Translating teamwork behaviours from aviation to healthcare: development of behavioural markers for neonatal resuscitation

E J Thomas, J B Sexton, R L Helmreich

....

Improving teamwork in healthcare may help reduce and manage errors. This paper takes a step toward that goal by (1) proposing a set of teamwork behaviours, or behavioural markers, for neonatal resuscitation; (2) presenting a data form for recording observations about these markers; and (3) comparing and contrasting different sets of teamwork behaviours that have been developed for healthcare. Data from focus groups of neonatal providers, surveys, and video recordings of neonatal resuscitations were used to identify some new teamwork behaviours, to translate existing aviation team behaviours to this setting, and to develop a data collection form. This behavioural marker audit form for neonatal resuscitation lists and defines 10 markers that describe specific, observable behaviours seen during the resuscitation of newborn infants. These markers are compared with those developed by other groups. Future research should determine the relations among these behaviours and errors, and test their usefulness in measuring the impact of team training interventions.

recording observations about these markers; and we compare and contrast different sets of behavioural markers for teamwork that have been developed for healthcare. These behavioural markers for teamwork could be used to help measure baseline team behaviour and the effectiveness of team training interventions for neonatal resuscitation. Furthermore, our comparisons with other sets of behavioural markers suggest that many of these markers are generalisable to settings other than neonatal resuscitation.

....

H ealthcare organisations and researchers[1–4] are trying to measure and improve teamwork among providers. Despite a large amount of experience in other industries,[5] relatively little is known about how to measure and improve teamwork in healthcare.[6] Measuring teamwork is a critical issue because regulators of healthcare—such as the US Joint Commission for Accreditation of Healthcare Organizations—are requiring organisations to implement some type of team training activities, and hospitals are spending large sums of money paying consultants for team training seminars. These regulators and purchasers will want measurable results, as will patients and their families.

In this paper we describe our initial research efforts to measure teamwork during neonatal resuscitation, a patient care process that requires teamwork and one for which improvements in quality of care could benefit patients and families. We describe the development of behavioural markers (observable teamwork behaviours) for neonatal resuscitation (University of Texas behavioural markers for neonatal resuscitation, UTBMNR); we present a data form for

The aviation experience with teamwork measurement

The Institute of Medicine and others have encouraged healthcare providers to look to the aviation industry because of its long history of measuring and improving teamwork to prevent and mitigate errors.[7 8] One method used in the aviation industry is the Line Operations Safety Audit (LOSA).[9] Over the last 17 years the University of Texas Human Factors Research Project has developed this observational methodology which uses expert observers in the cockpit during normal flights to record errors, threats to safety, and behaviours used to prevent and manage error. Data have been collected on more than 5000 domestic and international airline flights on which expert observers systematically record the behaviour of flight crews during normal operations under conditions of absolute confidentiality (an approach supported by the Federal Aviation Administration and the International Civil Aviation Organisation).

The LOSA team behaviours, also called behavioural markers, are associated with aircrew performance,[10 11] although they have not been studied employing experimental study designs. These behavioural markers are also used in some forms of Crew Resource Management, an aviation program for improving teamwork. Identification of a similar set of markers for healthcare may, after thorough evaluation, assist with the improvement of teamwork and perhaps even inform medical and nursing curriculum reforms that could help meet the Institute of Medicine's goal to ''improve team function''[12] and thereby improve the quality of care for patients.

See end of article for authors' affiliations
....

Correspondence to:
Dr Eric J Thomas, The University of Texas, Houston Medical School, 6431 Fannin MSB 1.122, Houston, TX 77030, USA; Eric.Thomas@uth.tmc.edu
....

....

Abbreviations: LOSA, line operations safety audit; NICU, neonatal intensive care unit; UTBMNR, University of Texas behavioural markers for neonatal resuscitation

A major concern was to avoid direct borrowing and application of the LOSA behaviours to healthcare because the two work environments have many important differences.[13] Therefore, data from focus groups of neonatal providers, surveys, and observation of resuscitations were used to identify some new teamwork related behaviours, and to translate the existing LOSA team behaviours to this setting. The ultimate purpose of this research is to develop tools to measure these behaviours, to determine the relations between team behaviours and errors, and to improve teamwork in order to reduce and better manage error.

DEVELOPMENT OF BEHAVIOURAL MARKERS FOR TEAMWORK IN NEONATAL RESUSCITATION
Overview
Although there are generalisable aspects of team behaviour in managing safety critical situations,[5] it is premature to assume that the same behaviours that manage threat and error in a cockpit will also manage error while resuscitating an infant. Therefore we tried to translate the behaviours to this new setting instead of just directly applying them. Subsequent research will determine which markers are useful for preventing and managing errors. This "translation" process relied upon focus groups of neonatal providers, survey data, and review of video-recorded neonatal resuscitations.

