Title
Effect of Nursing Interventions on Physiologic Monitor Alarm Rates

Author
Mammone, Tina

Publication Date
2014

Peer reviewed|Thesis/dissertation

Effect of Nursing Interventions on Physiologic Monitor Alarm Rates in a Neuroscience Intensive Care Unit

By

Tina Mammone, RN, MS, $PhD_{(c)}$

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Nursing

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Dedication

My first and foremost thanks go out to my family—Mom, Dad, Levac, and my sister Silvia. My heart is filled with gratitude for your unconditional love, words of encouragement, and unwavering support over the years. Thank you for believing in me and supporting me.

To Tracey Mulholland, my best friend for over twenty years: thank you for the late night phone calls, making me laugh, numerous reality checks, and cheering me on every step of the way.

Thank you all for understanding when I was not available as much as we would have liked throughout the past few years.

Acknowledgements

The following, mentors, friends, colleagues, co-workers, and students have made this five-year journey a rewarding learning experience; your patience, guidance, and unwavering support has meant the world to me.

Barbara J. Drew, Advisor, mentor, Chair of Dissertation Committee and Qualification Examination Committee Member. I cannot thank you enough for taking a chance with me when I was “homeless” doctoral student. I will never forget your welcoming me into the Drew Lab and the Department of Physiological Nursing on March 10, 2011—just over three years ago. Your steadfast guidance, invaluable instruction, tireless support, dedication to students, and your zest for life have deeply inspired me. You have left an incredible and permanent impression on me both professionally and personally—for which I will be forever thankful.

Xiao Hu, Qualification Examination Committee Member, for providing indispensable instruction and providing me the assistance of your talented bioengineering team—without your team’s skill, expertise, and hard work this analysis would not have been possible. We are all so fortunate that you chose to be a professor at the UCSF School of Nursing.

Daphne Stannard, Qualification Examination Committee Member, for her generosity, guidance, and encouragement both in my academic pursuits and in my position as the Patient Care Director of the Cardiovascular, Medicine, and Oncology Services at UCSF Medical Center.

Kathy Dracup, Chair of the Qualification Committee, for her wise counsel, kindness, and support over the years.

John Hoffman “Tweakmaster” for his incredible generosity and sponsorship in providing clinical application and technical support from Excel-Medical Electronics. This study would not have been feasible without his goodwill, collaboration, and quiet dedication to improving the care of monitored patients.

Stephen Treacy for sharing his wealth of knowledge and expertise regarding GE Solar 8000i monitors and physiologic monitor alarms. The guidance I received over countless emails, texts, and phone calls is much appreciated—your responsiveness 24/7 cannot be surpassed.

Steven Paul, who makes statistics sound so easy when I’m in his presence and for providing me invaluable instruction throughout the statistical analysis.

UCSF School of Nursing, Department of Physiological Nursing Faculty, Staff, students, and all members of the UCSF Alarm Study for your tremendous assistance, flexibility, patience, and the collaboration you have each shown me as I learned how to conduct nursing research.

UCSF Medical Center leadership community who have provided me support throughout my doctoral studies and are patiently waiting for me to finish.

UCSF Medical Center, Neuroscience ICU nursing leadership and nursing staff who welcomed me, my research study, and for the amazing support I received over the years.

The staff at H. M. Fishbon Memorial Library, UCSF Medical Center at Mount Zion for your incredible assistance in meeting my medical information needs. Special thanks to Gloria Won for helping me perform literature searches and tracking down hard to find research articles and books.

Bruce Smith, editor extraordinaire, for teaching me about em dashes, sentence logic, and avoiding the use of stand-alone pronouns and anthropomorphisms.

Effects of Nursing Interventions on Physiologic Monitor Alarm Rates in a Neuroscience Intensive Care Unit

Tina Mammone

Abstract

Introduction: Physiologic monitors play a vital role in saving patients' lives but expose clinicians to an overwhelming number of alarms, many of which are false. **Objective:** Our study aims were; a) to determine whether the mean hourly oxygen saturation (SpO_2) low-limit alarm rates could be reduced by modifying the default alarm setting and b) to determine whether there would be a reduction in mean hourly technical and critical arrhythmia alarm rates and the mean percentage of false-positive critical arrhythmia alarms following daily skin preparation and the application of high-quality ECG electrodes. **Methods:** We conducted a prospective randomized clinical trial in two neuroscience intensive care units, collecting data during two assessment periods. Each patient's alarm rate was calculated as the number of unique alarms divided by monitoring time. Critical arrhythmia alarms were determined (true vs. false) using a standardized protocol. Means and standard deviations of the hourly alarm rates and the mean percentage of false-positive critical arrhythmia alarms were determined during both assessments. A negative binomial regression was performed to test the main effect of unit, the main effect of assessment, and the unit by assessment interaction. **Results:** The combined use of a lower SpO_2 low-limit threshold and increased alarm delay resulted in a significant unit-by-assessment interaction ($p <$ .001). During Assessment 2, the experimental unit had a lower mean hourly SpO_2 alarm rate while in the control unit, the rate increased. No significant unit-by-assessment interactions were observed for the mean hourly technical alarm rates; during Assessment 2, both units experienced an increase in ECG lead fail alarm rates and although both units had a reduction in mean hourly

