

Assessing safety culture in NICU: psychometric properties of the Italian version of Safety Attitude Questionnaire and result implications

Alessandra Zenere MD,¹ M. Elisabetta Zanolin PhD,⁴ Roberta Negri ND,⁵ Francesca Moretti MD PhD,² Mario Grassi PhD⁶ and Stefano Tardivo MD³

¹Resident Doctor, ²Resident Doctor, ³Researcher, Unit of Hygiene and Preventive, Environmental and Occupational Medicine, Department of Diagnostics and Public Health, University of Verona, Verona, Italy

⁴Associate Professor, Unit of Epidemiology & Medical Statistics, Department of Diagnostics and Public Health, University of Verona, Verona, Italy

⁵Nurse, Neonatology and Neonatal Intensive Care Unit, Mantua Hospital 'C. Poma', Mantua, Italy

⁶Full Professor, Unit of Medical Statistics & Epidemiology, Department of Health Science, University of Pavia, Pavia, Italy

Keywords

neonatal intensive care unit, patient safety, questionnaire validation, Safety Attitude Questionnaire

Correspondence

Professor M. Elisabetta Zanolin
Unit of Epidemiology & Medical Statistics
Department of Public Health and
Community Medicine
University of Verona
Strada le Grazie 8
Verona 37135
Italy
E-mail: elisabetta.zanolin@univr.it

Accepted for publication: 23 September 2015

doi:10.1111/jep.12472

Abstract

Rationale, aims and objective Neonatal intensive care units (NICUs) are a high-risk setting. The Safety Attitude Questionnaire (SAQ) is a widely used tool to measure safety culture. The aims of the study are to verify the psychometric properties of the Italian version of SAQ, to evaluate safety culture in the NICUs and to identify improvement interventions.

Method A cross-sectional study was conducted in 6 level III NICUs. The SAQ was translated into Italian and adapted to the context, a confirmatory factor analysis (CFA) was performed to validate the questionnaire.

Results 193 questionnaires were collected. The mean response rate was 59.7% (range 44.5%–95.7%). The answers were analysed according to six factors: *f1 – teamwork climate*, *f2 – safety climate*, *f3 – job satisfaction*, *f4 – stress recognition*, *f5 – perception of management*, *f6 – working conditions*. The CFA indexes were adequate (McDonald's omega indexes varied from 0.74 to 0.94, the SRMR index was equal to 0.79 and the RMSEA index was 0.070, 95% CI = 0.063–0.078). The mean composite score was 57.6 (SD 17.9), ranging between 42.3 and 69.7 on a standardized 100-point scale. We highlighted significant differences among units and professions ($P < 0.05$).

Conclusions The Italian version of the SAQ proved to be an effective tool to evaluate and compare the safety culture in the NICUs. The obtained scores significantly varied both within and among the NICUs. The organizational and structural characteristics of the involved hospitals probably affect the safety culture perception by the staff.

Introduction

Safety culture is defined as 'the product of individual and group values, attitudes, perceptions, competencies, and patterns of behavior that determine the commitment to, and the style and proficiency of, an organization's health and safety management. Organizations with a positive safety culture are characterized by communications founded on mutual trust, by shared perceptions of the importance of safety, and by confidence in the efficacy of preventive measures' [1].

The spread and the progress of safety culture among organizations can significantly improve patient care and outcomes. Its importance is undeniable and widely recognized in international literature, as it is proved by the inclusion of patient safety in the main accreditation standards [2–5].

Neonatal intensive care units (NICUs) are a high-risk setting [6,7]. Previous studies reported that patients in NICUs experience a higher rate of medication error and potential adverse drug events (ADEs) than neonates in other wards and children of other age groups [8]. Furthermore, physiological immaturity and limited compensatory abilities make newborns more vulnerable to severe consequences of medical errors. Moreover, evidence shows that more than one-half of all adverse events are preventable, underlining the importance to widen the culture of safety in NICUs to minimize patients' harm [9].

Health organizations need specific tools to measure safety culture in order to identify weak points and plan tailored interventions. The Safety Attitudes Questionnaire (SAQ) is a widely used tool to assess patient safety. It was developed from the Flight

Management Attitude Questionnaire (FMAQ) by Sexton and colleagues at the University of Texas, Houston [4,10,11]. The SAQ has good psychometric properties and its results are associated with clinical outcomes. It also allows internal and external benchmarks [10,12,13].

It has been translated into different languages, such as Norwegian, Chinese, Turkish, Dutch and Swedish, maintaining acceptable psychometric properties [14–18]. Moreover, it has been adapted to be used in many settings, including intensive care units, operating rooms, emergency rooms, outpatient settings and, recently, in NICU settings [12,19–25]. However, no data are available on patient safety culture in the Italian NICUs nowadays.

