Telephone referrals by junior doctors: a randomised controlled trial assessing the impact of SBAR in a simulated setting

Neil James Cunningham,^{1,2,3} Tracey J Weiland,^{2,3} Julian van Dijk,² Paul Paddle,² Nicole Shilkofski,^{2,4} Nicola Yumei Cunningham^{1,3,5}

► Additional appendices are published online only. To view these files please visit the journal online (<http://dx.doi.org/10.1136/postgradmedj-2011-130719>).

¹Emergency Medicine, St. Vincent's Hospital, Melbourne, Victoria, Australia

²Education Centre, St. Vincent's Hospital, Melbourne, Victoria, Australia

³Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne, Victoria, Australia

⁴Anaesthesiology and Critical Care Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA

⁵Victorian Institute of Forensic Medicine, Melbourne, Victoria, Australia

Correspondence to

Dr Neil James Cunningham, Emergency Physician, Education Centre, PO Box 2900, St. Vincent's Hospital, 41 Victoria Parade, Fitzroy, Melbourne 3065, Australia; dr_cunningham@hotmail.com

Received 28 December 2011

Accepted 14 July 2012

Published Online First
14 August 2012

ABSTRACT

Objective To determine whether exposing junior doctors to Situation, Background, Assessment,

Recommendation (SBAR) improves their telephone referrals. SBAR is a standardised minimum information communication tool.

Methods A randomised controlled trial with participants and rating clinicians both blinded to group allocation. Hospital interns from a 2-year period (2006–2007) participated in two simulated clinical scenarios which required them to make telephone referrals. The intervention group was educated in SBAR between scenarios. Pre and post intervention telephone referrals were recorded, scored and compared. Six-month follow-up and year group comparisons were also made. An *objective rating* score measured the presence of specific 'critical data' communication elements on a scale of 1–12. Qualitative measures of *global rating* scores and participant self-rated scoring of performance were recorded. *Time to 'first pitch'* (the intern's initial speech) was also recorded.

Results Data were available for 66 interns out of 91 eligible. SBAR exposure did not increase the number of communication elements presented; *objective rating* scores were 8.5 (IQ 7.0–9.0) for SBAR and 8.0 (IQ 6.5–8.0) for the control group ($p=0.051$). Median *global rating* scores, designed to measure 'call impact', were higher following SBAR exposure (SBAR: 3.0 (IQR 2.0–4.0); control: 2.0 (IQ 1.0–3.0); $p=0.003$). *Global rating* scores improved as *time to 'first pitch'* duration decreased ($p=0.001$). SBAR exposure did not improve *time to 'first pitch'* duration.

Conclusion In this simulated setting exposure to SBAR did not improve telephone referral performance by increasing the amount of critical information presented, despite the fact that it is a minimum data element tool. SBAR did improve the 'call impact' of the telephone referral as measured by qualitative *global rating* scores.

INTRODUCTION

Telephone referral is a form of handover, which is defined as 'the transfer of professional responsibility and accountability for some or all aspects of care for a patient, or group of patients, to another person or professional on a temporary or permanent basis'.¹

Given the recognition of the importance of transferring critical information during a telephone referral it is surprising that this is a skill that many junior doctors are not taught at university.^{2–3} It is

not known how many medical students and doctors in Australia are currently trained in hand-over skills.⁴ Undergraduate medical students are commonly taught to present patients in a classical long format (the long case), including all available information. This is not an effective, efficient or acceptable method of transferring critical information in the workplace. The expectation from senior doctors is of a 'bullet point' approach, presenting only relevant information.

Telephone referrals have a clear association with patient safety. Effective communication during referrals, handovers and requests for consultation is fundamental to the safe delivery of care.⁵ Telephone referrals are common, everyday tasks. Junior doctors may perform this task several times a day depending upon their clinical placement. It is a task that a junior doctor would be expected to perform from their first day as an intern and typically involves referring up the hierarchical chain. There are identifiable systems issues associated with telephone referrals. These issues include the inexperience of junior staff new to hospital workplace environments, who may have knowledge deficits, confidence problems, language and culture challenges.

Errors in the clinical setting frequently occur at the junction between care providers,⁶ and the incidence of error is affected by human and situational factors. These factors include the challenges of working in a new, unknown environment and the inherent complexities of medical care. Thus it is critical that clinicians possess communication skills enabling them to 'speak the same language' when transferring critical information from one practitioner to another.

SBAR (Situation, Background, Assessment, Recommendation) is a minimum information tool that provides a structured and formalised model of communication between staff.⁷ This model of communication originated in the airline industry and military,⁸ and has been adapted for use in healthcare.^{6–10} SBAR is the most widely documented communication based cognitive aid in the clinical environment. There is a growing body of literature focusing on the use, implementation and evaluation of SBAR.^{7–11–13} In the clinical setting SBAR has the potential to improve the ability of 'pitcher' staff to collate and deliver critical information, improve the ability of 'catcher' staff to receive and interpret critical information, and improve safety by reducing errors occurring during referral.

