Validity of the patient safety climate items from the RN4CAST Survey compared with the SAQ

Internal consistency, concurrent validity and inter-rater reliability

Kristina Kymre

Master thesis
Department of Health Management and Health Economics
The Faculty of Medicine
UNIVERSITETET I OSLO
May 2012
Validity of the patient safety climate items from the RN4CAST survey compared with the SAQ

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo
Abstract

Background: The Norwegian Knowledge Centre for the Health Services participated in the multinational nursing survey RN4CAST as the Norwegian representatives. The items relating to nurses’ perceptions of patient safety climate have not been assessed in the Norwegian setting, which is needed if these items are to be used with certainty regarding their validity and reliability.

Objectives: The main objective of this thesis is to assess the concurrent validity and internal consistency of items relating to perceptions of patient safety climate in the Norwegian part of the multinational nursing survey RN4CAST. A secondary objective is to identify the main challenges associated with identifying causal relationships between safety climate and safety outcomes, and discuss these in relation to the survey undertaken in this study.

Methods: Nurses from two hospital wards completed both the SAQ and the items relating to patient safety climate from the RN4CAST survey. Two scaled indexes were created from the RN4CAST items, and three analyses were done: 1) Pearson’s correlation coefficient between the SAQ and the RN4CAST; 2) Chronbach’s Alpha for internal consistency within the RN4CAST items; and 3) Bland-Altman analysis for assessment of agreement.

Results: The assessment of the two scales created demonstrated high internal consistency (Chronbach’s Alpha of .82 and .80, respectively), moderate inter-item correlations, and good concurrent validity as measured by Pearson’s correlation coefficient ($r = .73$ and $.78$, respectively). The results from the Bland-Altman analysis demonstrated reasonably good agreement between the two measures of patient safety climate.

Conclusion: The result of the analyses demonstrates that the two scales are internally consistent, there is an association between the scales from RN4CAST and SAQ safety climate dimension, and that the level of agreement between the two methods is good. The two scales are both found to be validated measures of patient safety climate.
Acknowledgements

I would like to use this opportunity to thank the ones contributing to the realisation of this thesis. It has been a challenging process, as the intensions and objectives have been constantly changing. Sometimes it felt like the thesis had its own life, purposes and goals. Due to the never-ending willingness to share his vast knowledge and valuable time, providing guidance and comments, Professor John-Arne Røttingen has made the completion of this thesis possible. My sincere thanks go to him. Further, I would like to thank Ingeborg Strømseng Sjetne at the Norwegian Knowledge Centre for help and guidance towards a topic for my thesis. Christine Tvedt deserves many thanks for her invaluable help with data collection and processing, and for comments and assistance throughout the process.

To all my friends and fellow students at HLED: a heartfelt thank you for making these years here as joyful and awarding as they have turned out to be. And finally, to my Stein Arne - thank you for enduring my recent thesis-related meltdowns and general absentmindedness. Now, let’s go to Iceland!
# Table of contents

1. Introduction ........................................................................................................................................ 1
   1.1 Objections of the thesis ............................................................................................................ 2
   1.2 Thesis’ structure .................................................................................................................... 3

2. Theoretical framework ................................................................................................................... 4
   2.1 Organisational culture .......................................................................................................... 4
      2.1.1 Three layers or culture .................................................................................................. 4
   2.2 Organizational climate ......................................................................................................... 6
   2.3 Safety culture and climate ................................................................................................... 7
      2.3.1 Safety culture and climate in health care ...................................................................... 7
   2.4 Patient safety ........................................................................................................................ 9
      2.4.1 Patient safety and its relation to quality ...................................................................... 9
   2.5 Can safety climate predict patient safety outcomes? .............................................................. 12
      2.5.1 The level of aggregation ............................................................................................. 14
   2.6 Measuring safety culture/climate ............................................................................................ 14
      2.6.1 The Safety Attitudes Questionnaire (generic version) .................................................. 15
      2.6.2 The RN4CAST survey .................................................................................................. 17
      2.6.3 Hospital Survey on Patient Safety Culture (HSOPSC) ................................................ 18
      2.7 Safety Attitudes Questionnaire and Hospital Survey on Patient Safety Culture compared .......................................................................................................................... 20

3. Design .............................................................................................................................................. 22
   3.1 Ethical approval, consent and confidentiality .......................................................................... 22
   3.2 The survey ............................................................................................................................... 23
      3.2.1 The questionnaire .......................................................................................................... 23
   3.3 Data collection ......................................................................................................................... 25
   3.4 Sample .................................................................................................................................... 26
   3.5 Response rate .......................................................................................................................... 27
   3.6 Common method variance ....................................................................................................... 27

4. Analytical approach ....................................................................................................................... 29
   4.1 Validation of questionnaires ................................................................................................... 30
      4.1.1 Likert-type items and level of measurement .................................................................. 30
      4.1.2 Reliability ...................................................................................................................... 30
4.1.3 Validity .......................................................... 32
4.1.4 Assessment of agreement .......................................................... 33
4.2 Exploratory factor analysis .......................................................... 34
  4.2.1 Decisions in factor analysis .......................................................... 35
5 Results ..................................................................................... 38
  5.1 Sample descriptives .......................................................... 38
    5.1.1 Response rate .......................................................... 38
    5.1.2 Missing values .......................................................... 39
  5.2 Results SAQ ........................................................................ 40
  5.3 Principal component analysis .................................................. 41
  5.4 Reliability analysis .......................................................... 43
    5.4.1 Reliability of the composite scales from RN4CAST .......................................................... 43
  5.5 Concurrent validity assessment .................................................. 47
    5.5.1 Pearson's product-moment correlation coefficient .......................................................... 48
    5.5.2 Intraclass correlation .......................................................... 48
    5.5.3 Assessment of agreement .......................................................... 49
6 Discussion .................................................................................. 51
  6.1 General conclusions .......................................................... 51
  6.2 Methods and instruments .......................................................... 51
    6.2.1 Measurement caveats .......................................................... 52
    6.2.2 Outcome variables .......................................................... 54
  6.3 Safety climate and its relation to quality ........................................... 55
    6.3.1 Lack of explicit theoretical framework .......................................................... 56
    6.3.2 Direct causal relationship, circular causation or just co-variation? .......................................................... 57
  6.4 Strengths .......................................................... 57
  6.5 Limitations .......................................................... 58
7 Concluding remarks ........................................................................ 59
References .................................................................................. 60
Appendices .................................................................................. 66
List of Tables

Table 1: Factor definitions and item description SAQ (generic version) ........................................ 17
Table 2: Factor dimensions and item descriptions of HSOPSC .................................................. 19
Table 3: Items included in the RN safety composite scales, and the SAQ safety climate ........... 25
Table 4: Strength of linear relationship ..................................................................................... 32
Table 5: Types of validity .......................................................................................................... 33
Table 6: Descriptive statistics of the sample .............................................................................. 38
Table 7: Number and percent missing items ............................................................................. 39
Table 8: Number of respondents and results for each factor .................................................... 41
Table 9: Eigenvalues >.1 .......................................................................................................... 42
Table 10: Monte Carlo Parallel Analysis .................................................................................... 43
Table 11: Inter-item correlations RNsafety_I ........................................................................... 45
Table 12: Item-total correlation and Chronbach’s Alpha if item deleted ................................. 45
Table 13: Inter-item correlations RNsafety_II .......................................................................... 46
Table 14: Item-total correlation and Chronbach’s if item deleted ............................................. 47
Table 15: Spearman’s rho and Pearson’s correlation coefficients ............................................. 48
List of figures

Figure 1 Percentage positive response to each dimension ...............................................40
Figure 2 Scree plot ........................................................................................................42
Figure 3 Bland-Altman plot SAQ safety climate and Rnsafety_I ..................................49
Figure 4 Bland-Altman plot SAQ safety climate and Rnsafety_II ..................................49
Figure 5 Normal p-p plot of regression standardized residual ......................................50
# Acronyms and abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAQ</td>
<td>Safety Attitudes Questionnaire</td>
</tr>
<tr>
<td>HSOPSC</td>
<td>Hospital Survey on Patient Safety Culture</td>
</tr>
<tr>
<td>RN4CAST</td>
<td>Multinational survey on nursing</td>
</tr>
<tr>
<td>DK</td>
<td>Don’t Know (responses)</td>
</tr>
<tr>
<td>EUNetPAS</td>
<td>The European Network for Patient Safety</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>FA</td>
<td>Factor Analysis</td>
</tr>
</tbody>
</table>
1 Introduction

Since the American Institute of Medicine published the report “to Err Is Human”, where they estimated the number of adverse events and resulting deaths in the American health care service (Kohn, Corrigan and Donaldson, 2000), the interest for patient safety has risen. In Norway, it have been estimated that 16 per cent of all hospitalized patients were subject for damage acquired the hospital in 2010, and of these, 4700 patients died as a result of the damage (pasientsikkerhetskampanjen.no). At the same time, due to the demographic changes as ageing of the population, the Norwegian Statistical Agency have provided a conservative estimate of a shortage of 14,000 nurses in 2030 (Texmon and Stølen, 2009). Hence, there is a great need to increase the knowledge on how to plan and organize the health services in a way that is at the same time both effective, and ensures patient safety.

Several ongoing projects in Norway attempts to investigate the present status of patient safety, and increase the knowledge on this area. The Norwegian part of the multinational RN4CAST survey aim at developing knowledge about the relationship between hospitals’ organisational variables, the management of nursing resources, and patient safety in the Norwegian context (project plan Personnel and patient safety, Kunnskapssenteret 2009). Hence, it can be an important contributor to the development of knowledge about patient safety aspects in Norway, and eventually to contribute to a safer health care service.

The National Patient Safety Campaign was launched in January 2011, aiming at reducing patient damages and improve patient safety, both in primary and specialized care. The three main aims for this campaign are to reduce patient damages, establish lasting structures for patient safety, and to improve the patient safety culture in the health services. Hence, patient safety culture is considered a key aspect of safe care. Surveys demonstrate that the patient safety culture could affect the treatment of patients, and that it can vary considerably between units. Thus, it is considered important to measure how health care staff perceives the safety culture in their unit. The results from the patient safety climate assessment undertaken as part of the Safety Campaign will guide targeted interventions for improvement of patient safety culture and increase the patient safety in the Norwegian health care sector (pasientsikkerhetskampanjen.no).
1.1 Objections of the thesis

As the RN4CAST survey is part of a large, multinational study, not all aspects have been validated for the Norwegian setting. This thesis will contribute to this exactly, as it is an assessment of the concurrent validity, inter-item reliability and inter-rater reliability of the items from RN4CAST that relates to patient safety. It will allow the researchers working on this study to use these items as measures of nurses’ patient safety climate perceptions with a greater level of confidence. This work will contribute to fill the knowledge gap related to these items in the Norwegian branch of the RN4CAST survey. I will in the following investigate and discuss my main research question; “Can items from question B6 in the RN4CAST survey be used as measures of nurses’ perceptions of patient safety climate?”

The research question will be assessed using empirical analyses of survey data from a Norwegian hospital, containing the relevant items from the RN4CAST survey, as well as an already validated instrument for assessing health care workers’ perceptions of patient safety climate, the Norwegian version of the Safety Attitudes Questionnaire (Generic version, short form 2006).

In the process of gathering knowledge about research patient safety culture and climate measurements, it became increasingly clear that the notion of patient safety culture is not uncontested. For example, it is difficult to see what term best describes the phenomenon as there is no clear consensus in the use of concepts (The Health Foundation, 2011a). In many cases, safety climate and safety culture are used exchangeable, without discussion and definition of the underlying meaning. Further, the causal relationship found between safety climate and patient outcomes in single studies have proven difficult to establish in meta-studies (Ibid.). The Nordic Council of Ministers (2010) states that it is imperative that patient safety culture should be linked to other measures of patient safety, such as complications of treatment, mortality, and observable safety behaviour such as safe medication. It was thus perceived important to assess this issue more in relation to the validation of the patient safety climate items.
1.2 Thesis’ structure

The thesis is organized as follows. Section 1 has introduced the thesis. Second, the theoretical background will be presented, with emphasis on terms relevant to safety culture and climate, patient safety, quality of health care, and the measurement of patient safety climate, Third, the design and methods for data collection will be presented. Section 4 presents the analytical approach, and section 5 presents the results. Section 6 discusses the statistical findings as well as the findings from the literature. Finally, section 7 concludes the thesis.
2 Theoretical framework

In this chapter, the terms organisational culture and climate, patient safety culture and climate, and patient safety will be presented using relevant research and literature. Special attention will be given to the measurement of patient safety climate and its relationship to quality of health care services.

2.1 Organisational culture

According to Edgar Schein (2010), an organisational culture is defined as

“a pattern of shared basic assumptions learned by the group as it solved its problems of external adaptation and internal integration, which has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think and feel in relation to those problems” (p. 18).

The research and literature on organizational culture tends to fall within one of two main schools. If culture is regarded as something an organization has, it is often assumed that culture is something leaders can (and should) change and manipulate in order to achieve certain goals. However, if one regards culture as something an organization is, then culture is often conceptualized as something that evolves from the actions taken by the group members, which is grounded in unconscious values and therefore difficult to assess or change (Davies, Nutley and Mannion (2000). Schein’s definition addresses both these perspectives, as culture is both viewed as an expression of style, strategy documents, physical environment etc, as well as shared, unexpressed values and beliefs (which could be extremely difficult to change). Other, broader definitions emphasize that culture is a global phenomenon that encompasses the norms, values and beliefs and basic assumptions of an entire organisation (Blegen, Pepper and Ross, 2005).

2.1.1 Three layers or culture

Following Schein’s framework, any given culture is consisting of three layers. The layers refer to the degree of visibility to observers. The first cultural level is the level of artefacts, consisting of the observable phenomena that occur when one encounters a new, unfamiliar group. Artefacts are, among others, made up by the visible products of the group, i.e. the
physical surroundings, language, technology, the group’s style (clothing, behavioural manners), its myths and stories told about the organization, and last its expressed list of values and its observable rituals and ceremonies. Thus, artefacts are the organization’s observable and feelable structures and processes (Schein, 2010).

Among these artefacts is the group’s “climate”. There is extensive debate about the relationship between climate and culture, and whether these in fact are equivalent to each other (Zohar, Livne, Tenne-Gazit, Admi, & Donchin, 2007). According to Schein, the climate is a manifestation of the group’s culture, as it is better thought of as a product of the underlying basic assumptions (2010).

The second level consists of espoused beliefs and values, created via learning and problem solving by the group and its leader. If the solution to certain problems turns out to be successful, these solutions or values will gradually be transformed into shared values or beliefs, and finally into shared assumptions. One important aspect of this level is that the espoused beliefs and values continue to be conscious and are explicitly articulated as they are functioning as moral or normative “guidelines” in how the group best should solve key issues, and best to train new members of the group how to behave (Schein 2010). Further, one can observe deviations from the desired behaviour in the observable behaviour, if the beliefs and values providing comfort and meaning to the group are not congruent with effective performance.

The third and final level is the level of basic underlying assumptions. If the suggested solution to a problem continues to work repeatedly, what was first a hypothesis ends up being treated as reality. The degree of consensus within the group depends on previous success in implementing certain beliefs and values. Culture defined as a set of basic assumptions gives definitions on what to pay attention to, the meaning of things, the way to react emotionally towards events, as well as appropriate actions for situations. These underlying assumptions tend to be non-confrontable and non-debatable, and could thus be difficult to change. Thus, the essence of a given culture tends to lie in the pattern of underlying basic assumptions (ibid.).

