An Exploration of Contributing Factors to Patient Safety and Adverse Events

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy
In the Graduate School of The Ohio State University

By

Inga Mirdza Zadvinskis, M.S.N.

Graduate Program in Nursing

Ohio State University

2015

Dissertation Committee:

Pamela J. Salsberry, Ph.D., RN, Advisor

Laura A. Szalacha, Ph.D.

Emily S. Patterson, Ph.D.

Esther M. Chipps, Ph.D., RN

Abstract

More than 400,000 premature deaths per year occur due to preventable harm in U.S. hospitals, costing over $20 billion per year in healthcare expenses, lost worker productivity, and disability. Conceptual frameworks, such as the Generic Reference Model, contribute to a greater understanding of patient safety because they explain the context of patient harm. The healthcare context, including organizational factors such as strong safety culture and human factors like teamwork, may improve patient outcomes. Patient outcomes, such as adverse events, are more readily detected using instruments such as the Institute for Healthcare Improvement (IHI) Global Trigger Tool (GTT), which may detect up to ten times more adverse events than existing methods. The GTT uses keywords or triggers to guide chart reviews. Currently, relationships between safety culture and teamwork and adverse event detection using trigger-tools remain underexplored. The purpose of this study was to explore relationships between organizational and human factors with adverse events that result in patient harm detected using a modified trigger-tool methodology. The descriptive, cross-sectional design used the Safety Attitudes Questionnaire (SAQ) to measure interprofessional staff perceptions of safety culture using safety climate and teamwork climate subscales, and a retrospective, modified IHI GTT chart review methodology to measure patient outcomes at the unit level. The convenience sample was comprised of 32 nursing units/departments from one 750+-bed Midwestern U.S. regional acute care hospital that employed over 1000 nurses. Safety and teamwork climate percentage agreement averages were 75.61% and 70.07%, respectively. Medical surgical units reported the strongest safety climate whereas critical care units reported the strongest teamwork. An average of 69 adverse events occurred per 1,000 patient days, 21.83 adverse events per 100 admissions, and approximately 20% of admissions experienced an adverse event. The most frequently occurring adverse event was nausea. Medical surgical units experienced the greatest frequency

of adverse events compared to procedural, critical care, intermediate care, and OB gynecology unit types. Three GTT triggers had positive predictive values of 100%: healthcare-associated infections, injury/repair/removal of an organ, and over-sedation/ hypotension. Safety climate and teamwork did not have a statistically significant effect on frequency of GTT-identified adverse events. Unit type predicted about 30% of the variance in adverse events. This study provides preliminary evidence that researchers may use the GTT to detect unit-level adverse events. The GTT identifies adverse events not detected via other methods (such as nausea), and these adverse events affect patient outcomes, cost of care, and quality. The Generic Reference Model contains many contributing factors to patient safety, which were unmeasured, and these gaps provide opportunities for future research.

Dedication

This document is dedicated to my family. Thank you to my grandparents for having the courage to flee Soviet-occupied Latvia so that I could enjoy freedom, health, and education in the United States. This work honors them and the sacrifices they made. For my godmother Mirdza Gailitis, the first nurse I ever knew, thank you for sharing your faith. Mom and Paps, thank you for instilling in me the importance of education, integrity, and pride in our Latvian heritage, and for your love during the most difficult times. To my sister, Ivete, thank you for being such a good listener and dear friend today and always. Your family brings me so much joy! For Lukas, Nahuel, and Mateo, follow your dreams, as this journey was mine. Thank you, Pauline and Zigfrids Zadvinskis, for your dedication to family and philanthropy, which inspires people to be better human beings. Thank you to my terrific friends for making me laugh when I needed it most. Lastly, thank you, David, my incredible and supportive husband, for being so understanding during these 5 years of studies. You bring out the best in people. Without you, I could not have done this.

Acknowledgments

I am grateful for the support and encouragement that I have received during my doctoral education. Thank you, Dr. Pamela Salsberry, for being my advisor and encouraging me to develop essential skills in independent and creative thinking. I would like to thank my dissertation committee comprised of Drs. Szalacha, Patterson, and Chipps for their unique content expertise and guidance. Dr. Laura Szalacha generously shared her time and sense of humor to improve my understanding of statistics. Meetings with Dr. Emily Patterson were stimulating brainstorming sessions, and she intuitively knew when to provide extra encouragement throughout the process. Dr. Esther Chipps generously shared her research talents and exceptional contacts, and I am thankful for her mentorship and friendship. I was fortunate to experience an extraordinary research residency under the guidance of Dr. Po-Yin Yen, who opened my eyes to the world of nursing informatics and qualitative research. Dr. David Woods is one of the world's preeminent scholars in cognitive science engineering, and he taught me to think about patient safety and health care from a unique perspective. His enthusiasm for teaching made his class among my most memorable. My statistics professors, Dr. Richard Lomax and Dr. Jessica Logan, provided the foundation for analytic techniques that I will use throughout my research career. My Ph.D. student colleagues provided insights and inspiration throughout the program. I also would like to thank Kathy Crea, Pharm.D, for translating the textbook principles of patient safety to practice so seamlessly, and for all her donated time reviewing medical records. Dr. Darrell Spurlock epitomizes excellence in research, scholarship, and graduate nursing education. I am indebted to OhioHealth for awarding me scholarship funds for tuition and for their schedule flexibility during my program. Thank you, Ohio State University College of Nursing (CON), for the professional development opportunities including poster and podium presentations and interprofessional seminars that helped me grow as a nurse scholar. It was a privilege to be a student of such a

distinguished faculty. Thanks to all of you that have taught me from the first day of class to the last post-candidacy seminar: Dr. Pamela Salsberry, Dr. Edna Menke, Dr. Karen Ahijevych, Dr. Victoria von Sadovsky, Dr. Donna McCarthy, Dr. Celia Wills, Dr. Esther Chipps, Dr. Margaret Clark Graham, Dr. Barbara Polivka, Dr. Jodi Ford, Dr. Elizabeth Lenz, Dr. Laureen Smith, and Dr. Usha Menon. Finally, thanks to Drs. Gretchen Glasgow and Jeanette Chambers for encouraging me to pursue the Ph.D. in Nursing.

Vita

May 1995 ..... B.S.N., Michigan State University

May 2000 ..... M.S.N., Wayne State University

Publications

Buck, J., Chipps, E., Knupp, A. M., & Zadvinskis, I. M. (2015). Improving research quality in the area of patient safety: A call to action. In S. P. Stawicki, S. C. Galwankar, T. J. Papadimos, & S. D. Moffatt-Bruce (Eds.), *Fundamentals of patient safety in medicine and surgery* (pp. 303-308). New York, NY: Lippincott, Williams, and Wilkins.

Zadvinskis, I. M., Chipps, E., & Yen, P. Y. (2014). Exploring nurses' confirmed expectations regarding health IT: A phenomenological study. *International Journal of Medical Informatics, 83*(2), 89-98.

Zadvinskis, I. M., Wills, C. E., Chipps, E., & Patterson, E. (2013). Process measurement in safe patient handling and movement: A literature review. *American Journal of Safe Patient Handling and Movement, 3*, 55-63.

Field of Study

Major Field: Nursing

Table of Contents

Abstract ii

Dedication iv

Acknowledgments v

Vita vii

List of Tables xiii

List of Figures xvi

Chapter 1: Introduction 1
- Problem Statement 2
- Statement of Purpose 2
- Specific Aims 2
- Conceptual Framework for the Study 3
- Definition of Terms 3
- Assumptions 3
- Overview of Methodology 4
- Rationale and Significance 4
- Limitations 4
- Organization of the Dissertation 5

Chapter 2: Literature Review 6
- Causes of Adverse Events 7
- Relationship Between Safety and Teamwork With Adverse Events 7
  - Safety Culture/Climate and Adverse Events 8
    - Studies with significant relationships 8
    - Studies with non-significant relationships 8
  - Teamwork and Adverse Events 9
    - Studies with significant relationships 9
    - Studies with non-significant relationships 10
- Theoretical Approaches to Study Patient Safety 11
  - James Reason's Swiss Cheese Model 11
  - High Reliability Theory (HRT) 12
  - The Generic Reference Model (GRM) 13
- Rationale for the Selected Theoretical Approach 15
- GRM Contributing Factors and Hazards 16
  - Environmental Factors 16
  - Organizational Factors 17
    - Safety culture 17
    - Job satisfaction 18
    - Work stress 19
    - Management commitment 19
    - Resources 19

Environmental conditions 20
Human Factors 20
Teamwork 21
Health care worker behavior and performance 22
Communication 22
Subject of Incident Factors 23
Drugs, Equipment, and Documentation 24
Drugs 24
Equipment 25
Documentation 26
Electronic health records (EHRs) 26
Safety Incident 28
Outcomes and Consequences 28
Application of the GRM to This Study 29
Organizational Factors: Safety Climate 30
Measurement of Safety Climate 31
AHRQ Hospital Survey on Patient Safety (HSOPS) 31
Patient Safety Climate in Healthcare Organizations (PSCHO) 32
Safety Attitudes Questionnaire (SAQ) 33
Human Factors: Teamwork 35
Teamwork Culture Versus Teamwork Climate 35
Measurement of Teamwork Climate 36
Confounding Factors (Unit Characteristics) 37
Patient Outcomes: Adverse Events 38
Adverse Event Risk Factors 39
Preventability of Adverse Events 39
Measurement of Adverse Events 40
Voluntary error reporting systems 40
Administrative data analysis 41
Retrospective chart review 42
Measurement of Adverse Events Using the Global Trigger Tool (GTT) 43
GTT Triggers 43
Adaptations to the GTT 44
Validity of the GTT 47
Reliability of the GTT 48
GTT studies with Kappa .01-.20 (slight agreement) 48
GTT studies with Kappa .41-.60 (moderate agreement) 48
GTT studies with Kappa .61-.80 (substantial agreement) or 0.81-0.99 (almost perfect agreement**)** 49
Other methods for determining GTT reliability 49
Advantages of the GTT 49
Limitations of the GTT 50
Summary 50

Chapter 3: Methods 52
Design 52
Sample 53
Sample Size 53
Inclusion and Exclusion Criteria for Nursing Units/Departments 54
Inclusion and Exclusion Criteria for Patient Records 55
Measurement/Instrumentation 56
Safety Climate 57
Teamwork Climate 57
SAQ validity 59
Content validity 59

Construct validity 59
SAQ reliability 59
Confounding Variables: Unit Characteristics 60
Adverse Events 60
Global Trigger Tool 60
GTT validity 63
GTT reliability 63
Detailed Study Procedures 65
Independent Variables: Safety and Teamwork Climate 65
Confounding Variable: Unit Characteristics 65
Dependent Variable: Frequency of Adverse Events 65
Procedure 65
Accuracy of review 67
Data Analysis 69
Aim 1 69
Aim 2 70
Aim 3 70
Limitations 71
Internal Validity 71
External Validity 72

Chapter 4: Results 73
Descriptive Statistics 73
Control Variables: Unit Characteristics 73
Predictors: Teamwork Climate 75
Reliability 75
Teamwork climate as a continuous variable 75
Dichotomized teamwork climate 76
Predictors: Safety Climate 77
Reliability 77
Aim 1 78
Frequency of Teamwork Climate and Safety Climate by Unit Type 78
Teamwork climate by unit type 78
Safety climate by unit type 80
Unit size by unit type 80
Unit size by dichotomized teamwork climate 82
Unit size by safety climate 83
Aim 2 84
Dependent Variable: Frequency of Adverse Events 84
Frequency of Adverse Events by United Type 86
Aim 3 90
Bivariate Tests 91
Correlations 91
Unit size with safety climate, teamwork climate, and frequency of adverse events 91
Safety climate and teamwork climate 91
Safety climate with frequency of adverse events 91
Teamwork climate with frequency of adverse events 95
Teamwork by united type (medical surgical or other unit type) 96
Unit type and relationship with dependent variable frequency of adverse events: Results from a one-factor analysis of variance (ANOVA) fixed effects model 98
Testing of assumptions for analysis of variance 98
Analytic Strategy 105
Relationship between unit type with predictors dichotomized teamwork climate and safety climate: Results of an independent *t*-test 105

Relationship between low and high teamwork units and frequency of adverse events: Results from an independent *t*-test 107
Relationship between unit type and dichotomized teamwork climate with frequency of adverse events: Results from a two-factor analysis of variance (ANOVA) fixed effects model 109
Homogeneity of variance 109
Relationship between unit type, dichotomized teamwork climate and safety climate with frequency of adverse events: Hierarchical multiple regression results 111
Evaluation of assumptions 112
Normality 112
Outliers, linearity, homoscedasticity 113
Model summary 114
Model 1 (unit type) 115
Model 2 (unit type and dichotomized teamwork climate) 117
Model 3 (unit type, dichotomized teamwork climate and safety climate 117
Model 4 (unit type, dichotomized teamwork climate, safety climate, and interaction term 117
Alternative hypothesis of predicting adverse event occurrence 118
Summary 120

Chapter 5: Discussion 122
Summary of the Study 122
Findings in Context of Existing Literature 122
Aim One 122
Predictors: Teamwork climate 122
Predictors: Safety climate 123
Teamwork client and unit type 123
Safety climate and unit type 123
Aim Two 124
Adverse events 124
Adverse events by unit type 125
GTT chart review methodology 126
Aim Three 127
Control variable: Unit characteristics (size) 127
Safety climate and teamwork climate 127
Safety climate with frequency of adverse events 127
Teamwork climate with frequency of adverse events 128
Teamwork by unit type 129
Unit type of frequency of adverse events 129
Findings in Context of Hospital Environment 130
Limitations 130
Instrumentation 130
Global Trigger Tool 130
Safety Attitudes Questionnaire 132
Sectional and Cross-Sectional Data 132
External Validity 132
Interpretation of Findings 132
Small Variance 132
Measurement Error 133
Random error 133
Systematic error 133
SAQ items 134
Adverse event measurement 135
GTT measurement process 135

Sampling Error 137
Nonresponse Error 137
Insights regarding the Global Trigger Tool 138
Implications for Theory 139
Organizational Factors 139
Human Factors 140
Mediating or Moderating Relationships 141
Conclusions 141
Implications for Practice 142
Recommendations for Further Research 144
Concluding Remarks 145

References 146

Appendix: Clarification of GTT Definitions 172

List of Tables

Table 3.1. Safety Attitudes Questionnaire Safety Climate Domain and Items 57

Table 3.2. Safety Attitudes Questionnaire Teamwork Climate Domain and Items 58

Table 3.3. Triggers used in Global Trigger Tool 61

Table 3.4. Categories of Harm used in The Global Trigger Tool 62

Table 3.5. Summary of Measures 64

Table 4.1. Descriptives: Unit Size, Teamwork Climate, and Safety Climate 74

Table 4.2. Descriptive Statistics: Unit Type 75

Table 4.3. Teamwork Climate Categorized into Two Groups 76

Table 4.4. Descriptive Statistics of Teamwork Climate and Safety Climate by Unit Type 79

Table 4.5. Descriptive Statistics of Unit Size by Unit Type 81

Table 4.6. One-Factor Fixed ANOVA Table of Unit Size by Unit Type 81

Table 4.7. Bonferroni Multiple Comparison Procedure of Unit Size by Unit Type 82

Table 4.8. Descriptive Statistics of Unit Size by Dichotomized Teamwork Climate 82

Table 4.9. One-Factor Fixed ANOVA Table of Unit Size by Dichotomized Teamwork Climate 83

Table 4.10. Descriptive Statistics of Unit Size by Safety Climate 83

Table 4.11. One-Factor Fixed ANOVA Table of Unit Size by Safety Climate 84

Table 4.12. Descriptive Statistics of Dependent Variable: Frequency of Adverse Events 85

Table 4.13. Descriptive Statistics of Adverse Events by Unit Type 87

Table 4.14. Prevalence of Triggers, Adverse Events and Trigger Positive Predictive Value (PPV) Using Institute for Healthcare Improvement Global Trigger Tool (GTT) 89

Table 4.15. Correlations Between Unit Size, Teamwork Climate, Safety Climate, and Frequency of Adverse Events 92

Table 4.16. Teamwork Climate (TC) Dichotomized into Two Groups * MedSurg Dichotomized into Two Groups Crosstabulation 97

Table 4.17. Test of Homogeneity of Variances for the Dependent Variables Adverse Events per 1,000 Patient Days, Adverse Events per 100 Admissions, and Percentage of Admissions with an Adverse Event 99

Table 4.18. Tests of Normality for Adverse Events per 1,000 Patient Days by Unit Type 100

Table 4.19. Tests of Normality for Adverse Events per 100 Admissions by Unit Type 100

Table 4.20. Tests of Normality for Percentage of Admissions with an Adverse Event by Unit Type 101

Table 4.21. One-Factor Fixed Effects ANOVA: Unit Differences for Adverse Events per 1,000 Patient Days and Percentage of Admissions with an Adverse Event 101

Table 4.22. Bonferroni Multiple Comparison Procedure for Adverse Events per 1,000 Patient Days and Percentage of Admissions with an Adverse Event 103

Table 4.23. Group Statistics: Dichotomized Teamwork Climate and Safety Climate in Medical Surgical Units Versus Other Unit Types 106

Table 4.24. Independent Samples Test: Dichotomized Teamwork Climate and Safety Climate in Medical Surgical Units Versus Other Unit Types 106

Table 4.25. Group Statistics: Frequency of Adverse Events in Low and High Teamwork Units 107

Table 4.26. Independent Samples Test: Frequency of Adverse Events in Low and High Teamwork Units 108

Table 4.27. Descriptive Statistics for Dichotomized Teamwork Climate and Unit Type with Adverse Events per 1,000 Patient Days: Results of a Two-Factor ANOVA 109

Table 4.28. Levene's Test of Equality of Error Variances for the Dependent Variable: Adverse Events per 1,000 Patient Days 110

Table 4.29. Tests of Between-Subjects Effects for Dichotomized Teamwork Climate and Unit Type with Adverse Events per 1,000 Patient Days: Two-Factor ANOVA 110

Table 4.30. Variables Entered/Removed for Hierarchical Multiple Regression Analyses Predicting Frequency of Adverse Events per 1,000 Patient Days 114

Table 4.31. Hierarchical Multiple Regression Model Summary Predicting Frequency of Adverse Events per 1,000 Patient Days from Unit Type, Dichotomized Teamwork Climate, Safety Climate, and Interaction Term of Dichotomized Teamwork Climate and Safety Climate 115

Table 4.32. Hierarchical Multiple Regression Analyses Predicting Frequency of Adverse Events per 1,000 Patient Days from Unit Type, Dichotomized Teamwork Climate, Safety Climate, and Interaction Term of Dichotomized Teamwork Climate and Safety Climate 116

Table 4.33. Omnibus Tests of Model Coefficients predicting Likelihood of Units Experiencing an Adverse Event with Predictors Unit Type, Dichotomous Teamwork Climate, and Safety Climate 119

Table 4.34. Logistic Regression Results Predicting the Probability of a Unit Experiencing an Adverse Event 119

Table 5.1. Comparison of Adverse Event Frequency among Various Hospitals using the Institute for Healthcare Improvement Global Trigger Tool for Measuring Adverse Events 125

List of Figures

Figure 2.1. Schematic representation of Generic Reference Model 14

Figure 2.2. Schematic representation of the relationship between contributing factors, the safety incident, and patient outcomes 30

Figure 3.1. Bar chart of Safety Attitudes Questionnaire (SAQ) nursing unit response rate divided into three groups 54

Figure 4.1. Bar graph depicting the distribution of teamwork climate scores 76

Figure 4.2. Bar chart of dichotomized teamwork climate scores 77

Figure 4.3. Bar graph depicting the distribution of safety climate scores 78

Figure 4.4. Boxplot of teamwork climate by unit type 79

Figure 4.5. Boxplot of safety climate by unit type 80

Figure 4.6. Bar chart depicting distribution of harm by category for all charts reviewed 86

Figure 4.7. Boxplot of adverse events per 100 admissions by unit type 87

Figure 4.8. Scatterplot of unit size and safety climate 93

Figure 4.9. Scatterplot of unit size and teamwork climate 93

Figure 4.10. Scatterplot of unit size and adverse events per 1,000 patient days 94

Figure 4.11. Scatterplot of teamwork climate and safety climate 94

Figure 4.12. Scatterplot of safety climate and adverse events per 1,000 patient days 95

Figure 4.13. Scatterplot of teamwork climate and adverse events per 1,000 patient days 96

Figure 4.14. Bar chart comparing low versus high teamwork scores by unit type (medical surgical versus other unit type 98

Figure 4.15. Profile plot of two-factor ANOVA effect between unit type and low and high teamwork units 111

Figure 4.16. An expected normal P-P plot for adverse events per 1,000 patient days used to assess the assumption of normality in regression 113

Figure 4.17. Plot of the predicted values of the dependent variable, number of adverse events per 1,000 patient days 114

Figure 4.18. Bar chart depicting distribution of adverse events in two groups of units, with group zero not experiencing an adverse event and group one experiencing an adverse event..... 118

Chapter 1: Introduction

Evidence suggests that more than 400,000 premature deaths per year occur due to preventable harm in U.S. hospitals (James, 2013), costing $17-$29 billion per year in healthcare expenses, lost worker productivity, and disability (National Quality Forum, 2012). Although patient harm occurs frequently with significant costs, morbidity, and mortality, mandatory event reporting has not provided much insight regarding why adverse events occur (Nemeth et al., 2006). Lack of progress in preventing harm may be partially explained by lack of tools to measure improvements in patient safety (Shojania & Thomas, 2013) and absence of simple yet meaningful frameworks (Pronovost et al., 2009).

In 2009, the Institute for Healthcare Improvement (IHI) published the Global Trigger Tool (GTT) for Measuring Adverse Events (Griffin & Resar, 2009). The GTT found 10 times more adverse events than voluntary reporting and the Agency for Healthcare Research and Quality's Patient Safety Indicators (AHRQ PSIs; Classen et al., 2011). The GTT has become the most widely used global patient safety measure (Pronovost & Wachter, 2013). The GTT measures adverse event rates at a single point in time, but it does not explain the context of these events. Research is needed with theories that explain health care outcomes and the healthcare context (Agency for Healthcare Research and Quality [AHRQ], 2010).

The Generic Reference Model (GRM) explains what goes wrong when patients are harmed from health care (Runciman et al., 2006). An understanding of the context of patient safety may occur through areas of study categorized as contributing factors in the GRM (Runciman et al., 2008). For example, environmental, organizational, and human factors are contributing factors to a safety incident (Runciman et al., 2008), but research is needed to define each GRM concept (The World Alliance for Patient Safety Drafting Group et al., 2009). There is a need to know how to reduce harm in health care in order to improve quality (U.S. Department of Health and Human Services, 2013).

Problem Statement

Adverse events with patient harm continue to occur during hospitalization. An understanding of the healthcare context is critical to protecting patients from harm. Context consists of the environmental factors (e.g., physical space), organizational factors (e.g., culture), and human factors (e.g., teamwork). For example, an organizational factor, such as patient safety culture, can be conducive or detrimental to patient safety (Guldenmund, 2014). Human factors, such as teamwork, are vital to create a common mental model for ensuring patient safety (Leonard, Graham, & Bonacum, 2004). Past studies have shown strong teamwork appears to protect patients from adverse events with harm, but the effect of safety culture remains inconclusive (see Chapter 2 for details). Previous research in Norway indicates a statistically significant correlation with unit-level teamwork climate and GTT-identified adverse event occurrence, but the relationship between safety climate and adverse events was not statistically significant, possibly due to small sample size (n = 4; Deilkas & Hofoss, 2008). It is unknown if the relationship of safety and teamwork climates with GTT-detected adverse events is the same in a U.S. hospital, with a larger sample of nursing units, using advanced statistics.

Statement of Purpose

The purpose of this study is to explore relationships between organizational and human factors with adverse events that result in patient harm detected using a modified trigger-tool methodology.

Specific Aims

Aim 1. Describe self-reported safety climate and teamwork climate among an interprofessional group of providers working on 32 hospital inpatient units, measured using the Safety Attitudes Questionnaire (SAQ).

Aim 2. Explore the nature of adverse events identified using a modified IHI GTT chart review methodology.

Aim 3. Examine to what extent unit-level safety climate and teamwork climate (and their potential interaction) predict adverse events (per nursing unit) as detected via the modified IHI GTT chart review format, controlling for unit characteristics.

Conceptual Framework for the Study

The Generic Reference Model was the conceptual framework for this study. The GRM deconstructs patient safety into contributing factors, the incident, and outcomes (Runciman et al., 2006). Contributing factors relate to the precursors of an incident, and by addressing these, healthcare organizations may reduce patient risk (The World Alliance for Patient Safety Drafting Group et al., 2009). Contributing factors to safety incidents include: (a) environmental factors, (b) organizational factors, (c) human factors, (d) subject of incident factors, and (e) factors that include "drugs, equipment, and documentation." Failure of defenses leads to incidents, which are circumstances that could have resulted, or did result in unnecessary patient harm (Runciman et al., 2009). The patient outcome is the impact upon a patient that is wholly or partially attributable to an incident (Runciman et al., 2009).

Definition of Terms

Disagreement exists among researchers regarding the definition of safety culture and whether or not safety culture differs from safety climate (Halligan & Zecevic, 2011). When researchers use questionnaires to assess group-level perceptions of safety, they are assessing climate because climate is more readily measurable than culture (including behavior, values, and competencies) which can be difficult to assess via survey (Sexton et al., 2006). Safety climate is defined as the "perceptions of a strong and proactive organizational commitment to patient safety" (Sexton et al., 2006, p. 3). Teamwork is defined as the "perceived quality of collaboration between personnel" (Sexton et al., 2006, p. e). Patient outcomes are defined as adverse events with harm. An adverse event is unintended physical injury resulting from medical care that requires additional monitoring, treatment or hospitalization, or that causes death (Griffin & Resar, 2009).

Assumptions

It is assumed that the responses received on the SAQ from the participating respondents on each unit accurately reflected their perceptions.

Overview of Methodology

This study was conducted with a descriptive, cross-sectional design, using a retrospective, modified IHI GTT chart review methodology. The research setting was a 750+-bed Midwestern U.S. hospital. The convenience sample comprised 32 adult inpatient, direct-care nursing units from one acute care hospital. The organizational factor, safety culture, was measured through percentage agreement collected from the SAQ safety climate domain. Human factors, that is, teamwork, were measured through percentage agreement collected from the SAQ teamwork climate domain. Patient outcomes were measured through adverse events collected from the IHI GTT for measuring adverse events. Statistical analyses included descriptive statistics, correlation, and regression.

Rationale and Significance

This study sought to address a critical barrier to progress in the safety field by making connections between safety climate, teamwork climate, and adverse events detected using a modified GTT process. In addition, this is one of the few studies exploring the context of patient safety through the GRM conceptual framework. Because trigger tools are more sensitive for detecting adverse events, it may be possible to discover novel relationships between safety and teamwork climates with trigger-identified adverse events. These discoveries could drive changes in clinical practice by identifying nursing units at high risk for specific adverse events. The effect of this research will be to clarify contributing factor concepts and relationships among adverse events, which may in turn drive future unit-based patient safety endeavors such as preventative interventions for protecting patients from harm.

Limitations

The study has limited external validity (generalizability) because data were sampled from a single hospital. Safety, teamwork, and adverse event data from early 2013 represents a snapshot dependent on conditions that occurred during that time. Low response rates for the SAQ were problematic for some nursing units. Experts propose a minimum 60% SAQ response rate to ensure that the data represent safety culture not opinions (Pronovost & Sexton, 2005). SAQ data from units with a 40% response rate or more was used to maximize sample size. A

multicenter study used SAQ data with a low (47.9%) response rate, and found no significant difference between ICU-level response rate and clinical performance index, which is the difference between observed hospital survival rate and survival rate predicted by severity of illness at ICU admission (Huang et al., 2010). The GRM contains many subclasses of contributing factors and hazards to patient safety, and many GRM constructs remain unmeasured. These gaps provide opportunities for future research.

Organization of the Dissertation

The remainder of the dissertation is organized as follows. Chapter 2 provides a literature review, including theoretical approaches to the study of patient safety and details regarding GRM, organizational factors, human factors and adverse events. Chapter 3 describes the research methods. Chapter 4 presents the results, and Chapter 5 presents a discussion of the findings.

Chapter 2: Literature Review

Adverse events continue to occur in U.S. hospitals, despite concerted efforts to reduce patient harm. Approximately one out of 10 patients experience an adverse event during hospitalization (deVries, Ramrattan, Smorenburg, Gouma, & Boermeester, 2008). Although more than half of patients that experience an adverse event experienced no or minor disability, 7% of patients die (deVries et al., 2008). In addition to significant morbidity and mortality, adverse events also present a financial cost to the U.S. health system. Measurable medical error cost the U.S. health system $17.1 billion in 2008 (Van Den Bos et al., 2011). Medicare has also felt the financial impact of adverse events, spending approximately $324 million in October 2008 caring for older adults that experienced harm (Levinson, 2010). More recently, a study demonstrated that Medicare beneficiaries that underwent total knee arthroplasty and suffered an adverse event consumed significantly more unadjusted hospital resources ($3,110 cost) and had longer stays (1.3 days) than patients that did not experience an adverse event (Culler, Jevsevar, Shea, Wright, & Simon, 2015). Similar costs occur when patients with chronic conditions, such as congestive heart failure, suffer an adverse event. The actual direct cost of an adverse event for congestive heart failure patients was $1,029 per case and $903 for a surgical case (Pappas, 2008). Adverse drug events (ADEs) are also costly, with one study demonstrating that each ADE costs between $2,852 to $8,116,depending on ADE severity (Hug, Keohane, Seger, Yoon, & Bates, 2012). In addition, adverse events affect health care practitioners through psychological cost of causing harm to another human being. Professionals working in intensive care expressed feelings of guilt, shame, and concern for the patient after making an error (Laurent et al., 2014). Nurses report stronger negative feelings (feeling upset, worried, distressed, scared, or nervous) compared to physicians after making an error (Harrison et al., 2015). Adverse events are a significant issue due to their frequency, effect on patient morbidity and mortality, expense, and psychological effects on practitioners.

Causes of Adverse Events

Human factors are known to be one of the prime causes of adverse events (Smits et al., 2010; The Joint Commission, 2014). Human factors are "concerned primarily with the performance of one or more persons in a task-oriented environment interacting with equipment, other people, or both" (National Research Council, 1992, p. 9). Human-based adverse events occur through knowledge-based deficits, and rule-based failures regarding monitoring, intervention and/or verification (Smits et al., 2010). Organizations may also contribute to adverse events due to protocol failures and lack of knowledge transfer (Smits et al., 2010). Adverse events may also be caused by technical failures due to equipment design, construction, and materials (Smits et al., 2010). Hospitals strive to protect patients from adverse events by instilling a culture of safety and strong interprofessional teamwork.

Relationship Between Safety and Teamwork With Adverse Events

To improve patient safety, healthcare systems have looked for guidance from other high-risk industries such as aviation, nuclear power plants, and naval aircraft carriers (Wachter, 2012). Two defenses that organizations use to reduce risk of error are safety culture and teamwork. Safety culture is part of organizational culture (Smits et al., 2012) and is defined as "the product of individual and group values, attitudes, perceptions, competencies, and patterns of behavior that determine the commitment to, and the style and proficiency of, an organization's safety management" (Sorra & Nieva, 2004, p. 1). Some researchers treat safety culture and climate as synonymous (Jackson, Sarac, & Flin, 2010). Culture represents a more stable characteristic of an organization, such as its "personality," while climate represents a more visible mood state at a point in time (O'Connor, Buttrey, O'Dea, & Kennedy, 2011). Safety culture and climate are not separate entities, but different approaches of measuring the value of safety within an organization (Guldenmund, 2007). Because patient outcomes such as adverse events result from failure(s) in a healthcare system at a particular point in time, it is appropriate to use safety culture/climate interchangeably when summarizing relationships among these variables.

Safety Culture/Climate and Adverse Events

Studies with significant relationships. Some studies have demonstrated a significant relationship between safety culture and patient outcomes, including adverse events. A 2014 systematic review of 17 studies found mostly significant relationships between safety culture and patient outcomes such as decreased mortality (three studies), medication errors (four studies), AHRQ PSIs composite score (two studies), and fewer nurse-sensitive patient safety indicators (four studies; DiCuccio, 2014). Studies excluded from DiCuccio's review also showed similar results. A more positive safety culture was related to lower adverse event severity (Kline, Willness, & Ghali, 2008), fewer hospital acquired pressure ulcers (Brown & Wolosin, 2013), and decreased medication errors (Vogus & Sutcliffe, 2007; Zohar, Livne, Tenne-Gazit, Admi, & Donchin, 2007). Likewise, a study testing a structural equation modeling framework with data from 204 hospitals found a strong patient safety culture was negatively related to hospital-acquired conditions such as hospital associated urinary tract infections, falls, and catheter-associated infections (McFadden, Stock, & Gowen, 2015). A dissertation with cross-sectional, exploratory descriptive methodology found a low positive correlation between mean safety culture composite score and death in low mortality diagnosis-related groups (DRGs; Fagan, 2012).

Studies with non-significant relationships. On the other hand, no significant relationships of any size were found in a meta-analysis examining the relationship between safety culture and pressure ulcers, falls, medication errors, nurse sensitive outcomes, and post-operative outcomes (Groves, 2014). DiCuccio's (2014) review found five studies with either non-significant or unexpected results relating to patient safety culture and outcomes. Three of the five studies found non-significant relationships with nurse-sensitive patient safety indicators, which included falls, pressure ulcers, pulmonary embolism/deep vein thrombosis, and hospital-associated infection.

Non-significant relationships between safety culture and patient outcomes have also been found in countries outside of the United States. A Norwegian study did not find a statistically significant relationship between safety climate and adverse events (Deilkas & Hofoss, 2008). A Dutch study examining the relationship between clinical practice specialties, safety

culture, and patient outcomes found that safety culture did not mediate the relationship between clinical practice specialties and patient outcomes (Smits et al., 2012). In 35 Swiss acute care hospitals, safety culture was not a significant predictor of medication errors, pressure ulcers, patient falls, urinary tract infection, bloodstream infection, pneumonia, and patient satisfaction (Schwendimann, Zimmermann, Küng, Ausserhofer, & Sexton, 2013).

Findings regarding relationships between safety culture and patient outcomes are inconsistent. Reasons for the inconsistency may be due to methodological and population differences. Methodological concerns relate to measurement of safety climate and patient outcomes. Multiple survey measures of safety climate make comparisons challenging. Patient safety outcomes measured via the AHRQ PSIs are limited to a narrow set of particular outcomes. Hospital practitioners tend to underreport adverse events due to lack of time or perceived minor nature of the error (Kaldjian et al., 2008). Other methodological concerns are the self-selected (hospital) samples that choose to measure staff perceptions of safety climate, which may indicate a strong interest in safety that differs from hospitals that do not assess safety climate (Mardon, Khanna, Sorra, Dyer, & Famolaro, 2010). Perceptions of safety climate may differ depending on job class (senior versus non-senior managers) and clinician status (clinician versus non-clinician; Singer et al., 2003).

The two most scientifically rigorous studies appear to be the meta-analysis by Groves (2014) with 10 reports included in the analysis, and DiCuccio's (2014) review of 17 studies because reviews are trustworthy and represent the highest level of evidence (Greenhalgh, 2010). Groves' study is the first to use an objective meta-analytic method with reporting of effect size estimates for five outcomes. DiCuccio's systematic review provides many summary tables, allowing the reader to compare each study by setting, design, variables, measurement tools, and outcomes. Neither review contains a study that measures adverse events using a GTT methodology.

Teamwork and Adverse Events

Studies with significant relationships. Teamwork is "the perceived quality of collaboration between personnel" (Sexton et al., 2006, p. 3). The relationship between teamwork

and adverse events remains inconclusive, although there is more evidence supporting a significant relationship than not. A 2009 narrative review found that many contributing factors to adverse events originated from inconsistent teamwork rather than clinical skill deficits (Manser, 2009). Team system factors were reported in 30% to 50% of all ICU safety incidents (Sinopoli et al., 2007). Hospitals with better teamwork tend to have lower rates of patient safety indicators (Mardon et al., 2010) and adverse events (Deilkas & Hofoss, 2008). Teamwork appears to be associated with two nurse-sensitive indicators, falls, and pressure ulcers. Teamwork within units is inversely correlated with falls; units with stronger teamwork report fewer falls (Brown & Wolosin, 2013; Sammer, 2009) and stronger teamwork culture provides lower odds of developing a pressure ulcer (Taylor et al., 2012). Teamwork was significantly positively associated with perceived patient safety grades after controlling for other independent variables and demographic variables in a secondary analysis with over 100,000 RN respondents (Molloy, 2012).

Studies with non-significant relationships. Contrary to these results, there was no significant relationship between teamwork climate and risk-adjusted 30-day morbidity and mortality in Veterans Affairs Medical Centers and academic medical centers (Davenport, Henderson, Mosca, Khuri, & Mentzer Jr, 2007). Similarly, a secondary analysis of data from over 100,000 RN respondents found an inconclusive relationship between workplace climate factors like teamwork and number of reported safety events (Malloy, 2012). Reasons for the inconsistent evidence regarding the relationship between teamwork and adverse events may be due to methodological issues. Methodological concerns could be related to the measurement of teamwork climate and patient outcomes. For example, measurement of adverse events based upon self-report may underestimate their true frequency.

Previous research established a statistically significant correlation with unit-level teamwork climate and adverse events detected with a GTT methodology, but the relationship between safety climate and adverse events was not statistically significant (Deilkas & Hofoss, 2008). Multiple reasons may explain why. First, the purpose of the Deilkas and Hofoss (2008) study was to establish the psychometric properties of the SAQ Norwegian version, not to explore adverse events or patient harm. Second, the study used a sample of four hospital units/

departments, which is quite small, and validation is needed with larger samples. Third, correlation cannot be used to infer causation (Lomax, 2007), and correlation does not assume that one variable is influencing the other (Logan, 2013). More advanced statistics, such as regression analysis, will permit the ability to explore predictive relationships, which correlation analysis cannot do. Fourth, the Deilkas and Hofoss (2008) study was conducted in Norway, which has a universal, public health care system. Contributing factors to patient safety, such as environmental factors (health care payment), organizational factors (workplace social norms), and human factors (communication methods) may differ between these countries and health care systems. Research is needed to explore these relationships in a U.S. hospital system using a theoretical framework.

Theoretical Approaches to Study Patient Safety

Theoretical frameworks, or models, provide a mechanism for visualizing elements involved in patient safety and their interactions (Emanuel et al., 2008). Scholars develop and advance scientific knowledge by testing theory and observing the real world (Nemeth, 2012). Some experts assert that lack of underlying a conceptual framework to explain context undermines the advancement in patient safety practices (Dy et al., 2011). Theoretical frameworks explain the mechanisms by which perceptions affect behavior (Flin, 2007) and thus patient safety (Halligan & Zecevic, 2011; Reiman, Pietikäinen, & Oedewald, 2010).

James Reason's Swiss Cheese Model

Numerous patient safety theoretical frameworks exist. One of the most recognized patient safety theoretical frameworks is James Reason's Swiss Cheese Model, also known as the latent failures model. It describes how defenses, barriers, and safeguards may be penetrated to cause an accident (Peltomaa, 2012). In this model, an individual error is represented by one hole in a slice of cheese, and flawed systems are represented by holes in other pieces of cheese (Mansfield & Noria, 2015). In complex organizations like healthcare, an individual's error rarely causes harm (Wachter, 2012). Accidents happen when all of the layers are penetrated, for example, when all of the imperfections or holes line up (Dekker, 2011). The model represents the challenges in predicting the alignment of conditions and processes that make the patient

vulnerable to harm (Woods, Barnard, Sikka, & Dhindsa, 2015). Rather than focus on trying to perfect human behavior, Reason's model aims to shrink or plug the holes in the Swiss cheese (latent errors) to create layers of protection to decrease the probability that the holes will ever line up, let an error slip through, and harm a patient (Wachter, 2012). One of the key insights provided by Reason is that human error is inevitable, and a systems approach strives to catch or block human errors before they occur and cause harm (AHRQ, 2015b).

A strength of James Reason's Swiss Cheese Model is its system approach, focusing away from front-line workers (sharp-end practitioners) to conditions that influence and constrain work (Dekker, 2011). Due to its linear chain of events, it introduces the idea of a root cause; the first domino that falls sets everything in motion (Dekker, 2011). The model is limited by presenting events preceding the adverse event as occurring in a linear, fixed-order fashion, whereas in the real world, there are multiple sequences of events that lead to an adverse event (Dekker, 2011). Also, the broad definition of latent characteristics can include almost anything, which makes targeting failures difficult (Dekker, 2011). In addition, experts interpret the features of the Swiss Cheese model differently, and thus it may be too simplified to be effective for promoting patient safety (Perneger, 2005).

High Reliability Theory (HRT)

Another approach to studying patient safety is high reliability theory (HRT). High reliability theory identifies organizational properties or processes that reduce the risk of operating in a complex, interactive, and tightly coupled environment (Shrivastava, Sonpar, & Pazzaglia, 2009). It describes the ability of people at all levels within an organization to ensure consistent operations in high-hazard settings (Dekker, 2011; Tamuz & Harrison, 2006). Organizations that successfully operate and manage complex technological systems consistently and error-free are high-reliability organizations (HROs). Five HRO characteristics are preoccupation with failure, sensitivity to operations, reluctance to simplify, resilience, and deference to expertise (Weick & Sutcliffe, 2001).

The strength of HRT is its flexible structure, which facilitates swift response (Tamuz & Harrison, 2006). It also emphasizes a culture of mindfulness (Tamuz & Harrison, 2006), meaning

an event can happen at any time. On the other hand, HRT is limited because it is unknown if HRO features can persist throughout the entire lifetime of an organization (Boin & Schulman, 2008). Because failures occur so rarely in HROs (by definition), it is difficult to study the processes they manage successfully (Boin & Schulman, 2008). Researchers have studied HRO examples including air traffic controllers, nuclear power plants, and naval aircraft carriers (Wachter, 2012), but high reliability characteristics present in aviation or nuclear power industries may not apply to healthcare. Preoccupation with failure requires scrutiny of every error, and healthcare practitioners may fear liability and loss of job security (Portner, Fumanti, Bendas, & Cipolla, 2015). Plane crashes and nuclear accidents affect large groups of people while the devastating effect of an adverse event in healthcare affects one patient (Schulman, 2004).

While both the James Reason's Swiss Cheese Model and HRT have contributed to shifting from individual blame to a systems approach for ensuring patient safety, these models are limited for research exploring the potential effect of safety culture on patient outcomes. Notably, both models lack the ability to explain how the safety culture concept is tied to patient harm. They also appear to be process-oriented, describing how harm/failure may occur under particular environmental conditions, but patient outcomes are not emphasized. Additionally, these frameworks provide a high-level overview, but factors or processes that provide protection from harm are not described in detail. Alternative theoretical approaches to study contributing factors to patient safety may prove more useful.

The Generic Reference Model (GRM)

The GRM (Runciman et al., 2006) serves as a conceptual framework for researchers and clinicians to collect and organize information regarding safety incidents (see Figure 2.1). The GRM is a universal classification system designed to combine and compare information over time to facilitate sharing among individuals, organizations, and countries (Runciman et al., 2006). Runciman et al. (2006) intend for the model to be useful for a variety of health care areas, so that safety information is translatable into a common language for record sharing and analysis

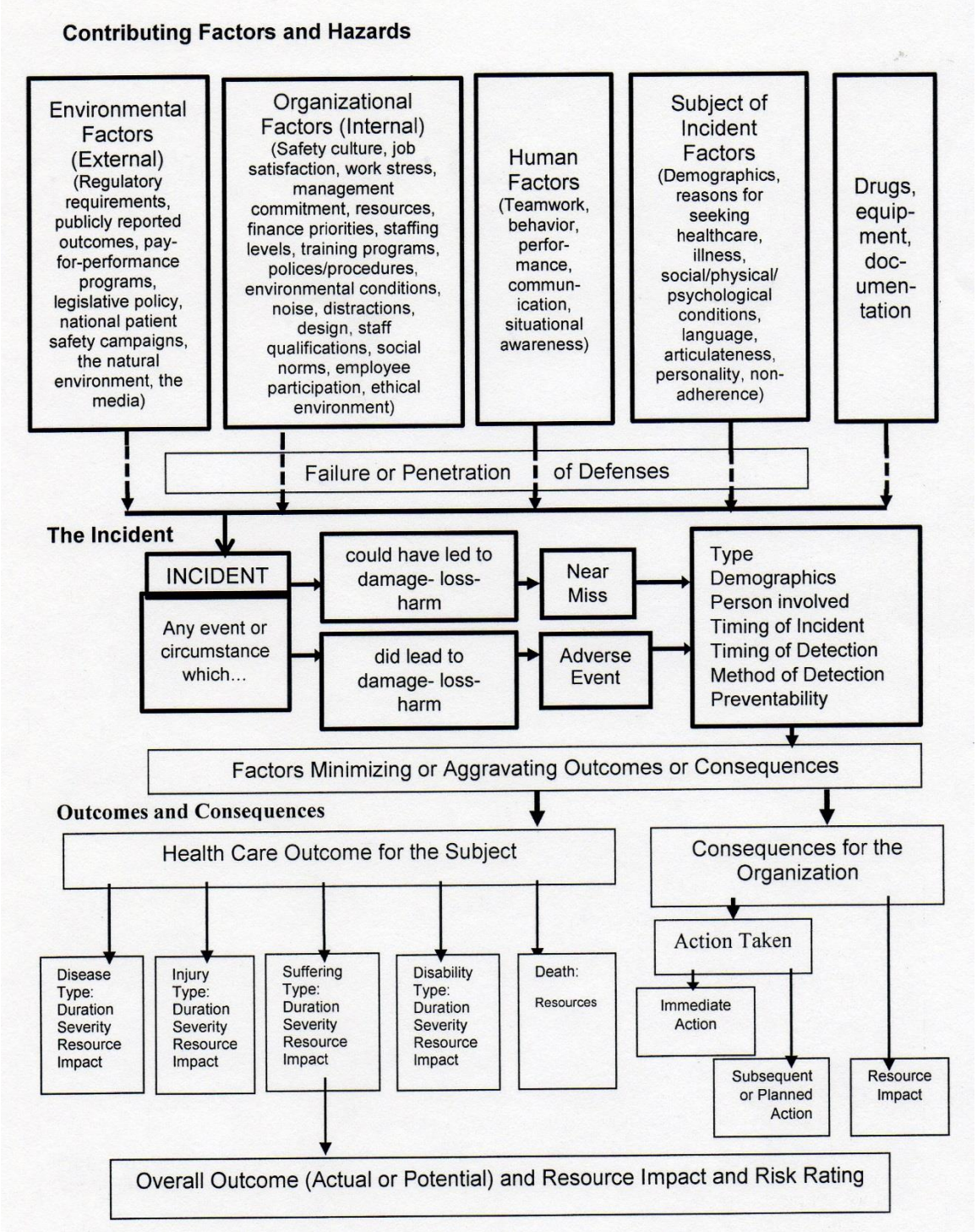

Figure 2.1. Schematic representation of Generic Reference Model. Schematic representation of the relationship between contributing factors and hazards, the safety incident, and outcomes or consequences. Reproduced from "An Integrated Framework for Safety, Quality and Risk Management: An Information and Incident Management System Based on a Universal Patient Safety Classification" by W. B. Runciman et al., 2006, *Quality and Safety in Health Care, 15*(Supplement 1), p. i82-90, with permission from BMJ Publishing Group Ltd.

(Runciman et al., 2006). The model is suitable for deconstructing patient safety incidents to elicit

The GRM contains three major sections: (a) contributing factors and hazards, (b) the incident, and (c) outcomes and consequences. Contributing factors/hazards are circumstances, actions or influences thought to have played a part in the origin, development or increased risk of a patient safety incident (The World Alliance for Patient Safety Drafting Group et al., 2009). Failure or penetration of defenses (contributing factors or hazards) plays a part in the development of a patient safety incident. There are five types of contributing hazards: environmental factors, organizational factors, human factors, subject of incident factors, and drugs, equipment, documentation, represented by boxes in Figure 2.1. Typically, more than one contributing factor is involved in a single patient safety incident (The World Alliance for Patient Safety Drafting Group et al., 2009). Failure or penetration of defenses plays a part in the development of a patient safety incident. A patient safety incident is "an event or circumstance that could have resulted, or did result, in unnecessary harm to a patient" (Runciman et al., 2009, p.19).

Rationale for the Selected Theoretical Approach

The GRM holds numerous strengths for understanding the phenomenon of patient safety and harm. Simply stated, the GRM helps explain what goes wrong when patients are harmed from health care (Runciman et al., 2006). The GRM accepts the importance of safety culture to prevent patient harm. An assumption of the GRM is that an understanding of culture is "an essential prerequisite for understanding how and why things go wrong in healthcare" (Runciman et al., 2008, p. 478). Because the GRM includes individual and organization consequences of failed defenses, it is consistent with an orientation towards outcomes. Improving patient outcomes (by reducing patient harm/adverse events) in health care settings is fundamental for improving quality (U.S. Department of Health and Human Services, 2013).

The GRM is quite detailed with inductive origins, based upon safety experts sharing their observations of relationships from data, existing literature, and clinical practice. Theory details are important to promote effective design of patient safety interventions and understand mechanisms of action for replication (Shekelle, Pronovost, & Wachter, 2010). There are several

contributing factors that acknowledge multifactorial causes of safety incidents, including structure and process failures (Runciman et al., 2010). Safety and teamwork climates effortlessly fit within the GRM as contributing factors to patient safety. Thus, the GRM is the most promising framework for research examining the effect of safety culture and teamwork on patient outcomes.

