

ORIGINAL ARTICLE

Teamwork and quality during neonatal care in the delivery room

EJ Thomas¹, JB Sexton^{1,2}, RE Lasky^{1,3}, RL Helmreich^{1,2}, DS Crandell³ and J Tyson³

¹Division of General Medicine, Department of Medicine, University of Texas Houston Medical School, Houston, TX, USA; ²Department of Psychology, The University of Texas at Austin, Human Factors Research Project, Houston, TX, USA and ³Department of Pediatrics, The University of Texas at Austin, Houston, TX, USA

Objective: Experts believe good teamwork among health care providers may improve quality. We sought to measure the frequency of team behaviors during delivery room care and to explore how these behaviors relate to the quality of care.

Study design: We video recorded neonatal resuscitation teams then used independent observers to measure teamwork behaviors and compliance with Neonatal Resuscitation Program (NRP) guidelines (a measure of quality of care).

Results: Observer agreement was either fair or good for all teamwork behaviors except workload management, vigilance, and leadership, for which agreement was slight. All teams ($n = 132$) exhibited the behaviors information sharing and inquiry, and all but one team exhibited vigilance and workload management. Other behaviors were present less often: assertion in 19.9% of teams, teaching 16.7%, leadership 19.7%, evaluation of plans 12.9%, and intentions stated 9.1%. Factor analysis identified three fundamental components of teamwork: communication (comprised of information sharing and inquiry); management (workload management and vigilance); and leadership (assertion, intentions shared, evaluation of plans, and leadership). All three components were weakly but significantly correlated with independent assessments of NRP compliance and an overall rating of the quality of care.

Conclusion: Most team behaviors can be reliably observed during delivery room care by neonatal resuscitation teams, and some are infrequently used. We found weak but significant and consistent correlations among these behaviors with independent assessments of NRP compliance and an overall rating of the quality of care. These findings support additional efforts to study team training for delivery room care and other areas of healthcare.

Journal of Perinatology (2006) 26, 163–169. doi:10.1038/sj.jp.7211451; published online 26 January 2006

Keywords: patient safety; neonatal resuscitation; neonatal resuscitation program; communication

Introduction

Approximately 10% of newborns require assistance to begin breathing and about 1% are critically ill and may need lifesaving therapies such as ventilation, chest compression, and medications to support heart rate.¹ Other aspects of delivery room care (warming, stimulation) are also important.² The Neonatal Resuscitation Program (NRP) is the curriculum used to teach providers how to care for newborns in the delivery room. The lifesaving potential of NRP is not as well established as Adult Cardiac Life Support, but worldwide, the NRP could possibly improve outcomes of over 1 million newborns per year.¹

Studies have identified shortcomings in providers' adherence to the NRP guidelines.³ It is widely accepted that professionals need to work as teams to accomplish their duties safely and efficiently. In healthcare, improving teamwork may improve quality by helping to prevent and manage errors,^{4,5} and team training has been recommended by the Institute of Medicine.⁶ However, there is very little research about how to measure teamwork among healthcare providers.⁷ For example, it is not known if team behaviors can be reliably measured, or how they relate to the process of care during resuscitations. Without such research there is little empirical support for developing programs to improve teamwork.

To address these knowledge gaps, we video recorded and assessed resuscitations of infants born by cesarean section under routine clinical conditions. We used the NRP as the standard of care and we report the frequencies of team behaviors, the reliability of the measurements, the associations of individual behaviors with each other, and the relationships among team behaviors and noncompliance with NRP. Finally, we discuss implications of these findings for neonatal resuscitation.

Methods

Overview

We, like the NRP, defined 'resuscitation' to include the full range of delivery room care, from initial assessment, warming, drying, and stimulation, to ventilation, chest compressions, and medication administration. Our general approach was to independently measure teamwork and the process of care and then

evaluate correlations among the two. We used one pair of trained observers to measure teamwork behaviors, and a different pair of observers to measure compliance with NRP guidelines. Previously, we adapted a teamwork model from commercial aviation,^{8,9} identified team behaviors (or behavioral markers), and developed a Behavioral Marker Audit Form for neonatal resuscitation.⁸ The form lists and defines 10 teamwork behaviors (Box 1) and allows observers to rate the frequency of these behaviors.