Definitions
In aviation, these teamwork related behaviours are called "threat and error management behaviours" because they have been shown to help manage threats and errors in commercial aviation cockpits. In this paper we use the term "behavioural markers" instead of "threat and error management behaviours". Behavioural markers are observable, non-technical behaviours that contribute to performance within a work environment.[3 14] This terminology is more generic and does not assume that these behaviours have been proven useful for managing threat and error during the resuscitation of newborns.

Focus groups
Focus groups were used because of the paucity of previous research on working together in neonatal intensive care units (NICUs) and our desire to understand provider perspectives. To our knowledge there have not been previous studies that ask NICU providers for their impressions of working together. Focus groups are an accepted method of initial research in areas where little or no research has been conducted. Our primary goal was to elicit provider perceptions about how they worked together and what influenced their work. This starting point was based upon our belief that any list of behavioural markers should arise from, and therefore be relevant to, the complex and dynamic environments in which healthcare providers work. The results of the focus groups are reported in detail elsewhere.[15]

We conducted seven focus groups using a series of open ended questions to elicit descriptions of how providers work together in an NICU. The terms "teamwork", "errors", or "error management behaviours" were not used. We convened one group each of transport nurses, staff nurses, residents, fellows, attending physicians, and two groups comprising multiple providers. With three to seven participants per group, the study included 36 subjects.

We then used qualitative research methods to analyse the focus group transcripts.[16] Qualitative methods help to develop themes that can be used to build a theoretical foundation for the phenomenon being studied, as well as providing the details for developing instruments for future study. Two investigators carefully read the transcripts to comprehend the whole and highlight relevant data bits. Synthesis revealed three common themes (reported below) emerging from the data. Further analysis identified descriptive elements for each theme. Themes and their descriptive elements were revised as analysis proceeded, to give an accurate presentation of the participants' perspectives. Investigators searched for and considered data to compare and contrast with the established themes to verify inclusion. Given that some groups were composed of just one provider type and others comprised several, themes that arose from the different groups could be compared and contrasted in order to derive more valid and generalisable themes on how the providers worked together. The validity of the analysis was enhanced by other investigators' independent reviews of the transcripts, which resulted in revision of the themes.

Providers in the NICU identified three broad themes that affect the way they work together: the providers, workplace factors, and group influences. The provider theme was related to provider personalities, reputations, and "egos". Workplace factors included issues such as staffing and the organisation of care processes. Group influences included communication styles, relationships, and team functioning. Notably, teamwork was mentioned infrequently and there were contrasting views about it. Many providers talked about teams as being organised along provider types (a nursing team) or care processes (resuscitation team). But others talked about teams as being more inclusive and functioning as a "family" and a "well oiled machine".

Focus group participants did not generate behavioural marker themes per se; however, numerous participants mentioned several LOSA behavioural markers. These included: information sharing, teamwork, assertion, inquiry, workload distribution, leadership, and evaluation of plans (all defined below). These findings supported inclusion of these behavioural markers, but also the focus groups emphasised the unique complexity of this work environment and the numerous factors that influence how providers work together.

Surveys
We also reviewed 10 years of survey data from healthcare providers.[17 18] These surveys measured provider attitudes about topics relevant to patient safety such as teamwork, safety climate, working conditions, stress recognition, job satisfaction, and the management of their organisation. Significant percentages of providers consistently reported that it was difficult to speak up; that conflicts were not appropriately resolved; that it was not easy to ask questions; that it was difficult to disagree "up the hierarchy"; and that staffing levels were not sufficient to handle the number of patients. These attitudes supported the inclusion of behavioural markers such as inquiry, assertion, leadership, and workload management.

VIDEO RECORDING OF NEWBORN RESUSCITATION AND DEVELOPMENT OF THE BEHAVIOURAL MARKER AUDIT FORM
Next, we video-recorded consecutive resuscitations of infants born by caesarean section. Video recording allowed images to be stored for viewing at a later time, and for multiple reviews of each event. A team of MD and PhD researchers reviewed an initial group of five videos and determined that behaviours could indeed be identified on the videotapes. Subsequently, two investigators (EJT, JBS) viewed additional video recordings. While viewing each event the investigators discussed the care process and looked for observable behaviours. During this process we did not explicitly refer to LOSA error management behaviours, although both investigators were familiar with them. We then developed a list of candidate behaviours that was refined as additional

events were observed. The initial list was derived from the focus group and survey results, and initial viewings of the videos. After reviewing 20 video recordings (in addition to the initial five recordings) we were able to define and list examples of the behavioural markers (text box) and to create the University of Texas Behavioral Marker Audit Form (appendix 1). The form was then used to measure these behaviours in 113 additional video recordings of neonatal resuscitation.