Although the SBAR model appears to be valid on an intuitive level,¹³ and clinicians respond favourably to incorporating this format in practice, few studies have systematically assessed the efficacy of SBAR use in improving referral¹² or changing individual or group behaviour.^{9 11 12 14–16} Furthermore, there exists no tool to objectively assess the performance of requests for consultations, and it is not clear whether apparent improvements in performance for those exposed to SBAR result from changes in knowledge (expected data transfer requirements in referral), skills (structured verbal delivery), or attitudes (such as confidence). By identifying the possible mechanisms involved in producing improvements in the communication skills required for referral, future educational interventions may be structured in a way to facilitate the change required, and target those most in need of communication skills training.

The aim of this study was to determine whether exposure of junior doctors to an SBAR educational intervention would improve their performance of telephone referral. The primary outcome measurement was an *objective rating* provided by researchers blinded to the group allocation. Other outcome measures included *global rating* of telephone referral performance, participant *self-rating* scores, and *time to 'first pitch'* (time taken to present the referral) measurements.

METHODS

This study covered a 2-year period between 2006 and 2007 and was conducted at the Education Centre, St. Vincent's Hospital, Melbourne. In this randomised controlled trial both participants and rating clinicians were blinded to group allocation.

All interns employed at St. Vincent's Hospital Melbourne in the year groups beginning 2006 and 2007 were eligible for participation. These first year clinicians undergo a 1-year intern programme. This consists of working for 1 year as junior doctors under supervision in medical and surgical hospital wards or emergency departments. Participants were identified for study eligibility using a hospital database and were invited to participate via email.

Pre-test questionnaires

Participants attended the study venue and confirmed their consent to participate. After consenting and enrolling in the study, participants were requested to complete a questionnaire (online supplementary appendix 1). This purpose-designed questionnaire included basic demographics (age, sex, whether English was their primary language); self-perceived communication skills (scored using a Likert scale); and also assessed their level of assertiveness using the previously validated Rathus Assertiveness Scale.¹⁷ The Rathus Assertiveness Scale is a 30 item scale which, following item reversal, sums to form a total score. The questionnaire was delivered as an online survey using Survey Monkey.

Test scenario

This involved a case review and subsequent telephone referral.

Two simulated clinical scenarios were available, comprising of a case summary and a case file including supporting documentation (emergency department and admitting registrar admission notes, medication chart, ECG, chest X-Ray and pathology results). The cases reflected situations that the doctors would realistically encounter in the clinical environment, and were weighted to be of similar difficulty. 'Case 1' was a medical patient with chest pain, 'Case 2' was a surgical patient with abdominal pain. A computer-generated randomisation sequence

was used to randomly allocate participants to receive either Case 1 or Case 2 first.

Participants were given 4 min to review the allocated case summary and file, prior to contacting a senior member of staff via telephone, either the medical or surgical registrar. Two senior clinicians played the roles of the registrars remotely via telephone, and were instructed to use a semi-structured script with clear parameters and set prompts (online supplementary appendix 2). Each clinician was instructed not to deviate from the script provided.

The case scenarios were devised by the research team which included three emergency physicians, a surgical registrar and a critical care nurse. The scenarios were not modified during the study period and the registrar receiving calls was blinded to group allocation of the caller. All telephone conversations were audio recorded for later scoring. A review of audio recordings confirmed no evidence of deviation from the script.

The test scenario was immediately followed by: completion by the participant of a self-rating performance score (online supplementary appendix 3); and a debrief. This standardised 'emotion-centred' debrief involved a discussion of how participants felt about the scenario rather than clinical (diagnostic/management) or performance issues. A series of standard questions guided this debrief to ensure consistency (online supplementary appendix 4), and the duration of each debriefing was approximately 5 min. This debrief was devised by the research team which included two members with extensive experience in simulation debriefing.

Intervention

The SBAR educational intervention was a one-to-one structured didactic programme introducing and explaining the SBAR method of clinical handover, and its application in telephone referrals.

A series of points were always made to provide consistent delivery—the difference between the expectations of senior staff regarding clinical information presented in examinations, ward rounds, telephone referrals; the role and structure of the SBAR tool; the importance of planning a referral prior to picking up a telephone; making a decision regarding an acceptable outcome prior to picking up the telephone.

The duration of the intervention was approximately 10 min and the same researcher delivered every intervention. Participants were not provided with an opportunity to practice SBAR during the educational intervention. An SBAR handout (online supplementary appendix 5) was provided and the interns were encouraged to refer to this during future referrals. Immediately following the intervention participants were asked if they had been previously exposed to SBAR.

Randomisation

All study participants underwent baseline testing with a standardised scenario. Following this, the participants were randomly allocated into one of two groups (figure 1):

1. Intervention group—standardised debriefing + SBAR education or;
2. Control group—*standardised debriefing only*

Randomisation was achieved by using a computer generated block randomisation sequence placed in a bank of opaque envelopes. The researcher became unblinded at this point, opening the pre-allocated envelope to reveal group allocation. The participants remained unaware of group allocation.

Figure 1 Comparison between control and intervention Situation, Background, Assessment, Recommendation (SBAR) groups.

Participants randomised to the intervention group received the standardised debriefing and the SBAR educational intervention, before moving on to the alternate case scenario.