Similar to previous researchers,³ we used the NRP as the standard for measuring the process of care during neonatal resuscitation. We then trained NRP instructors to view the video recordings to look for instances of noncompliance with the NRP guidelines.

Setting

We video recorded resuscitations of infants born by cesarean section at a 500 bed urban teaching hospital with an 80 bed neonatal intensive care unit and approximately 3400 deliveries during the study year. The care occurred in the resuscitation room that adjoined two labor and delivery operating rooms.

Subjects

Our subjects were the resuscitation teams. Every Cesarean Section was attended by a resuscitation team that include at least two providers, a physician (usually a pediatric resident) and one of a specially trained group of neonatal nurses who routinely attended all cesarean section deliveries and high-risk vaginal deliveries. Respiratory therapists as well as fellows or faculty were also team members when high-risk deliveries occurred. None of the providers had received team training. The study was approved by the institution's IRB and resuscitation team members gave consent to be video recorded. One hundred and twenty subjects were approached and two residents and one faculty member declined to participate. Parents were not asked for consent.

Data collection

Taping occurred 24 h a day, Monday–Friday between 1 January and 31 December 2002. There were two radiant warmers in the resuscitation room and two cameras were focused on each warmer. One camera was in the corner of the ceiling to capture all the providers caring for the infant; the other camera was attached to the warmer and it provided a close up view of the infant. Nurses started and stopped the cameras to videotape the care provided to each infant, but they frequently forgot to initiate recording (characteristics of recorded and nonrecorded infants are presented in the results and Table 1). For twin births, both were recorded. Each day a research assistant removed the tapes, and downloaded the file onto a secure computer.

Measurement of quality: NRP compliance

Like previous investigators,³ compliance with NRP guidelines was used as our measure of the quality of the resuscitation and it was assessed independently of teamwork (see below). We adapted a data collection form³ to be consistent with the most recent NRP guidelines.¹ To help assure that these evaluations were unbiased and confidential, we trained two nurses who were NRP instructors from another hospital to perform the evaluations and complete the form. They were not aware of the teamwork behaviors that were being measured by other reviewers, and they were not asked to make any assessments of teamwork. Each one independently reviewed all video recordings and used the form to record their observations. The form allowed the observer to indicate if the step was applicable to the infant and whether or not it was completed correctly. After the independent viewing, the observers jointly reviewed all the recordings and completed a form that represented their consensus about NRP compliance. These consensus judgments were used in all analyses.

Box 1 Definitions of team behaviors for neonatal resuscitation

1. *Information sharing.* Providers verbally share information that relates to the assessment and treatment of the baby.
2. *Inquiry.* Providers question each other about the infant's clinical status, their assessment, and treatment plans.
3. *Assertion.* An individual provider asserts his or her 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, color, and respirations (these are 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 Program guidelines are not examples of intentions shared.
5. *Teaching.* Teaching is observed during the observation. This may be short and informal information exchanges. It may occur between any of the providers (e.g., 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 prioritized.
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 behaviors 1–8 above. It may also include other observations such as dynamic 'give and take' among team members and nonverbal communication that are not explicitly defined in the behavioral markers.
10. *Leadership.* Leadership activities may include sharing of a mental model, assigning tasks, sharing of information and opinion. This may be rated for any provider at the resuscitation.

Table 1 Characteristics of the infants cared for by the resuscitation teams

	Video recorded (n = 132)	Not video recorded (n = 980)
Median gestational age	38 weeks (range 24–42)	38 weeks (range 20–44)
Median birth weight	2967 g (range 657–4445)	3110 g (range 230–4815)
<i>Race</i>		
Black	46%	44%
White	31%	28%
Hispanic	16%	20%
Asian	7%	8%
<i>Gender</i>		
Male	47%	51%
Female	53%	49%
<i>Median Apgar scores</i>		
1 min	8 (range 0–10)	8 (range 0–10)
5 min	9 (range 4–9)	9 (range 0–10)
1 min Apgar less than 4	7%	7%
5 min Apgar less than 7	5%	6%