GRM Contributing Factors and Hazards

Environmental Factors

Environmental (external) factors are not under control of health care organizations, yet play a part in the development of a safety incident. Examples of environmental factors are regulatory requirements, publicly reported outcomes (i.e., pay-for-performance programs), legislative policy, national patient safety campaigns, the natural environment, and the media. These factors exist within the macro-level system of health care, the external environment in which the health care organization exists (Shekelle et al., 2011). Experts agree that the environmental context for patient safety is important, but there is minimal evidence regarding which elements are most influential (Shekelle et al., 2011).

Regulatory requirements dictated by The Joint Commission or liability insurers may require certain safety practices to reduce incident risk (Hoffmann & Rohe, 2010; Shekelle et al., 2011), thus driving patient-safety initiatives (Devers, Pham, & Liu, 2004). Publicly reported patient outcomes/safety metrics, in conjunction with pay-for-performance (Shekelle et al., 2011), may influence practitioners' adoption of safety behaviors. Legislative policy may be a contributing factor in patient safety incidents (Duckers et al., 2009) through state regulations related to mandatory reporting, peer review protection, and event measurement (Weinberg, Hillborne, & Nguyen, 2005). However, a cross-sectional retrospective study found no effect of state legislation reporting requirements on central line-associated bloodstream infection (CLABSI) rates (Pakyz & Edmond, 2013). The handling of other rules and regulations (Hoffmann & Rohe, 2010), such as mandated work hour limitations for resident physicians (Singer & Vogus, 2013), may also reduce error risk by preventing residents' fatigue. Working extended shifts may decrease attention, memory, cognitive processing, and ability to learn (Mansukhani, Kolla, Surani, Varon, & Ramar, 2012).

National patient safety campaigns, the natural environment, and the media are other examples of environmental factors that influence patient safety. National patient safety campaigns may originate from larger social influences such as government, policy, and economic structures that influence individual behaviors (Fisher, 2008). The natural environment (Duckers et al., 2009) of the hospital (i.e., location in a flood plain subject to natural disaster or metropolitan area subject to terrorist attacks) is a potential environmental factor affecting staff's ability to respond to incidents. News media increases the public's awareness of health care dangers when they publicize sentinel safety events (Shekelle et al., 2011). Shekelle et al. (2013) stated regulatory requirements, public reporting, pay-for-performance, and local sentinel events are high priority contexts in patient safety research.

Organizational Factors

Organizational factors, internal to the system, affect the work environment (Duckers et al., 2009). Examples of organizational factors that may contribute to safety incidents are safety culture, resources, restrictions, structure of the practice (i.e., organizational hierarchy of nursing/medicine), handling of rules, regulations, and priorities (Hoffmann & Rohe, 2010). Powell-Cope, Nelson, and Patterson (2008) described organizational factors that influence the use of technology as culture, resources, policies, social norms, management commitment, training programs, employee participation/ empowerment, and the ethical environment. Shekelle et al. (2013) recommended organizational factors such as size, complexity, and financial status as high-priority contexts in patient safety research.

Safety culture. A safety culture exists when individuals feel comfortable highlighting potential safety threats or breakdowns without management disapproval (Wachter, 2012). A qualitative meta-analysis of the literature determined that leadership, teamwork, evidence-based practice, communication, learning, just culture, and patient-centered care are the subcultures of safety culture (Sammer, Lykens, Singh, Mains, & Lackan, 2010). Similarly, other experts state the most commonly measured safety features are teamwork, safety attitudes/behaviors, job demands, personal resources (stress), management, safety systems, risk perception, reporting, communication, and organizational factors (Flin, Burns, Mearns, Yule, & Robertson, 2006).

Researchers disagree how to define safety culture, and whether safety culture differs from safety climate (Halligan & Zecevic, 2011). Safety culture is usually assessed through caregiver and manager self-report via a validated instrument (Wachter, 2012). There is uncertainty regarding whether existing patient safety surveys address all relevant aspects of culture, but the core dimensions appear to be teamwork, communication, and leadership support (Singla, Kitch, Weissman, & Campbell, 2006).

Because safety culture is a mostly a local phenomenon (Wachter, 2012), there can be wide variability across different work units within a single hospital (Pronovost & Sexton, 2005). Experts recommend assessing safety culture at the unit (or department) level (Pronovost & Sexton, 2005), where practitioners provide care, for example, as close to the patient as possible (Deilkas & Hofoss, 2010). When researchers assess safety culture via validated survey, experts recommend aggregating responses at the unit level because safety climate is a shared belief about the work environment (Etchegaray & Thomas, 2014). Nursing units with positive safety climate scores may be able to compensate for poor hospital climate scores (Zohar et al., 2007).

Although safety culture is part of organizational culture, individual characteristics may also influence safety perceptions. Safety attitudes might vary according to job type/role and age (Gallego, Westbrook, Dunn, & Braithwaite, 2012). A descriptive, Australian study found that executives, senior managers, and older health care workers (especially staff members aged over 60) express more favorable safety attitudes (Gallego et al., 2012). When hospital staff expresses positive views regarding patient safety, patients report greater positive experiences of care via the Consumer Assessment of Healthcare Providers and Systems Hospital Survey (CAHPS; Sorra, Khanna, Dyer, Mardon, & Famolaro, 2012).

Job satisfaction. A second type of organizational factor is job satisfaction, which experts define as “positivity about the work experience” (Sexton et al., 2006, p. 3). Nurses who have a heavy workload may feel job dissatisfaction, which affects motivation for high-quality performance and attitudes (Carayon & Gurses, 2008). For example, a high workload increases error rates in ICUs (Steyer, Schiffinger, Huber, Valentin, & Strunk, 2012). Nurses and other direct care providers may be less satisfied with work than non-medical workers in hospitals

(Listyowardojo, Nap, & Johnson, 2012), but they have the closest contact with patients and are most likely to catch errors and prevent adverse events. When nurses feel job satisfaction, they are motivated to engage in error prevention techniques inherent in a safe healthcare organization.

Work stress. Nurses with an arduous amount of work may experience stress and burnout, which can affect their performance (Carayon & Gurses, 2008) in a healthcare organization. One study found that each additional patient per nurse (increased workload) increased the odds of poor or failing safety grades (Aiken et al., 2012). Nurses experiencing high workload may not be able to perform because of reduced physical and cognitive resources (Carayon & Gurses, 2008). Under stress, workers may revert to previous individual history, anti-socialization, and teaching, which conflict with teamwork principles (Knox & Simpson, 2004). When nurses recognize their stress, they can consciously slow their actions and use resources such as other team members to delegate or double-check care, and request minimal disruptions during cognitive tasks. With stress recognition, nurses adopt strategies to ensure optimal performance for safe patient care.

Management commitment. Managers exhibit core organizational values and demonstrate commitment to patient safety initiatives, modeling behavior for staff. They establish open communication expectations so nurses feel they can speak up and be heard, which influences nursing satisfaction, teamwork and patient safety (Garon, 2012). Nurse leaders, such as management, encourage the bedside nurse's expert voice, provide resources, encourage education, coordinate initiatives to advance safety initiatives and build partnerships (Thompson, Navarra, & Antonson, 2005). When nurses feel they have management support with adequate resources to provide quality patient care, they work in an organization that prioritizes patient safety.

Resources. The manner in which an organization establishes and uses its resources, like finances and funding priorities, contributes to safety incidents. Decisions concerning human resources, such as staffing, overtime, wages, and the working environment are aspects of work that may relate to patient safety outcomes (Stone et al., 2007). Evidence supports a relationship between higher (better) nurse staffing (i.e., a higher proportion of hours of care per day provided

by RNs, or having an additional nurse per shift) and patient outcomes such as lower inpatient mortality (Aiken, Clarke, Sloane, Sochalski, & Silber, 2002; Brooks Carthon, Kutney-Lee, Jarrín, Sloane, & Aiken, 2012; Shekelle, 2013), lower rates of urinary tract infection (Needleman, Buerhaus, Mattke, Stewart, & Zelevinsky, 2002), lower rates of upper gastrointestinal bleeding (Needleman et al., 2002), lower rates of pneumonia (Needleman et al., 2002), lower rates of shock or cardiac arrest (Needleman et al., 2002), failure to rescue (Aiken et al., 2002; Needleman et al., 2002; Park, Blegen, Spetz, Chapman, & De Groot, 2012), and decreased probability of an adverse event (Voepel-Lewis, Pechlavanidis, Burke, & Talsma, 2012). Nurse staffing is a financial resource that provides a defense barrier against patient safety incidents. Other resources that create optimal healthcare organizations are training programs and policies/procedures congruent with prioritization of safety.

Environmental conditions. Environmental conditions are an example of a GRM organizational factor affecting patient safety. Nurses that report a more positive work environment also report higher levels of patient safety (Kirwan, Matthews, & Scott, 2013). Deficiencies in the design of healthcare environments contribute to adverse events during care (Barach, 2008). Some experts believe that it is possible to improve patient safety by decreasing adverse event precursors, like latent conditions and cognitive conditions through improved facility design and working conditions (Reiling, Hughes, & Murphy, 2008). Nurses identified important physical working environmental factors that contribute to medication errors as lack of privacy, unsuitable space layout in the nursing unit, inadequate documentation and medication room space, a noisy environment, defective medication dispensation equipment, and nursing station location (Mahmood, Chaudhury, & Valente, 2011). Lighting and noise may act as a mediating factor influencing the use of technology for preventing adverse events (Powell-Cope et al., 2008).

Human Factors

Human factors are contributing factors to patient safety because they concern “the performance of one or more persons in a task-oriented environment interacting with equipment, other people, or both” (National Research Council, 1992, p. 9). The Joint Commission (2014) found human factors were the most frequently identified root cause of sentinel events from 2012-

2014. Similarly, a Dutch structured record review study found that human factors were involved in the causation of 65% of surgical adverse events (Zegers et al., 2011). Human factors that may contribute to a safety incident are behavior, performance or communication (Duckers et al., 2009; The World Alliance for Patient Safety Drafting Group et al., 2009).

Researchers engaged in human factors research design interactive systems of people, equipment and environment, building upon human strengths and limitations to guarantee safety, effectiveness, and utility (Henriksen, Dayton, Keyes, Carayon, & Hughes, 2008). Human factors science does not view the failure of people as the cause of adverse events (Russ et al., 2013). Rather, human factors specialists attempt to design systems that support human performance and resiliency to unanticipated events (Russ et al., 2013). A resilient healthcare system will allow practitioners to identify, adjust, and recoup from errors (Carayon, Xie, & Kianfar, 2014).

Teamwork. Human factors may affect the behavior of individuals or teams (Runciman et al., 2010). Human factors work ranges from the individual practitioner to an organizational level, which may include how design and policies affect teams (Russ et al., 2013). For example, policies requiring communication tools such as safety huddles at the beginning of every shift may positively affect how teams work together (Leonard et al., 2004). Experts have proposed using teams to organize health care work and manage care to improve quality and safety (Carayon et al., 2013).

Teamwork is defined as “the perceived quality of collaboration between personnel” (Sexton et al., 2006, p. 3). Patient safety is created by teams (Knox & Simpson, 2004) because members share responsibility, capitalize on individual talents, trust one another, and use open communication skills to accomplish goals. Reliable care processes and effective teamwork are necessary for consistent, high-quality health care (Leonard & Frankel, 2011). Through teamwork, staff shares responsibility for maintaining patient safety. The basic components of teamwork and communication share assertion language, psychological safety, situation awareness, and effective leadership behaviors (Leonard & Frankel, 2011).

Researchers have designed various interventions to improve teamwork for patient safety. The implementation of a comprehensive unit-based safety program (CUSP), part of the Michigan

Keystone ICU Project to standardize evidence-based interventions, demonstrated improved teamwork climate scores (Pronovost et al., 2008) one year after the intervention. Similarly, the AHRQ developed the TeamSTEPPS (Team Strategies and Tools to Enhance Performance and Patient Safety) program to improve team performance (AHRQ, 2008). Other researchers have used educational interventions to improve teamwork climate in the operating room (Bleakley, Boyden, Hobbs, Walsh, & Allard, 2006). Teamwork is an example of human factors contributing to patient safety because teams direct and coordinate patient care.

Health care worker behavior and performance. Patient safety is influenced by the health care worker's physical, cognitive, and social/behavioral performance processes (Carayon et al., 2014). Specifically, attention, clinical skills, education/training, fatigue, goals, personal health, knowledge, level of experience, motivation, perception, self-efficacy, sensory input, and stress levels may all affect practitioner performance or behavior (Hoffmann & Rohe, 2010; Powell-Cope et al., 2008). A prospective cohort study found that medical error risk was affected by practitioners' (fewer) years of experience, feeling unskilled, and decreased attention (Tanaka et al., 2012).

Human factors ergonomists use strategies to maximize effective performance processes. For example, medication carts with easy-to-maneuver wheels and adjustable height minimize nurses' physical exertion. Adjusting alarm parameters to reduce false alarms minimizes cognitive load (Carayon et al., 2014). Practitioners can assist one another by identifying cognitive bias during clinical decision-making (Croskerry, 2013). Health care organizations influence social and behavioral performance processes through standardization, such as implementing checklists to ensure safety steps (AHRQ, 2012b).

Communication. Communication is another human factor contributing to patient safety. A systematic review identified communication as one of the top five contributing factors to safety incidents (Lawton et al., 2012). In fact, communication was the number one identified root cause for delay in treatment sentinel events (resulting in death or permanent loss of function) occurring between 2004-2012 (The Joint Commission, 2014).

Effective interpersonal communication results in problem solving, decision-making, goal accomplishment, team building, and the exchange of ideas (Burger, 2013). Teams may use a variety of tools to facilitate structured communication such as briefings, multidisciplinary rounds, huddles, checklists, and situational briefing models such as SBAR (situation, background, assessment, and recommendation; Leonard & Frankel, 2011). Structured communication creates predictability and agreement for communication expectations among all team members (Leonard & Frankel, 2011). An SBAR pre-post implementation study found that nurses perceived improved communication and collaboration with SBAR, and patients experienced a decrease in unexpected deaths (DeMeester, Verspuy, Monsieurs, & VanBogaert, 2013). Some researchers have proposed the integration of critical conversations during non-procedural hospital junctures, similar to the concept of a procedural "time out," during times where the patient is at high-risk for harm, like the time of admission, changes in the clinical condition, and hospital discharge (Sehgal et al., 2011). Other safety programs such as crew resource management (CRM) include communication and cross-check techniques to improve communication among team members (Haerkens, Jenkins, & VanderHoeven, 2012). The absence of effective communication creates conditions where medical mistakes can happen (O'Daniel & Rosenstein, 2008).

Subject of Incident Factors

The "subject of incident factors" are patient characteristics contributing to the patient safety incident within the GRM. Patient characteristics are patient qualities and attributes, like demographics or reasons for seeking healthcare treatment (Runciman et al., 2009). Older age (Aranaz-Andrés et al., 2011; Kable, Gibberd, & Spigelman, 2008; Naessens et al., 2012) and greater number of comorbidities (Aranaz-Andrés et al., 2011; Kim, Capezuti, Kovner, Zhao, & Boockvar, 2010; Naessens et al., 2012) increase the odds of developing an adverse hospital event. Patients who need to travel further than a county for health care also have increased odds of developing an adverse hospital event (Naessens et al., 2012). There is evidence to support racial disparities in patient safety, as well. Hospitalized blacks are at higher risk than white patients are for experiencing patient safety events such as nosocomial infections and ADEs

(Metersky et al., 2011). Other personal characteristics, such as (longer) operation duration were strongly predictive of adverse events in Australian surgical patients (Kable et al., 2008).

Illness, social/physical/psychological conditions, the relationship between the patient and the hospital, language, articulateness, and personality are other examples of patient factors contributing to an incident (Hoffmann & Rohe, 2010). Patients with low English proficiency and or individuals at the end of life are at risk for medical error (Mattox, 2010). Personality factors such as patient non-adherence to the treatment plan may also contribute to a safety incident (Runciman et al., 2009).

Drugs, Equipment, and Documentation

The fifth type of GRM contributing factors are drugs, equipment, and documentation, which are essential technical components for much of healthcare (Runciman et al., 2008). The GRM labels "drugs, equipment, and documentation" as a separate type of contributing factors to patient safety, but other researchers such as Hoffmann and Rohe (2010) classify equipment under organizational factors. Identifying "drugs, equipment and documentation" as a separate group of contributing factors may be to highlight the significance of how medication errors, equipment and documentation contribute to patient safety incidents.

Drugs. Adverse drug events result in patient injury from medical care involving medication use, but their occurrence does not necessarily indicate an error or poor care quality (AHRQ, n.d.). When an ADE involves an error element such as omission or commission, experts refer to these as preventable ADEs (AHRQ, n.d.). Errors of commission involve direct causation (i.e., drug toxicity) while omission errors involve indirect causation, like lack of access to drugs or therapeutic failure (Kanjanarat et al., 2003).

Adverse drug events occur frequently during hospitalization, with researchers reporting the proportion of preventable ADEs ranging from 15 to 50% (Institute of Medicine Committee on Identifying and Preventing Medication Errors, 2006). These incidents have a significant economic and patient safety impact. One study estimated that ADEs in community hospitals cost more than $3,000 dollars on average, with an average increase in length of stay of 3.1 days (Hug et al., 2012).

ADEs occur due to prescribing errors, dispensing errors, medication administration errors, and patient compliance errors (American Society of Hospital Pharmacists, 1993). Drug errors most often originate in the administration phase of medication in both the ICU and non-ICU settings, with errors of omission being the most common type of error (Latif, Rawat, Pustavoitau, Pronovost, & Pham, 2013). Experts report ADEs most commonly for narcotic (Beckett, Sheehan, & Reddan, 2012; Mills, Neily, Kinney, Bagian, & Weeks, 2008), chemotherapeutic (Mills et al., 2008), diabetic (Beckett et al., 2012; Kale, Keohane, Maviglia, Gandhi, & Poon, 2012; Mills et al., 2008; Nobre & McKay, 2012) , cardiovascular (Beckett et al., 2012; Kale et al., 2012; Mills et al., 2008), and anti-infective (Beckett et al., 2012) medication classes. Heparin, Ativan (lorazepam) and Solu-Medrol (methylprednisolone sodium succinate) have also been noted as drugs causing patient harm (Kale et al., 2012).

Equipment. Health care providers use medical equipment, such as specialty beds, infusion pumps, and monitoring devices for improving patient care and outcomes. The literature shows conflicting results regarding the impact of equipment on safety incidents, with some studies showing a positive impact and other showing harm. Equipment may cause harm, such as when patients have been caught in bedside rails, incorrect medication doses have been delivered via infusion devices, and automated blood pressure devices have been inaccurate leading to unnecessary treatment (Swayze & Rich, 2012). Immediate-use steam sterilization (IUSS) for surgical instruments, when appropriately used, prevents surgical delay and subsequent patient harm. However, when practitioners use IUSS inappropriately and too frequently, it may actually increase the risk of surgical site infection (Young, 2013; Zuckerman, Parikh, Moore, & Talbot, 2012), a patient safety event. Medical equipment may be a source of bacteria when it is not properly cleaned (Havill, Havill, Mangione, Dumigan, & Boyce, 2011). Device manufacturers may be the most suited to detect equipment trends that would be difficult to detect from a single hospital's data (Lipschultz, 2013).

Nurses frequently use monitoring systems, such as electrocardiogram (ECG) alarms for heart rate monitoring, for patient safety. The Joint Commission approved a National Patient Safety Goal (NPSG) on alarm management (effective January 1, 2014) and for hospitals to

establish alarm safety as an organizational priority (American Association for the Advancement of Medical Instrumentation, 2013). In 2016, The Joint Commission will require policies to determine who has the authority for setting alarm settings, changing alarm parameters, turning parameters "off," and checking signals for setting accuracy (American Association for the Advancement of Medical Instrumentation, 2013), to minimize patient risk if practitioners do not attend to a signal or it malfunctions.

Documentation. Known causes of electronic-health record (EHR)-related patient safety incidents are documentation errors, such as wrong data entry or the failure to enter data (Sparnon & Marella, 2012). Failure to document fluid balance, weight, and vital signs may contribute to adverse events (Zimmermann et al., 2010). There is a relationship between documentation and ADEs, as well. Accurate documentation of drug allergies and medication reconciliation (Nobre & McKay, 2012) provide a defense against ADEs. In a Finnish retrospective study, documentation was the leading cause of information management errors related to medication administration (Jylhä, Saranto, & Bates, 2011). Among harmful drug errors, computer-entry errors and errors due to illegible handwriting are less likely to be associated with the ICU versus non-ICU hospital setting (Latif et al., 2013), because ICU settings are more likely to use EHRs and computerized provider order entry.

Electronic health records (EHRs). Nurses perceive the value of EHRs for patient safety. A 2011 cross-sectional, secondary analysis found that nurses who worked in hospitals with basic EHRs were less likely to report missing information or that "things fell 'between the cracks" during patient transfers between units (Kutney-Lee & Kelly, 2011, p. 470). Hospitals with a basic EHR are associated with superior patient outcomes, "independently of nurse staffing" (Kutney-Lee & Kelly, 2011, p. 470).

A secondary data analysis found associations between U.S. hospitals that adopted an EHR and a (small) significant reduction in length of stay and 30-day mortality, but an increase in 30-day re-hospitalization (Lee, Kuo, & Goodwin, 2013). The authors speculate that the EHR facilitates faster test/procedure/medication ordering and better care coordination, which leads to reduced length of stay. However, shorter length of stay may contribute to an increase in 30-day

re-hospitalization (Lee et al., 2013). Another study found that while EHRs did not reduce the rate of patient safety events, once an event occurred, the EHR reduced death by 34%, readmissions by 39%, and spending by $4,850 (16%; Bae & Encinosa, 2011). EHRs may have an impact on nurse-sensitive patient outcomes, as well. For example, researchers noted a reduction in the number of hospital-acquired pressure ulcers (HAPUs) with the introduction of an integrated EHR (Dowding, Turley, & Garrido, 2012) and better documentation of borderline low blood pressures and a decrease in patient falls within a Connecticut hospital (Nobre & McKay, 2012).

Researchers are exploring how greater EHR documentation frequency may signal patient safety risk. Nurses' EHR documentation (beyond required) and optional nursing text documentation were related to patient mortality and outcomes, demonstrating that patients who died had more vital signs documented and free-text comment frequency than patients with the same comorbidity risk who survived (Collins et al., 2013). The authors controlled for acuity levels through risk stratification and the age-adjusted Charleston comorbidity index. A higher likelihood of cardiac arrest was associated with a higher frequency of comment and vital sign documentation, possibly reflecting a nurse's concern regarding the patient's health status (Collins et al., 2013). Collins et al. (2013) stated that increased optional EHR documentation is a sign of nursing clinical judgment to increase surveillance and vigilant monitoring for patients who may be at risk of clinical deterioration.

Numerous health-care technologies exist to reduce ADEs in the medication-use process, such as the EHR, computerized provider order entry (CPOE) with decision support, electronic medication reconciliation, automated dispensing systems, bar code labeling of medications, and smart (intravenous) pumps (Cheng, 2011). The use of CPOE for drug order processing decreases the likelihood of error on that order by 48%, which researchers estimate in 2008, contributed to a 12.5% reduction in medication errors, or the prevention of over 17 million medication errors in the United States in that year (Radley et al., 2013). On the other hand, other researchers identified an increase in duplicate medication order errors with the implementation of an EHR with CPOE, contributing to a system vulnerability (Wetterneck et al., 2011). Wetterneck et al. (2011) stated the duplicate orders may have been be due to restricted visibility of previously

ordered and administered medications, limiting provider situational awareness. Another study found no significant association between the use of CPOE, EHRs, or bar coding and medication error rates (Flynn, Liang, Dickson, Xie, & Suh, 2012).

Safety Incident

The middle section of the GRM presents the incident, which arises from either unintended or intended acts (Runciman et al., 2009). An incident is an event that results in, or may result in, patient harm (The World Alliance for Patient Safety Drafting Group et al., 2009). An incident that could have led to damage, loss, or harm but did not reach the patient is a near miss, whereas an adverse event results in patient harm (Runciman et al., 2009).

Incidents are divided into classes, which are descriptive terms to group similar things based on shared features (The World Alliance for Patient Safety Drafting Group et al., 2009). The idea is to classify incidents based upon similarity, to allow production of quantitative reports (Runciman et al., 2006). For example, patient demographics are personal attributes that may have contributed to the incident, such as uncooperative behavior. The timing of the incident (night shift) may be important for some events such as delirium. The method of detection is the condition that leads to discovery of the incident, for example, a practitioner noticing an error, a patient monitor or alarm, or a risk assessment tool (Runciman et al., 2009). To illustrate, an infusion pump that is set up incorrectly delivering a sedative overdose causing respiratory depression would be classified as a medication and equipment incident type (Runciman et al., 2009).

Outcomes and Consequences

Patient safety incidents have personal and organizational outcomes. A patient outcome is the effect on the patient that is attributable to a safety incident. Harm involves an impairment of the structure or body function, including injury, disease, disability, suffering, and death (Runciman et al., 2009). The level of harm includes its severity and duration, and implications for treatment. An organizational outcome is the effect on an organization attributable to a safety incident, such as adverse publicity or increased resource use for patient care, or legal consequences (Runciman et al., 2009; The World Alliance for Patient Safety Drafting Group et al., 2009).

Potential consequences for the healthcare organization are immediate or planned action, and an evaluation on resource impact.

Application of the GRM to This Study

Although a researcher could study any number of contributing factors to patient safety, this study includes one example each of organizational and human factors within the GRM. Safety culture is the ideal organizational factor because of its widespread study and importance for nursing. Experts consider the inclusion of safety culture as a high priority in patient safety research (Shekelle et al., 2013). Teamwork serves as a human factor within the GRM model for this study because it is a common cause of inadvertent patient harm (Leonard et al., 2004). In health care, one of the most frequently measured safety features is teamwork (Flin et al., 2006). Nurses contribute to patient safety by coordinating and integrating aspects of care quality provided by the interdisciplinary team (Mitchell, 2008). Ineffective teamwork dynamics and ineffective communication may contribute to patient harm (Reid & Bromiley, 2012). This study will explore outcomes through adverse events that result in patient harm, which differs from the GRM definition because the outcome is actual harm, excluding near misses or good catches. Adverse events are the health care patient outcome in this study due to their significance, and high morbidity and mortality.

The previous section described GRM concepts. The following literature review will be organized according to the conceptual framework presented in Figure 2.2. The first section will discuss safety climate as an organizational factor. The second section will discuss teamwork as a human factor, and then unit size as a potential confounding factor when studying patient safety in hospitals. The third section will discuss literature related to adverse events, as an example of patient outcomes within the GRM. The final section will provide a brief synthesis of the literature, highlighting important unanswered questions, leading to the specific aims.

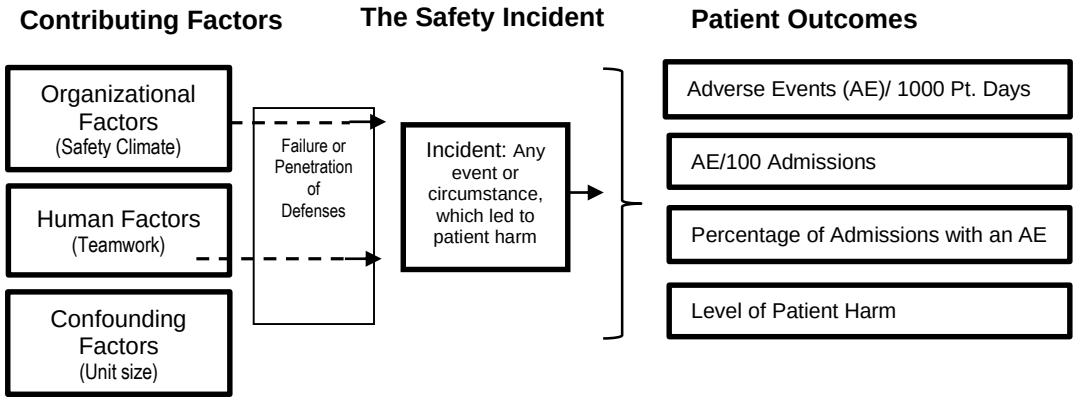

Figure 2.2. Schematic representation of the relationship between contributing factors, the safety incident, and patient outcomes. Adapted from "An Integrated Framework for Patient Safety, Quality and Risk Management: An Information and Incident Management System Based on a Universal Patient Safety Classification," by W. B. Runciman et al., *Quality and Safety in Health Care, 15*(Supplement 1), p. i82-90.

Organizational Factors: Safety Climate

Safety climate refers to perceptions regarding procedures and behaviors in the work environment that indicate the importance of safety compared to other goals (Flin et al., 2006; Zohar, 1980). When researchers use questionnaires to assess group-level perceptions of safety, they are assessing climate because climate is more readily measurable than culture, which includes behavior, values, and competencies, which can be difficult to assess via survey (Sexton et al., 2006). A methodological study comparing survey-based and observation-based evaluations of safety culture found that observation-based assessments of culture were limited with insufficient score differentiation between units (Freeth et al., 2012). Thus, while some experts recommend qualitative approaches like ethnography for studying safety culture (Halligan & Zecevic, 2011), doing fieldwork via observation may not provide greater differentiation between units than a safety climate survey. Freeth et al. (2012) recommended using validated climate surveys because survey-based assessments of safety climate appear to be in close agreement with summary audit scores representing quality of care. The National Quality Forum (2009) recommended that healthcare organizations measure their safety culture (presumably via a climate survey) because this determines how much risk an individual practitioner will take to protect patient safety. Numerous survey measures exist to measure the safety culture concept

via a safety climate. The following section will review three commonly used survey measures of safety climate.

Measurement of Safety Climate

AHRQ Hospital Survey on Patient Safety (HSOPS). The AHRQ Hospital Survey on Patient Safety (HSOPS) measures staff perceptions of patient safety culture in their unit or department, and for the hospital as a whole (AHRQ, 2015a). The HSOPS may be used to measure patient safety culture, trace changes over time, and evaluate cultural impact of patient safety interventions (AHRQ, 2015c). The authors developed the HSOPS through literature review, appraising existing safety culture surveys, and conducting interviews with hospital staff (Sorra & Nieva, 2004). Researchers pre-tested the survey with hospital staff to ensure understandability, and then pilot tested it with 1,400 hospital employees from 21 hospitals in six states across the United States. Analyses conducted during the pilot study led to dropping 26 of the original items (Sorra & Nieva, 2004). The developers concluded that the psychometric testing of the HSOPS provides solid evidence supporting the final 12 dimensions and 42 retained items due to the confirmatory factor analysis results, reliabilities, intercorrelations among the dimensions, and analysis of variance results (Sorra & Nieva, 2004).

HSOPS strengths are its comprehensive and specific nature, ease of use, and free access to the public (AHRQ, 2015a). The psychometrics are rigorous at the individual, unit, and hospital analytic levels (Sorra & Dyer, 2010). The AHRQ provides a comparative database with a central repository to encourage benchmarking between institutions (AHRQ, 2012a). Additionally, researchers have translated it into 18 languages.

On the other hand, the HSOPS has some limitations. First, although a comparative database is available, participating hospitals are not randomly selected to participate (Sorra et al., 2014). Non-random (i.e., non-probability) sampling may introduce bias because the results may not represent the population as a whole (Kerlinger & Lee, 2000). Second, hospitals may administer the survey in a variety of ways (paper only, Web-only surveys, combination methods), which can lead to differences in survey responses (Sorra et al., 2014). Web-based data collection may present problems such as incomplete responses, multiple submissions, and

security/data integrity issues (Waltz, Strickland, & Lenz, 2010a). Third, low response rates may be an issue, due to the perception that survey completion is excessively burdensome for staff (Adams-Pizarro, Walker, Robinson, Kelly, & Toth, 2008). In 2014, the average HSOPS hospital response rate was 53% (AHRQ, 2014). Low response rates make comparisons difficult.

Patient Safety Climate in Healthcare Organizations (PSCHO). A second commonly used survey measure of safety climate is the Patient Safety Climate in Healthcare Organizations (PSCHO). The purpose of the PSCHO is to assess attitudes towards patient safety and organizational culture based upon theoretical support from high reliability organization (HRO) literature (Singer et al., 2003). The Patient Safety Center of Inquiry (PSCI) at the VA Palo Alto Health Care System constructed the PSCHO through adaptations of five existing surveys (with permission; Singer et al., 2003), and feedback from pilot testing. The PSCHO structure contains 5 safety dimensions: organization, department, production, reporting/seeking help, and shame/self-awareness (Pumar-Méndez, Attree, & Wakefield, 2014). It contains 82 closed-end items rated on a 5-point Likert scale, with the scale: always, frequently, sometimes, rarely, and never.

In addition to recognized validity and reliability (Singer et al., 2007), the PSCHO has numerous strengths. It is useful for comparing healthcare workers safety perceptions with workers from other high reliability industries, such as aviation (Gaba, Singer, Sinaiko, Bowen, & Ciavarelli, 2003). The PSCHO has been used to compare worker perceptions from different work areas (clinical and non-clinical) and disciplines (medicine, nursing) within hospitals. Safety climate differs by discipline, hospital, and between and within work areas (Singer, Gaba, et al., 2009). Research supports the relationship between safety climate measured by the PSCHO and patient safety indicators. For example, patients have a lower risk of decubitus ulcer when their healthcare workers perceive a better safety climate (Singer, Lin, Falwell, Gaba, & Baker, 2009).

There are a few limitations with the PSCHO survey. Relationships between some nurse-sensitive patient outcomes such as medication errors, falls and infections remain unknown (DiCuccio, 2014) with the PSCHO. The PSCHO survey is based upon high-reliability theory (HRT), which is not the selected theoretical framework for this study. A principle of HRT is

preoccupation with failure, which requires scrutiny of every error. This study is not focusing on error, but patient harm. Errors may or may not lead to patient harm.

Safety Attitudes Questionnaire (SAQ). A third survey for measuring safety culture is the Safety Attitudes Questionnaire (SAQ). The SAQ "elicits a snapshot of the safety culture through surveys of frontline worker perceptions" (Sexton et al., 2006, p. 2). Researchers from an academic setting (the University of Texas) developed the SAQ in the mid-2000s, with funding from the Robert Wood Johnson Foundation and AHRQ (University of Texas Center for Healthcare Quality and Safety, 2014). The academic researchers developed the SAQ in response to a growing need for a valid and reliable measure of safety culture in healthcare.

Researchers or administrators may use the SAQ for tracking changes over time, benchmarking, designing improvement interventions, diagnosing current safety culture status, evaluating the impact of patient safety interventions, and fulfilling regulatory requirements (Devers et al., 2004). The SAQ contains six dimensions: safety climate, teamwork climate, job satisfaction, perceptions of management, stress recognition, and working conditions (Sexton et al., 2006). The 30 items use an ascending 5-point Likert scale with anchors of one indicating, "strongly disagree" and five indicating, "strongly agree." Completion time of the SAQ is typically 10-15 minutes (Sexton, et al., 2006) and may be used in either paper or electronic format.

The SAQ has strong validity and reliability. The SAQ developers established the SAQ's construct validity using exploratory factor analysis during development. They also performed multilevel confirmatory factor analysis to test the fit of the expected six-factor structure with a large sample of healthcare workers (n = 10,843) from 203 clinical areas (Sexton, et al., 2006). The 30 items fit the final model satisfactorily (Sexton et al., 2006). Sexton et al. (2006) established criterion-related validity of the SAQ by correlating its scores with James Reason's Checklist for Assessing Institutional Resilience. The independent safety climate measurement using the Reason's instrument produced the expected results, supporting convergent validity (Sexton et al., 2006). Other researchers established SAQ criterion-related validity via predictive validity testing. University of Pittsburgh researchers found a relationship between perceptions of management and safety climate (SAQ-ICU version domains) and hospital mortality and length of

stay (Huang et al., 2010). When ICU staff perceptions of management decreased by 10%, there were increased odds of death (Huang et al., 2010). Also, SAQ dimensions predicted patient safety grade in a study of 12 ICUs (Etchegaray & Thomas, 2012). Evidence supports the SAQ's criterion validity due to convergent validity and predictive validity testing.

The internal consistency reliability of the SAQ is evident with Cronbach's alpha scores ranging from 0.68 to 0.81 (Colla, Bracken, Kinney, & Weeks, 2005) and 0.65 to 0.83 (Fleming, 2005). A number of researchers have declared that 0.7 is an acceptable reliability coefficient although there is no evidence to support an ideal number (Kerlinger & Lee, 2000). The original SAQ developers assessed overall scale reliability using Raykov's ρ coefficient because Cronbach α underestimates composite reliability at the population level (Raykov, 1997). The Raykov ρ coefficient was 0.90, which indicates strong reliability of the SAQ (Sexton et al., 2006).

A potential perceived limitation of the SAQ is confusion regarding terminology and content validity. If the SAQ measures the multidimensional concept of safety climate, it is confusing that one of the six domains is called "safety climate." Also, the SAQ was unable to predict the number of adverse events reported within the last 12 months, although the HSOPS could not predict this, either (Etchegaray & Thomas, 2012).

The SAQ is the most extensively used cultural assessment tool in health care (Pronovost & Sexton, 2005). The six SAQ dimensions provide targets for multifaceted interventions, with the ability to demonstrate change in multiple safety dimensions. SAQ dimensions demonstrated an association with nurse-sensitive outcomes such as the odds of decubitus ulcers and nurse injury (Taylor et al., 2012). Two dimensions of the SAQ, safety and teamwork climate, provide good measures of organizational and human factors within the GRM. The SAQ will be used as the instrument for quantifying safety climate in this study due to its strong validity and reliability, widespread use, association with outcomes, and congruence with the GRM conceptual framework.

The previous section described and compared three safety climate instruments, and the rationale for selecting the SAQ. The following section will present teamwork as a human factor,

rationale for using the SAQ teamwork climate scale, and unit characteristics as a potential confounder within the proposed conceptual framework.

Human Factors: Teamwork

Teamwork is significant in patient safety research because patient safety is created by teams (Knox & Simpson, 2004). Teamwork has a significant impact on patient outcomes. In a retrospective chart review examining the impact of team behaviors on 30-day outcomes, patients were more likely to die or suffer a major complication when surgical teams exhibited infrequent team behaviors, even after adjusting for anesthesia risk category (Mazzocco et al., 2009). Similarly, a cross-sectional study demonstrated that a strong teamwork was negatively associated with adverse patient outcomes such as decubitus ulcers (Taylor et al., 2012). A study using augmented trigger tool methodology found that communication and teamwork problems were identified in approximately one-third of adverse events (Wong et al., 2015).

Salas, Cooke, and Rosen (2008) stated that teamwork is "the interdependent components of performance required to effectively coordinate the performance of multiple individuals" (p. 541). Teamwork as a construct has a wide range of definitions. A recent systematic review found over 27 unique constructs of teamwork in the intensive care unit literature (Dietz, Pronovost, Mendez-Tellez, et al., 2014). The most commonly studied teamwork constructs were communication, leadership, collaboration, coordination, and team climate/culture (Dietz, Pronovost, Mendez-Tellez, et al., 2014). Effective teamwork requires a cluster of competencies, such as teamwork-related knowledge, skills, and attitudes (AHRQ, 2005). Teams are useful when errors lead to grave consequences, when tasks are stressful requiring multiple quick decisions, and the lives of others depend on collective insight (Salas et al., 2008). Teamwork is a human factor that contributes to patient safety because safety outcomes depend upon performance of multiple persons working together for optimal patient care.

Teamwork Culture Versus Teamwork Climate

Teamwork climate provides a way to measure surface features of teamwork from attitudes and individuals at a certain time point, similar to safety climate (Flin, Mearns, O'Connor, & Bryden, 2000; Halligan & Zecevic, 2011). Experts do not have consensus regarding how to

conceptualize and operationalize teamwork constructs (Dietz, Pronovost, Benson, et al., 2014). Teamwork can be difficult to measure, and methods of teamwork assessment include survey of team climate or culture, direct observation, self-assessment, peer assessment, and outcome measurement (O'Leary, Sehgal, Terrell, & Williams, 2012). Many behavioral marker systems to measure teamwork focus on observing a specific task (Dietz, Pronovost, Benson, et al., 2014). Because direct observation can be difficult in healthcare settings when not all team members are present at the same time, a survey instrument may be best for assessing teamwork climate for general medical units (O'Leary, Sehgal, et al., 2012).

Measurement of Teamwork Climate

In a review examining teamwork tools associated with patient outcomes in internal medicine, less than 20% of teamwork tools examined the effect of teamwork on patient outcomes (Havyer et al., 2014). Of the 13 tools that correlated with patient outcomes, the SAQ had the strongest validity evidence (Havyer et al., 2014). The teamwork climate scale within the SAQ has demonstrated good psychometric testing and correlation with patient outcomes.

Other teamwork surveys focus on particular groups/roles, such as nurses (Kalisch, Lee, & Salas, 2010) or healthcare students (Hollar, Hobgood, Foster, Aleman, & Sawning, 2012), or settings such as simulation laboratories (Shrader, Kern, Zoller, & Blue, 2013), medical units that conduct interprofessional rounds (O'Leary, Boudreau, Creden, Slade, & Williams, 2012), or intensive care units (Weller et al., 2011). Another option for measuring teamwork is the TeamSTEPPS Teamwork Perceptions Questionnaire (T-TPQ), which measures individual attitudes related to the core components of teamwork (AHRQ, 2008). The T-TPQ may be used for site assessment prior to TeamSTEPPS training. Since this study will not be testing a teamwork intervention, the T-TPQ would not be an ideal choice for operationalizing teamwork. In addition, this study is focusing on the entire team's perceptions of teamwork, so a survey designed for a particular healthcare specialty group is not indicated. The teamwork climate subscale within the SAQ was selected to operationalize teamwork because of its strong validity and reliability, interprofessional scope, and correlation with patient outcomes.

Confounding Factors (Unit Characteristics)

Researchers may select nursing unit/department size (i.e., number of patient beds) as a characteristic that may affect perceptions of safety, teamwork, and patient outcomes, and account for this within statistical models. A descriptive study with data from 286 medical-surgical units and 3689 survey responses found that safety climate differed depending on unit size (Hughes, Chang, & Mark, 2012). Nurses working on smaller units described their workgroup as more accepting of safe work practices, and more likely to reveal and share errors compared to nurses working on larger units (Hughes et al., 2012).

Unit size may also influence teamwork. A secondary analysis with over 2,000 nursing staff working on 53 medical-surgical, intermediate, intensive care, and rehabilitation units in four Midwestern hospitals found that nurses working on smaller units reported higher teamwork scores on the Nursing Teamwork Survey (Kalisch, Russell, & Lee, 2013). Results indicated that backup behaviors, leadership capabilities, and trust decrease as the number of beds get larger (Kalisch et al., 2013). Kalisch et al. (2013) cited evidence regarding potential reasons for these findings. Nursing units with more beds may have larger (physical) space between coworkers, which may reduce face-to-face interactions, closeness of workstations, and ability for team members to meet (Kalisch et al., 2013).

Unit size may also be related to patient outcomes. A study of ICU best practice examples of physical design features found that larger ICUs with more beds had more cases of infection (Rashid, 2014). A narrative review found that larger nursing units experience increased medication administration errors compared to smaller nursing units (A. M. Parry, Barriball, & While, 2015). On the other hand, a cross-sectional descriptive study found no relationship between unit size and nurses' self-reported medication errors (Hung, Lee, Tsai, Tseng, & Chang, 2013). Thus, there is conflicting literature regarding the effect of unit size on patient outcomes.

It is difficult to study the effect of unit characteristics (such as size) on quality of care because studies with unit-level analyses have small sample sizes, which may bias results (Brewer & Verran, 2013). Due to the need for continued study of this variable, nursing unit size is included as a potential confounding factor of adverse events.

Patient Outcomes: Adverse Events

Adverse events are injuries caused by medical care (Wachter, 2012). An adverse event is negative/undesirable, directly involves the patient, and is the result of the healthcare process (Walshe, 2000). Adverse events can be divided into "sins of omission" and "sins of commission" (West, 2000). Errors of omission may involve delayed diagnosis, subtherapeutic medication doses, and failure to provide indicated treatments (Hayward, Asch, Hogan, Hofer, & Kerr, 2005). Conversely, "sins of commission" events are traceable to specific interventions, such as medication errors or poor surgical technique (West, 2000).

Researchers may approach patient safety from either an error-oriented or an injury-oriented approach (Layde et al., 2002). An error-oriented approach may include mistakes that do not harm patients, termed near misses (Layde et al., 2002). On the other hand, an injury-oriented approach focuses on patient harm from healthcare procedures or therapies, which may not be associated with error (Layde et al., 2002). Layde et al. (2002) advocated for an injury-oriented approach, so researchers focus on actual patient harm because most injuries are not due to negligence or error. The study of adverse events may be best suited to an injury-oriented approach, due to its patient-centered perspective.

Practitioners may generate errors during diagnosis, treatment or prevention (Garrouste-Orgeas et al., 2012). In the old view of human error, accidents are seen as the result of people not performing well in a safe system (Dekker, 2006). In contrast, the New View of human error avoids judging individuals when errors occur, but seeks to understand why someone's assessment and actions made sense at the time in order to prevent reoccurrence. This is the local rationality principle: People do what makes sense given the situation, operational pressures, and organizational norms at the time (Dekker, 2011). Errors become a symptom of a system problem that anyone could be vulnerable to repeat.

Researchers are exploring the relationship between clinician-detected adverse events and patient-reported incidents. Clinicians may miss some symptomatic adverse events, and patient adverse event reporting may be more reliable than clinician assessment (Basch, 2014). The National Cancer Institute (NCI) is developing a library of patient-reported adverse events

during cancer treatment, which 93% of practitioners believe will improve their comprehension of the patient experience (Bruner et al., 2011). In a Norwegian study of 63 hospitals, there was a strong correlation between patient-reported incidents and harm measured through the GTT (Bjertnaes, Deilkas, Skudal, Iversen, & Bjerkan, 2015). Future adverse event reporting may routinely include patient-reported incidents.

Adverse Event Risk Factors

Longer length of stay is associated with GTT-identified adverse events (Classen et al., 2011; Hwang, Chin, & Chang, 2014; Rutberg et al., 2014) and GTT-identified ADEs (Härkänen et al., 2015). Chart reviews with a greater number of triggers are also associated with adverse events (Hwang et al., 2014). Patient preferences (non-compliance with recommended treatment) and uncooperative behavior (walking without assistance) are patient factors that place patients at risk for adverse events (Wong et al., 2015). Lastly, patients who experienced adverse events tend to be older and male compared to patients without an adverse event (Classen et al., 2011).

Preventability of Adverse Events

When adverse events occur due to medical error, they may be preventable (Garrouste-Orgeas et al., 2012). According to the Ohio Hospital Association (2008), a never event or a preventable adverse event is "a medical mistake that results in death, loss of a body part, disability or loss of bodily function." The National Quality Forum (2012) estimated that the United States has spent $17-$29 billion per year in healthcare expenses, disability, and lost worker productivity due to preventable errors. One of the challenges with preventability is clearly defining it. A research team conducted a systematic review of preventable harm in health care and found seven themes or definitions for preventable harm (Nabhan et al., 2012). The researchers were unable to find empirical evidence regarding the validity and reliability of available preventable harm definitions (Nabhan et al., 2012).

Many practitioners feel that some patient harm may be inevitable in health care, yet others disagree. Harms that appear non-preventable today may become preventable in the future (G. Parry, Cline, & Goldmann, 2012) due to advancements in technology and procedures, and development of new medications. United States hospitals are required to track and analyze

patient harm to participate in Medicare (U.S. Department of Health and Human Services, n.d.). Mandatory reporting provides data, but U.S. hospitals have not gained a tremendous amount of insight regarding why adverse events occur (Nemeth et al., 2006). Lack of insight regarding causes of adverse events may be due to measurement issues.

Measurement of Adverse Events

Numerous approaches to measure adverse events exist, and previous researchers have compared eight methods for measuring health care errors and adverse events (Thomas & Petersen, 2003). Of these eight methods, error-reporting systems, administrative data analysis, and chart review were viable measurement options for this study. The following section will present advantages and disadvantages of these approaches, discussing psychometrics within the context of validity and reliability.

Voluntary error reporting systems. The first approach to measuring adverse events is through the hospital voluntary error reporting system, or the healthcare incident reporting system. Voluntary reporting systems focus on near misses (errors that did not cause patient harm) to find system vulnerabilities before serious harm occurs (Institute of Medicine, 2000). The advantages of error reporting systems are its ability to detect latent (system) errors, to provide multiple perspectives, and ability to be incorporated into routine hospital operations (Thomas & Petersen, 2003). They are also inexpensive and relatively easy to obtain (Sharek, 2012). The disadvantages to error reporting are its reliance on practitioners who may be reticent to report because of being too busy, fear of lawsuits, or concern about impact on professional reputation (Thomas & Petersen, 2003). In addition, voluntary reporting does not reveal all diagnostic errors (DeFeijter, DeGrave, Muijtjens, Scherpbier, & Koopmans, 2012) or adverse events (Blais, Derson, Bartlett, & Tamblyn, 2008). Lastly, hospitals cannot use incident reporting systems to measure error rates due to under-reporting, and lack of clearly defined events (numerators) for populations at risk (denominators) to calculate rates (Pham, Girard, & Pronovost, 2013). Incident report systems may also generate too many reports that do not permit in-depth analyses of what went wrong (Pham et al., 2013).

A retrospective Dutch study assessed the reliability and validity of voluntary error reporting through sensitivity and specificity testing, by comparing it against the gold standard of patient record review and other types of reporting such as informal and formal patient and family complaints, and legal claims (Christiaans-Dingelhoff et al., 2011). The sensitivity (i.e., true positives) and positive predictive value of voluntary reports for adverse events was 3.6% and 4.6%, respectively. Healthcare professionals reported relatively few adverse events, and so the authors concluded that these data sources were not useful as a predictive method to detect adverse events (Christiaans-Dingelhoff et al., 2011). Another study found that incident reports only identified 15.5% of charts with adverse events (Blais et al., 2008). Sharek (2012) reported even lower validity of voluntary error reporting, that voluntary reports identify between 2 and 8% of harmful events (Sharek, 2012). Reasons for poor validity and reliability of voluntary error reporting systems may be: (a) practitioners' differing personal definitions of an adverse event, and (b) varied degrees of commitment to the self-reporting mechanism (Walshe, 2000).