In order to investigate and measure organizational culture, one would have to assess the group’s culture at all three levels. The different levels will require different research methods to be used, as it could be difficult, for example, to assess unexpressed and unarticulated
beliefs and values via a structured survey tool. To do this, qualitative and observational methods would be required (Ashkanasy, Broadfoot and Falkus (2000). The most common method for assessing organizational culture is using survey tools. When questionnaires are used to measure safety culture, the different items measured are organized into scales that reflect different dimensions or traits relevant to safety culture. These include among other safety climate, teamwork climate and perceptions of management (Deilkås, 2010). As have been noted above, it is the climate which is most accessible to observation and measurement. Hence, survey measurement of culture can be contested – in fact, it might not be measuring culture, but climate. However, in many of these instruments, the term climate is replaced with “culture” (Olsen, 2008)

2.2 Organizational climate

In Guldenmund’s review of research on (safety) culture and climate (2000), attempts to identify systematic differences between organizational culture and climate were made. In his review, special attention was given to the presence of an eventual underlying theoretical model, as this should form the basis for any scientific enterprise. These two terms have been used in research with little attention to the concepts’ theoretical and practical underlying meaning. His findings demonstrated that historically, the term organizational climate was referring to a global, integrating concept, which underlie most organizational events and processes (similar to Schein’s definition of culture). Further, the term organizational climate has come to mean the more open manifestation of the organization’s underlying culture – climate is an expression of the culture (Guldenmund, 2000).

According to Gershon, Stone, Bakken and Larson (2004), aspects of organisational climates are easier to measure due to their tangible nature. This is in accordance with Schein’s definition of culture, as climate is an observable expression of the underlying culture. Hence, to measure climate has been regarded as a way of measuring culture, as a given group’s climate could be perceived as a manifestation of the group’s culture.
2.3 Safety culture and climate

According to Zohar et al (2007), safety climate can be defined as shared employee perceptions of how their unit and the organisation at large prioritises safety, especially in situations where safety competes with other performance dimensions (e.g. care speed or quality of care). As will be described in following parts of this chapter, there is no shared or commonly used term for neither patient safety culture nor climate (National Patient Agency (2004). Guldenmund (2010,p. 1466), claims that safety culture “..has become a term used. to explain everything related to safety failures that cannot be explained in any other way (Guldenmund, 2010) . Further, there has been a continuing debate about the fundamental dimensions of which constitutes safety climate (Flin, 2007) . Hence, the generality of these constructs have not been established.

2.3.1 Safety culture and climate in health care

The efforts to link culture and performance can be linked back to the Hawthorne studies, aspects of the social dimensions in groups of workers could be manipulated in order to affect employee efforts and performance (Davies, Mannion, Jacobs, Powell, & Marshall, 2007). In the 60s and early 70s, this topic received little attention. This changed when explanations for the worldwide success of Japanese factories were sought, and safety culture was identified as a major contributor for this success (Wilderom, Glunk and Maslowski (2000).

The interest for safety culture and climate research has risen since the 1980s. One of the focusing events was the Chernobyl accident in 1986, where a lack of safety culture was identified as one of the causes for the accident (Mearns and Flin (1999). Since, a plethora of researchers have tried to connect safety culture and climate with the organization’s performance and quality (Wilderom et al, 2000). The same challenges as in general organizational climate and culture research are applicable to research on safety culture and climate. Safety culture and safety climate have been used synonymously and interchangeably in many of the studies, since the concepts of climate and culture have not been operationalised or defined in a proper way, and the construct validity have not been thoroughly established (Olsen, 2010), Mearns and Flin 1999), which in turn have resulted in questionable methodological quality of some studies (Wilderom et al, 2000). A vast number of different definitions, and methods for measuring, safety culture and climate, have resulted in a mass of
research lacking a coherent theoretical framework (Wakefield, McLaws, Whitby, & Patton, 2010).

According to the UK Health and Safety Commission (1993), safety culture can be defined as

The safety culture of an organization is the product of the individual and group values, attitudes, competencies and patterns of behaviours that determine the commitment to, and the style and proficiency of, an organization’s health and safety programmes. 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 (cited from Vincent, 2010 p.273).

This definition is not directly addressing patient safety culture, which is defined by The European Network for Patient Safety (EUNetPAS) as

“An integrated pattern of individual and organisational behaviours, based upon shared beliefs and values that continuously seeks to minimise patient harm, which may result from the process of care delivery” (Christensen and Bartels, 2010, p.4)

Common definitions of the term patient safety climate seems to emphasize shared employee perceptions of how their unit and organisation prioritises patient safety aspects (see for example Zohar et al (2007), or The Health Foundation (2011b). But - there seems to be a vast array of different definitions, also in the health care sector (Halligan and Zecevic (2011). The many definitions of safety culture could imply that first, the concept of safety culture is not fully understood or agreed, and second, that there is a number of important aspects to the culture of safety. According to Vincent (2010), these aspects include both how the organization reacts to errors after they have occurred, the practice environment in which the errors are made (if there is a culture of blame, guilt etc that would impede learning from errors as well as effective actions to prevent future errors), if errors are anticipated beforehand instead of rather reacted to after they have occurred, and if problems are denied instead of acknowledged.
2.4 Patient safety

Safety is considered a core dimension of quality. This is, however, a new way of thinking about quality and safety. For example, Maxwell (1984) identified six core dimensions of quality, of which safety was not mentioned. In Norway, safety is included in the definition of quality, as one of the six core elements that constitutes health care services of high quality (Helsedirektoratet, 2010).

WHO defines patient safety as “freedom for a patient from unnecessary harm or potential harm associated with healthcare” (ICPS, 2004)

The by now well-known report “To Err is Human” contributed to a large degree to the agenda setting of patient safety in health care. This report from the American Institute of Medicine estimated that between 44000 and 98000 patients died every year in the US, resulting from adverse events and healthcare errors (Kohn, Corrigan and Donaldson, 2000). Following this report were a number of studies from both the US, UK and other countries, demonstrating that between 3 and 17 % of hospital admissions resulted in an adverse event (Flin, 2007). In Norway there is a lack of systematic reporting and measurement of adverse events, hence the estimates for the numbers of deaths and injuries have been based on international studies. Peder F. Hjort estimated in the book “Uheldige hendelser i helsetjenesten” that the corresponding figure for Norway in 2005, is 4000 deaths (2007).

Initially, the solution to these problems focused on administrative and technical aspects of the care process, such as the development of patient safety units in hospitals, but since these measures did not result in the expected changes, a call for systemic and organisation-level changes were raised (Zohar et al., 2007). There is now a widespread focus on patient safety culture, and attempts to link patient safety culture to safety outcomes are made.

2.4.1 Patient safety and its relation to quality

Patient safety can be viewed as an aspect of quality, or it can be viewed as something resulting from work to improve quality. It could also be regarded as a prerequisite for quality (Kunnskapssenteret, 2010). The relationship between quality and patient safety is not well defined or explored, and there is thus a vast array of definitions. Many describe this relationship as different endpoints in a continuum. According to Brown et al., (2008), the extent that an event will be viewed as a quality or safety issue will depend on the strength of
causation between event and the result, and the immediacy of the resulting harm. Vincent (2010) argues, based on Brown et al’s article, that events that clearly cause harm and have a clear relationship to specific errors or events are more likely to be labelled as patient safety issues. According to Brown et al (2008), as a generic rule for distinguishing safety and quality, very rare errors with high immediacy and high causality generally has been concerned with safety, while more frequent events with corresponding low immediacy and low causality has been concerned with quality. However, the authors do not follow this distinction, as there is no clear point as for when quality issues turns into safety issues. Further, they do not believe that safety can be distinguished from quality merely based on the causal link between error and outcome.

The term “risk management” is established in industrial safety research. In patient safety, this term attempts to establish a relationship between the understanding of individual’s errors and harms that could cause damage to the patient and a wider focus that includes employees, visitors, and the society as a whole. In this sense, the term enables patient safety to be a part of the wider healthcare system and its safety (Kunnskapssenteret, 2010). Patient safety is thus not primarily regarded to be the result of individuals’ actions, but is a result from multiple actors, interactions with different layers in the provision of care, and of the institutional and organizational context of which patient care is provided

According to Vincent (2010, p.31), patient safety can be defined as

“the avoidance, prevention and amelioration of adverse outcomes or injuries stemming from the process of health care”

This definition only addresses adverse outcomes or injuries stemming from the process of health care, and is thus separated from the wider quality perspective, which includes process and structure, as well as near-misses, errors and adverse events (Vincent, 2010). It is now argued that patient safety is also dependent on safe structures of which care is produced and delivered. However, whether patient safety is best assessed using structural, process or outcome measures is not agreed upon (ibid.). If structure and process measures are to be credible measures of patient safety, the link between these measures and the outcomes has to be established (Brook, McGlynn and Cleary, 1996, Pronovost, Miller and Wachter, 2006). A review prepared by Hearld, Alexander, Fraser and Jiang (2008) further demonstrated that the link between structure and outcome variables is not well established, as more than 45 % of the
studies included in their review had non-significant results between the organizational characteristics and the outcome variables. Of the remaining studies, a mix between positive and negative effects on quality was found (35% positive and 24% negative impact on quality). The authors of this review state that there is a need for theories and more research to incorporate all the three components structure, process and outcomes, for example by identifying situations where process variables moderate the relationship between structure and outcome. They found that most studies were focusing on linking structure and outcome, but the results were in many cases non-significant. However, the studies assessing the relationship between structure and process variables, and process and outcome variables, had more robust and consistent positive significant findings (Ibid.).

The use of outcome measures as a reflection of quality and safety is contested (Salzer, Nixon, Schut, Karver and Bickman, 1997). Patient outcomes may be determined by a combination of the patient’s underlying condition and the care the patient actually receives. Some outcome indicators are only a vague reflection of the safety and quality of care provided, such as wound infections (Vincent, 2010). Not all bad outcomes can be prevented, and some are side effects from treatment. As Lilford, Mohammed, Braunholtz and Hofer (2003) points to, if the factors that affect outcomes, but cannot be controlled by the organisation or the individual health care workers, vary systematically between units, organisations or professionals, the comparisons of the entities might be biased.

A problem with using adverse events and errors as the main end point for measuring patient safety, as in the definition cited from Vincent (2010), is that a number of harms can occur during the course of health care. Malfunctioning medical equipment, postoperative infections, complications of surgery, harms from overcrowded hospitals – they are all important harms to the patient, but are not errors stemming from the process of health care. Thus, when applying a narrow definition only focusing on errors, these events are not included. They may even be the most important ones for the patient. The difficulties of linking errors to patient harm may be a reflection of the more general issue of how to link process and structure measures to outcomes as described above, and is thus not limited to patient safety (Vincent, 2010). The important question will be how one can identify the extent to which differences in outcomes are a reflection of differences in the quality of care delivered to patients, and how outcomes correlates to quality at the organisational level (Lilford et al., 2003).
What constitute good patient safety climate?

Some key features of a good safety culture/ climate includes a blame-free environment, where individuals are encouraged to report both errors and near misses without fear of blame or punishment. Further, it entails promotion of collaboration between the different disciplines in the clinical area, allocation of resources for patient safety initiatives, as well as an understanding that the organization is operating in a high-risk area (Ross, 2011). Positive patient safety climate is characterised by high levels of communication, mutual trust, and a shared perception of the importance of safety, as well as trust in the institution’s efforts in the continuous improvement and commitment to safety (Nordic Council of Ministers, (2010). Finally, confidence in the efficacy of preventive measures is important (Halligan and Secevic 2011).

2.5 Can safety climate predict patient safety outcomes?

Research undertaken in other high risk industries (for example aviation and nuclear industry) has shown that safe performance is dependent upon staff attention to reducing risk for error. Thus, in order to improve patient safety, it is important to assess staff attitudes and perceptions on safety. According to the UK’s National Patient Safety Agency (2004), there is evidence that when staff are encouraged to report adverse events and near-events, and analysis of causes and learning is encouraged, it can have a positive and measurable effect on the organization’s performance. In their seven-step program for patient safety, the promotion of a safety culture that is open and fair, and attempts to ensuring lessons are learned as well as for sharing information, is the first step towards creating safe health care (ibid.)

In Deilkås’ PhD dissertation, a number of studies that demonstrate the effect of safety climate on a number of patient outcomes are referred to, such as urinary tract infections, medication errors, fewer blood stream infections, lower ventilator associated pneumonias and shorter ICU lengths of stay (Deilkås, 2010). Although the evidence linking patient safety culture and safety outcomes are still limited, these studies might demonstrate reasons to assume that safety culture could be an important aspect in both patient safety research and health care quality research. According to Gershon et al (2004), it is important to measure safety climate,
because there is increasing evidence that aspects of safety climate and culture are associated with a number of worker and quality of care outcomes. However, this suggestive evidence has failed to be present in meta-analytic studies – the association seems to be present merely in single studies which indicates that many studies have not found any effects or even negative effects (Hoff, Jameson, Hannan, & Flink, 2004). Others conclude that there is limited data of support to improving patient safety outcomes (Pronovost, Holzmuller, Ennen and Fox, 2011). Moreover, as Singer, Lin, Falwell, Gaba and Baker (2009) points to, while most researchers presume that the presence of a strong safety culture will lead to fewer errors and better safety outcomes, there is little available evidence supporting this link (Singer et al., 2009; Nordic Council of Ministers, 2010). Little empirical research has aimed at testing this causal relationship in detail (Health Foundation 2011a), and most of the studies undertaken use perceptual or self-reported measures of clinical quality, instead of objective derived measures, which could impose uncertainties on the results (Singer et al 2009).

Many studies attempting to establish an association between safety climate and patient outcomes use a cross-sectional design (Hearld et al 2008). Hence, it is not possible to determine an eventual causal relationship between climate and outcomes, nor the direction of such an eventual causal relationship. It is possible that there may be a circular relationship, where changes in safety behaviours and outcomes also improves safety climate (Health Foundation 2011a). For example, in a study by Pettker et al., (2009), 10 obstetrics-specific outcomes were prospectively tracked, while repeatedly measured safety attitudes using the SAQ questionnaire. They demonstrated a significant decrease in the adverse outcomes measured and an increase in the safety climate scores. However, they do not mention the hypothesized theoretical relationship between safety scores and outcomes. Further, they have not included a control group, and are hence not able to assess causality. Just to demonstrate that there is a co-variation does imply causality (Nordic Council of Ministers, 2010). Finally, the level of aggregation is contested. Some studies aggregate safety climate scores to the organisational level or even aggregates different hospitals, due to the lack of patient outcome data on unit or hospital level (Flin, 2007). This could be problematic, as the safety climate could vary considerably between units (Deilkas & Hofoss, 2010). As Pronovost et al (2011) point to, culture have been shown to vary six times more within units in a hospital than between hospitals.
The construct validity of the concept of organizational safety culture/climate has been questioned. Cooper (2000) claims that safety culture probably is made up of many dimensions, which may not change simultaneously, as one dimension may change independently, leaving the others unaffected. It is not given that improvements in one aspect of care correlates to other aspects, and different dimensions of safety culture may correspond to different clinical practices, and hence have different end points. These are good arguments for measuring the dimensions of patient safety cultures independently rather than relying on a single, global measure.

### 2.5.1 The level of aggregation

The discussion of outcome measures directly relates to the discussion of levels of aggregation of safety climate survey data. As previously described, the safety climate can vary considerably more between units and wards in hospitals than between health care institutions. But, there is not always feasible to obtain valid data on safety outcomes on the unit level. In some cases, this can only be found at the organisational level. Thus, safety climate data have in some studies been aggregated to institutional level, and studied in relation to safety data at the same level. This could be inappropriate and lead to false conclusions. If one assumes - as many studies in fact do- (Hoff et al (2004), Hearld et al (2008), the Health Foundation (2011b), that there is a causal relationship from safety climate to safety outcomes, important differences in safety climate could be masked if the climate is aggregated to the hospital or institutional level, as much of the variation is found at unit/ ward level. One is hence not able to distinguish between the “problematic” units and those with a good safety climate. Further, one has no way to identify the units with the highest number of problematic events. This approach does not allow for direct comparisons of unit with low rates of problematic events, and units with high rates, to the safety climate. Conclusions based on inappropriate levels of aggregation should thus be interpreted with caution.