Measurement of team behaviors

Two investigators (EJT, a general internist and JBS, a social psychologist) used a data form⁸ to rate the frequency of team behaviors. Although they had some knowledge of NRP, they did not make any assessments about the quality of the resuscitation process. EJT and JBS rated and discussed the first four records to facilitate consistent scoring. The next 54 records were then independently rated by EJT and JBS. After rating the frequency of team behaviors, EJT and JBS then discussed differences and agreed upon a final, consensus rating. The consensus ratings for the first 58 resuscitations were used in all subsequent data analysis. The remaining video recordings were randomly distributed and rated by only one investigator (37 for EJT and 37 for JBS).

Data analysis

NRP compliance. We calculated percent noncompliance with NRP as the number of items not performed or performed incorrectly, divided by the total number of applicable items multiplied by 100. Summary variables defined by the noncompliance percentages for each NRP stage were also calculated: preparation and initial steps (15 items), oxygen administration (three items), bag/mask ventilation (seven items), chest compression (five items), intubation (14 items), and drugs administered (only one drug, epinephrine was administered (four items)). In addition, the nurses rating NRP compliance gave an overall rating from 1 (poor) to 5 (excellent) reflecting their overall impression of the quality of the resuscitation. We do not

report the agreement between the two nurse raters' initial independent ratings because we relied upon their consensus ratings for all analyses.

Teamwork behaviors. We estimated the frequency of each behavior during the resuscitation by using a scale that ranged from 0 (none – the behavior was not observed during the resuscitation) to 4 (consistent – there were frequent and consistent examples of the behavior observed throughout the resuscitation). The five point range of this scale was based upon our many years of observing airline cockpit crews. It allowed both differentiation of crew performance and good reliability among observers. We report the percentage of resuscitations for which each behavior was observed at each frequency level. We used the weighted kappa statistic to measure the inter-rater reliability of the teamwork behavior measurements. The kappa statistic indicates the agreement between the raters beyond the agreement expected by chance alone. A weighted kappa was calculated for each behavior because it considers near misses in scoring on our five point scale rather than only counting perfect agreement.

The behavioral marker variables are ordinally scaled. Nevertheless, we used exploratory factor analysis to understand how the behavioral markers relate to each other and how they might reflect underlying teamwork constructs. We did not include the teamwork overall behavioral marker because it was an overall summary measure. We used the principal axis method with varimax rotation.

Correlations between teamwork summary variables and NRP noncompliance. The relationships between the teamwork summary variables and the summary variables characterizing deviations from the NRP protocol were analyzed by Spearman correlation coefficients because the teamwork variables were ordinally scaled and because of concern about the assumptions required by Pearson correlations. The teamwork behaviors were summarized by the three summary variables described below, communication, management, and leadership. The quality of the resuscitation was summarized by the percent noncompliance for all the NRP items administered (total) and by the overall rating of the quality of the resuscitation. Noncompliance percentage scores for preparation and initial steps and oxygen administration were also analyzed. Owing to the small number of infants administered bag/mask ventilation, chest compression, intubation, and drugs, summary variables reflecting noncompliance with those NRP stages are not reported. An $\alpha = 0.05$ was adopted. No adjustments for multiple comparisons were made because of the exploratory nature of the study.¹⁰

Results

The teams ($n = 132$ (118 unique teams)) cared for infants who tended to be healthy, but there was a broad range of gestational age, weight, and Apgar scores. The infants of video recorded teams were similar to those of nonvideo recorded teams. (Table 1) All teams included one or more residents and all but one also included a nurse. The one exception was a team comprised of a fellow and two residents. Eighteen (13.6%) of the teams had a fellow present and five (3.8%) included an attending (one team had both an attending and a fellow).

NRP noncompliance

All 132 infants received the NRP preparation and initial steps; 117 had oxygen administered; 20 had bag/mask ventilation; 11 were intubated; two received chest compressions; and one received epinephrine. The mean NRP noncompliance rate ranged from 15.9% for the 15 items in preparation and initial steps to 54.5% for infants requiring multiple attempts before being successfully intubated (14 items per intubation) (Table 2). Two of the six infants requiring multiple intubation attempts were successfully intubated on the second attempt, one infant required three attempts, two required four attempts, and one infant was successfully intubated after seven attempts.

