



Editor's choice  
Scan to access more  
free content

# Errors in after-hours phone consultations: a simulation study

Erel Joffe,<sup>1,2</sup> James P Turley,<sup>1</sup> Kevin O Hwang,<sup>3</sup> Todd R Johnson,<sup>1</sup> Craig W Johnson,<sup>1</sup> Elmer V Bernstam<sup>1,3</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjqs-2013-002243>).

<sup>1</sup>School of Biomedical Informatics, The University of Texas Health Science Center, Houston, Texas, USA

<sup>2</sup>Department of Hematology and Bone Marrow Transplantation, Tel Aviv Medical Center, Tel Aviv, Israel

<sup>3</sup>Department of Internal Medicine, Medical School, The University of Texas Health Science Center, Houston, Texas, USA

## Correspondence to

Dr Elmer Bernstam, School of Biomedical Informatics, The University of Texas Health Science Center, 7000 Fannin, Suite 600 Houston, TX 77030, USA; [Elmer.V.Bernstam@uth.tmc.edu](mailto:Elmer.V.Bernstam@uth.tmc.edu)

Received 19 June 2013

Revised 31 October 2013

Accepted 9 November 2013

Published Online First

13 December 2013

## ABSTRACT

**Background** After-hours out-of-hospital phone consultations require physicians to make decisions based on information provided by a nurse over the phone.

**Methods** We conducted a simulation study to evaluate physicians' actions following communication of key information. 22 nurses were asked to call physicians with six cases based on the six most common reasons for after-hours phone calls. We evaluated physicians' actions following the communication of key clinical information: A *situation cue* described a patient's problem (eg, confusion). A *background cue* described a specific clinical finding regarding the cause of the problem (eg, patient's sodium is low). For each cue we defined a list of indicators, based on the medical literature, to ascertain whether physicians acted upon the provided information (which was defined as addressing at least one of the indicators).

**Results** A total of 108 phone consultations (containing 88 situation and 93 background cues) were analysed. Situation cues were communicated in 90% (79/88) of the calls and background cues in 33% (31/93). Physician acted upon the provided information in 57% (45/79) and 48% (15/31) of the communicated situation and background cues, respectively. When the background cues were not communicated, physicians asked questions expected to elicit the cue in 12% of the cases. Responding to the situation cue was associated with longer conversations and active inquiry by the physician.

**Conclusions** After-hours phone calls are error prone. Both nurse communication and physician decision-making are problematic. Efforts to improve patient safety in this setting must address both communication and decision-making.

## INTRODUCTION

Discontinuities in care due to communication failures have been associated with preventable adverse events.<sup>1</sup> These failures may occur at any time, but the risk is

higher at night or during weekends (after-hours).<sup>2</sup> Inpatient after-hours phone communications are a clinical scenario in which a nurse consults a physician (who is on-call but may be outside of the hospital) regarding an acute patient problem using a telephone. These calls are usually limited to verbal communication, take place in a setting where scarce resources and fatigue are the rule, and are usually characterised by a paucity of information.<sup>2</sup> For example, often the consulted physician is not the primary physician responsible for the patient, and may have received only a very brief 'sign out' (eg, '40-year-old with pneumonia, doing well') or is entirely unfamiliar with the patient. Furthermore, nurse-physician communication is error prone due to different communication cultures specific to each profession (regardless of whether it takes place over the phone).<sup>3 4</sup> Considering the frequency of on-call phone communications, there is surprisingly little research regarding potential risks and adverse outcomes resulting from these calls.

In previous work, we found that a limited set of problems account for the majority of after-hours phone calls.<sup>2</sup> We attempted to improve nurse-physician communication by providing a problem-specific communication tool that lists the necessary data to be communicated under these common clinical scenarios.<sup>5</sup> We based the tool on the Situation-Background-Assessment-Recommendation (SBAR) framework, which is the most commonly implemented communication framework in healthcare.<sup>6</sup> Notably, previous evaluations of SBAR tools in healthcare have yielded mixed results.<sup>7</sup> Several studies proclaim great success in institution-wide implementations of the SBAR framework,<sup>8 9</sup> while other studies found no effect and even worse performance.<sup>10 11</sup> We found that a



To cite: Joffe E, Turley JP, Hwang KO, et al. *BMJ Qual Saf* 2014;**23**:398–405.

problem-specific SBAR tool did not improve communication of key information between nurses and physicians.<sup>5</sup> Specifically, in the majority of cases nurses stated the reason for the call (eg, patient is confused) but failed to communicate pertinent information about the cause (eg, patient's sodium is 119). This was observed regardless of whether they used the SBAR forms (that were designed to guide data extraction and communication). In this study, we addressed physician decision-making. Specifically, we evaluated physicians' actions following the communication of key information by nurses and whether physicians were able to elicit the information when not provided by the nurse.