Administrative data analysis. A second approach to measuring adverse events is through administrative data analysis. Administrative data results from running the healthcare system, such as enrolling people in health care plans, paying claims, determining reimbursement amounts, certifying coverage, tracking service utilization, and monitoring costs and performance (Kane & Radosevich, 2011). The advantages of administrative data analysis are that it has a large number of subjects, is representative, facilitates tracking people over time, allows direct comparisons between hospitals, uses available chart data, is inexpensive, and is less susceptible to selection biases compared to other methods (Kane & Radosevich, 2011; Sharek, 2012; Thomas & Petersen, 2003). For example, billing data provides an opportunity to quantify the costs of adverse events via increased length of stay in intensive care (Kaushal, Bates, Franz, Soukup, & Rothschild, 2007).

The disadvantages of administrative data analysis are that data may be incomplete, and there may be bias to "up code" certain conditions to increase reimbursement payment (Thomas & Petersen, 2003). A sensitive billing strategy may only detect less than half of adverse events (Bates, O'Neil, Petersen, Lee, & Brennan, 1995). Administrative data sets are large and clinical

information may be missing (Kane & Radosevich, 2011). Some experts state that administrative data identifies less than 10% of all harms (Sharek, 2012). The content validity of administrative data for detecting adverse events may be less than ideal because the information within data sets does not adequately reflect the scope of necessary information required for adverse event detection. Administrative data has poor sensitivity and specificity for identifying harm (Sharek, 2012).

Retrospective chart review. A third approach to measuring adverse events is though chart review. Retrospective chart review is considered the gold standard for error/adverse event detection (Blais et al., 2008). The 1991 Harvard Medical Malpractice Study (Brennan et al., 1991) established the standard by which to measure adverse events (Baker, 2004). It uses a two-stage chart review, with nurses to screen patient records that likely contain an adverse event, and physician review to confirm the presence of these identified adverse events. Advantages of chart review are its usage of existing data and common practice (Thomas & Petersen, 2003). Moreover, chart review measures "all cause" harm with the ability to provide a rate, such as the number of harms per 100 admissions or per 1,000 patient days (Sharek, 2012). The disadvantages of chart review are its expense, requirement for judgment about adverse event occurrence, incomplete records, and hindsight bias (Thomas & Petersen, 2003).

Some experts perceive retrospective chart review as obsolete due to newer, more efficient and more sensitive methods such as trigger methodologies (Sharek, 2012; Sharek & Classen, 2006). Triggers are sentinel words or events, such as certain drugs, antidotes, lab values, or stop orders that prompt reviewers to begin a more detailed chart audit (Resar, Rozich, & Classen, 2003). The advantages of trigger tool methodology are similar to retrospective chart review. Trigger tools offer the advantages of greater efficiency (20 minutes per chart) and ability to be population-specific (i.e., pediatric-focused; Kirkendall et al., 2012) with excellent specificity and good sensitivity (Sharek, 2012). The disadvantages of trigger tools are its demanding training requirement, intense use of resources, non-automation, and risk of incomplete documentation (Naessens et al., 2010; Sharek, 2012).

Measurement of Adverse Events Using the Global Trigger Tool (GTT)

In 2003, the IHI developed the GTT for measuring adverse events to improve hospital safety and track improvement (Griffin & Resar, 2009). Early adopters of the GTT are the Mayo Clinic, Baylor Health System (Dallas, Texas) and Adventist Health System (Orlando, Florida; Garrett Jr et al., 2013). The GTT uses an injury-oriented approach to patient safety because it measures adverse events with harm. Harm is defined as "unintended physical injury resulting from or contributed to by medical care that requires additional monitoring, treatment or hospitalization, or that results in death" (Griffin & Resar, 2009, p. 5). The GTT concentrates on adverse events resulting from direct care (commission) rather than issues related to substandard care (omission; Griffin & Resar, 2009). For example, a patient that is taking anticoagulants who develops a stroke from an intracellular bleed would have suffered an adverse event because the anticoagulant (commission) caused the event. On the other hand, a person with high blood pressure not treated with an antihypertensive that develops a stroke has experienced the devastating effects of poor care, but this would not be an adverse event according to the GTT because the event is due to omission of care (Griffin & Resar, 2009).

The GTT focuses on measuring harm to be congruent with the overall goal of patient safety to reduce patient injury or harm (Griffin & Resar, 2009). This broad definition of adverse events not requiring preventability or major disability (in contrast to previous studies) increases the sensitivity of the tool to detect greater adverse events (Classen et al., 2011). The IHI specifically requests that chart reviewers do not to attempt to determine preventability. Rather, users of the GTT record all events that are unintentional effects of medical care, regardless of perceived preventability.

GTT Triggers

Triggers serve as clues regarding the potential occurrence of an adverse event. Reviewers may find a number of positive triggers, but that does not indicate with certainty the presence of an adverse event (Griffin & Resar, 2009). For example, anti-emetic use is the most frequently identified trigger (Hwang et al., 2014; Naessens et al., 2010), but its identification does not necessarily indicate that an adverse event occurred. The focus is on finding adverse events,

not triggers. Some researchers have attempted to determine which triggers are most frequently associated with adverse events. Baylor measured adverse event yield, the percentage of all identified adverse events found positive for the particular trigger (Kennerly et al., 2013). The two triggers with the highest adverse event yield were occurrence of any operative complication and healthcare associated infection (Kennerly et al., 2013). Other researchers determined trigger positive predictive value, the rate at which a trigger is associated with an adverse event (Naessens et al., 2010). Triggers with high predictive values are return to surgery (Hwang et al., 2014; Naessens et al., 2010), in-unit procedure (Naessens et al., 2010), Clostridium difficile positive culture (Carnevali et al., 2013; Naessens et al., 2010; Unbeck et al., 2013) intubation/re-intubation (Naessens et al., 2010), healthcare associated infection (Hwang et al., 2014), any procedure complication (Hwang et al., 2014), "medication, other" (Hwang et al., 2014), occurrence of any operative complication (Hwang et al., 2014; Unbeck et al., 2013), intubation/reintubation (Hwang et al., 2014), admission to intensive care post-operatively (Unbeck et al., 2013). An average of 25-40% of patients experience at least one GTT-detected adverse event during hospitalization (Garrett Jr et al., 2013; Good, Saldana, Gilder, Nicewander, & Kennerly, 2011; Kennerly et al., 2014; Rutberg et al., 2014).

Adaptations to the GTT

Healthcare organizations have made adaptations to the GTT review process. Baylor Health System has revised the suggested medical record review order, based on reviewer experience (Good et al., 2011). They use nurse reviewers working for an external company that specializes in medical record reviews (Kennerly et al., 2013). Baylor and Adventist Health System have changed to a single person reviewer for conducting GTT reviews. Adventist moved to a single record review after 17 months' experience with a double record review (Garrett Jr et al., 2013). The percentage of harm severity rated *F* or greater was 40% 6 months before and 49% after the single reviewer process, which led to adoption of the single-reviewer process (Garrett Jr et al., 2013). Adaptations to the GTT review process are designed to achieve sustainable costs for an ongoing measurement system (Kennerly et al., 2014).

In addition to review process adaptations, healthcare organizations have also modified GTT data characteristics. The GTT detects adverse events that may have occurred prior to coming to the hospital because the GTT measures "what the patient experienced, not what happened within the hospital" (Griffin & Resar, 2009, p. 30). Some organizations seek to differentiate adverse events present on admission versus hospital-associated adverse events, so that organizations may prioritize safety efforts on hospital-associated events (Good et al., 2011). The GTT developers encourage hospitals to collaborate with long-term care facilities, physician offices, and clinics to address adverse events acquired outside the hospital (Griffin & Resar, 2009), as one organization determined that 38.8% of all adverse events were present on admission (Good et al., 2011).

Determination of adverse event preventability, omission of care, and customized triggers are other GTT adaptations. Some GTT review teams determine the preventability of an adverse event (Hwang et al., 2014; Kennerly et al., 2013; Landrigan et al., 2010; Schildmeijer, Nilsson, Arestedt, & Perk, 2012) or quality of documentation within the medical record to determine missing information (Sharek et al., 2011). Review teams are also documenting adverse events due to omission of care, which the original GTT excludes. Baylor's GTT reviews include omission of care, where the practitioner does not take action, and the patient may have benefitted, regardless of intention (Kennerly et al., 2013). Omission of care is part of a mortality review system at the Mayo Clinic due to discontent with trending GTT adverse event rates and lack of specificity to guide process improvement initiatives (Huddleston, Diedrich, Kinsey, Enzler, & Manning, 2014). The Hamilton (Ontario) Health Sciences department adopted triggers such as the Canadian Adverse Events Study screening tool criteria and customized criteria identified by physician stakeholders (Zimmermann et al., 2010). Researchers from the University of Toronto have added the following triggers: functional decline, in-hospital malnutrition, and hospital-acquired delirium (Wong et al., 2015). GTT triggers and data characteristics have evolved.

Healthcare systems are using the GTT methodology in concert with health information technology. Data collection and analysis may be facilitated through the electronic health record. A narrative review of 15 articles found that most automated trigger systems use a combination of

data from laboratory values, pharmacy data, patient demographics, and clinical care data, and tend to focus on detecting ADEs (Doupi, 2012). Sweden is using an automated version of the GTT, the Modified Automated GTT (MAG), which searches for triggers in medication, laboratory, ICD-9 codes, and unstructured text via text mining (Doupi et al., 2015). Compared to manual GTT review, the sensitivity of detecting adverse events with text mining was 60 to 100%, depending on trigger type, and the specificity between the triggers varied from 80 to 98% (Ohman, Keski-Kuha, Kaartinen, & Kujansuu, 2011). To facilitate data mining analysis, Baylor requires a structured description of the GTT-identified adverse event using the Situation-Background-Assessment format (Good et al., 2011). Northwestern Hospital combined the GTT with data warehouse mining, but found poor agreement between traditional trigger tool and data warehouse screening with only one-third of all adverse events detected by both methods (O'Leary et al., 2013). Health information technology will continue to advance the utility of the GTT.

As health information technology improves, GTT methodology will evolve to detecting triggers in real time, and perhaps preventing adverse events altogether. Researchers from the University of Toronto are using an augmented GTT methodology with prospective clinical surveillance to identify factors contributing to adverse events (Wong et al., 2015). An advanced practice nurse screens patient electronic health records daily, attends multidisciplinary rounds, and interacts with front-line staff daily to identify potential triggers. Medical record review occurs within 48 hours of an identified adverse event, and includes debriefing of staff involved in the case. This methodology has identified common contributing factors as policy and procedures, communication and teamwork problems, and medication process problems (Wong et al., 2015).

GTT usage has expanded to various population types such as patients with longer length of stays, pediatrics, surgery, and primary care, as well as international patients. Baylor's inclusion criteria for required length of stay is three days or more, which is two days longer than the original GTT requirement (Kennerly et al., 2013). The rationale is based upon 2008 data. Patients with a length of stay three days or more were substantially more likely to experience an adverse event than patients with a shorter length of stay (Kennerly et al., 2013). Cincinnati

Children's Hospital has adapted the GTT for pediatric populations, demonstrating that 25% of children experience at least one adverse event during hospitalization, which is two to three times higher than previous pediatric harm rates (Kirkendall et al., 2012). Researchers in the United Kingdom have developed a GTT screening process for primary care (de Wet & Bowie, 2009). For the outpatient surgery population, five surgical triggers were effective in identifying adverse events, with the pulmonary embolism/deep vein thrombosis (PE/DVT) having the highest positive predictive value of 58% (Rosen et al., 2011). Lastly, the GTT has gained international adoption. The GTT has been translated into the Danish (vonPlessen, Kodal, & Anhoj, 2012), Swedish (Schildmeijer et al., 2012), and Korean (Hwang et al., 2014) languages. Health care systems have adapted the GTT to make it congruent with their safety needs.

Validity of the GTT

Validity occurs when an instrument measures the characteristics it was intended to measure (Stommel & Wills, 2004), and includes three types: construct, content, and criterion-related (Soeken, 2010). To assess criterion-related validity, one would infer the probable standing of adverse events detected from the GTT compared to some other criterion (Soeken, 2010). Classen et al. (2011) evaluated the ability of the GTT to detect the incidence of adverse events compared to the AHRQ PSI and voluntary sentinel event reporting systems. The GTT had a sensitivity to detect patients with an adverse event of 94.9% and a specificity to detect patients with no events of 100% (Classen et al., 2011). In contrast, the patient safety indicator method had a sensitivity of 8.5% and specificity of 98.5%, and a local hospital's voluntary reporting system had a sensitivity of 0% and a specificity of 100% (Classen et al., 2011). The GTT had greater sensitivity to detect greater adverse events due to its broad definition of adverse events not requiring preventability or leading to major patient disability, in contrast to previous studies (Classen et al., 2011). Because researchers compared the GTT to existing measures (i.e., AHRQ patient safety indicator and voluntary reporting) and were able to demonstrate its ability to detect adverse events as well (if not better) than existing criteria (measures), the GTT has adequate criterion-related validity. The GTT is a valid measure for detecting adverse events.

Reliability of the GTT

Numerous studies have examined the GTT's interrater reliability. Factors that may affect GTT reliability are reviewer training (Classen, Lloyd, Provost, Griffin, & Resar, 2008), with highly experienced reviewers detecting more harm than newly trained GTT reviewers (Landrigan et al., 2010; vonPlessen et al., 2012), variation in definition of harm among reviewers and/or reviewer teams (Deilkas, 2013), and reviewer role, with physician agreement higher than non-physician reviewers (Classen et al., 2008). Internal review teams tend to perform reviews more reliably than external review teams (Landrigan et al., 2010; Sharek et al., 2011).

The Kappa coefficient, or Cohen's kappa, is a non-parametric measure of interrater reliability that determines the proportion of events consistently classified in the same category on both occasions, beyond that expected by chance (Vogt & Johnson, 2011; Waltz, Strickland, & Lenz, 2010b). Applied to the GTT, Kappa tests the amount of agreement between pairs of reviewers to see if the agreement could have occurred by chance (Classen et al., 2008). Experts have provided guidelines for the interpretation of Kappa (Landis & Koch, 1977; Viera & Garrett, 2005) using the following scale: less than 0: chance agreement, 0.01-0.20: slight agreement, 0.21-0.40: fair agreement, 0.41–0.60: moderate agreement, 0.61-0.80: substantial agreement and 0.81-0.99: almost perfect agreement.

GTT studies with Kappa .01-.20 (slight agreement). A Swedish study with five GTT teams demonstrated large variation between teams, with a combined unweighted Kappa of 0.20 (Schildmeijer et al., 2012), which indicates only slight reliability (Waltz et al., 2010b). However, one of the five teams documented more than three times as many adverse events as the others because they had a different definition of harm than the GTT proposes (Deilkas, 2013).

GTT studies with Kappa .41-.60 (moderate agreement). Other researchers found moderate interrater agreement ($\kappa = .45$) on the GTT (Mattsson, Knudsen, Lauritsen, Brixen, & Herrstedt, 2013). Mattsson et al. (2013) stated that the ability of the GTT to detect change in harm rates (i.e., responsiveness) is poor due to the moderate interrater agreement. Agreement among nurse reviewers at the Mayo Clinic also varied for adverse events, with Kappa coefficients ranging from 0.40 - 0.60 (Naessens et al., 2010).

GTT studies with Kappa .61-.80 (substantial agreement) or 0.81-0.99 (almost perfect agreement). A Korean study assessed agreement using a random sample of 60 records and found substantial agreement among reviewers (κ = .735; Hwang et al., 2014). At Cincinnati Children's Hospital, nurse reviewers had substantial agreement (κ = .63) on the presence or absence of an adverse event in the pediatric population (Kirkendall et al., 2012). In North Carolina hospitals, internal GTT review teams performed more reliably (κ = .64-.93) than external reviewers (κ = .40-.72; Landrigan et al., 2010; Sharek et al., 2011). Other GTT teams have reported very good reliability (91.8% reproducibility between nurse reviewers; Adler et al., 2008) and 78% agreement between analysts for harm identification (Garrett Jr et al., 2013).

Other methods for determining GTT reliability. Baylor Health System asked their corporate patient safety team to review a list of GTT-detected adverse events to establish reliability. Communications from the team indicated strong agreement that the GTT-identified events did represent adverse events (Kennerly et al., 2013). The Adventist Health System asked an independent analyst to re-review adverse events to test interrater reliability more informally, which resulted in 78% agreement between analysts (Garrett Jr et al., 2013).

Advantages of the GTT

The GTT holds numerous advantages for measuring adverse events. Random record selection generates a sampling approach that permits hospitals to establish harm rates and monitor improvement (Griffin & Resar, 2009). The GTT discovers more adverse events than voluntary reporting, such as incident reports (Nilsson, Pihl, Tagsjo, & Ericsson, 2012). The GTT is useful due to its ability to identify patient harm, short time limit, intuitive triggers, and interdisciplinary representation to consider patient care from different perspectives (Schildmeijer, Nilsson, Perk, Arestedt, & Nilsson, 2013).

Hospitals have modified and adapted the GTT to increase its usefulness for tracking adverse events. Changes have been made to the number of reviewers, review process, and triggers. The GTT may be used with an electronic health record. It has been translated into multiple languages supporting international use. Hospitals have modified the GTT to make it congruent with their safety needs.

Limitations of the GTT

The GTT also holds some limitations for measuring adverse events. The GTT may duplicate existing adverse event data available from other sources. In Denmark, a team established that the GTT duplicates existing data within the Danish Lung Cancer Registry (Lipczak, Neckelmann, Steding-Jessen, Jakobsen, & Knudsen, 2011). Furthermore, results may differ depending on team training, review processes, and documentation (vonPlessen et al., 2012). Some critics state that the GTT underestimates harm because it does not detect diagnostic errors or errors of omission (G. Parry et al., 2012). The GTT may also inflate harm rates because it includes non-severe temporary harms (G. Parry et al., 2012). GTT limitations are its orientation to physician actions, lack of omission in care as a trigger, imprecise triggers, and missing nursing triggers (Schildmeijer et al., 2013).

Despite the GTT's limitations, it has greater sensitivity and specificity than the AHRQ PSI method (Griffin & Resar, 2009), voluntary reporting systems (Griffin & Resar, 2009), and administrative database algorithms (Sharek, 2012). The GTT is the most widely used global measure of patient safety (Pronovost & Wachter, 2013). It represents significant progress in adverse event detection methodology because it finds ten times more confirmed, serious events than voluntary reporting and AHRQ PSIs (Classen et al., 2011). Using the GTT, experts estimate that adverse events occur in one-third of hospital admissions (Classen et al., 2011). For these reasons, adverse events will be measured using a modified GTT procedure.

Summary

The U.S. healthcare system continues to be afflicted by adverse events, which may be caused by human factors, organizations, and technical failures (Smits et al., 2010). Researchers have developed numerous conceptual frameworks for studying patient safety. The GRM is a comprehensive framework for exploring contributing factors to patient safety. Because the GRM includes contributing factors to patient safety (which may include safety culture and teamwork) and the consequences of failed defenses (adverse events), it is the ideal framework for this study. The evidence regarding the effect of contributing factors to patient safety outcomes remains conflicting. Lack of statistically significant relationships of safety and teamwork culture with

adverse events may be due to measurement issues. This study will address these gaps in knowledge by exploring the relationship between unit-level contributing factors to patient safety and GTT-identified patient harm. The project will use a descriptive, cross-sectional design, with a retrospective, modified IHI GTT chart review methodology to describe relationships among safety and teamwork climates with GTT-identified adverse events in a Midwestern U.S. hospital. Safety and teamwork climate data will be collected from the SAQ (Sexton et al., 2006). Patient harm will be assessed using the IHI GTT for Measuring Adverse Events (Griffin & Resar, 2009).

The specific aims of the study are:

Aim 1. Describe self-reported safety climate and teamwork climate among an interprofessional group of providers working on 32 hospital inpatient units, measured using the SAQ.

Aim 2. Explore the nature of adverse events identified using a modified Institute for IHI GTT chart review methodology.

Aim 3. Examine to what extent unit-level safety climate and teamwork climate (and their potential interaction) predict frequency of adverse events (per nursing unit) as detected via the modified IHI GTT chart review format, controlling for unit characteristics.

Chapter 3: Methods

The purpose of the study was to explore relationships between organizational and human factors with adverse events detected using a modified trigger-tool methodology. The specific aims were:

Aim 1. Describe self-reported safety climate and teamwork climate among an interprofessional group of providers working on 32 hospital inpatient units, measured using the SAQ.

Aim 2. Explore the nature of adverse events identified using a modified Institute for Healthcare Improvement (IHI) Global Trigger Tool (GTT) chart review methodology.

Aim 3. Examine to what extent unit-level safety climate and teamwork climate (and their potential interaction) predict frequency of adverse events as detected via the modified IHI GTT chart review format, controlling for unit characteristics (size and type).

This chapter organization includes six sections. The first three sections present the research design, sample, and measurement with associated instruments. Sections four through six outline the study procedures, data analysis, and study limitations.

Design

This study used a descriptive, non-experimental, cross-sectional design, with a retrospective chart review methodology. A descriptive design was selected because its main objective is to portray characteristics of individuals, situations, or groups accurately (Polit & Hungler, 1995). In this study, characteristics included safety and teamwork culture (measured via a climate survey) and potential relationships with GTT-identified adverse events. Safety and teamwork culture were pre-existing unit characteristics, and were not manipulated in this study. Non-experimental designs do not involve manipulation of variables or group assignment (Kerlinger & Lee, 2000). Nurses and health care workers were already assigned to groups, i.e.,

the units or departments where they chose to work. Variation stemmed from differences between units at a single time point, i.e., a between-participants (units) approach.

Sample

The study population was inpatient, nursing units/departments that served adult patients. The sampling approach was a non-probability, convenience sample. The convenience sample was comprised of 32 nursing units/departments from one 750+-bed Midwestern U.S. regional acute care hospital that employed over 1,000 nurses. Thirty-two of the 79 hospital inpatient units/departments with greater than 40% SAQ response rate participated in the study. SAQ respondents included nurses, nursing assistants, therapists (respiratory, physical, occupational, etc.), radiology technologists, pharmacists, technicians (pharmacy and operating room), physicians (mostly residents), and "other" team members when there were less than five similar respondents per group. (Access to SAQ scores by provider type was unavailable.) Individuals self-selected the unit where they worked most frequently. The unit types were critical care, intermediate, medical-surgical, OB (obstetrics) gynecology and procedural units. Procedural units included multiple cardiac departments (such as catheterization lab, testing, post procedure recovery, imaging), radiology, CT (computed tomography), endoscopy, inpatient dialysis, and surgery. The unit size (number of beds per unit) ranged from five to 86 beds.

Sample Size

The sample size, 32 units, was derived from an a priori power analysis using PASS 11.0 software to define the unit sample size necessary for linear multiple regression analyses with three predictors in the model (safety climate, teamwork climate, and potential interaction). A sample size of 32 achieved >88% power to detect an R-squared of .20 attributed to one independent variable using an F-test with a significance level (alpha) of .05. The variables tested were adjusted for an additional two independent variables with an R-squared of .20. A sample size of 32 achieves >80% power to detect a medium effect (R^2 =.20) attributed to three predictors at alpha = .05 (Cohen, 1988).

Inclusion and Exclusion Criteria for Nursing Units/Departments

This study was designed to determine unit-level differences in contributing factors to patient outcomes. Organizational (safety climate) and human factors (teamwork) were measured at the unit level. To be included in the study, units were required to have a response rate of 40% or greater on the Safety Attitudes Questionnaire (SAQ).

Response rates ranged from 40.22% to 100%. The mean response rate was 62.75%, and the median was 64.50%. The 50th percentile response rate was 64.50%. To determine the distribution of scores, the skewness ratio was calculated by dividing the skewness statistic by the standard error (.342 ÷ .414), which was 0.826. This number 0.826 falls within the range of -2 to +2, so response rate was not skewed, indicating that response rate data followed a normal distribution. There were 15 units with less than 60% response rate. (See Figure 3.1.)

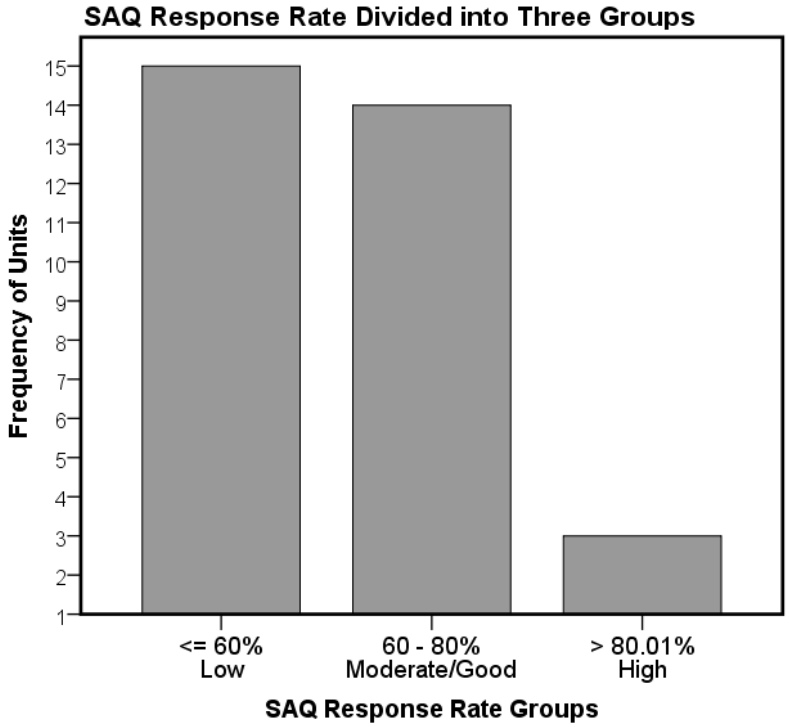

Figure 3.1. Bar chart of Safety Attitudes Questionnaire (SAQ) nursing unit response rate divided into three groups. Low response rate was less than 60%, moderate/good was 60-80%, and high response rate was greater than 80%. (*n* = 32 units).

Although some researchers propose a minimum 60% SAQ response rate to ensure that the data represent safety culture not opinions (Pronovost & Sexton, 2005), a 40% response rate maximized sample size. Significant relationships may still occur with SAQ response rates less than 60%. A cohort study with over 30 intensive care units had a 47.9% SAQ response rate, and found significant relationships between safety climate and patient length of stay (Huang et al., 2010). In a pre-post study implementing a hospital-based safety program, improvements were seen in all seven SAQ domains with a 43.6% SAQ response rate (Je et al., 2014). Studies outside the U.S. have also used the SAQ with lower response rates. For example, an Australian study established demographic differences in SAQ safety attitudes with a 46% RN response rate and 35% physician response rate (Gallego, Westbrook, Dunn, & Braithwaite, 2012). Translations of the SAQ have occurred with a 52% response rate (Devriendt et al., 2012). Researchers have established limited confounding based on SAQ response rates, except for the stress recognition domain (Watts, Percarpio, West, & Mills, 2010), which was not used in this study.

For procedural areas such as interventional radiology, CT, MRI, endoscopy, surgery, cardiac catheterization lab, arrhythmia services, etc. the patient must have had a procedure in the department during their hospital stay. Two data analysts and an expert billing coder developed a random chart selection process using billing codes for the inpatient dialysis department. This was to ensure that patients physically traveled to the inpatient dialysis area for treatment, rather than receive dialysis in their hospital room.

Inclusion and Exclusion Criteria for Patient Records

Patient outcome data were collected from individual patient records to determine the occurrence of adverse events. Patient records were randomly selected at the unit level for review. A random selection process (SAS code) selected ten medical record numbers for review for each unit/department. The patient's unit was selected based upon the first place the patient went after leaving the emergency department or immediately after surgery.

The GTT inclusion criteria were adapted with specifications unique to the proposal. The inclusion criteria were: 1) patient age greater than 18 years old, 2) length of stay at least 24 hours and admitted to the hospital, January 1- 31, 2013, 3) no psychiatric or addictive disease

admission, or admission to a rehabilitation program, 4) closed and completed medical record, with completed discharge and coding summaries. Patients received care in the hospital between January 1 and January 31, 2013. SAQ administration occurred between February 11, 2013, and March 15, 2013, and by selecting records from the preceding month, the team avoided the potential effect of SAQ survey administration on adverse event prevalence.

Inpatient or observation status did not influence inclusion criteria. During hospital registration, patients were assigned a status upon admission, such as “inpatient” or “observation.” Inpatient status occurred with severe health issues and the need for technical skilled care. Observation status occurred when the patient was not sick enough to warrant hospital admission, but was too sick to receive treatment in the provider’s office. As long as the patient had a length of stay of 24 hours or more, and met the other inclusion criteria above, the patient record could have been potentially included in the sample, regardless of status type.

The university and hospital data collection’s site Institutional Review Boards reviewed and approved the study.

Measurement/Instrumentation

The independent variables were safety climate and teamwork climate, and the dependent variable was frequency of adverse events per unit. Potential confounding variables were unit characteristics, size and type, which were controlled for in the analyses. Numerous survey measures exist to assess safety climate. The Joint Commission expects accredited hospitals and other organizations to evaluate safety climate using valid and reliable tools (The Joint Commission, 2012) but do not recommend one survey over another. The SAQ was selected to quantify safety climate in this study due to its strong validity and reliability, widespread use (Pronovost & Sexton, 2005), association with outcomes, and congruence with the GRM conceptual framework.

The SAQ provides a snapshot of safety culture (Sexton et al., 2006) through six domains. Two domains are termed “safety climate” and “teamwork climate.” These two domains are consistent with GRM organizational and human factors (contributing factors) to patient safety within the GRM conceptual framework.

Safety Climate

Safety climate is the "perceptions of a strong and proactive organizational commitment to safety" (Sexton et al., 2006, p. 3), which practitioners assess using seven items on a 1-5 Likert scale, with anchors 1 = "disagree strongly" and 5 = "agree strongly." For example, the first safety client item is "I would feel safe being treated here as a patient" and respondents indicate their level of agreement with statements. See Table 3.1 for SAQ safety climate items.

Table 3.1

Safety Attitudes Questionnaire Safety Climate Domain and Items

SAQ domain	SAQ items
Safety climate	1. I would feel safe being treated here as a patient. 2. Medical errors are handled appropriately in this work setting. 3. I know the proper channels to direct questions regarding patient safety in this work setting. 4. I receive appropriate feedback about my performance 5. In this work setting, it is difficult to discuss errors. 6. I am encouraged by others in this work setting to report any patient safety concerns I may have. 7. The culture in this work setting makes it easy to learn from the errors of others.

Note. Adapted from "The Safety Attitudes Questionnaire: Psychometric Properties, Benchmarking Data, and Emerging Research" by J.B. Sexton et al., 2006, *BMC Health Services Research*, *6*, p. 7.

Teamwork Climate

Teamwork is measured using the SAQ teamwork climate domain. Teamwork climate is the "perceived quality of collaboration between personnel" (Sexton et al., 2006, p. 3), which practitioners assess with six items. See Table 3.2. For example, the first teamwork climate item is "Nurse input is well received in this setting." Positive teamwork climate is the percentage of caregivers who agreed slightly or agreed strongly with the teamwork climate subscale items (i.e. ability to speak up, support from others, ability to ask questions; Rose, Thomas, Tersigni, Sexton, & Pryor, 2006). Previous research supports moderately good agreement with an observation-based measure of safety-related teamwork and survey scores (Freeth et al., 2012). Safety

climate and teamwork climate were measured at the unit/department level via positive percentage agreement.

Table 3.2

Safety Attitudes Questionnaire Teamwork Climate Domain and Items

SAQ domain	SAQ items
Teamwork climate	1. Nurse input is well received in this work setting. 2. In this work setting, it is difficult to speak up if I perceive a problem with patient care. 3. Disagreements in this work setting are resolved appropriately (i.e., not who is right, but what is best for the patient). 4. I have the support I need from others in this work setting to care for patients. 5. It is easy for personnel here to ask questions when there is something that they do not understand. 6. The physicians and nurses here work together as a well-coordinated team.

Note. Adapted from "The Safety Attitudes Questionnaire: Psychometric Properties, Benchmarking Data, and Emerging Research" by J.B. Sexton et al., 2006, *BMC Health Services Research*, *6*, p. 7.

SAQ domain scores are provided via positive percentage agreement. Scores for each SAQ domain reflect the percentage of respondents, who on average, responded positively to the items in that domain (Pascal Metrics, 2013). The SAQ domain score is the percentage of individuals whose responses across all items in the domain average out to be equal to greater than four (Pascal Metrics, 2013). The average across items for all respondents is examined, and only those individuals with a four or five average are counted. The numerator is the number of respondents with an average four or five across all items of the domain, and the denominator is the number of respondents. For example, the safety climate domain score is the percentage of individuals whose responses across all safety climate items average out to be equal to greater than four ("agree"; Pascal Metrics, 2013). Percentage agreement broadens the definition of agreement by including the adjacent scoring categories on the rating scale, which is advantageous because it relaxes the strict criterion that individuals agree exactly (Stemler, 2004). Previous researchers have proposed a minimum threshold scale of 60% positive percentage

agreement to indicate good teamwork and safety climate over time (Paine et al., 2011), while other organizations strive for 80% positive agreement (Pronovost et al., 2006).

The following subsection discusses SAQ validity and reliability.

SAQ validity.

Content validity. Content validity refers to the extent to which the content of the questionnaire represents the content domain (Soeken, 2010). SAQ items adequately represent safety culture because a review of measurement tools for patient safety culture determined that the SAQ covered 19 dimensions, the most of any survey in their review (Singla, Kitch, Weissman, & Campbell, 2006). Since 2006, other researchers have established content validity using a content validity index (CVI), which quantifies the extent of agreement among reviewers regarding representativeness of the items, stated as the proportion of items given a rating of quite/very relevant by involved raters (Waltz, Strickland, & Lenz, 2010). Researchers tested the content validity of a Dutch translation of the SAQ and calculated the average CVI (total scale) as 0.83, with the six subscale ratings ranging from 0.55 to 0.97 (Devriendt et al., 2012).

Construct validity. Construct validity refers to the extent to which SAQ items are consistent with the theory and concepts as operationally defined (Soeken, 2010). Researchers used exploratory factor analysis to establish SAQ construct validity during survey development, resulting in six attitudinal domains, including three targeted themes, safety climate, teamwork climate, and stress recognition (Sexton, et al., 2006).

SAQ reliability. Internal consistency reliability is concerned with the consistency of performance of one group of individuals across the items on a single measure (Waltz, Strickland, & Lenz, 2010). In reviews of safety surveys, researchers reported the overall SAQ Cronbach's alpha from 0.68 to 0.81 (Colla, Bracken, Kinney, & Weeks, 2005) or 0.65 to 0.83 (Fleming, 2005). A number of researchers state that 0.7 is an acceptable reliability coefficient although there is no evidence to support an ideal number (Kerlinger & Lee, 2000). The original SAQ developers assessed overall scale reliability using Raykov's ρ coefficient because Cronbach α underestimates composite reliability (CR) at the population level (Raykov, 2001). The Raykov ρ coefficient was .90, which indicates strong reliability of the SAQ (Sexton, et al., 2006).

The SAQ teamwork climate validity and reliability are established. The convergent validity of the SAQ teamwork climate subscale was established via correlation with the Nursing Teamwork Survey ($r = .76$, $p < .01$; Kalisch, Lee, & Salas, 2010). The inter-item reliability alpha of the teamwork climate subscale was established at 0.76 (Pronovost et al., 2008).

Confounding Variables: Unit Characteristics

Unit size was operationalized as the number of patient beds per unit or department. Unit types were critical care, intermediate, medical-surgical, OB Gynecology and procedural units.

Adverse Events

Adverse events are injuries caused by medical interventions (Institute of Medicine, 2000) not due the patient's underlying disease process (deVries, Ramrattan, Smorenburg, Gouma, & Boermeester, 2008). Conceptually, adverse events include many different types, such as adverse drug events, healthcare-associated infections, falls, pressure ulcers, etc. (Shojania & Thomas, 2013). Measurement of adverse events may occur through various forms, ranging from voluntary error reporting systems (incident reports), administrative data analysis, and retrospective chart review, and there are advantages and disadvantages to each approach. Evidence suggests that chart reviews using trigger tools identify a larger number of adverse events than other measurement methods (The Health Foundation Inspiring Improvement, 2010). Trigger tools also appear to detect events that may not be detected with other methods, such as incident reports or pharmacy interventions (Doupi, 2012).

Global Trigger Tool. In this study, adverse events were measured using the Institute for Healthcare Improvement Global Trigger Tool for Measuring Adverse Events (Griffin & Resar, 2009). GTT developers do not provide a definition of adverse event, but harm because the GTT counts only adverse events with harm to the patient, not near misses. An error could reach the patient but not cause harm, so this would not be an adverse event according to the GTT. The GTT defines harm as the unintended physical injury resulting from medical care that requires additional monitoring, treatment or hospitalization, or that causes death (Griffin & Resar, 2009). It is helpful to consider unintended harm from the viewpoint of the patient (Griffin & Resar, 2009).

The GTT is comprised of six modules (care, medication, surgical, intensive care, perinatal, emergency department) with 53 triggers. See Table 3.3 for GTT triggers. Triggers are sentinel words or events, such as certain drugs, antidotes, lab values, or stop orders that prompt reviewers to initiate a more detailed chart audit (Resar, Rozich, & Classen, 2003). Chart reviewers used two documents for clarifying GTT trigger definitions and adverse events: the 2012 New Zealand Health Quality & Safety Commission website (Health Quality and Safety Commission New Zealand, 2014) and the Florida Hospital frequently asked questions (Florida Hospital, 2009). See Appendix. Once reviewers identified an adverse event, they assigned a severity rating of harm (harm category) per GTT guidelines. See Table 3.4.

Table 3.3.

Triggers used in Global Trigger Tool

	Cares module triggers
C1	Transfusion or use of blood products
C2	Code/arrest/rapid response team
C3	Acute dialysis
C4	Positive blood culture
C5	X-ray or Doppler studies for emboli or DVT
C6	Decrease of > 25% in hemoglobin or hematocrit
C7	Patient fall
C8	Pressure ulcer
C9	Readmission within 30 days
C10	Restraint use
C11	Healthcare-associated infection
C12	In-hospital stroke
C13	Transfer to a higher level of care
C14	Any procedure complication
C15	Other
	Surgical module triggers
S1	Return to surgery
S2	Change in procedure
S3	Admission to intensive care post-op
S4	Intubation/reintubation/BiPap in Post Anesthesia Care Unit (PACU)
S5	X-ray intra-op or in PACU
S6	Intra-op or post-op death
S7	Mechanical ventilation greater than 24 hours post-op
S8	Intra-op epinephrine, norepinephrine, naloxone, or romazicon
S9	Post-op troponin level greater than 1.5 ng/ml
S10	Injury, repair, or removal of organ
S11	Any operative complication

Continued

Continued Table 3.3

	Medication module triggers
M1	Clostridium difficile positive stool
M2	Partial thromboplastin time greater than 100 seconds
M3	International normalized ratio (INR) > 6
M4	Glucose < 50 mg/dl
M5	Rising BUN or serum creatinine greater than 2 time baseline
M6	Vitamin K administration
M7	Benadryl (Diphenhydramine) use
M8	Romazicon (Flumazenil) use
M9	Naloxone (Narcan) use
M10	Anti-emetic use
M11	Over-sedation/hypotension
M12	Abrupt medication stop
M13	Other
	Intensive care module triggers
I1	Pneumonia onset
12	Readmission to intensive care
13	In-unit procedure
	Perinatal module triggers
P1	Terbutaline use
P2	3rd or 4th-degree lacerations
P3	Platelet count less than 50,000
P4	Estimated blood loss > 500 ml (vaginal) or > 1,000 ml (C-section)
P5	Specialty consult
P6	Oxytocic agents
P7	Instrumented delivery
P8	General anesthesia
	Emergency department module triggers
E1	Readmission to ED within 48 hours
E2	Time in ED greater than 6 hours

Note. Adapted from "IHI Global Trigger Tool for Measuring Adverse Events (second edition) Innovation Series White Paper" by F.A. Griffin and R.K. Resar, 2009, available at www.IHI.org, p. 37. Copyright 2009 by the Institute for Healthcare Improvement.

Table 3.4.

Categories of Harm used in The Global Trigger Tool

Category of harm	Label
E	Temporary harm to the patient and required intervention
F	Temporary harm to the patient and required initial or prolonged hospitalization
G	Permanent patient harm
H	Intervention required to sustain life
I	Patient death

Note. Adapted from "IHI Global Trigger Tool for Measuring Adverse Events (second edition) Innovation Series White Paper" by F.A. Griffin and R.K. Resar, 2009, available at www.IHI.org, p.6. Copyright 2009 by the Institute for Healthcare Improvement.

GTT validity. Researchers have established the GTT validity. The GTT has adequate criterion-related validity because it was compared to the AHRQ Patient Safety Indicators (PSI) and voluntary reporting, and the GTT demonstrated its ability to detect adverse events as well (if not better) than existing criteria (Classen et al., 2011). The GTT had a sensitivity to detect patients with an adverse event of 94.9% and a specificity to detect patients with no events of 100% (Classen et al., 2011). In contrast, the PSI method had a sensitivity of 8.5% and specificity of 98.5%, and a local hospital's voluntary reporting system had a sensitivity of 0% and a specificity of 100% (Classen et al., 2011). The GTT had greater sensitivity to detect greater adverse events due to its broad definition of adverse events not requiring preventability or leading to major patient disability, in contrast to previous studies (Classen et al., 2011). The GTT is a valid measure for detecting adverse events.

GTT reliability. Numerous studies have examined the GTT's interrater reliability. Factors that may affect GTT reliability are reviewer training (Classen, Lloyd, Provost, Griffin, & Resar, 2008), with highly experienced reviewers detecting more harm than newly trained GTT reviewers (Landrigan et al., 2010; von Plessen, Kodal, & Anhoj, 2012). Variation in definition of harm among reviewers and reviewer teams may also affect interrater reliability (Deilkas, 2013). Reviewer role may also influence reliability, with physicians demonstrating greater agreement than non-physician reviewers (Classen et al., 2008). Numerous studies have established substantial agreement between GTT reviewers (Kappa .61 - .80) (Hwang, Chin, & Chang, 2014; Kirkendall et al., 2012). Other GTT teams have reported very good reliability (91.8% reproducibility between nurse reviewers) (Adler et al., 2008) and 78% agreement between analysts for harm identification (Garrett Jr et al., 2013). Table 3.5 summarizes information regarding instrumentation.

Table 3.5.

Summary of Measures

Variable	Definition	Instru-ment	Number of items	Operationalization	Validity	Reliability
Safety climate (SC)	Perceptions of a strong and proactive organizational commitment to safety (Sexton et al., 2006)	Safety Attitudes Ques-tionnaire (Sexton et al., 2006)	Seven items on a 1-5 Likert scale, with anchors 1 = "disagree strongly" and 5 = "agree strongly"	Percentage of respondents in a unit that report "agree slightly" or "agree strongly" for each of the items within a scale are documented as percentage positive (Sexton et al., 2006)	SAQ-Construct validity: Factor analysis (Sexton et al., 2006)	Internal consistency reliability- The Raykov ρ coefficient was .90, which indicates strong reliability of the SAQ (Sexton et al., 2006)
Team-work climate (TC)	The perceived quality of collaboration between personnel (Sexton et al., 2006)	Safety Attitudes Ques-tionnaire (Sexton et al., 2006)	Six items on a 1-5 Likert scale, with anchors 1 = "disagree strongly" and 5 = "agree strongly"	Percentage of respondents in a unit that report "agree slightly" or "agree strongly" for each of the items within a scale are documented as percentage positive (Sexton et al., 2006)	Conver-gent validity: Correlation w/ Nursing Teamwork Survey ($r = .76$, $p < .01$) (Kalisch et al., 2010)	Inter-item reliability alpha of 0.76 for the teamwork climate subscale of the SAQ (Pronovost et al., 2008)
Adverse events (AEs)	Unintended harm to a patient from the viewpoint of the patient (Griffin & Resar, 2009)	Institute for Health-care Improve-ment (IHI) Global Trigger Tool (GTT) (Griffin & Resar, 2009)	53 triggers. Begin with Care and Medication module triggers, and proceed to other module triggers as time permits	No. of AEs/1,000 pt. days: Add total # of AEs, ÷ by total length of stay for all records reviewed, and × by 1,000. No. of AE/100 admissions: Add total # of AEs, ÷ by total records reviewed, and ×by 100. % of admissions with an AE: Add # of admissions with at least one AE, ÷ by total records reviewed, and × by 100.	Adequate criterion-related validity compared to AHRQ PSI and voluntary reporting (Classen et al., 2011)	GTT Kappa .61 - .80 (Hwang et al., 2014; Kirkendall et al., 2012). Very good reliability (91.8% reproducibility between nurse reviewers) (Adler et al., 2008) & 78% agreement b/w analysts for harm identification (Garrett Jr et al., 2013).
Patient harm	"Unintended physical injury resulting from or contributed to by medical care that requires additional monitoring, treatment or hospital-ization, or that results in death" (Griffin & Resar, 2009, p.5)	IHI Global Trigger Tool (GTT) (Griffin & Resar, 2009)	5 Categories: E: Temporary harm requiring intervention F: Temporary harm requiring initial or prolonged hospitalization G: Permanent harm H: Intervention required to sustain life (w/i 1 hour or less in order to prevent death) I: Patient death	The category of harm percentage per number of adverse events.		

Detailed Study Procedures

Independent Variables: Safety and Teamwork Climate

Independent variable data (safety and teamwork climate scores) were previously collected during electronic SAQ administration between February 11, 2013, and March 15, 2013. Staff were informed of the study via unit/department managers and the organization's email system. Staff accessed the SAQ link via email, which contained detailed instructions. Survey participation was voluntary with no compensation. Staff were encouraged to complete the SAQ at work. This secondary data served as the source for safety climate and teamwork climate. SAQ climate scores were aggregated to the unit-level. This respected SAQ respondents' privacy because unit-level data did not contain any individual identifying information. Access to individual SAQ items or responses by job type were not available.

Confounding Variable: Unit Characteristics

Hospital documents provided the number of patient beds for nursing units (unit size) and unit type. For procedural areas, the manager provided the number of pre-procedure and post-procedure bays (beds), or (operating, procedure) rooms.

Dependent Variable: Frequency of Adverse Events

Procedure. Procedures outlined in the second edition IHI Global Trigger Tool for Measuring Adverse Events white paper (Griffin & Resar, 2009) guided adverse event data collection. Some modifications were necessary to assess unit-level outcomes. The primary reviewer (doctoral student) reviewed all 317 medical records and the secondary reviewer reviewed 32 records (10% of intended record sample). The primary reviewer has been an RN for 20 years, has critical care experience, which experts recommend (Adler et al., 2008), and was a former adult health Nurse Practitioner. The secondary reviewer is a pharmacist with a Pharm.D, and an expert in patient safety, and is board certified as a pharmacotherapy specialist. Both reviewers have ties to the hospital. Internal review teams tend to perform reviews more reliably than external review teams (Landrigan et al., 2010; Sharek et al., 2011). Both reviewers

completed the IHI GTT training process and participated in a series of six IHI webinars to learn how to use the GTT with practice chart reviews. The webinars were taught by the original GTT developers, Griffin and Resar, who shared challenging case studies as learning opportunities. The case studies facilitated development of trigger and adverse events definitions prior to beginning independent chart review for the two reviewers. GTT training improves reviewer agreement (Classen et al., 2008). The criteria for competence in GTT procedures was based upon both reviewers completing the webinar series for GTT training. Adverse event data collection began in December 2014 and ended in April 2015.

Hospital leadership assisted with securing experts to retrieve an administrative dataset containing medical record numbers. A medical record number is affiliated with a particular facility and is unique for that admission. A data analyst used SAS code for random record selection at the unit/department level.

The GTT dictates that record reviews are limited to 20 minutes, regardless of length of stay or complexity of care. When a medical record was randomly selected for review, the reviewer scanned the record using the GTT list of triggers. If a trigger was identified, the reviewer examined other parts of the chart to determine if an adverse event occurred. A reviewer could identify a positive trigger without the occurrence of an adverse event.

Presence of an adverse event required documentation that the patient experienced harm from medical care (Griffin & Resar, 2009). The reviewer considered harm from the viewpoint of the patient, i.e. would you be happy if this happened to you. If an adverse event occurred, the reviewer recorded the adverse event, and rated its severity according the established GTT harm severity scale (Griffin & Resar, 2009). The harm scale is lettered E-I, with increasing severity. Category E involves temporary harm to the patient and requires intervention. Category F entails temporary harm to the patient and requires initial or prolonged hospitalization. Category G involves permanent patient harm. Category H includes an intervention required to sustain life (i.e. must be provided in one hour or less in order to prevent death), and the lastly, the most severe category is category I: patient death.

Patients may potentially stay in a number of nursing units during a hospital admission, and thus, determining the unit that corresponded with the adverse event may be challenging. For patients that moved back and forth between multiple units, the reviewer determined the date of transfer based on provider orders. For example, consider a patient being transferred from the ICU to an intermediate (step-down) unit. If the record was selected for the intermediate unit, the reviewer focused on reviewing the record sections for triggers based on care provided in the intermediate unit, not the ICU.

Accuracy of review. To determine interrater reliability (IRR), the team compared agreement between the primary reviewer and secondary reviewer regarding presence of adverse events. The IRR sample comprised a random selection of 32 charts (10% of the intended overall sample size of 320). The Kappa coefficient, or Cohen's kappa, is a non-parametric measure of interrater reliability, that determines the proportion of events consistently classified in the same category on both occasions, beyond that expected by chance (Vogt & Johnson, 2011; Waltz et al., 2010). Applied to the GTT, Kappa tests the amount of agreement between pairs of reviewers to see if the agreement could have occurred by chance (Classen et al., 2008).

