

A strategy to develop tailored patient safety culture improvement programs with latent class analysis method

Abstract

As patient safety can be built only on the bedrock of safety culture, many hospitals have designed and implemented their own safety culture improvement programs. Although carefully tailored clinical area specific programs would be most effective, lack of resources precludes this. This study applied a latent class analysis method and classified 72 clinical areas in a large tertiary hospital into four classes according to their score patterns in the six domains of the Korean version of Safety Attitudes Questionnaire (SAQ-K). Bayesian information criterion (BIC) and bootstrapped likelihood ratio test (BLRT) were used in deciding the number of classes. We named the best class estimate in each of the SAQ-K domains the 'current maxima,' the goals of which other classes aim to improve. Because the job satisfaction domain cannot be directly improved, we excluded it from the goal setting process. One class dominated the other three in most domains. The difference between the best class and others, the 'room for change,' can be used in deciding which clinical areas should be focused first and how much resources should be invested in developing safety culture improvement programs. Stress recognition domain, which showed only negligible room for change, was therefore excluded from the program development phase in this particular hospital. We expect this article to help hospitals to better design and implement their safety culture improvement programs, despite a chronic lack of resources.

Keywords: safety culture, safety attitudes questionnaire, patient safety, program development, latent class analysis

Volume 2 Issue 2 - 2015

Heon-Jae Jeong,¹ Minji Kim,² EunAe An,³ So Yeon Kim,⁴ Byungjoo Song⁴

¹Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, USA

²Annenberg School for Communication, University of Pennsylvania, USA

³Leadership Development Division, The Catholic Education Foundation, Korea

⁴Performance Improvement Team, Seoul St. Mary's Hospital, The Catholic University of Korea, Korea

Correspondence: Heon-Jae Jeong, Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, 624 North Broadway, Rm. 455, Baltimore, MD, 21205, USA, Tel 410-955-5315, Fax 410-955-6959, E-mail hjeong1@jhu.edu

Received: April 02, 2015 | **Published:** April 11, 2015

Introduction

This short article is the last in a series of articles on measuring and analyzing patient safety culture in the hospital setting. We have already covered the development of an instrument to measure the safety culture in Korea and a methodology to obtain accurate estimates of safety culture.^{1,2} In this article, we explain how to use these safety culture scores and propose a practical way of developing more effective safety improvement programs in hospitals. Because we used the results of the previous articles as the dataset of this study, detailed information of the raw data is not provided except when required.

Patient safety rests on the solid foundation of safety culture.³ Improving safety culture requires measuring it first, and therefore many resources have been invested in the development of instruments to gauge the culture.⁴⁻⁶ The Safety Attitudes Questionnaire (SAQ) by Sexton et al. is one of the most popular instruments⁷ and has been used in various healthcare settings including the intensive care units, emergency departments and operating rooms⁸⁻¹² of many countries.¹²⁻¹⁷

To provide an established safety culture measurement tool that allows international comparison, we developed and tested the Korean version of Safety Attitudes Questionnaire (SAQ-K).¹ Like the original SAQ does, SAQ-K is comprised of six domains, and contains 34 items shown in Table 1.⁷

Even with such validated instruments as SAQ-K, hospital safety managers are highly likely to encounter a scarcity of resources. To illustrate, they are tempted to develop programs for each clinical area: if a certain clinical area has low TC score, then they would come up with a program to improve the domain for that area, and iterate this

step for the other domains, and also for all the other clinical areas. Since programs tailored to a certain group are known to bring better results,^{18,19} this strategy is very attractive, but it requires unobtainable resources. Most large hospitals have scores, sometimes more than a hundred clinical areas, and developing and providing the most suitable program for each of them simultaneously is practically impossible. In many cases, safety managers return to a single generic program for the entire hospital.

Is it possible to stratify clinical areas into a manageable number of categories or classes with similar characteristics? If so, we would be able to tailor safety culture improvement programs for each category. This will maximize the effectiveness of available resources. The question, then, is how to categorize tens of clinical areas according to their SAQ-K profiles.