### 2.6 Measuring safety culture/climate

Historically, safety culture measurement has been most prominent within the organizational psychology research, and the methodology has in large part been developed within this
research tradition (Deilkås, 2010). The main focus has been industries and high reliability organizations. Surveys attempts to quantify and measure staff perceptions, which have been referred to as organizational climates. The use of safety attitude measurements in health care has as previously described, gained increased focus in health care settings recently. The first step to develop a safety culture is to assess and measure the organization’s present culture (National Patient Safety Agency 2004). There are a number of available instruments for measuring patient safety culture, and both qualitative and quantitative methods could (and should) be used. However, while many of these instruments share some of the dimensions, the terminology used to describe these dimensions vary greatly across instruments (Gershon et al., 2004). is in line with Schein’s definition of safety climate as an observable artefact of the underlying, non-observable culture.

2.6.1 The Safety Attitudes Questionnaire (generic version)

The Safety Attitudes Questionnaire (generic version) (SAQ) derives from a questionnaire used in commercial aviation, the Flight Management Attitudes Questionnaire (FMAQ). This questionnaire measures flight crew attitudes towards various interpersonal aspects of crew performance. These include items such as teamwork, the ability to speak up, leadership, communication, and joint decision-making. About 25% of the items in the FMAQ questionnaire demonstrated utility in the medical setting. It was therefore further developed into the Intensive Care Management Attitudes Questionnaire (Sexton et.al., 2006). A further refinement of this questionnaire, based on discussions with experts and health care providers, and two conceptual models (Vincent’s framework for risk and safety analysis, and the Donabedian conceptual model for quality assessment) (Deilkås, 2010), led to the identification of more than 100 items, covered by four themes (safety climate, teamwork climate, stress recognition and organisational climate). From this, the analytically derived structure of six factors, containing a total of 40 items, was developed. Three of the themes were identified as factors in the analysis, while the fourth (organisational climate) was further divided into three new factors; perceptions of management, working conditions and job satisfaction. An additional 20 items were included in the questionnaire, as they were perceived interesting and useful to the health care providers to whom the results of the pilot study were reported. The resulting questionnaire have been further refined into specific questionnaires for operating rooms, intensive care units, general inpatient wards, and
ambulatory clinics. The wording has been slightly changed to match the specific clinical setting, but the items are the same (Sexton et al., 2006).

According to Deilkås (2010), the SAQ is the most widely used instrument for the measurement of patient safety culture. It has been translated into seven languages, and has been administered in over 2000 hospitals in eleven countries. Further, it is a well-documented instrument for measuring patient safety culture, as it has been published benchmarking scores for 203 clinical areas in tree countries, and the overall response rate of these is 67%. Finally, SAQ may be the only questionnaire with results that link patient safety culture scores to patient safety outcomes (ibid.).

**Validation of SAQ in Norway**

The SAQ have been validated for use in a number of countries, and have been translated to many languages. In Norway this was done in relation to the PHD dissertation by Ellen Tveter Deilkås (Deilkas & Hofoss, 2008). It was validated using a back-translation method, where it was first translated to Norwegian and then translated back to English by a second, blinded researcher. It was then compared to the original English version, and any the re-translation were discussed with one of the authors of the American questionnaire (ibid.)

The SAQ was chosen for validation because the authors found this instrument to be the best-validated instrument, both with regards to internal and external validity. Internationally, the construct validity have been documented with published data from more than 200 sites in the UK, US and New Zealand, in different clinical areas, ICUs, ORs and ambulatory clinics (Sexton et al., 2006). With regards to the external validity, this have been studied via investigations of the relationship between safety culture scores and patient outcomes (Deilkas & Hofoss, 2008).

Deilkås and Hofoss demonstrated that the SAQ is a reliable and valid instrument for assessing patient safety cultures at wards and units also in the Norwegian health care setting. The confirmatory factor analysis undertaken in Norway gave a somewhat different factor structure than other assessments, with perceptions of management divided in two factors – perceptions of unit management and perceptions of hospital management (Deilkas & Hofoss, 2008). The SAQ (short form 2006, Norwegian version) consists of 36 questions, which relates to seven dimensions of patient safety (safety climate, teamwork climate, stress recognition, working
conditions, perceptions of hospital management, perception of unit management, and job satisfaction).

Table 1: Factor definitions and item description SAQ (generic version)

<table>
<thead>
<tr>
<th>Factor dimensions and descriptions</th>
<th>Items and examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Teamwork climate</strong></td>
<td>6 items</td>
</tr>
<tr>
<td>Perceived quality of collaboration between personnel</td>
<td>Nurse input is well received in this unit</td>
</tr>
<tr>
<td></td>
<td>In this unit, it is difficult to speak up if I perceive a problem with patient care</td>
</tr>
<tr>
<td><strong>Safety climate</strong></td>
<td>6 items</td>
</tr>
<tr>
<td>Perceptions of a strong and proactive organisational commitment of safety</td>
<td>I would feel safe being treated here as a patient</td>
</tr>
<tr>
<td></td>
<td>Medical errors are handled appropriately in this unit</td>
</tr>
<tr>
<td></td>
<td>I know the proper channels to direct questions regarding patient safety in this unit</td>
</tr>
<tr>
<td><strong>Stress recognition</strong></td>
<td>4 Items</td>
</tr>
<tr>
<td>Acknowledgement of how performance is influenced by stressors</td>
<td>When my workload becomes excessive, my performance is impaired</td>
</tr>
<tr>
<td></td>
<td>I am less effective at work when fatigued</td>
</tr>
<tr>
<td><strong>Working conditions</strong></td>
<td>4 items</td>
</tr>
<tr>
<td>Perceived quality of the ward work environment and the logistical support (staffing, equipment etc)</td>
<td>The level of staffing is sufficient to adequately treat the number of patients.</td>
</tr>
<tr>
<td></td>
<td>This hospital does a good job of training new personnel</td>
</tr>
<tr>
<td><strong>Job satisfaction</strong></td>
<td>5 items</td>
</tr>
<tr>
<td>Positivity about the work experience</td>
<td>I like my job</td>
</tr>
<tr>
<td></td>
<td>Working here is like being part of a large family</td>
</tr>
<tr>
<td><strong>Perceptions of hospital management</strong></td>
<td>4 items</td>
</tr>
<tr>
<td>Approvements of hospital management action</td>
<td>Hospital management supports my daily efforts</td>
</tr>
<tr>
<td></td>
<td>Hospital management does not knowingly compromise the safety of patients</td>
</tr>
<tr>
<td><strong>Perceptions of unit management</strong></td>
<td>4 items</td>
</tr>
<tr>
<td>Approvements of unit management action</td>
<td>Unit management supports my daily efforts</td>
</tr>
<tr>
<td></td>
<td>Unit management does not knowingly compromise the safety of patients</td>
</tr>
</tbody>
</table>

(Deilkas & Hofoss, 2008; Sexton et al., 2006)

2.6.2 The RN4CAST survey

The RN4CAST survey is a large, multinational cross-sectional survey conducted in 12 countries (with 3 “satellite” countries, including Norway), in 486 hospitals, 2169 units and 33700 nurses, in 2009. The survey gathered information on several levels of the health care service, including hospital profiles (such as number of beds and patient days), patient outcomes (in Norway from Norwegian Patient Register) and survey data from nurses. All participating countries have followed the same procedures for data collection, and used the same survey instruments (Aiken et al., 2012). The Norwegian survey aims at developing knowledge about the relationship between hospital staffing, management of nursing
resources, job satisfaction and patient safety in the Norwegian context (project plan Personnel and patient safety, Kunnskapssenteret, 2009).

2.6.3 Hospital Survey on Patient Safety Culture (HSOPSC)

Items from Hospital Survey on Patient Safety Culture (HSOPSC) are part of the questionnaire in RN4CAST. The HSOPSC was developed in the US, funded by the National Agency for Healthcare Research and Quality (AHRQ), based on the recognition of the need for a measurement tool for assessing patient safety cultures in health care organizations. It was developed based on literature reviews, including safety culture surveys, and interviews with health care personnel. After the questionnaire was designed, it was pilot tested on hospital employees from 21 American hospitals. Statistical analysis of the data were performed in order to assess psychometric properties of the instrument (item statistics, reliability and validity of the scales, as well as exploratory and confirmatory factor analysis for examination of the factor structure) (Sorra and Nieva, 2004). HSOPSC has been found to have good psychometric properties, where statistical testing includes item analysis, exploratory and confirmatory factor analysis, as well as inter-correlation and reliability analysis. The instrument has been used in a number of different countries, all over the world (Chen and Li, 2010).

Validation of HSOPSC in Norway

The HSOPSC instrument was translated and tested with regards to validity and reliability in Norwegian health care by Olsen (2008). The author chose this instrument because, at the time of the study, it was demonstrated to have met more psychometric criteria compared to other instruments, and because benchmarking data were available from the Internet. The instrument was translated to Norwegian and then back to English again by two independent researchers, and the instrument was pilot tested in a small group of health care workers. The study included 1919 workers from a Norwegian hospital, with a response rate of 55%. The author found that the factorial structure of the instrument was reproduced in this Norwegian hospital, with ten safety culture dimensions and four outcome-measure dimensions. The results generally met the reliability and validity criteria as have been demonstrated internationally (ibid.)
Respondents are asked of their overall perception of safety, the frequency and number of event reporting and the overall grade of patient safety (Nieva & Sorra, 2003). The items are rated on a 5-item Likert-type scale. The first seven dimensions measure unit-level aspects of safety culture, the three following factors measures hospital-level aspects of safety culture, and finally, four outcome variables are included (AHRQ webpage, accessed 05.04.2012)

<table>
<thead>
<tr>
<th>Factor dimensions and descriptions</th>
<th>Items and examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unit-level aspects of safety culture</strong>&lt;br&gt;Supervisor/Manager Expectations &amp; Actions Promoting Safety&lt;br&gt;Supervisors/managers consider staff suggestions for improving patient safety, praise staff for following patient safety procedures, and do not overlook patient safety problems</td>
<td>4 items&lt;br&gt;My supervisor/manager says a good word when he/she sees a job done according to established patient safety procedures.&lt;br&gt;My supervisor/manager seriously considers staff suggestions for improving patient safety.</td>
</tr>
<tr>
<td><strong>Organizational Learning—Continuous Improvement</strong>&lt;br&gt;Mistakes have led to positive changes and changes are evaluated for effectiveness</td>
<td>3 items&lt;br&gt;We are actively doing things to improve patient safety. Mistakes have led to positive changes here</td>
</tr>
<tr>
<td><strong>Teamwork Within Units</strong>&lt;br&gt;Staff support each other, treat each other with respect, and work together as a team</td>
<td>4 items&lt;br&gt;People support one another in this unit. When a lot of work needs to be done quickly, we work together as a team to get the work done.</td>
</tr>
<tr>
<td><strong>Communication Openness</strong>&lt;br&gt;Staff freely speak up if they see something that may negatively affect a patient and feel free to question those with more authority</td>
<td>3 items&lt;br&gt;Staff will freely speak up if they see something that may negatively affect patient care. Staff feel free to question the decisions or actions of those with more authority.</td>
</tr>
<tr>
<td><strong>Feedback and Communication About Error</strong>&lt;br&gt;Staff are informed about errors that happen, given feedback about changes implemented, and discuss ways to prevent errors</td>
<td>3 items&lt;br&gt;We are given feedback about changes put into place based on event reports. We are informed about errors that happen in this unit.</td>
</tr>
<tr>
<td><strong>Nonpunitive Response to Error</strong>&lt;br&gt;Staff feel that their mistakes and event reports are not held against them and that mistakes are not kept in their personnel file</td>
<td>3 items&lt;br&gt;Staff feel like their mistakes are held against them. (negatively worded)&lt;br&gt;When an event is reported, it feels like the person is being written up, not the problem. (negatively worded)</td>
</tr>
<tr>
<td><strong>Staffing</strong>&lt;br&gt;There are enough staff to handle the workload and work hours are appropriate to provide the best care for patients</td>
<td>3 items&lt;br&gt;We have enough staff to handle the workload. Staff in this unit work longer hours than is best for patient care. (negatively worded)</td>
</tr>
</tbody>
</table>
### Hospital-level aspects of safety

<table>
<thead>
<tr>
<th>Hospital Management Support for Patient Safety</th>
<th>3 items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital management provides a work climate that promotes patient safety and shows that patient safety is a top priority</td>
<td></td>
</tr>
<tr>
<td>The actions of hospital management show that patient safety is a top priority</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Teamwork Across Hospital Units</th>
<th>4 items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital units cooperate and coordinate with one another to provide the best care for patients</td>
<td></td>
</tr>
<tr>
<td>There is good cooperation among hospital units that need to work together. Hospital units work well together to provide the best care for patients.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hospital Handoffs and Transitions</th>
<th>4 items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Important patient care information is transferred across hospital units and during shift changes</td>
<td></td>
</tr>
<tr>
<td>Things &quot;fall between the cracks&quot; when transferring patients from one unit to another. (negatively worded) Important patient care information is often lost during shift changes. (negatively worded)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome variables</th>
<th>4 items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Perceptions of Safety</td>
<td></td>
</tr>
<tr>
<td>Procedures and systems are good at preventing errors and there is a lack of patient safety problems</td>
<td></td>
</tr>
<tr>
<td>Our procedures and systems are good at preventing errors from happening.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency of Event Reporting</th>
<th>3 items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mistakes of the following types are reported: (1) mistakes caught and corrected before affecting the patient, (2) mistakes with no potential to harm the patient, and (3) mistakes that could harm the patient</td>
<td></td>
</tr>
<tr>
<td>When a mistake is made, but is caught and corrected before affecting the patient, how often is this reported? When a mistake is made, but has no potential to harm the patient, how often is this reported?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patient Safety Grade (of the Hospital Unit)</th>
<th>1 item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please give your work area/unit in this hospital an overall grade on patient safety.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Events Reported</th>
<th>1 item</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the past 12 months, how many event reports have you filled out and submitted?</td>
<td></td>
</tr>
</tbody>
</table>

(Sorra and Nieva, 2004; AHRQ http://www.ahrq.gov/qual/patientsafetyculture)

### 2.7 Safety Attitudes Questionnaire and Hospital Survey on Patient Safety Culture compared

The Norwegian version of the Safety Attitudes Questionnaire (SAQ), Generic version (Short Form 2006) consist of seven dimensions, as can be seen from table 1. The dimensions relates to aspects of safety climate, teamwork climate, stress recognition, working conditions, job satisfaction, and perceptions of unit and hospital management. The HSOPSC instrument contains of a total of ten dimensions relating to patient safety aspects (see table 2). Seven of these dimensions regard the unit-level and the remaining three regards the hospital-wide level. One important difference between SAQ and HSOPSC is that the HSOPSC also includes outcome measures. Two single-item and two and multiple-item scales containing self-reported measures of the number of events reported, overall patient safety grade, overall perceptions of safety and the frequency of events reported are included.
**Conversion of dimension scores of the two instruments**

A study published in April 2012 investigated whether scores on dimensions from the HSOPSC instrument could be converted to scores to the corresponding dimension in a second instrument, the SAQ (Etchegaray & Thomas, 2012). They focused on the dimensions teamwork and safety climate, and found that while they were able to convert scores from HSOPSC into scores for the safety climate and teamwork dimensions of the SAQ instrument; the per cent of variance explained was moderate. Further, they stated that it is not self-evident that the scales match each other, as the items included in the dimensions in the two instruments are not the same. They concluded that the surveys probably cannot be converted (ibid.).
3 Design

According to Kumar (2005), research is the process of collecting, analysing and interpreting information, in order to answer questions. This process has certain requirements; it must be controlled, rigorous, systematic, valid, empirical and critical. These requirements are the same whether the research undertaken is qualitative or quantitative. What distinguish qualitative and quantitative research are the methods for data collection, procedures for analysing the data, and how the findings are communicated. Where quantitative research emphasizes measurement and classification of variables, aiming at drawing inferences about associations, qualitative research aims at describing variation in the research topic using fewer cases but in-depth assessments of these (Kumar, 2005).