Thus, the purposes of the present study are twofold: (1) to confirm the psychometric properties of the Italian translation of the SAQ (Short form); and (2) to evaluate the safety culture in six Italian NICUs.

Materials and methods

Study design and sample

This cross-sectional study has been performed by administering the SAQ (Short Form Version) in the NICUs of six public hospitals in north-eastern Italy. The questionnaire was administered to medical staff, including resident doctors, and nursing staff with a working experience of at least one month in the participating NICU. The characteristics of the six participating NICUs are shown in Table 1.

Questionnaire

The SAQ (Short Form Version) consists of 32 items that examine a caregiver's attitude towards six dimensions (latent factors, *f*) of safety culture: *f1* – *teamwork climate*, *f2* – *safety climate*, *f3* – *job satisfaction*, *f4* – *stress recognition*, *f5* – *perception of management*, *f6* – *working conditions*.

All responses were recorded using a 5-point Likert scale (from 1 = strongly disagree to 5 = strongly agree). Negatively worded items were reverse scored before analysis so that their valence matched the positively worded items [10]. Higher scores indicate a good safety culture in the evaluated dimensions.

The questionnaire was translated into Italian by two independent Italian translators with working experience in the medical field. Some words were slightly changed to contextualize the SAQ to the NICU setting, such as 'clinical area' was changed into 'NICU', 'patient' into 'new-born'. The item 'I would feel safe being treated here as a patient' was re-phrased into 'I would feel safe if my child was treated in this NICU'.

In the original SAQ version, the items referring to the *f5* – *perception of management* are measured both at a unit and hospital management level. However, in the Italian version, these items were only considered at a department level. This decision was based on the hierarchical model of most Italian hospitals.

A group of doctors and expert nurses reviewed the translated version to evaluate the appropriateness of the wording for a NICU setting and cultural context. After that, a native English translator performed a back translation from Italian to English. The translated version showed a good correspondence of the meanings of the items.

Data collection

The data collection was performed between March and May 2012. The questionnaires were filled in anonymously, voluntarily and in a confidential form.

The questionnaire was administered in paper format during routine staff meetings.

The following socio-demographical characteristics of the responders were also collected: age, sex, total working experience, years of working experience in NICUs, years of working experience in the participating NICU, profession and type of job contract.

Statistical analysis

Psychometric analyses were performed to assess the fit of the expected factor structure and to test validity and reliability of the Italian version of the SAQ.

A Homogeneity Analysis (HOMALS) was conducted to test the internal homogeneity of the items and the Likert equidistance of item option responses for each dimension described by the SAQ developers. A Guttman's eta (the first eigenvalue) higher than 0.2 and a straight line in the transformation plot for each item indicate an adequate item homogeneity and equidistance, respectively.

The factor structure of the questionnaire was analysed using the confirmatory factor analysis (CFA), using available data from all the subjects. The missing values referring to any single item were assigned the value of the median of all the subjects. We reported the following fit indexes: chi-square, chi-square/degrees of freedom ratio, root mean square error of approximation (RMSEA), standardized root mean square residual index (SRMR).

A model is considered acceptable if RMSEA < 0.08 [26], SRMR < 0.08 [27], chi-square/(degrees of freedom ratio, d.f.) < 2 [28].

Item-factor correlation, that is, 'factor loading' and McDonald's omega reliability were also computed for each dimension [29].

	NICU A	NICU B	NICU C	NICU D	NICU E	NICU F
Beds	44	28	25	18	20	20
Admissions	606	707	389	371	442	582
Doctors	18	10	7	8	11	11
Nurses	58	36	43	26	36	27
Other personnel	25	2	5	2	4	5
Staff (total)	101	48	63	36	57	43
Staff/bed ratio	2.30	1.71	2.52	2.00	2.85	2.15

Table 1 Characteristics (numbers) of the six participating NICUs

The number of admissions refers to 2011.

Correlations >0.45 in absolute value and omega >0.70 were the chosen cut offs to corroborate the structure model.

Once the construct validity was assessed, we calculated the mean score of the Likert items for each dimension and the overall composite score ($F7$) as the mean of the dimension scores. All the scores were standardized to a 100-point scale.

The Kolmogorov–Smirnov test was performed to verify the normal distribution of the scale scores. Descriptive statistics were used to summarize the population characteristics and scores of the items and scales. To test for external validity, the influence of the respondents' characteristics on the mean scores was evaluated using ANOVA for the normally distributed variables and using the non-parametric Kruskal–Wallis test for the not normally distributed ones. The statistical significance was set at $P < 0.05$. The statistical analysis and the CFA were performed using SPSS v.20 (IBM Corp., Armonk, NY, USA) and Mplus software v.1.04 (Muthèn & Muthèn, Los Angeles, CA, USA).