The observed percentage agreement regarding presence of an adverse event or not was 93.9%. The Cohen's Kappa score was 0.835. Experts have provided guidelines for the interpretation of Kappa (Landis & Koch, 1977; Viera & Garrett, 2005), and a Kappa score of .835 may be interpreted as almost perfect agreement. Each adverse event was counted as a separate data point because a participant could experience more than one adverse event during a hospitalization. Disagreements were settled through discussion between reviewers. During the independent review, the primary reviewer found two adverse events that the secondary reviewer did not. After reviewing the records together, the secondary reviewer agreed that these two cases were both adverse events.

The primary record reviewer reexamined all adverse events at the end of data collection to verify data accuracy, because the reviewer's skills may have improved over the four-month data collection period. The primary reviewer deleted three adverse events of nausea because administration of the two or more doses of anti-emetic were not administered on the nursing unit

selected for review. The reviewer deleted two adverse events of postoperative anemia, because the postoperative patients received only two units of blood, which is within normal limits of expected blood administration after surgery. Both reviewers discussed 23 complicated cases of adverse events to verify agreement prior to data analysis. The following text provides details regarding the changes made during this data cleaning process.

In the beginning of the study, the review team decided that the operational definition of nausea as an adverse event required two or more doses of an antiemetic to be considered an adverse event. Multiple patients experienced a positive trigger, “anti-emetic use,” but only one dose of an anti-emetic was administered (not two) on a nursing unit selected for record review. In other words, the patient may have received four doses of anti-emetic drugs during the entire hospitalization, but if only one dose was administered on that unit or procedural area, this did not constitute an adverse event for that unit. Three cases of nausea as an adverse event were deleted because the nausea (and corresponding trigger definition of two or more doses of an anti-emetic) were administered at locations that did not correspond with the unit selected for review.

Blood loss and corresponding anemia may occur with surgery. The quantity of blood administration after surgery may depend on surgery type. The review team did not realize in advance that determining the presence of an adverse event (postoperative acute blood loss anemia) would require a judgment regarding a normal or expected quantity of blood products administered after surgery to assist with determining the presence of an adverse event. For example, a patient may have positive triggers “decrease of > 25% in hemoglobin” and “transfusion or use of blood products,” but these triggers alone do not necessarily indicate that an adverse event occurred, as some blood loss and anemia may occur with any surgery. The primary reviewer found an on-line document that clarified expected blood administration after surgery. Experts state that two units of packed red blood cell (PRBC) administration is reasonable to expect after most surgeries, and up to four units may be expected after cardiothoracic surgery (Florida Hospital, 2009; Resar & Griffin, 2014). Thus if a patient is anemic and receives blood after surgery, this treatment may be part of routine post-operative care. Upon learning expected amounts of blood administration after surgery, the reviewer deleted two

adverse events of postoperative anemia, because the patients received only two units of blood, which is within normal limits of expected blood administration after surgery. Although these patients were anemic, the two units of blood would not be considered "additional treatment" beyond what is expected after surgery. To review, the GTT definition of patient harm is "unintended physical injury resulting from or contributed to by medical care that requires additional monitoring, treatment or hospitalization, or that results in death" (Griffin and Resar, 2009, p. 5).

In one case, the reviewers were unable to determine if a patient's skin condition was a Stage 1 or Stage 2 pressure ulcer. The ability to determine pressure ulcer staging based on documentation was important because the team decided that pressure ulcers as adverse events would include the more advanced stages of skin breakdown, classified as Stage 2, 3 or 4 pressure ulcers. In one patient, there was documentation of "sustained deep tissue injury," but no staging. The reviewer consulted with a Master's-prepared wound ostomy nurse with 30+ years nursing experience for expertise. She referred to the international classification system of pressure ulcers, which clarified that this was a Stage 2 pressure ulcer (National Pressure Ulcer Advisory Panel, European Pressure Ulcer Advisory Panel, & Pan Pacific Pressure Injury Alliance, 2014). In addition, the wound ostomy nurse noted that the documentation occurred within 24 hours of hospital admission, so this was considered a "present on admission" pressure ulcer. Because the GTT includes events present on admission, this patient experienced an adverse event.

Data Analysis

IBM SPSS Statistics version 22 and Microsoft Office Excel 2013 served as the software programs for data analyses.

Aim 1

Aim one was to describe self-reported safety climate and teamwork climate among an interprofessional group of providers working on 32 hospital inpatient units, measured using the SAQ. Aim 1 required the following analyses. Descriptive statistics (means and standard deviations) determined the distribution and variability of safety climate and teamwork climate.

Cronbach's alpha assessed the internal consistency reliability of safety climate and teamwork climate subscales. A one-way analysis of variance (ANOVA) determined if safety climate and teamwork climate scores varied by unit type and size, with safety climate/teamwork climate as the dependent measure and unit type or size as the independent measure.

Aim 2

Aim 2was to explore the nature of adverse events identified using a modified IHI GTT chart review methodology. Aim 2 required the following analyses. Descriptive statistics determined the distribution (percentage) of identified adverse events by level of harm. A one-way analysis of variance (ANOVA) determined if frequency of adverse events varied by unit type, with frequency of adverse events as the dependent measure and unit type as the independent measure. To determine the positive predictive value of each trigger, the number of adverse events identified with the trigger was divided by the number of triggers found in the patient charts (Carnevali et al., 2013).

Aim 3

Aim 3 was to examine to what extent unit-level safety climate and teamwork climate (and their potential interaction) predict frequency of adverse events as detected via the modified IHI GTT chart review format, controlling for unit characteristics (size and type). We were prepared for multicollinearity as a potential problem because previous research indicated that safety climate and teamwork climate were highly correlated with one another, $r = 0.73$ (Taylor et al., 2012), and $r = 0.8$ (Speroff et al., 2010). When two variables have a correlation of .70 or more, multicollinearity can cause both logical and statistical problems (Tabachnick & Fidell, 2013). Logical problems involve the idea that the two variables may be redundant if they are highly correlated, and statistical problems include inflated error terms with multicollinearity (Tabachnick & Fidell, 2013).

Correlations of safety climate and teamwork climate were estimated using two-tailed Pearson's product moment coefficient. The following adjustments were made to the original analytic plan. Because safety climate and teamwork climate were highly correlated with one another, and teamwork exhibited skewness, teamwork climate was transformed to a dichotomous

variable, comprised of two groups, consisting of low and high teamwork units. Teamwork was split into two groups based upon a dividing point of .60 (or 60% agreement) indicating that over half of respondents on average agreed with items in that domain (Pascal Metrics, 2013). Unit size was not correlated with any of the variables, so unit type was controlled for as a unit characteristic confounding variable (rather than unit size).

Aim 3 required the following analyses. Contingency table analyses were conducted to determine whether there was an association between dichotomized teamwork climate and unit type. A one-way ANOVA examined differences in unit type on frequency of adverse events. Independent *t*-tests determined: 1) the extent to which medical surgical versus other unit type means differed for low and high teamwork units and safety climate, and 2) the extent to which low and high teamwork units differed for frequency of adverse events. A two-factor ANOVA tested the effects of medical surgical units (versus other unit types) and low and high teamwork units on adverse events per 1,000 days.

We estimated a hierarchical regression model with adverse events per 1,000 patient days as the outcome, and dichotomized teamwork climate, safety, and a term interacting teamwork climate and safety climate, controlling for unit type. Hierarchical regression analysis was selected because it allows the researcher to determine the effect of independent variables after the effect of other variables has been controlled (Polit, 2010), by reviewing the R square change as each independent variable joins the others (Cohen, Cohen, West, & Aiken, 2003). Lastly, because so many of the units did not experience any adverse events, an alternative hypothesis was developed to examine the likelihood that a unit would experience an adverse event. A logistic regression analysis was conducted to predict the likelihood that a unit would experience an adverse event, with three predictor variables: unit type, dichotomized teamwork climate, and safety climate.

Limitations

Internal Validity

Researchers assert that threats to internal validity are important to consider in survey (non-experimental) research (Yiannakis, 1997). This study used a non-experimental, descriptive,

cross-sectional design, with a retrospective chart review methodology. The instrumentation threat to internal validity may be a concern due to the method used to detect adverse events. Some experts believe that the GTT is unreliable. A Swedish study with five GTT teams demonstrated large variation between teams, with a combined unweighted Kappa of 0.20 (Schildmeijer, Nilsson, Arestedt, & Perk, 2012), which indicates only slight reliability (Waltz et al., 2010). This may have occurred due to differing harm definitions than the GTT proposes (Deilkas, 2013). Other researchers found moderate inter-rater agreement (κ = .41- .60) on the GTT (Mattsson, Knudsen, Lauritsen, Brixen, & Herrstedt, 2013; Naessens et al., 2010). GTT results may differ depending on team training, review processes, and documentation (von Plessen, Kodal, & Anhoj, 2012). In this study, chart reviewers undertook specialized training to protect against investigator bias for adverse event data collection. The two record reviewers participated in a series of IHI-sponsored webinars to learn how to use the GTT from February- June 2014. The reviewers completed IHI assignments to learn the GTT review process.

Data regarding safety and teamwork climates were limited to participants' self-reported perceptions aggregated to the unit level. These perceptions were based on voluntary survey responses to the SAQ, and do not necessarily reflect non-respondents' views. We included units or departments with a least a 40% SAQ response rate in our sample, but some researchers propose a minimum 60% SAQ response rate to ensure that the data represent safety culture not opinions (Pronovost & Sexton, 2005). Previous research demonstrates limited confounding based on SAQ response rates (Watts et al., 2010).

External Validity

This study has limited generalizability because medical record reviews were limited to a single hospital site. However, we employed a random selection process from this site to ensure representativeness, in order to facilitate generalizability of the findings beyond the sample studied.

Chapter 4: Results

The purpose of the study was to explore relationships between organizational and human factors with adverse events detected using a modified trigger-tool methodology. The specific aims were:

Aim 1. Describe self-reported safety climate and teamwork climate among an interprofessional group of providers working on 32 hospital inpatient units, measured using the Safety Attitudes Questionnaire (SAQ).

Aim 2: Explore the nature of adverse events identified using a modified IHI GTT chart review methodology.

Aim 3. Examine to what extent unit-level safety climate and teamwork climate (and their potential interaction) predict frequency of adverse events as detected via the modified IHI GTT chart review format, controlling for unit characteristics (size and type).

The variables of study were (nursing/department) unit characteristics, safety climate, teamwork climate, and adverse events. Unit characteristics, such as size and type, were selected as control variables for statistical analyses. The predictors (independent variables) were safety climate and teamwork climate. The dependent variable was frequency of adverse events. The following section presents descriptive statistics regarding unit characteristics, teamwork climate as a continuous variable, teamwork climate as a dichotomized variable, and safety climate.

Descriptive Statistics

Control Variables: Unit Characteristics

The study setting was a 750+-bed Midwestern U.S. regional acute care hospital, and 32 of the 79 hospital inpatient units/departments with greater than 40% SAQ response rate participated in the study. The emergency department was excluded due to less than 40% response rate on the SAQ. The mean SAQ response rate was 62.75% (*SD* 14.37). This

sample's response rate was consistent with benchmarking data collected from over 1500 employees from 11 inpatient units in U.S. hospitals, with a reported response rate of 65.7% (Sexton et al., 2006). However, as explained in Chapter 3, 15 of the 32 units experienced a response rate of 40% - 60% (range 19.78).

The average unit size (number of beds per unit) was 28 beds (*SD* = 17.57). See Table 4.1. There were three intermediate care units, three OB/Gynecology units, five critical care units, 10 medical-surgical units, and 11 procedural areas in the sample of 32 units. See Table 4.2. Procedural areas included multiple cardiac departments (such as catheterization lab, testing, post procedure recovery, imaging), radiology, CT (computed tomography), endoscopy, inpatient dialysis, and surgery.

Table 4.1

Descriptives: Unit Size, Teamwork Climate, and Safety Climate

		Unit size (# of beds)	Teamwork climate	Safety climate
N	Valid	32.0	32.00	32.00
	Missing	0.0	0.0	0.0
Mean		28.00	.70	.76
Median		26.50	.75	.79
Std. Deviation		17.57	.19	.16
Variance		308.65	.04	.02
Skewness		1.23	-.95	-.70
Std. Error of Skewness		.41	.41	.41
Kurtosis		2.58	.68	.86
Std. Error of Kurtosis		.81	.81	.81
Range		81.00	.83	.70
Minimum		5.00	.17	.30
Maximum		86.00	1.00	1.00

Table 4.2

Descriptive Statistics: Unit Type

		Frequency	%	Valid %	Cumulative %
Valid	Critical care	5	15.6	15.6	15.6
	Intermediate care	3	9.4	9.4	25.0
	Medical surgical	10	31.3	31.3	56.3
	OB gynecology	3	9.4	9.4	65.6
	Procedural area	11	34.4	34.4	100.0
	Total	32	100.0	100.0	

Predictors: Teamwork Climate

Reliability. Pascal Metrics estimated the reliability of teamwork climate via Cronbach's alpha with raw data from 1206 individual respondents using six teamwork climate items. The Cronbach's alpha for teamwork climate was .80. A number of researchers state that 0.7 is an acceptable reliability coefficient although there is no evidence to support an ideal number (Kerlinger & Lee, 2000).

Teamwork climate as a continuous variable. Scores for each SAQ domain reflect the percentage of respondents who, on average, responded positively to the items in that domain (Pascal Metrics, 2013). For the teamwork climate domain, percentage agreement scores ranged from 17% to 100%. The mean was 70.07% (*SD* = .19) and the median was 75.2%. See Table 4.1. To determine distribution of scores, the skewness ratio was calculated by dividing the skewness statistic by the standard error (-.95 ÷ .41), which was -2.32. This number -2.32 falls outside the range of -2 to +2, so teamwork climate was negatively skewed, indicating that the data does not follow a normal distribution. Respondents, on average, responded more negatively to teamwork climate than safety climate. See Table 4.1 and Figure 4.1.

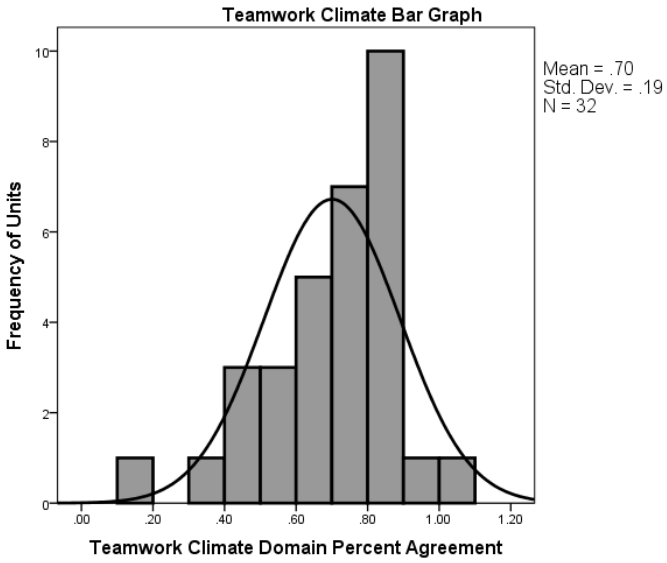

Figure 4.1. Bar graph depicting the distribution of teamwork climate scores.

Dichotomized teamwork climate. Teamwork climate was dichotomized into two groups due to its skewness. The two groups represented nursing units with low versus high teamwork climate scores, based upon a dividing point of .60 (or 60% agreement). Experts recommend that each SAQ domain should have a score of at least 60% to be considered "good" because this indicates that over half of respondents on average agreed with items in that domain (Pascal Metrics, 2013). Eight units or 25% of the sample reported less than good or low teamwork climate scores. Nearly all subsequent analyses with teamwork climate were performed using the dichotomized teamwork climate variable. See Table 4.3 and Figure 4.2.

Table 4.3.

Teamwork Climate Categorized into Two Groups

		Frequency	%	Valid %	Cumulative %
Valid	Low teamwork, < .6001	8	25.0	25.0	25.0
	High teamwork, > .6001	24	75.0	75.0	100.0
	Total	32	100.0	100.0	

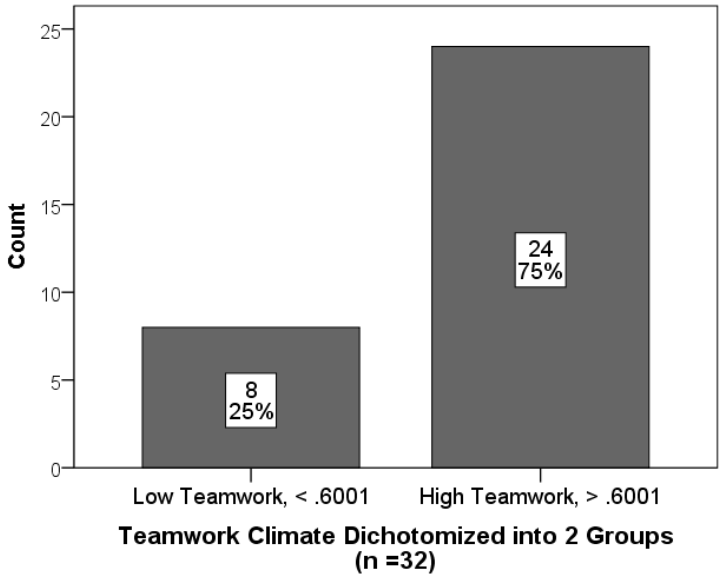

Figure 4.2. Bar chart of dichotomized teamwork climate scores. .

Predictors: Safety Climate

Reliability. Pascal Metrics estimated the reliability of safety climate via Cronbach's alpha with raw data from 1206 individual respondents using seven safety climate items. The Cronbach's alpha for safety climate was .82. A number of researchers state that 0.7 is an acceptable reliability coefficient although there is no evidence to support an ideal number (Kerlinger & Lee, 2000).

For the safety climate domain, percentage agreement scores ranged from 30% to 100%. The mean was 75.61% (*SD* = .16) and the median was 78.55%. See Table 4.1. To determine the distribution of scores, the skewness ratio was calculated by dividing the skewness statistic by the standard error (-.70 ÷ .41), which was -1.70. This number -1.70 falls within the range of -2 to +2, so safety climate was not skewed, indicating that safety climate data followed a normal distribution. See Table 4.1 and Figure 4.3.

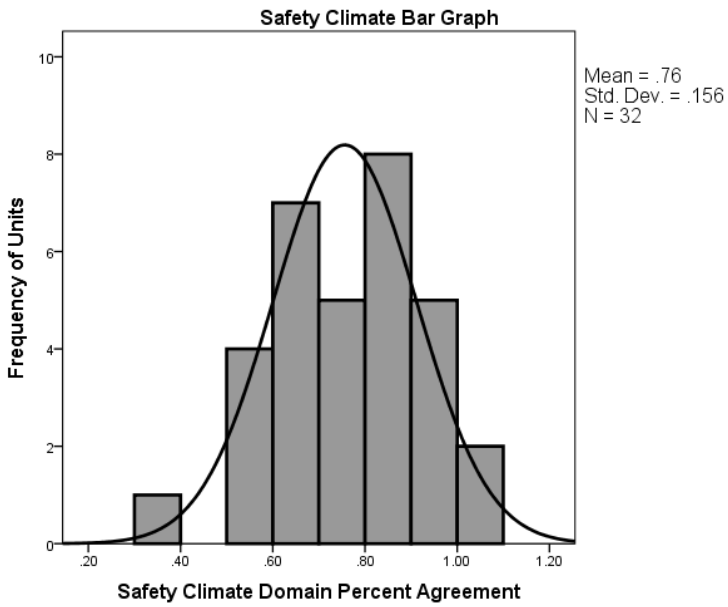

Figure 4.3. Bar graph depicting the distribution of safety climate scores.

Aim 1

This section presents descriptive statistics (related to aim 1) regarding teamwork and safety climate by unit type, and statistics regarding unit size by unit type.

Frequency of Teamwork Climate and Safety Climate by Unit Type

Teamwork climate by unit type. Unit types with the strongest teamwork climate were critical care (*M* = .77, *SD* = .17), intermediate care (*M* = .77, *SD* = .10) and medical surgical (*M* = .74, *SD* = .13). Procedural units had widest range (.78) with one unit experiencing very low teamwork (.17) and another unit experiencing high teamwork (.94). OB gynecology units experienced the lowest mean teamwork climate of all unit types (*M* = .62, *SD* = .18). The cut point for "good" teamwork climate was .60, because this indicates that over half of respondents on average agreed with items in that domain (Pascal Metrics, 2013). The 61.8% agreement for OB/gynecology units was barely greater than the .60 threshold, although it does still constitute "good" teamwork. See Table 4.4 and Figure 4.4.

Table 4.4.

Descriptive Statistics of Teamwork Climate and Safety Climate by Unit Type

		N	Mean	Std. deviation	Std. error	Minimum	Maximum
Teamwork climate	Critical care	5	.77	.17	.08	.57	1.00
	Intermediate care	3	.77	.10	.06	.66	.85
	Medical surgical	10	.74	.13	.04	.50	.89
	OB gynecology	3	.62	.18	.10	.42	.77
	Procedural area	11	.63	.25	.08	.17	.94
	Total	32	.70	.19	.03	.17	1.00
Safety climate	Critical care	5	.77	.15	.07	.63	1.00
	Intermediate care	3	.78	.06	.03	.71	.81
	Medical surgical	10	.80	.11	.03	.66	.94
	OB gynecology	3	.67	.11	.06	.55	.77
	Procedural area	11	.73	.22	.07	.30	1.00
	Total	32	.76	.16	.03	.30	1.00

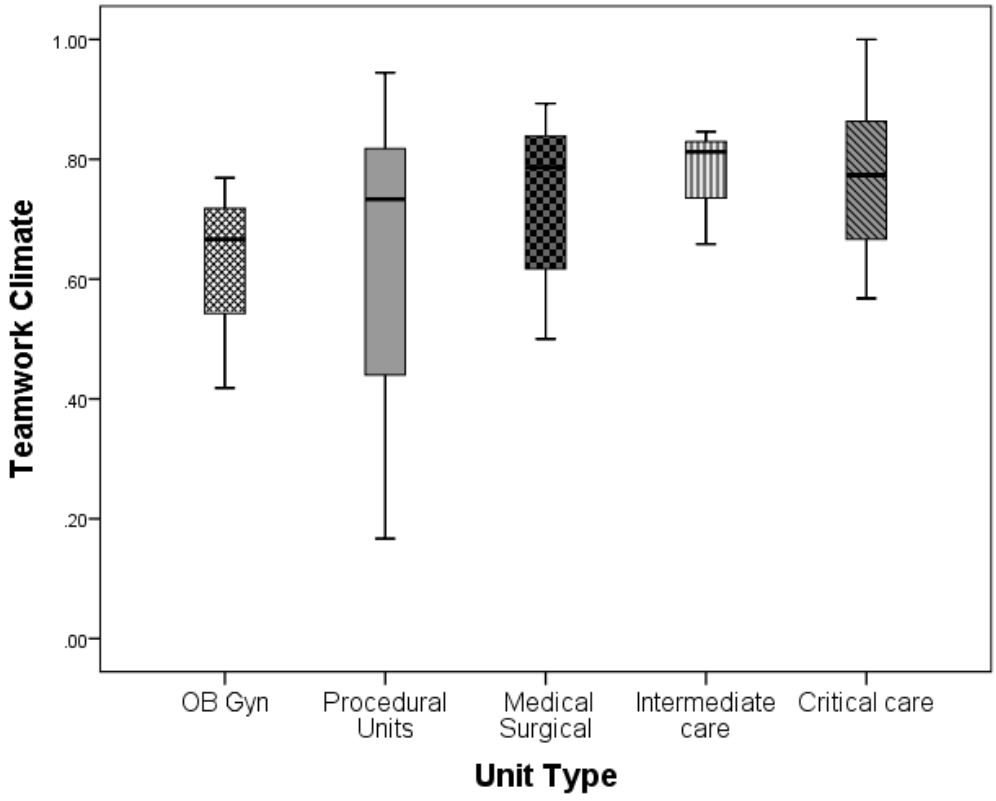

Figure 4.4. Boxplot of teamwork climate by unit type.

Safety climate by unit type. For safety climate, unit types with the strongest safety climate were medical surgical units (*M* = .80, *SD* = .11), followed by intermediate care (*M* = .78, *SD* = .06) and critical care (*M* = .77, SD = .15). OB gynecology units experienced the lowest mean safety climate of all unit types (*M* = .67, *SD* = .11), although this was still indicative of a "good" safety climate. Table 4.4 and Figure 4.5.

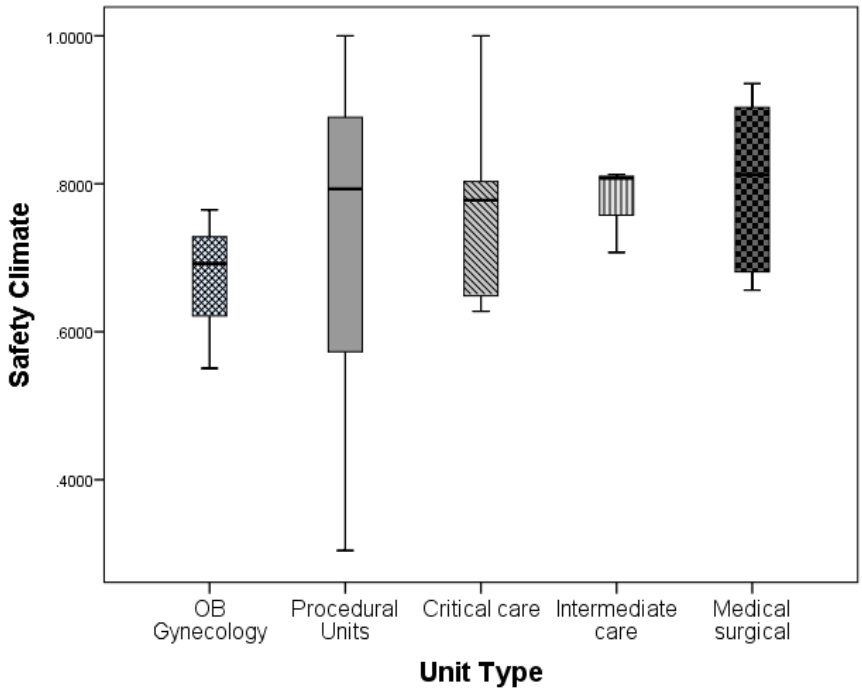

Figure 4.5. Boxplot of safety climate by unit type.

Unit size by unit type. The largest unit types were OB/gynecology with an average of 45 beds per unit (*SD* = 36.12), followed by intermediate care and medical surgical units (*M* = 40, *SD* = 17.50; *M* = 37, *SD* = 10.11). Procedural units were the smallest of all unit types, averaging about 15 beds per unit (*SD* = 8.72). See Table 4.5.

Table 4.5

Descriptive Statistics of Unit Size by Unit Type

	N	Mean	Std. deviation	Std. error	Minimum	Maximum
Critical care	5	21	10.82	4.84	9	32
Intermediate care	3	40	17.50	10.11	22	57
Medical surgical	10	37	10.11	3.20	24	57
OB gynecology	3	45	36.12	20.85	17	86
Procedural area	11	15	8.72	2.63	5	28
Total	32	28	17.57	3.11	5	86

The five unit types differed significantly in unit size (F = 5.28, df = 4, 27, p = .003). See Table 4.6. The Bonferroni post-hoc test determined that the average medical surgical unit size was statistically significantly larger than the average procedural unit size (21.91, p = .014). The average OB/ gynecology unit size was also statistically significantly larger than the average procedural unit size (30.24, p = .028). See Table 4.7.

Table 4.6

One-Factor Fixed ANOVA Table of Unit Size by Unit Type

	Sum of squares	df	Mean square	F	Sig.
Between groups	4197.76	4	1049.44	5.28	.003
Within groups	5370.24	27	198.90		
Total	9568.00	31			

Table 4.7

Bonferroni Multiple Comparison Procedure of Unit Size by Unit Type

(I) Unit Type	(J) Unit type	Mean difference (I-J)	Std. error	Sig.
Critical care	Intermediate care	-18.67	10.30	.811
	Medical surgical	-16.00	7.72	.480
	OB gynecology	-24.33	10.30	.256
	Procedural area	5.91	7.61	1.000
Intermediate care	Critical care	18.67	10.30	.811
	Medical surgical	2.67	9.28	1.000
	OB gynecology	-5.67	11.52	1.000
	Procedural area	24.58	9.19	.125
Medical surgical	Critical care	16.00	7.73	.480
	Intermediate care	-2.67	9.28	1.000
	OB gynecology	-8.33	9.28	1.000
	Procedural area	21.91*	6.16	.014
OB gynecology	Critical care	24.33	10.30	.256
	Intermediate care	5.67	11.52	1.000
	Medical surgical	8.33	9.28	1.000
	Procedural area	30.24*	9.19	.028
Procedural area	Critical care	-5.91	7.61	1.000
	Intermediate care	-24.58	9.19	.125
	Medical surgical	-21.91*	6.16	.014
	OB gynecology	-30.24*	9.19	.028

Note. * indicates the mean difference is significant at the 0.05 level.

Unit size by dichotomized teamwork climate. Units with low teamwork (less than 60% agreement) had an average size of 31.25 beds (*SD* = 27.67). Units with high teamwork had an average size of 26.92 beds per unit (*SD* = 13.35). Low and high teamwork units did not differ significantly in unit size (*F* = .36, *df* = 1,30, *p* = .554). See Tables 4.8 and 4.9.

Table 4.8

Descriptive Statistics of Unit Size by Dichotomized Teamwork Climate

Dichotomized teamwork climate	N	Mean	Std. deviation	Std. error	Minimum	Maximum
Low teamwork, < .6001	8	31.25	27.67	9.78	5	86
High teamwork, > .6001	24	26.92	13.35	2.73	5	57
Total	32	28.00	17.57	3.11	5	86

Table 4.9

One-Factor Fixed ANOVA Table of Unit Size by Dichotomized Teamwork Climate

	Sum of squares	df	Mean square	F	Sig.
Between groups	112.67	1	112.67	.36	.554
Within groups	9455.33	30	315.18		
Total	9568.00	31			

Unit size by safety climate. Table 4.10 contains the safety climate scores by unit size. There were two units with eight beds that both had 100% percentage agreement on safety climate.

Table 4.10

Descriptive Statistics of Unit Size by Safety Climate

Safety climate	N	Unit size mean (no. of beds)	Std. deviation	Std. error	Minimum	Maximum
.304	1	6	.	.	6	6
.500	1	5	.		5	5
.551	1	86	.	.	86	86
.558	1	28	.	.	28	28
.588	1	16	.	.	16	16
.628	1	32	.	.	32	32
.649	1	20	.	.	20	20
.656	1	57	.	.	57	57
.660	1	32	.	.	32	32
.681	1	38	.	.	38	38
.692	1	17	.	.	17	17
.696	1	28	.	.	28	28
.707	1	22	.	.	22	22
.765	1	33	.	.	33	33
.773	1	50	.	.	50	50
.778	1	12	.	.	12	12
.793	1	16	.	.	16	16
.803	1	32	.	.	32	32
.805	1	34	.	.	34	34
.808	1	57	.	.	57	57
.813	1	40	.	.	40	40
.820	1	38	.	.	38	38
.824	1	18	.	.	18	18
.841	1	36	.	.	36	36
.846	1	16	.	.	16	16
.903	1	25	.	.	25	25

Continued

Continued Table 4.10

.929	1	24	.	.	24	24
.933	1	22	.	.	22	22
.936	1	36	.	.	36	36
.957	1	5	.	.	5	5
1.000	2	8	2.12	1.50	6	9
Total	32	28	17.57	3.11	5	86

Safety climate did not differ significantly by unit size (F = 70.84, df = 30,1, p = .094). See Table 4.11.

Table 4.11

One-Factor Fixed ANOVA Table of Unit Size by Safety Climate

	Sum of squares	df	Mean square	F	Sig.
Between groups	9563.50	30	318.78	70.84	.094
Within groups	4.50	1	4.50		
Total	9568.00	31			

Aim 2

The following section presents descriptive statistics related to aim two. To reiterate, aim two was to explore the nature of adverse events identified using a modified IHI GTT chart review methodology.

Dependent Variable: Frequency of Adverse Events

Patient outcome data were collected from individual patient records to determine the occurrence of adverse events. A total of 317 records were reviewed (approximately 10 records for each of the 32 units). (For one unit, only seven records were available for review.) The mean patient age was 61.4 (*SD* 19.04*)* and 182 (57.4%) of the patients were women. The mean length of hospital stay was 4.1 days (*SD* 4.62). The team detected the following harm data. Means of approximately 69 adverse events occurred per 1,000 patient days (*SD* 75.82), and 21.83 adverse

events occurred per 100 admissions (*SD* 17.84), respectively. Approximately 20% of admissions experienced an adverse event. See Table 4.12.

Table 4.12.

Descriptive Statistics of Dependent Variable: Frequency of Adverse Events

		AEs per 1,000 patient days	AEs per 100 admissions	% of admissions with an AE
N	Valid	32	32	32
	Missing	0	0	0
Mean		69.20	21.83	19.64
Median		40.83	20.00	20.00
Std. deviation		75.82	17.84	15.52
Variance		5748.68	318.27	240.95
Skewness		1.46	.43	.23
Std. error of skewness		.41	.41	.41
Kurtosis		1.72	-.51	-.93
Std. error of kurtosis		.81	.81	.81
Range		285.71	60.00	50.00
Minimum		.00	.00	.00
Maximum		285.71	60.00	50.00

Descriptive statistics determined the distribution (percentage) of identified adverse events by level of harm. See Figure 4.6. By reviewing the medical records of 317 patients, 199 triggers and 69 adverse events were found. Fifty-three adverse events were associated with a harm category of E (76.81%), which represented temporary harm to the patient that required intervention. Category E events required additional treatment, such as an additional dose of anti-emetic medication. Sixteen events were associated with a harm category of F (23.19%), which represented temporary harm to the patient that required initial or prolonged hospitalization. Some category F events were related to procedure complications, medication side effects, infection, or post-operative ileus. An example is crepitus or subcutaneous air after chest tube insertion.

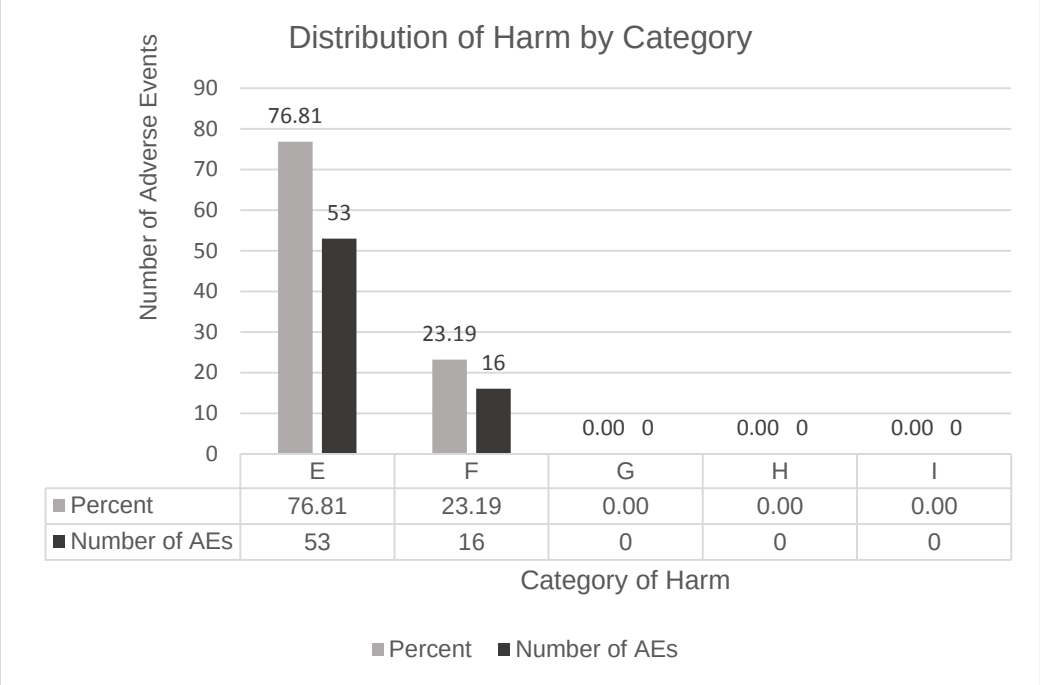

Figure 4.6. Bar chart depicting distribution of harm by category for all charts reviewed. Sixty-nine adverse events were found through review of 317 records for 32 units. Category E harm represents temporary harm to the patient that required intervention. Category F harm represents temporary harm to the patient and required initial or prolonged hospitalization.

Frequency of Adverse Events by Unit Type

Unit types with the greatest frequency of adverse events per 1,000 patient days were medical surgical, (*M* = 125.67, *SD* = 81.86), procedural (M = 58.93, *SD* = 79.63) and critical care (*M* = 38.57, SD = 9.97). Medical surgical units had widest range with some units experiencing zero events and others experiencing 285.7 adverse events per 1,000 patient days. Medical surgical units had the greatest number of adverse events per 100 admissions (*M* = 36.0, *SD* = 18.38), followed by critical care (*M* = 26.0, *SD* = 8.94) and procedural units (*M* = 14.42, SD = 15.59). Three unit types, medical surgical, OB gynecology, and procedural units experienced a minimum frequency (zero) of adverse events. OB gynecology units experienced the fewest adverse events of all unit types. See Figure 4.7 and Table 4.13.

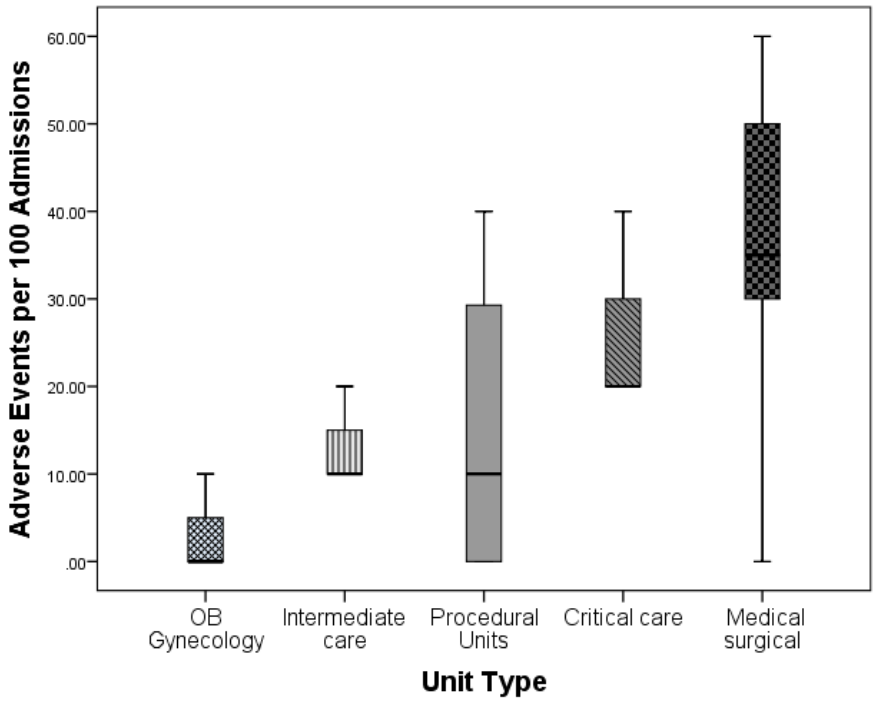

Figure 4.7. Boxplot of adverse events per 100 admissions by unit type.

Table 4.13

Descriptive Statistics of Adverse Events by Unit Type

		N	Mean	Std. dev-iation	Std. error	Min	Max
Adverse events per 1,000 patient days	Critical care	5	38.57	9.97	4.47	28.57	49.38
	Intermediate care	3	24.97	8.51	4.91	15.63	32.26
	Medical surgical	10	125.67	81.86	25.89	.00	285.71
	OB gynecology	3	13.89	24.06	13.89	.00	41.67
	Procedural area	11	58.93	79.63	24.01	.00	250.00
	Total	32	69.20	75.82	13.40	.00	285.71
Adverse events per 100 ad-missions	Critical care	5	26.00	8.94	4.00	20.00	40.00
	Intermediate care	3	13.33	5.77	3.33	10.00	20.00
	Medical surgical	10	36.00	18.38	5.81	.00	60.00
	OB gynecology	3	3.33	5.77	3.33	.00	10.00
	Procedural area	11	14.42	15.59	4.70	.00	40.00
	Total	32	21.83	17.84	3.15	.00	60.00
% of admissions with an adverse event	Critical care	5	20.00	7.07	3.16	10.00	30.00
	Intermediate care	3	13.33	5.77	3.33	10.00	20.00
	Medical surgical	10	32.00	14.76	4.67	.00	50.00
	OB gynecology	3	3.33	5.77	3.33	.00	10.00
	Procedural unit	11	14.42	15.59	4.70	.00	40.00
	Total	32	19.64	15.52	2.74	.00	50.00

To determine the positive predictive value of each trigger, the number of adverse events identified with the trigger was divided by the number of triggers found in the patient charts (Carnevali et al., 2013). See Table 4.14. One adverse event was found without a trigger. Frequently identified triggers were anti-emetic administration (*n* = 44), diphenhydramine (Benadryl) administration (*n* = 23), transfusion or use of blood products (*n* = 21), decrease of greater than 25% in hemoglobin or hematocrit (*n* = 13), restraint use (*n* = 10), and any procedure complication (*n* = 10*).* Six triggers had positive predictive values of greater than 75%: healthcare-associated infections (100%), injury, repair, or removal of organ (100%), over-sedation/ hypotension (100%), any procedure complication (90%), anti-emetic use (84%), and any operative complication (75%).

Numerous triggers were not identified: in-hospital stroke, clostridium difficile positive stool, romazicon (Flumazenil) administration, naloxone (Narcan) administration, return to surgery, change in procedure, intubation/re-intubation or BiPap use in the Post Anesthesia Care Unit (PACU), x-ray intra-operatively or in PACU, intra-operative or post-operative death, mechanical ventilation longer than 24 hours post-operatively, intraoperative administration of epinephrine, norepinephrine, Naloxone, or Romazicon, post-operative troponin level greater than 1.5 ng/mL, readmission to intensive care unit (ICU),intubation or reintubation in ICU, Terbutaline use, 3rd- or 4th- degree (perineal) lacerations, platelet count less than 50,000 (perinatal), specialty consult (perinatal), and general anesthesia (perinatal). The emergency department (ED) triggers were not applicable in this study because ED records were not reviewed due to low SAQ response rate.

Table 4.14

Prevalence of Triggers, Adverse Events and Trigger Positive Predictive Value (PPV) Using Institute for Healthcare Improvement Global Trigger Tool (GTT)

Triggers	No. of triggers found in the records	All AEs	
		n	PPV
C1: Transfusion or use of blood products	21	0	N/A
C2: Code/ arrest/ rapid response team activation	7	1	0.142
C3: Acute dialysis	1	0	N/A
C4: Positive blood culture	1	0	N/A
C5: X-ray or Doppler studies for emboli or deep vein thrombosis	8	1	0.125
C6: Decrease of greater than 25% in hemoglobin or hematocrit	13	0	N/A
C7: Patient fall	1	0	N/A
C8: Pressure ulcers	4	2	0.5
C9: Readmission within 30 days	8	2	0.25
C10: Restraint use	10	1	0.1
C11: Healthcare-associated infections	5	5	1.0
C12: In-hospital stroke	0	0	N/A
C13: Transfer to higher level of care	4	0	N/A
C14: Any procedure complication	10	9	0.9
C15: Other	1	1	1.0
M1: *Clostridium Difficile* positive stool	0	0	N/A
M2: Partial Thromboplastin Time > 100 seconds	7	0	N/A
M3: International Normalized Ratio (INR) > 6	1	0	N/A
M4: Glucose less than 50 mg/dL	3	0	N/A
M5: Rising BUN or serum creatinine > 2 times baseline	5	0	N/A
M6: Vitamin K (Phytonadione) administration	2	0	N/A
M7: Benadryl (Diphenhydramine)/ anti-histamines)/ IV corticosteroid use)	23	3	0.130
M8: Romazicon (Flumazenil) use	0	0	N/A
M9: Naloxone (Narcan) use	0	0	N/A
M10: Anti-emetic use	44	37	0.841
M11: Over-sedation/ hypotension	1	1	1.0
M12: Abrupt medication stop	2	1	0.5
M13: Other	2	1	0.5
S1: Return to surgery	0	0	N/A
S2: Change in procedure	0	0	N/A
S3: Admission to intensive care post-operatively	1	0	N/A
S4: Intubation/ re-intubation/ BiPap in Post Anesthesia Care Unit (PACU)	0	0	N/A
S5: X-ray intra-op or in Post Anesthesia Care Unit	0	0	N/A
S6: Intra-op or post-op death	0	0	N/A
S7: Mechanical ventilation longer than 24 hours post-op	0	0	N/A
S8: Intra-op epinephrine, nor-epinephrine, Naloxone, or Romazicon	0	0	N/A
S9: Post-op troponin level greater than 1.5 ng/mL	0	0	N/A
S10: Injury, repair, or removal of organ	1	1	1.0
S11: Any operative complication	4	3	0.75

Continued

Continued Table 4.14

I1: Pneumonia onset	1	0	N/A
I2: Readmission to intensive care unit (ICU)	0	0	N/A
I3: In-unit procedure	1	0	N/A
I4: Intubation/ reintubation	0	0	N/A
P1: Terbutaline use	0	0	N/A
P2: 3rd- or 4th- degree lacerations	0	0	N/A
P3: Platelet count less than 50,000	0	0	N/A
P4: Estimated blood loss > 500 ml (vaginal) or > 1,000 ml cesarean section	2	0	N/A
P5: Specialty consult	0	0	N/A
P6: Oxytocic agents (such as oxytocin, methylergonovine, and 15-methyl-prosto-glandin in the post-partum period)	1	0	N/A
P7: Instrumented delivery	4	0	N/A
P8: General anesthesia	0	0	N/A
E1: Readmission to ED within 48 hours	N/A	N/A	N/A
E2: Time in ED greater than 6 hours	N/A	N/A	N/A

Note. Emergency department (ED) records were not reviewed due to low SAQ response rate. Triggers are arranged according to GTT categories. C comprises the cares module triggers, M comprises the medication module triggers, S comprises the surgical module triggers, I comprises the intensive care module triggers, P comprises the perinatal module triggers, and E comprises the emergency department module triggers. N/A is not applicable.

Aim 3

This section presents results related to aim three. Aim 3 was to examine to what extent unit-level safety climate and teamwork climate (and their potential interaction) predicted adverse events (per nursing unit) as detected via the modified IHI GTT chart review format, controlling for unit characteristics. The section begins with a presentation of bivariate tests because ordinary least squares regression requires a testing of the linear assumption since it is based on a measure of linear association, Pearson's *r* (Hayes, 2013). Correlations among the control variable (unit size), predictors (safety climate and teamwork climate), and frequency of adverse events follow. Because unit size was not correlated with any of the variables, unit type was explored as a unit characteristic control variable. A contingency table analysis was conducted to determine whether there was an association between dichotomized teamwork climate and unit type. The bivariate subsection concludes with a one-way analysis of variance exploring differences in frequency of adverse events based upon unit type.

Bivariate Tests

Correlations.

Unit size with safety climate, teamwork climate, and frequency of adverse events. Pearson product-moment correlation coefficients were computed to assess relationships between the control variable (unit size), predictors (teamwork climate and safety climate), and dependent variables (number of adverse events per 1,000 patient days, frequency of adverse events per 100 admissions, and percentage of admission with an adverse event). There was no statistically significant correlation between unit size and teamwork climate (r = -.07, n = 32, p = .724) or safety climate (r = -.12, n = 32, p = .508). See Table 4.15. In addition, the correlations between unit size and frequency of adverse events were not significant. The correlation between unit size and number of adverse events per 1,000 patient days was .052, which was positive but not statistically different from zero (r = .05, n = 32, p = .776). Similarly, unit size was not correlated with adverse events per 100 admissions (r = .08, n = 32, p = .659) or percentage of admissions with an adverse event (r = .10, n = 32, p = .603).

From viewing the scatterplots, one can see that unit size and safety climate along with unit size and teamwork climate did not have linear relationships. See Figures 4.8 and 4.9. Safety climate and number of adverse events per 1,000 patient days also did not have a linear relationship. See Figure 4.10.

Safety climate and teamwork climate The Pearson correlation between safety climate and teamwork climate was .89, which was positive, had a strong effect size, and was statistically different from zero (r = .89, n = 32, p = < 0.001). See Table 4.7 and Figure 4.11. Cohen advised using r as a measure of effect size, using the subjective standard of r = .1 as a weak effect, r = .3 as a moderate effect, and r = .5 as a strong effect (Cohen, 1988). The between-unit correlation between safety climate and teamwork climate was .72 in a sample of over 203 clinical areas (Sexton et al., 2006).

Safety climate with frequency of adverse events. The correlations between safety climate and frequency of adverse events were not significant. The correlation between safety

climate and number of adverse events per 1,000 patient days was .23, which was positive, had a weak effect size, and was not statistically different from zero (r = .23, n =32, p = .214). From viewing the scatterplot, one can see that safety climate and number of adverse events per 1,000 patient days did not have a linear relationship. See Figure 4.12.

Table 4.15.

