

# Assessing the Safety of a New Clinical Decision Support System for a National Helpline

Nirvana LUCKRAJ<sup>a,1</sup>, Renee STRAZZARI<sup>a</sup>, Enrico COIERA<sup>b</sup> and Farah MAGRABI<sup>b</sup>

<sup>a</sup>*Healthdirect Australia, Australia*

<sup>b</sup>*Australian Institute of Health Innovation, Macquarie University, Australia*

ORCID ID: Nirvana Luckraj, Renee Strazzari, Enrico Coiera: <https://orcid.org/0000-0002-6444-6584>, Farah Magrabi <https://orcid.org/0000-0002-8426-5588>

**Abstract.** We assessed the safety of a new clinical decision support system (CDSS) for nurses on Australia's national consumer helpline. Accuracy and safety of triage advice was assessed by testing the CDSS using 78 standardised patient vignettes (48 published and 30 proprietary). Testing was undertaken in two cycles using the CDSS vendor's online evaluation tool (Cycle 1: 47 vignettes; Cycle 2: 41 vignettes). Safety equivalence was examined by testing the existing CDSS with the 47 vignettes from Cycle 1. The new CDSS triaged 66% of vignettes correctly compared to 57% by the existing CDSS. 15% of vignettes were overtriaged by the new CDSS compared to 28% by the existing CDSS. 19% of vignettes were undertriaged by the new CDSS compared to 15% by the existing CDSS. Overall performance of the new CDSS appears consistent and comparable with current studies. The new CDSS is at least as safe as the old CDSS.

**Keywords.** Clinical decision support, telephone triage, consumer, evaluation, symptom checker, patient safety, health self-management

## 1. Introduction

Despite the widespread availability and use of online health information, there is little evaluation of the quality and safety of triage advice provided to consumers. As Australia's national public health information service, *Healthdirect Australia* provides consumers with easy access to trusted, quality health information and advice online and via a telephone helpline [1]. The Healthdirect helpline is available to all Australians 24 hours a day in all states and territories except Queensland. In 2022 the helpline received over 1.4 million calls. Callers to the helpline are triaged by registered nurses who, following exclusion of a life-threatening emergency, are assisted by a clinical decision support system (CDSS) to reach an initial disposition. The final triage outcome is determined by application of the nurse's clinical judgment on the initial disposition from the CDSS and the caller's individual circumstances.

While the use of CDSS improves decision-making, poorly designed, implemented or used systems can pose risks to patient safety [2]. Studies have shown that safety

---

<sup>1</sup> Corresponding Author: Nirvana Luckraj, email: [Nirvana.Luckraj@healthdirect.org.au](mailto:Nirvana.Luckraj@healthdirect.org.au).

risks increase during IT system transitions including migration to new systems and routine software updates [3]. Healthdirect are transitioning to a new software platform and having procured a new CDSS, needed to confirm its accuracy, safety and usability. Accordingly, a 3-part evaluation is being undertaken to assess the CDSS [4]. Parts 1 and 2 test the CDSS in a laboratory setting to examine algorithm performance and user interaction. Part 3 will monitor routine use of the CDSS post-deployment. Here we report the results of Part 1 which examined the performance of the CDSS algorithm in providing safe and accurate advice using standardised patient cases or vignettes – a widely accepted method for assessing the performance of CDSS for triage advice (*triage performance*) [5]. A safety equivalence study was also conducted to assess if the new CDSS is at least as safe as the existing one [6].

## 2. Methods

### 2.1. Patient Vignettes

Two cycles of testing were undertaken to assess the triage performance of the CDSS. In Cycle 1, the CDSS performance was benchmarked against 47 vignettes from the study by Hill et al. (2020) [7]. Hill et al. (2020) adapted 30 patient vignettes from the study by Semigran et al. (2015) [8] and supplemented them with 18 new symptom-based vignettes reflecting Australia-specific illnesses. One vignette involving a blue bottle jellyfish sting was excluded as the CDSS had not yet been updated for this condition at the time of testing. The safety equivalence study was conducted by running the same 47 vignettes (i.e. from Cycle 1) through the old CDSS.

In Cycle 2, a sample of 41 vignettes were used, including 11 vignettes from Cycle 1 (10 re-tested + blue bottle jelly fish sting) and 30 vignettes selected by Healthdirect's medical team to cover the following criteria: 1/most common clinical reasons for calling the helpline; 2/ rare conditions where a delay in diagnosis and triage could lead to harm; 3/ International diseases not found in the Australian context e.g. Malaria; 4/ New risk factors e.g. Indigenous population; 5/ Range of age groups e.g. Paediatric and geriatric; and 6/ Diseases localized for Australia.