Ethical considerations

This study was based on data regarding patient safety culture among health care providers. It was conducted in compliance with the ethical guidelines of the Helsinki Declaration. All the participants received written information about the purpose of the study, and they were informed that the data would be collected anonymously and treated confidentially. This study was exempt from the need of approval by the local ethics committee.

Results

Response rates and demographics

A total of 193 questionnaires were collected among the six NICUs. One questionnaire was excluded because more than 50% of the answer were missing.

NICU F registered a very low response rate ($<40\%$), owing to an earthquake that involved the area during the administration period. NICU F was evacuated and partially closed for a few months because of severe structural damages. The questionnaires collected from this NICU were considered valid for the psychometric analysis, but were excluded when comparing scores among NICUs because the sample was not representative of the entire NICU.

After excluding NICU F, the mean response rate was 59.7% with a wide variation across all NICUs (range 44.5–95.7%) and across professions (48.1% among doctors and 65.8% among nurses, $P = 0.010$ chi-square test).

Among responders, 91.6% were female, 72.8% were nurses and 11.1% were doctors, the mean age was 36.7 years (standard deviation, $SD = 8.9$ years).

The working experience in a NICU was less than 10 years in 65% of the responders, between 10–20 years in 25.6% and >20 years in 9.4%.

Psychometric analysis

The internal item homogeneity within scales were adequate (from 0.478 to 0.680). The transformation plots of the optimal weights of HOMALS showed that the majority of the items of each dimension described a straight line, except for some items. Thus, accord-

ing to these preliminary results, the latter items were changed in the following way: (i) for items q2, q5, q9, the answer '1' was grouped together with '2'; (ii) for q11, '2' was grouped together with '3'; and (iii) for q25, q27, q32, '4' was grouped together with '5'. Even after changes in response scaling, q7 and q15 did not produce a straight line. Moreover, q15 showed a low item-factor correlation ($r = 0.18$), and so both q7 and q15 were excluded.

Figures 1a–f show the final transformation plots for each dimension (safety climate, job satisfaction, teamwork climate, perception of management, stress recognition, working conditions). Almost all the items were in a linear form and therefore the Likert equidistance assumption was respected. Thus, the items were processed as continuous variables in the CFA phase. The McDonald's omega indexes varied from 0.74 to 0.94, indicating a good reliability for each dimension, the SRMR index was equal to 0.079 (<0.08) indicating a small residual between observed and fitted covariances. The results of the CFA satisfied the chosen cut offs, indicating an acceptable model: the chi-square/d.f. ratio was 1.95, the RMSEA was equal to 0.070 (95% CI = 0.063–0.078) less than the cut off of 0.08.

Table 2 shows the McDonald's omega for the SAQ dimensions and item-factor correlations. Except for item q25, all the items showed factor loadings higher than the chosen cutoff (>0.45), indicating that the allocation of the items in the six factors was acceptable. The six dimensions also correlated with a latent factor ($F7$) and this made it possible to compute the overall composite score.

The stress recognition domain ($f4$) showed a low negative factor loading. Thus, the stress recognition score was considered separately and it was excluded when computing the SAQ composite score ($F7$).

SAQ response pattern

The scores of the validated scales were then computed on our sample. The median of the missing values (MV) was 0.52% (range 0–3.13%), the median of the 'not applicable' response was 1.04% (range 0–7.81%). The 'not applicable' responses were treated as missing values, and so the maximum number of MV was 10.4%.

All the scores were normally distributed except for the $f3$ – job satisfaction and $f4$ – stress recognition score. The mean composite score was 57.6 ($SD 17.9$). The higher scores were reported in the $f4$ – stress recognition domain (mean 77.5, $SD 21.1$, median 81.3, IQR 68.7–93.7), the worst emerged in the $f5$ – perception of management domain (mean 53.8, $SD 21.2$, median 52.8, IQR 37.8–71.9; Table 3).

As found in other studies, we also observed significant differences in all the scores among the units (Table 3), and in three factors ($f4$ – stress recognition, $f5$ – perception of management and $F7$ – SAQ composite score) among professions (Table 4).