Correlations Between Unit Size, Teamwork Climate, Safety Climate, and Frequency of Adverse Events

		Unit size	Teamwork climate	Safety climate	AEs per 1,000 patient days	AEs per 100 admissions	% of admissions with an AE
Teamwork climate	Pearson Correlation	-.07					
	Sig. (2-tailed)	.724					
	N	32					
Safety climate	Pearson Correlation	-.12	.89**				
	Sig. (2-tailed)	.51	.000				
	N	32	32				
AEs per 1,000 patient days	Pearson Correlation	.05	.21	.226			
	Sig. (2-tailed)	.776	.250	.214			
	N	32	32	32			
AEs per 100 admissions	Pearson Correlation	.08	.11	.08	.85**		
	Sig. (2-tailed)	.659	.542	.646	.000		
	N	32	32	32	32		
% of admissions with an AE	Pearson Correlation	.10	.04	.05	.86**	.98**	
	Sig. (2-tailed)	.603	.843	.809	.000	.000	
	N	32	32	32	32	32	

Note. ** indicates that the correlation is significant at the 0.01 level (2-tailed). AE is adverse event(s).

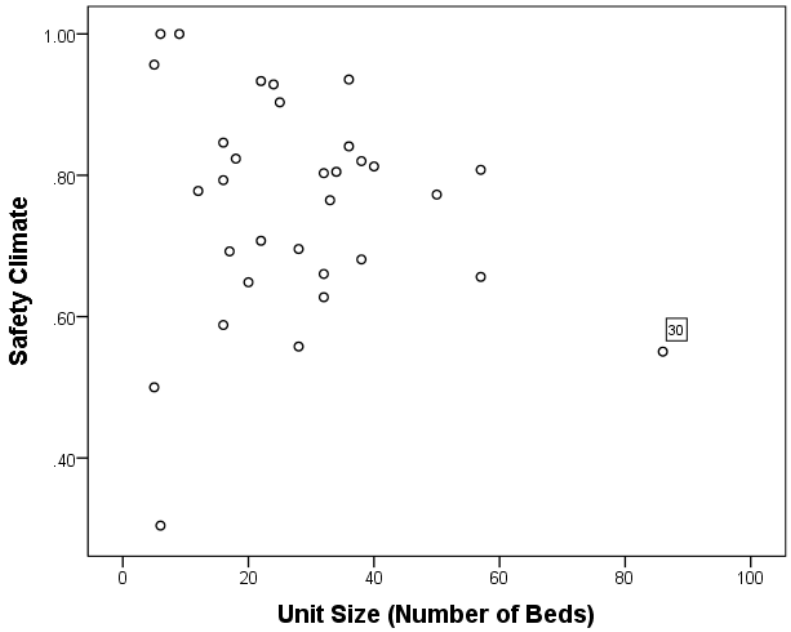

Figure 4.8. Scatterplot of unit size and safety climate. The marker labelled #30 (which appears to be an outlier regarding unit size) is one unit spread over two physical locations with the same staff and management serving both areas.

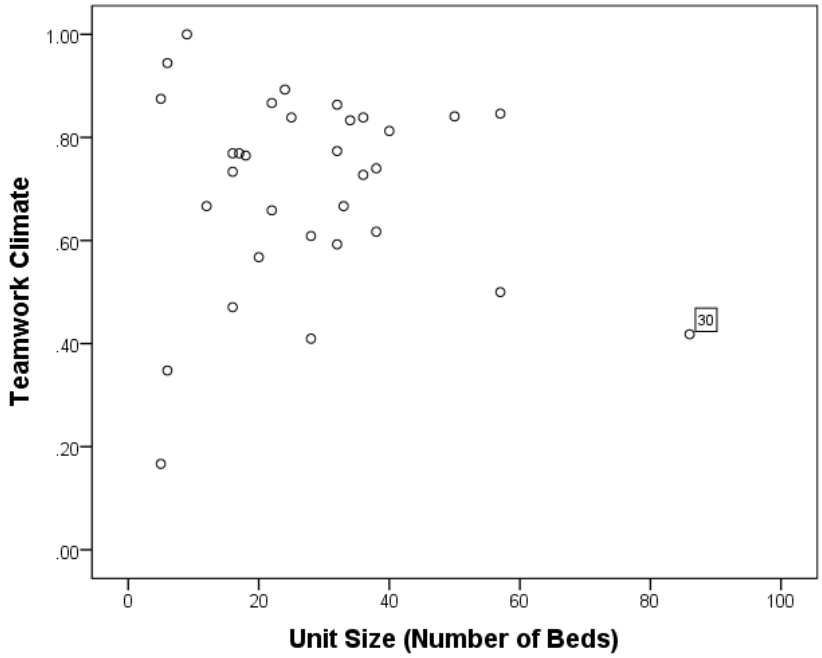

Figure 4.9. Scatterplot of unit size and teamwork climate. The marker labelled #30 (which appears to be an outlier regarding unit size) is one unit spread over two physical locations with the same staff and management serving both areas.

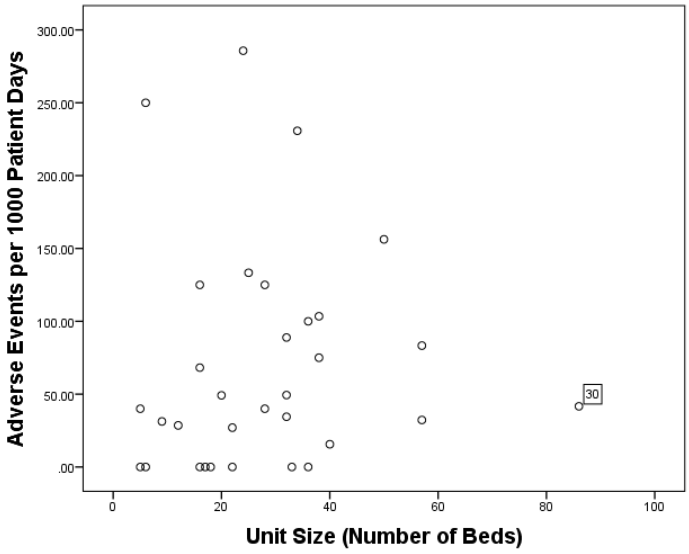

Figure 4.10. Scatterplot of unit size and adverse events per 1,000 patient days. The marker labelled #30 (which appears to be an outlier regarding unit size) is one unit spread over two physical locations with the same staff and management serving both areas.

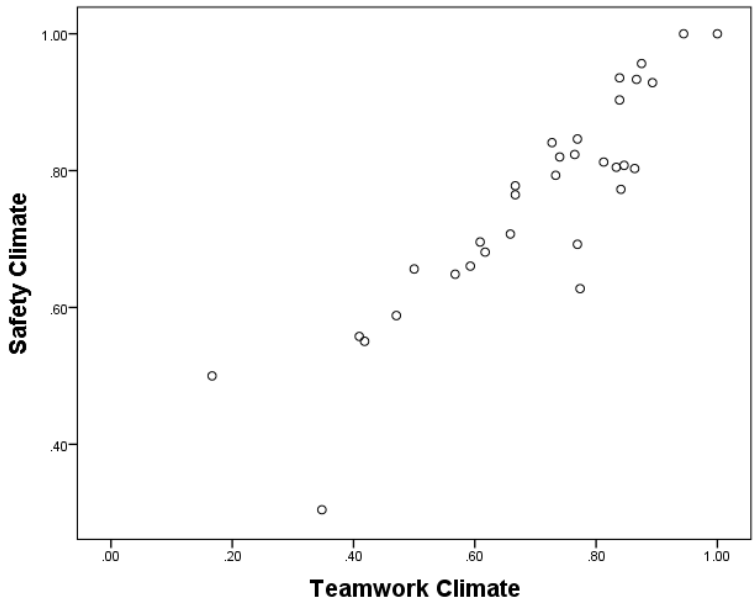

Figure 4.11. Scatterplot of teamwork climate and safety climate.

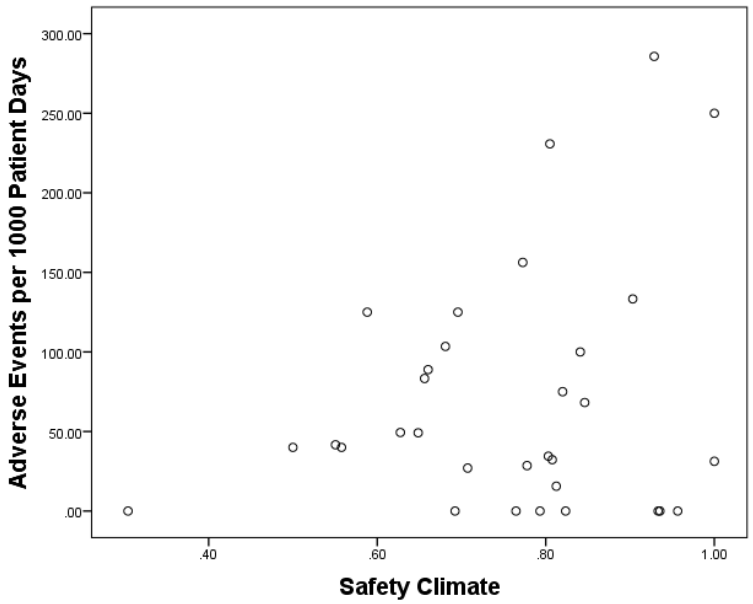

Figure 4.12. Scatterplot of safety climate and adverse events per 1,000 patient days.

The correlation between safety climate and number of adverse events per 100 admissions was .08, which was positive, had a very weak effect size, and was not statistically different from zero (r = .08, n = 32, p = .646). The correlation between safety climate and percentage of admissions with an adverse event was .05, which was positive, had a very weak effect size, and was not statistically different from zero (r = .05, n = 32, p = .809). There was no association between safety climate and frequency of adverse events.

Teamwork climate with frequency of adverse events. The Pearson correlations between teamwork climate and frequency of adverse events were not significant. The correlation between teamwork climate and number of adverse events per 1,000 patient days was .21, which was positive, had a weak effect size, and was not statistically different from zero (r = .21, n = 32, p = .250). From viewing the scatterplot, one can see that teamwork climate and number of adverse events per 1,000 patient days did not have a linear relationship. See Figure 4.13. The correlation between teamwork climate and number of adverse events per 100 admissions was 0.11, which was positive, had a weak effect size, and was not statistically different from zero (r = .11, n = 32, p = .542). The correlation between safety climate and percentage of admissions with

an adverse event was .04, which was positive, had a very weak effect size, and was not statistically different from zero (r = .04, n = 32, p = .843). There was no association between teamwork climate and frequency of adverse events.

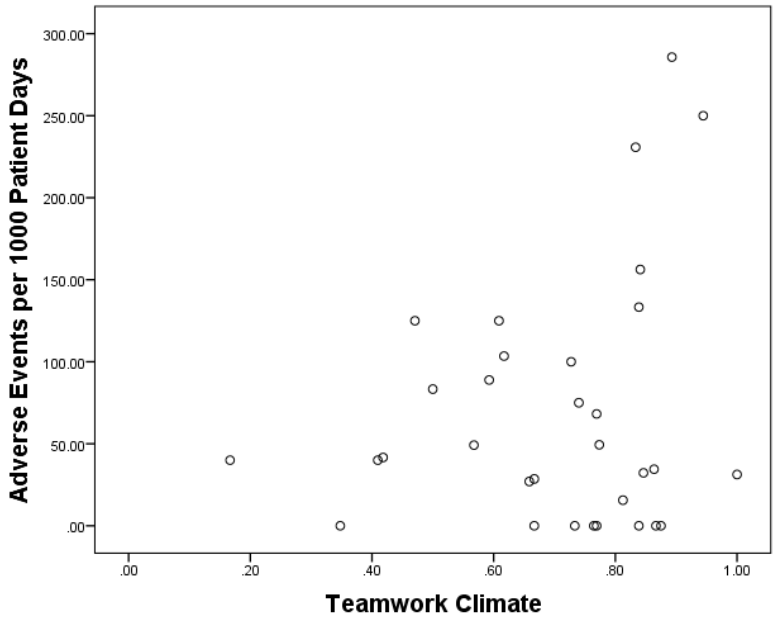

Figure 4.13. Scatterplot of teamwork climate and adverse events per 1,000 patient days.

Teamwork by unit type (medical surgical or other unit type). Contingency Table analyses were conducted to determine whether there was an association between dichotomized teamwork climate (0 = low teamwork, 1 = high teamwork) and unit type (0 = other unit type, 1 = medical surgical). From Table 4.16 it is apparent from the row marginal that 25% of the units had low teamwork. Two of the ten medical surgical units reported low teamwork. A greater frequency of other unit types experienced high teamwork compared to medical surgical units. See Figure 4.14. There was not a statistically significant association or relationship between dichotomized teamwork climate and unit type (X^2 = .19, df = 1, p = .660). See Table 4.9.

Table 4.16.

*Teamwork Climate (TC) Dichotomized into Two Groups * MedSurg Dichotomized into Two Groups Crosstabulation*

			MedSurg dichotomized into two groups		
			Other unit types	Medical Surgical units	Total
Teamwork climate (TC) dichotom-ized into two groups	Low teamwork, < .6001	Count	6	2	8
		% within TC categorized into 2 groups	75.0%	25.0%	100.0%
		% within MedSurg dichotomized into two groups	27.3%	20.0%	25.0%
		% of total	18.8%	6.3%	25.0%
	High teamwork, > .6001	Count	16	8	24
		% within TC categorized into 2 groups	66.7%	33.3%	100.0%
		% within MedSurg dichotomized into 2 groups	72.7%	80.0%	75.0%
		% of total	50.0%	25.0%	75.0%
Total		Count	22	10	32
		% within TC categorized into 2 groups	68.8%	31.3%	100.0%
		% within MedSurg dichotomized into 2 groups	100.0%	100.0%	100.0%
		% of total	68.8%	31.3%	100.0%

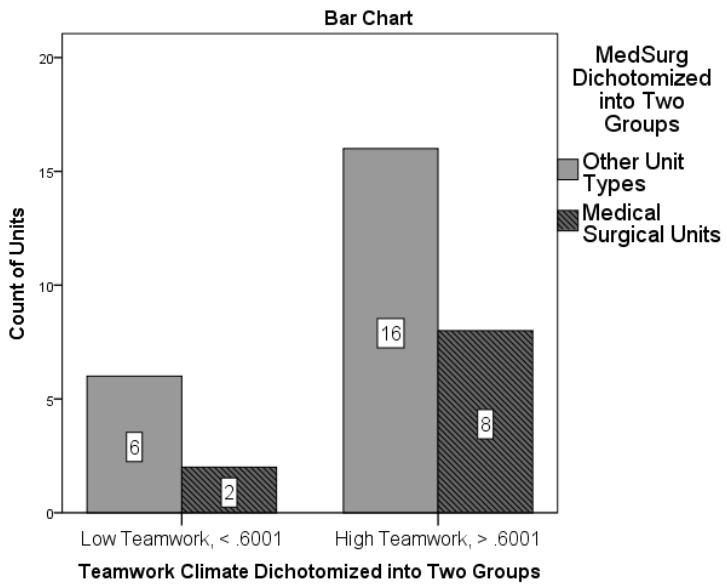

Figure 4.14. Bar chart comparing low versus high teamwork scores by unit type (medical surgical versus other unit type).

Unit type and relationship with dependent variable frequency of adverse events: Results from a one-factor analysis of variance (ANOVA) fixed effects model. A one-way analysis of variance was used to examine differences in unit type on frequency of adverse events. The independent variable was unit type with five different categories of grouping variable, and the dependent variable was frequency of adverse events (measured three ways). In the fixed effects model, all levels (unit types) are included in the design and analysis in the study (Lomax, 2007).

Testing of assumptions for analysis of variance*.*

Independence. The first ANOVA assumption is that each sample is an independent random sample from their representative population (Lomax, 2007). This can be tested by examining residual plots by group (Lomax, 2007). Scatterplots of adverse event residuals by unit demonstrated that the residuals fell into a random display of points for each group, which indicated that the assumption of independence was satisfied. See Figure 4.7 for a visual depiction the frequency of adverse events by unit type.

Homogeneity of Variance. The second assumption is homogeneity of variance, which assumes that the variances of each population are equal, which may be checked using Levene's test for equality of error variances (Lomax, 2007). Levene's test for equality of variances is not significant for adverse events per 1,000 patient days (p =.057), adverse events per 100 admissions (p = .077), or for percentage of admissions with an adverse event (p = .052). See Table 4.17. Levene's test suggests that the variances are not different, thus meeting the assumption of homogeneity of variance, although percentage of admissions with an adverse event was close to violating the assumption (p = .052). The effect of violating the homogeneity of variance assumption in a one-factor ANOVA is bias in the sum of squares within, which is the variability of observations within a group (i.e. unit type) combined across groups (Lomax, 2007). In addition, violation of the assumption can lead to an increase in Type I and/or Type II error (Lomax, 2007).

Table 4.17.

Test of Homogeneity of Variances for the Dependent Variables Adverse Events per 1,000 Patient Days, Adverse Events per 100 Admissions, and Percentage of Admissions with an Adverse Event

	Levene statistic	df1	df2	Sig.
AEs/1,000 patient days	2.63	4	27	.057
AEs/100 admissions	2.37	4	27	.077
% of admissions with an adverse event	2.70	4	27	.052

Note. The Levene tests the null hypothesis that the error variance of the dependent variable is equal across groups. The design is intercept + unit (types).

Normality. The normality assumption was checked via skewness and kurtosis statistics for frequency of adverse events, operationalized as adverse events per 1,000 patient days, adverse events per 100 admissions, and percentage of admissions with an adverse event. The skewness and kurtosis ratios for the five unit types fell within the range of -2 to +2, so it appears that the normality assumption was met. However, the Shapiro-Wilk statistics demonstrate that the intermediate care, OB/gynecology, and procedural units may not follow a normal distribution because their p values are less than .05 in Tables 4.18, 4.19, and/or 4.20. For example, in Table

4.18 for adverse events per 1,000 days, OB/gynecology and procedural units have values less than .05. Tables 4.19 and 4.20 demonstrate that intermediate care, OB/gynecology, and procedural units all have Shapiro-Wilk statistics less than .05, indicating non-normal distribution. For the intermediate care (*n* = 3) and OB/gynecology units (*n* =3), the non-normal distribution is likely due to the small sample size of this unit type. There may be increased likelihood of Type I or Type II errors due to unequal *n*'s among groups (Lomax, 2007). Because four out of the five unit types do not exhibit a normal distribution for the adverse events per 100 admissions, the ANOVA will be conducted using the dependent variables adverse events per 1,000 days and percentage of admissions with an adverse event (only).

Table 4.18.

Tests of Normality for Adverse Events per 1,000 Patient Days by Unit Type

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Type of unit	statistic	df	Sig.	statistic	df	Sig.
Adverse events per 1,000 patient days	Critical care	.26	5	.20*	.83	5	.127
	Intermediate care	.26	3	.	.96	3	.597
	Medical surgical	.21	10	.20*	.93	10	.414
	OB gynecology	.39	3	.	.75	3	.000
	Procedural area	.23	11	.11	.78	11	.005

Note. * indicates a lower bound of the true significance. [a] indicates a Lilliefors Significance Correction.

Table 4.19.

Tests of Normality for Adverse Events per 100 Admissions by Unit Type

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Type of unit	statistic	df	Sig.	statistic	df	Sig.
Adverse events per 100 admissions	Critical care	.35	5	.05	.77	5	.046
	Intermediate care	.39	3	.	.75	3	.000
	Medical surgical	.17	10	.20*	.94	10	.569
	OB gynecology	.39	3	.	.75	3	.000
	Procedural area	.28	11	.02	.82	11	.017

Note. * indicates a lower bound of the true significance. [a] indicates a Lilliefors Significance Correction.

Table 4.20.

Tests of Normality for Percentage of Admissions with an Adverse Event by Unit Type

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Type of unit	statistic	df	Sig.	statistic	df	Sig.
% of admissions with an adverse event	Critical care	.30	5	.16	.88	5	.325
	Intermediate care	.39	3	.	.75	3	.000
	Medical surgical	.25	10	.09	.90	10	.202
	OB gynecology	.39	3	.	.75	3	.000
	Procedural area	.28	11	.02	.82	11	.017

Note. [a] indicates Lilliefors Significance Correction.

A one-way ANOVA was conducted to determine if the five unit types differed significantly in adverse events per 1,000 patient days and percentage of admissions with an adverse event. The ANOVAs were statistically significant for number of adverse events per 1,000 patient days and percentage of admissions with an adverse event ($F = 2.84$, $df = 4,27$, $p = .044$; $F = 3.92$, $df = 4,27$, $p = .012$). See Table 4.21. Because the results of the ANOVA were significant, it was necessary to conduct a multiple comparison procedure (MCP) to determine which means or combination of means were different among the unit types (groups).

Table 4.21

One-Factor Fixed Effects ANOVA: Unit Differences for Adverse Events per 1,000 Patient Days and Percentage of Admissions with an Adverse Event

		Sum of squares	df	Mean square	F	Sig.
Adverse events per 1,000 patient days	Between groups	52790.78	4	13197.69	2.84	.044
	Within groups	125418.36	27	4645.13		
	Total	178209.13	31			
% of admissions with an adverse event	Between groups	2745.63	4	686.41	3.92	.012
	Within groups	4723.76	27	174.95		
	Total	7469.39	31			

An overall omnibus test assesses the equality of all the means simultaneously while controlling the error rate and maximizing power (Lomax, 2007). There are many types of multiple comparison procedures (MCPs), each with advantages and disadvantages. The Bonferroni MCP

tests pairwise or complex contrasts for balanced or unbalanced designs and the alpha level is split among the set of planned contrasts (Lomax, 2007). The Bonferroni procedure is slightly conservative (not as powerful) because the family-wise error rate may be less than the alpha level (Lomax, 2007). It is one of the MCPs least likely to make a Type I error, where the researcher incorrectly rejects a true null hypothesis (Lomax, 2007). Applied to this analysis, a Type I error would be stating that there is a significant difference in adverse events among unit types when in fact there is no significant difference between units (the null hypothesis). In this study, there was an unbalanced design because there were different sample sizes for each unit type, and thus, the Bonferroni MCP was an appropriate post hoc MCP for this study. The Bonferroni post-hoc test revealed that the frequency of adverse events measured via adverse events per 1,000 days did not differ between unit types. See Table 4.22. The Bonferroni post-hoc test revealed that the percentage of admissions with an adverse event was statistically significantly less in OB/gynecology units (-28.667, p = .028) compared to medical surgical units. See Table 4.22.

Table 4.22

Bonferroni Multiple Comparison Procedure for Adverse Events per 1,000 Patient Days and Percentage of Admissions with an Adverse Event

Dependent variable	(I) Unit type	(J) Unit type	Mean difference (I-J)	Std. error	Sig.	95% Confidence interval Lower bound	95% Confidence interval Upper bound
AEs per 1,000 patient days	Critical care	Intermediate care	13.60	49.77	1.000	-138.53	165.74
		Medical surgical	-87.10	37.33	.273	-201.20	27.00
		OB gynecology	24.69	49.77	1.000	-127.45	176.82
		Procedural unit	-20.35	36.76	1.000	-132.71	92.01
	Intermediate care	Critical care	-13.60	49.77	1.000	-165.74	138.53
		Medical surgical	-100.70	44.87	.332	-237.84	36.43
		OB gynecology	11.08	55.65	1.000	-159.01	181.17
		Procedural unit	-33.96	44.39	1.000	-169.64	101.73
	Medical surgical	Critical care	87.10	37.33	.273	-27.00	201.20
		Intermediate care	100.70	44.87	.332	-36.43	237.84
		OB gynecology	111.79	44.87	.192	-25.35	248.92
		Procedural unit	66.75	29.78	.334	-24.27	157.77
	OB Gynecology	Critical care	-24.69	49.77	1.000	-176.82	127.45
		Intermediate care	-11.08	55.65	1.000	-181.17	159.01
		Medical surgical	-111.79	44.87	.192	-248.92	25.35
		Procedural unit	-45.04	44.39	1.000	-180.72	90.65
	Procedural unit	Critical care	20.35	36.76	1.000	-92.01	132.71
		Intermediate care	33.96	44.39	1.000	-101.73	169.64
		Medical surgical	-66.75	29.78	.334	-157.77	24.27
		OB gynecology	45.04	44.39	1.000	-90.65	180.72

Continued

Continued Table 4.22

% of admissions with an AE	Critical care	Intermediate care	6.67	9.66	1.000	-22.86	36.19
		Medical surgical	-12.00	7.25	1.000	-34.14	10.14
		OB gynecology	16.67	9.66	.959	-12.86	46.19
		Procedural unit	5.58	7.13	1.000	-16.22	27.39
	Intermediate care	Critical care	-6.67	9.66	1.000	-36.13	22.86
		Medical surgical	-18.67	8.71	.412	-45.28	7.95
		OB gynecology	10.00	10.80	1.000	-23.01	43.01
		Procedural unit	-1.08	8.62	1.000	-27.42	25.25
	Medical surgical	Critical care	12.00	7.25	1.000	-10.14	34.14
		Intermediate care	18.67	8.71	.412	-7.95	45.28
		OB gynecology	28.67*	8.71	.028	2.05	55.28
		Procedural unit	17.58	5.78	.052	-.08	35.25
	OB Gynecology	Critical care	-16.67	9.66	.959	-46.19	12.86
		Intermediate care	-10.00	10.80	1.000	-43.01	23.01
		Medical surgical	-28.67*	8.71	.028	-55.28	-2.05
		Procedural unit	-11.08	8.62	1.000	-37.42	15.25
	Procedural unit	Critical care	-5.58	7.13	1.000	-27.39	16.22
		Intermediate care	1.08	8.62	1.000	-25.25	27.42
		Medical surgical	-17.58	5.78	.052	-35.25	.08
		OB gynecology	11.08	8.62	1.000	-15.25	37.42

Note. * indicates the mean difference is significant at the 0.05 level.

The following subsection presents the analytic strategy related to aim three. To determine if safety climate and teamwork climate varied by unit size, there was a plan to estimate ordinary least squares (OLS) with safety climate and/or teamwork climate as the dependent

measure and unit size as the independent measure. Due to lack of correlation and linear relationship between unit size and safety climate, ordinary least squares (linear regression) analyses were not run with unit size as a control variable because the assumption of linearity was violated. Violating the assumption of linearity may jeopardize the meaningfulness of the interpretation of the regression coefficient in linear regression (Hayes, 2013). Instead, unit type was selected as a control variable for the analyses.

The subsection begins with an independent t-test to examine the extent to which medical surgical unit versus other unit type means differed for dichotomized teamwork climate. Next, an independent t-test determines the extent to which low and high teamwork units means differ with respect to adverse event frequency. Following that, a two-factor ANOVA fixed model explores the relationship between unit type and dichotomized teamwork climate with adverse event occurrence. A hierarchical multiple regression examines how predictive unit type, teamwork climate, safety climate, and the interaction between teamwork and safety climate were regarding adverse event frequency. The section concludes with a logistic regression analysis that predicts the likelihood that a unit would experience an adverse event in this sample of 32 units.

Analytic Strategy

Relationship between unit type with predictors dichotomized teamwork climate and safety climate: Results of an independent *t*-test An independent t-test determined the extent to which medical surgical unit versus other unit type means differed for dichotomized teamwork climate (TC) and safety climate (SC). As shown in Table 4.23, medical surgical units had an average safety climate of .80 (*SD* = .11) and other unit types had a lower mean of 0.74 (*SD* = .17).

The independence assumption was checked by plotting residual (for teamwork climate and safety climate) by group (medical surgical units versus other unit types). Residual scatterplots demonstrated a random display of points, indicating that the independence assumption was satisfied. According to the Levene's test, the homogeneity of variance assumption was satisfied for dichotomized teamwork climate and safety climate ($F = .82$, $p = 0.372$; $F = 1.99$, $p = 0.168$). The independent *t* test indicated that the dichotomized teamwork

climate and safety climate means were not statistically significant (t = -4.28, df = 30, p = .672; t = -1.09, df = 30, p = .286). See Table 4.24. Thus, the null hypothesis that the teamwork and safety climate means were the same by unit type was not rejected at the .05 level of significance. Dichotomized teamwork climate and safety climate means did not statistically differ between medical surgical units and other types of units.

Table 4.23

Group Statistics: Dichotomized Teamwork Climate and Safety Climate in Medical Surgical Units Versus Other Unit Types

	MedSurg dichotomized into two groups	N	Mean	Std. deviation	Std. error Mean
TC categorized into 2 groups	Other unit types	22	.73	.46	.10
	Medical surgical units	10	.80	.42	.13
Safety climate	Other unit types	22	.74	.17	.04
	Medical surgical units	10	.80	.11	.03

Note. TC is teamwork climate and SC is safety climate.

Table 4.24

Independent Samples Test: Dichotomized Teamwork Climate and Safety Climate in Medical Surgical Units Versus Other Unit Types

		Levene's test for equality of variances		t-test for equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean diff-erence	Std. error diff-erence	95% confidence interval of the difference Lower	Upper
TC dicho-tomized into 2 groups	Equal variances assumed	.82	.372	-.43	30.00	.672	-.07	.17	-.42	.28
	Equal variances not assumed			-.44	18.83	.664	-.07	.17	-.42	.27
Safety climate	Equal variances assumed	1.99	.168	-1.09	30.00	.286	-.06	.06	-.19	.06
	Equal variances not assumed			-1.29	26.78	.208	-.06	.05	-.17	.04

Note. TC is teamwork climate.

Relationship between low and high teamwork units and frequency of adverse events: Results from an independent *t*-test. An independent t-test determined the extent to which low and high teamwork units means differed for frequency of adverse events. Higher teamwork units experienced an unexpected greater average amount of adverse events per 1,000 patient days compared to low teamwork units (*M* = 72.76, *SD* = 85.09; *M* = 58.51, *SD* = 38.62). However, mean adverse events per 100 admissions were nearly the same for low teamwork units (*M* = 22.32, *SD* = 13.77) and high teamwork units (*M* = 21.67, *SD* = 19.26). Mean percentage of admissions with an adverse event were also quite similar for low teamwork units (*M* = 21.07, *SD* = 13.44) and high teamwork units (*M* = 19.17, *SD* = 16.40). See Table 4.25.

Table 4.25

Group Statistics: Frequency of Adverse Events in Low and High Teamwork Units

	TC Categorized into 2 Groups	N	Mean	Std. deviation	Std. error Mean
AEs per 1,000 patient days	Low teamwork, < .6001	8	58.51	38.62	13.65
	High teamwork, > .6001	24	72.76	85.10	17.37
AEs per 100 admissions	Low teamwork, < .6001	8	22.32	13.79	4.87
	High teamwork, > .6001	24	21.67	19.26	3.93
% of admissions with an adverse event	Low teamwork, < .6001	8	21.07	13.44	4.75
	High teamwork, > .6001	24	19.17	16.40	3.35

Note. TC is teamwork climate. AEs is adverse events.

According to the Levene's test, the homogeneity of variance assumption was satisfied for adverse events per 1,000 patient days ($F = 3.916$, $p = .057$), adverse events per 100 admissions

(*F* = .796, *p* = .380), and percentage of admissions with an adverse event (*F* = .491, *p* = .489). See Table 4.26.

The independent *t* test indicated that the mean adverse events per 1,000 patient days, adverse events per 100 admissions, and percentage of admissions with an adverse event were not statistically different between low and high teamwork units (*t* = .455, *p* = .653; *t* = .088, *p* = .930, *t* = .296, *p* = .769). There was no statistically significant difference in mean adverse events between low and high teamwork units. See Table 4.26.

Table 4.26

Independent Samples Test: Frequency of Adverse Events in Low and High Teamwork Units

		Levene's test for equality of variances		*t*-test for equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean diff-erence	Std. error diff-erence	95% confidence interval of the difference Lower	Upper
AEs/ 1,000 patient days	Equal variances assumed	3.92	.057	-.46	30	.653	-14.25	31.36	-78.29	49.79
	Equal variances not assumed			-.65	26.71	.524	-14.25	22.10	-59.61	31.10
AEs per 100 ad-missions	Equal variances assumed	.80	.380	.09	30	.930	.66	7.40	-14.46	15.77
	Equal variances not assumed			.11	16.90	.918	.66	6.26	-12.56	13.87
% of Ad-missions with an AE	Equal variances assumed	.491	.489	.27	30	.769	1.90	6.43	-11.23	15.04
	Equal variances not assumed			.33	14.58	.748	1.90	5.81	-10.51	14.32

Note. AEs is adverse events.

Relationship between unit type and dichotomized teamwork climate with frequency of adverse events: Results from a two-factor analysis of variance (ANOVA) fixed effects model. A two-way ANOVA examined the effects of medical surgical unit (zero = other unit type, one = medical surgical unit) and dichotomized teamwork climate (zero = low teamwork, one = high teamwork) on adverse events per 1,000 patient days. Among units with high teamwork, medical surgical units experienced greater mean adverse events (per 1,000 patient days) than other unit types. Non-medical surgical (other) units with low teamwork experienced greater mean adverse events than high teamwork units did (*M* = 49.31, *SD* = 40.97; *M* = 41.36, *SD* = 64.72). See Table 4.27.

Table 4.27

Descriptive Statistics for Dichotomized Teamwork Climate and Unit Type with Adverse Events per 1,000 Patient Days: Results of a Two-Factor ANOVA

Teamwork climate dichotomized into two groups	Unit type dichotomized into two groups	Mean	Std. deviation	N
Low teamwork, < .6001	Other unit types	49.31	40.97	6
	Medical surgical units	86.11	3.93	2
	Total	58.51	38.62	8
High teamwork, > .6001	Other unit types	41.36	64.72	16
	Medical surgical units	135.56	89.75	8
	Total	72.76	85.10	24
Total	Other unit types	43.53	58.35	22
	Medical surgical units	125.67	81.86	10
	Total	69.20	75.82	32

Homogeneity of variance. An ANOVA assumption is homogeneity of variance, which assumes that the variances of each population are equal (Lomax, 2007). Levene's test for equality of variances was not significant for adverse events per 1,000 patient days (*p* =.249), indicating that the variances were equal and the homogeneity of variance assumption was satisfied. See Table 4.28.

Table 4.28.

Levene's Test of Equality of Error Variances for the Dependent Variable: Adverse Events per 1,000 Patient Days

F	df1	df2	Sig.
1.45	3	28	.249

Note. Tests the null hypothesis that the error variance of the dependent variable is equal across groups. The design was Intercept + MEDSURG + TC_Categ + MEDSURG * TC_Categ.

Table 4.29 demonstrates that there was a significant main effect for unit type (F = 4.41, df = 1,28, p = .045). Medical surgical units experienced significantly more adverse events per 1,000 patient days than other unit types. The effect size for unit type was large (partial η^2 = .14) per Cohen's standards that indicate a large effect with partial η^2 = .14 (Cohen, 1988). Observed power was .53. There was not a significant main effect of dichotomized teamwork climate on frequency of adverse events per 1,000 patient days (F = .44, df = 1,28, p = .511). The interaction between medical surgical units and dichotomized teamwork climate was not significant (F = .85, df = 1,28, p = .365). See Figure 4.15 for a profile plot of two-factor ANOVA effect between unit type and low and high teamwork units.

Table 4.29

Tests of Between-Subjects Effects for Dichotomized Teamwork Climate and Unit Type with Adverse Events per 1,000 Patient Days: Two-Factor ANOVA

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared	Noncent. parameter	Observed power[b]
TC_Categ	2016.94	1	2016.94	.44	.511	.02	.44	.10
MEDSURG	20092.92	1	20092.92	4.41	.045*	.14	4.41	.53
TC_Categ * MEDSURG	3857.27	1	3857.27	.85	.365	.03	.85	.14
Error	127629.12	28	4558.18					
Total	331440.52	32						

Continued

Continued Table 4.29

Corrected Total	178209.14	31

Note. Computed using alpha = 0.05. *$p < .05$. MEDSURG is the variable that dichotomized unit types into two groups, 0 = other unit type and 1 = medical surgical unit. TC_Categ is teamwork climate dichotomized into two groups, low and high teamwork. MEDSURG * TC_Categ is the interaction term between MEDSURG and teamwork climate dichotomized into two groups.

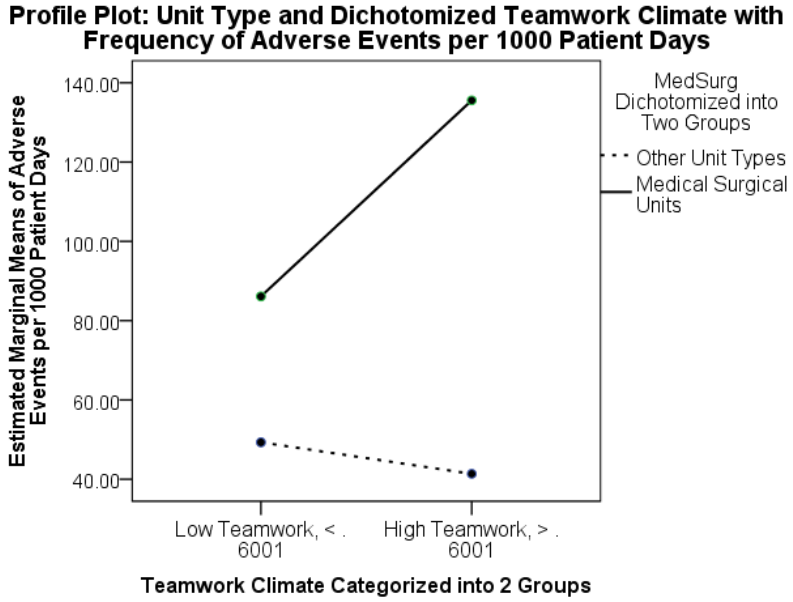

Figure 4.15. Profile plot of two-factor ANOVA effect between unit type and low and high teamwork units. There appears to be a main effect for unit type (medical surgical unit solid line is greater than the dashed "other unit" line regarding Y-axis adverse events per 1,000 patient days), but no main effect for teamwork climate. There is no interaction effect because the lines do not cross.

Relationship between unit type, dichotomized teamwork climate and safety climate with frequency of adverse events: Hierarchical multiple regression results. When a researcher uses hierarchical multiple regression, each independent variable is entered into the model in a series of steps, and the researcher controls the variable entry, based upon logic or theory (Polit, 2010). A reason for conducting hierarchical regression analysis is to determine the

effect of crucial independent variables after the effect of other variables has been controlled (Polit, 2010), by reviewing the R square change as each independent variable joins the others (Cohen et al., 2003). A hierarchical multiple regression was performed to examine how predictive unit type, teamwork climate, safety climate, and the interaction between teamwork and safety climate were on frequency of adverse events, measured via adverse events per 1,000 patient days. The rationale underlying the hierarchical order for variable entry was the causal priority and structural property of adverse events. First, patients and health care workers are assigned to hospital unit type. Teams organize health care work on units to manage care (Carayon et al., 2013). Teams create patient safety (Knox & Simpson, 2004), and safety climate is a shared belief about the work environment on a unit level (Etchegaray & Thomas, 2014). Hospitals establish units and develop teams that establish safety culture, which influences care, resulting in patient outcomes (adverse events).

Evaluation of assumptions.

Normality. Hierarchical multiple regression requires an evaluation of the assumption of normal distribution (Tabachnick & Fidell, 2013). Unit type and dichotomized teamwork climate were categorical variables, so an examination of normality was not indicated. Safety climate was normally distributed, as described in the descriptive statistics section. See Figure 4.3.

In an expected normal probability plot, in a normal distribution the points for the cases fall along the diagonal running from lower left to upper right (Tabachnick & Fidell, 2013). In Figure 4.16, it appears that at low adverse event frequency, there are too many cases above the diagonal, and at high adverse event frequency, there are too many cases below the diagonal, which may reflect patterns of skewness and kurtosis (Tabachnick & Fidell, 2013).

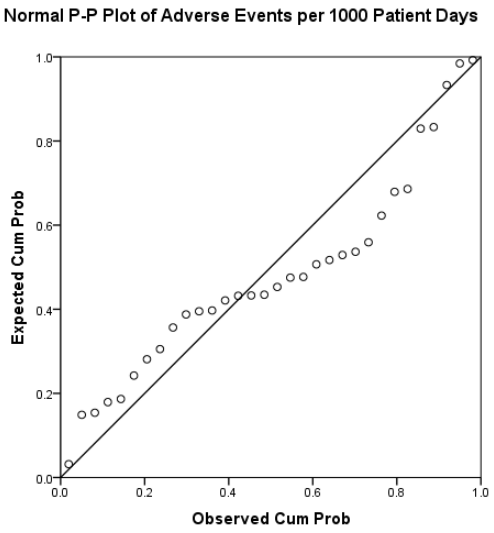

Figure 4.16. An expected normal P-P plot for adverse events per 1,000 patient days used to assess the assumption of normality in regression..

Outliers, linearity, homoscedasticity.. Residual scatterplots test of the assumptions of normality, linearity, and homoscedasticity between the predicted dependent variable (frequency of adverse events) and errors of prediction (Tabachnick & Fidell, 2013). Normally distributed residuals should line up along a horizontal line (Tabachnick & Fidell, 2013). It is somewhat difficult to interpret the scatterplot in Figure 4.17 due to the small amount of cases. Some points appear to fan out to the right, which may indicate heteroscedasticity (Tabachnick & Fidell, 2013).

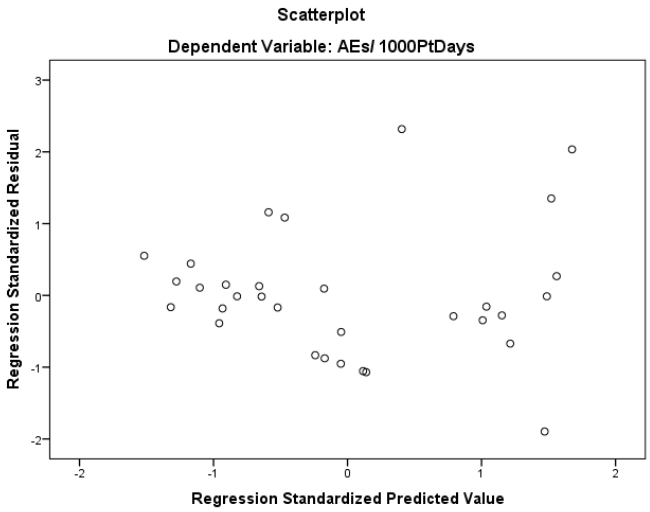

Figure 4.17. Plot of the predicted values of the dependent variable, number of adverse events per 1,000 patient days..

Model summary The dependent variable was adverse events per 1,000 days. Model one included unit type, and unit types were dummy coded such that zero was other unit type and one was equal to the selected unit type. Medical surgical unit type was left out of the coding schema so that it became the comparison unit. Model two added dichotomized teamwork climate. Model three added safety climate. Model four added the interaction term between dichotomized teamwork climate and safety climate. See Table 4.30 for variables entered in each model.

Table 4.30

Variables Entered/Removed for Hierarchical Multiple Regression Analyses Predicting Frequency of Adverse Events per 1,000 Patient Days

Model	Variables entered	Variables removed	Method
1	Procedural unit, OB gynecology, intermediate care, critical care		. Enter
2	Teamwork dichotomized		. Enter
3	Safety climate		. Enter
4	Teamwork dichotomized X Safety climate interaction		. Enter

Note. All requested variables were entered.

Model 1 (unit type). Unit size was not used as a control variable because it was not correlated with frequency of adverse events. Instead, unit type served as a control variable unit characteristic.

Model 1 suggested that 29.6% of the variance in adverse events was predicted by unit type (R^2= .30, $F_{4, 27}$ = 2.84, p = .044). See Table 4.31. The Model 1 regression equation was adverse events = 125.67 – 87.1(CC) – 100.70(IC) -111.76(OB) -66.74(PR). (CC is critical care, IC is intermediate care, OB is OB gynecology, and PR is procedural units). See Table 4.32 for regression coefficients.

The mean number of adverse events per 1,000 patient days for medical surgical units (comparison group) was 125.67. See Table 4.32. Changing from a medical surgical unit to a critical care unit corresponded with a significant 87.1 reduction in adverse events per 1,000 patient days (t = -2.33 p = .027). Changing from a medical surgical unit to an intermediate care, OB/gynecology or procedural unit corresponded with a significant 100.70, 111.79, 66.75 reduction in adverse events per 1,000 patient days (t = -2.25, p = .033; t = -2.49, p = .019; t = -2.24, p = .033). See Table 4.32.

Table 4.31

Hierarchical Multiple Regression Model Summary Predicting Frequency of Adverse Events per 1,000 Patient Days from Unit Type, Dichotomized Teamwork Climate, Safety Climate, and Interaction Term of Dichotomized Teamwork Climate and Safety Climate

					Change statistics				
Model	R	R square	Adjusted R square	Std. error of the estimate	R square change	F change	df1	df2	Sig. F change
1	.54[a]	.30	.19	68.16	.30	2.84	4	27	.04
2	.55[b]	.30	.17	69.20	.01	.19	1	26	.67
3	.56[c]	.31	.15	69.94	.01	.46	1	25	.51
4	.56[d]	.32	.12	71.30	.00	.05	1	24	.83

Continued

Continued Table 4.31

Note. In model 1, the predictors are unit type, comprised of procedural unit, OB gynecology, intermediate care, critical care, and medical surgical (comparison group). In model 2, the predictors are unit type and teamwork dichotomized. In model 3, the predictors are unit type, teamwork dichotomized, and safety climate. In model 4, the predictors are unit type, teamwork dichotomized, safety climate, and an interaction term with teamwork dichotomized and safety climate. The dependent variable is adverse events per 1,000 patient days.

Table 4.32

Hierarchical Multiple Regression Analyses Predicting Frequency of Adverse Events per 1,000 Patient Days from Unit Type, Dichotomized Teamwork Climate, Safety Climate, and Interaction Term of Dichotomized Teamwork Climate and Safety Climate

		Unstandardized coefficients		Standardized coefficients			Correlations		
Model		B	Std. error	Beta	t	Sig.	Zero-order	Partial	Part
1	(Constant)	125.67	21.55		5.83	.000			
	Critical care	-87.10	37.33	-.42	-2.33	.027	-.18	-.41	-.37
	Intermediate care	-100.70	44.87	-.39	-2.25	.033	-.19	-.40	-.36
	OB gynecology	-111.79	44.87	-.44	-2.49	.019	-.24	-.43	-.40
	Procedural unit	-66.75	29.78	-.43	-2.24	.033	-.10	-.40	-.36
2	(Constant)	115.49	32.02		3.61	.001			
	Critical care	-87.10	37.90	-.42	-2.30	.030	-.18	-.41	-.38
	Intermediate care	-103.25	45.93	-.40	-2.25	.033	-.19	-.40	-.37
	OB gynecology	-110.09	45.72	-.43	-2.41	.023	-.24	-.43	-.40
	Procedural unit	-64.67	30.61	-.41	-2.11	.044	-.10	-.38	-.35
	TC dichotomized into 2 groups	12.73	29.21	.07	.44	.667	.08	.09	.07
3	(Constant)	65.26	81.05		.81	.428			
	Critical care	-84.63	38.48	-.41	-2.20	.037	-.18	-.40	-.36
	Intermediate care	-96.65	47.43	-.38	-2.04	.052	-.19	-.38	-.34
	OB gynecology	-101.91	47.76	-.40	-2.13	.043	-.24	-.39	-.35
	Procedural unit	-62.11	31.17	-.40	-2.00	.057	-.10	-.37	-.33
	Teamwork dichotomized	-9.83	44.55	-.06	-.22	.827	.08	-.04	-.04
	Safety climate	85.31	126.21	.18	.68	.505	.23	.13	.11
4	(Constant)	38.00	148.05		.26	.800			
	Critical care	-85.07	39.28	-.41	-2.17	.041	-.18	-.40	-.37
	Intermediate care	-96.95	48.38	-.38	-2.00	.056	-.19	-.38	-.34
	OB gynecology	-102.21	48.72	-.40	-2.10	.047	-.24	-.39	-.35
	Procedural unit	-59.58	33.77	-.38	-1.76	.090	-.10	-.34	-.30
	Teamwork dichotomized	32.87	197.68	.19	.17	.869	.08	.03	.03
	Safety climate	132.03	246.77	.27	.54	.598	.23	.11	.09

Continued

Continued Table 4.32

Teamwork dichotomized X SC interaction	-66.23	298.49	-.33	-.22	.826	.12	-.05	-.04

Model 2 (unit type and dichotomized teamwork climate). In Model 2, teamwork climate was entered after controlling for unit type. Model two as a whole explained that 30.1% of the variance in adverse events was predicted by dichotomized teamwork climate after controlling for unit type (R^2= .30, $F_{1, 26}$ = 2.24, p = .08). See Table 4.31. The inclusion of dichotomized teamwork in the model resulted in a non-significant 0.5% increase in the amount of variance in adverse events explained above and beyond that accounted for by unit type (ΔR^2 = .01, p = .667). Taken together, unit type and teamwork climate were not significantly related to adverse events. See Table 4.31.

Model 3 (unit type, dichotomized teamwork climate and safety climate). In Model 3, safety climate was entered after controlling for unit type and teamwork. Model three suggested that 31% of the variance in adverse events was predicted by safety climate after controlling for unit type and teamwork (R^2= .31, $F_{1, 25}$ = 1.91, p = .119). See Table 4.31. The inclusion of safety climate in the model resulted in a non-significant 1% increase in the amount of variance in adverse events explained above and beyond that accounted for by unit type and teamwork (ΔR^2 = .01, p = .505). Taken together, unit type, teamwork climate and safety climate were not significantly related to adverse events.

Model 4 (unit type, dichotomized teamwork climate, safety climate, and interaction term). In Model 4, an interaction term between teamwork and safety climate was entered after controlling for unit type, teamwork, and safety climate. Model four suggested that 32% of the variance in adverse events was predicted by the interaction term of dichotomized teamwork and safety climate after controlling for unit type, dichotomized teamwork, and safety climate (R^2= .32, $F_{1, 24}$ = 1.58, p = .19). See Table 4.31. The inclusion of the interaction term in the model resulted in a non-significant 0% increase in the amount of variance in adverse events explained above and beyond that accounted for by unit type, teamwork and safety climate (ΔR^2 = .00, p = .826).

Taken together, the interaction, unit type, teamwork climate and safety climate were not significantly related to adverse events.