Team behaviors

Agreement was either 'fair' ($\kappa = 0.41$ – 0.60) or 'good' ($\kappa = 0.61$ – 0.80) for all individual teamwork behaviors except workload management, vigilance, and leadership. Agreement was 'slight' (0.21 – 0.40) for those three behaviors¹¹ (Table 3). Information sharing and inquiry were always observed and

Table 2 Noncompliance with the neonatal resuscitation program

Neonatal resuscitation program section	Mean %	Range
Preparation and Initial steps (15 items)	16	0–50
Oxygen administered (3 items)	30	0–67
Bag/Mask ventilation (7 items)	26	0–100
First Intubation (14 items)	24	0–44
Second intubation ^a	55	8–40

^aSix infants had a second intubation attempt.

Table 3 Reliability of teamwork behavior measurements

Teamwork behavior	Percent agreement	Weighted kappa	Standard error of kappa
Information sharing	66.7	0.653	0.038
Inquiry	63.0	0.685	0.043
Assertion	83.3	0.582	0.016
Plans stated	92.6	0.509	0.004
Teaching	90.7	0.798	0.009
Evaluation of plans	96.3	0.711	0.004
Workload management	88.9	0.232	0.006
Vigilance	64.8	0.315	0.029
Teamwork overall	70.4	0.492	0.032
Leadership	79.6	0.373	0.015

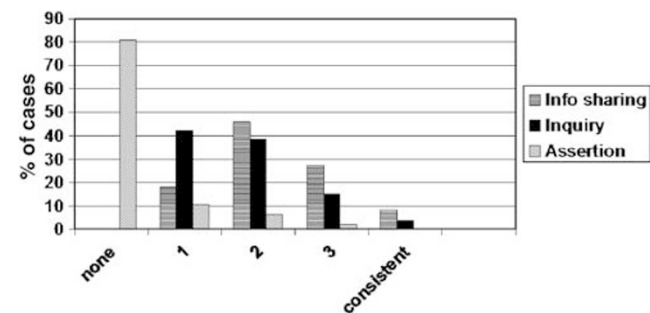


Figure 1 Frequency of behavioural markers.

vigilance and workload management were observed for all but one resuscitation. The other behaviors were present less often: assertion was observed in 19.9% of teams, teaching 16.7%, leadership 19.7%, evaluation of plans 12.9%, and intentions stated 9.1% (Figures 1–3).

We identified three factors that explained virtually all (99.96%) of the variation in the team behaviors. (Table 4) Each described a fundamental component of teamwork: *communication*, *management*, and *leadership*. Communication was defined by two markers: information sharing and inquiry; management was defined by two markers: workload management and vigilance; and leadership was defined by four markers: assertion, intentions

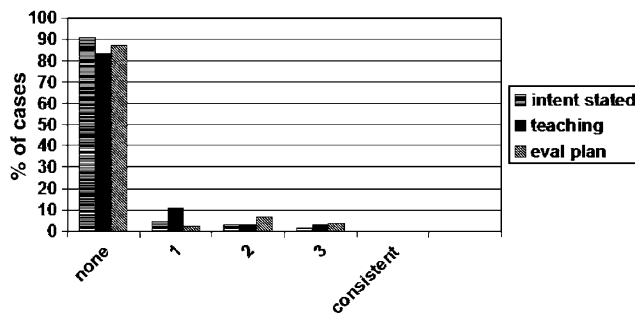


Figure 2 Frequency of behavioural markers.

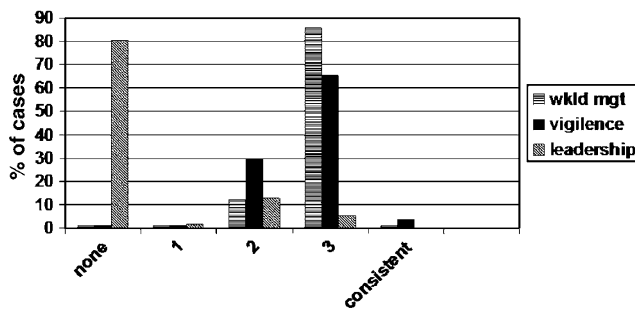


Figure 3 Frequency of behavioural markers.