Participants randomised to the control group did not receive the intervention at this point and proceeded to the alternate case scenario. To ensure that the control group was not deprived of an educational opportunity the SBAR education intervention was provided at the end of the session.

Scoring methods

The researchers used several methods of scoring as outlined below. Scoring was undertaken after completion of all data collection and was performed on numbered audio data files by a single senior clinician (emergency physician) who was blinded to group allocation (SBAR or control), case order and time/date of recording. The recordings were reviewed in random order and each one was reviewed twice, and on separate occasions—once to complete the *objective rating* score, and once to complete the *subjective global* score. All data were entered directly into an excel spreadsheet for later analysis.

A second senior clinician rated 30% of files drawn equally from pre-test and post-test observations assessment of consistency. Inter-rater reliability assessments revealed adequate consistency between raters for subjective and objective rating scales (Cohen's $\kappa > 0.7$). Additionally, both the *objective rating* and *global rating* scales demonstrated adequate internal consistency (Cronbach's $\alpha > 0.7$) at each time point, thereby permitting scale summation.

Objective rating score (primary measure)

This scoring method utilised a pre-determined rating tool to record the presence or absence of specific data items presented in the recorded conversation (online supplementary appendix 6). The form was devised by the research team using Delphi methodology involving five clinicians and four iterations, and was designed to analyse 12 expected critical data components of a clinical referral of this type. The form was modified during a preliminary trial period, no study participants were exposed during this time. The version used in the study was tested between two raters to ensure reliability. The final objective

rating score out of twelve was calculated by the presence or absence of the predetermined 'critical' data components as noted in the 'Objective data tick box' (online supplementary appendix 6). Additionally, the percentage of critical information conveyed was calculated from this total.

Global rating score

Raters were asked to measure the call impact—the ability to 'get the message across.' They were instructed not to assess or comment on any specific clinical data components or adherence to SBAR structure. A free text option was provided to allow the scorer to write their reasons for the allocated score. Scores were rated as poor, fair, good or excellent (online appendix 7).

Time to 'first pitch'

This was defined as the 'time taken from when the intern started talking, to when there was a clear finish and an expectation of a response.' The finish was marked by either by a closing statement, direct question or excessively long pause. This definition was determined by the two coding researchers prior to review of audio recordings. Minor introductory exchanges were not viewed as a termination of the pitch. This time was measured in seconds and was obtained from the recorded files retrospectively. We did not undertake reliability testing of time to first pitch.

SELF-RATING SCORE

The participant completed a self-rating of their own performance immediately after each case scenario (online supplementary appendix 3).

ASSOCIATIONS

We looked for associations between performance scores (*objective rating*, *global rating* and *self-rating*) and: *time to 'first pitch'*; baseline assertiveness scores (Rathus); gender; whether English was a first language or not; scenario type (medical or surgical); individual data elements presented in recorded telephone referral conversations.

FOLLOW-UP TESTING

The main comparison was between the Control and SBAR intervention groups (figure 1). This allowed for an assessment of the immediate effect of the SBAR intervention. Additional follow-up testing was performed which was not part of the randomised controlled trial, this is outlined below.

The sequential intern year groups were all first year clinicians who were 'SBAR naïve'—they had no previous exposure to SBAR training. They were tested at different times during the year, which would likely have had a bearing upon their level of 'workplace experience' and familiarity with the task being studied. The **2006** interns were tested in the final quarter of their intern year. This meant that they were 'experienced.' Their baseline level of telephone referral skills had developed via the traditional 'learn as you go' method and they were not exposed to SBAR until the end of their intern year. The **2007** interns were tested at two points, early in the year when they were 'inexperienced.' They had additional 6-month follow-up testing 6 months later when they were both 'experienced' and had been exposed to SBAR (6 months earlier)—this follow-up testing was not part of the randomised controlled trial.

The following sub-group comparisons were made: —'Early year' testing and '6-month follow-up' (2007 interns only). This assessed skill retention/progress. All 2007

participants were invited to a re-testing session 6-months following their initial testing. They were tested using the same method as described above with one of the two standardised patient case scenarios. The case was randomly assigned using a computer-generated sequence. There was no specific follow-up SBAR training, and at the time of the study there was no systematic support or reinforcement of the application of SBAR in the clinical environment.

—'2006 end of year' and '2007 end of year' intern groups. This was a comparison between one group whose referral skills developed via the traditional 'learn as you go' method ('2006 end of year'), and a group who were exposed to SBAR early in their intern year ('2007 end of year').

DATA ANALYSIS

Data were analysed using SPSS V.15.0. Frequencies and percentages were calculated for categorical data, and mean (SD) and median (IQR) were calculated for continuous variables. Preliminary tests of assumptions revealed violations of normality for all data except assertiveness data. The Mann–Whitney U Test was used to compare differences between two independent groups for all outcomes except assertiveness, which was assessed using independent samples t-test. Wilcoxon signed rank test was used to assess changes over time (pretest-posttest). Fisher's exact test was used 2X2 contingency tables. Spearman's correlation was used to assess bivariate associations. For all inferential tests, two-tailed tests of significance were used and α was set at 0.05.