## METHODS

A simulation study was conducted at the Texas Medical Center (Houston, Texas, USA) from May 2010 to May 2011. In all, 22 pairs, each consisting of a registered nurse and an internal medicine physician (attending or senior resident responsible for fielding after-hours calls from nurses on a regular basis), were enrolled in the study. Both nurses and physicians had to be practicing on general internal medicine wards at the time of the study.

We presented each nurse with six clinical scenarios (table 1). We based the scenarios on actual patient records from a local tertiary-care hospital. Cases were selected by two experts (internists) (EVB, KOH) for representing a moderate diagnostic challenge in the context of the six most common reasons for after-hours calls (fever, glucose management, behaviour problems, medication prescription, blood pressure and pain).<sup>2</sup> The rationale was that in order to evaluate whether physicians considered the communicated information in diagnosis and management of patients we must present some challenge. For example, a common reason for an after-hours call is blood glucose management. The reflexive response of a physician in such a case might be to ask about the glucose level and decide about insulin dosage. In our experiment, we presented a case where there was an order for insulin in the presence of normal blood glucose and a concomitant order for glucose (ie, a treatment for elevated blood potassium level). Consideration of the communicated information would direct the physician away from simple management of blood glucose to the evaluation and management of elevated blood potassium level.

Each case had at least one cue that was required in order to resolve the clinical scenario. There were two types of cues. A *situation cue* answered the question 'what is wrong with the patient that is prompting the call?' For example, the patient is disoriented and pulled out his intravenous line. These cues were used to evaluate the physician's understanding of the general situation and generate a differential diagnosis. A *background cue* was a specific clinical finding that

answered the question 'why does this particular patient suffer from this problem?' For example, very low sodium level in a patient with acute confusion. These were used to determine whether the physician understood the aetiology of the patient's problem and was able to act on this understanding. We asked nurses to review the patient records, extract information (with or without the aid of an SBAR form) and then call the physician requesting instructions on how to manage the clinical scenario. For a detailed description of the SBAR forms and methodology for evaluating their utility in communicating information see<sup>5</sup>.

For each call we noted whether the situation/background cues were communicated. Then, for cases where the information was communicated, we evaluated the appropriateness of physicians' responses. We considered physician responses to be 'appropriate' responses if there was any indication that the physician had considered and acted upon the provided information.

We defined the set of appropriate responses using a guide-book for the on-call physician and the UpToDate knowledge base (see online supplementary appendix A).<sup>12 13</sup> For the situation cues we defined the elements of an appropriate differential diagnosis or what further questions/diagnostic workup was indicated. For example, in a young patient with an acute confusion (situation cue) a physician is expected to consider infection, metabolic or electrolyte abnormalities, stroke or other organic brain damage and medication/drug/alcohol related confusion.<sup>14 15</sup> Indication that the physician has indeed considered the situation cue would therefore include questions regarding a possible infection, medications and substance abuse, laboratory tests and head imaging.<sup>14 15</sup> For the background cues we established what the appropriate orders should be. For example, given confusion which is associated with low sodium (background cue), a physician could either repeat blood tests to validate the diagnosis; give intravenous fluids with sodium; order a limit on free water intake; or admit to the intensive care unit.<sup>16</sup>

We defined physicians' responses as appropriate if, following a communicated situation cue, they had asked about *any* of the possible diagnoses on the differential diagnosis, and if following a communicated background cue, they had ordered *any* of the indicated orders (table 1, see online supplementary appendix A). In real clinical practice, a physician is expected to consider *all* major diagnoses on the differential, and address *all* active problems associated with these diagnoses. However, since there are often multiple reasonable strategies for a given clinical scenario, and since it is difficult to precisely define 'appropriate response' we used a much more lenient measure, focusing on whether physicians acted upon the communicated information rather than the adequacy of their medical decision. Figure 1 presents the study design.