Table 4 Factor loading coefficients after varimax rotation for each of the three factors

Variables	Factor 1 (communication)	Factor 2 (management)	Factor 3 (leadership)
Information sharing	-0.766	-0.296	-0.202
Inquiry	-0.807	-0.015	-0.329
Assertion	-0.375	-0.066	-0.581
Plans stated	-0.133	-0.141	-0.446
Teaching	-0.242	-0.135	-0.119
Evaluation of plans	-0.351	0.005	-0.611
Workload management	-0.066	-0.403	-0.025
Vigilance	-0.154	-0.606	-0.250
Leadership	-0.086	-0.209	-0.505

Loadings > absolute 0.400 are bolded indicating the variables contributing most to each factor.

shared, evaluation of plans, and leadership. The scores on the markers contributing most (factor loading coefficients >0.400) to each factor were summed to generate three ordinally scaled summary variables labeled communication, management, and leadership. Higher scores for each of these summary variables indicated more frequently observed teamwork behaviors.

Correlations between teamwork and NRP noncompliance

All three scales (communication, management, leadership) were weakly but significantly correlated with independent measures of

quality. Communication was correlated with the nurses' rating of overall quality of the resuscitation (Spearman $\rho = -0.236$, $P = 0.007$), the total noncompliance with all NRP steps administered (Spearman $\rho = -0.214$, $P = 0.014$), and noncompliance during preparation and initial steps (Spearman $\rho = -0.230$, $P = 0.008$). Management was significantly correlated with the total noncompliance with all NRP steps administered (Spearman $\rho = -0.201$, $P = 0.021$) and with noncompliance during the NRP stage preparation and initial steps (Spearman $\rho = -0.252$, $P = 0.003$). Leadership was significantly correlated with the nurses' rating of overall quality of the care (Spearman $\rho = -0.288$, $P < 0.001$). These analyses were repeated on the 118 unique teams and all the previously significant correlations remained significant and all the nonsignificant correlations remained nonsignificant with very little difference in the magnitude of the correlations.

Discussion

We found that team behaviors adapted from commercial aviation can be reliably observed during delivery room care by neonatal resuscitation teams in a teaching hospital, and the behaviors group into three fundamental aspects of teamwork: communication, management, and leadership. Furthermore, these behaviors were significantly correlated with independent assessments of the technical quality of the resuscitation.

Team behaviors and measures of technical quality were weakly but significantly correlated. Although these correlations do not confirm a causal relationship, they raise several important possibilities to be explored in future research. The behaviors may prevent errors and thereby improve quality; the behaviors may be used by teams in response to errors (error management behaviors); the behaviors may be correlated with some other causal factor which then affects quality; or the behaviors may cause errors. There is no conceptual or empirical evidence to support the last possibility. However, the suggestion that these behaviors may be used to prevent and manage errors is consistent with the aviation model of threat and error management developed by Helmreich *et al.*,¹² and breakdowns in teamwork and communication have been identified as causal factors in aviation accidents and emergencies.^{13,14}

Despite the substantial interest in teamwork in healthcare, we are the first to report that these team behaviors can be measured with moderate reliability during real (not simulated) patient care, and to find that the behaviors are correlated with independent assessments of the technical quality of care. Other groups have made substantial progress in developing sets of teamwork behaviors for anesthesia¹⁵ and neonatology¹⁶ and we have compared our teamwork behaviors with those developed by other groups.⁸ In addition, Carthey *et al.*¹⁷ have related the behaviors to the process of surgical care. Our work adds to theirs by using video recordings

to allow independent assessments of team behaviors and technical quality. This step is important because if an observer sees high technical quality they could be biased to rate teamwork as high and vice versa. Our independent assessments increase the likelihood that these team behaviors are in fact used by providers in a way that is tied to the actual process of technical care.