artifact alarms, it was insignificant. Similarly, no significant unit-by-assessment interaction in the mean hourly critical arrhythmia alarm rate was observed; no reduction in critical arrhythmia alarm rates was found in the experimental unit. Likewise, the intervention did not reduce the mean percentage of false-positive critical arrhythmia alarms. **Conclusion:** A lower SpO_2 alarm limit and increased alarm delay safely reduces non-actionable alarms. However, our novel electrode regimen does not reduce critical arrhythmia and technical alarm rates or false-positive arrhythmia alarms.

Table of Content

Dedication……..... iii

Acknowledgements…….. iv

Abstract…… vii

Chapter 1 *Introduction*……..... 1

Statement of the Problem……... 2

Purpose of the Study……... 4

Dissertation Chapters……. 8

References…….. 9

Chapter 2 Part 1 *Literature Review*…….. 10

Introduction……..... 10

Methods…….. 13

Search Strategies……. 13

Review of Current Research……... 14

Physiologic Monitor Alarms……..... 14

Causes of Alarms…… 15

Pediatric Studies…….. 15

Adult Studies……... 16

Parameter Alarms……... 17

Technical Alarms…… 23

Arrhythmia Alarms……..... 25

Sensitivity and Specificity…….. 31

Strategies for Optimizing Physiologic Monitor Alarms……. 32
Alarm Threshold Settings…… 33
Alarm Delay Setting…… 38
Patient Monitoring Supplies and Clinical Practice…….. 40
Skin Preparation…….. 40
ECG Electrodes……... 42
Impact of Alarm Fatigue……..... 43
Impact of Alarms on Clinicians……... 44
Impact of Alarms on Patients…….. 48
Impact of Alarms on Health Care Organizations……..... 50
Limitations of Current Research……..... 52
Suggestions for Future Research…… 53
Conclusion…….. 54
Summary Table…….. 57
Abbreviations……. 86
References…….. 87
Chapter 2 Part 2 *Theoretical Considerations*…….. 99
Background……..... 99
Vigilance Research……..... 102
Theoretical Framework…….. 105
Habituation: A Dual Process Theory…… 105
Effect of Stimulus Frequency…… 107

Effect of Stimulus Change……. 109

Effect of Stimulus Intensity……... 111

Effect of Time…… 112

Vigilance Research……... 113

Signal Salience……... 113

Signal Modality…….. 116

Stimulus Complexity……. 117

Conclusion…… 119

References……. 124

Chapter 3 *Design and Measurement Issues*……..... 134

Background……... 134

Study Designs……... 135

Qualitative Research…….. 135

Quantitative Research…… 136

Observational Study Designs ……..…….. 136

Cohort…….. 136

Cross-sectional……. 139

Quasi-experimental…… 142

Randomized Clinical Trial……..…….. 143

Summary of Study Designs…….. 144

Dimensions of Data Collection Approaches……. 145

Structure …… 145

Quantifiability…… 146
Researcher obtrusiveness……... 147
Objectivity…….. 147
Data Collection Methods…….. 148
Observation…… 148
Observer Bias…… 150
Hawthorne Effect……... 150
Ethical Issues……. 151
Information Technology…… 152
Researcher and Vendor Supplied Software……... 152
Vendor Supplied Software with Visual Records……... 153
Integration Software…….. 156
BedMasterEx……..... 157
Methodological Limitations…….. 159
Design and Instrument Selection…….. 163
Contextual Issues…….. 164
Compensation Equalization and Rivalry……..... 164
Diverse Participants and Patient Placement …….. 166
Conclusion…… 167
References……. 169
Chapter 4 *Method*s…….. 176
Introduction……... 176