Description of the Behavioral Marker Audit Form

The Behavioral Marker Audit Form for neonatal resuscitation is a one page form with three sections (event demographics, behavioural markers, and threats to patient care). Two scales—observability and frequency—are used to rate each behaviour. The observability scale allowed us to indicate how well a behaviour could be observed, and the frequency scale was used to indicate how often a behaviour occurred. More detailed definitions of the observability and frequency scales are provided in the next two sections of the paper. There is no scale to rate the quality of the behaviour because we do not yet know whether or how these behaviours prevent or manage errors. We used the form to record observations made while viewing the video recordings.

Definitions of the 10 behavioural markers for neonatal resuscitation

1. *Information sharing*. Providers verbally share information that relates to the assessment and treatment of the baby.
2. *Inquiry*. Providers ask questions of each other.
3. *Assertion*. An individual provider asserts their opinion (through questions or statements of opinion) during *critical times*. Assertion *does not* include routine statements or questions about a baby's heart rate, tone, colour, and respirations (these form part of information sharing or inquiry).
4. *Intentions shared*. A provider states their intentions before deviating from the norm. Statements about following routine neonatal resuscitation programme guidelines are not examples of intentions shared.
5. *Teaching*. Teaching is observed during the observation. This may be in the form of short and informal information exchanges. It may occur between any of the providers (for example, nurses can teach residents and vice versa).
6. *Evaluation of plans*. An explicit discussion about the status of the baby and the decisions made to get to the current situation.
7. *Workload management*. The workload is distributed among those present at the resuscitation. Tasks are appropriately prioritised.
8. *Vigilance/environmental awareness*. Providers remain alert and focused on the resuscitation.
9. *Teamwork overall*. This is a global assessment based upon the ratings of behaviours 1–8 above. It may also include other observations such as dynamic ''give and take'' among team members and non-verbal communications that are not explicitly defined in the behavioural markers.
10. *Leadership*. Leadership activities may include sharing of a mental model, assigning tasks, and sharing of information and opinion. This may be rated for any provider at the resuscitation.

Observability rating scale

We could not always hear everything that providers were saying because multiple providers may have been talking at once, or because of background noise from alarms or pagers. Therefore the observability rating allowed us to indicate how well we could observe the behaviour. If the behaviour was absent, observability was 0; if the behaviour was present, a rating of 1, 2, 3, or 4 was used to indicate how observable the behaviour was. For example, poor audio or a simple inability to interpret verbalisations may lead a rater to choose an observability rating of 1 or 2 for a given behaviour. If the behaviour was explicit and easily observed, the reviewer may choose a rating of 3 or 4 for observability. These ratings allow identification of clear examples of certain behaviours. A rating or 3 or 4 means it would serve as an example of the behaviour (not necessarily a ''good'' or ''bad'' example, just an observable example).

The observability rating is a marked departure from LOSA, where after thousands of observations a definitive list of behaviours exists. In this setting, we wanted to make sure that observers had the opportunity to note that a behaviour may not be present or observable. The observability rating also allows observers to qualify the confidence of their observations. This may be done either because of technical problems (poor sound) or because of uncertainty about whether a specific behaviour meets study definitions.

Each behaviour's observability is rated using the following scale:

- 0: None. The behaviour was not observed.
- 1: Poor observability, either because of consistently poor audio quality or an inability to determine whether verbalisations were related to the behaviour.
- 2: Minimum/inconsistent. There may be inconsistent audio quality or inconsistent ability to determine if verbalisations were related to the behaviour.
- 3: Standard observability. Almost all verbalisations can be understood and the observation could be used as an example of the behaviour being rated.
- 4: Outstanding observability. The behaviour is unusually explicit and clear. Would make an excellent example of the behaviour.

Frequency rating scale

Each behaviour is also rated on a 1–4 scale based upon its frequency. Unlike LOSA, these ratings are not anchored at the upper or lower end by any reference either to the quality of the behaviour or to safety or errors. For example, LOSA uses a rating of 1 to indicate poor performance that is a clear detriment to safety. In the setting of neonatal resuscitation there are no empirical data to tell us what a ''poor'' behaviour is. We do not know if a poor rating of one of these behaviours is correlated with decreased safety (increased errors). Therefore, instead of anchoring the scale as either being detrimental to safety (score of 1) or beneficial to safety (score of 4), we rated the frequency of the behaviour. If there were examples of a behaviour carried out in an obviously ''poor'' or unhelpful manner, this was noted in the comments section or in the individual rating section.