A questionnaire can both be quantitative and qualitative, depending on the purpose of the research and the design of the questionnaire. Quantitative methods are useful when one wants to compare groups or units, and the sample size is large. If the aim is to assess opinions and attitudes, or testing an eventual relationship between variables, one can use a structured, close-ended questionnaire. The data collected can thus be suitable for quantitative statistical analyses.

3.1 Ethical approval, consent and confidentiality

This survey does not include patients or patient related information, and are thus not subject for restrictions following the Norwegian law on medical and health research, § 2 (Helseforskningsloven).

However, any collection, recording or storing of personal data with use of a computer requires the research project to notify the Data Protection Official for Research (Lov om behandling av personopplysninger). The use of personal data in this project implies that the project was required to submit a notification form to the Data Protection Official for Research. As this survey is part of a larger project, the approval for this survey has been given in relation to this larger project (Appendix I).
The cover letter in the questionnaire clearly stated that participation was voluntarily, and that
the responses would be treated with confidentiality by all researchers. When the respondents
submitter their questionnaire, it was regarded as consent to participate in the survey. The Data
Protection Official for Research stated that the list connecting names and questionnaires-IDs
had to be deleted when data collection was ended. The researchers have signed a
confidentiality agreement, and all data were stored and analysed in the Norwegian Knowledge
Centre for Health Services’ secure database.

3.2 The survey

This study is designed as a quantitative survey with a cross-sectional design. A survey
collects information from the study population through their responses to standardized
questions (Chambliss and Schutt 2010). As the researcher is not present to clarify and answer
questions, it is essential that the questions are clear and easy to understand. To increase the
response rate, the distribution form of the questionnaire is important. A collective distribution,
such as distribution in the work place, is a way that usually gives highest response rate,
compared to distribution via mail or in a public place (Kumar, 2005).

As the aim of this survey was to investigate if the questions from RN4CAST can be used as
measures of patient safety attitudes, it was considered a great advantage to have respondents
answering both questionnaires at once. Thus, it is possible to estimate measures of correlation
and agreement between the two instruments. The aim is to investigate whether there is an
association between questions in the two instruments. The survey is thus correlational. The
study is also descriptive – it can be used to describe how the respondents perceive the patient
safety culture at their ward at the time they responded to the survey. Both instruments consist
of closed-ended questions, and they include both positive and negative worded items. Space
for comments on patient safety issues at the work place/clinical area was included.

3.2.1 The questionnaire

The questionnaire had already been designed when I entered the project. It is a combination of
the Norwegian version of Safety Attitudes Questionnaire (SAQ (2006 short form, general
version), and questions regarding patient safety attitudes from the RN4CAST survey. These
questions are from the Hospital Survey on Patient Safety Culture (HSOPSC). There is also included some demographic questions relating to gender, work experience and occupational group.

**Safety Attitudes Questionnaire (SAQ)**

The first 41 items in the questionnaire consists of the entire Safety Attitudes Questionnaire (2006 short form, Norwegian version). All questions are closed-ended with response along a six—point Likert-type scale (strongly disagree, slightly disagree, neutral, slightly agree, strongly agree, don’t know). The sixth response option is not included in Sexton et al’s article from 2006. The instrument has been further described in chapter 2.6

**Hospital Survey on Patient Safety Culture (HSOPSC)**

The items from RN4CAST on patient safety culture are items from the Hospital Survey on Patient Safety Culture (HSOPSC). The questions are seven items from the five dimensions communication openness, feedback and communication about error, nonpunitive response to error, hospital management support for patient safety, and hospital handoffs and transitions. All questions are closed-ended with response along a five—point Likert-type scale (strongly disagree, disagree, neither, agree, strongly agree). See chapter 2.6 for a more thorough description of the instrument.

**The composite scales**

As will be further described, two composite safety climate scales are created from the RN4CAST items relating to patient safety. They differ somewhat in the items included. Table 3 gives an overview of the items included in the two scales, and which HSOPSC dimension the items belong to.
Table 3: Items included in the RN safety composite scales, and the SAQ safety climate

<table>
<thead>
<tr>
<th>Feedback and communication about error</th>
<th>Composite scale RN safety II</th>
<th>Composite scale RN safety I</th>
<th>SAQ safety climate items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback and communication about error</td>
<td>We are given feedback about changes put into place based on event reports</td>
<td>We are given feedback about changes put into place based on event reports</td>
<td>I would feel safe being treated here as a patient</td>
</tr>
<tr>
<td></td>
<td>In this unit, we discuss ways to prevent errors from happening again</td>
<td>In this unit, we discuss ways to prevent errors from happening again</td>
<td>Medical errors are handled appropriately in this unit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>I know the proper channels to direct questions regarding patient safety in this unit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>I receive appropriate feedback about my performance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>In this unit, it is difficult to discuss errors</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>I am encouraged by my colleagues to report any patient safety concerns I may have</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The culture in this unit makes it easy to learn from the errors of others</td>
</tr>
<tr>
<td>Communication openness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonpunitive response to error</td>
<td>Staff feel like their mistakes are held against them</td>
<td>Staff feel like their mistakes are held against them</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handoffs and Transitions</td>
<td>Things “fall between the cracks” when transferring patients from one unit to another</td>
<td>Important patient care information is often lost during shift changes</td>
<td></td>
</tr>
</tbody>
</table>

The items from the RN4CAST survey are patient safety items relating to perceptions of how errors are treated, how feedback is given and how patient safety concerns are treated. The different dimensions in SAQ have been assessed, and the items from RN4CAST have been found to be most similar to the items in the SAQ safety climate dimension. The items included in the RN safety composite scale are hence assumed to relate to the same underlying constructs as the SAQ safety climate dimension.

### 3.3 Data collection

Initially, this questionnaire was supposed to be distributed to nine Norwegian hospitals in a number of wards. However, the national Patient Safety Campaign were planning on distributing the same SAQ questionnaire the autumn of 2011, and most of the hospitals were reluctant to make their employees respond to questionnaires with similar themes twice in a short time. Thus, eventually we were left with only one hospital, and two wards. The wards...
chose to participate because they were already involved in a research project at the Norwegian Knowledge Centre for the Health Care Services.

The questionnaires were distributed to health care personnel at two wards in one Norwegian hospital during the summer of 2011. Before distribution, a list of names and addresses of the employees had been collected, and this was used to label the questionnaires with a personal id-number. A questionnaire was then placed in a named envelope together with a pre-paid, addressed return envelope, which the respondents were supposed to return to the Knowledge Centre via mail. A researcher from the Knowledge Centre and the author of this thesis travelled to the hospital to distribute the questionnaires. Here we met with the hospital’s infection control nurse as well as one nurse from each of the two participating wards. This contact aimed at both summing up the data collection of the previous project, as well as to increase the support for this one. It was seen as a great advantage to be able to build upon an established relationship with a motivated and interested employee, as this could increase the support for this questionnaire and eventually the final response rate.

When the questionnaires were returned, the names would be crosschecked and removed from the list, so that there would be no way to trace the answers back to the individuals. After 6 weeks, a reminder was sent out to the non-respondents, together with a new questionnaire. When the questionnaires arrived at the Knowledge Centre the data were entered into PAWS statistics (IBM SPSS 18.0 (Statistical Package for Social Sciences)). The Data Protection Official for Research approved the method of data collection and handling.

### 3.4 Sample

The first participating ward is a surgical ward, with a total of 21 beds and 23 nurses, as well as 13 nurse aid staff and 14 orthopedists. The second ward is an intensive care unit with a total of 71 nurses, as well as 4 nursing aid personnel and 18 anesthesiologists. As only nurses were asked to respond to the RN4CAST survey, only the responses from nurses in these two wards are included in the analysis. This gives a sample of 94 nurses. 27 of the nurses did not complete the survey, giving a number of respondents of 67 and a response rate of 71%. As the aim was not to investigate the patient safety climate at the different wards, the two wards are treated as one sample, and all statistical analyses are performed using this one sample.
approach. The demographic composition of the sample will be presented more closely in chapter five.

3.5 Response rate

A common problem with surveys is a low response rate. There are a number of steps available to increase the response rate, but ultimately it can depend on how interesting the survey appears for the respondents, the length of the questionnaire and the methodology used to distribute the questionnaire. Chambliss and Schutt (2010) claims one should be content with a response rate of 50%

Since only two units participated in this survey, it was important to ensure a high response rate. One way of achieving this, while at the same time ensuring high quality of the data collected, is by ensuring that the purpose and relevance of the survey is explained to the potential respondents (Kumar, 2005). Hence, the cover letter of the survey is could play an important role in ensuring high response rate and high data quality. The cover letter should be interesting to read and it must explain issues such as voluntary participation and confidentiality (Chambliss and Schutt (2010).

3.6 Common method variance

Common method variance is defined as the variance attributable to the method of measurement rather than the construct the measures represent (Podsakoff, MacKenzie, Lee, & Podsakoff (2003). Method biases are said to be one of the main sources of measurement error – and can hence threaten the conclusions about underlying relationships between variables measured. They can have both a random and a systematic component. The latter is the most threatening to the validity of the conclusions, because it provides an alternative reason for the observed relationships between measures of different constructs than the one hypothesized by the researchers (ibid.). In this survey, there is a hypothesized correlational relationship between one construct, namely safety climate, measured by two different instruments. If the
measures of the two constructs share the same method of measurement, these methods may exert a systematic effect of the observed correlation between the measures.

The systematic effect may increase if the same respondents rate the different constructs. This could be due to a cognitive need for appearing coherent (Podsakoff et al., 2003) – that is, if one has responded in a certain way to the predictor variable, one can automatically and unknowingly respond in a way that supports this response in the criterion variable. If one gives an overall positive response to a question of grading the safety at one’s hospital, one might respond to items regarding aspects of safety in a more positive manner. Further, any observed relationship between predictive and outcome variable could be distorted by how the respondents perceive the relationship between these items in the questionnaire, and their implicit assumptions about this relationship. If they implicitly assume that there should be a relationship between how they rate the patient safety state, and how the work environment should support the patient safety state, it could affect the way they respond. Finally, just the fact that predictor and criterion variables are measured at the same point of time could be a source of mono-method bias.

In a review by Meade, Watson and Kroustalis (2007), attempts were made to identify the magnitude of common method bias in organisational research. They found that even if the presence of common method variance in organisational research is pervasive, the magnitude of the common method bias is rather small in many cases. They claim that use of negative worded items; randomized item order and multiple methods and raters where this is feasible would in many cases be enough to counteract common method bias. Further, common methods may not even result in problematic common method bias – and is not necessarily representing a threat to the validity of study conclusions in every case (ibid.).
4 Analytical approach

The questionnaires were entered into PAWS statistics (SPSS) as soon as they arrived the Knowledge Centre. A reminder was sent to the ones not responding, together with a new copy of the questionnaire and a pre-paid, addressed envelope, 6 weeks after the initial distribution. This was the only reminder sent – and if the questionnaire still was not returned, the individuals were labeled as non-respondents. From the Data Protection Office for Research, only one reminder was permitted. The period of data collection was extended by a few weeks, as the initial distribution of questionnaires took place during the summer period. In order to reach those on summer holidays, it was quite a long time period before the reminder was sent out.

The responses were recoded from a scale from 1-5 (6) to 0-100. This is in accordance with recommendations from Sexton et al’s article (2006). The 6th response option in the SAQ, “don’t know” (DK), was coded as missing. There is no unambiguous way to treat DK responses, but the survey methodology resources consulted recommended coding as missing, and it is thus the approach I have followed. This could pose difficulties in the analysis, as some of the items in the survey have a quite substantial number of don’t know-responses. The items with negative wording were then reversed, and the mean of each item were calculated. The items were then collapsed into scales (factors). The mean score and the percentage positive response for each factor, along with reference values from the literature, will be presented in the results section of this thesis. The percentage of respondents giving positive scores (> slightly agree) will also be presented. According to Sexton et al (2006), this is a measure of the strength of the safety climate. This is important, because, as Zohar et al points to in their article, the strength indicates the level of consensus among the individuals in the unit regarding the importance of patient safety (Zohar et al., 2007). The number of missing items was assessed before any statistical analyses were performed, and will be presented in the results section.
4.1 Validation of questionnaires

According to Prous et al (2009) a valid questionnaire is required to be feasible, reliable and precise, have content validity (be adequate for the problem it is intended to measure) and construct validity (adequate for the underlying concept or phenomenon), and be sensitive to change. To investigate an instrument’s psychometric properties is a process that involves assessment of all these different requirements.

4.1.1 Likert-type items and level of measurement

There is extensive debate as to whether data derived from Likert-type scales can be treated as interval level data. A Likert-type scale produces data on the ordinal level, and in order to perform parametric tests, to present results using means, standard deviation etc, data on the continuous level is required. Thus, in theory one cannot treat data from a Likert-type scale as continuous data. This debate is somewhat theoretical, however. In practice, parametric statistics based on the normal distribution is often applied. Some claim that the parametric statistics are robust enough to use, even if the underlying assumptions are not met. Further, according to Ringdal (2007), one can apply parametric statistics on ordinal level data, given that the data are assumed to measure some latent continuous variable, if it is reasonable to assume that the variable is continuous in the population, and have at least five categories. In this case, the data have been transformed from responses along a scale from 1-6, into 0-100, and the analysis is performed on factors that are averages of many items. Thus, the assumptions behind parametric statistics are found to be met, and the statistics performed in the analysis will be based on the normal distribution.

4.1.2 Reliability

An instrument’s reliability is the degree to which the instrument accurately measures something without error. Reliability measures the proportion of the total variance a variable adopts that is owed to true differences between the respondents. A reliable instrument is thus a precise instrument, given that the measurements are free of error. Eventual measurement errors could be caused by either a systematic error (bias) or a random error (Prous, de Yébenes, Salvanès and Ortels, 2009). The effect of systematic errors could result in erroneous
conclusions about correlations of variables in the instrument, and it is hence critical to reduce the chances of such errors to occur.

**Interitem reliability**

To assess the internal consistency, or the inter-item reliability, of the instrument used in this survey, the Chronbach’s alpha will be estimated for each factor. Cronbach’s Alpha is one of the most frequent statistics applied to assess internal consistency (Pallant, 2007). It is a group-level summary statistic or coefficient, that describes the degree to which measurements from a specific sample of respondents are replicated or consistent across a set of items or objects. It is used on continuous item-response formats (for example 1-5) (Helms, Henze, Sass, & Mifsud, 2006). It is important to note that even if Chronbach’s alpha is an indicator of the degree to which respondents’ answers to scale items in an instrument co-vary, it is not necessarily a measure of homogeneity (unidimensionality) of the responses. Thus, it should not be used to indicate the extent to which a single factor has been measured. The internal consistency depends on the number of items that makes up the factor and the mean correlation between these items. Thus, a conceptual framework for explaining the associations between items is required in order to avoid high alpha scale scores on scales with little conceptual meaning (Ibid.)

**Interrater reliability**

The intraclass correlation is a measure of agreement within raters for the same subject. The ICC is the proportion of total variability accounted for by the variability among subjects. A high ICC means that not much of the variability is due to variability in measurement on the two instruments; that the agreement between them is high (Prieto, Lamarca, Casado and Alonso, 1997)

An estimate of interrater reliability can give information on the proportion of ratings that represents an underlying construct. There are six different ICCs, with different interpretation.

**Pearson’s and Spearman’s correlation coefficient**

To further assess the relationship between the variables, the Pearson’s correlation coefficient will be calculated. This is a measure of strength of relationship between variables, and can
vary from -1 to 1. The former indicates a perfect negative correlation, the latter a perfect positive correlation. Values around 0 indicate no relation between the variables. It is important to note that to just “look around” in the data for correlations is usually a bad idea. Assessments of correlation should always be performed on basis of theoretical knowledge and hypotheses that are theory-derived.