Discussion

Psychometric properties of the Italian version of the SAQ in a NICU setting

Many data are available on psychometric properties of translated versions of the SAQ in different countries, but only a few studies

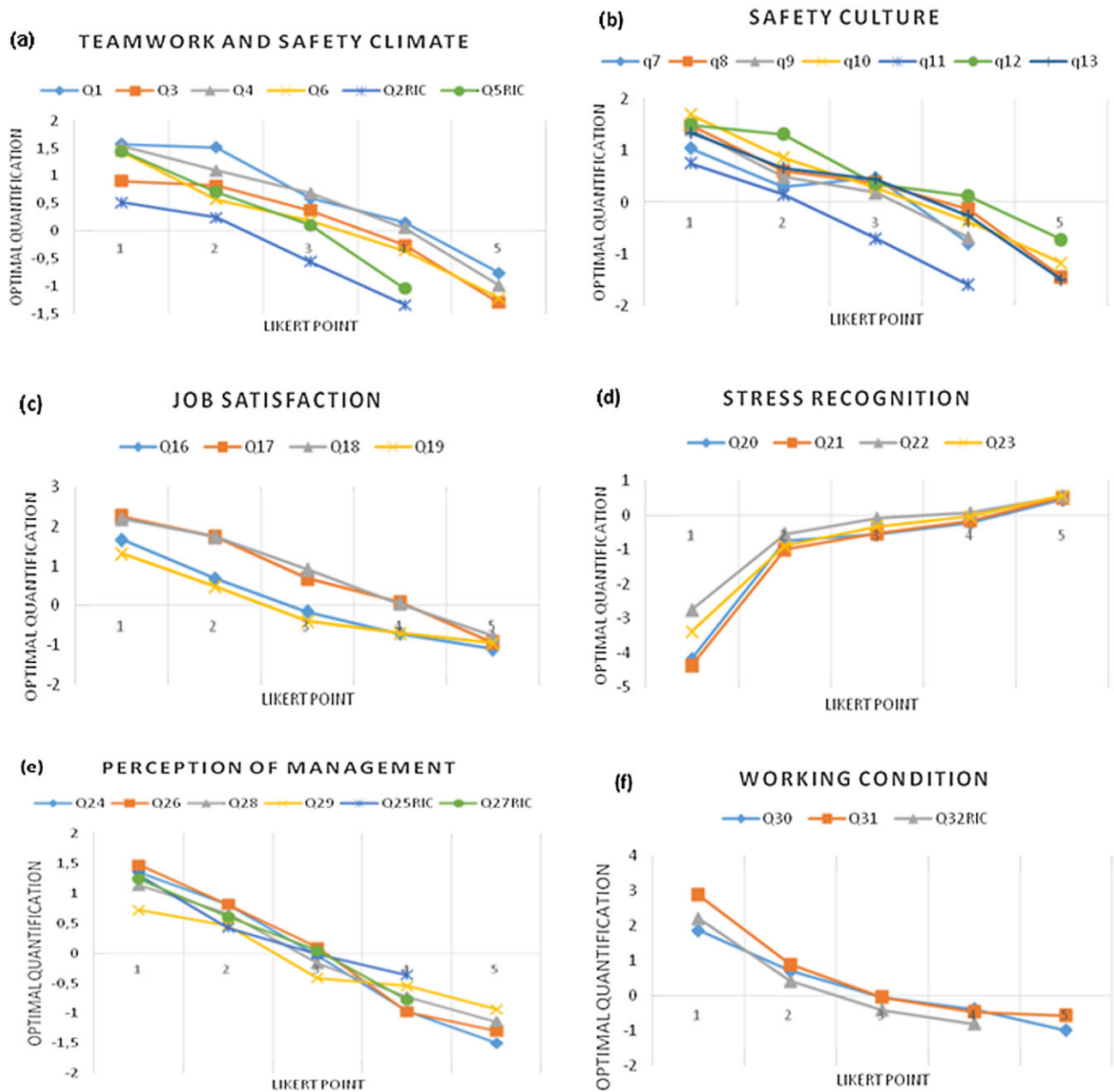


Figure 1 Transformation plots for each dimension.

have been conducted in Italy using this instrument and consequently limited information has been collected for the Italian version. This is the first study aiming to investigate safety culture in Italian NICUs with the SAQ.

Our results indicate that the SAQ maintains its psychometric properties when translated into Italian and administered in NICUs. The fit indexes of the model were acceptable, and the items showed high factor loadings. The two items, q7 and q15, were excluded from the original dimensions and were considered separately because the Likert equidistance assumption was not satisfied. However, the overall questionnaire structure fits and the SAQ

also shows good to excellent reliability indexes when examined on the whole ($F7 - SAQ$ composite score).