**Table 1** Clinical scenarios, cues and measures of evaluating appropriateness of physicians' actions

| Presented scenario  | Situation cue  | Expected action*  | Background cue  | Expected action*   |
|---|--|---|---|--|
| <i>Fever case:</i> A 43 y/o man admitted several days ago with a UTI, has a 101.5° fever and 2–3 loose stools             | Persistent fever in a hospitalised patient/patient under antibiotic treatment    | Evaluation for healthcare associated or complicated infection. Asked about <i>any of</i> : recent hospitalisation, recent antibiotics use, other source of infection, presence of a complicated UTI or non-infectious cause of fever† | Back surgery 1 month ago or prior hospitalisation until 4 days ago  | <i>Any of</i> : Broadened antibiotic coverage, or ordered imaging for epidural abscess or asked about evidence for surgical site infection |
| <i>Glucose case:</i> A 48 y/o man has a standing order for insulin but blood glucose is 90 mg/dL                          | There is also an order for glucose 50% (a standard Tx for high potassium levels) | Evaluation for the cause of elevated potassium. Asked about <i>any of</i> : current medications, renal function or acid/base values   | Patient is treated with tacrolimus (may cause elevated potassium levels)  | <i>Any of</i> : Asked for tacrolimus level, reduced dose or requested a nephrological consultation   |
| <i>Behaviour (Confusion) case:</i> A 19 y/o man admitted for sickle cell crisis is disoriented and pulled out his IV line | The patient is confused  | Evaluation for an acute change in mental status. Asked about <i>any of</i> recent opiate therapy, substance abuse or ordered <i>any of</i> : discontinue opiates, blood glucose, electrolytes or head CT                              | Patient has a new low sodium level (119 mg/dL) and a persistently high WBC count (23 000)   | <i>Any of</i> : Treated low sodium, evaluated and treated for an infection   |
| <i>High Blood Pressure case:</i> An 85 y/o woman has a high blood pressure of 180/90 mmHg                                 | None†  | None†   | Home treatment with clonidine was discontinued (causes rebound high blood pressure); the patient has received large volumes of fluids | <i>Any of</i> : Reinstated clonidine Tx., stopped fluid Tx. or ordered diuretic Tx   |
| <i>Medication case:</i> A 31 y/o woman has difficulty sleeping and asks for a sleeping pill                               | The patient was admitted for acute liver injury                                  | Refrained from benzodiazepines or asked about comorbidities (eg, respiratory compromise) or conflicting drugs   | The patient is treated at home with a CPAP for obstructive sleep apnoea   | Refrained from any sleep medication until CPAP treatment was reinstated  |
| <i>Chest Pain case:</i> A 61 y/o woman presents with chest pain   | The patient has swelling of her leg  | Suspicious for pulmonary embolism. <i>Any of</i> : CTA, D-dimer, V/Q scan, lower extremity venous US or dedicated examination   | None‡   | None‡  |

\*A complete list of the evaluation criteria can be found in online supplementary appendix A.

†There were no relevant 'appropriate responses' for the situation cue of the High Blood Pressure case (the blood pressure in itself was not high enough to warrant a comprehensive evaluation).

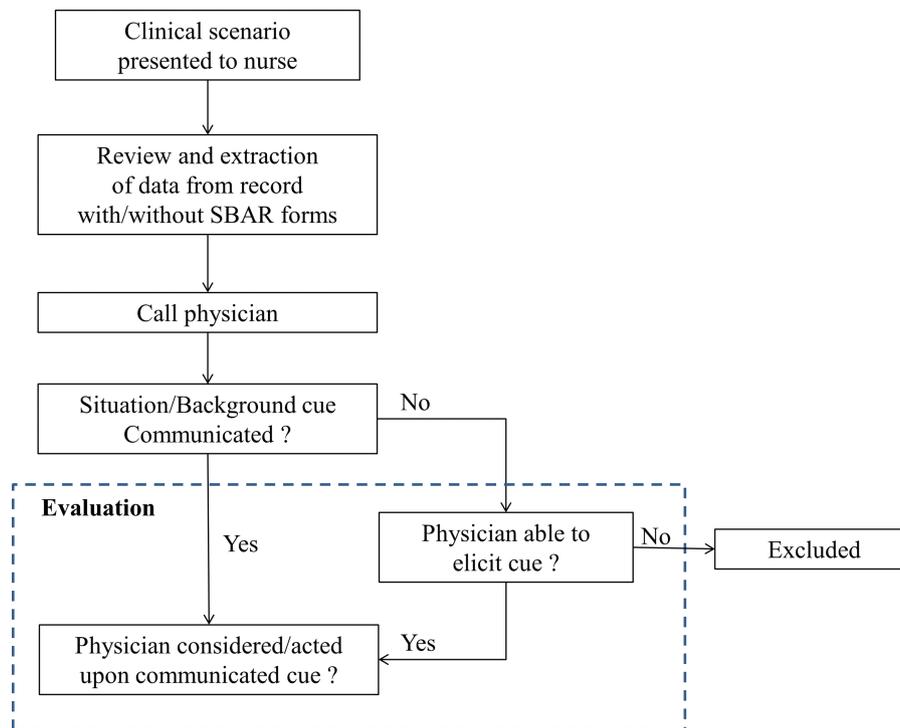
‡There were no relevant 'appropriate responses' for the background cue of the Chest Pain case (the patient had multiple comorbidities that could have been responsible for chest pain. Thus, we made the decision to present the background cue of leg-swelling to the nurse subjects and excluded it from the analysis).