As the Pearson’s correlation coefficient is a parametric measure, the non-parametric alternative (Spearman’s rho) will also be calculated. As previously described, scales that are assumed to reflect an underlying construct, containing items with more than five response categories can be treated as continuous and hence it is appropriate to perform parametric statistics. If the Pearson’s and the Spearman’s correlation coefficient provide similar estimates of the strength of the relationship between the two variables, parametric statistics will be used in the following analyses.

<table>
<thead>
<tr>
<th>Correlation Coefficient value</th>
<th>Strength of linear relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least 0.8</td>
<td>Very strong</td>
</tr>
<tr>
<td>0.6 up to 0.8</td>
<td>Moderately strong</td>
</tr>
<tr>
<td>0.3 to 0.5</td>
<td>Fair</td>
</tr>
<tr>
<td>Less than 0.3</td>
<td>Poor</td>
</tr>
</tbody>
</table>

Source: Chan (2003)

**4.1.3 Validity**

An instrument’s validity refers to how well it measures what it is intended to. There are four main different ways to assess the validity of an instrument (Chambliss and Schutt 2010). It is important to note that these four types of validity are not the only ones existing, they are not universally agreed upon, and there exists multiple terminologies depending on the research tradition. For example, construct and criterion validity can sometimes be used interchangeably. Table 5 presents an overview of the four types of validity.
Table 5: Types of validity

<table>
<thead>
<tr>
<th>Types of validity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Face validity</strong> (logic or apparent validity)</td>
<td>Refers to the degree an instrument appears to measure, in the opinion of experts and respondents</td>
</tr>
<tr>
<td><strong>Content validity</strong></td>
<td>Establishes whether or not the entire range of the concept’s dimensions have been covered by the instrument. This relates to the composition of the instrument, and is usually based on the expert judgements from many sources (e.g. literature reviews, pilot studies and expert opinions). It can also be explored using factorial analysis</td>
</tr>
<tr>
<td><strong>Criterion validity</strong></td>
<td>Established if the results from a new instrument matches an already validated instrument. They can both be measured at the same time (concurrent validity) or predict scores on a different instrument in the future (predictive validity)</td>
</tr>
<tr>
<td><strong>Construct validity</strong></td>
<td>Refers to the degree an instrument correlates to the theorized underlying construct it is intended to measure – if it measures what it is intended to measure</td>
</tr>
</tbody>
</table>

source: Prous et al 2009, Chambliss and Schutt 2010

Both face and content validity of the SAQ and the HSOPSC have been established (Sexton et.al, 2006; Blegen, Gearhart, O’Brien, Seghal and Alldregde, 2009), also in Norway (Deilkas & Hofoss, 2008), (Olsen, 2008). With regards to its construct validity, scores from SAQ have been associated with a number of different patient outcomes (Colla, Bracken, Kinney, & Weeks, 2005). Here, the concurrent validity of the RN4CAST patient safety attitudes questions will be assessed, in relation to the previously validated instrument, SAQ.

### 4.1.4 Assessment of agreement

The correlation coefficient measures linear relationships. However, it is possible that there exists a linear relationship between two variables, but there could still be large differences in the variables across their range of values. The Bland-Altman plot, or difference plot, is a graphic display of the agreement of the variables. It is a common method used to assess agreement between two methods that measures continuous variables measured on the same scale. The Bland-Altman plot displays the difference scores of the two variables - the bias (x-y), or the mean difference between the two measures, on the Y-axis, and the mean of the two variables (x+y/2) on the X-axis. It can also include the zero bias-line and the 95% upper and lower agreement lines (0+1.96diff and 0-1.96diff). If there is no systematic bias, the differences will be symmetric around the zero bias line (Dewitte, Fierens, Stöckl, & Thienpont, 2002)
A Bland-Altman assessment of agreement was used to compare the agreement between the two instruments. For this purpose, MedCalc software was used. A range of agreement was defined as mean bias ± 1.96 SD (Bland and Altman, 1986). The results from this assessment were compared to the results from the correlation analysis, and will be presented in the results chapter.

### 4.2 Exploratory factor analysis

Factor analysis is a multivariate statistical procedure most frequently used in psychology and education (Williams, Ownsman and Brown, 2010). It can be used to reduce a large number of variables into smaller numbers of coherent subscales (Pallant 2007) – which are often referred to as factors. Further, as it can contribute to establishing underlying dimensions between measured variables and latent constructs, it can contribute to formation and refining of theories. Finally, it can contribute to evidence assessment of the construct validity of self-reporting scales (Williams et al., 2010). There are two main approaches to factor analysis – exploratory and confirmatory factor analysis. The approach used in this thesis is a principal component analysis, which is a form of exploratory factor analysis. The reason for the choice of method relates to the different useability in the different stages of the research process of the two approaches. Exploratory factor analysis is most often applied in early stages of the research process, to gather information about the interrelationships of given set of variables. Confirmatory factor analysis is most dominantly used to test specific hypotheses and theories regarding the underlying structure of the given variables (Pallant 2007). Further, there are a variety of different but related statistical techniques underlying the factor analysis umbrella. The most common ones are principal component analysis (PCA) and factor analysis (FA). These are in many ways similar, and the terms are often used interchangeably in research. Both are tools in the process of producing a smaller number of linear combinations of the original variables in a way that accounts for most of the variability in the pattern of correlations.

The two approaches differ in the way they treat the variance in the variables. In PCA, the original variables are transformed to a smaller set of linear combinations. Here, all the variance in the variables is used. In FA, however, the factors are estimated via the use of
mathematical models, where only the shared variance is analysed. These two methods will in many cases produce similar results, but there is nonetheless extensive debate as to which method is recommended (ibid.).

### 4.2.1 Decisions in factor analysis

Before conducting a factor analysis, the sample size and adequacy must be assessed, and a number of decisions needs to be made about the procedures and techniques applied in the analysis. In the following sections, the different decisions will be presented and discussed.

#### Sample size and adequacy

The sample size in factor analyses should be given thorough considerations. According to Williams et al. (2010), there exists several rule of thumbs regarding size in the literature, and there are very varying opinions. However, there seems to be agreement on that the larger sample, the better. Some proposes a sample size of 300 or more, while others claim that the ratio of subjects to items is of greater concern than the sample size itself. Here, the recommendations vary from 10 cases per item to be factored, to a ratio of 5:1 (Pallant 2007).

Measures of sample adequacy give information on how strong the items in the analysis are correlated, and it does help the researcher assess whether there is an underlying relationship between the items, and thus the grouping of the items in the instrument into coherent subscales (Burton and Mazerolle 2011). An inspection of the correlation matrix for correlation coefficients of greater value than .03 is recommended in the literature. If few are found, it could imply that a factor analysis is not appropriate on the data (Pallant 2007). To assess sampling adequacy, the Kaiser-Meyer-Olkin measure of sampling adequacy needs to be examined. A KMO correlation above .60 is generally assumed to be adequate for analysing of the PCA output. Further, Bartlett’s test of sphericity should be significant if the factor analysis to be undertaken. If the test does not produce significant results, it is an indication of no relationship between the variables.

#### Factor extraction – number and methods

The method for extracting factors needs to be determined next. Factor extraction (identification) is the process of determining the smallest number of factors that best can
represent the interrelations in the given variables (Pallard, 2007). Other than drawing on a priori knowledge about the number of factors needed to explain the underlying latent constructs, there are a number of approaches available to guide this process. The most common is the previously mentioned Principal Component Analysis. The need for simplicity in the number of factors must be balanced with the need for explaining as much of the variance in the variables as possible. The first factor is the combination of items whose shared correlations explain the most of the variance, and the following factors are the ones that explain the greatest of the remaining total variance (Burton and Mazerolle, 2011). There is no given rule for balancing these needs; the researcher needs to make judgements in this process. However, Pallard (2007) describes three techniques that can provide assistance when deciding on how many factors to keep from the analysis. The first is Kaiser’s criterion, or the eigenvalue rule. This specifies that only factors with an eigenvalue of 1.0 or more should be retained. The eigenvalue is the amount of the total variance explained by the factor. Second, Catell’s scree test, where one plots each of the eigenvalues of the factors, and inspects this plot for when the shape of the curve changes direction and becomes horizontal. It is then easy to see which factors contribute most to the explanation of the variance in the variables and thus should be retained – these are the ones above the elbow, or break, in the plot. The final technique is parallel analysis, where the factors’ eigenvalues are compared to randomly generated eigenvalues from data sets with similar size. Here, only those eigenvalues that exceeds the corresponding values from the random data set are retained. This has been shown to be the most accurate method for factor retention, as both the two previously described approaches have tendencies to overestimate the number of components (Pallard, 2007). However, this is the method with the least reported use in the literature, perhaps because it is not available in standard statistical programs such as PAWS or SAS (Williams et al., 2010).

Factor rotation

The next step in the analysis is the selection method for factor rotation. This is a tool for easing the interpretation of factor analysis, and involves presenting the patterns of loadings in a more easily interpretable way, as it shows how the different variables ‘clump together’ (Pallard, 2007). Rotation maximises the high item loadings and minimizes the low item loadings (Williams et al., 2010). There are two techniques to choose between, which results in
either orthogonal (uncorrelated) or oblique (correlated) factor solutions. The latter are more difficult to interpret, describe and report, and in practice they often result in similar solutions. Hence, many researchers apply both orthogonal and oblique rotations and report the one clearest and easiest to interpret (Pallard, 2007). A different approach is to assess the inter-correlations among the factors. If this is low, one can use an orthogonal rotation method.

**Interpretation**

At least two or three items should load on a factor in order to give it a meaningful interpretation (Williams et al., 2010). Further, instrument items should only load to a single factor at .50 or higher, with no cross loadings (loading on more than one factor) on more than >.30 (Burton and Mazerolle 2011). The factor loadings should then be given a name or a theme that reflects an underlying theory or concept.
5 Results

5.1 Sample descriptives

5.1.1 Response rate

As 67 nurses out of a total of 94 completed and returned the questionnaire, the final response rate was 71 %. This is a high response rate, and is in large part due to the efforts made by the hospital hygiene nurse and the nurse representatives from the two wards, as they clearly communicated the importance of completing the questionnaire to their colleagues. The response rate is somewhat higher than the ones Sexton et al (2006) presents in their article on benchmarking data for SAQ. See table 6 for demographic description of the sample in this survey. I have chosen to aggregate these two wards to one sample. However, to treat them as one might pose challenges. As previously described, the level of aggregation might be of significance. As no attempt is made here to examine eventual relationships between safety climate and safety outcomes, the dangers of incorrectly aggregation of data have been considered to be small.

Demographic description of sample

Table 6: Descriptive statistics of the sample (n=94 non-respondents n=27)

<table>
<thead>
<tr>
<th>Demographic variable</th>
<th>Frequency n, (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5, (7.5%)</td>
</tr>
<tr>
<td>Female</td>
<td>60, (89.6%)</td>
</tr>
<tr>
<td>Missing</td>
<td>2 (2.9%)</td>
</tr>
<tr>
<td>Position</td>
<td></td>
</tr>
<tr>
<td>Nurse</td>
<td>64, (95.5%)</td>
</tr>
<tr>
<td>Nurse manager</td>
<td>3, (4.5%)</td>
</tr>
<tr>
<td>Years in speciality/clinical area</td>
<td></td>
</tr>
<tr>
<td>&lt; 1 mnd</td>
<td>0, (0%)</td>
</tr>
<tr>
<td>1 mth - 2 years</td>
<td>8, (12%)</td>
</tr>
<tr>
<td>3-4 years</td>
<td>4, (6%)</td>
</tr>
<tr>
<td>&gt; 4 years</td>
<td>55, (82%)</td>
</tr>
<tr>
<td>Non-respondents</td>
<td>27, (29.3%)</td>
</tr>
</tbody>
</table>
5.1.2 Missing values

The number of non-responses is never higher than two per item. However, all the responses from the category “don’t know” were coded as missing. Thus, the number of missing values is for some items quite substantial. The percentage of missing (both DK responses and non-response) varies from 1- 28.3% on single items, and for the dimension “perceptions of hospital management”, the percentage of missing items is as large as 34.4%, when missing responses for all items are cumulated. Missing items have been excluded pairwise, and have hence been assessed separately in each item.

The items with the most missing values are the items from the dimensions “perceptions of hospital management”. Similarly, the other dimension with the highest number of missing values is the “perception of unit management” dimension. See table 7 for the number and percentage DK and non-response items, and total missing items.

<table>
<thead>
<tr>
<th>Item</th>
<th>Don’t Know (n, %)</th>
<th>Non-response (n,%</th>
<th>Total missing (n, %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I get adequate, timely information about events that might affect my work from hospital management</td>
<td>17, 25.3%</td>
<td>2, 3%</td>
<td>19, 28.3%</td>
</tr>
<tr>
<td>Hospital management is doing a good job</td>
<td>16, 23.9%</td>
<td>2, 3%</td>
<td>18, 26.3%</td>
</tr>
<tr>
<td>Problem personnel are dealt with constructively by our hospital management</td>
<td>13, 19.4%</td>
<td>2, 3%</td>
<td>15, 22.4%</td>
</tr>
<tr>
<td>Hospital management supports my daily efforts</td>
<td>9, 13.4%</td>
<td>2, 3%</td>
<td>11, 16.4%</td>
</tr>
<tr>
<td>Hospital management doesn't knowingly compromise patient safety</td>
<td>8, 11.9%</td>
<td>2, 3%</td>
<td>10, 14.9%</td>
</tr>
<tr>
<td>Unit management doesn't knowingly compromise patient safety</td>
<td>6, 8.9%</td>
<td>2, 3%</td>
<td>8, 11.9%</td>
</tr>
<tr>
<td>Problem personnel are dealt with constructively by our unit management</td>
<td>6, 8.9%</td>
<td>2, 3%</td>
<td>8, 11.9%</td>
</tr>
</tbody>
</table>

As one can see from table 7, there is a quite substantial number of missing values for items in the SAQ factors “perception of hospital management” and for two of the items in the “perceptions of unit management”. This could be due to the proximity between front line workers and the management in the organization, and will be discussed in following chapters.
5.2 Results SAQ

SAQ scores are presented as the percentage of positive responses for each dimension. Presented as reference values for this sample are results from Sexton et al (2006) and Deilkås and Hofoss (2008). The SAQ scores are not the main focus of this thesis, and are hence presented only briefly here. However, as one can see from figure 1, the results are quite similar. There is no value for the perception of hospital management from Sexton et al 2006. This is due to the different factor structure of SAQ in Norway, where the dimension “perception of management” have been divided into two dimensions. Table 8 provides more information on the mean, SD and Chronbach’s Alpha for each SAQ dimension, as well as for the two composite RN safety climate scales.