SAMPLE SIZE CALCULATIONS

In order to detect a difference with a moderate effect size (Cohen's $d=0.50$) on the objective rating (primary outcome) using a two-tailed test, we calculated that a sample size of 34 would be required assuming power of 80% with α set at 0.05. This was based on estimates only in the absence of pilot data.

RESULTS

Sample size and participant demographics

There were 91 interns eligible for enrolment over the 2-year period of study. Seventy-two took part, 31/45 (69%) from 2006 and 41/46 (89%) from 2007, 19 interns were excluded due to logistical reasons such as country rotations, a further three interns declined to take part in testing. In a further three cases there were incomplete data due to audio recording problems resulting in a full set of baseline data (including audio recording) for 66/91 interns (72.5%).

Only two participants had previously heard of SBAR, one from each experimental group. Both were from the 2007 year group (1 control, 1 intervention) and both were included in the analysis.

The sample demographics of age, gender and presence of 'English as a first language' were similar to those eligible for participation (table 1).

Comparison between control and intervention (SBAR) groups

Objective rating score

At post-test, *objective rating* scores (presence of data items) were 8.5 (IQR 7.0–9.0) and 8.0 (IQR 6.5–8.0) for SBAR and control groups respectively ($p=0.051$).

Global rating score

At post-test, *global rating* scores were 3.0 (IQR 2.0–4.0) and 2.0 (IQR 1.0–3.0) for SBAR and control groups respectively ($p=0.003$).

Time to 'first pitch'

At post-test, median *time to 'first pitch'* was 83 s (IQR 71–110) and 80 s (IQR 61–114) for SBAR and control groups respectively ($p=0.852$). Those exposed to SBAR showed a significant reduction in median *time to 'first pitch'* between baseline and post-test. There was no significant difference between groups in baseline assertiveness scores and *time to 'first pitch'* at any point (table 2).

Self-rating score

No significant differences were detected in mean subjective *self-rating* scores from baseline to post-test, between groups. No significant differences were seen at 6-month follow-up or in association with baseline assertiveness scores (table 2).

Associations

Neither *objective rating* nor *global rating* scores (dichotomised as poor/fair vs good/excellent) varied according to baseline level of assertiveness; gender; English as a first language; or scenario type (medical/ surgical).

Time to 'first pitch'—Across the whole 2-year sample *global rating* scores (poor/fair/good/excellent) were negatively correlated with *time to 'first pitch'* with scores improving as duration decreased ($\rho=-0.488$, $p<0.001$, table 3).

Individual 'critical information' data elements—26 of 30 (86.6% (95% CI 70% to 95%)) participants who received a good/excellent *global rating* score provided a 'diagnosis (or diagnosis unknown)' during the referral conversation compared with 15/37 (41% (95% CI 26% to 57%)) who received a fair/poor *global rating* score ($p<0.001$). Eleven of 30 (37% (95% CI 22% to 55%)) who received a good/excellent *global rating* score compared with 5/37 (14% (95% CI 5.4% to 28%)) who received a fair/poor *global rating* score stated 'further resource requirements' ($p=0.043$). Other *objective rating* components were not significantly associated with *global rating* scores.

Weak correlations were found between *objective rating* and *global rating* scores, and *objective rating* scores and *time to 'first pitch'* (table 4).

Follow-up testing

'Early year' testing and '6-month follow-up' (2007 interns only)

2007 *Early year* interns showed no change between baseline (pre SBAR exposure) and '6-month follow-up' (SBAR exposed) for objective rating scores or time to 'first pitch'. Global and self-

Table 1 Comparison of demographics between study participants randomised to control or SBAR intervention, and to those eligible for study participation

Variable	SBAR	Control	Total	P	Eligible (n=91)
Age, mean (SD)	25.8 (2.2)	25.5 (1.9)	25.6 (2.1)	NS	25.5 (2.1)*
Gender, male	19 (54.3%)	12 (36.4%)	31 (45.6%)	NS	38/88*
1st language English	31 (88.6%)	28 (84.8%)	59 (86.8%)	NS	Unknown

*Data unavailable for 1 person.

SBAR, Situation, Background, Assessment, Recommendation.

Table 2 Objective rating, global rating, subjective self-rating, assertiveness and time to 'first pitch' total scores group