All the dimensions had a very high correlation with the $F7 - SAQ$ composite score except for the stress recognition domain. Furthermore, the stress recognition dimension differentiated from other factors, since it reported an inverse correlation with the composite score. This finding is consistent with other studies that highlighted the dissonance of this scale [30]. A possible explanation can be a different interpretation of these items. Staff may consider these items to be a measure of staff stress level, instead of a measure of the staff ability to recognize how stressors may have

Table 2 The McDonald's omega for the SAQ dimensions and item-factor correlations*

	McDonald's omega	Item	Item-factor correlation
<i>f1 – teamwork and safety climate</i>	0.94	q1. Nurse input is well received in this clinical area.	0.52
		q2. In this clinical area, it is difficult to speak up if I perceive a problem with patient care.	0.44
		q3. Disagreements in this clinical area are resolved appropriately (i.e. not who is right, but what is best for the patient).	0.73
		q4. I have the support I need from other personnel to care for patients.	0.57
		q5. It is easy for personnel here to ask questions when there is something that they do not understand.	0.67
		q6. The doctors and nurses here work together as a well-coordinated team.	0.77
<i>f2 – safety culture</i>	0.86	q7. I would feel safe being treated here as a patient.	Excluded
		q8. Medical errors are handled appropriately in this clinical area.	0.68
		q9. I know the proper channels to direct questions regarding patient safety in this clinical area.	0.53
		q10. I receive appropriate feedback about my performance.	0.67
		q11. In this clinical area, it is difficult to discuss errors.	0.50
		q12. I am encouraged by my colleagues to report any patient safety concerns I may have.	0.61
<i>f3 – job satisfaction</i>	0.84	q13. The culture in this clinical area makes it easy to learn from the errors of others.	0.70
		q15. I like my job.	Excluded
		q16. Working here is like being part of a large family.	0.78
		q17. This is a good place to work.	0.76
		q18. I am proud to work in this clinical area.	0.66
		q19. Morale in this clinical area is high.	0.72
<i>f4 – stress recognition</i>	0.86	q20. When my workload becomes excessive, my performance is impaired.	0.69
		q21. I am less effective at work when fatigued.	0.84
		q22. I am more likely to make errors in tense or hostile situations.	0.66
		q23. Fatigue impairs my performance during emergency situations (e.g. emergency resuscitation seizure).	0.54
<i>f5 – perception of management</i>	0.83	q24. Management supports my daily efforts.	0.80
		q25. Management does not knowingly compromise patient safety.	0.31
		q26. Management is doing a good job.	0.82
		q27. Problem personnel are dealt with constructively.	0.64
		q28. I get adequate, timely info about events that might affect my work.	0.60
		q29. The levels of staffing in this clinical area are sufficient to handle the number of patients.	0.48
<i>f6 – working condition</i>	0.74	q30. This hospital does a good job of training new personnel.	0.70
		q31. All the necessary information for diagnostic and therapeutic decisions is routinely available to me.	0.61
		q32. Trainees in my discipline are adequately supervised.	0.70
<i>F7 – SAQ composite score</i>	0.91	<i>f1 – teamwork and safety climate</i>	0.95
		<i>f2 – safety culture</i>	0.94
		<i>f3 – job satisfaction</i>	0.92
		<i>f4 – stress recognition</i>	-0.25
		<i>f5 – perception of management</i>	0.79
		<i>f6 – working condition</i>	0.92

*Question 14 is not part of the scales listed above as indicated by authors of the original SAQ.

Table 3 Mean factors and composite scores for each NICU

	NICU A	NICU B	NICU C	NICU D	NICU E	Total	<i>P</i>
<i>f1 – teamwork and safety climate</i>	69.9	43.2	55.4	64.0	60.7	58.2	<0.001
<i>f2 – safety culture</i>	67.1	42.8	58.4	63.5	60.1	57.8	<0.001
<i>f3 – job satisfaction</i>	74.3	40.8	53.6	68.3	66.5	59.9	<0.001
<i>f4 – stress recognition</i>	72.8	80.1	88.6	64.3	79.5	77.5	<0.001
<i>f5 – perception of management</i>	66.0	42.6	40.8	67.4	54.6	53.8	<0.001
<i>f6 – working condition</i>	71.4	42.0	48.6	67.6	65.1	58.2	<0.001
<i>F7 – SAQ composite score</i>	69.7	42.3	51.4	66.2	61.4	57.6	<0.001

P < 0.05 are in bold type. *P*-value at ANOVA for normally distributed variables and Kruskal–Wallis for not normally distributed ones (*f3* and *f4*).

	Doctors	Nurse	Other	Total	<i>P</i>
<i>f1 – teamwork and safety climate</i>	59.5	56.9	67.2	58.2	0.173
<i>f2 – safety culture</i>	57.2	57.1	66.4	57.9	0.199
<i>f3 – job satisfaction</i>	55.6	59.7	68.5	59.7	0.228
<i>f4 – stress recognition</i>	86.2	76.8	67.9	77.8	0.041
<i>f5 – perception of management</i>	49.9	52.9	71.8	53.9	0.003
<i>f6 – working condition</i>	57.6	56.6	72.5	58.0	0.061
<i>F7 – SAQ composite score</i>	56.0	56.6	69.3	57.6	0.038

P < 0.05 are in bold type. *P*-value at ANOVA for normally distributed variables and Kruskal–Wallis for not normally distributed ones (*f3* and *f4*).