CPAP, continuous positive airway pressure; CTA, CT angiogram; IV, intravenous; Tx, treatment; US, ultrasound; UTI, urinary tract infection; V/Q, ventilation/perfusion scan; WBC, white blood cell; y/o, year old.

As we were using cases based on actual patient records, there were no relevant 'appropriate responses' for the situation cue of the High Blood Pressure case (the blood pressure in itself was not high enough to warrant a comprehensive evaluation) and for the background cue of the Chest Pain case (the patient had multiple comorbidities that could have been responsible for chest pain. Thus, we made the decision to present the background cue of leg-swelling to the nurse subjects and excluded it from the analysis).

The Fever case included a misleading detail, a description of 2–3 loose stools suggesting the possibility of a *Clostridium difficile* infection. However, there were no other data to support the diagnosis (ie, overt diarrhoea, leucocytosis, etc). In this case, physicians' actions were deemed adequate only if they had entertained an alternative diagnosis to that suggested by the misleading detail.

Nurses were asked to come to the laboratory (ie, simulated internal medicine ward), while physicians were contacted by phone (ie, simulated 'out of hospital' call). Physicians did not receive any sign-out information about the patients prior to the experiment. Subjects were told we are conducting a study to evaluate the communication between nurses and physicians without any additional details. Hence, both nurses and physicians were blinded to the objectives of the study and to the evaluation measures. Experiments were scheduled when it was convenient for the subjects. We provided nurses with a medical record (including an admission note, progress notes, medical orders, medication, nursing notes, laboratory and imaging results) and a bedside (nursing) chart (with vital signs, intake/output, etc). We used actual hospital forms to ensure that nurses worked with records that were as close as possible to records used in routine clinical practice. The expert panel made sure that the records covered all the



**Figure 1** Study design.

pertinent data for the evaluation of each clinical scenario. Records were minimally modified to fit the clinical scenario, and to comply with the requirements for de-identification. The nurse subject could ask the nurse conducting the experiment (JPT) about physical exam findings for which there were scripted answers. If the question fell outside of the scripted answers, no information was given. The cumulative time to review the six records was limited to 2 h. Calls were not time-limited and were recorded by video and audio (MP3) and analysed by a single non-blinded reviewer (EJ). To avoid possible biases we refrained from using subjective measures in the evaluation (table 1).

We noted what information was communicated by the nurse, including erroneous information, and what additional information the physician requested. Evaluation was based on the data elements listed on the SBAR forms (see online supplementary appendix B). The rationale was that the SBAR forms listed a set of data elements that an expert panel deemed to be required for an appropriate evaluation of the type of case.<sup>5</sup> We recorded the time that elapsed between the beginning of the phone conversation and when the reason for the call was communicated; the total length of the call; number and types of data items communicated regarding the patient's situation other than the situation cue (eg, patient identification and location, whether the problem was urgent); and the number and type of data items communicated regarding the patient's background other than the cue (eg, reason for hospitalisation, prior medical history, vital signs, medication). Physicians were also evaluated for their ability to elicit required information regarding

the situation and background cues when these were not provided by the nurses (eg, asking for the most recent laboratory results in the confused patient). Online supplementary appendix C presents a transcription of one session including comments to indicate the steps of the study, the positions where the situation and background cues were provided as well as communication of situation data elements, background data elements, assessment and plan.

#### Statistical analysis

Randomisation and allocation of cases to subjects were based on a Latin Square randomisation table.<sup>17</sup> Statistical analyses were performed using SPSS (V.20, IBM Inc., Chicago, Illinois, USA). We used generalised estimating equations (GEEs) to evaluate the association between adequacy of physicians' actions and properties of the communication. We chose GEEs due to repeated measures within subjects and cases, missing observations and non-normal distribution of our data.<sup>18</sup> We conducted the analysis sequentially. First, we assessed models based on different distributions (normal, Poisson and negative-binomial distributions for numerical variables, and binomial distribution for binary variables). Then, we found the best fitting correlation matrix (unstructured, independent or compound symmetry). We chose the model with the lowest Quasi likelihood under Independence Criterion.

#### Ethical considerations

This study was approved by the Committee for the Protection of Human Subjects (the UTH IRB). All the

participants gave written informed consent and received a US\$50 gift card. Any potentially identifying information in the patients' records was erased.