Use of this frequency rating emphasises that the Behavioral Marker Audit Form for neonatal resuscitation is currently a screening instrument in early stages of development. More detailed analyses of the video recordings, including comparisons with independent ratings of the events for the presence of errors, will help us to identify behaviours that can be called error management behaviours and subsequently taught to providers.

Frequency rating definitions are:

1: Rare. One or two *rare examples* of the behaviour occurred.
2: Isolated/minimum. There were *isolated examples* of the behaviour throughout the observation.
3: Intermittent/standard. There were *intermittent examples* throughout the observation.
4: Consistent. There were *consistent examples* of the behaviour throughout the observation.

The frequency scale was not used to count the absolute number of times a behaviour was observed.

Behavioural markers

The next section of the behavioural marker audit form for neonatal resuscitation contains the 10 behavioural markers we are currently rating. Below, each marker is defined and compared with LOSA markers.

1. Information sharing

This includes verbalisation of information that relates to the assessment and treatment of the baby. For example, verbalisation of heart rate, colour, tone, vocal cord visualisation, statements of opinion, advocating of views in noncritical moments, and other relevant observations or impressions about the baby's status. Discussions about Apgar scores are examples of information sharing. (Apgar scores are used to rate the health of a newborn infant.) Information sharing may take the form of ''baby talk''. For example, ''we like it when you cry'', ''you are nice and pink'' (assessment of colour), ''you are a strong little baby'' (assessment of tone).

Information sharing is usually observable and therefore the observability rating is often a 2 or 3. There were several LOSA markers that appear to be either not applicable to this setting or were so infrequent that we collapsed them into a new marker called ''information sharing''. Information sharing is not a LOSA marker, but it is a more general concept that includes LOSA behaviours such as monitoring and cross checking, contingency management, briefings, automation management, and communication environment.

2. Inquiry

Providers ask questions of each other. For example, ''What's the heart rate?'' ''What do you think of his tone?'' ''Should we intubate her?'' Like information sharing, inquiry is usually observable (2 or 3). Statements may also serve as questions depending upon the inflection of the voice. This marker exists in LOSA and its definition was not modified.

3. Assertion

An individual provider asserts their opinion (through questions or statements of opinion) during *critical times*. Assertion *does not* include routine statements or questions about a baby's heart rate, tone, colour, and respirations (these are part of information sharing or inquiry). The observability rating for assertion is often 0, meaning that the majority of the events observed did not contain examples of assertion. Assertion may be indirect, through the use of repeated questions (for example, ''do you really think the endotracheal tube is in the oesophagus?''). Assertion can overlap with other behaviours. For example, the previous quote is also an example of inquiry. An example of assertion and workload management is ''Is that your phone ringing? Let me help.'' This marker exists in LOSA and its definition was not modified.

4. Intentions shared

A provider states their intentions before deviating from the norm. Statements about following routine neonatal resuscitation programme guidelines are not examples of intentions shared. Thus defined, this behaviour may be infrequently observed (observability rating may often be 0). The LOSA marker is called ''plans stated''. We changed it to ''intentions shared'' because it more accurately described what we saw. ''Plans stated'' implies statement of an explicit plan. In neonatal resuscitation, we saw much more sharing of general intentions/next steps, but not explicit statements of multistepped plans.

5. Teaching

Teaching is assessed during the observation. This may be short and informal information exchanges. It may occur between any of the providers (for example, nurses can teach residents and vice versa). Teaching is often not seen (observability = 0).

There is no comparable LOSA marker because rates of on-the-job training in commercial aviation are far lower than in medicine, especially in hospitals with residency programmes and those affiliated to medical schools.

6. Evaluation of plans

An explicit discussion about the status of the baby and the decisions made to get to the current situation. This is often not observed, probably because of the standardisation of neonatal resuscitation. This marker exists in LOSA and its definition was not modified.

7. Workload management

The workload is distributed among those present at the resuscitation. Tasks are appropriately prioritised. This behaviour is usually observed, but not heard on tape. Therefore observability is usually a 2, 3, or 4, with 4 reserved for examples where there are explicit verbal decisions made about workload management.

This marker exists in LOSA and its definition was not modified.

8. Vigilance/environmental awareness

Providers remain alert and focused on the resuscitation. This is usually observable on the basis of non-verbal behaviours (rated 2–4), with 4 reserved for some examples where vigilance is explicitly addressed. The frequency rating is usually a 3 or 4. It is 1 or 2 where there are clear examples of loss of focus on the task, often because of interruptions such as social conversations, phone calls, twin births, and parental interruptions. These interruptions are recorded separately below. The same marker exists in LOSA.