Table 4 Mean factors and composite scores among professions

a negative impact on patient safety (as in the intention of the scale authors). Nevertheless, the stress recognition domain does not affect the reliability of the entire questionnaire. However, because of these factors, it was excluded when computing *F7 – SAQ composite score*.

Centres and profession comparison

The response rate was acceptable, nearly 60% as recommended by the developers. The response rate varies widely across the NICUs and this can be due to the way the questionnaire was administered among the centres, considering that the SAQ was given at the beginning of routine staff meetings, but in some cases, the staff was asked to fill it in after the meeting. Moreover, people who were not at work that day received the questionnaire from the nurse coordinators but the response rate was much lower in this case.

Similar to many other studies, we found that the response rate was lower among doctors than nurses, which justifies the sample characteristics, meaning that the study population was mostly made up of young females.

The missing values analysis revealed a low proportion of missing values (MV). The number of MV was lower than 5% for all the items, except for *q25* (MV = 10.4%), which was anyway lower than in other studies [14,31,32]. The high MV at *q25* may be due to a difficulty in interpreting the item meaning, so a better translation or wording may be helpful to reduce the MV.

The external validity of the Italian version of SAQ was confirmed, enabling benchmarking among organizations. The analysis highlighted significant differences among NICUs when considering both patient safety as a whole and in its different aspects.

Interestingly, the NICU with the worse SAQ composite score (and the lowest score on the dimension *f3 – job satisfaction*) had the lowest staff/bed ratio, whereas NICU A had a high staff/bed ratio and registered the highest SAQ composite score. Literature shows a significant relationship between inadequate staffing levels and negative outcomes such as burnout and job dissatisfaction [33,34], it also demonstrates that a lack of team collaboration and weak leadership are associated with high nurse workload [35,36]. Therefore, working conditions, teamwork and management are important factors that can influence personnel stress levels. In our study, these dimensions are positively correlated with the SAQ composite score.

These observations suggest that organizational and structural characteristics of the involved hospitals can significantly influence the awareness of safety culture issues. Further investigations are needed to identify specific determinants of the differences

retrieved among the centres. Such information may guide hospital managements to identify the positive and negative organizational aspects affecting safety issues and climate.

We also reported significant differences among professions in two dimensions (*f4 – stress recognition* and *f5 – perception of management*). In other words, doctors showed the highest scores in the former and the worst in the latter. Results regarding stress recognition suggest that doctors are able to identify that a stressful environment may lead to unsafe work conditions. However, as previously described, our data analysis raised doubts about a possible misinterpretation of these items, because the staff may have evaluated their level of stress. This observation lead to difficulties in correctly interpreting this outcome. The perception of management was rated lower for nurses or doctors than for other staff members. This result may have different explanations:

- 1 Doctors and nurses are much more likely to directly collaborate with the management than other staff members. Therefore, gaps in management performance may be perceived much more distinctly by nurses and doctors;
- 2 The duties of staff members, who are not doctors and nurses, affect patient safety much less and so they may feel more satisfied with the work of the management, and need less support than other professions, and
- 3 Other staff members may have a stronger perception of hierarchy and they may therefore have rated the management with higher scores to avoid disappointment.

There were also significant differences in the *F7 – SAQ composite score*, because it was lower among doctors in comparison with nurses. Sexton *et al.* highlighted differences among professions, although they found higher scores among doctors [24]. The cultural context influences the clinical approach of professions in NICUs. In fact, this has been shown when the clinical activities and outcomes in an Italian NICU were compared with those recorded in the Vermont Oxford Network database [37]. Our findings suggest that the cultural context is crucial to determinate safety attitudes and perceptions as well.

Approaches for improvement

The fragility of patients who are treated in NICUs requires a high standard of care. Developing an adequate safety culture may be a key issue to guarantee such a high standard.

The dimensions and SAQ composite scores highlight the possibility of improving patient safety. Overall, almost all the dimensions can be significantly improved. However, the already mentioned differences among the centres require the implementa-

tion of tailored interventions. Each participating NICU needs a personalized feedback and the results need to be discussed to reach a better understanding of both the positive and negative evaluation. Context-specific factors (such as the hypothesis regarding the lower staff/bed ratio for the negative evaluation of the NICU B) may then be analysed and other centres may provide solutions for improvement. For example, centres with higher teamwork and safety climate scores may be asked for advice regarding their positive results. Continuous staff training combined with communication and non-technical skills promotion are required to implement safety culture in NICUs. The management should endorse the improvement process with specific actions aiming at the weakest areas pointed out in the questionnaire results.