## RESULTS

Of the 132 (22 nurse–physician pairs×6) phone calls, 12 were cancelled (two pairs) due to a no-show of the nurse or inability to contact the physician, nine were cancelled by the nurse conducting the experiment due to time constraints, and three were excluded from analysis due to errors in the case presentation (eg, background cue inadvertently presented with the case). A total of 108 phone consultations by 20 nurse–physician pairs were analysed. Of these, 88 cases contained a situation cue and 93 cases contained a background cue. In all, 57 cases were delivered without the SBAR form and 51 cases with the SBAR form. (In previous work we demonstrated that there was no difference in the communication between the groups.<sup>5</sup>)

In 14% of the cases (12/88), nurses failed to communicate the situation cue (table 2). Of these, in 42% (5/12) the nurses actually reported a misleading finding (eg, new onset fever rather than persistent fever). In 58% (7/12), the physicians asked questions aimed at eliciting the situation cue, but received the correct answer in only three cases. In summary, the situation cue was communicated in a total of 79 cases (independently by the nurse in 76 and elicited by the physician in three).

In 72% (67/93) of the cases, nurses failed to provide the background cue. Of these, in 7% (5/67) the nurses actually reported incorrect information (eg, a normal sodium level when the sodium was actually low). The physicians asked questions meant to elicit the cue in 12% (8/67) of the unreported background cues, and received the correct answer in five of the cases. In summary, the background cue was communicated in a total of 31 cases (independently by the nurse in 26 and elicited by the physician in five).

Physicians acted upon the communicated information (ie, appropriately) following 57% (45/79) of the communicated situation cues and 48% (15/31) of the communicated background cues. Providing an

appropriate response was case dependent (30%–88%,  $p=0.001$ ). In 25 cases, both the situation and the background cues were communicated. Of these, physicians acted appropriately regarding the situation cue in 72% (18/25). There was no association between nurses' use of a problem-specific SBAR form and the appropriateness of physicians' actions ( $p=0.5$  for the situation cues,  $p=0.14$  for the background cues). Online supplementary appendix D presents examples of various errors encountered in physicians' responses.

After controlling for the differences between subjects and cases, better performance regarding the situation cue was associated with a more active inquiry by the doctor regarding background information ( $p<0.001$ , table 3). Communication of the background cue was not associated with a significant improvement of physician performance regarding the situation cue ( $p=0.41$ ); however, the sample size was small.

## DISCUSSION

In nearly half of the cases (34/79, 43%), physicians failed to recognise and respond to the presentation of common serious clinical situations (eg, change in mental status, high potassium level). In a significant minority of cases (7/25, 28%), physicians failed to address the reason for the call even when presented with both the situation and background cues (eg, failing to address a case of an acute confusion despite the nurse's description of a patient with a behavioural change and a low sodium level). In over half of the cases (16/31, 52%), physicians failed to identify and treat the cause for the clinical condition. This was observed when the background cue involved non-trivial knowledge (eg, tacrolimus associated with high potassium levels) as well as when the knowledge was straightforward (eg, persistent fever in a patient who underwent surgery recently). Failure to communicate the necessary information accounted for a minority of the missed situation cues. On the other hand, inadequate reporting of information accounted for the majority of missed background cues. When nurses did not report the cues, physicians often failed to elicit

**Table 2** Rates of provided cues and appropriate actions

| Scenario                     | Situation cue provided | Appropriate actions* | Background cue provided | Appropriate actions* |
|------------------------------|------------------------|----------------------|-------------------------|----------------------|
| Behaviour (confusion) (n=17) | 100% (17)              | 88% (15)             | 53% (8)                 | 75% (6)              |
| Chest pain (n=15)            | 87% (13)               | 54% (7)              | None                    | None                 |
| Fever (n=17)                 | 77% (13)               | 31% (4)              | 30% (5)                 | 20% (1)              |
| Glucose (n=19)               | 90% (17)               | 30% (5)              | 26% (5)                 | 20% (1)              |
| High blood pressure† (n=20)  | None                   | None                 | 45% (6)                 | 67% (4)              |
| Medication (n=20)            | 95% (19)               | 74% (14)             | 35% (7)                 | 43% (3)              |
| Total                        | 90% (79/88)            | 57% (45)             | 33% (31/93)             | 48% (15)             |

\*Percentages presented are from the cases in which the cue was provided.

†The 'High Blood Pressure' and 'Chest Pain' cases had only one cue.

