Healthcare professionals’ perception of hospital- and unit-Level managers’ contribution to improving safety

Abstract
The Safety Attitudes Questionnaire (SAQ) is a popular instrument to measure safety culture; however, its six domains have not been equally analyzed and used. Perception of management (PM), one of the underutilized domains, consists of two sets of the same items: one set for unit-level managers and the other for hospital-level managers. The SAQ was administered in a large tertiary hospital in Seoul, with 1,381 questionnaires being returned, including approximately 74% from women and 54% from nurses, which reflects Korea’s healthcare professional composition well. Respondents were asked to score management’s behavior in improving quality and safety. To calculate the score difference (unit managers’ score less hospital managers’ score), the generalized estimating equation was used to take the clinical unit’s clustering effects into account. In all subgroups and all PM items, the unit managers’ score was higher than that of hospital managers; most differences were statistically significant. On a scale of 0 to 100, the greatest difference was observed in the pharmacist group (14.5). In most cases, the score difference was around four to six. Various hypothetical explanations were offered. In Korea, many hospital managers are evaluated by hospitals’ financial performance and, quite often, monetary compensation for adverse events costs less than investing in improving safety, although there is no concrete evidence for this yet. In addition, hospital management’s term lasts around two to three years, which is too short of a time for a hospital’s reputation to drop in Korea’s healthcare environment. Consequently, hospital managers naturally put less emphasis on preventing medical errors. Another explanation arises from healthcare professionals’ fear of being reprimanded after giving a low score to unit managers. Although this survey was administered anonymously, respondents could have felt uncomfortable being critical of their unit managers, who will supervise respondents for a long time. These reasons are all conjecture. Further study is needed.

Introduction
Since the end of the 20th century, when the Institute of Medicine’s (IOM) seminal report “To Err is Human: Building a Safer Health System” came out, the issue of patient safety has never left the center stage of healthcare. “Saving lives by preventing medical errors” certainly deserved the top priority amongst all medical procedures. Several factors that influence safety have been considered, with safety culture being identified as a prerequisite for improving safety or a determinant of the level of safety that patients are assured.1 Yet what exactly is safety culture? Every quality and safety expert likely has their own definition and a list of domains that they think comprise it. The difficulty of coming to a consensus on a definition is a natural phenomenon, meaning safety culture is still a vague concept. Among the various definitions, we generally agree with Wachter’s version that “a culture of safety is where individuals feel comfortable drawing attention to potential hazards or actual failures without fear of censure from management” as well as his more casual version that it is “the way we do things around here.”2 Both touch on the core concepts.

Based on the old saying that, “without measure, there is no improvement,” researchers have come up with instruments measuring what they asserted was the safety culture and its components.3 The Safety Attitudes Questionnaire (SAQ), one of the most popular instruments that has been used globally, is an example. It has six domains (components) that might constate safety culture (Table 1). Each domain of the original SAQ developed by Bryan Sexton includes four to ten items.4 Among the domains, this study focused on one that has been less utilized—namely, perception of management (PM), which is defined as “healthcare professionals’ approval of managerial action.”

Using the PM items, we tried to determine how differently healthcare professionals perceive (or give scores to) the level of managerial action for improving safety provided by their unit-level managers and hospital-level managers. To address this issue, we need to understand the unique structure of PM. First, SAQ’s PM domain contains ten items, which is around twice that of the other domains. As shown in Table 3, these are two five-item sets where the items are identical between the sets. To illustrate, if an idea like “management support[s] my daily efforts” is presented to healthcare professionals, there must be two items: one about unit managers’ support and the other for hospital managers’ support. Thus, approval of managerial action at both levels of management can be measured simultaneously, and the results are comparable.