Alternative hypothesis of predicting adverse event occurrence. An alternative hypothesis was developed in order to examine the likelihood that a unit would experience an adverse event because 25% of the sample units (a somewhat large portion) did not experience any adverse events. Frequency of adverse events was recoded into a dichotomous variable.

A logistic regression analysis was conducted to predict the likelihood that a unit would experience an adverse event in this sample of 32 units. Three predictor variables, unit type (1= medical surgical, 2= critical care, 3= intermediate care, 4= OB/gynecology, 5 = procedural unit), dichotomized teamwork climate (0 = low teamwork, 1 = high teamwork) and safety climate (a continuous variable) were used in the analysis, with simultaneous entry of predictors. The outcome variable was units experiencing an adverse event (1 = yes, 0 = no). Eight of the 32 units did not experience any adverse events. See Figure 4.18.

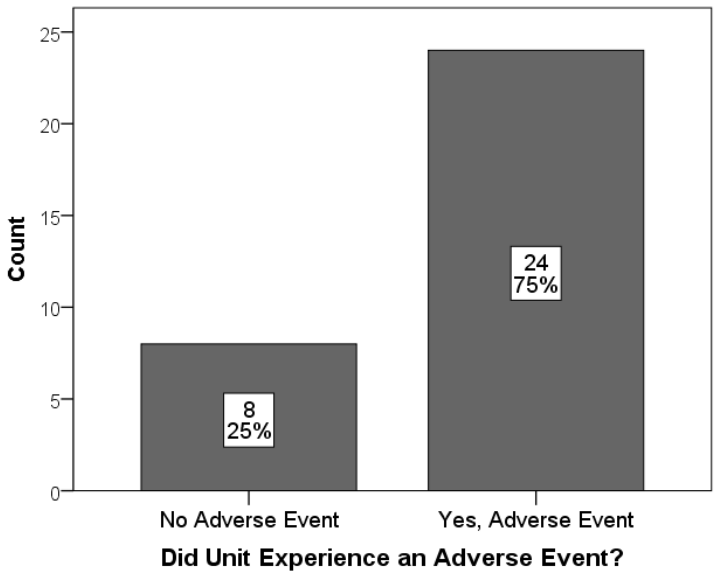

Figure 4.18. Bar chart depicting distribution of adverse events in two groups of units, with group zero not experiencing an adverse event and group one experiencing an adverse event.. *n* = 32 units.

As shown in Table 4.33, the omnibus test of the full model for predicting an adverse event was statistically significant (likelihood ratio chi-square = 14.26, *df* = 6, *p* = .027). We can reject the null hypotheses that all of the predictor effects are zero. However, an examination of Table 4.34 demonstrates that none of the independent variables were significant in predicting the likelihood of a unit experiencing an adverse event. Unit type, dichotomized teamwork climate, and safety climate were not related to whether an adverse event occurred. Reasons for the non-significant results may be due to small sample size. Some experts recommend a minimum of observation-to-predictor ratio of 10 to one, with a minimum sample size of 100 or 50 (Peng, Kuk, & Ingersoll, 2002).

Table 4.33

Omnibus Tests of Model Coefficients predicting Likelihood of Units Experiencing an Adverse Event with Predictors Unit Type, Dichotomous Teamwork Climate, and Safety Climate

		Chi-square	df	Sig.
Step 1	Step	.35	1	.554
	Block	.35	1	.554
	Model	14.26	6	.027

Table 4.34

Logistic Regression Results Predicting the Probability of a Unit Experiencing an Adverse Event

		B	S.E.	Wald	df	Sig.	Odds ratio Exp(B)	95% C.I. for EXP(B) Lower	Upper
Step 1[a]	Medical surgical			4.60	4	.331			
	Critical care	19.08	16962.57	.00	1	.999	193239644.51	.00	.
	Intermediate care	19.41	23154.78	.00	1	.999	269225040.20	.00	.
	OB gynecology	-3.45	1.93	3.19	1	.074	.03	.00	1.40
	Procedural	-2.56	1.32	3.77	1	.052	.08	.01	1.02
	High teamwork	3.28	2.35	1.96	1	.162	26.66	.27	2645.23
	Safety climate	3.20	5.39	.35	1	.553	24.61	.00	960129.05
	Constant	-.69	4.57	.02	1	.881	.50		

Summary

Aim one was to describe self-reported safety climate and teamwork climate among an interprofessional group of providers working on 32 hospital inpatient units, measured using the SAQ. The average unit size (number of beds per unit) was 28 beds. There were three intermediate care units, three OB/Gynecology units, five critical care units, 10 medical-surgical units, and 11 procedural areas in the sample of 32 units. The average teamwork and safety climate scores were 70.07% and 75.61%, respectively. Safety climate was normally distributed whereas teamwork climate was negatively skewed. Safety climate and teamwork climate were highly correlated with each another. Because teamwork climate was negatively skewed, it was recoded into a dichotomous variable (using a .60 or 60% agreement as a cut point) for subsequent analyses. Eight units or 25% of the sample reported less than good or low teamwork climate scores.

Aim two was to explore the nature of adverse events identified using a modified IHI GTT chart review methodology. Sixty-nine adverse events occurred per 1,000 patient days and 21 adverse events occurred per 100 admissions. About 20% of admissions experienced an adverse event. Fifty-three adverse events were associated with least severe harm category E, and 16 events were associated with harm category F, which represented temporary harm to the patient that required initial or prolonged hospitalization. The most frequently identified triggers were anti-emetic administration, diphenhydramine (Benadryl) administration, and transfusion or use of blood products. Three triggers had positive predictive values of 100%: healthcare-associated infections, injury/repair/removal of an organ, and over-sedation/hypotension.

Aim three was to examine to what extent unit-level safety climate and teamwork climate (and their potential interaction) predicted adverse events per unit detected via the modified IHI GTT chart review format. There was no statistically significant association between safety climate and frequency of adverse events. Likewise, there was no statistically significant association between teamwork climate and frequency of adverse events.

Because unit size was not correlated with any of the variables, unit type was explored as a unit characteristic control variable. Unit types were critical care, intermediate care, medical-

surgical, OB gynecology, and procedural units. According to the one-factor fixed ANOVA, there was no association between unit type and teamwork, and among unit types, there was only one statistically significant difference regarding adverse event frequency. OB gynecology units experienced a lower percentage of admissions with an adverse event compared to medical surgical units.

Teamwork and safety climate means did not statistically differ between medical surgical units and other types of units. Low and high teamwork units did not experience a statistically significant difference in mean adverse events. A two-way ANOVA was conducted to determine if combinations of unit type and teamwork strength differed significantly in terms of frequency of adverse events per 1,000 patient days. There were two unit types (medical surgical versus other unit type) and two categories of teamwork (low and high). There was a significant main effect of unit type on adverse events, but not for teamwork. Medical surgical units experienced significantly more adverse events per 1,000 patient days than other unit types. Teamwork strength (low versus high) did not have a main effect. There was not a statistically significant interaction effect between unit type and dichotomized teamwork climate.

In the hierarchical multiple regression model predicting frequency of adverse events per 1,000 patient days, the only statistically significant variable was unit type entered in model one. The rest of the variables, teamwork climate, safety climate, and the interaction between teamwork climate and safety climate were not statistically significant predictors of adverse events. Because 25% of the sample units did not experience an adverse event, frequency of adverse events was dichotomized for analysis with logistic regression. Unit type, dichotomized teamwork climate, and safety climate were not related to the likelihood that a unit would experience an adverse event.

Chapter 5: Discussion

This chapter is organized into the following sections. The first section summarizes the study. The second and third sections present findings in the context of the literature (organized by specific aim) and the hospital environment. Section four explains study limitations. The fifth section contains an interpretation of findings through an analysis of potential explanations for non-statistically significant results. The sixth, seventh and eighth sections present insights regarding the Global Trigger Tool, implications for theory, and conclusions. The final three sections include implications for practice, recommendations for further research, and concluding remarks.

Summary of the Study

The purpose of this study was to explore relationships between organizational (safety culture) and human factors (teamwork) with adverse events detected using a modified trigger-tool methodology. The major finding of this study was that there was no association between safety climate and teamwork climate with frequency of GTT-detected adverse events per nursing unit. Reasons for non-significant findings may be due to small variance between units, measurement error (random and systematic), sampling error, and non-response error. After considering possible explanations for non-significant findings, the most logical conclusion appears to be that the non-statistically significant findings are valid.

Findings in Context of Existing Literature

Aim One

The following section analyzes the findings related to the literature for aim one. Aim one was to describe self-reported safety climate and teamwork climate among an interprofessional group of providers working on 32 hospital inpatient units, measured using the SAQ.

Predictors: Teamwork climate. Teamwork climate was negatively skewed, with the mean 70.07%. Respondents, on average, responded more negatively to teamwork climate than

safety climate. However, this sample's mean was higher than the 64.3% mean reported from a national sample of over 1500 employees from 11 inpatient units in U.S. hospitals (Sexton et al., 2006). Although the teamwork climate score distribution "deviates from perfect symmetry," it is rare that sample data follow a perfect distribution (Lomax, 2007, p. 68). This sample's teamwork climate was higher than a national sample average, which provides perspective for interpretation.

Predictors: Safety climate. The mean safety climate was 75.61%, which is higher than the 60.5% mean from benchmarking data in a national sample of over 1500 employees from 11 inpatient units in U.S. hospitals (Sexton et al., 2006), as well as the 63% mean from 890 clinical areas in a descriptive study (Rose et al., 2006). The 75% mean is also greater than other units from outside the United States. A cross-sectional descriptive study of 54 U.S. and Swiss medical-surgical units found that the mean safety climate score was 66% (Schwendimann, Zimmermann, Küng, Ausserhofer, & Sexton, 2013). The 32 units in this convenience sample may not accurately represent the distribution of safety climate in all nursing units or departments.

Teamwork climate and unit type. Unit types with the strongest teamwork climate were critical care, intermediate care, and medical surgical, but there was very little variability between the top three unit types (77.4%, 77.2%, and 74.2%). In this sample, it appears that the teamwork climate domain of the SAQ was not able to discriminate differences in teamwork between units. Critical care units experienced the highest average teamwork climate of all unit types. Critical care unit teamwork was 77.7%, which is higher than the 72.2% average SAQ teamwork climate score in a study with 110 intensive care units from 61 hospitals (France et al., 2010). Many teamwork factors are being investigated and implemented in critical care, such as communication skills, interprofessional rounds, handoffs, and team training (Dietz, Pronovost, Mendez-Tellez, et al., 2014). Thus, it is not surprising that critical care units experienced the highest teamwork among unit types, because they were likely using strategies to improve teamwork. In addition, critical care units have more consistent providers (smaller number of intensivists) than medical surgical units with rotating coverage of internal medicine physicians.

Safety climate and unit type. Type of unit may also affect safety climate. Medical surgical units experienced the strongest safety climate scores in this study. These results differ

from a descriptive correlational study of 12 inpatient units that found critical care units had the strongest safety climate scores of varying unit types including critical care and general (medical surgical) units (Abdou & Saber, 2011). The average safety climate for critical care units was 77.1% in this study, which is slightly higher than a study with 110 intensive care units from 61 hospitals with average SAQ safety climate score of 74.2% (France et al., 2010). OB gynecology units experienced the lowest mean safety climate (66.9%) of all unit types, but this was higher than the 33.3 to 55.4% range experienced by an OB service located within a tertiary-level academic center (Pettker et al., 2009). It appears that the effect of unit type on safety climate differs across samples.

Aim Two

The following section analyzes the findings related to the literature for aim two. Aim two was to explore the nature of adverse events identified using a modified Institute for Healthcare Improvement Global Trigger Tool chart review methodology.

Adverse events. The Global Trigger Tool is not designed to compare adverse event rates between organizations due to varying skill among reviewers and other GTT aspects (Griffin & Resar, 2009). The GTT developers assume that this bias is relatively stable over time in a particular health care organization. Nonetheless, it may be helpful to use national data to determine if adverse events rates are in the general range of other healthcare organizations (Griffin & Resar, 2009). Based upon data from hundreds of organizations using the GTT, most organizations experience about 90 adverse events per 1,000 patient days and 40 adverse events per admissions (Resar, 2009). Typically, about 30% of hospital admissions experience an adverse event (Resar, 2009).

Sixty-nine adverse events occurred per 1,000 patient days in this sample of 32 nursing units/departments. A comparison of these results to more recent literature is warranted. This frequency is smaller than one large health care system (Garrett Jr et al., 2013), very similar to another system (Good, Saldana, Gilder, Nicewander, & Kennerly, 2011), and greater than two other healthcare systems outside the United States, in Sweden (Rutberg et al., 2014) and Korea

(Hwang et al., 2014). It appears that the adverse frequency is within the ballpark of expected results. See Table 5.1.

Table 5.1

Comparison of Adverse Event Frequency among Various Hospitals using the Institute for Healthcare Improvement Global Trigger Tool for Measuring Adverse Events

		All adverse events (AEs) included		
	Level of measurement	AEs/1,000 patient days	AEs/100 admissions	% of admissions with an AE
Adventist Health System (Garrett et al., 2013)	Hospital-level	85	38	26%
This sample	Unit-level	69.20	21.83	19.64%
Baylor Health Care System (Good et al., 2011)	Hospital-level	68.1	50.8 (per 100 encounters)	39.8%
Rutberg et al. (2014)	Hospital-level	33.2		20.5%
Hwang et al. (2013)	Hospital-level	12.38	7.79	7.2%

Note. Adapted from 'Large-scale deployment of the Global Trigger Tool across a large health care system,' by V.S. Good et al., 2011, *BMJ Quality and Safety, 20*, p. 29, and 'Developing and implementing a standardized process for Global Trigger Tool application across a large health care system,' by P.R. Garrett et al., 2013, *The Joint Commission Journal on Quality and Patient Safety, 39*(7), p. 294, and 'Characterizations of adverse events detected at a university hospital' by H. Rutberg et al., 2014, *BMJ Open, 4*(e004879), p. 1.

Over three-quarters (76.81%) of the adverse events were associated with the least severe harm category of E. Other GTT studies have also reported that the most frequently identified harm category was E (Garrett Jr et al., 2013; Good et al., 2011; Suarez et al., 2014). In contrast, other GTT studies identified harm category F most commonly (Kennerly et al., 2014; Landrigan et al., 2010; Rutberg et al., 2014) among all harm categories.

Adverse events by unit type. Unit types with the greatest frequency of adverse events per 1,000 patient days were medical surgical, procedural, and critical care unit types. A prospective study with over 292 wards (units) in 71 French hospitals reported 6.6 adverse events per 1,000 patient days in hospital wards (Michel, Quenon, Djihoud, Tricaud-Vialle, & de Sarasqueta, 2007), which is considerably less than the average of 125 adverse events per 1,000 patient days in medical surgical units within this study. The difference is likely due to the

methodological differences for adverse event detection. Michel et al. (2007) used external physician and nurse investigators to assess occurrence of adverse events through chart review screening followed by discussion with hospital staff, whereas in the present the study, reviewers used the GTT, which is a highly sensitive tool for detecting adverse events.

In general, there is little published information regarding adverse event frequency by unit type. A prospective observational study of one intensive and one coronary care unit established an adverse event rate of 80.5 adverse events per 1,000 patient days (Rothschild et al., 2005), which is considerably higher than the mean of 38.57 adverse events per 1,000 patient days in the five critical care units in this study. Reasons for the higher adverse events rate in the Rothschild et al. (2005) study may be due to their more broad definition of adverse event and thorough data collection methods with both direct continuous observation and voluntary/solicited reports.

One retrospective case series within a 715-bed university-based tertiary care center found that adverse events occurring in the radiology department, while rare, were primarily due to anaphylactic and drug reactions (Tindel, Darby, & Simmons, 2014). Thus, if a GTT reviewer were focusing on detecting adverse events related to the radiology department, the medication trigger, "Benadryl (Diphenhydramine)/ Anti-histamines)/ IV Corticosteroid use)" may be particularly useful as a clue that an anaphylactic reaction occurred due to a contrast agent.

GTT chart review methodology. Frequently identified triggers were anti-emetic administration, diphenhydramine (Benadryl) administration, transfusion or use of blood products, decrease of greater than 25% in hemoglobin or hematocrit, restraint use, and any procedure complication. It is expected that reviewers find many positive triggers, but fewer adverse events (Griffin & Resar, 2009). Six triggers had positive predictive values of greater than 75%: healthcare-associated infections (100%, *n* = 5), injury, repair, or removal of organ (100%, *n* = 1), over-sedation/ hypotension (100%, *n* = 1), any procedure complication (90%, *n* = 9), anti-emetic use (84%, *n* = 37), and any operative complication (75%, *n* = 3). In the literature, triggers with a greater tendency to detect true adverse events are health-care associated infections (Hwang et al., 2014; Kennerly et al., 2013), any procedure complication (Hwang et al., 2014), and any

operative complication (Kennerly et al., 2013). Triggers with high positive predictive values were consistent with literature findings.

Aim Three

The following section analyzes the findings related to the literature for aim three. Aim three was to examine to what extent unit-level safety climate and teamwork climate (and their potential interaction) predict adverse events per nursing unit as detected via the modified IHI GTT chart review format, controlling for unit characteristics.

Control variables: Unit characteristics (size). There was no statistically significant correlation between unit size and teamwork climate, safety climate or frequency of adverse events. Previous research did not find a relationship with unit size and medication administration errors, either (Hung, Lee, Tsai, Tseng, & Chang, 2013). On the other hand, research shows that smaller units may have stronger safety climate (Hughes, Chang, & Mark, 2012) and stronger teamwork (Kalisch, Russell, & Lee, 2013) compared to larger units. Previous research has also demonstrated that larger units may experience more negative patient safety indicators and infections than smaller units do (Rashid, 2014; Rivard, Elixhauser, Christiansen, Zhao, & Rosen, 2010).

Safety climate and teamwork climate. Safety climate and teamwork climate were highly correlated with each another, with a Pearson's *r* of .89. This correlation is considerably higher than the .72 correlation from benchmarking data in a national sample of over 1500 employees from 11 inpatient units in U.S. hospitals (Sexton et al., 2006), and the .80 correlation from 40 ICUs in U.S. hospitals (Speroff et al., 2010). Safety climate and teamwork may be highly correlated with one another because teamwork may represent a subculture of safety culture, as explained in a qualitative meta-analysis (Sammer, Lykens, Singh, Mains, & Lackan, 2010). In the current climate of increased patient complexity and treatment, healthcare organizations require strong teamwork to attain a safe culture (Sammer et al., 2010).

Safety climate with frequency of adverse events. The correlation between safety climate and frequency of adverse events was not significant. The non-significant relationship is consistent with level I evidence from a meta-analysis by Groves (2014), and five out of 17 studies

within DiCuccio's (2014) review. These results differ from previous research that demonstrated a relationship between positive safety culture and adverse event severity (Kline, Willness, & Ghali, 2008), fewer hospital acquired pressure ulcers (Brown & Wolosin, 2013), and decreased medication errors (Vogus & Sutcliffe, 2007; Zohar, Livne, Tenne-Gazit, Admi, & Donchin, 2007). The non-significant relationship also differs from a study that found strong safety culture was negatively related to hospital-associated urinary tract infections, falls, and catheter-associated infections (McFadden, Stock, & Gowen, 2015).

A recent study explored the association between safety climate (measured via the HSOPS) and GTT-identified adverse events in 272 Norwegian patient records. This cross-sectional study found the unit with the highest (best) self-assessed safety climate experienced more adverse events than a unit with lower self-assessed safety climate (Farup, 2015). These unexpected findings raise questions regarding the validity and reliability of safety climate instruments (Farup, 2015). A mixed methods study designed to explore the validity of the AHRQ patient safety indicators by interviewing Veteran's Administration hospital staff found similar results. Hospital staff working in both low and high performing hospitals (based on a composite PSI score) reported positive safety culture (Shin et al., 2014). Hospital staff perceptions of safety culture may not accurately reflect patient outcomes such as adverse events or patient safety indicators.

The mean safety climate was 75.61%, which is higher than means from other samples (Rose et al., 2006; Schwendimann et al., 2013; Sexton et al., 2006). It is possible that the sample had self-selection bias because people chose to complete the SAQ voluntarily, and respondents may possess different characteristics than non-responders (Kerlinger & Lee, 2000). However, because safety climate exhibited a normal distribution (whereas teamwork climate did not) it appears that the sample represented healthcare workers with both favorable and unfavorable perceptions regarding safety.

Teamwork climate with frequency of adverse events. The correlation between teamwork climate and frequency of adverse events was not significant. Previous research found a non-significant relationship between teamwork climate and morbidity and mortality (Davenport,

Henderson, Mosca, Khuri, & Mentzer Jr, 2007), and number of reported safety events (Malloy, 2012). On the other hand, other studies have demonstrated relationships between strong teamwork and lower rates of patient safety indicators (Mardon, Khanna, Sorra, Dyer, & Famolaro, 2010), adverse events (Deilkas & Hofoss, 2008), fewer falls (Brown & Wolosin, 2013; Sammer, 2009) and lower odds of developing a pressure ulcer (Taylor et al., 2012). Of these results, it is noteworthy that the Deilkas and Hofoss (2008) found a significant relationship between teamwork and GTT-identified adverse events, whereas this study did not find such a relationship. The Deilkas and Hofoss (2008) study was conducted in Norway that has a universal, public health care system, and it is possible that human factors like communication methods, in addition to teamwork, may differ in this system, and have a greater effect on adverse events.

Teamwork by unit type. There was not a statistically significant relationship between dichotomized teamwork climate and unit type. Negative skewness of teamwork led to its dichotomization into two groups comprised of low and high teamwork. The loss of information, going from a continuous variable to a dichotomous one, may result in a Type II error in bivariate analyses or the one-way ANOVA (Owen & Froman, 2005). In this analysis, a Type II error could mean overlooking a relationship between teamwork and a variable like unit type or frequency of adverse events that possibly occurs in the population (Owen & Froman, 2005).

Unit type and frequency of adverse events. OB gynecology units experienced statistically significantly fewer adverse events than medical surgical units did. Women cared for in labor and delivery and post-partum units are usually younger (less than 40 years old). Previous research indicates that patients who experience GTT-identified adverse events tend to be older and male compared to patients that do not experience an adverse event (Classen et al., 2011). Thus, the finding that OB/gynecology units experienced significantly less adverse events than medical surgical units that care for both genders and older patients with chronic disease is not surprising.

The GTT contains eight perinatal triggers specific to OB gynecology units, which may be effective for measuring adverse events for this unit type. Within an urban teaching hospital OB gynecology service, physicians voluntarily reported seven adverse events and 38 potential

adverse events (November, Chie, & Weingart, 2008). The GTT triggers may have also detected these same physician-reported adverse events. Adverse event examples reported by physicians that would also have been GTT-detected include over-sedation or hypotension, drop in Hemoglobin by 25%, hypoglycemia, and excessive blood loss. Thus, it appears that GTT triggers could be effective for calculating unit-level adverse event rates for OB/gynecology units, but further testing is necessary.

Findings in Context of Hospital Environment

Respondents, on average, reported teamwork climate more negatively than safety climate. A discussion with hospital leadership provided insight regarding potential reasons for low perceived teamwork. Some units with low reported teamwork were in the midst of leadership transitions. Another unit's personnel policies appeared unpredictable to staff, which caused perceived lack of control and a high stress work environment. Management commitment and work stress are examples of additional internal organizational factors that contribute to patient safety within the GRM.

Limitations

Instrumentation

Global Trigger Tool. Record reviewers collected patient outcome data from individual electronic patient records to determine the occurrence of adverse events. Health record data may have issues with validity and reliability due to variations in recording, clinical competence, situational factors, and data type (Aaronson & Burman, 1994). GTT reviewers' skill may have improved over time, although reviewers dispersed the 10 records for each unit over a few months so that frequency of adverse events did not differ by time reviewed. It is also possible that reviewer fatigue could make adverse event detection less accurate (Kerlinger & Lee, 2000). Reviewers limited chart reviews to a maximum of ten per day to control fatigue. Due to the 20-minute time limit for record reviews, it is likely that adverse events were missed. Some records with short lengths of stay only required five minutes to review. Although all units used electronic records, one specialty area used a different digitized record system, which may have influenced data type and ability to detect events with the GTT.

Reliability of the health record data and corresponding assessment of adverse events occurred though interrater reliability of GTT reviewers. The observed percentage agreement regarding presence of an adverse event or not was 93.9%, and the Cohen's Kappa score was 0.835, which may be interpreted as almost perfect agreement (Landis & Koch, 1977; Viera & Garrett, 2005). However, if practitioners did not document completely in the electronic record, then reviewers may not have detected all adverse events with the GTT.

Looking collectively at the adverse event data from all units combined, the most frequently identified adverse event was nausea. Nausea, as a gastrointestinal adverse event, may be related to the use of opioids for pain control, and opioid-related events have been identified as one of the top ten safety concerns for healthcare organizations (ECRI Institute, 2015). One could argue the definition of nausea as an adverse event based upon two or more doses of anti-emetic medication was too sensitive. Using this definition, there were 37 cases of nausea as an adverse event. The frequency decreases to 20 nausea adverse events if the definition is changed to three or more anti-emetic doses or 12 cases if the definition is four or more doses of anti-emetic medication. Clearly, it is important for the team to define GTT triggers and corresponding adverse events because this affects the adverse event rate.

Although the National Cancer Institute has developed terminology criteria for adverse events with severity ratings, including nausea (National Cancer Institute, 2015), number of antiemetic doses is not part of the nausea severity definition, so there is little guidance regarding what quantity of anti-emetic doses is acceptable for cancer treatment, or what medication quantity constitutes an adverse event. For post-operative nausea and vomiting, current guidelines recommend treatment based upon risk stratification, and prophylaxis of high risk patients with combination therapy (Gan et al., 2014). The GTT developers state that a determination of adverse event occurrence is facilitated by thinking from the viewpoint of the patient, meaning how would you feel if the event happened to you (Griffin & Resar, 2009). Patients with nausea expect relief with one dose of an anti-emetic medication, and they will likely not be happy if nausea persists. Thus, the definition of nausea requiring two or more doses of antiemetic medication appears reasonable.

Safety Attitudes Questionnaire. Responses to the SAQ or other safety climate surveys may vary by job type. Physicians have more positive perceptions of safety than nurses do (Singer et al., 2009; Speroff et al., 2010). In this study, if more physicians responded to the SAQ than other health care professionals on certain units, then it is likely that the safety climate positive percentage agreement would be higher on those units compared to other units. Because we do not have access to SAQ responses by job type, we do not know how job type may have affected the results, which is a study limitation.

Selection and Cross-Sectional Data

The selection of units did not occur randomly, and measurement of variables occurred only once. This convenience sample may possess characteristics that differ from other units not in the sample, and this could account for differences in adverse event frequency among unit types. Cross sectional studies, by definition, measure both the independent and dependent variables at the same time (Grimes & Schulz, 2002). Because the researcher identifies both the exposure and outcome concurrently, it is difficult to determine the temporal sequence (Grimes & Schulz, 2002). The cross-sectional design provides a snapshot (Stommel & Wills, 2004) of safety and teamwork climate in January 2013 only.

External Validity

A single hospital served as the data collection site for 32 units. The findings may not be generalizable to other hospitals of different size or non-Midwestern U.S. location.

Interpretation of Findings

Small Variance

Safety climate and teamwork climate demonstrated little variance (.024 and .036, respectively), which may explain the lack of statistically significant findings. Lack of variability made it difficult to detect statistically significant differences in safety climate and teamwork between units. Similarities across units may make one consider if safety and teamwork climates are an organizational construct, but safety culture surveys are designed to assess culture in individual units (Singla, Kitch, Weissman, & Campbell, 2006). There is debate among scholars regarding group level analyses of safety climate surveys, and whether they should be aggregated

at the unit, organization, managerial position, or professional background level (Pumar-Méndez, Attree, & Wakefield, 2014). Specific to the SAQ, experts note greater variability between work units within a hospital than between hospitals, and thus recommend culture assessments across all work units in a hospital (Pronovost & Sexton, 2005). Although units possessed similar safety and teamwork climates, these perceptions are probably an accurate representation of safety culture at the unit level.

The social desirability of projecting favorable self-assessments of safety and teamwork climates may explain the small variance and overall positive safety and teamwork perceptions. Safety and teamwork climates were highly correlated with one another (r = .889), which may be from sharing social desirability variance (Nunnally & Bernstein, 1994). A larger sample size with units from multiple hospitals may demonstrate greater variability, improving the ability to detect safety and teamwork climate variation between units.

Measurement Error

Random error. Measurement error in any of the variables may have contributed to non-statistically significant results. In classical test theory, random error is the difference between an individual's true score and observed score (Osterlind, 2010), and it affects the reliability of the score (Waltz, Strickland, & Lenz, 2010). Random error may alter a respondent's optimal performance, with potential causes being environmental distractions, motivation, or anxiety (Osterlind, 2010). Applied to this study, sources of random error may have been a noisy environment and lack of privacy for respondents completing the SAQ. Respondents may have been more or less motivated to respond depending on perceived importance of the SAQ, time of day (before or after working a shift), and fatigue. Some units may have encouraged SAQ completion as a group (after a staff meeting) whereas other units may have encouraged independent completion. This study used secondary SAQ data collected in 2013, and so the ability to control extraneous variance related to SAQ administration procedures was not possible.

Systematic error. Systematic error leads to consistent different responses among groups that are unrelated to the construct being measured (Osterlind, 2010). All members of the sample are predisposed to respond similarly (Osterlind, 2010). Sources of systematic error may

be the measurement tool or process, and do not fluctuate from one measurement situation to the next (Waltz et al., 2010).

SAQ items. Applied to this study, systematic error may be due to inclusion of items irrelevant to the concept being measured (Waltz et al., 2010). Some SAQ items seem as if they do not correspond well with their corresponding climate definitions. For example, the fourth item within the safety climate domain is, "I receive appropriate feedback about my performance" (Sexton et al., 2006, p. 7). This item does not seem to match well with the definition safety climate, "perceptions of a strong and proactive organizational commitment to safety" (Sexton et al., 2006, p.3). Feedback about performance does not seem to be a safety issue, but a human resources issue.

There are also similar concerns with items two and four not corresponding well to the teamwork climate domain. Item two within the teamwork climate domain is "In this work setting, it is difficult to speak up if I perceive a problem with patient care" (Sexton et al., 2006, p. 7). Speaking up about a patient concern seems to correspond better with safety climate rather than teamwork, which is "perceived quality of collaboration between personnel" (Sexton et al., 2006, p. 3). Item four within teamwork is "I have the support I need from others in this work setting to care for patients" (Sexton et al., 2006, p. 7). A respondent could interpret support in terms of staffing, resources (supplies for patient care), or managerial commitment, and not think of collaboration, which is what the SAQ developers intended to measure in the teamwork climate domain. Perceptions regarding ability to speak up and resources may not be relevant to teamwork climate, and this could serve as a potential source of systematic error.

Responses to each SAQ item are valuable, and one divergent item within a domain may misrepresent perceptions of safety or teamwork climate. Safety and teamwork climate domain scores are based upon the percentage of individuals whose responses across all of the items in the domain average out to be equal to or greater than four (Pascal Metrics, 2013). The safety climate domain is comprised of seven items and teamwork climate is comprised of six items. If a respondent misinterprets one item and records an inaccurate perception, the domain average may be affected. Although some SAQ items appear to not correspond with their domains, the

SAQ underwent confirmatory factor analysis with a satisfactory model fit, which demonstrates its construct validity (Sexton et al., 2006). Thus, incongruent SAQ items as a systematic error source are unlikely to explain the non-statistically significant results.

Adverse event measurement. Another source of systematic error in the study may have been the measurement instrument for adverse events. An alternative measurement instrument for adverse events could be the AHRQ patient safety indicators (PSIs). In 2013 (the same year of selected record reviews), examples of PSIs were pressure ulcers, mortality, retained surgical item, iatrogenic pneumothorax rate, central venous catheter-related blood stream infections, postoperative hip fracture rate (due to a fall in hospital), and transfusion reactions (Agency for Healthcare Research and Quality, 2013). The GTT triggers would detect these adverse events, as well as others like procedure complications, allergic medication reactions, and nausea, which the PSIs do not detect. In fact, the ability to detect statistically significant unit-level differences in safety and teamwork climates with PSIs may have been less likely compared to GTT-detected events because the GTT detects a greater variety of events. The GTT has excellent sensitivity and specificity for detecting adverse events (Classen et al., 2011) to discover statistically significant differences between units. Use of the GTT as an adverse event measurement tool was not the reason for the non-statistically significant findings.

GTT measurement process. GTT guidelines state that ten patient records should be sampled every two weeks, or 20 per month, and twelve data points are necessary to establish a baseline adverse event rate (Griffin & Resar, 2009). In this study, only ten records from each unit at one time point were reviewed. This may not accurately reflect a unit's adverse event rate because of the extremely small sample and cross-section of data at that single time point. Of course, this potential systematic error applied to all units.

The GTT developers did not design the GTT for unit-level analyses, and so concerns about validity and reliability are expected. For unit-level analyses, one may ask, "Are we measuring what we think we are measuring?" (Kerlinger & Lee, 2000, p. 666). The GTT may measure adverse events well at the hospital-level, but it may not accurately measure adverse events at the unit level. This unit-level review process requires replication 11 more times to

establish an accurate baseline rate of adverse events at a unit level in this sample. However, because GTT-identified adverse events are calculated using a standard denominator (per 1,000 days, per 100 admissions, percentage of admissions with an adverse event), this study demonstrates that it is possible to calculate a unit-level adverse event rate.

The GTT does not measure all adverse events, and it can miss harms due to diagnostic errors and omitted care (Pronovost & Wachter, 2014). The GTT does not detect preventable harm (Lipczak, Neckelmann, Steding-Jessen, Jakobsen, & Knudsen, 2011). Variations in certain events, such as pneumothorax secondary to central line placement, thoracentesis, or biopsy may occur more frequently in larger hospitals (Pronovost & Wachter, 2014) or certain units that perform many procedures. Hospitals that routinely screen all patients for deep vein thrombosis will have inflated harm rates compared to hospitals that do not perform such screenings because they will detect more deep vein thrombosis cases (Pronovost & Wachter, 2014). It is acknowledged that the GTT did not detect all adverse events in this sample, but the systematic error from non-detected adverse events was unlikely to influence the association between safety climate and teamwork climate with adverse events.

Scholars have expressed concerns regarding the reliability of the GTT for detecting adverse events due to wide variations in interrater reliability (Deilkas, 2013; Schildmeijer, Nilsson, Arestedt, & Perk, 2012; von Plessen, Kodal, & Anhoj, 2012). Reasons for variations are lack of training sessions and clearly defining harm/adverse events (Deilkas, 2013; Schildmeijer et al., 2012; von Plessen et al., 2012). Both reviewers completed the IHI GTT training process and participated in a series of IHI webinars to learn how to use the GTT with practice chart reviews. Other hospitals use multiple GTT teams comprised of multiple reviewers, which may contribute to poor interrater agreement. The use of only two record reviewers may have been advantageous for reliability. The webinars, adverse event definition list, and two-reviewer team size likely contributed to the very good interrater reliability among the reviewers in this study. Due to consistent GTT procedures, the lack of statistically significant findings between safety climate and teamwork climate with adverse events is unlikely to be from systematic error within the GTT measurement process for adverse events.

Sampling Error

A source of survey error is sampling error due to variability that occurs by chance because a sample rather than an entire population was surveyed (U.S. Federal Committee on Statistical Methodology, 2001). This sample of 32 nursing units from one hospital may not reflect the population of U.S. nursing units. There was no association between safety climate and teamwork climate with frequency of adverse events in these 32 units, but other samples may have different findings.

Nonresponse Error

A potential error source is nonresponse error due to the inability to obtain desired information from an eligible unit (U.S. Federal Committee on Statistical Methodology, 2001). Low response rates may have biased the safety and teamwork climate domain scores because persons that completed the survey may differ systematically from non-responders (Stommel & Wills, 2004). The effect of low response rate is a smaller sample size, resulting in increased variance, which may introduce bias in survey estimates (U.S. Federal Committee on Statistical Methodology, 2001).

Non-responder perceptions may be more positive than responder perceptions, and because they perceive that everything is fine, they may not have felt compelled to complete the SAQ. Conversely, SAQ responders may have had stronger, more negative perceptions than non-responders. In the case of teamwork, intimidated staff with challenging relationships may have been more likely to complete the survey than staff with strong partnerships, which may have negatively skewed an entire unit's score. In this sample, the 17 units with a high response rate (greater than 60%) reported an average teamwork climate of 68.69%, which was less than the average 71.6% for the 15 units with low response rate. According to an independent t-test, safety and teamwork climate means did not statistically differ between low response rate and high response rate units (t = -2.40, df =30, p = .812; t = .429, df = 30, p = .671). Low response rates were experienced among the five unit types in different proportions: four (80%) out of five critical care units, two (66.7%) out of three intermediate care units, two (20%) out of ten medical surgical units, two (66.7%) out of three OB gynecology units and five (45.5%) out of 11 procedural units

experienced low response rates. There was not a statistically significant association between low and high response rate and unit type (X^2 = 6.05, *df* = 4, *p* = .195).

The previous section reviewed potential reasons for lack of association between safety climate and teamwork climate with frequency of GTT-detected adverse events per unit including small variance between units, measurement error (random and systematic), sampling error, and non-response error. Small variance between units remains a viable reason for lack of statistically significant findings. Although unlikely, random error from environmental distractions may have contributed to non-statistically significant findings by altering SAQ respondent data. Systematic error is unlikely to have influenced the results because of the SAQ's established validity and reliability, and the consistency of GTT procedures with good interrater reliability. Sampling error remains a possible explanation for non-significant findings, and replication is recommended. Low response rates do not appear to have influenced safety and teamwork climate perceptions, but non-responder perceptions remain unknown. Taking these considerations into account, the most likely explanation for no association between safety climate and teamwork climate with frequency of GTT-detected adverse events per unit is that such a relationship does not exist in this sample.

Insights regarding the Global Trigger Tool

The GTT developers recommend that the two primary record reviewers have a clinical background, knowledge about the record system and the provision of care (Griffin & Resar, 2009). They state that the best reviewers are experienced nurses, but pharmacists and respiratory therapists are also valuable team members. In this study, the primary record reviewer was a former advanced practice nurse and the second reviewer was a pharmacist. The pharmacist was instrumental for her knowledge regarding pharmacotherapy, medication adverse events, and patient safety. For the nurse reviewer, extensive nursing experience was helpful to ascertain usual care processes in critical care, surgery, and procedural areas, which enhanced the ability to recognize surgical and procedural adverse events. Regardless of clinical specialty, reviewers need knowledge about disease to differentiate patient care due to underlying illness or disease, or if treatment was related to health care.

Reviewers need access to hospital content experts. For example, in this study, there was one case where the reviewer contacted a wound ostomy nurse for assistance with differentiating stage I versus Stage II pressure ulcer due to unclear documentation in the record. It is also recommended to have access to an infection control expert to be up-to-date with definitions regarding hospital-associated conditions from the Centers for Disease Control and other agencies. Hospital content experts working in specialized units such as labor and delivery are also useful. Reviewers without specialized clinical experience may need to ask experts for their opinions of narrative notes to determine if clinical deterioration is related to care or is it truly progression of disease.

Baylor and Adventist Health Systems use a single person reviewer for conducting GTT reviews. Some adverse events were straightforward and easy to detect, while others were not. At the end of data collection, both reviewers examined adverse event cases involving a judgement, and met to discuss these cases. For example, clinical judgment was required to determine if a rash was due to medications, or if a treatment was related to a procedure complication rather than the health condition. The opportunity to discuss difficult cases with a second reviewer was beneficial to ensure the validity of adverse events.

Implications for Theory

Organizational Factors

The GRM suggests that contributing factors like organizational and human factors may provide a defense against safety incidents. In this study, organizational factors were operationalized through safety climate. However, numerous other organizational factors, not measured in this study, may influence patient outcomes. A common measure of unit-level workload in nursing is the nurse-patient ratio (Carayon & Gurses, 2008), and staffing levels are an internal organizational factor that may affect patient safety. A 2007 systematic review found that higher nurse staffing (increasing one nurse full-time equivalent) was associated with decreased odds of patient adverse events (Kane, Shamliyan, Mueller, Duval, & Wilt, 2007). More recently, a multicenter longitudinal study found that there was increased patient mortality in ICUs when the patient-to-nurse ratio exceeded 2.5 (Neuraz et al., 2015). In contrast to these results, a

2012 literature review did not find a statistically significant association between nurse staffing and morbidity and mortality (McGahan, Kucharski, & Coyer, 2012).

The benefits of increased nurse staffing may differ across unit types. For example, increasing one nurse full-time equivalent (FTE) per patient day reduced 16% of hospital-related mortality in surgical patients, but this percentage dropped to 5.6% for medical patients (Kane et al., 2007). A 2014 literature review concluded that there is limited evidence that better staffing improves patient outcomes in medical surgical units (Pannick, Beveridge, Wachter, & Sevdalis, 2014). The impact of RN skill mix (the proportion of hours of care provided by nurse) per unit type may also differentially affect patient outcomes. For example, RN skill mix was associated with reduced failure to rescue in general medical surgical units whereas in critical care it was associated with fewer cases of sepsis (Blegen, Goode, Spetz, Vaughn, & Park, 2011).

Other organizational factors that were not measured in this study include job satisfaction, work stress, management commitment, resources, finance priorities, training programs, polices/procedures, environmental conditions, noise, distractions, design, staff qualifications, social norms, employee participation, and the ethical environment. The measurement of additional variables would have required a much larger sample size, which was not possible to obtain due to limited resources in this small pilot study. Experts state that definitions are needed in the Generic Reference Model (The World Alliance for Patient Safety Drafting Group et al., 2009), and these unmeasured variables provide opportunities for future research.

Human Factors

Similar research opportunities exist for all of the types of GRM contributing factors to patient safety. Specific to human factors, this study did not measure situational awareness, communication, behavior, or performance. Of these human factors, situational awareness, the ability of nurses to switch attention between varying information sources, is an area with limited research (Sitterding, Broome, Everett, & Ebright, 2012). High cognitive workload may contribute to decision-making errors from reduced attention with errors, slips, lapses, or mistakes as the result (Carayon & Gurses, 2008). Nurses' situational awareness paired with timely and appropriate responses may result in reduced patient harm (Sitterding et al., 2012).

Communication is another human factor (in addition to teamwork) that contributes to patient safety, and it is one of the top five contributing factors to safety incidents, according to a 2012 systematic review (Lawton et al., 2012). A time motion study with 36 medical surgical units within 17 health systems found that about one-fifth of nursing practice time is spent on care coordination, communicating information about the patient (Hendrich, Chow, Skierczynski, & Lu, 2008). Increased efficiency of information flow and status updates for medical surgical nurses could decrease wasted time and the risk of errors (Hendrich et al., 2008).

Mediating or Moderating Relationships

Another possibility to consider is mediating or moderating relationships among contributing factors and adverse events within the GRM. Adverse events may be more or less likely in one set of circumstances than another (Hayes, 2013). Failure of defenses for patient safety may occur through a mediator, a mechanism through which contributing factors influence adverse events. Voepel-Lewis et al. (2012) demonstrated that the association between nurse staffing and adverse events depends upon surveillance level. A moderator variable (or interaction) influences the magnitude of the causal effect of contributing factors on adverse events (Hayes, 2013). Previous research found a statistically significant relationship between increasing nurse staffing and adverse events (measured via ratio of actual to expected deaths) for hospitals located in environments with managed care penetration greater than 7.5% (Mark, Harless, & McCue, 2005). The inclusion of additional GRM contributing factor variables, and explanation of conditions under which the relationships occur, will contribute to successful theory development (Mark, Hughes, & Jones, 2004) and greater understanding of patient safety.

Conclusions

Safety climate and teamwork did not have a statistically significant effect on frequency of GTT-identified adverse events in this sample of 32 units. However, the study provides preliminary evidence that researchers may use the GTT to detect unit-level adverse events. Unit-level outcomes have special importance for nursing because the National Database of Nursing Quality Indicators® records nursing quality indicators on a unit level, such as pressure ulcer prevalence, restraint prevalence, and hospital-associated infections (Montalvo, 2007). The GTT

is able to detect differences in adverse event rates among unit types, and unit type explains approximately 30% of the variance in adverse event frequency.

Implications for Practice

This study demonstrates that medical surgical units experience a statistically significantly greater frequency of adverse events compared to other unit types. A cross-sectional survey of over 30,000 European medical surgical nurses found that nurses ration patient care due to high workloads, and they are concerned about the quality of care deteriorating (Aiken, Sloane, Bruyneel, Van den Heede, & Sermeus, 2013). Medical surgical nurses' delayed ability to detect clinical deterioration may contribute to poor patient outcomes (De Meester, Van Bogaert, Clarke, & Bossaert, 2013). Because medical surgical nurses care for as many as six to seven patients, they may not be able to detect deterioration due to heavy workload. In California, there is a minimum nurse staffing requirement where medical surgical nurses care for two fewer patients per shift (5 patients per nurse) compared to New Jersey and Pennsylvania, two states without a minimum staffing requirement (Aiken et al., 2010). Aiken et al. (2010) demonstrated that a lower number of patients per nurse was associated with significantly lower mortality. Improved nurse staffing and decreased workload may decrease frequency of adverse events in medical surgical units.

Dr. Melnyk encourages Ph.D. students to focus on the "so what" factors within their research programs, addressing the potential impact on quality of care, costs and patient outcomes (Melnyk, 2014). Applied to this study, one can estimate the cost of adverse events. The GTT measures "unintended consequences of medical care, whether preventable or not" (Griffin & Resar, 2009, p. 5). In this study, 37 of the 69 adverse events detected were nausea, which one could contend is a preventable adverse event. Experts recommend prophylactic treatment of post-operative nausea and vomiting in high-risk patients such as females, non-smokers, and those with a history of motion sickness, etc. (Gan et al., 2014). A 2014 European study estimated the economic burden of each episode of nausea and vomiting to be €31 ± 22, based upon staff wages and cost of materials (medications, supplies) to treat nausea (Eberhart et al., 2014). Using a 2014 United States dollar and Euro exchange rate history, this is equivalent to

$37.50 ± 26.62 (X-Rates, 2015). These 37 nausea adverse events, which were likely preventable, cost this hospital $1,387.50 ± 984.94 in January 2013. Multiply the cost times 12 (months), the hospital may have spent over $16,000 in one year for treating nausea, that was likely preventable. This does not include unmeasured costs like nursing time directed away from monitoring and assisting other patients, the cost of nausea on patient satisfaction, and the potential impact on lengthened hospital length of stay and reduced quality of life. Hospitals will not discover the economic burden of nausea without using the GTT because staff do not report it via voluntary event reporting and the AHRQ PSIs do not include it. Nausea is a silent and potentially preventable adverse event with considerable economic and patient burden.

The prevention of nausea may be a prime opportunity for nurses to influence the quality and safety of healthcare. For example, interventions such as repositioning (Fathi, Nikbakht Nasrabadi, & Valiee, 2014), hydration (Gan et al., 2014), ginger (Montazeri et al., 2013; Saberi, 2014), acupressure (Saberi, 2014), or aromatherapy (Hodge, McCarthy, & Pierce, 2014) may provide some relief for nausea. In addition, there may be a need to consider alternative anti-emetic medications to treat nausea, such as aprepitant (Emend) or palonosetron (Aloxi), which have demonstrated greater efficacy than ondansetron (Zofran) for decreasing emesis and post-operative nausea and vomiting (Gan et al., 2014).

Thirty-two additional adverse events occurred in addition to nausea. An in-hospital potentially preventable complication (adverse event) cost Maryland hospitals an average of $2,326 (per patient) in 2008 (Fuller, McCullough, Bao, & Averill, 2009). If one multiplies 32 x $2,326, this hospital may have incurred over $74,000 in expenses related to non-nausea adverse events in one month. Admittedly, some of these adverse events may not have been preventable, and some adverse events may have cost less than $2,326 to treat. On the other hand, the 16 category F adverse events that required initial or prolonged hospitalization may have cost much more than $2,326 per patient to treat. The “so what” factor of this study is that the GTT identifies adverse events not detected via other methods, and these adverse events cost hospitals financially, affecting quality of care and patient outcomes.

Recommendations for Further Research

Unmeasured variables within the GRM provide opportunities for future research. This study operationalized one (internal) organizational factor and one human factor, but it did not account for other GRM contributing factors like external environmental factors, “subject of incident factors,” and “drugs, equipment, documentation” (Runciman et al., 2006). Measurement of these additional factors may be necessary to understand circumstances that lead to patient harm.

Because the 32 units originated from one hospital, generalizability is limited, and replication with units from multiple hospitals is recommended. The GTT lacks triggers for inpatient psychiatric and rehabilitation patients, and trigger development may assist with detecting adverse events in these specialized clinical areas. Trigger detection via automated information technology already exists (Li et al., 2014), especially for clear-cut triggers such as medication and laboratory values. However, detection of harm due to health care from other triggers is less concrete and requires clinical expertise and judgement.

Automated data collection techniques will enhance the utility of the GTT. An automated report could detect triggers, and then a review team would review only health records with positive triggers (Griffin & Resar, 2009). The information used to make a judgement regarding patient harm is often located in the narrative textual (progress or nursing) notes of the electronic health record. In the future, completely automated GTT record reviews may be possible with advancements in natural language processing. Natural language processing uses automated methods to represent textual information into a computable format for analysis (Friedman & Elhadad, 2014). Due to the 20-minute time limit for record reviews, it is likely that the GTT misses adverse events, but automated reviews using natural language processing may overcome this by prioritizing record reviews beginning with triggers with high positive predictive values for detecting adverse events, such as healthcare-associated infections or anti-emetic administration. There are concerns with accuracy of automated review systems, with one study demonstrating that only one-third of all adverse events were detected using data warehouse mining compared to the standard GTT review process (O'Leary et al., 2013). Future research is needed to enhance the sensitivity and specificity of GTT record reviews using automated methods and big data

techniques. In addition, the ability to detect triggers in real time, and prevent adverse event occurrence is an exciting area open to research exploration.