Figure 1 Percentage positive response to each dimension

Reliability of SAQ dimensions

All dimensions from SAQ demonstrated satisfying internal consistency as measured by the Chronbach’s Alpha. It ranged from .721 to .865, with the highest Chronbach’s Alpha for safety climate (see table 8). This result is similar to reference values from the literature (Sexton et al 2006, Deilkås and Hofoss 2008). Hence, the dimensions seem to be applicable for my sample as well.
Table 8 Number of respondents and results for each factor

<table>
<thead>
<tr>
<th>Factor</th>
<th>N</th>
<th>valid, missing</th>
<th>Mean</th>
<th>SD</th>
<th>Percentage positive</th>
<th>Chronbach's Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teamwork climate</td>
<td>61</td>
<td>7</td>
<td>75</td>
<td>16.2</td>
<td>76.1</td>
<td>.725 (n=6)</td>
</tr>
<tr>
<td>Safety climate</td>
<td>58</td>
<td>10</td>
<td>71</td>
<td>19.4</td>
<td>69.6</td>
<td>.865 (n=7)</td>
</tr>
<tr>
<td>Job satisfaction</td>
<td>64</td>
<td>4</td>
<td>79</td>
<td>19.5</td>
<td>82.9</td>
<td>.855 (n=5)</td>
</tr>
<tr>
<td>Stress recognition</td>
<td>63</td>
<td>5</td>
<td>73</td>
<td>20.5</td>
<td>74.4</td>
<td>.748 (n=4)</td>
</tr>
<tr>
<td>Perception of unit management</td>
<td>53</td>
<td>15</td>
<td>54</td>
<td>21.7</td>
<td>42.4</td>
<td>.832 (n=5)</td>
</tr>
<tr>
<td>Perception of hospital management</td>
<td>44</td>
<td>24</td>
<td>47</td>
<td>18.1</td>
<td>26.4</td>
<td>.788 (n=5)</td>
</tr>
<tr>
<td>Working conditions</td>
<td>62</td>
<td>6</td>
<td>58</td>
<td>18</td>
<td>55.5</td>
<td>.721 (n=4)</td>
</tr>
<tr>
<td>Rnsafety_I</td>
<td>62</td>
<td>6</td>
<td>66</td>
<td>18</td>
<td>66.2</td>
<td>.815 (n=4)</td>
</tr>
<tr>
<td>Rnsafety_II</td>
<td>61</td>
<td>7</td>
<td>65</td>
<td>16.5</td>
<td>68.4</td>
<td>.804 (n=5)</td>
</tr>
</tbody>
</table>

5.3 Principal component analysis

The principal component analysis (PCA) analysis was initially attempted with all the items from the SAQ instrument, and the safety climate items from the RN4CAST instrument. The aim was to investigate whether the factor structure of the SAQ could be replicated in this sample, and to see if the items from the RN4CAST survey were included in the hypothesized SAQ factors. All items were included in a principal component analysis with oblimin rotation and listwise exclusion of missing items. However, the correlation matrix from this analysis turned out not to be positive definite. This might be a result of the large amount of missing data in this data set, or due to the relatively small sample size.

Hence, an attempt to perform the analysis again, now with pairwise exclusion of cases, was made. The result was the same for this attempt as well. The small sample size and the large number of missing data for some items imply that a factor analysis cannot be performed on all items in the questionnaire.
Thus, an attempt with the dimensions from SAQ that were assumed to have the most similar items, and that correlated the most with the composite scales from RN, were included in a Principal Component Analysis. In this analysis 19 items, of which 12 items of the SAQ instrument and 7 items from the RN4CAST survey, were included.

As described in chapter 4.2, the first step in a factor analysis is to assess the sample adequacy. The results form the Kaiser-Meyer-Olkin test for sample adequacy was .844, and the Bartlett’s test of sphericity was significant. This implies an adequate sample, and the analysis was continued.

To assess how many factors should be retained, the factors’ eigenvalues were assessed. As can be seen from table 9, five factors had an eigenvalue higher than 1. According to the Kaiser criterion, only the factors with an eigenvalue higher than 1 should be retained. Thus, according to this criterion, five factors could be retained. These five factors accounted for 42.3 %, 6.8 %, 6.4 %, 6.2 % and 5.2 % of the variance respectively. The Kaiser criterion will often result in too many factors being retained. The scree plot was thus inspected (figure 2).

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial Eigenvalues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>1</td>
<td>8.044</td>
</tr>
<tr>
<td>2</td>
<td>1.300</td>
</tr>
<tr>
<td>3</td>
<td>1.227</td>
</tr>
<tr>
<td>4</td>
<td>1.182</td>
</tr>
<tr>
<td>5</td>
<td>1.004</td>
</tr>
</tbody>
</table>

Figure 2 Scree plot
As one can see from the scree plot (figure 2), there is a clear break after the first factor. This could imply that only one factor should be retained. This was further assessed using a parallel analysis (table 10). The results confirmed that the eigenvalue of only one factor exceeded the corresponding value for a randomly generated data matrix of the same size (n=19 variables, n=53 respondents). To assess if it was possible to perform a PCA analysis on only the items in the safety culture dimension from SAQ and the items from RN4CAST, a new parallel analysis were performed (n=15 variables, n=53 respondents). The results from this confirmed that it is not possible to extract more than one factor with the limited amount of observations present (see table 10). This excludes the possibility for assessing whether the factor structure in my sample is similar to the factor structure found in other SAQ validation studies, and to assess how the RN4CAST items would relate to the SAQ factors.

<table>
<thead>
<tr>
<th>Number of variables</th>
<th>19</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of subjects</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>Number of replications</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvalue</th>
<th>Random eigenvalue</th>
<th>SD</th>
<th>Eigenvalue</th>
<th>Random eigenvalue</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2,2781</td>
<td>.1557</td>
<td>1</td>
<td>2,0381</td>
<td>.1410</td>
</tr>
<tr>
<td>2</td>
<td>1,9643</td>
<td>.1008</td>
<td>2</td>
<td>1,7833</td>
<td>.0969</td>
</tr>
</tbody>
</table>

### 5.4 Reliability analysis

#### 5.4.1 Reliability of the composite scales from RN4CAST

Composite scales of some of the items from RN4CAST/HSOPSC have been created. As described in chapter 4, variables measured on a Likert-type scale can be treated as interval-level data if some assumptions are found to be met. Hence, the creation of the composite scales is in large part due to the possibility to treat the data as continuous measures and use parametric statistics in the data analysis. The scales are based on own my judgements about
the relationship between the items that make up the SAQ dimensions, and a suggestion from a researcher at the Knowledge Centre working on the same project this thesis is a part of. The two different scaled indexes will both be subject for the same analyses, to see if one scale is a better reflection of patient safety climate as measured by the RN4CAST items in relation to the SAQ dimensions than the other. As will be described, the composition of the indexes varies somewhat, which gives slightly different alpha scores and also correlation to the corresponding SAQ dimension. The two scales both demonstrated satisfying internal consistency as measured by the Chronbach’s Alpha (see tables 12 and 14). Thus, the scale valued highest will then be a matter of how one values internal consistency relative to correlation with the criterion variable one wants to compare the scale to. See the discussion for a more thorough elaboration of this topic.

**Composite safety climate scale nr 1: RNsafety_1**

The first composite scale is made up of the HSOPSC items “Staff feel like their mistakes are held against them”; “Staff feel free to question the decisions or actions of those with more authority”; “In this unit, we discuss ways to prevent errors from happening again” and “We are given feedback about changes put into place based on event reports”. This scale is hypothesized to correlate to the SAQ dimensions “safety climate”, as the items from the two instruments both addresses perceptions of how errors are treated, how feedback is given and how patient safety concerns are treated.

The relationships between the items in the scale have been assessed with correlation analysis. According to Ringdal (2007), one wants to find correlation coefficients of medium strength (.30-.60). As one can see from the inter-item correlation matrix (table 11), there is a positive but not too strong correlation between the items in the scale. The correlation coefficients range from .458-.591
Table 11: Inter-item correlations RNsafety_1

<table>
<thead>
<tr>
<th></th>
<th>Staff feel like their mistakes are held against them</th>
<th>Staff feel free to question the decisions or actions of those with more authority</th>
<th>In this unit, we discuss ways to prevent errors from happening again</th>
<th>We are given feedback about changes put into place based on event reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff feel like their mistakes are held against them</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff feel free to question the decisions or actions of those with more authority</td>
<td>.577</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In this unit, we discuss ways to prevent errors from happening again</td>
<td>.458</td>
<td>.517</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>We are given feedback about changes put into place based on event reports</td>
<td>.532</td>
<td>.483</td>
<td>.591</td>
<td>1.000</td>
</tr>
</tbody>
</table>

**Internal consistency**

Further, the correlation between the individual items and the composite scale has been assessed. The correlation is ranging from .627-.652 (see table 12). The Chronbach’s Alpha for the scale is .815, and as one can see from the table, it would not increase if any of the items in the scale were removed.

Table 12: Item-total correlation and Chronbach’s Alpha if item deleted

<table>
<thead>
<tr>
<th></th>
<th>Corrected Item-Total Correlation</th>
<th>Cronbach’s Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff feel like their mistakes are held against them</td>
<td>.627</td>
<td>.772</td>
</tr>
<tr>
<td>Staff feel free to question the decisions or actions of those with more authority</td>
<td>.631</td>
<td>.769</td>
</tr>
<tr>
<td>In this unit, we discuss ways to prevent errors from happening again</td>
<td>.634</td>
<td>.768</td>
</tr>
<tr>
<td>We are given feedback about changes put into place based on event reports</td>
<td>.653</td>
<td>.758</td>
</tr>
</tbody>
</table>

Chronbach’s Alpha scale: .815

Average correlation between items in scale: .63
**Composite safety climate scale nr 2: RNsafety_II**

The second composite scale, which also is hypothesized to correspond to the SAQ dimension “safety climate”, contains the items “Staff feel like their mistakes are held against them”; “Important patient care information is often lost during shift changes”; “Things fall between the cracks when transferring patients from one unit to another”; “In this unit, we discuss ways to prevent errors from happening again” and “We are given feedback about changes put into place based on event reports”. Here, one item, which regards the information flow between shifts, has been included. Further, one item included in the previously described scale is removed, namely “Staff feel free to question the decisions or actions of those with more authority”.

As can be seen from table 13, most items have a moderate correlation to each other, which is what one would want to find. One item however, demonstrates weak correlation to two other items, namely “Things fall between the cracks when patients are transferred from one ward to another”. These are highlighted in bold in the table. This item is from the HSOPSC dimension “Handoffs and Transitions”. This item relates more to communication between wards, which might be the reason for its low correlation to the items relating to communication about error, and the presence of a guilt-free environment.

<table>
<thead>
<tr>
<th></th>
<th>Staff feel like their mistakes are held against them</th>
<th>Important patient care information is often lost during shift changes</th>
<th>Things “fall between the cracks” when patients are transferred from one ward to another</th>
<th>We are given feedback about changes put into place based on event reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff feel like their mistakes are held against them</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Important patient care information is often lost during shift changes</td>
<td>.563 1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Things “fall between the cracks” when patients are transferred from one ward to another</td>
<td>.222 .414 1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In this unit, we discuss ways to prevent errors from happening again</td>
<td>.459 .462 .416 1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>We are given feedback about changes put into place based on event reports</td>
<td>.533 .455 .222 .589 1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Internal consistency**

The correlation between the individual items and the composite scale has been assessed also for this scale (see table 14). The correlation is fairly good, ranging from .606-.648. However, one item demonstrates a lower correlation, at .406. Chronbach’s Alpha for the scale is .796. The item demonstrating lowest correlation would also increase the Chronbach’s Alpha to .804 if the item were removed. This is the same item as having low correlation in the inter-item correlation matrix seen above (table 14). Hence, this item is removed in the following analyses.

<table>
<thead>
<tr>
<th>Corrected Item-Total Correlation</th>
<th>Cronbach's Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff feel like their mistakes are held against them</td>
<td>.606</td>
</tr>
<tr>
<td>Important patient care information is often lost during shift changes</td>
<td>.624</td>
</tr>
<tr>
<td>Things “fall between the crack” when patients are transferred from one ward to another</td>
<td>.406</td>
</tr>
<tr>
<td>In this unit, we discuss ways to prevent errors from happening again</td>
<td>.648</td>
</tr>
<tr>
<td>We are given feedback about changes put into place based on event reports</td>
<td>.612</td>
</tr>
</tbody>
</table>

**Chronbach’s Alpha scale**: .796

**Average correlation between items in scale**: .601

### 5.5 Concurrent validity assessment

Concurrent validity refers to the degree that the results from the instrument match results from an already validated instrument. This will be assessed using the Pearson’s correlation coefficient and the intraclass correlation coefficient ($ICC_2$)
5.5.1 Pearson's product-moment correlation coefficient

The relationship between patient safety climate measured by SAQ and by the two composite safety scales I constructed from RN4CAST was investigated using the Pearson's product-moment correlation coefficient, and Spearman's rho. Preliminary analyses were performed to assess violations of the assumptions of normality, linearity and homoscedasticity (constant variance). The results demonstrated a strong, positive correlation between the SAQ safety dimension and the two RN safety composite scales (see table 15). The correlation appears to be larger using the parametric alternative than the non-parametric. Hence, the chance of underestimating the strength of the relationship between the safety climate scales due to non-linearity seems to be small.

<table>
<thead>
<tr>
<th>Table 15: Spearman's rho and Pearson's correlation coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RNsafety_I</strong></td>
</tr>
<tr>
<td>SAQ safety climate</td>
</tr>
<tr>
<td>Sig. (2.tailed)</td>
</tr>
<tr>
<td>Pearson's correlation coefficient</td>
</tr>
<tr>
<td>Sig. (2.tailed)</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2.tailed)

The coefficient of determinant \( r^2 \) was then calculated for the two variables, using the Pearson's \( r^2 \) = .54 and \( r^2 \) = .61 This gives that the SAQ safety climate variable accounts for 54% and 61% of variability in RNsafety_I and RNsafety_II variables, respectively.

5.5.2 Intraclass correlation

The intraclass correlation coefficient was estimated in order to assess the degree of consistency between how safety climate is judged in the two instruments. The range for an intraclass correlation coefficient is similar to other coefficients, it can vary between 0.0 and.0. An ICC value close to 1 implies that the values in the two variables or scales are similar.
ICC2 SAQ safety climate and RNsafety_I

The intraclass correlation coefficient with a two-way random model, using single measures (ICC\textsubscript{2}), gave an intraclass correlation coefficient of .73 (95% C.I .583-.836). The consistency of raters are assessed, not the absolute agreement.

ICC2 SAQ safety climate and RNsafety_II

The same model was applied to the second scale from the RN4CAST survey, RNsafety_II. Here, the ICC2 for single measures were .78 (95% C.I .642-.864).

5.5.3 Assessment of agreement

Figure 3 Bland-Altman plot SAQ safety climate and Rnsafety_I

As can be seen from the two scatter plots above (figure 3 and 4), there is a quite evenly distribution of observations across the plots. The RNsafety climate scale II has a somewhat higher mean difference than the RNsafety_I (6.9 and 4.8, respectively). It seems that there is a
tendency to score the SAQ safety climate dimension somewhat higher than both the RNsafety composite scales (as the mean for both is a positive value). The limits of agreement for the two measures are quite high (-22.6 to 32.2 and -17.1 to 30.8, respectively). This could imply that the variation between individual scores of the two scales is somewhat high. However, as very few data points fall outside the limits of agreement, and as the data points are evenly distributed across the plots, the agreement between the two measurement instruments are found to be good.

In order to investigate whether variables assumed to predict safety climate in fact would do so in this sample as well, a regression analysis was attempted with the different scales as dependent variable and individual items from the survey as independent variables. However, as one can see from the p-p plot of standardized residuals (figure 5), there is a clear tendency for column distribution. This can imply that the assumption of continuity is not met when single items are used in the analysis, and a linear regression model cannot be performed on this data set.

Based on the assessment of internal consistency, reliability and concurrent validity in relation to the SAQ safety climate dimensions, the two scales have been validated for use as measures for safety climate, and are valid measures of nurses’ perceptions of patient safety climate in the RN4CAST survey. The level of agreement between the SAQ safety climate dimension and the two scales are reasonably good. One scale (RNsafety_I) demonstrates higher internal consistency and agreement with SAQ safety climate, whereas the other (RNsafety_II) are found to correlate to SAQ somewhat higher.
6 Discussion

In the following sections, the statistical findings will be discussed in relation to the research question presented in the introduction, and some general conclusions regarding the research question will be drawn. Then, the relationship between safety climate surveys and safety performance will be discussed in relation to the aims and results of this study. Finally, some concluding comments will be made on the second topic assessed in this thesis, based on the theoretical findings from the literature.