Research Design……..... 176

Research Setting……..... 176

Research Sample……... 176

Timeline…… 177

Human Subjects Assurance……... 179

Confidentiality…….. 179

Data Safety Monitoring Board……. 180

Risks to Human Subjects…….. 180

Intervention: SpO_2 Low-limit Threshold Alarm and SpO_2 Alarm Delay……. 181

SpO_2 Low-Limit Threshold Alarm…… 181

SpO_2 Alarm Delay…….. 185

Intervention: Skin Preparation and Electrodes…….. 186

Technical Alarms……... 186

Arrhythmia Alarms…… 188

Electrode-site Skin Preparation…….. 190

ECG Electrodes…….. 191

Application……. 192

Storage……... 192

Availability of New Patient Monitoring Supplies……. 193

Training ……. 195

Data Collection…….... 195

Instruments……. 195

Alarm Data……..... 197

Arrhythmia Annotation…….. 198

Patient Demographic Variables……. 199

Patient Outcome Variables……..... 200

Administrative information…… 201

Training Plan…….. 201

Missing Data……..... 202

Data Analysis and Statistics…….. 203

Demographic and Outcome Variables …….. 207

Administrative Information……... 208

Funding…… 208

Summary……... 208

References……. 210

Chapter 5 *Results*…… 213

Patient Demographics……... 213

Cardiac Monitoring Time……. 216

Compliance with New Default SpO_2 Low-Limit Alarm Settings …… 216

Compliance with Novel Electrode-site Skin Preparation and Electrodes……. 218

Patient Acuity……..... 221

Patient Outcome Data……... 224

SpO_2 Low-Limit Alarms……... 226

Technical Alarms…….. 228

ECG Lead Fail Alarms……. 228

Artifact Alarms…….. 230

Arrhythmia Suspend Alarms…….. 232

Arrhythmia Alarms…… 232

Audible Alarms…….. 236

Physiologic Monitor Alarms…….. 237

False-positive Cardiac Arrhythmia Alarms……..... 239

Power Analysis…….. 245

Reference……... 247

Chapter 6 *Discussion*……... 248

Objective of Dissertation ……... 248

Purpose of Intervention…….. 248

Technology-based Nursing Intervention…… 249

Practice-based Nursing Intervention…….. 254

Summary of Mean Hourly Alarm Outcome Variables…… 267

Review of Study Findings……. 271

Limitations……. 273

Implications for Nursing……...…. 275

Implications for Research…….. 276

Conclusion……. 279

References…….. 282

Glossary of Alarm Terminology…… 287

References…….. 289

List of Tables

Chapter 2 Part 1

Table 2.1. Physiologic Monitor Alarms: Summary …… 57

Chapter 4

Table 4.1. Physiologic Monitor Alarms by Alarm Type (Pilot Data)……. 178

Table 4.2. Alarm Study Timeline…… 179

Table 4.3. Parameter Alarms (Pilot Data)…… 182

Table 4.4. 8 and 11 NICU Default Parameter Alarm Threshold Limits……. 184

Table 4.5. 8 and 11 NICU Default Arrhythmia and Parameter Alarm Severity Levels……. 184

Table 4.6. Summary of Intervention: Modification of SpO_2 Alarm Settings……. 185

Table 4.7. Smart Lead Feature……...…. 187

Table 4.8. Common Technical Alarms (Pilot Data)…… 188

Table 4.9. Arrhythmia Alarms (Pilot Data)…….... 190

Table 4.10. Use and Storage of Patient Monitoring Supplies……. 193

Table 4.11. Summary of Intervention: Skin Preparation and Electrode Application……. 193

Table 4.12. Alarm Study: Patient Variables……...……. 200

Chapter 5

Table 5.1. Patient Demographics……..... 215

Table 5.2. Total Cardiac Monitoring Time in Hours for Patients in the Experimental and Control Unit……. 216

Table 5.3. Mean Cardiac Monitoring Time in Hours for Patient in the Experimental and Control Unit……. 216

Table 5.4. Compliance with New Default SpO_2 Alarm Settings ……..... 217

Table 5.5. Compliance with Novel Skin Preparation and Daily Application of ECG Electrodes…… 220

Table 5.6. Patient Outcomes for the Experiment Unit……...……. 225

Table 5.7. Patient Outcomes for the Control Unit……... 225

Table 5.8. Mean Hourly SpO_2 Low-limit Alarm Rates for Patients in the Experimental Unit and Control Unit……. 226

Table 5.9. Prevalence of SpO_2 Low-limit Alarms for Patients in the Experimental and Control Unit…… 227

Table 5.10. Mean Hourly Technical Alarm Rates for Patients in the Experimental and Control Unit…… 229

Table 5.11. Prevalence of Technical Alarms for Patients in the Experimental and Control Unit……..... 232

Table 5.12. Mean Hourly Arrhythmia Alarm Rates for Patients in the Experimental and Control Unit…… 233