9. Teamwork overall

This is a global assessment based upon the ratings of behaviours 1–8 above. It may also include other observations such as dynamic ''give and take'' among team members and non-verbal communications that are not explicitly defined in the behavioural markers. In LOSA the comparable marker is ''overall crew effectiveness rating''. Because the goal of the current study was exploratory and descriptive, we were not able to judge effectiveness, so the LOSA term was not appropriate.

10. Leadership

Leadership activities may include sharing of a mental model, assigning tasks, and sharing of information and opinion. This may be rated for any provider at the resuscitation. There is usually not a clear leader (either in deed or word), so the observability rating is often 0. Lack of leadership was not obviously detrimental to the process of care. This may be related to the standardised nature of these resuscitations and because most of the babies were not very ill. Leadership is often attributed to one person, but two people may show leadership; if so, the rating would represent the combined

leadership activities. The LOSA marker focuses on leadership shown by the captain. But the neonatal resuscitations were notable for not having one explicitly identified leader. Moreover, leadership roles were fluid and highly dependent upon team composition and experience level.

After the 10 behavioural markers, the form allows reviewers to rate individuals who differed from the team, to rate threats, and make additional comments.

Individual ratings

Was there an individual who differed in one of the above markers in a manner that set them apart from the rest of the providers? Note that one person may show leadership, but that is rated under "leadership" and not as an individual rating.

LOSA provides an opportunity to rate each member of the crew on technical proficiency and a composite assessment of crew resource management skills. In the past, LOSA collected individual ratings if an individual differed substantially from the rest of the crew, but current LOSA methodology has removed this step.

Threats and other comments

These may include under- or overstaffing; environmental distractions such as phone calls, twin birth, social conversations, alarms, parental interruptions; and other (including the complexity of the medical situation, illness of infant, twins, congenital malformations). Each threat is rated for observability. It is also rated (yes/no) for whether it was managed. The threat or problem can be managed without being solved. For example, the providers may believe they are understaffed and an attending is needed. The attending may never come, but the threat (understaffing) could still be managed. Recognising a threat as a threat (for example, telling a caller you will call back when done with the resuscitation) can qualify as managing the threat. Threat severity is used to indicate the degree to which the threat interfered with the resuscitation. Although we use the term "threat," it is not yet clear that these events can be empirically demonstrated to "threaten" the infant.

Other comments

Space for any observations not already included on the data form.

Candidate for additional review

Events with good examples of behaviours, or complicated events (for example, babies who require intubation) with a lot of interaction among team members, are candidates for additional review. Observers then write in the comment section why it is a candidate for additional review.

COMMENT

We have described the process we undertook to identify some new teamwork related behaviours (behavioural markers) for neonatal resuscitation and to translate the existing LOSA team behaviours to this setting. This process resulted in the development of the UTBMNR form (appendix). This form has been developed after careful review of decades of research on teamwork in aviation, recent survey and focus group data from healthcare providers, and video recordings of neonatal resuscitations. It represents our first attempt to translate (not borrow) and identify new behavioural markers relevant to the resuscitation of newborn infants. The form is being used as a screening instrument to rate the observability and frequency of 10 behavioural markers and observable threats to patient care.

These behavioural markers describe specific, observable behaviours (not attitudes, personality traits, or organisational factors) related to the experiences of frontline providers and are labelled using language relevant to the environment to which they are being applied. For example, we changed the name of several LOSA aviation markers because of their unclear application to the resuscitation of newborns. Of the 10 current markers, four are renamed or redefined (information sharing, intentions stated, teaching, and leadership).

Other aviation markers were excluded because they were not relevant to video recordings in this setting, or unobservable when using them. These include briefings, monitoring and crosschecking, and automation management. Our exclusion of these markers here does not preclude their use in behavioural observation in other clinical settings. Briefings were not seen on our video recordings but they do occur in some cases before arrival at the bedside. Because they were not observable they are not included on the UTBMNR form. However, briefings may be very important for neonatal resuscitation, as in other areas of medicine. For example, a neonatal resuscitation team often knows in advance about the scheduled delivery of a high risk infant. This advance notice gives the team an opportunity to discuss the resuscitation plans and think about complications that may arise and how they will be handled. In LOSA terminology this process would include both briefings and contingency management.