Measure	Group	Baseline	Post test	p Value*	6 m follow-up	p Value*
Objective rating score (median, IQR)	SBAR	8.0 (7.0–8.25)	8.5 (7.0–9.0)	0.056	8.0 (7.0–9.5), n=13	0.253
	Control	7.0 (6.5–8.0)	8.0 (6.5–8.0)	0.229	8.0 (6.5–9.0), n=13	0.655
p Value†		0.169	0.051		0.210	
Objective rating percentage of critical information conveyed (median, IQR)	SBAR	66.6 (58.3–68.8)	75.0 (58.3–75.0)	0.056	66.6 (58.3–79.2)	0.253
	Control	62.5 (58.3–66.7)	66.7 (58.3–75.0)	0.084	66.7 (54.2–75.0)	0.655
p Value†		0.311	0.088		0.223	
Global rating score (median, IQR)	SBAR	2.0 (2.0–3.0)	3.0 (2.0–4.0)	0.043	4.0 (3.0–4.0), n=13	0.024
	Control	2.0 (1.5–3.0)	2.0 (1.0–3.0)	0.626	3.0 (3.0–3.5), n=13	0.004
p Value†		0.670	0.003		0.113	
Subjective self-rating score (median, IQR)	SBAR	28.0 (26.0–32.8)	29.0 (26.0–32.0)	0.204	32.0 (28.0–33.5), n=13	0.009
	Control	27.5 (25.0–31.0)	28.5 (27.0–30.3)	0.459	31.0 (28.5–32.0), n=13	0.075
p Value†						
Assertiveness score (Mean, 95% CI)	SBAR	4.47 (–2.7 to 11.7)	Not collected		Not collected	
	Control	10.0 (–8.6–8.7)				
p Value‡		0.852				
Time to 'first pitch' (median, seconds)	SBAR	101 (72–127)	80 (71–110)	0.017	80 (58–135), n=12	0.480
	Control	81 (68–117)	80 (61–114)	0.376	80 (68–104), n=12	0.136
p Value†		0.238	0.638		0.799	

*Comparison to baseline.

†Mann–Whitney U test.

‡Student t test.

SBAR, Situation, Background, Assessment, Recommendation.

rating scores for these interns had improved between 'early year' baseline and '6-month follow-up' (table 5).

'2006 end of year' and '2007 end of year' intern groups

This was a comparison between two intern groups who were all considered workplace 'experienced', nearing the end of their first year as a doctor. One group's referral skills had developed via the traditional 'learn as you go' method ('2006 end of year.') The other group had additionally been exposed to SBAR early in their intern year and were being tested at the 6-month follow-up stage ('2007 end of year.') Objective rating scores and median time to 'first pitch' did not differ between these groups. Experienced interns exposed to SBAR ('2007 end of year') had significantly better global rating scores (n=26) compared with experienced interns who had not been exposed to SBAR ('2006 end of year') tested at baseline (n=31) (table 6).

DISCUSSION

This randomised controlled trial attempted to assess whether exposure to brief training in the structured communication tool, SBAR, would improve the transfer of critical information between two individuals: the 'pitcher' (intern) and the 'catcher' (registrar). Despite SBAR being a minimum information tool we found no improvement in our primary measure—an *objective rating* score measuring the presence of specific 'critical data' communication elements presented by the participants during the test referrals.

Table 3 Time to 'first pitch'

Global rating score group	Time to 'first pitch' median (seconds)	Range (seconds)
Poor	122	94–136
Fair	94	76–120
Good	76	71–101
Excellent	65	57–77

p=–0.489, p<0.001.

SBAR is widely regarded as a useful communication tool in clinical handover, and we were able to show improvements in blinded *global rating* scores following exposure to SBAR. Our primary measure results suggest that either SBAR is not working the way we thought it would (increasing critical data transfer); we were unable to measure what was 'useful' versus 'not useful' data; or, more likely, improvements in telephone referral skills come about by a combination of problem recognition, strategy employment and skill practice and feedback. These latter factors are hard to quantify and depend upon individual and systemic factors such as availability of targeted skill teaching, using SBAR or other communication strategies.

All interns presented a similar proportion of the expected critical data elements of a telephone referral—SBAR did not affect the *amount* of critical data relayed. However, the *type* of critical data was affected. Interns with *global rating* scores of 'excellent' and 'good' were more likely than others to have stated a '*diagnosis (or diagnosis not known)*' and '*resources required*' during the referral conversation. Highlighting the importance of these particular data elements during handover education may be useful but further study would be warranted.

Exposure to just 10 min of one-to-one structured training in SBAR improved qualitative *global rating* scores (the ability to 'get the message across') immediately compared with controls receiving a debrief only. This immediate effect was observed irrespective of whether SBAR training was received at the start or the end of the internship year. The improvement persisted at 6 months but this may represent improving data prioritisation, delivery skills, and confidence in referral, and is likely to be attributable to practice in the clinical environment rather than solely the educational intervention itself. However, when comparing 'end of year' intern groups, the participants who were exposed to SBAR early in their intern year outperformed (*global rating*) those who had not—interns undergoing the traditional method of 'learning as you go.' There is the possibility that these improved *global rating* scores seen in the '2007 end of year' group were due to participation in the study and being made to think about the cognitive processes of a good telephone referral rather than SBAR specifically.

Table 4 Correlation matrix for main outcomes at post-test

	Objective rating score	Global rating score	Subjective self-rating score	Time to 'first pitch'
Objective rating score	—	—	—	—
Global rating score	$\rho=0.278$ $p=0.023$, $n=67$	—	—	—
Subjective self-rating score	$\rho=0.085$ $p=0.486$, $n=69$	$\rho=0.101$ $p=0.414$, $n=67$	—	—
Time to 'first pitch'	$\rho=0.231$ $p=0.058$, $n=69$	$\rho=-0.489$ $p<0.001$, $n=67$	$\rho=0.101$ $p=0.414$, $n=67$	—

Although SBAR did not affect the amount of 'critical data' relayed, it did reduce the time taken to present the referral. This is consistent with other studies involving the use of handover or referral tools and suggests that control participants may have presented extra 'non-critical' information.¹⁸ This may be mistimed, poorly organised data, or the insertion of 'pseudo-information'.¹⁹ While we have not explored the amount or nature of this extra information, this is worth further investigation. *Global rating* scores varied significantly according to *time to 'first pitch'* with scores improving as duration decreased. Thus conciseness appears to be a feature of effective telephone referral as determined by clinician raters of this study.