This study showed some limits. First, the sample size was relatively small and limited the analysis. Secondly, the Italian version of SAQ-SF was tested in a specific context (NICU) and thus it requires caution when extending the results to other clinical areas, especially as regard the exclusion of some items from the scale.

Conclusions

In the present study, we confirmed the psychometric properties of the Italian version of SAQ and it proved to be an effective tool to evaluate and compare safety culture perceptions in NICUs.

This study can be a starting point for further investigations to consolidate the results in Italian NICUs and to extend the use of this tool to other hospital units. Moreover, it will be important to explore the effect of organizational characteristics on safety culture and the correlations between SAQ scores and clinical outcomes in NICUs.

References

1. Health and Safety Executive (1993) ACSNI Study Group on Human Factors: Third Report – Organizing for Safety. London: HSE Books.
2. Australian College of Neonatal Nurses Inc. (2012) Australian Standards for Neonatal Nurses, 3rd edn. Camperdown, NSW, Australia: ACNN.
3. Joint Commission International (2013) Accreditation Standards for Hospitals, 5th edn. Oakbrook Terrace, IL, USA: JCI.
4. Sexton, J. B., Thomas, E. J. & Helmreich, R. L. (2000) Error, stress, and teamwork in medicine and aviation: cross sectional surveys. *BMJ (Clinical Research Ed.)*, 320, 745–749.
5. Weaver, S. J., Lubomski, L. H., Wilson, R. F., Pfoh, E. R., Martinez, K. A. & Dy, S. M. (2013) Promoting a culture of safety as a patient safety strategy a systematic review. *Annals of Internal Medicine*, 158, 369–374.
6. Snijders, C., van Lingen, R. A., Molendijk, A. & Fetter, W. P. (2007) Incidents and errors in neonatal intensive care: a review of the literature. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, 92, 391–398.
7. Dean, G. E., Scott, L. D. & Rogers, A. E. (2006) Infants at risk: when nurse fatigue jeopardizes quality care. *Advances in Neonatal Care*, 6 (3), 120–126.
8. Kaushal, R., Bates, D. W., Landrigan, C., McKenna, K. J., Clapp, M. D., Federico, F. & Goldmann, D. A. (2001) Medication errors and adverse drug events in pediatric inpatients. *JAMA: The Journal of the American Medical Association*, 285, 2114–2120.
9. Sharek, P. J. & Classen, D. (2006) The incidence of adverse events and medical error in pediatrics. *Pediatric Clinics of North America*, 53 (6), 1067–1077.
10. Sexton, J. B., Helmreich, R. L., Neilands, T. B., Rowan, K., Vella, K., Boyden, J., Roberts, P. R. & Thomas, E. J. (2006) The safety attitude questionnaire: psychometric properties benchmarking data, and emerging research. *BMC Health Service Research*, 6, 44.
11. Sexton, J. B., Berenholtz, S. M., Goeschel, C. A., Watson, S. R., Holzmüller, C. G., Thompson, D. A., Hyzy, R. C., Marsteller, J. A., Schumacher, K. & Pronovost, P. J. (2011) Assessing and improving safety climate in a large cohort of intensive care units. *Critical Care Medicine*, 39, 934–939.
12. Profit, J., Etchegaray, J., Petersen, L. A., Sexton, J. B., Hysong, S. J., Mei, M. & Thomas, E. J. (2012) The Safety Attitudes Questionnaire as a tool for benchmarking safety culture in the NICU. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, 97, F127–F132.
13. Colla, J. B., Bracken, A. C., Kinney, L. M. & Weeks, W. B. (2005) Measuring patient safety climate: a review of surveys. *Quality & Safety in Health Care*, 14 (5), 364–366.
14. Deilkås, E. T. & Hofoss, D. (2008) Psychometric properties of the Norwegian version of the Safety Attitudes Questionnaire (SAQ), Generic version (Short Form 2006). *BMC Health Service Research*, 8, 191.
15. Lee, W. C., Wung, H. Y., Liao, H. H., Lo, C. M., Chang, F. L., Wang, P. C., Fan, A., Chen, H. H., Yang, H. C. & Hou, S. M. (2010) Hospital safety culture in Taiwan: a nationwide survey using Chinese version safety attitude questionnaire. *BMC Health Service Research*, 10, 234.
16. Kaya, S., Barsbay, S. & Karabulut, E. (2010) The Turkish version of the safety attitudes questionnaire: psychometric properties and baseline data. *Quality & Safety in Health Care*, 19, 572–577.
17. Devriendt, E., Van den Heede, K., Coussemont, J., et al. (2012) Content validity and internal consistency of the Dutch translation of the Safety Attitudes Questionnaire: an observational study. *International Journal of Nursing Studies*, 49 (3), 327–337.
18. Göras, C., Wallentin, F. Y., Nilsson, U. & Ehrenberg, A. (2013) Swedish translation and psychometric testing of the safety attitudes questionnaire (operating room version). *BMC Health Service Research*, 13, 104.
19. Nordén-Hägg, A., Sexton, J. B., Källemark-Sporrong, S., Ring, L. & Kettis-Lindblad, A. (2010) Assessing safety culture in pharmacies: the psychometric validation of the Safety Attitudes Questionnaire (SAQ) in a national sample of community pharmacies in Sweden. *BMC Clinical Pharmacology*, 10, 8.
20. Makary, M. A., Sexton, J. B., Freischlag, J. A., Millman, E. A., Pryor, D., Holzmüller, C. & Pronovost, P. J. (2006) Patient safety in surgery. *Annals of Surgery*, 243 (5), 628–632.
21. Modak, I., Sexton, J. B., Lux, T. R., Helmreich, R. L. & Thomas, E. J. (2007) Measuring safety culture in the ambulatory setting: the safety attitudes questionnaire – ambulatory version. *Journal of General Internal Medicine*, 22 (1), 1–5.
22. Patterson, P. D., Huang, D. T., Fairbanks, R. J. & Wang, H. E. (2010) The emergency medical services safety attitudes questionnaire. *American Journal of Medical Quality*, 25 (2), 109–115.
23. Raftopoulos, V., Savva, N. & Papadopoulou, M. (2011) Safety culture in the maternity units: a census survey using the Safety Attitudes Questionnaire. *BMC Health Service Research*, 11, 238.
24. Profit, J., Etchegaray, J., Petersen, L. A., Sexton, J. B., Hysong, S. J., Mei, M. & Thomas, E. J. (2011) Neonatal intensive care unit safety culture varies widely. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, 97 (2), 120–126.
25. Hamdan, M. (2013) Measuring safety culture in Palestinian neonatal intensive care units using the Safety Attitudes Questionnaire. *Journal of Critical Care*, 28 (5), 886.e7–886.e14.
26. Vandenberg, R. J. & Lance, C. E. (2000) A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3 (1), 4–70.