**Table 3** Call properties: comparison of appropriate to inappropriate actions

| Call properties   | Situation inappropriate action* (n=34) | Situation appropriate action* (n=45) | p Value | Background inappropriate action* (n=16) | Background appropriate action* (n=15) | p Value |
|---|--|--------------------------------------|---------|---|---------------------------------------|---------|
| Information communicated regarding the patient's situation (other than the situation cue) |  |                                      |         |   |                                       |         |
| Time to state reason for call (s)   | 44±117                                 | 36±39                                | 0.60    | 84±164                                  | 27±29                                 | 0.21    |
| Total no. of situation data items   | 2.3±0.8                                | 2.8±0.8                              | 0.32    | 2.6±1.0                                 | 2.7±1.0                               | 0.64    |
| Independently provided by the nurse   | 2.2±0.8                                | 2.7±0.8                              | 0.12    | 2.5±1.0                                 | 2.6±0.9                               | 0.75    |
| Situation data items physician asked about  | 0.1±0.3                                | 0.1±0.2                              | 0.32    | 0.2±0.5                                 | 0.3±0.7                               | 0.47    |
| Information communicated regarding patient's background (other than the background cue)   |  |                                      |         |   |                                       |         |
| Total no. of background data items  | 5.6±1.9                                | 6.4±2.9                              | 0.92    | 7.2±2.6                                 | 6.8±2.4                               | 0.22    |
| Independently provided by the nurse   | 4.8±1.9                                | 5.0±2.9                              | 0.41    | 6.2±2.9                                 | 5.3±2.2                               | 0.13    |
| Background data items physician asked about   | 0.7±1.1                                | 1.3±1.5                              | <0.001  | 0.9±1.4                                 | 1.4±1.8                               | 0.59    |
| Wrong information reported  | 9% (3)                                 | 9% (4)                               | 0.20    | 6% (1)                                  | 7% (1)                                | 0.96    |
| Reason for admission  | 88% (30)                               | 87% (39)                             | 0.81    | 100% (16)                               | 87% (13)                              | NS†     |
| Medical history   | 88% (30)                               | 80% (36)                             | 0.90    | 100% (16)                               | 87% (13)                              | NS†     |
| Home medications  | 12% (4)                                | 22% (10)                             | 0.80    | 44% (7)                                 | 40% (6)                               | 0.84    |
| Associated signs and symptoms   | 35% (12)                               | 18% (8)                              | 0.94    | 31% (5)                                 | 20% (3)                               | NS†     |
| Vital signs   | 24% (8)                                | 40% (18)                             | 0.98    | 50% (8)                                 | 40% (6)                               | 0.62    |
| Nurse gave an assessment/recommendation   | 24% (8)                                | 27% (12)                             | 0.98    | 18% (3)                                 | 28% (5)                               |         |
| Nurse gave a correct assessment/recommendation  | 6% (2)                                 | 16% (7)                              | 0.34‡   | 6% (1)                                  | 33% (5)                               | 0.34    |
| Length of talk (min)  | 3.5±2.4                                | 6.1±3.6                              | 0.04‡   | 5.6±2.4                                 | 6.3±4.5                               | 0.58    |

Communication of data items (other than the cues) regarding the patient's situation (eg, patient identification and location, whether the problem was urgent) and the patient's background (eg, reason for hospitalisation, prior medical history, vital signs, medication).

When applicable values are mean±SD.

\*Regardless of whether the background cue was communicated.

†Non-significant—statistical analysis not feasible due to a small number of observations.

‡Non-significant due to the number of variables evaluated (ie, type I error inflation).

the relevant information. Appropriate action regarding the situation cue was associated with active inquiry by the physician ( $p<0.001$ ).

Our study has several limitations. First, we evaluated only a limited number of non-trivial clinical scenarios. All cases were based on real patients, concerned the most common after-hours clinical problems and our experts considered the diagnostic challenge in each of these to be typical of those encountered in routine care. Nonetheless, it is possible that some of the cases were challenging resulting in a pessimistic estimation of physician performance. On the other hand, to reduce the possible bias from case selection, and since it is difficult to precisely define 'appropriate response,' we used unrealistically lenient evaluation measures based on an objective list of published indicators. These measures may have resulted in an over-optimistic estimation of performance. Consequently, we could not quantitate the risk to patient safety posed by after-hours calls. Notably, our objective was not to quantify the extent of errors, but rather to evaluate the critical thinking of physicians following the communication of key information in this setting.

A second limitation is that when designing the study, we did not anticipate that nurses would often not communicate important cues, in particular, background cues. Thus, we did not have adequate

statistical power to compare physician performance when given all required information (ie, both the situation and the background cues) with cases with missing information. However, a practicing physician is expected to elicit the necessary information from the reporting nurse. This was clearly not observed in our study.

A final limitation stems from conducting this study in a laboratory environment. Nurses and physicians had no prior knowledge of the patients. Further, nurses could not see the patients and were therefore deprived of important information. We cannot exclude the possibility that nurses would have acted differently when an actual physical patient was present for their evaluation and when actual patients depended on their actions. On the other hand, unlike real life settings, nurses were afforded ample time to review the patients' records in a distraction-free environment, and physicians were not sleep deprived when responding to the call.