Table 1 SAQ Domain Definitions

SAQ Domain	Definition (number of items)
Teamwork Climate (TC)	Perceived quality of collaboration between personnel (5)
Safety Climate (SC)	Perception of a strong and proactive organizational commitment to safety (6)
Job Satisfaction (JS)	Positivity about the work experience (5)
Stress Recognition (SR)	Acknowledgment of how performance is influenced by stressors (4)
Perception of Management (PM)	Approval of managerial action (10)
Working Conditions (WC)	Perceived quality of the work environment and logistical support (4)

Another hardship that safety managers face is that there are few ways to take advantage of all the domains of SAQ-K simultaneously; this means safety culture has not been understood as a whole but rather as an entity segmented by SAQ domains. In this article, we used latent class analysis (LCA), also known as the finite mixture model, to categorize all clinical areas into a few classes, using clinical area SAQ estimates of all domains. The premise of LCA is quite straightforward: “there might be different distinctive but latent classes that affect the SAQ-K score patterns.”²⁰ For a hypothetical example, a group of clinical areas may show a pattern of very high TC, SC, and JS, but low in SR and mid PM and WC scores, another group has low scores in all domains but SR, and so forth. LCA can detect such patterns, and determine which clinical areas belong to which class. Statistically speaking, LCA is “identifying unobserved heterogeneity in a population.”²¹ Therefore, with LCA we can consolidate all six domains of SAQ-K scores and reveal the patterns across classes, or categories, of clinical areas. We can then analyze class-specific SAQ-K estimates and develop a safety improvement intervention.

This study, therefore, has a two-fold objective: to apply LCA and find distinctive classes of clinical areas based on SAQ-K domain scores; and to introduce a way to set goals of a hospital-wide safety culture improvement programs for each class based upon the LCA results. However, describing how to develop a safety improvement program is too ambitious of an undertaking in a single article. The second part of this article therefore focuses on how to prioritize the SAQ domains to be improved through a program.

Methods

Data preparation for LCA

We used the same dataset for the previous articles in this series. Since SAQ-K validation and its psychometric property were already addressed in those articles, we do not include them here. SAQ-K was administered to healthcare workers (HCWs) in a large tertiary hospital in Seoul from October to November 2013, and a total of 1,381 questionnaires were returned. After excluding questionnaires with too many missing values, 1,142 questionnaires from 72 clinical areas were included in the analysis. Each item of SAQ-K was measured with five-point Likert-type scales from *disagree strongly* (1) to *agree strongly* (5) and then converted into a 0 to 100 scale.^{1,7} SAQ-K Domain score of each respondent was the arithmetic mean of the item scores in the domain. Once we calculated person-level SAQ-K domain scores, we applied empirical Bayes method and obtained clinical area-specific SAQ-K domain scores. Since most safety improvement programs would take place within a clinical area,²²⁻²⁴ this study also used clinical areas as unit of analysis. Detailed information of this step can be found in the previous article.¹

Classifying clinical areas with LCA

With the six SAQ-K domain estimates from 72 clinical areas, we began by developing a two-class LCA model. Then we increased the number of classes in a model, one by one. Each time, we tested whether the addition of another class significantly improved the model fit with bootstrapped likelihood ratio test (BLRT).²⁵ We repeated this incremental increase in the number of classes until there ceased to be additional value with the addition of another class. Computationally, though it was done by software package, BLRT was conducted in three steps. First, we ran k and $k-1$ class models and calculate twice the difference in log-likelihoods (Likelihood-ratio (LR) test statistic).

Second, we used the $k-1$ model estimates from the first step to generate data and obtain LR test statistics, iterating this step to achieve solid bootstrap distribution of LR test statistics. Third, we compared the LR test statistic from the first step with the bootstrap distribution from the second step, to test the $k-1$ class model against k class model, estimating p -value.²⁵

We chose BLRT among other test statistics, because we used only 72 clinical area-level observations as units of analysis, and BLRT was a consistent method for such a relatively small sample. We also referred to Bayesian information criterion (BIC) to solidify our model decision.²¹

Determining the number of classes does not depend solely on a single statistic, rather, researchers use multiple statistics simultaneously to find the best model,²¹ and in many cases a priori information on the study topic should be considered. The result section describes how we did this. Once we arrived at the model of our choice, we calculated SAQ-K domain estimates for each class and observed patterns among the classes.

Exploring SAQ-K domain scores to develop class-specific safety culture improvement programs

This section does not require much statistical background, but is critical in designing hospital-wide safety improvement program. We applied agreed-upon strategies from the communication and public health fields.^{26,27} However, we also encourage readers to be creative, and to come up with best strategy for their own workplace.