Interestingly, providers did not consistently use all the behaviors. Information sharing, inquiry, vigilance, and workload management were almost always observed, but assertion and leadership was observed in only about 20% of cases, evaluation of plans was observed in 13% of cases, and intentions stated were observed in 9% of cases. More behaviors were used in more complicated cases, such as those that required intubation. In other more routine cases, we hypothesize that the standardized nature of neonatal resuscitation may not require use of all these behaviors. In particular, leadership skills are less likely to be needed during standardized processes. Teaching was rarely observed. This is in part due to our definition that required 'exchanges' to occur among providers, such as the typical questions and answers of attending round, but in an abbreviated format. In our ongoing analysis of the resuscitations we will consider changing 'teaching' to 'advice'. This is found more frequently because it includes single short statements such as 'tilt the head some more' and would not require exchanges of information among providers.

The grouping of our behaviors into communication, management, and leadership is consistent with previous literature on teamwork in neonatal resuscitation¹⁸ and other areas¹⁹. The difference between management and leadership has been popularized by Peter Drucker and Warren Benis who note that management is 'doing things right' and leadership is 'doing the right things'. The behaviors we call workload management and vigilance should help providers to manage or 'do things right'. The behaviors assertion, intentions shared, and evaluation of plans can be used by team leaders to make sure the team is 'doing the right things'. For endeavors involving a team composed of participants with different levels of experience and differing expertise, enactment of all three components is essential when complications arise.

Our findings can be used to justify and design studies that incorporate these behaviors into neonatal resuscitation training and to evaluate their impact on the quality of care (as has been performed for leadership training in Advanced Life Support²⁰). For example, a study could take the standard NRP guidelines and add a team training component. Providers could be told that in addition to following the NRP treatment algorithm, it is important to communicate with each other, manage the task, and assure appropriate leadership. These teamwork domains can then be further defined by using our specific team behaviors: communication is comprised of information sharing and inquiry (asking questions); management is comprised of workload management and vigilance; and leadership is comprised of assertion, intentions shared, and evaluation of plans. These team

behaviors could be practiced in either low or high fidelity simulators. Subsequently, process of care and infant outcomes could then be assessed to measure the impact of team training.

Limitations

Most of our video recordings were of residents and nurses caring for relatively healthy infants in a teaching hospital. It is possible that our findings are not applicable to community-based settings or sicker infants. Many problems with neonatal resuscitation occur with very low birth weight infants and there were very few of these infants in our study. Generalizability may also be limited if other institutions' neonatal resuscitation teams are comprised of different provider types than in this study. We relied on nurses to start the cameras so even though the observed characteristics of the recorded infants are similar to those not recorded, nurses may have chosen to record only certain types of infants. Provider behavior may have been affected by the presence of cameras. All of these issues could limit the generalizability of our findings.

The data support, but do not prove the hypothesis that these team behaviors result in better quality of care. Our significant correlations between team behaviors and measures of quality can be interpreted in several ways, as discussed above. It will be important to test this hypothesis in future studies. The moderate reliability of the observations of team behaviors makes it more difficult to detect correlations between the behaviors and measures of quality of care. Future research should attempt to improve the reliability, perhaps by clarifying definitions, reducing the number of behaviors that observers are asked assess, or training observers differently.

We present a limited view of teamwork. For example, we did not measure team orientation (attitudes that team members have towards one another)¹⁹; shared mental models²¹; motivational and coordination processes²²; nonverbal communication; or team activities that may have occurred before arrival in the delivery room. Although important, these aspects of teamwork are not readily measured or taught to individuals, as are the team behaviors we studied.

Conclusions

Our data describe selected dimensions of teamwork during delivery room care and neonatal resuscitation and compliance with the NRP. Some discreet team behaviors can be measured and there are correlations among these behaviors and the process of care. Thus, our findings should encourage research on incorporating team training into neonatal resuscitation, but some team behaviors are difficult to reliably measure and cause-effect relationships among team behaviors and the process of care are unclear. Experience in aviation suggests that team training programs should be based on careful analysis of the culture and practices in the local environment. One cannot assume that approaches can be applied without change across professional domains or institutions.

Acknowledgments

Funding provided by Agency for Healthcare Research and Quality (#U18HS1116401).