Concluding Remarks

Adverse events cause significant patient morbidity, mortality, and increased health care costs. Due to alarming adverse event rates, hospitals promote safety culture and teamwork as defenses to reduce risk of patient harm, but this study demonstrated no statistically significant association between unit-level safety and teamwork climates with adverse event frequency. Unit type may be an important consideration for understanding safety incidents, as it predicted about 30% of the variance in adverse events. Medical-surgical units experience more adverse events than other unit types, which may be due to numerous contributing factors. Researchers may consider using the GTT to detect unit-level adverse events not detected via other methods. Future research and advancements in health information technology will contribute to GTT automation and might detect adverse events in real time to improve patient safety.

References

Aaronson, L. S., & Burman, M. E. (1994). Use of health records in research: Reliability and validity issues. *Research in Nursing and Health, 17*, 67-73.

Abdou, H. A., & Saber, K. M. (2011). A baseline assessment of patient safety culture among nurses at student university hospital. *World Journal of Medical Sciences, 6*(1), 17-26.

Adams-Pizarro, I., Walker, Z., Robinson, J., Kelly, S., & Toth, M. (2008). Using the AHRQ Hospital Survey on Patient Safety Culture as an intervention tool for regional clinical improvement collaboratives. In K. Henriksen, J. B. Battles, & M. A. Keyes (Eds.), *Advances in patient safety: New directions and alternative approaches* (Vol. 2: Culture and Redesign). Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK43728/

Adler, L., Denham, C. R., McKeever, M., Purinton, R., Guilloteau, F., Moorhead, J. D., & Resar, R. (2008). Global Trigger Tool: Implementation basics. *Journal of Patient Safety, 4*(4), 245-249.

Agency for Healthcare Research and Quality. (2005). *Medical teamwork and patient safety: The evidence-based relation*. Retrieved from http://archive.ahrq.gov/research/findings/final-reports/medteam/ medteamwork.pdf

Agency for Healthcare Research and Quality. (2008). *TeamSTEPPS™: Teamwork attitudes questionnaire manual*. Retrieved from http://teamstepps.ahrq.gov/taq_index.htm

Agency for Healthcare Research and Quality. (2010). *Chapter 7: The importance of theory: Assessing the evidence for context-sensitive effectiveness and safety* (11-0006-EF). Retrieved from http://www.ahrq.gov/ research/findings/final-reports/contextsensitive/ context7.html

Agency for Healthcare Research and Quality. (2012a). *Hospital survey on patient safety culture: 2012 user comparative database report*. Rockville, MD: Author.

Agency for Healthcare Research and Quality. (2012b). *Patient safety primers: Human factors engineering*. Retrieved from http://psnet.ahrq.gov/primer.aspx?primerID=20

Agency for Healthcare Research and Quality. (2013). *Patient Safety Indicators technical specifications- version 4.5 May 2013*. Retrieved from http://www.qualityindicators .ahrq.gov/Archive/IQI_TechSpec_V45.aspx

Agency for Healthcare Research and Quality. (2014). *Hospital survey on patient safety culture: 2014 user comparative database report: Table 2-2. Response rate statistics, 2014 database hospitals*. Retrievedfrom http://www.ahrq.gov/professionals/quality-patient-safety/patientsafetyculture/hospital/2014/hosp14table2-2.html

Agency for Healthcare Research and Quality. (2015a). *Frequently asked questions: Surveys on patient safety culture*. Retrieved from http://www.ahrq.gov/professionals/quality-patient-safety/patientsafetyculture/pscfaq.html#Q13

Agency for Healthcare Research and Quality. (2015b). *Patient safety primers: Systems approach*. Retrieved from http://psnet.ahrq.gov/primer.aspx?primerID=21

Agency for Healthcare Research and Quality. (2015c). *Surveys on patient safety culture*. Retrieved from http://www.ahrq.gov/professionals/quality-patient-safety/patientsafetyculture/index.html

Agency for Healthcare Research and Quality. (n.d.). *AHRQ PSNet glossary: Adverse drug events (ADE)*. Retrieved August 12, 2013, from http://psnet.ahrq.gov/popup_glossary.aspx?name=adversedrugevent

Aiken, L. H., Clarke, S. P., Sloane, D. M., Sochalski, J., & Silber, J. H. (2002). *Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction*. *Journal of the American Medical Association, 288*(16), 1987-1993. doi:10.1001/jama.288.16.1987

Aiken, L. H., Sermeus, W., VandenHeede, K., Sloane, D. M., Busse, R., McKee, M., . . . Kutney-Lee, A. (2012). Patient safety, satisfaction, and quality of hospital care: Cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. *BMJ, 344*, e1717. doi:10.1136/bmj.e1717

Aiken, L. H., Sloane, D. M., Bruyneel, L., Van den Heede, K., & Sermeus, W. (2013). Nurses' reports of working conditions and hospital quality of care in 12 countries in Europe. *International Journal Of Nursing Studies, 50*(2), 143-153. doi:http://dx.doi.org/10.1016/j.ijnurstu.2012.11.009

Aiken, L. H., Sloane, D., Cimiotti, J. P., Clarke, S. P., Flynn, L., Seago, J. A., . . . Smith, H. L. (2010). Implications of the California nurse staffing mandate for other states. *Health Services Research, 45*(4), 904-921. doi:10.1111/j.1475-6773.2010.01114.x

American Association for the Advancement of Medical Instrumentation. (2013). Joint Commission approves 2014 alarms goal. *AAMI News, 48*(8). Retrieved from http://www.aami.org/publications/AAMINews/Aug2013/Joint_Commission_2014_Alarms_Goal.html

American Society of Hospital Pharmacists. (1993). ASHP guidelines on preventing medication errors in hospitals. *American Journal of Hospital Pharmacists, 50*, 305-314.

Aranaz-Andrés, J. M., Limón, R., Mira, J. J., Aibar, C., Gea, M. T., & Agra, Y. (2011). What makes hospitalized patients more vulnerable and increases their risk of experiencing an adverse event? *International Journal for Quality in Health Care, 23*(6), 705-712. doi:10.1093/ intqhc/mzr059

Bae, J., & Encinosa, W. E. (2011). Health information technology and its effects on hospital costs, outcomes, and patient safety. *Inquiry, 48*(4), 288-303.

Baker, G. R. (2004). Commentary: Harvard Medical Practice Study. *Quality and Safety in Health Care, 13*(2), 145-152.

Barach, P. (2008). Strategies to reduce patient harm: Understanding the role of design and the built environment. *Studies In Health Technology And Informatics, 132*, 14-22.

Basch, E. (2014). New frontiers in patient-reported outcomes: Adverse event reporting, comparative effectiveness, and quality assessment. *Annual Review of Medicine, 65*(1), 307-317. doi:10.1146/annurev-med-010713-141500

Bates, D. W., O'Neil, A. C., Petersen, L. A., Lee, T. H., & Brennan, T. A. (1995). Evaluation of screening criteria for adverse events in medical patients. *Medical Care, 33*(5), 452-462. doi:10.2307/3766540

Beckett, R. D., Sheehan, A. H., & Reddan, J. G. (2012). Factors associated with reported preventable adverse drug events: A retrospective, case-control study. *Annals of Pharmacotherapy, 46*(5), 634-641. doi:10.1345/aph.1Q785

Bjertnaes, O., Deilkas, E. T., Skudal, K. E., Iversen, H. H., & Bjerkan, A. M. (2015). The association between patient-reported incidents in hospitals and estiamted rates of patient harm. *International Journal for Quality in Health Care, 27*(1), 26-30. doi:10.1093/intqhc/mzu087

Blais, R., Derson, B., Bartlett, G., & Tamblyn, R. (2008). Can we use incident reports to detect hospital adverse events? *Journal of Patient Safety, 4*(1), 9-12.

Bleakley, A., Boyden, J., Hobbs, A., Walsh, L., & Allard, J. (2006). Improving teamwork climate in operating theatres: The shift from multiprofessionalism to interprofessionalism. *Journal of Interprofessional Care, 20*(5), 461-470. doi:10.1080/13561820600921915

Blegen, M. A., Goode, C. J., Spetz, J., Vaughn, T., & Park, S. H. (2011). Nurse staffing effects on patient outcomes: Safety-net and non-safety-net hospitals. *Medical Care, 49*(4), 406-414.

Boin, A., & Schulman, P. (2008). Assessing NASA's safety culture: The limits and possibilities of high-reliability theory. *Public Administration Review, 68*(6), 1050-1062.

Brennan, T. A., Leape, L. L., Laird, N. M., Hebert, L., Localio, A. R., Lawthers, A. G., . . . Hiatt, H. H. (1991). Incidence of adverse events and negligence in hospitalized patients: Results of the Harvard Medical Practice Study I. *New England Journal of Medicine, 324*(6), 370-376. doi:doi:10.1056/NEJM199102073240604

Brewer, B. B., & Verran, J. A. (2013). Measuring nursing unit environments with four composite measures. *Nursing Economic$, 31*(5), 241-249.

Brooks Carthon, J. M., Kutney-Lee, A., Jarrín, O., Sloane, D., & Aiken, L. H. (2012). Nurse staffing and postsurgical outcomes in black adults. *Journal of the American Geriatrics Society, 60*(6), 1078-1084. doi:10.1111/j.1532-5415.2012.03990.x

Brown, D. S., & Wolosin, R. (2013). Safety culture relationships with hospital nursing sensitive metrics. *Journal for Healthcare Quality, 35*(4), 61-74. doi:10.1111/jhq.12016

Bruner, D. W., Hanisch, L. J., Reeve, B. B., Trotti, A. M., Schrag, D., Sit, L., . . . Basch, E. M. (2011). Stakeholder perspectives on implementing the National Cancer Institute's patient-reported outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). *Translational Behavioral Medicine, 1*(1), 110-122. doi:10.1007/s13142-011-0025-3

Burger, J. (2013). Communication. In P. A. Potter, A. G. Perry, P. A. Stockert, & A. M. Hall (Eds.), *Fundamentals of nursing* (8th ed., pp. 309-327). St. Louis, MO: Elsevier Mosby.

Carayon, P., & Gurses, A. P. (2008). Nursing workload and patient safety: A human factors engineering perspective. In R. G. Hughes (Ed.), Patient safety and quality: An evidence-based handbook for nurses (vol. 2, pp. 203-216). Rockville, MD: Agency for Healthcare Research and Quality.

Carayon, P., Wetterneck, T. B., Rivera-Rodriguez, A. J., Hundt, A. S., Hoonakker, P., Holden, R. J., & Gurses, A. P. (2013). Human factors systems approach to healthcare quality and patient safety. *Applied Ergonomics, 45*(1), 14-25. doi:10.1016/j.apergo.2013.04.023

Carayon, P., Xie, A., & Kianfar, S. (2014). Human factors and ergonomics as a patient safety practice. *BMJ Quality & Safety, 23*, 196-205. doi:10.1136/bmjqs-2013-001812

Carnevali, L., Krug, B., Amant, F., Van Pee, D., Gérard, V., de Béthune, X., & Spinewine, A. (2013). Performance of the adverse drug event trigger tool and the Global Trigger Tool for identifying adverse drug events: Experience in a Belgian hospital. *Annals of Pharmacotherapy, 47*(11), 1414-1419. doi:10.1177/1060028013500939

Cheng, C. M. (2011). Hospital systems for the detection and prevention of adverse drug events. *Clinical Pharmacology and Therapeutics, 89*(6), 779-781. doi:10.1038/clpt.2010.356

Christiaans-Dingelhoff, I., Smits, M., Zwaan, L., Lubberding, S., van der Wal, G., & Wagner, C. (2011). To what extent are adverse events found in patient records reported by patients and healthcare professionals via complaints, claims and incident reports? *BMC Health Services Research, 11*(49), 1-10.

Chung, S. K., Ahn, M. J., Yoo, J. Y., Choi, M., Hyang, N., Woo, S. R., . . . Oh, E. G. (2011). Implementation of best practice for chemotherapy-induced nausea and vomiting in an acute care setting. *International Journal of Evidence-Based Healthcare, 9*(1), 32-38. doi:10.1111/j.1744-1609.2010.00198.x

Classen, D. C., Lloyd, R. C., Provost, L., Griffin, F., & Resar, R. K. (2008). Development and evaluation of the Institute for Healthcare Improvement Global Trigger Tool. *Journal of Patient Safety, 4*(3), 169-177.

Classen, D. C., Rescar, R., Griffin, F., Federico, F., Frankel, T., Kimmel, N., . . . James, B. C. (2011). “Global Trigger Tool” shows that adverse events in hospitals may be ten times greater than previously measured. *Health Affairs, 30*(4), 581-588.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). New York, NY: Routledge.

Colla, J. B., Bracken, A. C., Kinney, L. M., & Weeks, W. B. (2005). Measuring patient safety climate: A review of surveys. *Quality and Safety in Health Care, 14*, 364-366.

Collins, S. A., Cato, K., Albers, D., Scott, K., Stetson, P. D., Bakken, S., & Vawdrey, D. K. (2013). Relationship between nursing documentation and patients’ mortality. *American Journal of Critical Care, 22*(4), 306-313. doi:10.4037/ajcc2013426

Croskerry, P. (2013). From mindless to mindful practice: Cognitive bias and clinical decision making. *New England Journal of Medicine, 368*(26), 2445-2448. doi:10.1056/ NEJMp1303712

Culler, S. D., Jevsevar, D. S., Shea, K. G., Wright, K. K., & Simon, A. W. (2015). The incremental hospital cost and length-of-stay associated with treating adverse events among Medicare beneficiaries undergoing TKA. *Journal of Arthroplasty, 30*(1), 19-25. doi:http://dx.doi.org/ 10.1016/j.arth.2014.08.023

Davenport, D. L., Henderson, W. G., Mosca, C. L., Khuri, S. F., & Mentzer Jr, R. M. (2007). Risk-adjusted morbidity in teaching hospitals correlates with reported levels of communication and collaboration on surgical teams but not with scale Measures of teamwork climate, safety climate, or working conditions. *Journal of the American College of Surgeons, 205*(6), 778-784. doi:http://dx.doi.org/10.1016/j.jamcollsurg.2007.07.039

De Meester, K., Van Bogaert, P., Clarke, S. P., & Bossaert, L. (2013). In-hospital mortality after serious adverse events on medical and surgical nursing units: a mixed methods study. *Journal of Clinical Nursing, 22*(15-16), 2308-2317. doi:10.1111/j.1365-2702.2012 .04154.x

de Wet, C., & Bowie, P. (2009). The preliminary development and testing of a global trigger tool to detect error and patient harm in primary-care records. *Postgraduate Medical Journal, 85*(1002), 176-180. doi:10.1136/pgmj.2008.075788

DeFeijter, J. M., DeGrave, W. S., Muijtjens, A. M., Scherpbier, A. J. J. A., & Koopmans, R. P. (2012). A comprehensive overview of medical error in hospitals using incident-reporting systems, patient complaints and chart review of inpatient deaths. *PLoS One, 7*(2), e31125. doi:10.1371/journal.pone.0031125.

Deilkas, E. T. (2013). Imprecision regarding the Global Trigger Tool. *BMJ Quality & Safety, 22*, 271. doi:10.1136/bmjqs-2012-001584

Deilkas, E. T., & Hofoss, D. (2008). Psychometric properties of the Norwegian version of the Safety Attitudes Questionnaire (SAQ), generic version (short form 2006). *BMC Health Services Research, 8*(191), 1-10. doi:10.1186/1472-6963-8-191

Deilkas, E. T., & Hofoss, D. (2010). Patient safety culture lives in departments and wards: Multilevel partitioning of variance in patient safety culture. *BMC Health Services Research, 10*(85). doi:10.1186/1472-6963-10-85

Dekker, S. (2006). *The field guide to understanding human error*. Burlington, VT: Ashgate Publishing.

Dekker, S. (2011). *Patient safety: A human factors approach*. Boca Raton, FL: CRC Press.

DeMeester, K., Verspuy, M., Monsieurs, K. G., & VanBogaert, P. (2013). SBAR improves nurse–physician communication and reduces unexpected death: A pre and post intervention study. *Resuscitation, 84*(9), 1192-1196. doi:http://dx.doi.org/10.1016/j.resuscitation .2013.03.016

Devers, K. J., Pham, H. H., & Liu, G. (2004). What is driving hospitals' patient-safety efforts? *Health Affairs, 23*(2), 103-115. doi:10.1377/hlthaff.23.2.103

Devriendt, E., Van den Heede, K., Coussement, J., Dejaeger, E., Surmont, K., Heylen, D., . . . Milisen, K. (2012). Content validity and internal consistency of the Dutch translation of the Safety Attitudes Questionnaire: An observational study. *International Journal Of Nursing Studies, 49*(3), 327-337.

deVries, E. N., Ramrattan, M. A., Smorenburg, S. M., Gouma, D. J., & Boermeester, M. A. (2008). The incidence and nature of in-hospital adverse events: A systematic review. *Quality and Safety in Health Care, 17*, 216-223. doi:10.1136/qshc.2007.023622

DiCuccio, M. H. (2014). The relationship between patient safety culture and patient safety outcomes: A systematic review. *Journal of Patient Safety*. doi:10.1097/PTS .0000000000000058. Advance online publication. Retrieved from http://journals.lww.com/journalpatientsafety/Abstract/publishahead/The_Relationship_Between_Patient_Safety_Culture.99779.aspx

Dietz, A. S., Pronovost, P. J., Benson, K. N., Mendez-Tellez, P. A., Dwyer, C., Wyskiel, R., & Rosen, M. A. (2014). A systematic review of behavioural marker systems in healthcare: what do we know about their attributes, validity and application? *BMJ Quality & Safety, 23*(12), 1031-1039. doi:10.1136/bmjqs-2013-002457

Dietz, A. S., Pronovost, P. J., Mendez-Tellez, P. A., Wyskiel, R., Marsteller, J. A., Thompson, D. A., & Rosen, M. A. (2014). A systematic review of teamwork in the intensive care unit: What do we know about teamwork, team tasks, and improvement strategies? *Journal of Critical Care, 29*(6), 908-914. doi:10.1016/j.jcrc.2014.05.025

Doupi, P. (2012). Using EHR data for monitoring and promoting patient safety: Reviewing the evidence on trigger tools. *Studies In Health Technology And Informatics, 180*, 786-790.

Doupi, P., Svaar, H., Bjørn, B., Deilkås, E., Nylén, U., & Rutberg, H. (2015). Use of the Global Trigger Tool in patient safety improvement efforts: Nordic experiences. *Cognition, Technology & Work, 17*(1), 45-54. doi:10.1007/s10111-014-0302-2

Dowding, D. W., Turley, M., & Garrido, T. (2012). The impact of an electronic health record on nurse sensitive patient outcomes: An interrupted time series analysis. *Journal of the American Medical Informatics Association, 19*(4), 615-620. doi:10.1136/amiajnl-2011-000504

Duckers, M., Faber, M., Cruijsberg, J., Grol, R., Schoonhoven, L., & Wensing, M. (2009). Safety and risk management interventions in hospitals: A systematic review of the literature. *Medical Care Research and Review, 66*(6), s90-119. doi:10.1177/1077558709345870

Dy, S. M., Taylor, S. L., Carr, L. H., Foy, R., J., Pronovost, P. P., Ovretveit, J., . . . Shekelle, P. G. (2011). A framework for classifying patient safety practices: Results from an expert consensus process. *BMJ Quality & Safety, 20*, 618-624. doi:10.1136/bmjqs.2010.049296

Eberhart, L., Koch, T., Kranke, P., Rüsch, D., Torossian, A., & Nardi-Hiebl, S. (2014). Activity-based cost analysis of opioid-related nausea and vomiting among inpatients. *Journal of Opioid Management, 10*(6), 415-422. doi:10.5055/jom.2014.0238

ECRI Institute. (2015). Top 10 patient safety concerns for healthcare organizations. Plymouth Meeting, PA: Author.

Emanuel, L., Berwick, D., Conway, J., Combes, J., Hatlie, M., Leape, L., . . . Walton, M. (2008). What exactly is patient safety? New directions and alternative approaches (vol. 1). Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK43629/pdf/advances-emanuel-berwick_110.pdf

Etchegaray, J. M., & Thomas, E. J. (2012). Comparing two safety culture surveys: Safety Attitudes Questionnaire and Hospital Survey on Patient Safety. *BMJ Quality & Safety, 21*(6), 490-498. doi:10.1136/bmjqs-2011-000449

Etchegaray, J. M., & Thomas, E. J. (2014). Safety Attitudes Questionnaire: Recent findings and future areas of research. In P. Waterson (Ed.), *Patient safety culture: Theory, methods, and application* (pp. 285-298). Burlington, VT: Ashgate.

Fagan, A. T. (2012). *The relationship of safety culture and patient outcomes in U.S. academic health science centers* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses A&I; ProQuest Dissertations & Theses Full Text database. (3529679 Ph.D.).

Farup, P. G. (2015). Are measurements of patient safety culture and adverse events valid and reliable? Results from a cross sectional study. *BMC Health Services Research, 15*(186). doi:10.1186/s12913-015-0852-x

Fathi, M., Nikbakht Nasrabadi, A., & Valiee, S. (2014). The effects of body position on chemotherapy-induced nausea and vomiting: A single-blind randomized controlled trial. *Iranian Red Crescent Medical Journal, 16*(6), e17778. doi:10.5812/ircmj.17778

Fisher, E. B. (2008). The importance of context in understanding behavior and promoting health. *Annals of Behavioral Medicine, 35*, 3-18. doi:10.1007/s12160-007-9001-z

Fleming, M. (2005). Patient safety culture measurement and improvement: A "how to" guide. *Healthcare Quarterly, 8*(Sp), 14-19. doi:10.12927/hcq.2005.17656

Flin, R. (2007). Measuring safety culture in healthcare: A case for accurate diagnosis. *Safety Science, 45*(6), 653-667. doi:10.1016/j.ssci.2007.04.003

Flin, R., Burns, C., Mearns, K., Yule, S., & Robertson, E. M. (2006). Measuring safety climate in health care. *Quality and Safety in Health Care, 15*(2), 109-115. doi:10.1136/qshc.2005 .014761

Flin, R., Mearns, K., O'Connor, P., & Bryden, R. (2000). Measuring safety climate: Identifying the common features. *Safety Science, 34*(1-3), 177-192.

Florida Hospital. (2009). *Florida Hospital frequently asked questions* [re: IHI Global Trigger Tool White Paper]. Retrieved from http://www.safetyleaders.org/pages/sitesearch.jsp

Flynn, L., Liang, Y., Dickson, G. L., Xie, M., & Suh, D.-C. (2012). Nurses' practice environments, error interception practices, and inpatient medication errors. *Journal of Nursing Scholarship, 44*(2), 180-186.

France, D. J., Greevy, R. A., Liu, X., Burgess, H., Dittus, R. S., Weinger, M. B., & Speroff, T. (2010). Measuring and comparing safety climate in intensive care units. *Medical Care, 48*(3), 279-284.

Freeth, D., Sandall, J., Allan, T., Warburton, F., Berridge, E. J., Mackintosh, N., . . . Abbott, S. (2012). A methodological study to compare survey-based and observation-based evaluations or organizational and safety cultures and then compare both approaches with markers of the quality of care. *Health Technology Assessment, 16*(25), 1-8. doi:10.3310/ hta16250

Friedman, C., & Elhadad, N. (2014). Natural language processing in health care and biomedicine. In E. H. Shortliffe & J. J. Cimino (Eds.), Biomedical informatics computer applications in health care and biomedicine (4th ed.). New York: Springer. doi:10.1007/978-1-4471-4474-8

Fuller, R. L., McCullough, E. C., Bao, M. Z., & Averill, R. F. (2009). Estimating the costs of potentially preventable hospital acquired complications. *Health Care Financing Review, 30*(4), 17-32.

Gaba, D. M., Singer, S. J., Sinaiko, A. D., Bowen, J. D., & Ciavarelli, A. P. (2003). Differences in safety climate between hospital personnel and naval aviators. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 45*(2), 173-185. doi:10.1518/hfes .45.2.175.27238

Gallego, B., Westbrook, M. T., Dunn, A. G., & Braithwaite, J. (2012). Investigating patient safety culture across a health system: Multilevel modelling of differences associated with service types and staff demographics. *International Journal for Quality in Health Care, 24*(4), 311-320. doi:10.1093/intqhc/mzs028

Gan, T. J., Diemunsch, P., Habib, A. S., Kovac, A., Kranke, P., Meyer, T. A., . . . Philip, B. K. (2014). Consensus guidelines for the management of postoperative nausea and vomiting. *Anesthesia and Analgesia, 118*(1), 85-113. doi:10.1213/ANE .0000000000000002

Garon, M. (2012). Speaking up, being heard: Registered Nurses' perceptions of workplace communication. *Journal of Nursing Management, 20*(3), 361-371. doi:10.1111/j.1365-2834.2011.01296.x

Garrett Jr, P. R., Sammer, C., Nelson, A., Paisley, K. A., Jones, C., Shapiro, E., . . . Housman, M. (2013). Developing and implementing a standardized process for global trigger tool application across a large health system. *Joint Commission Journal on Quality and Patient Safety, 39*(7), 292-297.

Garrouste-Orgeas, M., Philippart, F., Bruel, C., Max, A., Lau, N., & Misset, B. (2012). Overview of medical errors and adverse events. *Annals of Intensive Care, 2*(2). doi:10.1186/2110-5820-2-2

Good, V. S., Saldana, M., Gilder, R., Nicewander, D., & Kennerly, D. (2011). Large-scale deployment of the Global Trigger Tool across a large hospital system: Refinements for the characterization of adverse events to support patient safety learning opportunities. *BMJ Quality & Safety, 20*, 25-30. doi:10.1136/bmjqs.2008.029181

Greenhalgh, T. (2010). *How to read a paper: The basics of evidence-based medicine* (4th ed.). Hoboken, NJ: Blackwell.

Griffin, F. A., & Resar, R. K. (2009). *IHI global trigger tool for measuring adverse events. IHI Innovation Series White Paper*. Retrieved from www.IHI.org

Grimes, D. A., & Schulz, K. F. (2002). Descriptive studies: What they can and cannot do. *Lancet, 359*(9301), 145.

Groves, P. S. (2014). The relationship between safety culture and patient outcomes: Results from pilot meta-analyses. *Western Journal of Nursing Research, 36*(1), 66-83. doi:10.1177/ 0193945913490080

Guldenmund, F. W. (2007). The use of questionnaires in safety culture research: An evaluation. *Safety Science, 45*(6), 723-743. doi:10.1016/j.ssci.2007.04.006

Guldenmund, F. W. (2014). Organisational safety culture principles. In P. Waterson (Ed.), *Patient safety culture: Theory, methods, and application* (pp. 15-41). Burlington, VT: Ashgate.

Haerkens, M., Jenkins, D., & VanderHoeven, H. (2012). Crew resource management in the ICU: The need for culture change. *Annals of Intensive Care, 2*(1), 39.

Halligan, M., & Zecevic, A. (2011). Safety culture in healthcare: A review of concepts, dimensions, measures and progress. *BMJ Quality & Safety, 20*, 338-343. doi:10.1136/bmjqs.2010 .040964

Härkänen, M., Kervinen, M., Ahonen, J., Voutilainen, A., Turunen, H., & Vehviläinen-Julkunen, K. (2015). Patient-specific risk factors of adverse drug events in adult inpatients: Evidence detected using the Global Trigger Tool method. *Journal of Clinical Nursing, 24*(3-4), 582-591. doi:10.1111/jocn.12714

Harrison, R., Lawton, R., Perlo, J., Gardner, P., Armitage, G., & Shapiro, J. (2015). Emotion and coping in the aftermath of medical error: A cross-country exploration. *Journal of Patient Safety, 11*(1), 28-35.

Havill, N. L., Havill, H. L., Mangione, E., Dumigan, D. G., & Boyce, J. M. (2011). Cleanliness of portable medical equipment disinfected by nursing staff. *American Journal of Infection Control, 39*(7), 602-604. doi:10.1016/j.ajic.2010.10.030

Havyer, R. A., Wingo, M., Comfere, N., Nelson, D., Halvorsen, A., McDonald, F., & Reed, D. (2014). Teamwork assessment in internal medicine: A systematic review of validity evidence and outcomes. *Journal of General Internal Medicine, 29*(6), 894-910. doi:10.1007/s11606-013-2686-8

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: The Guilford Press.

Hayward, R. A., Asch, S. M., Hogan, M. M., Hofer, T. P., & Kerr, E. A. (2005). Sins of omission: Getting too little medical care may be the greatest threat to patient safety. *Journal of General Internal Medicine, 20*(8), 686-691. doi:10.1111/j.1525-1497.2005.0152.x

Health Quality and Safety Commission New Zealand. (2014). *Revised Global Trigger Tool implementation guide 2014*. Retrieved from http://www.hqsc.govt.nz/our-programmes/other-topics/publications-and-resources/publication/690/

Hendrich, A., Chow, M. P., Skierczynski, B. A., & Lu, Z. (2008). A 36-hospital time and motion study: How do medical-surgical nurses spend their time? *The Permanente Journal, 12*(3), 25-34.

Henriksen, K., Dayton, E., Keyes, M. A., Carayon, P., & Hughes, R. (2008). Understanding adverse events: A human factors framework. In R. G. Hughes (Ed.), *Patient safety and quality: An evidence-based handbook for nurses*. Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK2666/pdf/ch5.pdf

Hodge, N. S., McCarthy, M. S., & Pierce, R. M. (2014). A prospective randomized study of the effectiveness of aromatherapy for relief of postoperative nausea and vomiting. *Journal of PeriAnesthesia Nursing, 29*(1), 5-11. doi:http://dx.doi.org/10.1016/j.jopan.2012.12.004

Hoffmann, B., & Rohe, J. (2010). Patient safety and error management: What causes adverse events and how can they be prevented? *Dtsch Arztebl Int, 107*(6), 92-99. doi:10.3238/arztebl.2010.0092

Hollar, D., Hobgood, C., Foster, B., Aleman, M., & Sawning, S. (2012). Concurrent validation of CHIRP, a new instrument for measuring healthcare student attitudes towards interdisciplinary teamwork. *Journal of Applied Measurement, 13*(4), 360-375.

Huang, D. T., Clermont, G., Kong, L., Weissfeld, L. A., Sexton, J. B., Rowan, K. M., & Angus, D. C. (2010). Intensive care unit safety culture and outcomes: A U.S. multicenter study. *International Journal for Quality in Health Care, 22*(3), 151-161. doi:10.1093/intqhc/mzq017

Huddleston, J. M., Diedrich, D. A., Kinsey, G. C., Enzler, M. J., & Manning, D. M. (2014). Learning from every death. *Journal of Patient Safety, 10*(1), 6-12. doi:10.1097/PTS.0000000000000053

Hug, B. L., Keohane, C., Seger, D. L., Yoon, C., & Bates, D. W. (2012). The costs of adverse drug events in community hospitals. *The Joint Commission Journal on Quality and Patient Safety, 38*(3), 120-126.

Hughes, L. C., Chang, Y., & Mark, B. A. (2012). Quality and strength of patient safety climate on medical-surgical units. *Journal of Nursing Administration, 42*(10), S27-S36.

Hung, C.-C., Lee, B.-O., Tsai, S.-L., Tseng, Y. S., & Chang, C.-H. (2013). Structure determines medication errors in nursing units: A mechanistic approach. *Western Journal of Nursing Research, 37*(3), 299-319.

Hwang, J.-I., Chin, H. J., & Chang, Y.-S. (2014). Characteristics associated with the occurrence of adverse events: A retrospective medical record review using the Global Trigger Tool in a fully digitalized tertiary teaching hospital in Korea. *Journal of Evaluation in Clinical Practice, 20*(1), 27-35. doi:10.1111/jep.12075

Institute of Medicine Committee on Identifying and Preventing Medication Errors. (2006). *Preventing medication errors.* P. Aspden, J. A. Wolcott, J. L. Bootman, & L. R. Cronenwett (Eds.). Retrieved from http://books.nap.edu/openbook.php?record_id=11623

Institute of Medicine. (2000). *To err is human: Building a safer health system*. Washington, DC: National Academies Press.

Jackson, J., Sarac, C., & Flin, R. (2010). Hospital safety climate surveys: Measurement issues. *Current opinion in critical care, 16*(6), 632-638.

James, J. T. (2013). A new, evidence-based estimate of patient harms associated with hospital care. *Journal of Patient Safety*. doi:10.1097/PTS.0b013e3182948a69

Je, S. M., Kim, J. S., You, J. S., Chung, S. P., Cho, J., Lee, J. H., . . . Chung, H. S. (2014). Assessing safety attitudes among healthcare providers after a hospital-wide high-risk patient care program. *Yonsei Medical Journal, 55*(2), 523-529.

Jylhä, V., Saranto, K., & Bates, D. W. (2011). Preventable adverse drug events and their causes and contributing factors: The analysis of register data. *International Journal for Quality in Health Care, 23*(2), 187-197. doi:10.1093/intqhc/mzq085

Kable, A., Gibberd, R., & Spigelman, A. (2008). Predictors of adverse events in surgical admissions in Australia. *International Journal for Quality in Health Care, 20*(6), 406-411. doi:10.1093/intqhc/mzn032

Kaldjian, L. C., Jones, E. W., Wu, B. J., Forman-Hoffman, V. L., Levi, B. H., & Rosenthal, G. E. (2008). Reporting medical errors to improve patient safety: A survey of physicians in teaching hospitals. *Archives of Internal Medicine, 168*(1), 40-46. doi:10.1001/archinternmed.2007.12

Kale, A., Keohane, C. A., Maviglia, S., Gandhi, T. K., & Poon, E. G. (2012). Adverse drug events caused by serious medication administration errors. *BMJ Quality & Safety, 21*(11), 933-938. doi:10.1136/bmjqs-2012-000946

Kalisch, B. J., Lee, H., & Salas, E. (2010). The development and testing of the nursing teamwork survey. *Nursing Research, 59*(1), 42-50.

Kalisch, B. J., Russell, K., & Lee, K. H. (2013). Nursing teamwork and unit size. *Western Journal of Nursing Research, 35*(2), 214-225. doi:10.1177/0193945912439107

Kane, R. L., & Radosevich, D. M. (2011). *Conducting health outcomes research*. Sudbury, MA: Jones and Bartlett Learning.

Kane, R. L., Shamliyan, T. A., Mueller, C., Duval, S., & Wilt, T. J. (2007). The association of Registered Nurse staffing levels and patient outcomes: Systematic review and meta-analysis. *Medical Care, 45*(12), 1195-1204.

Kanjanarat, P., Winterstein, A. G., Johns, T. E., Hatton, R. C., Gonzalez-Rothi, R., & Segal, R. (2003). Nature of preventable adverse drug events in hospitals: A literature review. *American Journal of Health-System Pharmacy, 60*(14). Retrieved from http://www .medscape.com/viewarticle/460564_2

Kaushal, R., Bates, D. W., Franz, C., Soukup, J. R., & Rothschild, J. M. (2007). Costs of adverse events in intensive care units. *Critical Care Medicine, 35*(11), 2479-2483.

Kennerly, D. A., Kudyakov, R., da Graca, B., Saldaña, M., Compton, J., Nicewander, D., & Gilder, R. (2014). Characterization of adverse events detected in a large health care delivery system using an enhanced Global Trigger Tool over a five-year interval. *Health Services Research, 49*(5), 1407-1425. doi:10.1111/1475-6773.12163

Kennerly, D. A., Saldaña, M., Kudyakov, R., daGraca, B., Nicewander, D., & Compton, J. (2013). Description and evaluation of adaptations to the Global Trigger Tool to enhance value to adverse event reduction efforts. *Journal of Patient Safety, 9*(2), 1-9.

Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research* (4th ed.). Belmont, CA: Cengage Learning.

Kim, H., Capezuti, E., Kovner, C., Zhao, Z., & Boockvar, K. (2010). Prevalence and predictors of adverse events in older surgical patients: Impact of the present on admission indicator. *The Gerontologist, 50*(6), 810-820. doi:10.1093/geront/gnq045

Kirkendall, E. S., Kloppenborg, E., Papp, J., White, D., Frese, C., Hacker, D., . . . Kotagal, U. (2012). Measuring adverse events and levels of harm in pediatric inpatients with the global trigger tool. *Pediatrics, 130*(5), e1206-e1214. doi:10.1542/peds.2012-0179

Kirwan, M., Matthews, A., & Scott, P. A. (2013). The impact of the work environment of nurses on patient safety outcomes: A multi-level modelling approach. *International Journal Of Nursing Studies, 50*(2), 253-263.

Kline, T. J. B., Willness, C., & Ghali, W. A. (2008). Determinants of adverse events in hospitals-The potential role of patient safety culture. *Journal for Healthcare Quality, 30*(1), 11-17.

Knox, G. E., & Simpson, K. R. (2004). Teamwork: The fundamental building block of high-reliability organizations and patient safety. In B. J. Youngberg & M. J. Hatlie (Eds.), *The patient safety handbook* (pp. 379-414). Sudbury, MA: Jones and Bartlett.

Kutney-Lee, A., & Kelly, D. (2011). The effect of hospital electronic health record adoption on nurse-assessed quality of care and patient safety. *Journal of Nursing Administration, 41*(11), 466-472. doi:10.1097/NNA.0b013e3182346e4b

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Landrigan, C. P., Parry, G. J., Bones, C. B., Hackbarth, A. D., Goldmann, D. A., & Sharek, P. J. (2010). Temporal trends in rates of patient harm resulting from medical care. *New England Journal of Medicine, 363*(22), 2124-2134. doi:10.1056/NEJMsa1004404

Latif, A., Rawat, N., Pustavoitau, A., Pronovost, P. J., & Pham, J. C. (2013). National study on the distribution, causes, and consequences of voluntarily reported medication errors between the ICU and non-ICU settings. *Critical Care Medicine, 42*(2), 389-398. doi:10.1097/CCM .0b013e318274156a

Laurent, A., Aubert, L., Chahraoui, K., Bioy, A., Mariage, A., Quenot, J.-P., & Capellier, G. (2014). Error in intensive care: Psychological repercussions and defense mechanisms among health professionals. *Critical Care Medicine, 42*(11), 2370–2378. doi:10.1097/CCM .0000000000000508

Lawton, R., McEachan, R. R. C., Giles, S. J., Sirriyeh, R., Watt, I. S., & Wright, J. (2012). Development of an evidence-based framework of factors contributing to patient safety incidents in hospital settings: A systematic review. *BMJ Quality & Safety, 21*(5), 369-380. doi:10.1136/bmjqs-2011-000443

Layde, P. M., Maas, L. A., Teret, S. P., Brasel, K. J., Kuhn, E. M., Mercy, J. A., & Hargarten, S. W. (2002). Patient safety efforts should focus on medical injuries. *Journal of the American Medical Association, 287*(15), 1993-1997. doi:10.1001/jama.287.15.1993

Lee, J., Kuo, Y. F., & Goodwin, J. S. (2013). The efect of electronic medical record adoption on outcomes in U.S. hospitals. *BMC Health Services Research, 13*, 39. Retrieved from doi:10.1186/1472-6963-13-39

Leonard, M. W., & Frankel, A. (2011). Role of effective teamwork and communication in delivering safe, high-quality care. *Mount Sinai Journal of Medicine, 78*(6), 820-826. doi:10.1002/msj .20295

Leonard, M. W., Graham, S., & Bonacum, D. (2004). The human factor: The critical importance of effective teamwork and communication in providing safe care. *Quality and Safety in Health Care, 13*(suppl 1), i85-i90. doi:10.1136/qshc.2004.010033

Levinson, D. R. (2010). *Adverse events in hospitals: National incidence among Medicare beneficiaries*. (OEI-06-09-00090). Washington, DC: Department of Health and Human Services.

Li, Q., Melton, K., Lingren, T., Kirkendall, E. S., Hall, E., Zhai, H., . . . Solti, I. (2014). Phenotyping for patient safety: Algorithm development for electronic health record based automated adverse event and medical error detection in neonatal intensive care. *Journal of the American Medical Informatics Association, 21*(5), 776-784. doi:10.1136/amiajnl-2013-001914

Lipczak, H., Neckelmann, K., Steding-Jessen, M., Jakobsen, E., & Knudsen, J. (2011). Uncertain added value of Global Trigger Tool for monitoring of patient safety in cancer care. *Danish Medical Bulletin, 58*(11), A4337.

Lipschultz, A. (2013). Adverse patient events involving medical devices. *Biomedical Instrumentation & Technology, 47*(2), 152-156. doi:10.2345/0899-8205-47.2.152

Listyowardojo, T. A., Nap, R. E., & Johnson, A. (2012). Variations in hospital worker perceptions of safety culture. *International Journal for Quality in Health Care, 24*(1), 9-15. doi:10.1093/ intqhc/mzr069

Logan, J. (2013). *Introduction to regression*. Class lecture. Center for Cognitive and Brain Sciences. Ohio State University. Columbus, OH.

Lomax, R. G. (2007). *An introduction to statistical concepts* (2nd ed.). New York, NY: Routledge.

Mahmood, A., Chaudhury, H., & Valente, M. (2011). Nurses' perceptions of how physical environment affects medication errors in acute care settings. *Applied Nursing Research, 24*(4), 229-237. doi:10.1016/j.apnr.2009.08.005

Malloy, P. A. (2012). *Examining the relationship between work climate and patient safety among nurses in acute care settings.* Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3503408)

Manser, T. (2009). Teamwork and patient safety in dynamic domains of healthcare: A review of the literature. *Acta Anesthesiologica Scandinavica, 53*(2), 143-151. doi:10.1111/j.1399-6576.2008.01717.x

Mansfield, S. A., & Noria, S. (2015). Human error in the context of health systems approach. In S. P. Stawicki, S. C. Galwankar, T. J. Papadimos, & S. D. Moffatt-Bruce (Eds.), *Fundamentals of patient safety in medicine and surgery* (pp. 69-76). Philadelphia, PA: Lippincott, Williams, and Wilkins.

Mansukhani, M., Kolla, B. P., Surani, S., Varon, J., & Ramar, K. (2012). Sleep deprivation in resident physicians, work hour limitations, and related outcomes: A systematic review of the literature. *Postgraduate Medicine, 124*(4), 241-249. doi:10.3810/pgm.2012.07.2583

Mardon, R., Khanna, K., Sorra, J., Dyer, N., & Famolaro, T. (2010). Exploring relationships between hospital patient safety culture and adverse events. *Journal of Patient Safety, 6*(4), 226-232.

Mark, B. A., Harless, D. W., & McCue, M. (2005). The impact of HMO penetration on the relationship between nurse staffing and quality. *Health Economics, 14*(7), 737-753. doi: 10.1002/hec.988

Mark, B. A., Hughes, L. C., & Jones, C. B. (2004). The role of theory in improving patient safety and quality in health care. *Nursing Outlook, 52*, 11-16.

Mattox, E. A. (2010). Identifying vulnerable patients at heightened risk for medical error. *Critical Care Nurse, 30*(2), 61-70. doi:10.4037/ccn2010102

Mattsson, T. O., Knudsen, J. L., Lauritsen, J., Brixen, K., & Herrstedt, J. (2013). Assessment of the global trigger tool to measure, monitor and evaluate patient safety in cancer patients: reliability concerns are raised. *BMJ Quality & Safety, 22*(7), 571-579. doi:10.1136/bmjqs-2012-001219

Mazzocco, K., Petitti, D. B., Fong, K. T., Bonacum, D., Brookey, J., Graham, S., . . . Thomas, E. J. (2009). Surgical team behaviors and patient outcomes. *The American Journal of Surgery, 197*(5), 678-685. doi:10.1016/j.amjsurg.2008.03.002

McFadden, K. L., Stock, G. N., & Gowen, C. R. (2015). Leadership, safety climate, and continuous quality improvement: Impact on process quality and patient safety. *Health Care Management Review, 40*(1), 24-34. doi:10.1097/HMR.0000000000000006

McGahan, M., Kucharski, G., & Coyer, F. (2012). Nurse staffing levels and the incidence of mortality and morbidity in the adult intensive care unit: A literature review. *Australian Critical Care, 25*(2), 64-77. doi:10.1016/j.aucc.2012.03.003

Melnyk, B. M. (2014). Speeding the translation of research into evidence-based practice and conducting projects that impact healthcare quality, patient outcomes and costs: The 'so what' outcome factors. *Worldviews on Evidence-Based Nursing, 11*(1), 1-4. doi:10.1111/wvn.12025

Metersky, M. L., Hunt, D. R., Kliman, R., Wang, Y., Curry, M., Verzier, N., . . . Moy, E. (2011). Racial disparities in the frequency of patient safety events: Results from the national Medicare patient safety monitoring system. *Medical Care, 49*(5), 504-510. doi:10.2307/23053809

Michel, P., Quenon, J. L., Djihoud, A., Tricaud-Vialle, S., & de Sarasqueta, A. M. (2007). French national survey of inpatient adverse events prospectively assessed with ward staff. *Quality and Safety in Health Care, 16*(5), 369-377. doi:10.1136/qshc.2005.016964

Mills, P. D., Neily, J., Kinney, L. M., Bagian, J., & Weeks, W. B. (2008). Effective interventions and implementation strategies to reduce adverse drug events in the Veterans Affairs (VA) system. *Quality and Safety in Health Care, 17*(1), 37-46. doi:10.1136/qshc.2006.021816

Mitchell, P. H. (2008). Defining patient safety and quality care. In R. G. Hughes (Ed.), Patient safety and quality: An evidence-based handbook for nurses (pp. 1-5). Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from http://www-ncbi-nlm-nih-gov.proxy.lib.ohio-state.edu/books/NBK2681/

Molloy, P. A. (2012). *Examining the relationship between work climate and patient safety among nurses in acute care settings* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses A&I; ProQuest Dissertations & Theses Full Text database. (3503408 Ph.D.).

Montalvo, I. (2007). The national database of nursing quality indicators. *OJIN: The Online Journal of Issues In Nursing, 12*(3). doi:0.3912/OJIN.Vol12No03Man02

Montazeri, A. S., Hamidzadeh, A., Raei, M., Mohammadiun, M., Montazeri, A. S., Mirshahi, R., & Rohani, H. (2013). Evaluation of oral ginger efficacy against postoperative nausea and vomiting: A randomized, double - blinded clinical trial. *Iranian Red Crescent Medical Journal, 15*(12), e12268. doi:10.5812/ircmj.12268

Nabhan, M., Elraiyah, T., Brown, D. R., Dilling, J., LeBlanc, A., Montori, V. M., . . . Murad, M. H. (2012). What is preventable harm in healthcare? A systematic review of definitions. *BMC Health Services Research, 12*(128). doi:10.1186/1472-6963-12-128

Naessens, J. M., Campbell, C. R., Shah, N., Berg, B., Lefante, J. J., Williams, A. R., & Culbertson, R. (2012). Effect of illness severity and comorbidity on patient safety and adverse events. *American Journal of Medical Quality, 27*(1), 48-57. doi:10.1177/1062860611413456

Naessens, J. M., O'Byrne, T. J., Johnson, M. G., Vansuch, M. B., McGlone, C. M., & Huddleston, J. M. (2010). Measuring hospital adverse events: Assessing inter-rater reliability and trigger performance of the Global Trigger Tool. *International Journal for Quality in Health Care, 22*(4), 266-274. doi:10.1093/intqhc/mzq026

National Pressure Ulcer Advisory Panel, European Pressure Ulcer Advisory Panel, & Pan Pacific Pressure Injury Alliance. (2014). *Prevention and treatment of pressure ulcers: Quick reference guide*. Retrieved from http://www.npuap.org/wp-content/uploads/2014/08/Updated-10-16-14-Quick-Reference-Guide-DIGITAL-NPUAP-EPUAP-PPPIA-16Oct2014.pdf

National Cancer Institute. (2015, June 8). Protocol development: Common terminology criteria for adverse events (CTCAE) v5.0 draft review. Retrieved June 15, 2015, from http://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm

National Quality Forum. (2009). *Safe practices for better healthcare: 2009 update: A consensus report*. Washington, DC: Author.

National Quality Forum. (2012). *National voluntary consensus standards for patient safety: A consensus report*. Retrieved from http://www.qualityforum.org/Publications/2012/03/National_Voluntary_Consensus_Standards_for_Patient_Safety__A_Consensus_Report.aspx

National Research Council. (1992). *Human factors specialists' education and utilization: Results of a survey*. Washington, DC: The National Academies Press.

Needleman, J., Buerhaus, P., Mattke, S., Stewart, M., & Zelevinsky, K. (2002). Nurse-staffing levels and the quality of care in hospitals. *New England Journal of Medicine, 346*(22), 1715-1722. doi:10.1056/NEJMsa012247

Nemeth, C. P. (2012). Process without progress: Emergency medicine, patient safety, and the need for science. *Annals of Emergency Medicine, 59*(2), 112-114. doi:10.1016/j.annemergmed.2011 .08.016

Nemeth, C. P., Cook, R. I., Donchin, Y., Dierks, M., Patterson, E., Bitan, Y., . . . Powell, T. (2006). *Learning from investigation: Experience with understanding healthcare adverse events*. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(9), 914-917. doi:10.1177/154193120605000911

Neuraz, A., Guerin, C., Payet, C., Polazzi, S., Aubrun, F., Dailler, F., . . . Duclos, A. (2015). Patient mortality is associated with staff resources and workload in the ICU: A multicenter observational study. *Critical Care Medicine*. Advance online publication. doi:10.1097/CCM.0000000000001015

Nilsson, L., Pihl, A., Tagsjo, M., & Ericsson, E. (2012). Adverse events are common on the intensive care unit: Results from a structured record review. *Acta Anaesthesiologica Scandinavica*, n/a-n/a. doi:10.1111/j.1399-6576.2012.02711.x

Nobre, C., & McKay, C. (2012). Surveillance of adverse drug events in a large tertiary-care hospital. *Connecticut Medicine, 76*(2), 91-94.