First, for each SAQ-K domain, we picked the highest estimates among the classes; we call it the ‘current maximum (optimum)’ of the domain within the hospital. These were used as the realistic goals for each class, considering its infrastructure or the hospital environment. Then, for each class, we calculated the difference between its SAQ-K score and that of the current maximum for each domain; this difference was the ‘room for change,’ and was used to prioritize what domains in which class should be addressed.

We used Stata 13.1 (StataCorp, College Station, Texas) to obtain empirical Bayes estimates and Mplus 7.3 (Muthén & Muthén, Los Angeles, California) for LCA.

Results

Respondents’ characteristics and psychometric property of SAQ-K is available in our previous articles.¹

Decision on the number of classes

We began with a two-class model (72 areas divided into two classes), and then increased the number of classes until the p -value from BLRT was not significant. We stopped at the six-class model, which means that dividing 72 clinical areas into six classes did not add statistical value, when compared to dividing them into five classes. Then, we compared BIC and the number of clinical areas in the smallest class from each model. As shown in Table 2, the smallest class of the five-class model contains only one clinical area, which would not help develop hospital-wide safety culture improvement programs. BIC-wise, the four-class model showed the best result, the smallest value. The four-class model was thus identified as the best for developing safety programs in this hospital.

Table 2 Model Decision Criteria and Results

Model decision criteria used	Number of classes in a model				
	2	3	4	5	6
Bootstrapped likelihood-ratio test (BLRT, <i>p</i> -value)	< 0.00	< 0.00	< 0.00	< 0.00	0.67
Bayesian Information Criterion (BIC)	2097.69	2077.01	2074.16	2078.50	2086.64
Number of the clinical areas in the smallest class	27	12	4	1	1

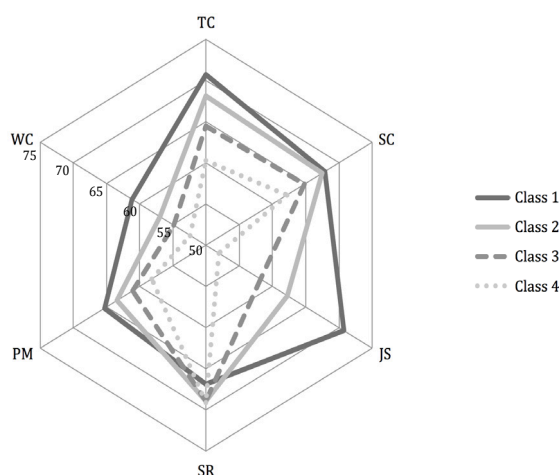
Table 3 Characteristics of the Classes

Class	Number of Clinical Areas	SAQ-K Domain Scores (Standard Error)					
		TC	SC	JS	SR	PM	WC
Class 1	4	70.69 (2.06)	67.89 (0.39)	70.75 (1.43)	66.87 (1.06)	65.30 (0.55)	61.15 (0.51)
Class 2	29	68.13 (0.47)	67.48 (0.30)	62.27 (0.86)	68.96 (0.21)	63.38 (0.21)	56.93 (0.23)
Class 3	27	64.43 (0.53)	64.85 (0.23)	58.17 (0.88)	68.92 (0.33)	61.07 (0.27)	54.85 (0.35)
Class 4	12	60.30 (0.85)	62.20 (0.62)	51.90 (1.59)	69.09 (0.88)	58.20 (0.58)	52.38 (0.48)

Characteristics of the classes

The SAQ-K domain estimates of each class are shown in Table 3; the same information was also plotted on a radar chart to visualize the topography of class-specific SAQ-K scores (Figure 1). Four clinical areas were categorized as class 1, and their domain estimates were higher than those of other classes except for SR domain. In class 1 clinical areas, the domain estimates spanned from 61.15 for WC to 70.75 for JS. Taking the most clinical areas (29), class 2 domain estimates were spread from 56.93 for WC to 68.96 for SR. Compared to class 1 estimates, JS and WC of class 2 were notably lower. Class 3 had 27 clinical areas and the domain estimates ranged from 54.85 for WC and 68.92 for SR. Twelve clinical areas fell under class 4, in which the highest and lowest WC and SR estimates were 52.38 and 69.09, respectively.