**Box 1:** Example vignettes used to assess CDSSs for the national helpline (after [7, 8]).

*Myocardial infarction (heart attack):* 64-year-old male. 1 day of chest pain (8/10 pain). Pain does not move elsewhere. Sweating. Breathless. Feels tightness in mid chest. Chief complaint = chest pain  
Triage: emergency

*Queensland tick typhus:* 25-year-old female. Been unwell for 9 days. Fever (38.5°C), headache, dry cough, widespread rash. Muscle weakness. Painful upper left abdomen, with nausea and vomiting. Painful joints. Has been camping and gone on bush walks. Chief complaint = fever  
Triage: urgent

*Blue bottle jellyfish sting:* 16-year-old female. Swimming at beach. Stung by a jellyfish on arm. Immediate pain, slowly fading. Has some redness and swelling on arm, itchy. No trouble breathing, no abdominal pain. No nausea or vomiting. If asked- jellyfish was blue. Chief complaint = jellyfish sting  
Triage: selfcare

### 2.2. Laboratory Testing Procedure

The CDSS was tested via an online evaluation tool that was provided by the vendor. The correct diagnosis and triage advice were kept hidden. In each cycle, the vignettes

were randomly divided between two testers; two clinicians in Cycle 1 and two members of the Healthdirect medical team in Cycle 2. For the safety equivalence study, the old CDSS was tested by a registered nurse who was familiar with the tool. Testers entered each patient vignette into the CDSS and recorded the *triage advice* and *task time* (i.e. time duration from when a user begins an assessment to being provided with triage advice) provided by the evaluation tool. The order in which vignettes were entered was randomised to minimise any learning effects. To ensure consistency, practice vignettes and a standardised protocol was developed. If required, vignettes were re-entered. Disagreements between the testers and the CDSS prompted a discussion with a member of the medical team for resolution.

### 2.3. Outcome Measures and Analysis

Triage accuracy and safety was examined over all vignettes (Cycle 1: n=47; Cycle 2: n=41) and by the four triage categories [7]: i/ *emergency*: requiring immediate medical care; ii/ *urgent*: requiring medical attention within 24 hours; iii/ *non-urgent*: see a healthcare provider in the near future; and iv/ *self-care*: assistance people provide themselves. Triage performance was compared with other CDSSs reported in the literature [5], as well as the old CDSS. A two-tailed t-test (assuming unequal variances) was used to compare the task times recorded by testers.

## 3. Results

### 3.1. Triage Accuracy

We found that the overall triage accuracy of the CDSS in both testing cycles (Table 1) was comparable with the performance of most symptom checkers reported in a recent systematic review (avg. triage accuracy 49-68% [5]) and the Hill et al. (2020) study (average 49%, range 17-61% [7]). Comparison of task times between testers showed there was no difference in the time (min:sec) taken to work through the vignettes in each cycle (Cycle 1: 3:50 (3:19 – 4:20),  $p=0.341$ ; Cycle 2: mean (95% CI) = 2:58 (2:42 – 3:19),  $p=0.811$ ).

**Table 1.** Triage accuracy of CDSS compared with old CDSS.

	New CDSS Cycle 1 (N=47)		Safety equivalence Old CDSS (N=47)		New CDSS Cycle 2 (N=41)	
	n	%	n	%	n	%
<b>Correctly triaged</b>	31	66	27	57	25	61
<b>Overtriage</b>	7	15	13	28	9	22
<b>Undertriage</b>	9	19	7	15	7	17

### 3.2. Triage Safety

The safety equivalence study showed that the new CDSS's triage accuracy was 9% higher but 4% more vignettes were undertriaged (Table 1). Accuracy by the four triage

categories is shown in Figures 1 and 2. At the conclusion of Cycle 1, a medical review of undertriaged and overtriaged vignettes was conducted by Healthdirect’s medical team. Of 14 undertriaged vignettes it was found that five flows from those vignettes had triaged appropriately based on user input. Similarly, seven of the nine overtriaged vignettes had triaged appropriately based on user input. All 11 undertriaged and overtriaged flows were referred to the CDSS vendor for further investigation (Table 2).

Expected triage	Actual Triage				Total
	Selfcare	Non-urgent	Urgent	Emergency	
Selfcare	5	3	1		9
Non-urgent	2	8	1		11
Urgent		4	8	2	14
Emergency			3	10	13
Total	7	15	13	12	47

Figure 1. Cycle 1 triage safety showing degree of undertriage and overtriage of the CDSS.