Because the original SAQ was not designed to be used this way, we had to check that this approach is scientifically sound, such as whether the original ten PM items deal with managerial support in one meaning or are asked in two sets of items whose contents differ. Thus, our “same items to different people approach” was tested first. In this way, the collected responses for the PM domain provided us with an opportunity to see how much different scores healthcare professionals give to their unit and hospital managers.
Methods

This study was conducted in South Korea; thus, we administered the Korean version of the SAQ (SAQ-K) that was developed and validated a few years ago. Due to the difficulty in translating the double negative sentence structure in English into Korean, two original SAQ items were dropped, but those were not PM items. In sum, for the PM domain, SAQ-K contains the same items as the original SAQ.²

Survey administration and data preparation

The SAQ-K was administered to healthcare workers at a tertiary hospital in Seoul, Korea, from October 2013 through November 2013. To prevent respondents from fearing being reprimanded for their responses, the survey was conducted in an anonymous manner.

SAQ-K responses were first measured on a 5-point Likert scale (1 = Disagree Strongly, 2 = Disagree Slightly, 3 = Neutral, 4 = Agree Slightly, 5 = Agree Strongly) and then converted into a 0 to 100 scale, with intervals of 25, as the original SAQ rubric suggested.

Instrument validation

When the SAQ-K was developed and validated, the correlated factor model was used, where all six SAQ domains were included and assumed to be related. However, in this study, we used only the PM response dataset. Therefore, we considered the PM domain of SAQ-K as an independent instrument. We then validated the ten PM items using internal consistency, Eigenvalue, and factor loadings. For in-depth investigation of the items’ characteristics, item information curves (IIC) were generated using item response theory (IRT) of the graded response model (GRM), as SAQ uses a Likert (ordinal) scale.⁶,⁷

Comparison of healthcare workers’ approval of unit and hospital managers

For all respondent groups, we calculated the PM score difference between unit and hospital management (mean unit management score less mean hospital management score: UM-HM). Then the hypothesis “UM-HM = 0” was tested. Because each unit naturally had clustering effects, we applied the generalized estimating equation (GEE) with an exchangeable correlation structure. The mean UM score was obtained by dividing the sum of responses of items PM1, 3, 5, 7, and 9 (Table 3). The mean HM score was estimated by dividing the sum of responses of PM2, 4, 6, 8, and 10. The responses were indicated on a scale of 0 to 100. All analyses were conducted using Stata 16.1 (StataCorp, College Station, Texas).

Results

Characteristics of respondents

In total, 1,381 questionnaires were returned. We excluded those with missing values in the PM domain or demographic information, including units. Ultimately, 1,139 questionnaires were analyzed. Table 2 summarizes the characteristics of the respondents. Most respondents were female (73.9%). In terms of job type, 53.5% were nurses while 33.2% were physicians. This split is understandable as the largest job type in Korean healthcare is nurses, most of whom are female. We had 132 supporting staff in the dataset as well as ten pharmacists and the same number of administrators in the survey. With regard to the work experience, healthcare professionals with 5–10 years of experience accounted for 25.4% of participants, followed by those with 3–4 years of experience (21.9%). People with fewer than 6 months or more than 21 years of experience accounted for 6.6% and 5.4%, respectively.

Table 2 Characteristics of Respondents

| Characteristics | N  | %   |
|-----------------|----|-----|
| Gender          |    |     |
| Male            | 297| 26.1|
| Female          | 842| 73.9|
| Work years      |    |     |
| Less than 6 months | 75 | 6.6 |
| 7 ~ 11 months   | 122| 10.7|
| 1 ~ 2 years     | 193| 16.9|
| 3 ~ 4 years     | 249| 21.9|
| 5 ~ 10 years    | 289| 25.4|
| 11 ~ 20 years   | 150| 13.2|
| More than 21 years | 61 | 5.4 |
| Job type        |    |     |
| Physician       | 378| 33.2|
| Nurse           | 609| 53.5|
| Pharmacist      | 10 | 0.9 |
| Supporting Staff| 132| 11.6|
| Administration  | 10 | 0.9 |
| Total           | 1,139| 100.0 |