Although duration of referral was comparable between experienced and inexperienced interns, participants exposed to SBAR were more effective than controls as indicated by *global rating* scores. This suggests that the effect of exposure to SBAR on effectiveness (*global rating*) was independent of the phenomenon of becoming more concise during telephone referral.

Improvements in performance for those exposed to SBAR may result from changes in knowledge, skills or behaviours. The mechanisms of improving referral effectiveness in this situation are likely to be a combination of all three. Other studies have suggested that use of an SBAR minimum information tool helps staff 'know what they should say' when making referrals.¹⁵ Although we expected to show improvement in knowledge (presence of expected critical data to be transferred during referral), this was not the case. Skill and behaviour improvement was noted, SBAR made a difference to *global rating* scores, these ratings could be considered to encompass what a senior clinician would expect from a typical telephone referral. The difference between the objective rating and global rating results suggest that senior clinician's perception of a 'good' referral is not just confined to the inclusion of an expected list of data to be transferred. This study's *global rating* results combined with the clinical experience of the research group would suggest that an effective referral encompasses an array of skills—choice and prioritisation of which data to present (and which to leave out), conciseness, structure, clarity of message, cross-checking of received instructions, etc.

The delivery of telephone referral up a hierarchical chain may affect an individual's confidence and assertiveness. Assertiveness was not shown in this study to be associated with improved scores (*objective rating* or *global rating*). An individual's confidence levels are likely to change during the intern year, most would be expected to improve slowly as a result of a combination of skill acquisition and positive feedback—the absence of these factors may limit or worsen confidence. Anecdotal feedback from participants included some who stated that they avoided calling registrars and had received repeated negative feedback from senior staff during these types of telephone referrals. The safety implications of avoiding communication relating to unwell patients are obvious. The psychological implications on an individual who receives repeated negative feedback might be less immediately noticed in a workplace where staff rotate frequently and work under limited supervision out of hours. Junior doctors are recognised to be a group working under considerable stress.²⁰ Specific communication training such as SBAR would provide an objective reference point for an intern for self-rating performance of telephone referrals that may mitigate specific incidents of negative feedback from other staff.

Further analysis of the underlying mechanisms involved in producing an improvement in telephone referral communication skills is needed to guide future education. The only previous study that has attempted to systematically test the impact of teaching a communication tool (ISBAR, a modification of SBAR) only tested one person who made a telephone call during a group simulated scenario.¹² This method resulted in only testing 20% of the available population and it is unclear whether that person self selected to make the call. The varied team dynamics of an immersive simulation may have affected some participants' decisions to either call or avoid calling, and the performance of that call. Our study attempted to test a specific skill in all of the interns in a population and to more closely analyse the way in which the tool improves performance.

The intervention used in the study was brief and therefore likely not the ideal method of training a junior doctor in this particular skill. Development of the complex skill set required for communication of critical information in a wide variety of

Table 5 Comparison of 'early year' testing (pre SBAR exposure) and '6-month follow-up' (exposed to SBAR)—2007 interns only

Measure	Group		p Value
	'Early year', Not SBAR exposed (tested at baseline)	'Six month follow-up', SBAR exposed (6-months post-test)	
Objective rating score (median, IQR)	8 (7–8), N=38	8 (7–9), N=26	0.251
Global rating score (median, IQR)	2 (1–3), N=38	3 (3–4), N=26	<0.001
Subjective self-rating score (median, IQR)	27.5 (25–31), N=38	31.0 (29–32), N=26	0.002
Time to 'first pitch' (median, IQR, seconds)	100.0 (72–128), N=37	78.0 (61–109), N=23	0.100

All Wilcoxon signed rank test.

SBAR, Situation, Background, Assessment, Recommendation.

Table 6 Comparison of '2006 end of year' group (experienced interns not exposed to SBAR) and '2007 end of year' group at 6-month post-test (experienced interns exposed to SBAR)

Measure	Group		p Value
	'2006 end of year', Not SBAR exposed (baseline testing)	'2007 end of year', SBAR exposed (6-months post-test)	
Objective rating score (median, IQR)	8.0 (7–8), N=30	8.0 (7–9), N=26	0.272
Global rating score (median, IQR)	2.0 (2–3), N=31	3.0 (3–4), N=26	<0.001
Subjective self-rating score (median, IQR)	29.0 (26–31), N=31	31.0 (29–32), N=26	0.024
Time to 'first pitch' (median, IQR, seconds)	76.0 (69–121), N=29	78.0 (61–109), N=23	0.632

All Mann–Whitney U test.