As shown in Figure 1, SAQ-K estimates across classes varied by domain, as they should, but there was very little difference in SR scores for all four classes, from 66.87 for class 1 to 69.09 for class 4.


Figure 1 Radar Chart of SAQ-K Domain Estimates by Class.

Determining the current maxima of the SAQ-K classes and setting the improvement goals

For each domain, we chose the best class estimate as the goal for the other classes. In this selection process, we excluded JS and SR domains. JS, the extent to which HCWs in a clinical area are satisfied with their jobs, is the results of various factors, and therefore it is not a domain that one can improve directly. The SR domain was also excluded from the selection, because there was too small a difference in SR domain estimates across the four classes, and thus there was practically no room for improvement in SR domain. Thus we used TC, WC, PM and WC domains to develop safety culture improvement programs.

In this hospital, class 1 clinical areas surprisingly outperformed the other three classes for all the four domains. (See Table 4 and Figure 2 for an illustration of room for change, calculated as the difference between class 1 and the remaining classes in each domain).

Table 4 Room for Change for SAQ-K Domains

SAQ-K Domain	Class 2	Class 3	Class 4
TC	2.57	6.26	10.40
SC	0.41	3.04	5.69
PM	1.92	4.23	7.10
WC	4.22	6.30	8.77

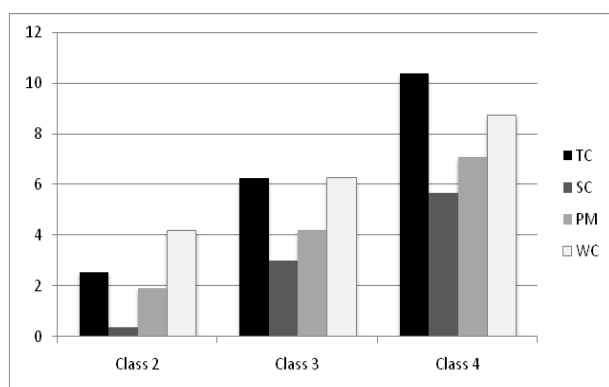


Figure 2 Room for Change in SAQ-K Domains.

Apparently, class 4 showed the greatest room for change for all four SAQ-K domains. Among them, TC showed the biggest, followed by WC, PM and then SC. Class 3, following Class 4, had the most room for change in WC and TC. PM and SC showed less room for change than the other domains but still they were visible. Class 2 showed the least room for change, the difference with current maxima. Note that in class 2, WC showed much more room for change, unlike the other classes. This room for change can guide the setting of goals for improvement, and the domain that draws more attention than the others in each class.

Discussion

There is an old saying in the management field, “if you can’t measure it, you can’t manage it.”²⁸ Though true, it does not mean that accurate measurements always bring the best management. As a sequel to our two articles that were dedicated to measuring safety culture, this study links measurement to its actual application in management. We have tried to keep this study very practical and easy for frontline safety managers to understand. However, this does not mean that we avoided solid statistical methodology. Au fond, seasoned patient safety specialists well know that successful programs should be data-driven, which means that they should be built on the rigorous analytical methods to derive the maximum buy-in from the management and also HCWs, and to yield the best possible results. As before, we provide a detailed step-by-step description on our methodology, so safety managers in the field can readily apply it.

This study assumed typical hospitals where a central quality and patient safety (QPS) department is in charge of planning hospital-wide safety culture improvement programs. In this structure, safety managers in the QPS department tend to develop and implement a single uniform program across the hospital. Although tailored messages and bespoke programs have been shown to bring superior results to a uniformed one,^{18,19,22} safety managers lack the resources to develop and implement programs for each clinical area. We used LCA and categorized 72 clinical areas into four distinct classes by their SAQ profiles. The beauty of using LCA is that it simultaneously takes into account all six domains of SAQ-K estimates to classify clinical areas and reveal their patterns, unlike most healthcare organizations that use one SAQ domain at a time. This approach provides a more holistic understanding of patient safety culture itself.