Although automation management may be important in anaesthesia or critical care medicine, it was not included as a behavioural marker here, simply because it was not observed to form a substantial part of the resuscitation of a preterm infant. In contrast, managing the "glass cockpit" and other technology in commercial aviation makes automation management very important. Similarly, fewer gauges and readouts made monitoring and crosschecking less relevant to our observations. If an observer saw a rare example of this behaviour it was considered part of information sharing.

These differences from LOSA and other behavioural markers highlight one of the limitations of our approach. We focused on a relatively limited aspect of the processes required to care for a newborn infant. Although we conducted focus groups and surveyed providers to obtain an understanding of the broad environment in which the providers work, our ultimate goal was to measure behaviour during resuscitation. In contrast, Carthey and colleagues[3] have developed a list of behavioural markers for paediatric cardiac surgery that include observations about organisational factors and other individual characteristics that may be important for patient care but are not readily observable. We strongly agree that organisational factors have powerful influences on teams, but we have decided to carve out organisational assessment as a separate area of research. This decision should result in a more parsimonious, reliable, sensitive, easier to define, and easier to observe set of markers that describe team behaviours.

The LOSA behavioural markers and the new markers we define here are tools—not theories or conceptual models—to help measure teamwork. However, the 10 behavioural markers we identified are consistent with conceptual models of teamwork.[19] After a review of the teamwork literature (mostly non-medical), Dickinson and McIntyre identified seven components of teamwork, six which are represented by our behavioural markers. The six (with our associated markers) are: team leadership (leadership); communication (information sharing, inquiry, assertion, intentions stated); monitoring (evaluation of plans, vigilance); feedback (information sharing, inquiry, assertion—if they are related to team members' performance); backup behaviour (workload distribution); and coordination (teamwork overall). The seventh component of teamwork identified by Dickinson and McIntyre is team orientation. This refers to attitudes that team members have towards one another. Our behavioural

Table 1 Comparison of healthcare behavioural markers with LOSA behavioural markers*

LOSA	UTBMNR	Halamek et al (neonatal resuscitation)[2]	Gaba et al (anaesthesia)[1]	Carthey et al (paediatric cardiac surgery)[3]††
Briefings	No comparable marker†	Know the environment	Orientation Feedback	Mental readiness (individual)
Plans stated	Intentions verbalised			
Workload assignment	No comparable marker‡	Distribute workload optimally	Workload distribution	
Contingency management	No comparable marker§	Anticipate and plan	Preparation / planning / anticipation	Anticipation
Monitor/crosscheck	No comparable marker¶			Coordination (team)
Workload management	Workload management			
Vigilance	Vigilance	Allocate attention wisely	Vigilance	Situational awareness (individual)
Automation management	No comparable marker**			
Evaluation of plans	Evaluation of plans		Re-evaluation	
Inquiry	Inquiry		Inquiry / advocacy / assertion	
Assertiveness	Assertion		Inquiry / advocacy / assertion	
Communication environment	Information sharing		Communications	Communication style (individual) Communication (team)
Leadership	Leadership	Assume the leadership role	Leadership/followership Overall hot-seat person effectiveness	Team leadership (team)
	Teamwork overall	Communicate effectively Use all available information	Group climate Overall anaesthesia crew effectiveness	Cognitive flexibility (individual)
		Use all available resources Call for help early enough Maintain professional behaviour		Team adaptation (individual) Safety awareness (individual)
				Experience (team) Redundancy (team) Adaptation (team) Policy (organisational) Planning and scheduling (organisational) Learning mechanisms (organisational)

*All groups of markers were originally based upon LOSA. Therefore the LOSA markers are listed first in the table. The table is a simplification of LOSA—the LOSA markers were developed for different phases of flight (predeparture/taxi, takeoff/climb, descent/approach/land, and global). For the UTBMNR, we are able to indicate why a marker is or is not comparable to LOSA, as indicated by the remaining footnotes.
†Not included because not observed on the video recordings.
‡Collapsed with workload management.
§Not included because not observed on the video recordings.
¶Included as part of workload management.
**Does not appear to be an important part of neonatal resuscitation.
††The markers from Carthey et al are grouped into team, individual, and organisational markers, as indicated in parentheses after each marker.
LOSA, Line Operations Safety Audit; UTBMNR, University of Texas behavioural markers for neonatal resuscitation.

markers do not account for this directly because we sought to measure behaviours and not attitudes. However, we did use attitudinal data when identifying and translating the markers.

Our process for identifying the behavioural markers is also similar to that proposed by Dickinson and McIntyre. We reviewed a series of videotapes (similar to their critical incident workshops) in which we clarified behavioural statements and matched them to team components. Our current data collection form is similar to their behavioural observation scales.

