



## Healthline: do primary care doctors agree with the advice?

Ian St George, Matthew Cullen, Michelle Branney

### Abstract

**Aims** To assess agreement between the advice to symptomatic callers to Healthline, and that advised by primary care doctors given the same clinical information, and thus to assess the safety of Healthline advice.

**Methods** Ninety records of symptomatic calls to Healthline were examined by three primary care specialists, blinded to the actual advice given. They independently recorded what they would have advised, and their advice was compared with that actually given by the Healthline nurse guided by Care Enhance Call Centre software.

**Results** Variation among the three doctors was greater than that between the median doctor and Healthline. In 82% of cases, the median doctor triaged to an endpoint close to (or lower than) Healthline. In all but one of the remainder, at least one doctor thought there was no risk to the patient (i.e. in 99% of cases). Review of that case indicated nurse error and the guideline itself was judged to be safe.

**Conclusion** New Zealand primary care specialists regarded the Care Enhance Call Centre decision support software used by Healthline as clinically safe.

In the 1990s, although telephone advice formed a significant part of the workload of US and UK emergency departments, it was often inaccurate or inadequate.<sup>1,2</sup> In 1995, Aitken and coworkers reported phoning 30 New Zealand public hospital emergency departments and 20 private accident and medical centres, role-playing the parent of a feverish infant. They received a wide range of responses. In 14 centres, the doctor on call was contacted at once; in 5 centres, the doctor gave advice; in 26 centres, the nurse gave advice; and in 5 centres, the advisers did not say who they were.

The authors judged the advice from 16 of the 36 centres to be inadequate.<sup>3</sup> Thus such observations led to a desire to standardise telephone advice for symptomatic callers by using electronic information systems.

In the United States during the 1990s, private insurance schemes began offering nurse-led, software-supported telephone triage and advice; and NHS Direct (a free, 24-hour advice and triage line) began in England.

Meanwhile, in New Zealand in 1994, Tisdale reported a pilot of a nurse-run telephone advice line,<sup>4</sup> and in 1998 Cameron and others reported high levels of satisfaction by those phoning the National Poisons Centre for advice, as well as considerable saving of public money.<sup>5</sup>

Furthermore, the repeated observation that resource-intensive primary medical services were being inappropriately used by those whose symptoms could have been managed with lower levels of care,<sup>6</sup> led to a desire for primary care demand management, and again, telephone triage seemed to supply an answer.

The two drivers for freely available telephone triage were thus standardisation of advice, and improved resource use by directing callers to lower levels of care.

The New Zealand Health Ministry funded Healthline pilot began in four regions in 2000. In 2004, it incorporated Plunket Line and in May 2005 it became a national, state-funded, 24x7, primary health service offering health information, well child, and parenting assistance, and symptom triage.

About 70% of Healthline callers seek advice on symptoms, and 70% of the calls are outside normal working hours. Many are triaged to lower levels of care than they had intended before calling Healthline.<sup>7</sup>

The question arises, then: is it safe? An independent study of the Healthline pilot, commissioned by the Ministry, exhaustively traversed the clinical quality activities of the service, and concluded, 'The Healthline service has operated at least as safely to date as similar overseas telephone services,'<sup>8</sup> but Moriarty and others expressed concern after a study of simulated callers.<sup>9</sup>

Since then, Healthline has changed the software from 'Personal Health Advisor™' to a new decision support package called 'Care Enhance Call Centre™.'

We decided then to examine the degree of concordance between primary medical care specialists with Healthline advice using the Care Enhance Call Centre™ software package.

## Method

Healthline nurses triage callers seeking advice on current symptoms by using a symptom-specific guideline to one of nine dispositions ('endpoints': Table 1). The guideline is computer-based, prompting the nurse to ask a series of questions designed to exclude the most serious potential causes of the symptom. The endpoint is reached when a specific cause cannot be excluded. The nurse makes a full electronic record including demographic data, free text on the presenting symptom, the responses to the guideline questions, the endpoint reached, and the advice and assistance given.

**Table 1. Triage endpoints for symptomatic callers**

- |   |
|---|
| <ol style="list-style-type: none"><li>1. Activate 111</li><li>2. Go to general practitioner or emergency department immediately</li><li>3. Call or go to other healthcare provider immediately (e.g. mental health team, poisons centre, dentist, lead maternity carer)</li><li>4. See doctor or other healthcare provider within 4 hours</li><li>5. See doctor or other healthcare provider within 8 hours</li><li>6. See doctor or other healthcare provider within 24 hours</li><li>7. See doctor or other healthcare provider within 72 hours</li><li>8. See doctor or other healthcare provider within 2 weeks</li><li>9. Home self-care</li></ol> |
|---|

Beginning from midnight 31 January 2005, we selected the first 10 cases triaged to each endpoint—and printed the call record, obliterating the actual endpoint reached, and the advice given. Three primary care specialists then independently examined the

90 clinical records thus generated and decided what endpoint they would have recommended, given the same information. There were two general practitioners (both women, one urban and one rural) and one accident and medical doctor with a hospital emergency department background: two were selected on the basis of personal acquaintance with the service, and one represented the Wellington after hours medical service.

Thus the 'gold standard' with which we compared the nurse decisions, was that of experienced primary care specialists, given the same clinical information.

We defined agreement as an endpoint at or immediately adjacent to that actually recommended by the nurse. We calculated concordance by crude percentage agreement, and by Cohen's Kappa (K) which provides a number between 0 and 1: a K of 0.7 or more is regarded as showing satisfactory inter-rater reliability.