However, note that there are dozens of cluster analysis methods with long history other than LCA; *k*-means first appeared in 1950’s,²⁹ and fuzzy c-means clustering algorithm in 1984.³⁰ Development

of clustering methods continued in the 21st century, such as cluster ensemble (2003),³¹ overlapping clustering (2005),³² and so forth. The problem is, as Jain et al. describes, there is no single best algorithm, indeed, “each algorithm, implicitly, or explicitly, imposes a structure on the data,”²⁹ and therefore, if the clustering meets the purpose, the algorithm can be said to be functioning. In this particular study, LCA sufficed our goal of clustering clinical areas to a manageable number of classes to develop safety improvement programs. Therefore, we did not try other clustering methods. However, we admit further study is necessary to test other clustering methods on SAQ data, the study that would add great value to knowledge base on how to get the most out of safety culture survey data, especially with a larger dataset such as the one that was collected from all the hospitals in a country.

In LCA, deciding the best number of classes of clinical areas was critical part; as Nylund et al. pointed out, there is no consensus on how to determine the number of classes especially in mixture model, albeit several suggestions have been made.²¹ Actually, statistical software packages like Mplus that we used provide various model fit indices, such as Akaike’s information criterion (AIC), Bayesian information criterion (BIC), Vuong-Lo-Mendell-Rubin likelihood ratio test and bootstrapped likelihood ratio test (BLRT). We based our decision on Nylund et al.’s simulation study with samples of various sizes. According to them, “BIC performed the best of the information criteria, and BLRT proved to be a very consistent indicator.”²¹ Note that even with only those two indicators, we were confronted with discrepancy: BIC showed the 4 class model was the best, but BLRT showed that 5 was. Therefore we considered practical application of the result of this study. Though BLRT showed the 5 class model is superior to the 4 class model, the additional class only included just one clinical area, and designing a program for one area is not consistent with the purpose of this study. Given such a flood of model fit indices, one might need to compromise to decide the best number of classes whenever running a mixture model, and the goal of the classification should guide the decision, as it did in this study.

We had expected each class to exhibit very different characteristics or patterns in SAQ-K domain estimates. For example, one class may dominate one domain, but be dominated in others. However, in this dataset, all domains except for SR showed their estimates in the same order, and therefore, Figure 1, the radar chart, showed hexagons becoming smaller from class 4 through class 1, without lines crossing each other. This finding should never be generalized to all hospitals, but such unexpected consistency in SAQ-K domain estimates patterns across classes led us to consider that safety culture might follow a growth curve. To date, few studies have looked at safety culture as an evolving or growing curve in a certain direction; further study from this perspective would add great value to the knowledge base on safety culture.

We used ‘room for change’ to set the goals of hospital-wide safety culture improvement programs. We set the goals of improvement at the highest estimates of any class for each domain, not the perfect score of 100. This decision is based upon the assumption that safety culture of each clinical area is affected by an entire hospital’s physical and cultural infrastructure, and therefore, aiming for 100 points from SAQ-K scores is neither possible nor logical. We borrowed the statistical term ‘local maxima’ and named those SAQ domain goals as ‘current maxima,’ the connotation of which is that there will be better status in the future, but presently and practically we pursue this goal. Each class, therefore, showed different size of room for change for each domain as depicted in Figure 2, and hospital safety managers

can develop and implement safety culture improvement program according to this room for change profile. Typically, hospitals administer SAQ regularly—no shorter than a year (e.g., every 18 months for Johns Hopkins Hospital), and thus hospitals can evaluate the effectiveness of their programs, and set new goals. Indeed, this approach is iterative, pushing the envelope of current maxima every time, providing momentum for every clinical area to reach the next level in its safety culture. Note that the scores of class 1 were the goals of all domains in this hospital, and according to this study, there might not be anything for class 1 areas to do. However, the purpose of this approach is to concentrate resources where they are most needed, and therefore investing resources in the other classes may be logical. However, since this process is iterative, the classes and goals will be different in the next cycle, and thus eventually all clinical areas are the targets of hospital-wise safety culture improvement programs.

One might think using the best clinical area score of each domain would be simpler and better. However, we suggest class level SAQ domain estimates, because there are many clinical areas with only a handful of HCWs working, and using estimates of such small areas would not be the best way. There might also be clinical areas with distinct characteristics that are not shared by most clinical areas (e.g., operating room or emergency department). To avoid picking such an area and setting its scores as the goal for the entire hospital, using class-level estimates, which were obtained from multiple clinical areas in the class, might be a reasonable way to develop more successful programs.