SBAR, Situation, Background, Assessment, Recommendation.

scenarios requires a constructivist approach.²¹ Real world, case-based learning and feedback promote the transfer of knowledge, skills and behaviours to novel situations. Identification of poor performers who may need additional assistance before and during intern year is also desirable to allow this group to attain a satisfactory level of critical communication, and avoid the detrimental effects of negative feedback in the workplace. The intervention and follow-up method used in the study was not designed to specifically identify these poor performers and we would recommend the development of institution specific plans for this purpose. In our hospital, we have used these findings to focus upon skill and scenario based handover teaching, with feedback on performance available from expert staff. These sessions allow junior medical staff who struggle with this skill to be identified and given extra assistance. Involvement of registrar staff in these teaching sessions has given them insight into, and practice with, this style of structured communication from the 'catcher's' perspective.

SBAR is an easily taught communication tool that can improve effectiveness of referral performance in junior doctors and provide them with a standardised approach to a complex task. This has clinical safety implications for junior doctors. Education and communication skills' training has the potential to reduce errors during telephone referral, a form of clinical handover which is a recognised area of risk in healthcare. The act of telephone referral requires a coordinated interaction between two persons, each of who are required to display an array of communication skills in order for safe transfer of critical information to occur. There are therefore numerous factors can lead to adverse events and risk reduction at telephone referral is difficult to measure. Error reduction and the important patient safety implications of this require further detailed study in order to search for methods of improving communication.

Strengths and Limitations

SBAR was not in common usage in this hospital environment during the testing period making it unlikely that the participants had reinforcement of learning in the clinical environment. However, we did not ascertain whether there was any sporadic support for the application of SBAR in the clinical environment.

This study is strengthened by the use of randomised controlled trial design with clinicians and participants blinded to allocation. The study is unlikely to have been affected by selection bias as no participants declined; reasons for non-participation were logistical; interns were rotated to external hospitals or unavailable due to night shift work during the testing phase. Despite this, the single site nature of this study may limit generalisability to other settings. The sample was however comparable to the population of interns eligible to participate in terms of demographics, however we cannot

exclude the possibility of bias resulting from loss to follow-up. The use of structured tools may also have minimised bias. Further bias at the 6-month follow-up time point may have occurred as the same case scenarios were used; the case may have either been discussed among participants following initial testing or partially recalled. One of the strengths of the study is that we attempted to standardise the referral interaction as would occur in a real setting, rather than expecting SBAR use as a referral monologue.

Post intervention testing did not assess whether or not the interns used the SBAR tool in the second scenario (or in the 6-month follow-up scenario). This was deliberate as the testing was designed to assess the presence of expected data components and effectiveness of the telephone referral, not whether the interns were able to use SBAR tool or not. This was intended to avoid a potential bias of the interns assuming that they were being assessed on their use of SBAR, rather than a broader assessment of their communication skills. Participants were exposed to the educational intervention and given a written SBAR tool to take with them but application and utilisation of SBAR in the clinical environment was not recorded.

Use of the same standardised cases were used at 6-month follow-up could potentially have affected the observed improvement in *global rating* scores. However, the cases were designed to be clinically straightforward with the score given based upon the ability to get the message across, rather than getting the clinical diagnosis right. Although the intern group tested were all in their first year of clinical work and diagnostic acumen per se was not being tested, there may be some bias in the quality of referral if an intern was unclear on the expected diagnosis for each clinical scenario. Every attempt was made to reduce this, the research team designed the scenarios with 'high signal' clinical cues, pitched at a diagnostic level at which a final year medical student would be expected to be.

Previous exposure to SBAR should have been a formal exclusion criterion. Only two participants had previous exposure 1 intervention, 1 control) and this should not have had a large affect upon analysis.

Potential bias may have been introduced due to testing of different groups at different times. Ideally, 'early year' and '6-month follow-up' testing would have occurred for both year groups. The study start date was delayed which precluded the 2006 'early year' testing, and SBAR education was introduced to all interns as part of the hospital teaching programme in 2008.

Additionally, although it was our intention to analyse differences at 6-month follow-up, analyses at this time point were affected by substantial loss to follow-up, thereby resulting in sample sizes that were too small for meaningful interpretation.

Finally our comparisons between 'end of year 2006' and 'end of year 2007' may be subject to low statistical power due to limited sample size.

Main messages

In this simulated setting **S**ituation, **B**ackground, **A**ssessment, **R**ecommendation (SBAR) did not increase the amount of critical information presented, despite the fact that it is a minimum data element tool.

SBAR did improve the 'call impact'—the ability to get the message across.

SBAR is an easily taught tool that can provide junior doctors with a standardised approach to a complex task, this may reduce errors during telephone referral.

Current research questions

Further research is needed to determine:

- the specific mechanisms of **S**ituation, **B**ackground, **A**ssessment, **R**ecommendation skill acquisition in order to tailor education.
- Whether specific training in telephone referral prior to intern year can reduce clinical risk in this handover task.
- Whether poor communicators can be identified prior to intern year in order to provide additional assistance in attaining satisfactory levels of critical communication skills.

Key references

- **BMA and NPSA**. *Safe Handover: Safe Patients*. London: National Patient Safety Agency. 2004:7.
- **Joint Commission on Accreditation of Healthcare Organizations (JCAHO)**. *Sentinel Event Statistics*. 2005.
- **Marshall S**, Harrison J, Flanagan B. The teaching of a structured tool improves the clarity and content of inter-professional clinical communication. *Qual Saf Health Care* 2009;**18**:137–40.