References

- 1 Kattwinkel J (ed). *Textbook of Neonatal Resuscitation*. American Academy of Pediatrics and American Heart Association, 2000.
- 2 Tyson JE. Immediate care of the newborn infant In: Sinclair JC, Bracken MB (eds), *Effective Care of Newborn Infant*. Oxford University Press: New York, 1992 pp 21–39.
- 3 Carbine DN, Finer NN, Knodel E, Rich W. Video recording as a means of evaluating neonatal resuscitation performance. *Pediatrics* 2000; **106**: 654–658.
- 4 Helmreich RL, Schaefer HG. Team performance in the operating room. In: Bogner MS (ed), *Human Error in Medicine*. Lawrence Erlbaum and Associates: Hillsdale, NJ, 1994 pp 225–253.
- 5 Morey JC, Simon R, Jay GD, Wears RL, Salisbury M, Dukes KA *et al*. Error reduction and performance improvement in the emergency department through formal teamwork training: evaluation results of the MedTeams project. *Health Services Research* 2002; **37**: 1553–1581.
- 6 Kohn LT, Corrigan JM, Donaldson MS eds. *To Err is Human. Building a Safer Health System*. Washington, DC: National Academy Press, 1999.
- 7 Shojania KG, Duncan BW, MacDonald KM (eds) *et al*. Making healthcare safer: a critical analysis of patient safety practices, 2001. Available at <http://www.ahrp.gov/clinic/ptsafety/spotlight.htm>.
- 8 Thomas EJ, Sexton JB, Helmreich RL. Translating teamwork behaviors from aviation to healthcare: development of behavioral markers for neonatal resuscitation. *Quality Safety Healthcare* 2004; **13**(Suppl 1): i57–64.
- 9 Thomas EJ, Sherwood GD, Mulhollem JL, Sexton JB, Helmreich RL. Working together in the neonatal intensive care unit: provider perspectives. *J Perinatol* 2004; **24**: 552–559.
- 10 Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology*. 1990; **1**(1): 43–46.
- 11 Byrt T. How good is that agreement? *Epidemiology* 1996; **7**: 561.
- 12 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*. Academic Press: San Diego, CA, 1993, pp 3–41.
- 13 Helmreich RL, Merritt AC. *Culture at Work in Aviation and Medicine: National, Organizational, and Professional Influences*. Ashgate: Aldershot, UK, 1998.
- 14 Helmreich RL. Managing human error in aviation. *Scientific American* 1997; 62–67.
- 15 Gaba DM, Howard SK, Flanagan B, Smith BE, Fish KJ, Botney R. Assessment of clinical performance during simulated crises using both technical and behavioral ratings. *Anesthesiology* 1998; **89**: 3–18.
- 16 Halamek LP, Kaegi DM, Gaba DM, Sowb YA, Smith BC, Smith BE *et al*. Time for a new paradigm in pediatric medical education: teaching neonatal resuscitation in a simulated delivery room environment. *Pediatrics* 2000; **106**: E45.
- 17 Carthey J, de Leval MR, Wright DJ, Farewell VT, Reason JT. Behavioral markers of surgical excellence. *Safety Science* 2003; **41**: 409–425.
- 18 Finer NN, Rich W. Neonatal resuscitation: toward improved performance. *Resuscitation* 2002; **53**: 47–51.
- 19 Dickinson TL, McIntyre RM. A conceptual framework for teamwork measurement. In: Brannick MT, Salas E, Prince C (eds), *Team Performance Assessment and Measurement*. Lawrence Erlbaum: Mahwah New Jersey, 1997; 19–44.
- 20 Cooper S. Developing leaders for advanced life support: evaluation of a training programme. *Resuscitation* 2001; **49**: 33–38.
- 21 Cannon-Bowers JA, Salas E, Converse S. Shared mental models in expert team decision making. In: Castellan NJ (ed), *Current Issues in Individual and Group Decision Making*. Lawrence Erlbaum: Hillsdale NJ, 1993, pp 221–246.
- 22 Zaccaro SJ, Rittman AL, Marks MA. Team leadership. *Leadership Quarterly* 2001; **12**: 451–483.