27. Hu, L. & Bentler, P. M. (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
28. Byrne, B. (1989) A primer of LISREL: Basic Applications and Programming for Confirmatory Factor Analytic Models. New York: Springer Verlag.
29. McDonald, R. P. (1999) Test Theory. A Unified Treatment. London: Lawrence Erlbaum Associates.
30. Taylor, J. A. & Pandian, R. (2013) A dissonant scale: stress recognition in the SAQ. *BMC Research Notes*, 6, 302.
31. Zimmermann, N., Küng, K., Sereika, S. M., Engberg, S., Sexton, B. & Schwendimann, R. (2013) Assessing the safety attitudes questionnaire (SAQ), German language version in Swiss university hospitals – a validation study. *BMC Health Services Research*, 13, 347.
32. France, D. J., Greevy, R. A. Jr, Liu, X., Burgess, H., Dittus, R. S., Weinger, M. B. & Speroff, T. (2010) Measuring and comparing safety climate in intensive care units. *Medical Care*, 48 (3), 279–284.
33. Aiken, L. H., Clarke, S. P., Sloane, D. M., Sochalski, J. & Silber, J. H. (2002) Hospital nurse staffing and patient mortality, nurse burnout and job dissatisfaction. *JAMA: The Journal of the American Medical Association*, 288 (16), 1987–1993.
34. Profit, J., Sharek, P. J., Amspoker, A. B., Kowalkowski, M. A., Nisbet, C. C., Thomas, E. J., Chadwick, W. A. & Sexton, J. B. (2014) Burnout in the NICU setting and its relation to safety culture. *BMJ Quality & Safety*, 23 (10), 806–813.
35. Spence, K., Tarnow-Mordi, W., Duncan, G., Jayasuryia, N., Elliott, J., King, J. & Kite, F. (2006) Measuring nursing workload in neonatal intensive care. *Journal of Nursing Management*, 14 (3), 227–234.
36. O'Brien-Pallas, L., Duffield, C. & Alksin, C. (2004) Who will be there nurse? Retention of nurses nearing retirement. *The Journal of Nursing Administration*, 34 (6), 298–302.
37. De Nisi, G., Berti, M., Malossi, R., Pederzini, F., Pedrotti, A. & Valente, A. (2009) Comparison of neonatal intensive care: Trento area versus Vermont Oxford Network. *Italian Journal of Pediatrics*, 35, 5–11.