To examine the distribution of healthcare professionals across a total of 72 different clinical units, we made a Stem-and-Leaf plot (Figure 1). In several units, fewer than ten healthcare professionals provided their services, and most units had ten to 30 healthcare workers. The largest unit had 53 staff members. As we were interested in the marginal difference between the PM scores of unit and hospital managers, we do not introduce the names of each unit.
Healthcare professionals’ perception of hospital- and unit-Level managers’ contribution to improving safety

The internal consistency among the ten PM items was favorable. Average interitem covariance was .44, and the scale reliability coefficient was 0.93. Although not shown here, Eigenvalues and other statistics strongly suggest that the ten items fell under a single concept. Table 3 shows that the factor loading for each item was satisfactory. Note that PM2 and PM7 showed relatively lower values; this information corresponds to the IIC in Figure 2, where the curves of the two items were exceptionally low. However, in terms of factor loadings, there were no issues (all standardized loadings were higher than .53). The meaning of the IIC will be discussed in a later section, but basically the items provide enough information for respondents in the range of -3 to +3 standard deviation, practically covering all respondents.

Table 3 Factor loadings of the PM items

| ID | Items | \( \beta \) | B | SE |
|----|-------|------------|---|----|
| PM1 | Unit management supports my daily efforts | .75 | .92 | .03 |
| PM2 | Hospital management supports my daily efforts | .53 | .64 | .03 |
| PM3 | Unit management doesn’t knowingly compromise patient safety | .78 | .93 | .03 |
| PM4 | Hospital management doesn’t knowingly compromise patient safety | .81 | .97 | .03 |
| PM5 | Unit management is doing a good job | .79 | .89 | .03 |
| PM6 | Hospital management is doing a good job | .79 | .97 | .03 |
| PM7 | Problem personnel are dealt constructively by our unit management | .62 | .72 | .03 |

PM domain and item characteristics

We calculated the score difference in healthcare professionals’ perception of management’s endeavor to improve safety between unit and hospital management (UM-HM). As shown in Table 4, all respondent groups gave a higher score to unit managers (UM-HM>0). In most cases, such a difference was statistically significant after controlling for unit-level clustering using GEE. The largest difference was observed among pharmacists (14.49). However, caution should be exercised because there was only one pharmacy in the hospital and all ten pharmacists were working in that pharmacy. A close relationship among the pharmacy manager and pharmacists might explain the high score that the pharmacists gave to the unit manager. Other than that, healthcare professionals with 5–10 years of work experience showed a difference of 6.49 with a large \( z (7.64) \). Note that this group is composed of people too young to take on a management position in Korean healthcare settings. Two groups showed statistically non-significant differences: respondents with fewer than 6 months of work experience and administration staff. The result from the administration group is understandable because there were only ten respondents across the hospital, which resulted in a high \( p \)-value. In general, besides a few exceptions, all groups gave around 5-point higher scores to unit managers.

Table 4 Difference in PM scores between unit and hospital management

| Characteristics of Respondents | UM-HM | SE | z | \( P>|z| \) | 95% CI |
|-------------------------------|-------|----|---|------------|-------|
| Gender                        |       |    |   |            |       |
| Male                          | 4.93  | 1.26 | 3.91 | 0.00       | 2.46  | 7.40 |
| Female                        | 5.10  | 0.48 | 10.69 | 0.00      | 4.17  | 6.04 |

Figure 1 Stem-and-Leaf plot of the number of respondents by clinical area.

Figure 2 Item information curves of the PM items.
Table 5 depicts expected differences calculated from the above GEE results. Therefore, the values should not be understood as the measured ones. The numbers, however, are worth reviewing as they help us understand the topography of marginal differences in scores between the two groups across units in the hospital. The pattern was quite similar to that observed in Table 4. The differences were widest in the pharmacist group for all work experience periods for both genders: from 11.93 in males with fewer than 6 months of work experience (smallest) to 16.43 in males with 5–10 years of work experience (largest). For other groups, the difference varied more widely than it did in Table 4.