We did not use the SR domain for setting the goal of improvement due to invariance in its estimates across the classes. This study does not provide any information from which such ignorable difference arose, but we speculate that some factors might influence each HCW's acknowledgement of how performance is influenced by stressors, factors that override the influence of clinical areas. Hospital- or even national-level cultural characteristics might be one possible reason,³ but we are rather cautious about making any specific interpretation. Further study on stress recognition and safety culture would clarify our understanding.

We did not go beyond setting the goals for classes, leaving the program development process at the discretion of the hospital safety managers. It is out of the scope of this study, but more importantly, each hospital has its own resources, governance structure and even preference for safety culture improvement programs. Despite such variability, as seasoned safety experts, we offer the following suggestions. HCWs and hospitals executives should never use SAQ scores as a grading instrument. Once HCWs perceive SAQ as a performance indicator, it ceases to reflect reality, though the scores might dramatically increase. The classification method introduced in this article should be understood only as a tool to decide how better to support HCWs and their working areas. In this regard, calling each class 'class 1,' 'class 2' might not be the best way, since it implies rankings (we did so only because this is an academic article). We suggest more positive wording, such as 'teamwork dominant group' for a class with higher TC estimates than other domains.

Conclusion

We are near the end of this journey. In this series of articles, we have developed the Korean version of Safety Attitudes Questionnaire and devised a way to obtain more accurate estimates of SAQ-K scores across tens of clinical areas in a hospital with empirical Bayes method.¹

We also revealed that clinical area and job type had combinational effects on healthcare professionals' safety attitudes, and developed a formula to estimate them with crossed random effects model [2]. We then applied latent class model to categorize clinical areas and suggested goals for the enhancement of safety culture improvement programs.

We hope and believe that this series of articles can help many hospitals and healthcare professionals to improve. Please ensure that your hospital is as safe as possible. It will save patients' lives.

Contact the corresponding author at hjeong1@jhu.edu for SAQ-K, and Stata and Mplus commands for the analyses in this article.

Acknowledgement

None.

Conflict of interest

The authors declare that they have no financial or non-financial competing interests.

References

1. Jeong HJ, Jung SM, An EA, et al. Development of the Safety Attitudes Questionnaire - Korean Version (SAQ-K) and Its Novel Analysis Methods for Safety Managers. *Biom Biostat Int J*. 2015;2(1):00020.
2. Jeong HJ, An EA, Kim SY, et al. Combinational Effects of Clinical Area and Healthcare Workers' Job Type on the Safety Culture in Hospitals. *Biom Biostat Int J*. 2015;2(2):00024.
3. Jeong HJ, Pham JC, Kim M, et al. Major Cultural-Compatibility Complex: Considerations on Cross-Cultural Dissemination of Patient Safety Programmes. *BMJ Qual Saf*. 2012;21(7):612–615.
4. Morello RT, Lowthian JA, Barker AL, et al. Strategies for Improving Patient Safety Culture in Hospitals: A Systematic Review. *BMJ Qual Saf*. 2013;22(1):11–18.
5. Etchegaray JM, Thomas EJ. Comparing Two Safety Culture Surveys: Safety Attitudes Questionnaire and Hospital Survey on Patient Safety. *BMJ Qual Saf*. 2012;21(6):490–498.
6. Colla J, Bracken A, Kinney L, et al. Measuring Patient Safety Climate: A Review of Surveys. *Qual Saf Health Care*. 2005;14(5):364–366.
7. Sexton JBI, Helmreich RL, Neilands TB, et al. The Safety Attitudes Questionnaire: Psychometric Properties, Benchmarking Data, and Emerging Research. *BMC Health Serv Res*. 2006;6(1):44.
8. Huang DT, Clermont G, Kong L, et al. Intensive Care Unit Safety Culture and Outcomes: A US Multicenter Study. *Int J Qual Health Care*. 2010;22(3):151–161.
9. Modak I, Sexton JB, Lux TR, et al. Measuring Safety Culture in the Ambulatory Setting: The Safety Attitudes Questionnaire-Ambulatory Version. *J Gen Intern Med*. 2007;22(1):1–5.
10. Patterson PD, Huang DT, Fairbanks RJ, The Emergency Medical Services Safety Attitudes Questionnaire. *Am J Med Qual*. 2010;25(2):109–115.
11. Makary MA, Sexton JB, Freischlag JA, et al. Operating Room Teamwork among Physicians and Nurses: Teamwork in the Eye of the Beholder. *J Am Coll Surg*. 2006;202(5):746–752.
12. Nordén-Hägg A, Sexton JB, Källemark-Sporrong S, et al. Assessing Safety Culture in Pharmacies: The Psychometric Validation of the Safety Attitudes Questionnaire (SAQ) in a National Sample of Community Pharmacies in Sweden. *BMC Clin Pharmacol*. 2010;10:8.