#### Why were physicians erring?

The observation that in almost half of the cases physicians failed to act upon the information provided to them is very worrisome. These were common cases (in particular the situation cues), presented to highly

trained physicians and evaluated against unrealistically lenient criteria.

In the following section, we discuss possible reasons for failure in the different cases. As was found in previous studies, the majority of errors we witnessed could be attributed to problems with cognitive processes.<sup>19–21</sup> These include lack of knowledge, failure to recognise the significance of data or failure to synthesise all available data supporting the correct diagnosis (see online supplementary appendix D). In the case of persistent fever and several loose stools for example, 78% of physicians considered the diagnosis of *C difficile* colitis, yet only 31% considered the possibility of other hospital acquired infections. This demonstrates an anchoring effect and how a single detail (eg, loose stools) can focus even the most experienced physicians on a wrong diagnosis.<sup>22</sup> Interestingly, there were no other data to support the diagnosis of *C difficile* colitis (overt diarrhoea, leucocytosis, etc), but it seems that once physicians formulated a diagnosis they stopped searching for additional information (premature closure).<sup>23</sup>

In the Chest Pain case, despite a very suggestive description, only 54% of physicians entertained the possibility of a pulmonary embolus (PE). While similar rates have been cited in the literature,<sup>24</sup> we believe that in this case the reason for the error was a diagnosis momentum. The patient had a previous history of congestive heart failure and evaluated as such, despite a lack of other findings to support the diagnosis.<sup>23</sup> Of the eight physicians who did consider the possibility of PE, only four requested an imaging study of the chest, and of these only two noticed the patient suffered from kidney disease and ordered the indicated ventilation/perfusion scan. These results are consistent with known problems with the management of PE.<sup>24</sup>

Problems with clinical judgment (ie, considering all the relevant information but coming to the wrong conclusion)<sup>20</sup> may have been responsible for lack of treatment in 37% of severe low sodium cases (confused patient), 80% of cases in which a hospital acquired infection (fever) was suspected and for prescribing benzodiazepines in 26% of acute liver injury cases (medication); however, this is unlikely as a sole explanation, considering that these were relatively straightforward cases. Further, it is not clear why in 70% of high potassium cases (glucose) and in 41% of acute confusion (behaviour) cases (both, situation cues that were communicated by most nurses), physicians proceeded directly to symptomatic treatment without attempting to identify a cause for the problem.

We suspect that our findings are attributable to the nature of after-hours phone calls rather than any characteristic of the participating professionals. In clinical practice, these calls are made at times when resources are limited and the physicians are often unfamiliar with the patient. It is also possible that there is a cultural component whereby nurses are looking for

‘quick fixes’ that would decrease their workload, and physicians favour symptomatic treatment that would suffice until morning when the primary physician resumes responsibility.<sup>25</sup>

#### Mitigating the risk of phone consultations

After-hours phone calls are potentially dangerous due to communication failures, cognitive limitations, and possibly the limited resources and limited responsibility of the on-call physician. Simple interventions, such as problem-specific templates for communicating patient data, may reduce communication failures.<sup>2</sup> Our results show, however, that such interventions are not effective in isolation.<sup>5 26</sup> It seems that improving communication is necessary, but not sufficient to improve outcomes.

Eliminating phone communications would require significant changes in healthcare processes. Alternative solutions might be to provide access to a shared electronic health record (ie, a comprehensive patient record accessible from outside the hospital)<sup>27</sup> or to develop computerised systems designed to support both communication and decision-making. Establishing the necessary knowledge base and developing practical systems will be challenging for multiple reasons. However, identifying the common reasons for after-hours calls as well as common communication and cognitive errors may guide such efforts. More importantly, recognising and documenting the risk to patient safety associated with after-hours phone consultations are a necessary step toward changing the organisational culture to reduce or eliminate the potential of harm associated with these communications.<sup>25</sup> Further research is needed to identify the extent and impact of adverse outcomes associated with care provided over the phone.

#### CONCLUSIONS

After-hours phone calls are error prone. Both nurse communication and physician decision-making are problematic. Efforts to improve patient safety in this setting must address both communication and decision-making.

**Contributors** EJ: analysed data, drafted manuscript, revised manuscript. JPT: helped plan the study, was primarily responsible for conducting the experiments, participated in data analysis, currently convalescing from critical illness, and therefore was not available to approve the final manuscript. KOH: helped plan the study, participated in data analysis, helped draft and revise the manuscript. TRJ: helped plan the study, participated in data collection and analysis, helped revise the manuscript. CWJ: helped plan the study, led the data analysis, helped revise the manuscript. EVB: conceived of the study, obtained funding, supervised data collection and analysis, helped draft and revise the manuscript.