In healthcare there are several studies comparable to ours. As noted above, Carthey and colleagues[3] have identified a much broader list of markers which includes organisational markers. Our work is similar to theirs in that both studies attempted to translate and apply, rather than just borrow from, markers from aviation. Simon and colleagues found that aviation crew resource management behaviours could have prevented some malpractice claims so they applied them to emergency medicine teams.[20] This method of identifying behaviours is very different from our process and makes comparisons difficult. Furthermore, the large number of team behaviours (48) used in their intervention study[4] meant that they could not be compared directly with ours. Of course, there are also substantial differences in the scope and location (all processes in the emergency department) of their research compared with ours.

Gaba and colleagues used the LOSA markers when developing their anaesthesia crisis resource management curriculum,[21] and they have studied observer reliability during simulated crises.[1] A related group has developed a similar curriculum for training providers in neonatal resuscitation.[2] There is a great deal of similarity among all sets of markers (table 1).

An important goal of future research is to determine the relations among these behavioural markers and independently measured errors. Do teams exhibiting these markers commit fewer errors or manage errors better than teams that do not exhibit these markers? There is some preliminary evidence that this is the case.[3] Other research questions relate to the reliability of observers, the effects of team composition on behaviour (including sex, provider types, experience working together, and experience of the individuals), and determining which markers are related to which errors. If relations like these can be elucidated then these behavioural markers may become true error management behaviours. Research can then address how to teach healthcare teams to use these behaviours appropriately to improve quality. Interventions and curriculum reform to improve teamwork will require substantial resources and should therefore be supported by valid evidence, perhaps even from clinical trials. Such efforts should proceed with sensitivity to organisational and professional cultures and the demanding environments in which providers work.

Behavioral Markers of Neonatal Resuscitations

| Date: _____ Time: _____ | Observer ID: _____ | Length of observation: _____ Minutes |
| (Month/Day/Year)   (24 hour) | Case#: _____ | Observation Start Time: _____ |

<table>
<tr><td rowspan="2">Observe</td><td>0 = None
The behavior was not observed</td><td>1 = Poor observability
This may be due to consistently poor audio quality or inability to determine if verbalizations were related to the behavior</td><td colspan="2">2 = Minimum/inconsistent
There may be inconsistent audio quality or inconsistent ability to determine if verbalizations were related to the behavior</td><td>3 = Standard observability
Almost all verbalizations can be understood and the observation could be used as an example of the behavior being rated</td><td>4 = Outstanding observability
The behavior is usually explicit and clear. Would make excellent example of the behavior</td></tr>
</table>

| Freq | **1 = Rare**
One or two *rare examples* of the behavior occurred | **2 = Isolated/minimum**
There were *isolated examples* of the behavior throughout the observation | **3 = Intermittent/standard**
There were *intermittent examples* throughout the observation | **4 = Consistent**
There were *frequent/explicit examples* of the behavior throughout the observation |

Behavioral Markers:	Observability 0–4	Frequency 1–4	Comments (provide examples, anchors, and definitions)
1. Information sharing (vebalization of heart rate, vocal cord visulization, color)			
2. Inquiry: Asks questions; Asks for input & other relevant info			
3. Assertion (views are advocated during critical times)			
4. Intensions verbalized (state intentions before taking action)			
5. Teaching appropriate to the situation			
6. Evaluation of plans			
7. Workload management:			
8. Vigilance/Environmental Awareness			
Overall Ratings			
9. Teamwork overall			
10. Leardership			
Individual Ratings			
Was there an individual who differed significantly from the rest of the team? What was the behavioral marker # and what would that rating be for the individual? #_____			

Threats	Observability	Managed? (Yes/No)	Explain (was threat resolved?)	Threat severity (1 = Low to 4 = High)
Staffing Level				
Environmental Distractions (twins, background conversations, phones, alarms, etc.)				
Other (please describe/indicate level of complexity)				
Any Other Comments (e.g. complications, other behavioral markers, etc.)?				
Candidate for additional review? Yes/No Explain:				

Figure 1   The University of Texas Behavioral Marker Audit Form.

....
Authors' affiliations
E J Thomas, J B Sexton, R L Helmreich, The University of Texas Center of Excellence for Patient Safety Research and Practice, Houston, Texas, USA
E J Thomas, The University of Texas Houston Medical School, Department of Medicine, Division of General Medicine

Key messages box

- Researchers and team training programmes in commercial aviation identify specific teamwork behaviours, or behavioural markers, that can be used to measure and improve teamwork among cockpit crews.
- Similarly, we used information from surveys, focus groups, and video recordings of neonatal healthcare providers to define a set of 10 behavioural markers for measuring teamwork during neonatal resuscitation.
- We then developed the behavioural marker audit form to assess the following behavioural markers for neonatal resuscitation: information sharing, inquiry, assertion, intentions verbalised, teaching, evaluation of plans, workload management, vigilance, leadership, and teamwork overall.