6.1 General conclusions

The initial aim of this study was to investigate the association between process variables measured as the number of hand hygiene situations performed satisfactorily (in the WHO’s Clean Care is Safer Care framework), and patient safety climate. However, due to the limitations imposed by the sample size, it was not possible to assess this further. Hence, the only assessment undertaken is whether items from the RN4CAST survey could be used as valid measures of Norwegian nurses’ perceptions of patient safety climate. The internal consistency and the concurrent validity of the two composite scales have been found to be satisfactory, as have the agreement in measurement between the criterion variable (SAQ safety climate) and the two composite scales.

However, the results come with a number of caveats. The sample size, instruments, and methods used all have challenges and shortcomings, which will be discussed in the following sections.

6.2 Methods and instruments

As the sample did not allow for an exploratory factor analysis, we were left with correlation analysis to assess whether there was an association between the items from the two surveys relating to patient safety climate. This gives a level of uncertainty regarding both whether the
SAQ factor structure would be replicated in this sample, and how the items from RN4CAST relate to the different dimensions in the SAQ.

Deilkås and Hofoss (2008) have distributed the SAQ in a Norwegian hospital using the same six-item Likert-type scale. They found a missing items rate at between 0-13%, with a mean of 2.9%. This is considerably lower than in this survey. As the purpose of this survey not was to assess the safety climate at the units, and as neither the dimensions “perceptions of unit management” nor “perceptions of hospital management” have been used in the statistical analyses, the problem of a vast number of missing items is not at considerable as it could have been. The reason for the high number of missing values in the dimension “perceptions of hospital management” is the response category “don’t know” (DK). The number of non-responses is never higher than two for any of the items. Thus, there was clearly a tendency to use the response option “don’t know” instead of making a judgment of the performance of the hospital management. The reason behind this reluctance to score the hospital management could be that it is perceived to be a remote entity, which has little perceived impact on the day-to-day activities at the unit level. It could also be that the frontline staff just does not have an opinion on, or enough information about, how the hospital management performs. Only 26.5% of the respondents gave positive responses to the dimension “perception of hospital management”, whereas 42.5 % gave positive responses to the dimension “perceptions of unit management”. The number of DK responses was lower for the latter dimension (n=15, 20%, and n=24, 34.4%, respectively). The unit management may be perceived as closer in the organisation, and there is likely more contact between the unit management and the frontline workers. This could be the reason both for the better assessment and the higher number of respondents reporting an opinion on these items.

6.2.1 Measurement caveats

The different rating scales in this survey could be a serious concern for the validity of the findings. For the purpose of comparing the two instruments, it would have been a great advantage to use the same rating scale in both. The items from the RN4CAST survey (from HSOPSC) have either 6 or 7.5% missing items. Here, responses are given along a 5-point Likert-type scale, and the sixth response option is not included. Hence, all missing values are non-responses. We do not know how respondents value the 5-point Likert-type scale in relation to the 6-point Likert-type scale prior to the RN4CAST items. Survey design and
response theory is a rather large field in psychology, which is outside the scope of this thesis to assess. However, the different rating scales should be kept in mind for the interpretation of the results. There could be systematic differences in how these scales are perceived when responding to the survey, and this is a shortcoming of the design of this survey. One cannot exclude the possibility of this being the cause of the somewhat different scores for the SAQ safety climate scale, and the two composite safety scales created. It could have been better to remove the sixth response option in this survey, as the purpose was to validate the items from RN4CAST, and to have all answers on the same scale could have reduced the level of uncertainty regarding the results somewhat. However, SAQ has been validated in Norway using the six-level response format, and it could be problematic to remove one response option without further assessment of how this could affect the results.

The way the instruments differentiate between unit, ward and hospital might be a problem, as we do not know what the respondents have in mind when they are responding to the items. The HSOPSC is divided into either unit or hospital-level dimensions. The generic version of the SAQ does not differentiate the dimensions in this way. Most of the items ask for perceptions of patient safety climate “here”. Hence, besides from the two dimensions of perceptions of unit and hospital management, one does not actually know what respondents are answering to. It could either be perception of unit or hospital-wide aspects of patient safety. This could be a second explanation for the differences in scores between the SAQ safety climate dimension and the two composite safety scales, as the items from HSOPSC does not contain the same level of uncertainty regarding the different levels in the organisation.

As described in section 3.6, an especially problematic aspect of using self-reported outcome measures collected in the same instrument as the safety climate variables relates to mono-method bias. This could be caused by an artificial covariance between the predictor variable and the criterion variable, as the respondents providing the measures of both variables are the same. In this study the most likely sources for mono method bias arises from the fact that the criterion variable and predictor variable is measured at the same point in time. Both SAQ and RN4AST items are included in a single questionnaire. To overcome this, one could have administered the different instruments at two points in time. This may have resulted in a lower response rate, however, and also risk the assessments to change. Due to the small sample size,
increasing the response rate is considered the most important aspect, and administration of the questionnaire once is chosen.

The respondents did not know that some items in the questionnaire would be used to validate other items. Hence, they had no reason to make inferences about how an eventual relationship between the items was assumed to be. Due to this fact, respondents’ implicit theories about relationships is not regarded a potential source for common method bias in this study. What could be a source of bias, however, is the respondents’ cognitive need to respond to items in a consistent way. The predictive and the criterion items relate to the same underlying construct. Unfortunately, as the methods for adjusting for mono method bias in large part involve factor analysis, it was not possible to adjust for this type of bias in this study. When interpreting the relationship between these scales, one should keep this in mind.

6.2.2 Outcome variables

One of the initial aims was to investigate the predictive validity of the different safety scales in relation to the outcome measures included in this survey. It would then have been possible to detect eventual differences between the SAQ safety climate scale and the two composite scales created, if one were better in predicting the outcome variables. However, as we would have to use single items from the SAQ instrument, and single outcome measures from RN4CAST, a linear regression analysis were not allowed for by the data available. Ideally, in order to treat the items as continuous data, there should have been more observations. The limitations of the sample size became increasingly evident when an attempt was made on regression analysis. The p-p plot and the scatter plot of the standardized residuals clearly indicated that the independent variables (which were single items) were not data measured on a continuous scale, as a distinct column pattern could be seen (see figure 4). Ideally, further investigation of hypothesized predictors for patient safety climate as well as the predictive validity of the different scales should have been carried out, for example using logistic regression to account for the ordinal level data.

However, even if we were able to use the nurse-reported outcome variables provided in this study, it might not have allowed us to draw valid conclusions about the predictive validity of the different scales. As previously described, a number of systematic measurement errors could be present when the predictive variable and the outcome variable are measured at the same point of time, using the same instrument.
Further, it might not be reasonable to use self-reported outcome measures in the analyses, neither for predictive validity nor for assessment of causality between safety climate and outcomes. Hard, clinical data from reliable sources would give the most accurate picture of the state of patient safety. However, as such high-quality reporting systems in many cases are lacking, or, as discussed, are present only for organisations as a whole, one would in many cases have to rely on less accurate and valid outcome measures, such as health-care worker-reported outcome measures.

The HSOPSC instrument, which is one of the most frequently used instrument for assessing safety climate, includes four outcome measures. The health-care workers are for example asked to give a score to the overall safety climate and to estimate how many errors have been reported. These self-reported outcome measures are then used in the analyses (see for example Sarac, Flin, Mearns and Jackson, (2011). In the RN4CAST survey, respondents are asked to score a number of different patient outcomes, such as the frequency of medication errors, the frequency of pressure ulcers, and how often various hospital-acquired infections occur. It could be problematic to use these assessments as measures of patient safety, and it should in any case be accompanied by a discussion of potential sources for mono method bias. It is not without problems to use self-reported outcome measures in quality and safety assessments. A number of aspects can result in unreliable outcome measures if this approach is used. Some have origins in psychological theory, where it is well known that underreporting of negative events is common (for an example in health care, see Thomas and Petersen, (2003). The same points discussed earlier about mono method bias apply for the use of outcome measures gathered in the same instrument and at the same point in time as explanatory variables. It could result in erroneous correlations if some of the common variance between these variables is in fact due to the measurement methods. While some call for the need of linking safety climate survey results to objective derived outcome measures (Blegen et al., 2009), none of the studies investigated here have discussed common method variance as a possible source of measurement error.

6.3 Safety climate and its relation to quality

It could easily be argued that quality of health care is something more than treating patients in a safe way. Safety is an important aspect of quality, but the nature of the relationship between these two terms is more contested. When do quality issues turn into safety issues? It is
important to have clear definitions and operationalisations if these terms are supposed to be used in research, but as of now there are not consensual agreements on these definitions. Nor is there agreement on what constitutes an adverse event, what an error is, and if harm needs to occur before the event is important (Lilford et al, 2003; Vincent, 2010).

As previously described, both the SAQ and the HSOPSC have been found to have good psychometric properties, their factor structure have been replicated in different samples (with minor modifications), they have been found to be reliable and valid in a number of countries and they are of the two most widely used instruments to measure health care workers’ perceptions of the safety climate they are a part of. However, few attempts to establish a link between the safety climate and safety outcomes, as measured either as patient outcomes, health care worker outcomes, or structural outcomes, have found a causal relationship between safety climate and outcomes. Does it matter if the questionnaires are reliable and valid, if patient safety climate cannot be linked to safety outcomes? Are they then useful? Many researchers have concluded their work by calling for more research on this topic, and as of now, these questions stand without an empirically derived answer.

### 6.3.1 Lack of explicit theoretical framework

As described in chapter 2, there is no consensus on which theoretical framework best explains the relationship between safety climate and safety outcomes. This implies that what is regarded the most important aspects of a good safety climate varies between studies. Hence, a number of explanatory variables and safety climate dimensions are included in the different studies. While it is possible to identify the main theoretical approach from the items included in the instruments, the lack of coherence can result in a fragmented and non-transparent body of research regarding the theoretical associations between safety climate and outcomes (Wakefield et al., 2010). This could be a reflection of the fact that the constitutive dimensions and most important facets of safety culture seems to not be fully understood. Very few of the studies have stated an hypothesized causal model that could contribute in explaining how this relationship actually is (Jackson, Sarac and Flin, 2010). As Hearld et al (2008) suggest, the lack of theory-driven research reduces the possibilities for the accumulation of knowledge and progression in the patient safety climate field. This lack of theoretic explicitness could hinder researchers and others working with patient safety improvements to effectively communicate
about strategies for safety improvements. It certainly can make this communication process less accurate and transparent.

6.3.2 Direct causal relationship, circular causation or just co-variation?

There are few studies that aim at directly investigating this link - how improving safety climate is assumed to improve patient outcomes (The Health Foundation, 2011a). Some studies even finds that improving certain process outcomes such as error reporting rate may well happen without a corresponding improvement in safety climate. It could be that the changes in patient outcomes are causing the change in safety climate, rather than the other way around. Further, there is possible that there is an interrelationship between patient and staff outcomes, and safety climate – they could all influence each other. (Ibid.). In many cases, while trying to establish a causal link between safety climate and patient outcomes, the result is limited to the establishment of simultaneously improvements of safety climate and the outcome in question. If we were able to continue with the initial aim of this thesis, it might be that the same criticism could have been applied to the results, as observations of hand hygiene and assessment of patient safety climate would have taken place simultaneously. Hence, we would only have been able to assess the number of opportunities for hand hygiene performed satisfactory, and the patient safety climate score. It would not have allowed us to assess an eventual causal relationship, as it would have been a cross-sectional study.

6.4 Strengths

The survey performed is, to my knowledge, the only study attempting to establish the validity of the RN4CAST questions that relates to patient safety culture. The HSOPSC has been validated in many settings, including Norway. Only one study aiming at assessing the relationship between SAQ and HSOPSC dimensions have been identified (Etchegaray & Thomas, 2012). They found that while the scores from the different instruments could be converted, they could not be used exchangeably. They focused their analysis on the different scales in the instruments, however, whereas the items included in the RN4CAST survey are not from one, single dimension of the HSOPSC instrument; rather it consists of items from
several dimensions (see chapter 3.2). This could be a strength in the sense that it includes items from HSOPSC that are similar to the SAQ safety climate items, but are not limited to the HSOPSC scale perceived to be most similar to the SAQ safety climate dimension. The assessment of the internal consistency and concurrent validity of a composite scale of these items have enabled researchers working with data from the RN4CAST survey to use these items as reliable measures of nurses’ perception of patient safety climate.

The response rate obtained in this survey is rather high, especially when one recognizes that this survey was distributed in the middle of the summer. This is in many aspects due to the efforts from the infection control nurse and the unit representatives, who did a great job at motivating the staff, and demonstrating the importance of completing the questionnaire. This is especially important due to the small sample in this study.

### 6.5 Limitations

As discussed in the previous sections, there are a number of shortcomings and caveats following the survey. We were not able to assess the initial aim for the thesis due to these limitations, nor were we able to perform a factor analysis. Hence, the statistical analyses have to a large degree relied on correlation methods. These methods are useful for measuring the general direction of relationships between variables. To account for this, an assessment of agreement was conducted. The results from this analysis further supported the conclusion that the two scaled indexes are valid measures of patient safety climate. However, this analysis, similar to the correlation analysis, requires that the variables be measured on a continuous scale. As demonstrated when attempts on linear regression analysis were made, this assumption might not be valid, and certainly not when single items are used as independent variables. To overcome the problem of continuity, the composite scales have been created. While the comparison of the parametric and non-parametric correlation coefficients contributed to the conclusion that it was justified to regard the scaled indexes as continuous level data, it nonetheless provides a level of uncertainty regarding the results. To overcome this, more research could be undertaken with a larger sample, or using categorical (logistic) regression analysis.
7 Concluding remarks

Based on results from the statistical analyses performed, and on the face validity assessment of the items composing the RN4CAST safety climate scales, it is judged that the items from RN4CAST that relates to patient safety aspects can be used as measures of patient safety climate, if they are combined into a composite scale. The two different scales created here are both reliable and demonstrates good internal consistency. Which one is chosen is dependent on how one makes the trade-off between correlation to the criterion measure, and internal consistency as measured by Chronbach’s Alpha. Based on the assessment of internal consistency, reliability and concurrent validity in relation to the SAQ safety climate dimensions, the two scales have been validated for use as measures for safety climate, and are valid measures of nurses’ perceptions of patient safety climate in the RN4CAST survey.

Regarding the second topic investigated in this thesis, it has been difficult to find conclusive evidence that improvement in safety climate will result in improved safety outcomes. It is proposed that the relationship in fact is not directly causational, but more complex, with all the variables influencing on each other. However, it may be that due to lack of coherence in definitions and terms used to assess and measure safety climate, the number of different instruments used for measurement, and the vast number of safety outcomes that have been used as end points, one has not been able to demonstrate this causal relationship. It have been said before, but more research on this topic is required if the link between safety climate and safety outcomes is supposed to be identified and understood. To develop a more coherent and consensual theoretical framework will be an ideal starting point, and research in this area is needed. Further, more research on the link between both processual, structural and outcome variables could be required if a coherent framework for patient safety climate and outcomes are to be developed.
References


lov om behandling av personopplysninger (LOV-2000-04-14-31).

Lov om medisinsk og helsefaglig forskning (helseforskningsloven) (LOV-2008-06-20-44).