### Table 5 Expected difference in PM score between unit and hospital management

| Characteristics of Respondents | UM-HM | SE | z   | P>|z| | 95% CI |
|-------------------------------|-------|----|-----|------|-------|
|                               | Low   | High |     |      |       |
| Experience                    |       |      |     |      |       |
| < 6 months                    | 1.42  | 1.21 | 1.17| 0.24 | -0.96 | 3.80 |
| 7 – 11 months                 | 4.38  | 1.05 | 4.18| 0.00 | 2.33  | 6.44 |
| 1 – 2 years                   | 4.75  | 0.98 | 4.83| 0.00 | 2.82  | 6.68 |
| 3 – 4 years                   | 4.59  | 0.76 | 6.07| 0.00 | 3.11  | 6.07 |
| 5 – 10 years                  | 6.49  | 0.85 | 7.64| 0.00 | 4.82  | 8.15 |
| 11 – 20 years                 | 4.87  | 1.16 | 4.21| 0.00 | 2.60  | 7.14 |
| > 21 years                    | 5.34  | 1.70 | 3.13| 0.00 | 2.00  | 8.68 |
| Job Type                      |       |      |     |      |       |
| Physician                     | 4.21  | 0.82 | 5.12| 0.00 | 2.60  | 5.82 |
| Nurse                         | 5.22  | 0.49 | 10.76| 0.00 | 4.27  | 6.17 |
| Pharmacist                    | 14.49 | 5.01 | 2.89| 0.00 | 4.66  | 24.32 |
| Supporting Staff              | 5.69  | 1.43 | 3.97| 0.00 | 2.88  | 8.50 |
| Administration                | 2.51  | 1.93 | 1.30| 0.19 | -1.27 | 6.30 |

Note Supporting means supporting staff; Admin. means people in charge of administrative staff; grey cells indicate statistically significant differences at alpha level=.05

### Discussion

This study began as a part of a series to search for new ways to utilize already collected SAQ data. Such methodological exploration led us to the idea that we could measure the difference in perceived activity of managers in improving patient safety between unit and hospital managers. At first, this study appeared to be very simple—namely, just a comparison of scores. However, while pursuing the goal in a scientifically sound way, we had to take several unexpected steps. First, we had to demonstrate that the PM domain is unidimensional. In other words, it measures only one vector, healthcare professionals’ approval of management, instead of the two different vectors, one...
for unit management and the other for hospital management. For example, if the items for hospital-level management were targeting relatively bigger tasks, such as resource allocation, among the various activities in a hospital and items for unit management were measuring very localized issues, such as dealing with disgruntled employees, we could not compare the scores of unit and hospital management; they would be apples and oranges, with any score difference potentially coming from the difference in contents of items for each group. Thus, the unidimensionality should be guaranteed to ensure the validity of this study.

As previously mentioned, by disposing of the other domains and pretending we have a ten-item PM survey questionnaire, we could test whether all the items pointed in the same direction. Of course, the test was both quantitative and qualitative. How well the instrument worked was checked. Internal consistency, although heavily related with the number of items, was satisfactory. Factor loadings were all good. However, those of PM2 “Hospital management supports my daily efforts” and PM7 “Problem personnel are dealt with constructively by our unit management” were lower than the others. Such a phenomenon was easily recognized by using IIC built on the IRT GRM framework. Indeed, they did not add a significant amount of information to the instrument. If the contents of these two items with lower loadings were the same only when targeting different groups, we would have removed them, leaving the instrument with eight items. Four items for each group seemed still valid considering the other SAQ domains. However, as we have different items with lower loadings, removing the two items would mean getting rid of a total of four of the ten PM items, leaving only a three-item questionnaire for each of the unit and hospital managers. Three was too small and, thus, we kept all ten PM items based on the fact that those items’ factor loadings are still high (>0.5) enough, albeit comparatively lower than others. At this point, a -equivalence issue arises—that is, if factor loadings vary widely, traditional mean scores may not reflect the real trait level of a respondent precisely. Thus, a weighted value like a factor score should work better. However, this approach was not available because PM2 and PM7 asked different contents of different groups. We had to stick to the simple mean score of the responses.