CONCLUSION

In this simulated setting SBAR did not increase the amount of critical information presented, as measured by *objective rating* scores, despite the fact that it is a minimum data element tool. SBAR did improve the 'call impact' of the telephone referral as measured by qualitative *global rating* scores. Concise presentation of information is important in effective referral. However, SBAR was not shown to be responsible for significant improvements in conciseness. Better telephone referrals contained data on the diagnosis (or lack of a clear diagnosis) and whether further resources were required. This study does not show increased data transfer as the sole mechanism for improvement in referral performance and further research in assessment of telephone referral communication skills is recommended. Practise with this tool and expert feedback provides an opportunity to obtain essential communication

skills in critical information transfer. Further research is needed to determine both the mechanisms of SBAR usage, and the most effective way of ensuring all junior doctors have appropriate skills in telephone referral.

Acknowledgements The authors are grateful to Dr Stuart Dille and Professor George Jelinek for comments and editorial assistance on drafts of this manuscript. The authors also acknowledge the contribution of the late A/Professor Andrew Dent in the study design.

Contributors All authors contributed extensively to the work presented in this paper. NJC, TW and JVD designed the experiment. NJC, PP, NS, NYC administered the experiment. All authors discussed the results and commented on the paper.

Funding This study was funded in part by The Windermere Foundation.

Competing interests None.

Ethics approval Ethics approval was provided by St. Vincent's Hospital Melbourne Human Research Ethics Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **BMA and NPSA**. *Safe Handover: Safe Patients*. London: National Patient Safety Agency. 2004:7.
2. **Bomba DT**, Prakash R. A description of handover processes in an Australian public hospital. *Aust Health Rev* 2005;**29**:68–79.
3. **Chandratilleke D**, Ong JA. Final Year Medical Students: Confidence in Providing Medical Handover. *Proceedings of Bridging the Gap: 16th Australasian Prevocational Medical Education Forum*; 6–9 November 2011, New Zealand.
4. **Arora VM**, Johnson JK, Meltzer DO, *et al*. A theoretical framework and competency-based approach to improving handoffs. *Qual Saf Health Care* 2008;**17**:11–4.
5. **Joint Commission on Accreditation of Healthcare Organizations (JCAHO)**. *Sentinel Event Statistics*. 2005.
6. **Leonard M**, Graham S, Bonacum D. The human factor: the critical importance of effective teamwork and communication in providing safe care. *Qual Saf Health Care* 2004;**13**:85–90.
7. **Hohenhaus S**, Powell S, Hohenhaus JT. Enhancing patient safety during hand-offs: standardized communication and teamwork using the 'SBAR' method. *Am J Nurs* 2006;**106**:72A–B.
8. **Weiner EL**, Kanki BG, Helmreich RL. *Cockpit Resource Management*. Massachusetts: Academic Press, 1995.
9. **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.
10. **Woodhall LJ**, Vertacnik L, McLaughlin M. Implementation of the SBAR communication technique in a tertiary centre. *J Emerg Nurs* 2008;**34**:314–17.
11. **Mikos K**. Monitoring handoffs for standardization. *Nurs Manage* 2007;**38**:16–20.
12. **Marshall S**, Harrison J, Flanagan B. The teaching of a structured tool improves the clarity and content of interprofessional clinical communication. *Qual Saf Health Care* 2009;**18**:137–40.
13. **France DJ**, Stiles R, Gaffney EA, *et al*. Crew resource management training - clinician' reactions and attitudes. *AORN J* 2005;**82**:213–24.
14. **Riesenberg LA**, Leitzsch J, Little BW. Systematic review of handoff mnemonics literature. *Am J Med Qual* 2009;**24**:196–204.
15. **Clark E**, Squire S, Heyme A, *et al*. The PACT Project: improving communication at handover. *Med J Aust* 2009;**190**(11 Suppl):S125–7.
16. **Thompson JE**, Collett LW, Langbart MJ, *et al*. Using the ISBAR handover tool in junior medical officer handover: a study in an Australian tertiary hospital. *Postgrad Med J* 2011;**87**:340–4.
17. **Rathus SA**. A 30-item schedule for assessing behavior. *Behav Ther* 1973;**4**:398–406.
18. **Catchpole KR**, De Leval MR, McEwan A, *et al*. Patient handover from surgery to intensive care: using Formula 1 pit-stop and aviation models to improve safety and quality. *Pediatr Anesth* 2007;**17**:470–8.
19. **Beach C**, Crockery P, Shapiro M; Center for Safety in Emergency Care. Profiles in patient safety: emergency care transitions. *Acad Emerg Med* 2003;**10**:364–7.
20. **Markwell AW**, Wainer Z. The health and well-being of junior doctors: insights from a national survey. *Med J Aust* 2009;**191**:441–4.
21. **Ertmer PN**, Newby T. Behaviourism, cognitivism, constructivism: comparing critical features from an instructional design perspective. *Perform Improv Q* 1993;**6**:50–70.