J B Sexton, R L Helmreich, The University of Texas at Austin, Department of Psychology, Human Factors Research Project, Austin, Texas, USA

APPENDIX

The University of Texas Behavioral Marker Audit Form
The form is reproduced in fig 1.

REFERENCES

1 **Gaba DM**, Howard SK, Flanagan B, *et al.* Assessment of clinical performance during simulated crises using both technical and behavioral ratings. *Anesthesiology* 1998;**89**:3–18.
2 **Halamek LP**, Kaegi DM, Gaba DM, *et al.* Time for a new paradigm in pediatric medical education: teaching neonatal resuscitation in a simulated delivery room environment. *Pediatrics* 2000;**106**:E45.
3 **Carthey J**, de Leval MR, Wright DJ, *et al.* Behavioural markers of surgical excellence. *Safety Science* 2003;**41**:409–25.
4 **Morey JC**, Simon R, Jay GD, *et al.* Error reduction and performance improvement in the emergency department through formal teamwork training: evaluation results of the MedTeams project. *Health Serv Res* 2002;**37**:1553–81.
5 **Brannick MT**, Salas E, Prince C, eds. *Team performance assessment and measurement.* Mahwah New Jersey: Lawrence Erlbaum, 1997.
6 **Shojania KG**, Duncan BW, MacDonald KM, *et al. Making healthcare safer: a critical analysis of patient safety practices,* 2001.Available at http://www.ahcpr.gov/clinic/ptsafety/spotlight.htm.
7 **Odegard S**. Safety management in civil aviation: a useful method for improved safety in medical care? *Safety Science Monitor* 2000;**4**:1–12.
8 **Helmreich RL**, Merritt AC. *Culture at work: national, organizational, and professional influences.* Aldershot: Ashgate, 1998.
9 **Helmreich RL**, Klinect JR, Wilhelm JA. Models of threat, error, and CRM in flight operations. In: *Proceedings of the Tenth International Symposium on Aviation Psychology.* Columbus, Ohio: The Ohio State University, 1999:677–82.
10 **Helmreich RL**, Foushee HC. Why crew resource management: empirical and theoretical bases of human factors training in aviation. In: Wiener EL, Kanki BG, Helmreich RL, eds. *Cockpit resource management.* San Diego: Academic Press, 1993.
11 **Helmreich RL**, Wilhelm JA, Gregorich SE, *et al.* Preliminary results from evaluation of cockpit resource management training: Performance ratings of flight crews. *Aviat Space Environ Med* 1990;**61**:576–9.
12 **Kohn LT**, Corrigan JM, Donaldson MS, eds. To err is human. In: *Building a safer health system.* Washington DC: National Academy Press, 1999.
13 **Thomas EJ**, Helmreich RH. Will airline safety models work in medicine? In: Rosenthal M, Sutcliffe KM, eds. *Medical error.* San Francisco: Josey Bass, 2002.
14 **Klampfer B**, Flin R, Helmreich RJ, *et al.* Enhancing performance in high risk environments: recommendations for the use of behavioral markers. Zurich: Proceedings of the Behavioral Markers Workshop, July 5–6, 2001.
15 **Thomas EJ**, Sherwood GD, Mulhollem JL, *et al.* Working together in the neonatal intensive care unit: provider perspectives. *J Perinatol*, (in press).
16 **Morse J**. Emerging from the data: the cognitive process of analysis in qualitative inquiry. In: Morse J, eds. *Critical issues in qualitative research methods.* Newbury Park, CA: Sage, 1994:23–43.
17 **Sexton JB**, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: data from the frontlines. *BMJ* 2000;**320**:745–9.
18 **Thomas EJ**, Sexton BJ, Helmreich RL. Discrepant attitudes about teamwork among critical care nurses and physicians. *Crit Care Med* 2003;**31**:956–9.
19 **Dickonson TL**, McIntyre RM. A conceptual framework for teamwork measurement. In: Brannick MT, Salas E, Prince C, eds. *Team performance assessment and measurement.* Mahwah, New Jersey: Lawrence Erlbaum, 1997:19–44.
20 **Simon R**, Morey JC, Locke A. Full scale development of the emergency team coordination course and evaluation measures. Andover MA: Dynamics Research Corporation, 1997.
21 **Gaba DM**, Fish KJ, Howard SK. *Crisis management in anesthesiology.* New York: Churchill-Livingstone, 1994.