (viewed May 2018).
 - 17 Australian Institute of Health and Welfare. Emergency department care 2016–17: Australian hospital statistics (Cat. no. HSE 194; Health services series no. 80). Canberra: AIHW, 2017.
 - 18 Cooke G, Valenti L, Glasziou P, Britt H. Common general practice presentations and publication frequency. *Aust Fam Physician* 2013; 42: 65–68.
 - 19 Britt H, Miller GC, Bayram C, et al. A decade of Australian general practice activity 2006–07 to 2015–16 (General Practice Series, no. 41). Sydney: Sydney University Press, 2016. https://ses.libra.ry.usyd.edu.au/bitstream/2123/15482/5/9781743325162_ONLINE.pdf (viewed June 2018).
 - 20 Middleton K, Butt M, Hammerla N, et al. Sorting out symptoms: design and evaluation of the “babylon check” automated triage system [preprint]. June 2016. <http://arxiv.org/abs/1606.02041> (viewed Apr 2018).
 - 21 NHS England. NHS111 online evaluation (draft). Dec 2017. https://askmygp.uk/wp-content/uploads/111-Online-Evaluation-DRAFT_.pdf (viewed Oct 2019).
 - 22 WebMD Health Corp. Searches for cold and flu symptoms up sharply on WebMD's Symptom Checker™ as consumers look for guidance; most recent data suggest flu rates are on the decline. *PR Newswire* [online], 28 Feb 2018. <https://www.prnewswire.com/news-releases/searches-for-cold-and-flu-symptoms-up-sharply-on-webmds-symptom-checker-as-consumers-look-for-guidance-most-recent-data-suggest-flu-rates-are-on-the-decline-300605273.html> (viewed Mar 2019). ■

Supporting Information

Additional Supporting Information is included with the online version of this article.