Appendices

Appendix I: Approval from the Norwegian Social Science Data Services

Ingeborg Strømseng Sjetne
Nasjonal kunnskapscenter for helsetjenesten
Postboks 7004 St. Olavs plass
0130 OSLO

Dato: 26.04.2010
Vår ref: 22537 KOHLR
Deres dato: Deres ref:

ENDRET JEMMELSGRUNNLAG FOR BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til endringsmelding mottatt 10. mars 2010 samt oppklarende e-postkorrespondanse og telefonsamtaler med prosjektleder Ingeborg Strømseng Sjetne og prosjektmedarbeider Christine Tvedt vedrørende prosjektet 22537 "Personell og pasientsikkerhet"

Endringen går ut på at det nå skal samles inn ytterligere opplysninger i forbindelse med prosjektet. Følgende tillegg opplysninger skal samles inn:

a) Registerdata fra Norsk pasientregister (NPR)
b) Registerdata fra Norsk overvåkningssystem for infeksjoner i helsetjenesten (NOIS)
c) Administrative opplysninger om de aktuelle sykehusene

I tillegg skal det gjennomføres et delprosjekt som omfattes av hovedprosjektets formål. I forbindelse med delprosjektet skal følgende opplysninger sammes inn:

d) Registerdata fra Nasjonal register for leddprotester (NRL) som er aggregert på sykehusnivå
e) Et registretningskjema der opplysninger om helsearbeideres etterlevelse av retningslinjer for håndhygiene registreres
f) Et spørreskjema med spørsmål om pasientsikkerhetskultur som fylles ut av helsearbeidere
g) Et spørreskjema med spørsmål av administrativ art som fylles ut av en representant ved hver av de aktuelle ortopediske enhetene

Opplysninger fra de ulike datakildene a-g skal ikke Kobles på individnivå, jf., e-post fra Christine Tvedt 15. april 2010. Personvernombudet forstår det slik at det bare er opplysninger fra følgende datakilder som ikke vil være anonyme:

a) Registerdata fra Norsk overvåkningssystem for infeksjoner i helsetjenesten (NOIS)
f) Et spørreskjema om pasientsikkerhetskultur som fylles ut av helsearbeidere

Ombudet legger til grunn at de øvrige opplysningene som innhentes er anonyme, og dermed ikke omfattes av mekylektitten i henhold til personopplysningsloven. Vi minner om å anmelde opplysninger om anonyme opplysninger om opplysninger som det ikke er mulig å føre tilbake til enkeltpersoner, hverken
direkte (via for eksempel navn eller fødselsnummer eller referanse til slike opplysninger) eller indirekte (via sammenstilling av bakgrunnsopplysninger som navn på institusjon, kjønn, alder, diagnose, stilling eller lignende).


På bakgrunn av ovennevnte momenter finner personvernombudet at samfunnets interesse i å behandle data finner sted klart overstiger ulempene den kan medføre den enkelte, og at behandlingen av opplysninger som utleveres fra NOIS kan hjemles i personopplysningsloven §§ 8 d) og 9 h). Grunnet utvalgssituasjonen og at prosjektleder ikke har tilgang til direkte personidentifiserbare opplysninger, anses det som uforholdsmessig vanskelig å gi informasjon til utvalget, jf. personopplysningsloven § 20 b).

Christine Tvødt opplyser i e-post 22. april 2010 at det vil bli sendt søknad til REK vedrørende dispensasjon fra tausehtsjekken. Ombudet legger til grunn at søknaden innvilges og ber om at svar fra REK etter sendes når det foreligger.


På bakgrunn av at det nå blir behandlet sensitive personopplysninger i forbindelse med prosjektet, finner ombudet at behandlingen av personopplysninger vil være regulert av personopplysningsforskriften § 7-27. Personvernombudet tiltråd at prosjektet gjennomføres. Det vises også til brev fra personvernombudet med ombudets opptrinnelige prosjektvurdering, datert 25. september 2009, og ombudet forutsetter at prosjektutkastet for øvrig er uendret.

Senest innen prosjektutslutt, 31. desember 2011, skal datamaterialet anonymiseres ved at direkte og indirekte identifiserende personopplysninger slettes eller omskrives/kategoriseres slik at det ikke lenger er mulig å fore opplysningene tilbake til enkeltpersoner.

Vennlig hilsen

Bjørn Henrichsen

Kontaktperson: Ragnhild.Haugland@nsd.uib.no, tlf.: 55 58 83 34
Hei,

Personvernombudet mottok 11.05.2011 endringsmelding angående prosjekt 22537 "Personell og pasientsikkerhet".

I følge endringsmeldingen gjaldt endringen at det var blitt gjort noen endringer angående punkt f) (Med henvisning til brev sendt fra personvernombudet 26.04.2010).

I følge endringsmeldingen gikk endringen ut på at det nå ville bli benyttet et annet spørreskjema, utvalget var blitt mindre, i tillegg til noen få endringer i informasjonsskrivet.

Gjennom saksbehandlingen fremkom det at:

- Spørreskjemaet i sin realitet er det samme, jf. telefonsamtale med forsker 14.06.2011 og 15.06.2011
- Utvalget er endret fra ca. 800 personer til ca. 300 personer.

Forsker opplyser på telefon 15.06.2011 at det vil kunne bli tilknyttet en prosjektmedarbeider, masterstudent, til prosjektet.

Personvernombudet har med dette registrert endringene i delprosjekt f (jf. brev fra personvernombudet 26.04.2011).

Personvernombudet minner om at øvrige godkjenninger i prosjektet (for dette delprosjektet og de øvrige delprosjektetene) fremdeles er gjeldende.

Vennlig hilsen

Marte Bertelsen
Fagkonsulent

Norsk samfunnsvitenskapelig datatjeneste AS
Personvernombud for forskning
Harald Hårfagres gate 29, 5007 BERGEN

Tlf. direkte: (+47) 55 58 33 48
Tlf. sentral: (+47) 55 58 21 17
Faks: (+47) 55 58 96 50
E-post: Marte.Bertelsen@nsd.uib.no
www.nsd.uib.no/personvern
Spørreundersøkelse om pasientsikkerhetskultur

Invitasjon til å delta i en spørreundersøkelse om pasientsikkerhetskultur og informasjon om undersøkelsen.

Med dette inviteres du til å besvare en spørreundersøkelse om pasientsikkerhetskultur som gjennomføres blant personale på to posten ved to sykehus, med til sammen 250-300 ansatte.

Spørreundersøkelsen ingår i forskning om pasientsikkerhet og sammenhenger mellom organisatoriske kjennetegn ved sykehusene, pasientsikkerhetskultur og praksisvilkår for helsearbeidere i norske sykehus. Det endelige målet er å utvikle kunnskap som kan danne grunnlag for målrettede intervensjoner for å styrke pasientsikkerhet i sykehus.

Svarerne fra undersøkelsen vil bli koblet til data fra andre undersøkelser på sykehuset og/eller posten og annet materiale Kunnskapssenteret har tilgang til, men ikke på individnivå. Posten/sykehuset skal være anonym i presentasjon av resultater.

Personvern og anonymitet
Spørreskjemaene er merket med et referansenummer for å koble opplysningene til post og sykehus. Idet svarkonvolutten returneres til Kunnskapssenteret vil opplysningene være anonyme.


Hvert eneste svar er viktig
Det er frivillig å delta, men vi oppfordrer alle til å svare og levere inn skjema. Skjemaene skal besvares individuelt. Høy svarprosent er viktig for å få god ytte av datamaterialet nå og i fremtiden.

Framgangsmåte

Dersom du ønsker mer informasjon om undersøkelsen er du velkommen til å kontakte undertegnede.

Christine Tvedt
cht@nokc.no

Tlf: 46 41 97 90
Pasientsikkerhetskulturen her hvor jeg arbeider

Når du fyller ut skjemaet, gjør det ut fra dine erfaringer i det kliniske området. Velg mellom disse svaralternativene:

Svært uenig  Litt uenig  Naytral  Litt enig  Svært enig  Vet ikke

Spørsmål:

1. Innspill blir godt mottatt her

2. Her er det vanskelig å gi ifra om jeg oppdager ett problem i pasientbehandlingen

3. Her blir uenighet håndtert riktig (dvs ikke ut fra hvem som har rett, men ut fra hva som er best for pasienten)

4. Jeg får den støtte jeg trenger fra andre sykehusansatte for å ta meg av pasientene

5. Her er det lett for ansatte å spare når det er nødvendig

6. Legene og sykepleierne* her samarbeider som et velkoordinert team (*kan byttes ut)
<table>
<thead>
<tr>
<th>Nr.</th>
<th>Sats</th>
<th>Svært uenig</th>
<th>Litt uenig</th>
<th>Nøytral</th>
<th>Litt enig</th>
<th>Svært enig</th>
<th>Vet ikke</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.</td>
<td>Jeg ville føle meg trygg hvis jeg var pasient her</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Her blir medisinske feil håndtert riktig</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Jeg vet hvilke kanaler jeg skal bruke for å stille spørsmål om pasientsikkerhet her</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>Jeg får passende tilbakemelding om arbeidet mitt (korrekt, tilstrekkelig og fra rett person)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>Det er vanskelig å diskutere feil her</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Kolleger oppmuntrer meg til å si fra om enhver sikkerhetsbekymring som jeg måtte ha</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>Kulturen her gjør det lett å lære av andres feil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td>Mine forslag om sikkerhet ville bli behandlet om jeg la dem fram for ledelsen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.</td>
<td>Jeg liker jobben min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.</td>
<td>Å arbeide her er som å være del av en stor familie</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.</td>
<td>Dette er et godt sted å arbeide</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>18.</td>
<td>Jeg er stolt av å arbeide her</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>19.</td>
<td>Innsatsviljen er stor her</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>20.</td>
<td>Når arbeidsbelastningen min blir for stor, arbeider jeg dårligere</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>21.</td>
<td>Jeg er mindre effektiv når jeg er sliten</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>22.</td>
<td>Det er mer sannsynlig at jeg gjør feil når situasjonen blir ansprent eller fiendtlig</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>23.</td>
<td>Jeg arbeider dårligere i kri sesituasjoner (som resuscitering, anfall o.l.) når jeg er sliten</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>24a.</td>
<td>Ledelsen her i enheten legger forholdene til rette for mitt daglige arbeid</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>24b.</td>
<td>Sykehusledelsen legger forholdene til rette for mitt daglige arbeid</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>25a.</td>
<td>Ledelsen her i enheten gjør aldri noe som de vet kan gå ut over pasientsikkerheten</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td></td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Nøytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-------------</td>
<td>------------</td>
<td>---------</td>
<td>-----------</td>
<td>------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>25. Sykehusledelsen gjør aldri noe som de vet kan gå ut over pasientsikkerheten</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. Ledelsen her i enheten gjør en god jobb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Sykehusledelsen gjør en god jobb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. Ledelsen her i enheten gir meg adekvat informasjon, i rett tid, om hendelser som kan ha betydning for mitt arbeid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. Sykehusledelsen gir meg adekvat informasjon, i rett tid, om hendelser som kan ha betydning for mitt arbeid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. Bemanningsnivået her er tilstrekkelig til at vi kan ta oss av det antall pasienter vi har</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31. Dette sykehuset gjør en god jobb med å lære opp nyansatte</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32. All den informasjon som jeg trenger til diagnose og terapeutiske beslutninger er rutinemessig tilgjengelig for meg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32.</td>
<td>De som er under opplæring på mitt fagfelt, får tilstrekkelig tilsyn</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>33.</td>
<td>Samarbeidet med sykepleierne er godt i denne enheten</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>34.</td>
<td>Samarbeidet med legene er godt i denne enheten</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>35.</td>
<td>Samarbeidet med farmasøytene er godt i denne enheten</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
<tr>
<td>36.</td>
<td>Kommunikasjonssvikt som fører til forsinkelser i pasientbehandlingen er vanlige</td>
<td>Svært uenig</td>
<td>Litt uenig</td>
<td>Naytral</td>
<td>Litt enig</td>
<td>Svært enig</td>
<td>Vet ikke</td>
</tr>
</tbody>
</table>

### Bakgrunnsoppspørsmål

**Hva er din stilling?**
(Sett kun ett kryss)

- Overlege/avd. overlege
- Lege i spesialisering
- Turnusleger
- Hjelpemidler/synepleier
- Avd. sykepleier/synepleiers lederansvar
- Off. godkjent sykepleier/Jordmor
- Bloinjen
- Radiograf
- Psykoterapeut
- Annet

**Kjønn:** Er du mann eller kvinne?

- Mann
- Kvinne

**Hvor lenge har du arbeidet i enheten/fagområdet?**

- Mindre enn 1 mnd
- 1 mnd - 2 år
- 3-4 år
- 4 år eller mer

**Skriv gjørne ned kommentarer vedrørende dine erfaringer med pasientsikkerhetskultur der du arbeider**

---

Copyright Bryan Sexton, Johns Hopkins University
Noen tilleggs-spørsmål til sykepleiere som deltar i undersøkelsen:

A. Generelt sett, hvordan vil du beskrive kvaliteten på sykepleien som blir gitt til pasientene på din post?
   - [ ] Lite bra
   - [ ] Noenlunde bra
   - [ ] Bra
   - [ ] Utmerket

B. Hvor sikker er du på at pasientene dine klarer seg selv etter utskrivelse?
   - [ ] Slott ikke sikker
   - [ ] Ganske sikker
   - [ ] Sikker
   - [ ] Svært sikker

C. Hvor sikker er du på at foretaksledelsen vil iverksette tiltak for å løse problemer angående pasientbehandlingen som du rapporterer?
   - [ ] Slott ikke sikker
   - [ ] Ganske sikker
   - [ ] Sikker
   - [ ] Svært sikker

D. Gi en generell vurdering av pasientsikkerheten i din avdeling:
   - [ ] Meget dårlig
   - [ ] Dårlig
   - [ ] Akseptabel
   - [ ] Meget god
   - [ ] Fremragende

E. I løpet av det siste året, vil du si at kvaliteten på pleie og behandling i ditt sykehus ...
   - [ ] er blitt dårligere
   - [ ] er uendret
   - [ ] er blitt bedre

F. De neste spørsmålene ber om din mening om pasientsikkerhetsaspekter i arbeidssituasjonen din:

1. Ansatte føler at feil blir brukt mot dem................................................................. [ ]
   - [ ] Heilt uenig
   - [ ] Uenig
   - [ ] Både - og
   - [ ] Enlig
   - [ ] Heilt enlig

2. Informasjon som er viktig i pasientbehandlingen går ofte tapt ved vaktskifte................................................................. [ ]
   - [ ] Heilt uenig
   - [ ] Uenig
   - [ ] Både - og
   - [ ] Enlig
   - [ ] Heilt enlig

3. Ting "faller mellom to stoler" når pasienter blir overflyttet fra en avdeling til en annen................................................................. [ ]
   - [ ] Heilt uenig
   - [ ] Uenig
   - [ ] Både - og
   - [ ] Enlig
   - [ ] Heilt enlig

4. Ansatte kan fritt slette spørsmål vedrørende beslutninger og handlinger tatt av personer med mer autoritet................................................................. [ ]
   - [ ] Heilt uenig
   - [ ] Uenig
   - [ ] Både - og
   - [ ] Enlig
   - [ ] Heilt enlig

5. I denne enheten diskuteres vi hvordan vi kan forebygge at de samme hendelsene gjentas................................................................. [ ]
   - [ ] Heilt uenig
   - [ ] Uenig
   - [ ] Både - og
   - [ ] Enlig
   - [ ] Heilt enlig

6. Vi får tilbakemeldinger om endringer som blir igangsatt basert på rapporterte uanskåde hendelser................................................................. [ ]
   - [ ] Heilt uenig
   - [ ] Uenig
   - [ ] Både - og
   - [ ] Enlig
   - [ ] Heilt enlig

7. Foretaksledelsens handlinger viser at pasientsikkerheten har topp prioriteter................................................................. [ ]
   - [ ] Heilt uenig
   - [ ] Uenig
   - [ ] Både - og
   - [ ] Enlig
   - [ ] Heilt enlig

G. Hvor ofte vil du si at de følgende hendelsene skjer med dine pasienter?

<table>
<thead>
<tr>
<th></th>
<th>Alte</th>
<th>Noon ganger i året eller mindre</th>
<th>En gang i måneden eller mindre</th>
<th>Noon ganger i uka</th>
<th>En gang i uka</th>
<th>Hver dag</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Pasient fikk feil medisin, feil dose eller til feil tidspunkt</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>2. Trykkårs oppstår under sykhusinnleggselven</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>3. Pasientfall med skade</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>4. Sykhusinfeksjon</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>4.1. Urinvisinfeksjon</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>4.2. Blodforfalling/Sepsis</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>4.3. Pneumoni</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>

Takk!