We then tapped into the score difference between the two groups. Although not described earlier, unit-level managers’ score for each item was higher than that of hospital managers’ with statistical significance. The differences were similar across items. Therefore, how the current overall domain score difference looked is understandable. We try to explain such a difference in various ways.

First, the lower score of hospital management might have come from Korea’s unique situation, where the term of hospital management, including the president, is only around two to three years. In addition, in many cases, they are evaluated by the hospital’s financial growth. The problem might originate here: Preventing medical errors may require more costs than just compensating the adverse events that already occurred and clearing them legally. Furthermore, generally losing a reputation at the hospital level takes a relatively long time, this phenomenon is expected to be more prominent with unit-level management. From the authors’ previous experience, regardless of data-collection methodology (i.e., paper or computer-based), such a fear does exist. Each of the above scenarios can explain the score difference to some degree, but we should admit that further study is certainly needed.

Exploring a new way of using the PM score from the original SAQ interpretation was the goal of this study as this domain has not drawn much attention. This aim meant that we desired to maximize the use of the SAQ. To our knowledge, teamwork climate and safety climate have thus far been the primary areas of interest in SAQ use, while few studies have put enough emphasis on the remaining domains. However, this is understandable and even inevitable. In many medical error cases, communication breakdown among healthcare professionals has been the most important cause or a contributing factor of the occurrence. To resolve such issues, encouraging barrierless communication should be guaranteed, underpinning safer care. This indicates the modus operandi of teamwork and safety climate: They make possible more active communication among healthcare professionals without any blame or barrier. All in all, teamwork certainly deserves the head table among various potential constituents of safety culture.

This does not, however, justify our ignorance of the other SAQ domains. If the other domains are really less important, they should not have even been asked from the beginning. Healthcare professionals’ time should be spent on addressing those domains. As shown in Table 3, the five PM item sets for each group put more emphasis on asking about unit-level activity. Thus, hospital-level management activities are not subject to be captured with the items. For example, it is a stretch to think that a senior manager is dealing with a problem person in a unit or small adverse events and the resulting risks in a unit. Although any hospital managers may deal with such unit-level issues, such actions are taken through the unit managers, not directly by hospital management. Such activities are not easily caught at the frontline in healthcare professionals’ eyes.

Yet there could be a completely different scenario: There was no actual difference in managers’ endeavor, but they were only scored differently. To collect candid responses, the SAQ was administered anonymously. However, as healthcare professionals, the respondents might fear being reprimanded for giving lower scores to their direct supervisors, with whom they will be working for a long time. This phenomenon is expected to be more prominent with unit-level management. From the authors’ previous experience, regardless of data-collection methodology (i.e., paper or computer-based), such a fear does exist. Each of the above scenarios can explain the score difference to some degree, but we should admit that further study is certainly needed.

Also, on the other hand, unit managers are specialists of a certain area and are usually trained in the very unit and tend to retire from the same place. In many cases, unit managers in Korea work in one or, at most, a few units throughout their professional life. Thus, unlike hospital management, unit managers’ reputations are destined to be affected by units’ performance in quality and safety. Few unit managers risk the reputation of the unit under their supervision. This train of thought might explain why unit managers work more aggressively to keep healthcare safer.