13. de Carvalho RE, Cassiani SH. Cross-cultural Adaptation of the Safety Attitudes Questionnaire-Short Form 2006 for Brazil. *Rev Lat Am Enfermagem*. 2012;20(3):575–582.
14. Zimmermann N, Küng K, Sereika SM, et al. Assessing the Safety Attitudes Questionnaire (SAQ), German Language Version in Swiss University Hospitals: A Validation Study. *BMC Health Serv Res*. 2013;13:347.
15. Deikås ET, Hofoss D. Psychometric Properties of the Norwegian Version of the Safety Attitudes Questionnaire (SAQ), Generic Version (short form 2006). *BMC Health Serv Res*. 2008;8:191.
16. Lee WC, Wung HY, Liao HH, et al. Hospital Safety Culture in Taiwan: A Nationwide Survey Using Chinese Version Safety Attitude Questionnaire. *BMC Health Serv Res*. 2010;10:234.
17. Raftopoulos V, Pavlakis A. Safety climate in 5 Intensive Care Units: A Nationwide Hospital Survey Using the Greek-Cypriot version of the Safety Attitudes Questionnaire. *J Crit Care*. 2013;28(1):51–61.
18. Sinnema H, Franx G, Volker D, et al. Randomised Controlled Trial of Tailored Interventions to Improve the Management of Anxiety and Depressive Disorders in Primary Care. *Implement sci*. 2011;6:75.
19. Hawkins RP, Kreuter M, Resnicow K, et al. Understanding Tailoring in Communicating about Health. *Health Educ Res*. 2008;23(3):454–466.
20. Vermunt JK, Magidson J. Latent Class Analysis. The sage encyclopedia of social sciences research methods. *Sage*. 2004;549–553.
21. Nylund KL, Asparouhov T, Muthén BO. Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural equation modeling*. 2007;14(4):535–569.
22. Timmel J, Kent PS, Holzmüller CG, et al. Impact of the Comprehensive Unit-based Safety Program (CUSP) on Safety Culture in a Surgical Inpatient Unit. *Jt Comm J Qual Patient Saf*. 2010;36(6):252–260.
23. Pronovost P, Needham D, Berenholtz S, et al. An Intervention to Decrease Catheter-Related Bloodstream Infections in the ICU. *N Engl J Med*. 2006;355(26):2725–2732.
24. das Neves ZC, Tipple AF, Silva e Souza AC, et al. Hand hygiene: The Impact of Incentive Strategies on Adherence among Healthcare Workers from a Newborn Intensive Care Unit. *Rev Lat Am Enfermagem*. 2006;14(4):546–552.
25. Asparouhov T, Muthén B. Using Mplus TECH11 and TECH14 to Test the Number of Latent Classes (Mplus Web Notes: No. 14). Los Angeles, CA: Mplus; 2012.
26. Jeong HJ, Kim M. A Practical Guide to Behavioral Theory-Driven Statistical Development of Quality and Safety Improvement Program in Health Care. *Biom Biostat Int J*. 2014;1(1):00001.
27. Fishbein M. The role of theory in HIV prevention. *AIDS Care*. 2000;12(3):273–278.
28. Broadbent J. If You Can't Measure It, How Can You Manage It? Management and Governance in Higher Educational Institutions. *Public Money & Management*. 2007;27(3):193–198.
29. Jain AK. Data clustering: 50 years beyond K-means. *Pattern recognition letters*. 2010;31(8):651–666.
30. Bezdek JC, Ehrlich R, Full W. FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*. 1984;10(2-3):191–203.
31. Strehl A, Ghosh J. Cluster ensembles-a knowledge reuse framework for combining multiple partitions. *The Journal of Machine Learning Research*. 2003;3:583–617.
32. Banerjee A, Krumpelman C, Ghosh J, et al. Model-based overlapping clustering. Paper presented at Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining: Chicago, USA; 2005. 532–537 p.