**Funding** This study was funded in part by a grant from the University of Texas System Patient Safety Committee (to EVB, JPT, KOH and TRJ) and by a training fellowship from the Keck Center Computational Cancer Biology Training Program of the Gulf Coast Consortium (CPRIT Grant No. RP101489 to EJ).

**Competing interests** None.

**Ethics approval** Committee for the Protection of Human Subjects at the University of Texas Health Science Center at Houston.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

- Petersen LA, Brennan TA, O'Neil AC, *et al*. Does housestaff discontinuity of care increase the risk for preventable adverse events? *Ann Intern Med* 1994;121:866–72.
- Bernstam EV, Pancheri KK, Johnson CM, *et al*. Reasons for after-hours calls by hospital floor nurses to on-call physicians. *Jt Comm J Qual Patient Saf* 2007;33:342–9.
- Tjia J, Mazor KM, Field T, *et al*. Nurse-physician communication in the long-term care setting: perceived barriers and impact on patient safety. *J Patient Saf* 2009;5:145–52.
- Manojlovich M. Nurse/physician communication through a sensemaking lens: shifting the paradigm to improve patient safety. *Med care* 2010;48:941–6.
- Joffe E, Turley JP, Hwang KO, *et al*. Evaluation of a problem-specific SBAR tool to improve nurse-physician phone communication in the after-hours setting: a randomized trial. *Jt Comm J Qual Patient Saf* 2013;39:495–501.
- Riesenberg LA, Leitzsch J, Little BW. Systematic review of handoff mnemonics literature. *Am J Med Qual* 2009;24:196–204.
- Cohen MD, Hillgoss PB. The published literature on handoffs in hospitals: deficiencies identified in an extensive review. *Qual Saf Health Care* 2010;19:493–7.
- Haig KM, Sutton S, Whittington J. SBAR: a shared mental model for improving communication between clinicians. *Jt Comm J Qual Patient Saf* 2006;32:167–75.
- Woodhall LJ, Vertacnik L, McLaughlin M. Implementation of the SBAR communication technique in a tertiary center. *J Emerg Nurs* 2008;34:314–17.
- Talbot R, Bleetman A. Retention of information by emergency department staff at ambulance handover: do standardized approaches work? *Emerg Med J* 2007;24:539–42.
- Field T, Tjia J, Mazor K. Randomized trial of a warfarin communication protocol for nursing homes: An SBAR-based approach. *Am J Med* 2011;124:179.e1–7.
- Marshall SA, Ruedy J. *On Call Principles and Protocols*. 5th edn. Philadelphia: Elsevier, 2011.
- Basow DS. ed. *UpToDate*. Waltham, MA, 2013.
- Confusion/decreased level of consciousness. In: Marshall SA, Ruedy J, eds. *On-call principles and protocols*. 5th edn. Philadelphia: Elsevier, 2011:69–81.
- Francis J, Young BG. Diagnosis of delirium and confusional states. In: Basow DS, ed. *UpToDate*: Waltham, MA. 2012. <http://www.uptodate.com/contents/diagnosis-of-delirium-and-confusional-states> (accessed 18 May 2012).
- Sterns RH. Overview of the treatment of hyponatremia. *UpToDate*, Basow DS. ed. *UpToDate*, Waltham, MA. 2012. <http://www.uptodate.com/contents/overview-of-the-treatment-of-hyponatremia> (accessed 21 May 2012).
- Dénes J, Keedwell A. *Latin squares and their applications*. New York: Academic Press, 1974.
- Pan W. Akaike's information criterion in generalized estimating equations. *Biometrics* 2001;57:120–5.
- Zwaan L, de Bruijne M, Wagner C. Patient record review of the incidence, consequences, and causes of diagnostic adverse events. *Arch Intern Med* 2010;170:1015–21.
- Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med* 2005;165:1493–9.
- Elstein AS. Thinking about diagnostic thinking: a 30-year perspective. *Adv Health Sci Educ Theory Pract* 2009;14:7–18.
- Arzy S, Brezis M, Khoury S. Misleading one detail: a preventable mode of diagnostic error? *J Eval Clin Pract* 2009;15:804–6.
- Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med* 2003;78:775–80.
- Roy PM, Meyer G, Vielle B, *et al*. Appropriateness of diagnostic management and outcomes of suspected pulmonary embolism. *Ann Intern Med* 2006;144:157–64.
- Barber N, Rawlins M, Franklin BD. Reducing prescribing error: competence, control, and culture. *Qual Saf Health Care* 2003;12:29i–32.
- Winters BD, Pham J, Pronovost PJ. Rapid response teams—walk, don't run. *JAMA* 2006;296:1645–7.
- O'Malley A, Samuel D, Bond A. After-hours care and its coordination with primary care in the US. *J Gen Intern Med* 2012;27:1406–15.