November, M., Chie, L., & Weingart, S. N. (2008). Physician-reported adverse events and medical errors in obstetrics and gynecology In K. Henriksen, J. B. Battles, & M. A. Keyes (Eds.), *Advances in patient safety: New directions and alternative approaches* (Vol. 1: Assessment). Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK43623/?report=reader#!po=5.55556.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

O'Connor, P., Buttrey, S. E., O'Dea, A., & Kennedy, Q. (2011). Identifying and addressing the limitations of safety climate surveys. *Journal of Safety Research, 42*, 259-265.

O'Daniel, M., & Rosenstein, A. H. (2008). Professional communication and team collaboration. In R. G. Hughes (Ed.), *Patient safety and quality: An evidence-based handbook for nurses* (vol. 2, pp. 271-284). Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from http://www.ahrq.gov/professionals/clinicians-providers/resources/nursing/resources/nurseshdbk/ODanielM_TWC.pdf

O'Leary, K. J., Boudreau, Y. N., Creden, A. J., Slade, M. E., & Williams, M. V. (2012). Assessment of teamwork during structured interdisciplinary rounds on medical units. *Journal of Hospital Medicine, 7*(9), 679-683. doi:10.1002/jhm.1970

O'Leary, K. J., Devisetty, V. K., Patel, A. R., Malkenson, D., Sama, P., Thompson, W. K., . . . Williams, M. V. (2013). Comparison of traditional trigger tool to data warehouse based screening for identifying hospital adverse events. *BMJ Quality & Safety, 22*(2), 130-138. doi:10.1136/bmjqs-2012-001102

O'Leary, K. J., Sehgal, N. L., Terrell, G., & Williams, M. V. (2012). Interdisciplinary teamwork in hospitals: A review and practical recommendations for improvement. *Journal of Hospital Medicine, 7*(1), 48-54. doi:10.1002/jhm.970

Ohio Hospital Association. (2008). *Preventable adverse events fact sheet*. Retrieved November 4, 2012, from http://www.ohanet.org/wp-content/uploads/2012/02/PreventableAdverseEvents.pdf

Ohman, J., Keski-Kuha, T., Kaartinen, M., & Kujansuu, E. (2011, January). *Use of Global Trigger Tool in detecting adverse events in neurosurgery: comparison of manual and text mining detection of triggers*. Paper presented at the Patient Safety Research Days, Helsinki, Finland. Retrieved from https://www.thl.fi/documents/10531/102913/potilasturvallisuustutkimuksen_paivat_2011.pdf

Osterlind, S. J. (2010). *Modern measurement: Theory, principles, and applications of mental appraisal*. Boston: Pearson.

Owen, S. V., & Froman, R. D. (2005). Why carve up your continuous data? *Research in Nursing & Health, 28*(6), 496-503. doi:10.1002/nur.20107

Paine, L. A., Rosenstein, B. J., Sexton, J. B., Kent, P., Holzmueller, C. G., & Pronovost, P. J. (2011). Republished paper: Assessing and improving safety culture throughout an academic medical centre: A prospective cohort study. *Postgraduate Medicine, 87*, 428-435. doi:10.1136/pgmj.2009.039347rep

Pakyz, A. L., & Edmond, M. B. (2013). Influence of state laws mandating reporting of healthcare-associated infections: The case of central line-associated bloodstream infections. *Infection Control and Hospital Epidemiology, 34*(8), 780-784. doi:10.1086/671280

Pannick, S., Beveridge, I., Wachter, R. M., & Sevdalis, N. (2014). Improving the quality and safety of care on the medical ward: A review and synthesis of the evidence base. *European Journal of Internal Medicine, 25*(10), 874-887. doi:10.1016/j.ejim.2014.10.013

Pappas, S. H. (2008). The cost of nurse-sensitive adverse events. *Journal of Nursing Administration, 38*(5), 230-236.

Park, S. H., Blegen, M. A., Spetz, J., Chapman, S. A., & De Groot, H. (2012). Patient turnover and the relationship between nurse staffing and patient outcomes. *Research in Nursing and Health, 35*(3), 277-288. doi:http://dx.doi.org/10.1002/nur.21474

Parry, A. M., Barriball, K. L., & While, A. E. (2015). Factors contributing to registered nurse medication administration error: A narrative review. *International Journal Of Nursing Studies, 52*, 403-420. doi:10.1016/j.ijnurstu.2014.07.003

Parry, G., Cline, A., & Goldmann, D. (2012). Deciphering harm measurement. *The Journal of the American Medical Association, 307*(20), 2155-2156. doi:10.1001/jama.2012.3649

Pascal Metrics (2013). [Safety culture data interpretation guide: SAQ].

Peltomaa, K. (2012). James Reason: Patient safety, human error, and swiss cheese. *Quality Management in Healthcare, 21*(1), 59-63. doi:10.1097/QMH.1090b1013e3182418294

Peng, C.-Y. J., Kuk, L. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research, 96*(1), 3-14.

Perneger, T. V. (2005). The Swiss cheese model of safety incidents: Are there holes in the metaphor? *BMC Health Services Research, 5*, 71-71. doi:10.1186/1472-6963-5-71

Pettker, C. M., Thung, S. F., Norwitz, E. R., Buhimschi, C. S., Raab, C. A., Copel, J. A., . . . Funai, E. F. (2009). Impact of a comprehensive patient safety strategy on obstetric adverse events. *American Journal of Obstetrics and Gynecology, 200*(5), 492.e491-.e498. doi:10.1016/j.ajog.2009.01.022

Pham, J. C., Girard, T., & Pronovost, P. J. (2013). What to do with healthcare incident reporting systems. *Journal of Public Health Research, 2*(3), e27. doi:10.4081/jphr.2013.e27

Polit, D. F. (2010). *Statistics and data analysis for nursing research* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

Polit, D. F., & Hungler, B. P. (1995). *Nursing research: Principles and methods* (5th ed.). Philadelphia: J. B. Lippincott.

Portner, M., Fumanti, B. J., Bendas, C., & Cipolla, J. (2015). Learning from others: Examples from air transportation and industrial realms. In S. P. Stawicki, S. C. Galwankar, T. J. Papadimos, & S. D. Moffatt-Bruce (Eds.), *Fundamentals of patient safety in medicine and surgery* (pp. 92-101). Philadelphia, PA: Lippincott, Williams, and Wilkins.

Powell-Cope, G., Nelson, A. L., & Patterson, E. S. (2008). Patient care technology and safety. In R. G. Hughes (Ed.) *Patient safety and quality: An evidence-based handbook for nurses* (vol. 3, pp. 207-219). Retrieved from http://www-ncbi-nlm-nih-gov.proxy.lib.ohio-state.edu/books/NBK2686/

Pronovost, P. J., & Sexton, J. B. (2005). Assessing safety culture: Guidelines and recommendations. *Quality and Safety in Health Care, 14*(4), 231-233.

Pronovost, P. J., & Wachter, R. M. (2014). Progress in patient safety: A glass fuller than it seems. *American Journal of Medical Quality, 29*(2), 165-169. doi: 10.1177/1062860613495554

Pronovost, P. J., Berenholtz, S. M., Goeschel, C., Thom, I., Watson, S. R., Holzmueller, C. G., . . . Sexton, J. B. (2008). Improving patient safety in intensive care units in Michigan. *Journal of Critical Care, 23*(2), 207-221. doi:10.1016/j.jcrc.2007.09.002

Pronovost, P. J., Goeschel, C. A., Marsteller, J. A., Sexton, J. B., Pham, J. C., & Berenholtz, S. M. (2009). Framework for patient safety research and improvement. *Circulation, 119*(2), 330-337. doi:10.1161/circulationaha.107.729848

Pronovost, P. J., Holzmueller, C. G., Needham, D. M., Sexton, J. B., Miller, M., Berenholtz, S. M., . . . Morlock, L. (2006). How will we know patients are safer? An organization-wide approach to measuring and improving safety. *Critical Care Medicine, 34*(7), 1988-1995.

Pumar-Méndez, M. J., Attree, M., & Wakefield, A. (2014). Methodological aspects in the assessment of safety culture in the hospital setting: A review of the literature. *Nurse Education Today, 34*(3), 162-170. doi:http://dx.doi.org/10.1016/j.nedt.2013.08.008

Radley, D. C., Wasserman, M. R., Olsho, L. E. W., Shoemaker, S. J., Spranca, M. D., & Bradshaw, B. (2013). Reduction in medication errors in hospitals due to adoption of computerized provider order entry systems. *Journal of the American Medical Informatics Association, 20*(3), 470-476. doi:10.1136/amiajnl-2012-001241

Rashid, M. (2014). Two decades (1993-2012) of adult intensive care unit design: A comparative study of the physical design features of the best practice examples. *Critical Care Nursing Quarterly, 37*(1), 3-32. doi:10.1097/CNQ.0000000000000002

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement, 21*(2), 173-184. doi:10.1177/01466216970212006

Raykov, T. (2001). Bias of coefficient alpha for fixed congeneric measures with correlated errors. *Applied Psychological Measurement, 25*(1), 69-76.

Reid, J., & Bromiley, M. (2012). Clinical human factors: The need to speak up to improve patient safety. *Nursing Standard, 26*(35), 35-40.

Reiling, J., Hughes, R. G., & Murphy, M. R. (2008). The impact of facility design on patient safety. In R. G. Hughes (Ed.), *Patient safety and quality: An evidence-based handbook for nurses* (vol. 2, pp. 167-192). Rockville, MD: Agency for Healthcare Research and Quality

Reiman, T., Pietikäinen, E., & Oedewald, P. (2010). Multilayered approach to patient safety culture. *Quality and Safety in Health Care, 19*(5), e20. doi:10.1136/qshc.2008.029793

Resar, R. K. (2009, February). *Reflections on the Institute for Healthcare Improvement (IHI) Global Trigger Tool: Triggers and targeted injury detection systems (TIDS)*. Retrieved from http://archive.ahrq.gov/news/events/other/triggers/triggers4.html

Resar, R. K., & Griffin, F. (2014). *Web and action: Using the IHI Global Trigger Tool for measuring patient safety: Session #2*. webinar. Institute for Healthcare Improvement.

Resar, R. K., Rozich, J. D., & Classen, D. (2003). Methodology and rationale for the measurement of harm with trigger tools. *Quality and Safety in Health Care, 12*(Suppl II), ii39-ii45.

Rivard, P. E., Elixhauser, A., Christiansen, C. L., Zhao, S., & Rosen, A. K. (2010). Testing the association between patient safety indicators and hospital structural charcteristics in VA and nonfederal hospitals *Medical Care Research and Review, 67*(3), 321-341.

Rose, J. S., Thomas, C. S., Tersigni, A., Sexton, J. B., & Pryor, D. (2006). A leadership framework for culture change in health care. *Joint Commission Journal on Quality and Patient Safety, 32*(8), 433-442.

Rosen, A. K., Mull, H. J., Kaafarani, H., Nebeker, J., Shimada, S., Helwig, A., . . . Itani, K. (2011). Applying trigger tools to detect adverse events associated with outpatient surgery. *Journal of Patient Safety, 7*(1), 45-59.

Rothschild, J. M., Landrigan, C. P., Cronin, J. W., Kaushal, R., Lockley, S. W., Burdick, E., . . . Bates, D. W. (2005). The critical care safety study: The incidence and nature of adverse events and serious medical errors in intensive care. *Critical Care Medicine, 33*(8), 1694-1700. doi:10.1097/01.CCM.0000171609.91035.BD

Runciman, W. B., Baker, G. R., Michel, P., Dovey, S., Lilford, R., Jensen, N., . . . Bates, D. W. (2010). Tracing the foundations of a conceptual framework for a patient safety ontology. *Quality and Safety in Health Care, 19*(e56). doi:10.1136/qshc.2009.035147

Runciman, W. B., Baker, G. R., Michel, P., Jauregui, I. L., Lilford, R. J., Andermann, A., . . . Weeks, W. B. (2008). The epistemology of patient safety research. *International Journal for Evidence-based Healthcare, 6*, 476-486. doi:10.1111/j.1479-6988.2008.00117.x

Runciman, W. B., Hibbert, P., Thomson, R. G., VanDerSchaaf, T., Sherman, H., & Lewalle, P. (2009). Towards an international classification for patient safety: Key concepts and terms. *International Journal for Quality in Health Care, 21*(1), 18-26. doi:10.1098/intqhc/mzn057

Runciman, W. B., Williamson, J. A. H., Deakin, A., Benveniste, K. A., Bannon, K., & Hibbert, P. D. (2006). An integrated framework for safety, quality and risk management: An information and incident management system based on a universal patient safety classification. *Quality and Safety in Health Care, 15*(suppl 1), i82-i90. doi:10.1136/qshc.2005.017467

Russ, A. L., Fairbanks, R. J., Karsh, B. T., Militello, L. G., Saleem, J. J., & Wears, R. L. (2013). The science of human factors: Separating fact from fiction. *Quality and Safety in Health Care*. doi:10.1136/bmjqs-2012-001450

Rutberg, H., Borgstedt Risberg, M., Sjödahl, R., Nordqvist, P., Valter, L., & Nilsson, L. (2014). Characterisations of adverse events detected in a university hospital: A 4-year study using the Global Trigger Tool method. *BMJ Open, 4*(e004879). doi:10.1136/bmopen-2014-004879

Saberi, F. (2014). Effect of ginger on relieving nausea and vomiting in pregnancy: A randomized, placebo-controlled trial. *Nursing & Midwifery Studies, 3*(1), e11841.

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On teams, teamwork, and team performance: Discoveries and developments. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50*(3), 540-547. doi:10.1136/bmjopen-2014-004879

Sammer, C. E. (2009). *Culture of safety in hospitals: A three-part analysis of safety culture, evidence-based practice guidelines, and patient outcomes* (Doctoral dissertation), Retrieved from ProQuest Dissertations & Theses A&I; ProQuest Dissertations & Theses Full Text database. (3396587)

Sammer, C. E., Lykens, K., Singh, K. P., Mains, D. A., & Lackan, N. A. (2010). What is patient safety culture? A review of the literature. *Journal of Nursing Scholarship, 42*(2), 156-165. doi:10.1111/j.1547-5069.2009.01330.x

Schildmeijer, K., Nilsson, L., Arestedt, K., & Perk, J. (2012). Assessment of adverse events in medical care: Lack of consistency between experienced teams using the Global Tigger Tool. *BMJ Quality and Safety, 21*, 307-314. doi:10.1136/bmjqs-2011-000279

Schildmeijer, K., Nilsson, L., Perk, J., Arestedt, K., & Nilsson, G. (2013). Strengths and weaknesses of working with the Global Trigger Tool method for retrospective record review: Focus group interviews with team members. *BMJ Open, 3*(e003131). doi:10.1136/bmjopen-2013-003131

Schulman, P. (2004). General attributes of safe organisations. *Quality & Safety in Health Care, 13*(Suppl 2), ii39-ii44. doi:10.1136/qshc.2003.009613

Schwendimann, R., Zimmermann, N., Küng, K., Ausserhofer, D., & Sexton, J. B. (2013). Variation in safety culture dimensions within and between US and Swiss Hospital Units: An exploratory study. *BMJ Quality & Safety, 22*(1), 32-41. doi:10.1136/bmjqs-2011-000446

Sehgal, N. L., Fox, M., Sharpe, B. A., Vidyarthi, A. R., Blegen, M. A., & Wachter, R. M. (2011). Critical conversations: A call for a nonprocedural “time out.” *Journal of Hospital Medicine, 6*(4), 225-230. doi:10.1002/jhm.853

Sexton, J. B., Helmreich, R. L., Neilands, T. B., Rowan, K., Vella, K., Boyden, J., . . . Thomas, E. J. (2006). The Safety Attitudes Questionnaire: Psychometric properties, benchmarking data, and emerging research. *BMC Health Services Research, 6*, 44-10. doi:10.1186/ 1472-6963-6-44

Sharek, P. J. (2012). The emergence of the Trigger Tool as the premier measurement strategy for patient safety. *Web M & M: Morbidity and mortality rounds on the web*. Retrieved from http://webmm.ahrq.gov/perspective.aspx?perspectiveID=120

Sharek, P. J., & Classen, D. C. (2006). The incidence of adverse events and medical error in pediatrics. *Pediatric Clinics of North America, 53*, 1067-1077.

Sharek, P. J., Parry, G., Goldmann, D., Bones, K., Hackbarth, A., Resar, R., . . . Landrigan, C. P. (2011). Performance characteristics of a methodology to quantify adverse events over time in hospitalized patients. *Health Services Research, 46*(2), 654-678. doi:10.1111/ j.1475-6773.2010.01156.x

Shekelle, P. G. (2013). Nurse-patient ratios as a patient safety strategy: A systematic review. *Annals of Internal Medicine, 158*(5_Part_2), 404-409. doi:10.7326/0003-4819-158-5-201303051-00007

Shekelle, P. G., Pronovost, P. J., & Wachter, R. M. (2010). *Assessing the evidence for context-sensitive effectiveness and safety of patient safety practices: Developing criteria*. Rockville, MD: Agency for Healthcare Research and Quality.

Shekelle, P. G., Pronovost, P. J., Wachter, R. M., McDonald, K. M., Schoelles, K., Dy, S. M., . . . Lohr, K. N. (2013). The top patient safety strategies that can be encouraged for adoption now. *Annals of Internal Medicine, 158*, 365-368.

Shekelle, P. G., Pronovost, P. J., Wachter, R. M., Taylor, S. L., Dy, S. M., Foy, R., . . . Haughom, J. (2011). Advancing the science of patient safety. *Annals of Internal Medicine, 154*(10), 693-W.248.

Shin, M. H., Sullivan, J. L., Rosen, A. K., Solomon, J. L., Dunn, E. J., Shimada, S. L., . . . Rivard, P. E. (2014). Examining the validity of AHRQ's patient safety indicators (PSIs): Is variation in PSI composite score related to hospital organizational factors? *Medical Care Research and Review, 71*(6), 599-618. doi:10.1177/1077558714556894

Shojania, K. G., & Thomas, E. J. (2013). Trends in adverse events over time: why are we not improving? *BMJ Quality & Safety, 22*(4), 273-277. doi:10.1136/bmjqs-2013-001935

Shrader, S., Kern, D., Zoller, J., & Blue, A. (2013). Interprofessional teamwork skills as predictors of clinical outcomes in a simulated healthcare setting. *Journal of Allied Health, 42*(1), 1E-6E.

Shrivastava, S., Sonpar, K., & Pazzaglia, F. (2009). Normal accident theory versus high reliability theory: A resolution and call for an open systems view of accidents. *Human Relations, 62*(9), 1357-1390. doi:10.1177/0018726709339117

Singer, S. J., Gaba, D. M., Falwell, A., Lin, S., Hayes, J., & Baker, L. (2009). Patient safety climate in 92 U.S. hospitals: Differences by work area and discipline. *Medical Care, 47*(1), 23-31. doi:10.1097/MLR.1090b1013e31817e31189d

Singer, S. J., Gaba, D. M., Geppert, J. J., Sinaiko, A. D., Howard, S. K., & Park, K. C. (2003). The culture of safety: Results of an organization-wide survey in 15 California hospitals. *Quality and Safety in Health Care, 12*(2), 112-118. doi:10.1136/qhc.12.2.112

Singer, S. J., Lin, S., Falwell, A., Gaba, D. M., & Baker, L. (2009). Relationship of safety climate and safety performance in hospitals. *Health Services Research, 44*(2, Part I), 399-421. doi:10.1111/j.1475-6773.2008.00918.x

Singer, S. J., Meterko, M., Baker, L., Gaba, D. M., Falwell, A., & Rosen, A. K. (2007). Workplace perceptions of hospital safety culture: Development and validation of the Patient Safety Climate in Healthcare Organizations survey. *Health Services Research, 42*(5), 1999-2021.

Singer, S. J., & Vogus, T. J. (2013). Reducing hospital errors: Interventions that build safety culture. *Annual Review of Public Health, 34*, 373-396. doi:10.1146/annurev-publhealth-031912-114439

Singla, A. K., Kitch, B. T., Weissman, J. S., & Campbell, E. G. (2006). Assessing patient safety culture: A review and synthesis of the measurement tools. *Journal of Patient Safety, 2*(3), 105-115.

Sinopoli, D. J., Needham, D. M., Thompson, D. A., Holzmueller, C. G., Dorman, T., Lubomski, L. H., . . . Pronovost, P. J. (2007). Intensive care unit safety incidents for medical versus surgical patients: A prospective multicenter study. *Journal of Critical Care, 22*(3), 177-183. doi:http://dx.doi.org/10.1016/j.jcrc.2006.11.002

Sitterding, M. C., Broome, M. E., Everett, L. Q., & Ebright, P. R. (2012). Understanding situation awareness in nursing work: A hybrid concept analysis. *Advances in Nursing Science, 35*(1), 77-92.

Smits, M., Wagner, C., Spreeuwenberg, P., Timmermans, D. R. M., van der Wal, G., & Groenewegen, P. P. (2012). The role of patient safety culture in the causation of unintended events in hospitals. *Journal of Clinical Nursing, 21*(23-24), 3392-3401. doi:10.1111/j.1365-2702.2012.04261.x

Smits, M., Zegers, M., Groenewegen, P. P., Timmermans, D. R., Zwaan, L., van der Wal, G., & Wagner, C. (2010). Exploring the causes of adverse events in hospitals and potential prevention strategies. *Qual Saf Health Care, 19*(5), e5. doi:10.1136/qshc.2008.030726

Soeken, K. L. (2010). Validity of measures. In C. F. Waltz, O. L. Strickland, & E. L. Lenz (Eds.), *Measurement in nursing and health* (pp. 163-201). New York, NY: Springer.

Sorra, J., & Dyer, N. (2010). Multilevel psychometric properties of the AHRQ hospital survey on patient safety culture. *BMC Health Services Research, 10*(1), 199. doi:10.1186/1472-6963-10-199

Sorra, J., Famolaro, T., Yount, N. D., Smith, S. A., Wilson, S., & Liu, H. (2014). *Hospital survey on patient safety culture 2014: User comparative database report*. Retrieved from http://www.ahrq.gov/professionals/quality-patient-safety/patientsafetyculture/hospital/2014/index.html

Sorra, J., Khanna, K., Dyer, N., Mardon, R., & Famolaro, T. (2012). Exploring relationships between patient safety culture and patients' assessments of hospital care. *Journal of Patient Safety, 8*(3), 131-139.

Sorra, J., & Nieva, V. (2004). *Hospital survey on patient safety culture*. AHRQ Publication No. 04-0041. Retrieved http://www.ahrq.gov/qual/patientsafetyculture/hospcult.pdf

Sparnon, E., & Marella, W. M. (2012). The role of the electronic health record in patient safety events. *Pennsylvania Patient Safety Advisory, 9*(4), 113-121. Retrieved from http://www.patientsafetyauthority.org/ADVISORIES/AdvisoryLibrary/2012/Dec;9(4)/Pages/113.aspx

Speroff, T., Nwosu, S., Greevy, R. A., Weinger, M. B., Talbot, T. R., Wall, R. J., . . . Dittus, R. S. (2010). Organisational culture: Variation across hospitals and connection to patient safety climate. *Quality and Safety in Health Care, 19*(6), 592-596. doi:10.1136/qshc.2009.039511

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical assessment, research and evaluation, 9*(4), 1-19. Retrieved from http://pareonline.net.proxy.lib.ohio-state.edu/getvn.asp?v=9&n=4

Steyer, J., Schiffinger, M., Huber, C., Valentin, A., & Strunk, G. (2012). Attitude is everything? The impact of workload, safety climate, and safety tools on medical errors: A study of intensive care units. *Health Care Management Review, 38*(4), 306-316. doi:10.1097/HMR.0b013e318272935a

Stommel, M., & Wills, C. E. (2004a). *Clinical research: Concepts and principles for advanced practice*. Philadelphia, PA: Lippincott.

Stommel, M., & Wills, C. E. (2004b). Judging the quality of measurement. In M. Stommel & C. E. Wills (Eds.), *Clinical research: Concepts and principles for advanced practice nurses* (pp. 208-228). Philadelphia, PA: Lippincott, Williams, and Wilkins.

Stone, P. W., Mooney-Kane, C., Larson, E. L., Horan, T., Glance, L. G., Zwanziger, J., & Dick, A. W. (2007). Nurse working conditions and patient safety outcomes. *Medical Care, 45*(6), 571-578.

Suarez, C., Menendez, M. D., Alonso, J., Castaño, N., Alonso, M., & Vazquez, F. (2014). Detection of adverse events in an acute geriatric hospital over a 6-year period using the Global Trigger Tool. *Journal of the American Geriatrics Society, 62*(5), 896-900. doi:10.1111/jgs.12774

Swayze, S. C., & Rich, S. E. (2012). Promoting safe use of medical devices. *Online Journal of Issues in Nursing, 17*(1). doi:10.3912/OJIN.Vol17No01PPT01

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.

Tamuz, M., & Harrison, M. I. (2006). Improving patient safety in hospitals: Contributions of high-reliability theory and normal accident theory. *Health Services Research, 41*(4 Pt 2), 1654-1676.

Tanaka, M., Tanaka, K., Takano, T., Kato, N., Watanabe, M., & Miyaoka, H. (2012). Analysis of risk of medical errors using structural-equation modelling: A 6-month prospective cohort study. *BMJ Quality & Safety, 21*(9), 784-790. doi:10.1136/bmjqs.2010.048330

Taylor, J. A., Dominici, F., Agnew, J., Gerwin, D., Morlock, L., & Miller, M. R. (2012). Do nurse and patient injuries share common antecedents? An analysis of associations with safety climate and working conditions. *BMJ Quality and Safety, 21*, 101-111. doi:10.1136/bmjqs-2011-000082

The Health Foundation Inspiring Improvement. (2010). *Global trigger tools. Evidence scans*. Retrieved from http://www.health.org.uk/public/cms/75/76/313/2601/Global%20trigger%20tools.pdf?realName=va7dtO.pdf

The Joint Commission. (2012) Improving patient and worker safety: Opportunities for synergy, collaboration and innovation. Oakbrook Terrace, IL: Author.

The Joint Commission. (2014). *Sentinel event data, root causes by event type, 2004: 2Q 2014*. Oakbrook Terrace, IL: Author.

The World Alliance for Patient Safety Drafting Group, Sherman, H., Castro, G., Fletcher, M., Hatlie, M., Hibbert, P. D., . . . Virtanen, M. (2009). Towards an international classification for patient safety: The conceptual framework. *International Journal for Quality in Health Care, 21*(1), 2-8. doi:10.1093/intqhc/mzn054

Thomas, E. J., & Petersen, L. A. (2003). Measuring errors and adverse events in health care. *Journal of General Internal Medicine, 18*(1), 61-67. doi:10.1046/j.1525-1497.2003.20147.x

Thompson, P. A., Navarra, M. B., & Antonson, N. (2005). Patient safety: The four domains of nursing leadership. *Nursing Economic$, 23*(6), 331-333.

Tindel, M. S., Darby, J. M., & Simmons, R. L. (2014). A retrospective review of crisis events in diagnostic radiology: An analysis of frequency, demographics, etiologies, and outcomes. *Journal Of Patient Safety, 10*(2), 111-116.

U.S. Department of Health and Human Services. (2013). *2013 annual progress report to Congress: National strategy for quality improvement in health care*. Retrieved from http://www.ahrq.gov/workingforquality/nqs/nqs2013annlrpt.htm

U.S. Department of Health and Human Services. (n.d.). *Spotlight on...Adverse events*. Retrieved from http://oig.hhs.gov/newsroom/spotlight/2012/adverse.asp

U.S. Federal Committee on Statistical Methodology. (2001). Measuring and reporting sources of error in surveys *Statistical policy*. Washington, D.C.: Office of Information and Regulatory Affairs Office of Management and Budget.

Unbeck, M., Schildmeijer, K., Henriksson, P., Jürgensen, U., Muren, O., Nilsson, L., & Härenstam, K. P. (2013). Is detection of adverse events affected by record review methodology? An evaluation of the "Harvard Medical Practice Study" method and the "Global Trigger Tool." *Patient Safety in Surgery, 7*(10). doi:doi:10.1186/1754-9493-7-10

University of Texas Center for Healthcare Quality and Safety. (2014). *Safety Attitudes and Safety Climate Questionnaire*. Retrieved from https://med.uth.edu/chqs/surveys/safety-attitudes-and-safety-climate-questionnaire/

Van Den Bos, J., Rustagi, K., Gray, T., Halford, M., Ziemkiewicz, E., & Shreve, J. (2011). The $17.1 billion problem: The annual cost of measurable medical errors. *Health Affairs, 30*(4), 596-603. doi:10.1377/hlthaff.2011.0084

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The Kappa statistic. *Family Medicine, 37*(5), 360-364.

Voepel-Lewis, T., Pechlavanidis, E., Burke, C., & Talsma, A. N. (2012). Nursing surveillance moderates the relationship between staffing levels and pediatric postoperative serious adverse events: A nested case-control study. *International Journal Of Nursing Studies, 50*(7), 905-913.

Vogt, W. P., & Johnson, R. B. (2011). *Dictionary of statistics and methodology: A nontechnical guide for the social sciences* (4th ed.). Washington DC: Sage.

Vogus, T. J., & Sutcliffe, K. M. (2007). The impact of safety organizing, trusted leadership, and care pathways on reported medication errors in hospital nursing units. *Medical Care, 45*(10), 997-1002.

vonPlessen, C., Kodal, A. M., & Anhoj, J. (2012). Experiences with global trigger tool reviews in five Danish hospitals: An implementation study. *BMJ Open, 2*(5), e001324. doi:10.1136/bmjopen-2012-001324

Wachter, R. M. (2012). *Understanding patient safety* (2nd ed.). New York, NY: McGraw Hill.

Walshe, K. (2000). Adverse events in health care: Issues in measurement. *Quality in Health Care, 9*(1), 47-52.

Waltz, C. F., Strickland, O. L., & Lenz, E. R. (2010). *Measurement in nursing and health research* (4th ed.). New York, NY: Springer.

Watts, B. V., Percarpio, K., West, P., & Mills, P. D. (2010). Use of the safety attitudes questionnaire as a measure in patient safety improvement. *Journal of Patient Safety, 6*(4), 206-209.

Weick, K. E., & Sutcliffe, K. M. (2001). *Managing the unexpected: Assuring high performance in an age of uncertainty*. San Francisco, CA: Jossey-Bass.

Weinberg, J., Hillborne, L. H., & Nguyen, Q. T. (2005). Regulation of health policy: Patient safety and the states. In K. Henriksen, J. B. Battles, & E. S. Marks (Eds.), *Advances in patient safety: From research to implementation* (Vol. 1: Research findings). Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from http://www.ncbi.nlm.nih .gov/books/NBK20479/?report=reader

Weller, J., Frengley, R., Torrie, J., Shulruf, B., Jolly, B., Hopley, L., . . . Paul, A. (2011). Evaluation of an instrument to measure teamwork in multidisciplinary critical care teams. *BMJ Quality & Safety, 20*(3), 216-222. doi:10.1136/bmjqs.2010.041913

West, E. (2000). Organisational sources of safety and danger: Sociological contributions to the study of adverse events. *Quality in Health Care, 9*, 120-126.

Wetterneck, T. B., Walker, J. M., Blosky, M. A., Cartmill, R. S., Hoonakker, P., Johnson, M. A., . . . Carayon, P. (2011). Factors contributing to an increase in duplicate medication order errors after CPOE implementation. *Journal of the American Medical Informatics Association, 18*(6), 774-782. doi:10.1136/amiajnl-2011-000255

Wong, B. M., Dyal, S., Etchells, E. E., Knowles, S., Gerard, L., Diamantouros, A., . . . Shojania, K. G. (2015). Application of a trigger tool in near real time to inform quality improvement activities: A prospective study in a general medicine ward. *BMJ Quality & Safety, 24*(4), 272-281. doi:10.1136/bmjqs-2014-003432

Woods, D., Barnard, C., Sikka, V. K., & Dhindsa, H. S. (2015). Risk mitigation: Implications for patient safety. In S. P. Stawicki, S. C. Galwankar, T. J. Papadimos, & S. D. Moffatt-Bruce (Eds.), *Fundamentals of patient safety in medicine and surgery* (pp. 14-25). Philadelphia, PA: Lippincott, Williams, and Wilkins.

X-Rates. (2015). *Historic lookup (exchange rates)*. Retrieved from http://www.x-rates.com/ historical/?from=EUR&amount=1&date=2015-01-01

Yiannakis, A. (1997). *Sources of invalidity in experimental and non-experimental designs*. Retrieved from http://www.sp.uconn.edu/~yian/reserve/invalid.htm

Young, M. (2013). Are you putting your patients at risk? *AAMI News, 48*(8). Retrieved from http://www.aami.org/publications/AAMINews/Aug2013/Immediate_Use_Steam_ Sterilization.html

Zegers, M., deBruijne, M. C., deKeizer, B., Merten, H., Groenewegen, P. P., VanderWal, G., & Cordula, W. (2011). The incidence, root-causes, and outcomes of adverse events in surgical units: Implication for potential prevention strategies. *Patient Safety in Surgery, 5*(13). doi:10.1186/1754-9493-5-13

Zimmermann, R., Pierson, S., McLean, R., McAlpine, S. A., Caron, C., Morris, B., & Lucas, J. (2010). Aiming for zero preventable deaths: Using death review to improve care and reduce harm. *Healthcare Quarterly, 13*, 81-87.

Zohar, D. (1980). Safety climate in industrial organizations: Theoretical and applied implications. *Journal of Applied Psychology, 65*(1), 96-102.

Zohar, D., Livne, Y., Tenne-Gazit, O., Admi, H., & Donchin, Y. (2007). Healthcare climate: A framework for measuring and improving patient safety. *Critical Care Medicine, 35*(5), 1312-1317.

Zuckerman, S. L., Parikh, R., Moore, D. C., & Talbot, T. R. (2012). An evaluation of immediate-use steam sterilization practices in adult knee and hip arthroplasty procedures. *American Journal of Infection Control, 40*(9), 866-871. doi:10.1016/j.ajic.2011.11.006

Appendix: Clarification of GTT Definitions

Clarification of GTT Definitions

Module	GTT Trigger	GTT Trigger	Definition/Clarification	Further Explanation
Cares	C1	Transfusion or use of blood products	Any transfusion of packed red blood cells (PRBCs) or whole blood should be investigated for causation, including excessive bleeding (surgical or anticoagulation-related), unintentional trauma of a blood vessels, etc. Transfusion of many units beyond expected blood loss within the first 24 hours of surgery, including intra-operatively and post-operatively will likely be related to a perioperative AE. Cases in which excessive blood loss occurs pre-operatively are not typically AEs. Patients receiving anticoagulants who require transfusions of fresh frozen plasma and platelets have likely experienced an AE related to the use of anticoagulants.	Florida Hospital FAQs document (2009) states 2 or more units of PRBCs is within normal limits for most surgeries; up to 4 units PRBCs for cardiovascular surgery
Cares	C2	Code/ arrest/ rapid response team activation	All emergency codes should be investigated, however not all codes are AEs. Some may be related to progression of disease. Check for medication-related issues. Cardiac or pulmonary arrest intra-operatively or in PACU should be considered an AE. However, a sudden cardiac arrhythmia resulting in a cardiac arrest may not be an AE but related to cardiac disease. Failure to recognize signs and symptoms would be an example of an error of omission and would not be counted as an AE unless the changes in patient condition were the result of some medical intervention.	
Cares	C3	Acute Dialysis	A new need for dialysis may be the course of a disease process of the result of an AE (i.e. drug-induced renal failure or reaction to the administration of a dye for radiological procedures).	
Cares	C4	Positive blood culture	A positive blood culture should be investigated as an indicator of an AE – specifically a healthcare associated infection. Generally AEs associated with this trigger include infections diagnosed 48 hours or more after admission: blood stream infections, sepsis from other device infections (catheter associated UTI) or any other healthcare associated infection. Positive blood cultures related to other diseases (Community Acquired Pneumonia) would not be considered to be an AE.	
Cares	C5	X-Ray or Doppler Studies for Emboli or Deep Vein Thrombosis	Trigger includes CT scan to diagnose pulmonary embolus. DVT/PE in most cases is an AE. Rare exceptions are those related to disease process (cancer/clotting disorders). Look for causation prior to admission that could be attributed to medical care such as prior surgical procedure. Lack of prophylaxis with no DVT or PE is not an AE (it is an error of omission).	
Cares	C6	Decrease of greater than 25% in hemoglobin or hematocrit	Look for potential causes such as excessive bleeding (surgical or anticoagulation-related), or unintentional trauma of a blood vessels etc. Did the decrease occur in a relatively short period of time (72 hours or less)? Bleeding events may be related to use of anticoagulants or aspirin or a surgical complication. The decrease in HgB or Hct in itself is not an AE unless related to some medical treatment. A decrease associated with a disease is not an AE.	
Cares	C7	Patient Fall	If the fall occurred as a result of administration of a sedative, analgesic, or muscle relaxant, an AE has occurred.	

Cares	C8	Pressure Ulcers	(Documented under "Tissue Integrity" in nurses notes). Pressure ulcers are AEs. Chronic decubiti are AEs if they occurred during a hospitalization.	Stage II or higher are AEs. Consult with wound ostomy RN if staging unclear in documentation, or refer to pressure ulcer practice guidelines.
Cares	C9	Readmission within 30 days	Any re-admission within 30 days of discharge could be an AE (such as a surgical site infection, DVT, PE).	
Cares	C10	Restraint use	Review the documented reasons for restraint use and evaluate the relationship between use of restraints and confusion from drugs etc. (which would indicate an AE).	
Cares	C11	Healthcare-Associated Infections	Any infection that starts 48 hours after a hospital admission is considered an AE (whether associated with a Foley catheter or not). Determine if the infection is related to medical care (prior procedure, urinary catheter at home or in long-term care) versus naturally occurring disease (community acquired pneumonia).	
Cares	C12	In-hospital Stroke	Evaluate the cause of the stroke to determine whether it is associated with a procedure (e.g., surgical, conversion of Atrial Fibrillation) or anticoagulation. When procedure or treatments have likely contributed to stroke this is an AE.	
Cares	C13	Transfer to Higher Level of Care	Transfer to a higher level of care/rapid response team/arrest. Transfer to a higher level of care includes transfers within the institution, to another institution from yours, or to yours from another institution. A higher level of care is indicated when a patient's clinical condition deteriorates and a rapid response team is called. Look for the reason for the transfer or the change in condition.	
Cares	C14	Any Procedure Complication	A complication resulting from any procedure is an AE. Procedure notes frequently do not indicate complications, especially if they occur hours or days after the procedure note has been dictated. Watch for complications noted in coding or discharge summary or other progress notes. Post-operative Ileus: If an intervention was required, and this delayed patient recovery (such as progression to eating, mobility, etc.), then it is an AE.	
Cares	C15	Other	Use this category if an AE is uncovered that does not fit a trigger. An event does not require a listed trigger to be counted as an event.	
Medica-tion	M1	*Clostridium Difficile* Positive Stool	A positive stool sample for C. difficile is a likely complication for patients on multiple antibiotics and an indication of an AE. Look for a history of antibiotic use.	
Medica-tion	M2	Partial Thrombo-plastin Time > 100 seconds	High PTT is a frequent occurrence but an AE has only occurred if evidence of a bleed is present when patients are on heparin. Look for evidence of bleeding to determine if an AE has occurred. Elevated PTT in itself is not an AE – there must be manifestation such as bleeding, drop in Hg or Hct or bruising.	
Medica-tion	M3	International Normalized Ratio (INR) > 6	High INR is also a frequent occurrence but as with heparin, an AE has only occurred if evidence of a bleed is present secondary to Warfarin.	
Medica-tion	M4	Glucose less than 50 mg/dL	Low serum glucose does not necessarily indicate an AE has occurred. Determine if hypoglycemia occurred secondary to insulin administration. Check nurses' notes for signs and symptoms (e.g., lethargy, shakiness), administration of glucose or orange juice or other interventions. If symptoms are present, look for	

			administration of insulin or oral hypoglycemic medications. If the patient is not symptomatic, there is no AE.	
Medica-tion	M5	Rising BUN or serum creatinine > 2 times baseline	Consider several sequential results to see whether serum creatinine levels rose more than twice from baseline. If this rise can be correlated to a nephrotoxic medication, and interventions were required to correct renal problems, then an AE has occurred. Review physician progress notes and the H & P for other causes of renal failure (such as pre-existing renal disease or diabetes) that could have put the patient at greater risk for renal failure. This would not be an AE but rather the progression of disease. A two-fold increase in serum creatinine is consistent with the (Risk, Injury, and Failure) RIFLE criteria for kidney 'injury.'	The AE must correspond with the department/ procedure area for which the record is selected, and where the contrast/ nephrotoxic medication was administered.
Medica-tion	M6	Vitamin K (Phyton-adione) admin-istration	If this trigger is found, check the progress notes for evidence of excessive bruising, GI bleed, hemorrhagic stroke, or large hematoma as examples of AEs. Any bleed related to anticoagulants (e.g. Warfarin, Enoxaparin, Heparin, Dabagatrin or even Aspirin) is an AE. An AE has likely if there are lab reports indicating a drop in Hematocrit or guaiac-positive stools.	
Medica-tion	M7	Benadryl (Diphen-hydramine)/ Anti-histamines)/ IV Cortico-steroid use)	Antihistamines can be used for conditions such as sinusitis, seasonal allergies, pre-op/pre-procedure medication, or as a sleep aid etc., so try to find the reason why this medication was administered. If the trigger denotes an allergic reaction to a medication, it may be prescribed as a stat dose. Review nurses notes to determine if the antihistamine was ordered secondary to an allergic reaction caused by a medication. Whether it is a ***one-time or scheduled*** antihistamine dose, if an allergic reaction secondary to a medication or blood transfusion has occurred, then it is counted as an AE.	
Medica-tion	M8	Romazicon (Flumazenil) use	Flumazenil reverses the effects of benzodiazepines such as Diazepam. Prolonged sedation or marked hypotension is the AEs caused by benzodiazepines.	
Medica-tion	M9	Naloxone (Narcan) use	Naloxone is a powerful narcotic antagonist. Usage likely represents an AE except in cases of drug abuse of self-inflicted overdose.	
Medica-tion	M10	Anti-emetic use	Nausea and vomiting commonly are the result of drug administration both in surgical and non-surgical settings (due to opiates and antibiotics). Nausea and vomiting that interferes with feeding, post-operative recovery or delayed discharge suggest an AE. One episode treated successfully with anti-emetics would suggest no AE. ***Two*** or more episodes are considered an AE, because it is ongoing, causing patient harm.	Nausea must correspond with two or more doses of anti-emetic on the unit for which the record is selected.
Medica-tion	M11	Over-sedation/ hypotension	Review the physician progress notes, nursing or multidisciplinary notes for evidence of over sedation and lethargy. Review vital signs records or graphics for episodes of hypotension related to the administration of a sedative analgesic or muscle relaxant. Intentional overdose is not considered an AE.	
Medica-tion	M12	Abrupt medication stop	Look for this trigger in the medication administration record (MAR). Look for the reason the medication has been stopped or withheld. A sudden change in patient condition requiring adjustment of medications is often related to an AE. 'Abrupt' is best described as an unexpected stop or deviation from typical ordering practice. For example, discontinuation of an IV antibiotic for switch to oral is not unexpected.	

Medica-tion	M13	Other	Use this trigger for ADEs detected but not related to one of the medication triggers listed above.
Surgical	S1	Return to Surgery	A return to surgery can either be planned or unplanned and both can be a result of an AE. An example of an AE would be a patient who had internal bleeding following the first surgery and required a second surgery to explore for the cause and stop the bleeding. Even if the second surgery is exploratory but reveals no defect, this should be considered an AE.
Surgical	S2	Change in Procedure	When the procedure indicated in the post-operative notes is different from the procedure planned in the pre-operative notes or documented in the surgical consent, a reviewer should look for details as to why the change occurred. An unexpected change in procedure due to complications or device or equipment failure should be considered an AE, particularly if LOS increases or obvious injury has occurred.
Surgical	S3	Admission to Intensive Care post-operatively	Admission to ICU can be either a normal post-operative journey or it may be unexpected. The unexpected admissions frequently are related to operative AEs. For example, admission to intensive care following aortic aneurysm repair may be expected, but admission following knee replacement would be unusual. The reviewer needs to determine why the admission occurred.
Surgical	S4	Intubation/ re-intubation/ BiPap in Post Anesthesia Care Unit	Anesthesia, sedatives or pain medications can result in respiratory depression requiring the use of BiPap or re-intubation post-operatively which would be an AE.
Surgical	S5	X-ray intra-op or in PACU	Imaging of any kind that is not routine for the procedure requires investigation. An x-ray taken due to suspicion of retained items or incorrect instrument count or sponge count would be a positive trigger. The identification of a retained item necessitating an additional procedure is an AE. If the retained item is identified and removed without additional evidence of harm or re-operation (i.e. while patient is “open”), this is not considered an AE.
Surgical	S6	Intra-op or post-op death	All deaths that occur intra-operatively should be considered AEs unless death is clearly expected and the surgery was of a heroic nature. Post-operative deaths will require review of the record for specifics, but in general all post-op deaths will be AEs.
Surgical	S7	Mechanical ventilation longer than 24 hours post-op	Short-term mechanical ventilation post-operatively for cardiac, major thoracic and certain abdominal procedures is planned. If the patient required mechanical ventilation beyond 24 hours, an intra-operative or post-operative AE should be considered. Patients with pre-existing pulmonary or muscular disease may experience more difficulty in quickly weaning from a ventilator post-operatively but this should not automatically exclude the possibility of an AE. Reviewers must use clinical judgment to determine whether the intra-operative and post-operative care was event-free or part of the disease process.
Surgical	S8	Intra-op epinephrine, nor-epinephrine, Naloxone, or Romazicon	These medications are not routinely administered intra-operatively. Review anesthesia and operative notes to determine the reasons for administration. Hypotension caused by bleeding or over-sedation are examples of AEs that might be treated with these medications.
Surgical	S9	Post-op troponin level	A postoperative increase in troponin levels may indicate a cardiac event. Reviewers will need to use clinical judgment as to whether a cardiac event has occurred.

		greater than 1.5 ng/mL	
Surgical	S10	Injury, repair, or removal of organ	Review operative notes and post-operative notes for evidence that the procedure included repair or removal of any organ. The removal or repair must be part of the planned procedure or this is an AE and likely the result of a surgical misadventure such as accidental injury.
Surgical	S11	Any operative complication	This refers to any number of complications, including but not limited to PE, DVT, decubiti, MI, renal failure etc.
ICU	I1	Pneumonia onset	If evidence suggests pneumonia started prior to admission to the hospital, there is no AE, but if the review suggests initiation in the hospital, it is an AE. In general, any infection starting in not only the ICU but in any hospital unit will be considered nosocomial. Re-admissions either to the hospital or the ICU could represent a nosocomial infection from a previous hospital admission.
ICU	I2	Readmission to ICU	Admission to ICU can be either a normal post-operative journey or it may be unexpected. The unexpected admissions frequently are related to operative AEs. For example, admission to intensive care following aortic aneurysm repair may be expected, but admission following knee replacement would be unusual. The reviewer needs to determine why the admission occurred.
ICU	I3	In-unit procedure	Any procedure occurring on a patient in the ICU requires investigation. Look at all bedside procedures and the procedures done while the patient was in ICU. Complications will commonly not be on the dictated procedure note, but may be evidence by the care required which might indicate an event has occurred.
ICU	I4	Intubation/ reintubation	Anesthesia, sedatives or pain medications can result in respiratory depression requiring the use of BiPap or re-intubation post-operatively which would be an AE.
Perinatal	P1	Terbutaline use	Use of Terbutaline could result in an unnecessary intervention of a C-section. Look for complicating factors. Use of Terbutaline in pre-term labor is *not* a positive trigger.
Peri-natal	P2	3rd- or 4th-degree lacerations	By definition a 3rd - or 4th- degree laceration is an AE. Look for additional events to the mother or child associated with the laceration as a part of a cascade to assess AE severity.
Peri-natal	P3	Platelet count less than 50,000	Look for AEs related to bleeding such as strokes, hematomas, and hemorrhage requiring blood transfusions. Look for information regarding why the platelet count dropped to see if it was a results of a medication. Usually, a platelet transfusion is an indication that the patient has a low platelet count. Events related to transfusions or bleeding may indicate that an AE has occurred.
Peri-natal	P4	Estimated blood loss > 500 ml (vaginal) or > 1,000 ml (C-section)	The accepted limit for “normal” blood loss after vaginal delivery is 500 ml, and a blood loss of 1,000 ml is considered WNL after a cesarean birth.
Peri-natal	P5	Specialty consult	May be an indicator of injury or other harm.
Peri-natal	P6	Oxytocic agents (such as oxytocin, methyl-ergonovine, and 15-methyl-	These are agents used to control post-partum hemorrhage (> 500 ml for vaginal, > 1,000 ml C-section). If standard administration of oxytocin occurs post-delivery, evaluate for administration amounts greater than 20 units in the immediate post-partum period.

		prosto-glandin in the post-partum period)	
Peri-natal	P7	Instrumented delivery	Instruments may cause injury to the mother, including bruising, trauma, and perineal lacerations.
Peri-natal	P8	General anesthesia	May be an indicator of harm resulting from poor planning or other sources of harm
Emer-gency	E1	Readmission to Ed within 48 hours	Look for drug reactions, infection or other reasons that events may have brought the patient back to the ED and then required admission.
Emer-gency	E2	Time in ED greater than 6 hours	Long ED stay can represent less than optimal care. Look for complications arising from the ED such as falls, hypotension, or procedure related complications.

Note. Modified from: New Zealand Health Quality & Safety Commission. (2012). *Global Trigger Tool implementation guide*. Wellington, New Zealand: Health Quality & Safety Commission. Retrieved from: http://www.hqsc.govt.nz/our-programmes/other-topics/publications-and-resources/publication/690/
AE is adverse event.