Expected triage	Actual Triage				Total
	Selfcare	Non-urgent	Urgent	Emergency	
Selfcare	4	4	0	1	9
Non-urgent	1	1	3	0	5
Urgent		5	8	1	14
Emergency			1	12	13
Total	5	10	12	14	41

Figure 2. Cycle 2 triage safety showing degree of undertriage and overtriage of the CDSS.

Table 2. Medical team review and follow-up of undertriaged and undertriaged vignettes.

Diagnosis	Expected triage outcome	Cycle 2 triage outcome	Comments
<i>Undertriage</i>			
Acute exacerbation	COPD emergency	emergency	
Kidney stone	emergency	urgent	two levels undertriage; no change
Malaria	emergency	emergency	
Migraine	urgent	urgent	
Queensland tick typhus	urgent	urgent	
Shingles	urgent	urgent	
Tonsillitis	urgent	urgent	
Bee sting	urgent	consultation	undertriage; no change
Trochanteric bursitis	consultation	self-care	undertriage; no change
<i>Overtriage</i>			
Cellulitis	urgent	urgent	
Threadworm	self-care	consultation	overtriage; no change

As with Cycle 1, undertriaged and overtriaged vignettes from Cycle 2 were reviewed by Healthdirect’s medical team. Of the eight undertriaged vignettes, one flow had triaged appropriately based on user input. Similarly, of 12 over-triaged vignettes it was determined that eight had triaged appropriately based on user input. All undertriaged and overtriaged flows were again referred to the vendor for further investigation.

#### 4. Discussion

Safety assessment of the new CDSS showed that its triage performance is comparable with current CDSSs for symptom assessment [5]. The safety equivalence study demonstrated that the new CDSS is at least as safe as the old CDSS. Nurses use their clinical judgment when assessing the initial disposition advice from the CDSS. A 2021 audit of the Healthdirect helpline found that, “under-triaging was rectified by nurses using their clinical judgment at the point of final disposition...indicating that nurse clinical judgment ensured a safer outcome for a significant proportion of callers”. This study has several limitations. Vignettes are not representative of real-world calls received on the helpline and do not contain all the details nurses would obtain in a discussion on a call. Each vignette was tested by one tester, inter-rater reliability between testers was not formally assessed and variability of tester input was observed. However, there was no difference in time between testers. While vignettes are useful for benchmarking user experience and performance of symptom checkers they should be regarded as a first step because they are not representative of real-world use. This study is being followed up by further evaluation with community users experiencing symptoms to examine nurses’ experience of using the CDSS (Part 2). Nurses receiving calls on the helpline will be interviewed to assess perceptions about the usability of the new CDSS, and its performance will be monitored in routine use (Part 3).

#### 5. Conclusions

Overall performance of the new CDSS appears consistent and comparable with current studies. The safety equivalence study demonstrates that the new CDSS is at least as safe as the old CDSS. The new CDSS has been determined to be a safe and appropriate triage tool to support nurse call agents on the helpline.

#### References

- [1] Healthdirect Australia helpline from: <https://www.healthdirect.gov.au/> (last accessed 27 November 2022).
- [2] Kim MO, Coiera E, Magrabi F. Problems with health information technology and their effects on care delivery and patient outcomes: a systematic review. *J Am Med Inform Assoc.* 2017;24(2):246-50, doi: 10.1093/jamia/ocw154.
- [3] Magrabi F, Liaw ST, Arachi D, Runciman W, Coiera E, Kidd MR. Identifying patient safety problems associated with information technology in general practice: an analysis of incident reports. *BMJ Qual Saf.* 2016;25(11):870-80, doi: 10.1136/bmjqs-2015-004323.
- [4] Fraser H, Coiera E, Wong D. Safety of patient-facing digital symptom checkers. *Lancet.* 2018;392(10161):2263-4, doi: 10.1016/S0140-6736(18)32819-8.
- [5] Wallace W, Chan C, Chidambaram S, Hanna L, Iqbal FM, Acharya A, Normahani P, Ashrafian H, Markar SR, Sounderajah V, Darzi A. The diagnostic and triage accuracy of digital and online symptom checker tools: a systematic review. *NPJ Digit Med.* 2022;5(1):118, doi: 10.1038/s41746-022-00667-w.
- [6] Leveson NG. *Engineering a safer world: systems thinking applied to safety.* USA: The Massachusetts Institute of Technology (MIT) Press; 2011.
- [7] Hill MG, Sim M, Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. *Med J Aust.* 2020;212(11):514-9, doi: 10.5694/mja2.50600.
- [8] Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ.* 2015;351:h3480, doi: 10.1136/bmj.h3480.