Another aspect worth considering lies in the contents of the PM items from the beginning. As shown in Table 3, the five PM item sets for each group put more emphasis on asking about unit-level activity. However, this is understandable and even inevitable. In many medical error cases, communication breakdown among healthcare professionals has been the most important cause or a contributing factor of the occurrence. To resolve such issues, encouraging barrierless communication should be guaranteed, underpinning safer care. This indicates the modus operandi of teamwork and safety climate: They make possible more active communication among healthcare professionals without any blame or barrier. All in all, teamwork certainly deserves the head table among various potential constituents of safety culture.
Conflicts of interest
None.

References
1. 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.

2. Wachter R. Understanding patient safety. 2012, McGraw Hill Professional.

3. Lee GS, Park MJ, Na HR, et al. A Strategy for Administration and Application of a Patient Safety Culture Survey. J of Quality Improvement in Health Care. 2015.21(1):80–95.

4. Sexton JB, Helmrreich RL, Neilands TB, et al. The Safety Attitudes Questionnaire: psychometric properties, benchmarking data, and emerging research. BMC Health Serv Res. 2006.6(1):44.

5. 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):8–17.

6. Jeong HJ, Lee WC. Item response theory–based evaluation of psychometric properties of the safety attitudes questionnaire—korean version (saq–k). Biom Biostat Int J. 2016;3(5):174–187.

7. Kim S–H, AS Cohen. Detection of differential item functioning under the graded response model with the likelihood ratio test. Applied Psychological Measurement. 1998;22(4):345–355.

8. Acocq AC. Discovering structural equation modeling using Stata. Stata Press Books. 2013, College Station, TX. Stata Press Books.

9. Jeong HJ, Jung SM, An EA, 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):45–51.

10. Lee G, Park MJ, Na HR, et al. Are healthcare workers trained to be impervious to stress? Biom Biostat Int J. 2015;2(2):78–79.

11. Jeong HJ, Kim M, An EA, et al. A strategy to develop tailored patient safety culture improvement programs with latent class analysis method. Biom Biostat Int J. 2015;2(2):71–76.

12. Jeong HJ, Lee WC. The pure and the overarching: an application of bifactor model to safety attitudes questionnaire. Biom Biostat Int J. 2016;4(6):222–232.

13. Jeong HJ, Park M, Kim CH, et al. Saving lives by saving time: association between measurement scale and time to complete safety attitudes questionnaire. Biom Biostat Int J. 2016;4(5):177–180.

14. Jeong HJ, Lee WC. Does differential item functioning occur across respondents’ characteristics in safety attitudes questionnaire? Biom Biostat Int J. 2016;4(3):103–111.

15. Lee WC, Han SH, Jeong HJ, Ensuring Taiwan patient safety culture survey is free from differential item functioning. Int J Health Plann Manage, 2019.

16. Jeong HJ, Lee WC, Liao HH, et al. The hospital patient safety culture survey: reform of analysis and visualization methods. Int J Environ Res Public Health. 2019;16(19):3624.

17. Jeong HJ, Xiao HH, Han SH, et al. An application of item response theory to scoring patient safety culture survey data. A practical approach to develop a parsimonious survey questionnaire–taiwanese patient safety culture survey as an example. Biom Biostat Int J. 2017;6(3):350–355.

18. Xiao HH, Lee WC, You YL, et al. Reducing diagnostic errors through effective communication: harnessing the power of information technology. J Gen Intern Med. 2008;23(4):489–494.

19. Singh H, Naik AD, Rao R, et al. Reducing diagnostic errors through effective communication: harnessing the power of information technology. J Gen Intern Med. 2008;23(4):489–494.

20. Khairat S, Gong Y. Understanding effective clinical communication in medical errors. Stud Health Technol Inform. 2010. 160(Pt 1):704–708.

21. Nance JJ. Why hospitals should fly: The ultimate flight plan to patient safety and quality care. 2008;Bozeman,MT: Second River Healthcare Press.

22. O’Daniel M, AH Rosenstein. Professional communication and team collaboration, in Patient safety and quality: An evidence–based handbook for nurses. 2008;Agency for Healthcare Research and Quality (US).

23. 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.

24. 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.