## Results

The endpoints advised by the three doctors varied: crude percentage agreement between the highest and lowest was only 51% and Cohen's K = 0.43.

For that reason we compared the *median* doctor-advised endpoint for each case with the endpoints reached by the Healthline nurses. For this comparison, crude percentage agreement was 70%, and Cohen's K = 0.78.

In 10 cases (11.1%), the median doctor triaged to two endpoints or more lower (i.e. to a lower level of care) than the Healthline nurse. In 64 cases (71.1%), the median doctor triaged to within one endpoint of that reached by the Healthline nurse. Thus, for 82.2% of cases, there were no safety concerns.

In the other 16 cases (17.8%), the median doctor triaged to two or more levels of care higher, but the three doctors were unanimous in only seven of these (7.8%). They were asked to review these seven cases, considering specifically whether the Healthline nurse endpoint was unsafe. In only one case (1.1% overall) did all three consider that the lower endpoint posed some risk to the patient. On detailed review of this case, the lower endpoint reached by the Healthline nurse was not a function of the software, but an error by the operator.

In only four cases, the doctors would have triaged to 111 when Healthline did not, and in all of these the Healthline nurse had recommended immediate medical care. The doctors considered there was no risk to the patients in these cases.

## Discussion

A Kappa of 0.78 between the median doctor and Healthline is reassuring, as is the perception that Healthline endpoints were clinically safe in 99% of cases, and the software clinically safe in all cases. The single case where disagreement was unanimous among the doctors was not a fault of the software, but an error by the nurse.

One doctor spontaneously commented:

The exhaustiveness of the questioning that the nurses do before reaching a conclusion is probably much more accurate in the end than I would do in the night for example when woken up. The computer prompts them to go into all possible urgent scenarios and thus I think it would be pretty unusual for them to miss anything. Having seen the system 'up close' it

appears excellent. The patients will appreciate receiving consistent advice rather than the enormous variety they receive from different individual doctors.

We observed wide variation among the three doctors (Kappa 0.43), wider than that between the median doctor and Healthline (Kappa 0.78). Indeed, one doctor wrote:

It could be that I am a little cautious or that I am used to dealing with patients who want immediate service/action.

Gribben similarly asked 12 doctors to assess whether they would have managed patients attending a city emergency department themselves, and found a Kappa of 0.34 among them; he remarked:

There was a surprisingly wide range of views on the proportion of cases that the GPs thought could be completely handled in primary care. We asked for personal views, and so naturally the skills and experience of individual GPs will have contributed to the range. These assessments did not appear to be related to age or gender and the scores of the A&M (accident and medical) doctors and academic GPs were distributed across the range<sup>11</sup>

The range of medical practice variation is wide, and with Marshall Marinker we offer no criticism of that: "Narrowing the range may give us the illusion of consensus, without telling us anything at all about whether the consensus is better than the diversity," he wrote.<sup>10</sup> Similarly, we make no judgement as to who is right: triage can be standardised, but without guidelines it is not a precise science.

**Note:** We acknowledge that three is a small group of doctors, and that a larger group may have provided more reliable data.

**Author information:** Ian St George, Medical Director, Healthline, Wellington; Matthew Cullen, Co-President, McKesson Asia-Pacific, Lane Cove, New South Wales, Australia; Michelle Branney, General Manager, McKesson NZ Ltd, Wellington

**Correspondence:** Ian St George, Healthline, PO Box 10643, Wellington. Fax: (04) 499 2357; email: [ian.stgeorge@healthline.co.nz](mailto:ian.stgeorge@healthline.co.nz)

## References:

1. Evans RJ, McCabe M, Allen H, et al. Telephone advice in the accident and emergency department: a survey of current practice. *Arch Emerg Med.* 1993;10:216–9.
2. Isaacman DJ, Verdile VP, Kohen FP, Verdile LA. Pediatric telephone advice in the emergency department: results of a mock scenario. *Pediatrics.* 1992;89:35–9.
3. Aitken ME, Carey MJ, Kool B. Telephone advice about an infant given by after-hours clinics and emergency departments. *N Z Med J.* 1995;108:315–7.
4. Tisdale G. The results of a telephone health advice and information service February-March 1993. *N Z Med J.* 1994;107:128–9.
5. Cameron S, Cheu B, Green J, et al. Client satisfaction with the New Zealand Poisons Information Service. *Vet Hum Toxicol.* 1998;40:163–5.
6. Hider P, Kirk R, Bidwell S, et al. Emergency Department Attendance. Christchurch: New Zealand Health Technology Assessment Clearing House; 1998.
7. St George IM, Cullen MJ. The Healthline pilot: call centre triage in New Zealand. *N Z Med J.* 2001;114:429–30.
8. Kalafatelis E, Fryer K, Harsant M, et al. The evaluation of the Healthline service: final evaluative report. Wellington: BRC Marketing and Social Research; 2002. Available online. URL: [http://www.brc.co.nz/papers/healthline\\_report.pdf](http://www.brc.co.nz/papers/healthline_report.pdf) Accessed October 2005.

9. Moriarty H, McLeod D, Dowell A. Mystery shopping in health service evaluation. *Br J Gen Pract.* 2003;53:942–6.
10. Marinker M. Don Alfonso's rehearsal; 1993. Available online. URL: [http://www.marinker.co.uk/don\\_alfonso.htm](http://www.marinker.co.uk/don_alfonso.htm) Accessed October 2005.
11. Gribben B. General practitioners' assessments of the primary care caseload in Middlemore Hospital Emergency Department. *N Z Med J.* 2003;21;116(1169). URL: <http://www.nzma.org.nz/journal/116-1169/329>