Table 5.13. Types of Patient Status Arrhythmia Alarms and Default Alarm Notification Levels……. 234

Table 5.14. Prevalence of Alarms for Patients in the Experimental and Control Unit…….. 235

Table 5.15. Mean Hourly Rates for Critical Arrhythmia Alarms for Patients in the Experimental and Control Unit…….. 235

Table 5.16. Types of GE Physiologic Monitor Alarms…….. 238

Table 5.17. Mean Percent of False-positive Critical Arrhythmia Alarms for Patients in the Experimental and Control Unit……..... 243

Table 5.18. Comparison of Mean Percent False-positive Critical Arrhythmia Alarms for Patients in the Experimental Unit during Assessment 1 and 2……... 243

Table 5.19. Comparison of Mean Percent False-positive Critical Arrhythmia Alarms for Patients in the Control Unit during Assessment 1 and 2……..... 244

Table 5.20. Prevalence of Critical Arrhythmia Alarms for Patients in the Experimental and Control Unit..... 244

List of Figures

Chapter 2 Part 1

Figure 2.1. Epidermis Cross Section…….. 85

Chapter 2 Part 2

Figure 2.2. Schematic Illustration of the Phenomenon of Habituation and Sensitization…… 122

Figure 2.3. Schematic Illustration of Anticipated Changes in Physiologic Monitor Alarm Characteristics Following an Alarm Management Intervention…… 123

Chapter 4

Figure 4.1. Alarm Notification and Severity Levels…….. 182

Figure 4.2 Instructions for Use of New Patient Monitoring Supplies……. 194

Figure 4.3. 8 and 11 NICU Alarm Study Connectivity Diagram……..... 196

Figure 4.4. Illustration of Process for Alarm Annotation Using BedMasterEx Software…….. 196

Figure 4.5 Physiologic Monitor Alarm Categories for 8 and 11 NICU Alarm Study……... 199

Figure 4.6. Patient Admit/Discharge Screen…….. 203

Chapter 5

Figure 5.1. Compliance with New Default SpO_2 Alarm Setting……..... 218

Figure 5.2. Compliance with Novel Skin Preparation and Daily Application of ECG Electrodes……... 221

Figure 5.3. Acuity Summary Report for Experimental Unit during Assessment 1…… 222

Figure 5.4. Acuity Summary Report for Experimental Unit during Assessment 2……..... 223

Figure 5.5. Acuity Summary Report for Control Unit during Assessment 1……. 223

Figure 5.6. Acuity Summary Report for Control Unit during Assessment 2……. 224

Figure 5.7. Mean Hourly Rates for SpO_2 Low-Limit Alarms…… 227

Figure 5.8. Mean Hourly Rates for ECG Lead Fail Alarms …….. 229

Figure 5.9. Mean Hourly Rates for Artifact Alarms…….. 230

Figure 5.10. Mean Hourly Rates for Arrhythmia Suspend Alarms…… 231

Figure 5.11. Mean Hourly Rates for Arrhythmia Alarms ……..... 234

Figure 5.12. Mean Hourly Rates for Audible Alarms ……... 236

Figure 5.13. Mean Hourly Rates for Physiologic Monitor Alarms ……... 239

List of Appendices

Appendix A. Physiologic Monitor Alarms: Annotation Plan……..…… 290

Appendix B. Alarm Study Plan for Counting, Analyzing, and Monitor Alarms……... 292

Chapter 1

Introduction

The nurse administrator of the future must have a diverse and multifaceted background. Through my experience in direct patient care, subspecialty nursing in multiple settings, international nursing assignments, management of multiple patient care units across two facilities, and oversight of large initiatives in patient care services, I have become a broadly qualified nurse. I have worked for years to develop the skills and competencies that are necessary to be a nurse leader. Through my graduate studies and work experience, I have gained deep appreciation of the importance of ongoing learning about research; for nurse leaders, this research knowledgeability is essential for ensuring high-quality patient care.

This year, the Emergency Care and Research Institute (ECRI, 2013) identified alarm hazards as the top health technology hazard for 2014, weighing factors such as severity, frequency, breadth, insidiousness, profile, and preventability. From an administrative perspective, I see this area of study as contributing to the safety of patients and clinicians, to the advancement of nursing research, and to the ongoing development of medical device alarm-management models. My research interest is to develop strategies to reduce non-actionable clinical alarms and hence, minimize alarm fatigue. After completing my doctoral studies, I aim to contribute to the body of nursing knowledge through my current and future research.

Various alarm studies have been conducted in general ICUs and progressive care units, but no studies have examined clinical interventions (pre- and post-intervention) to systematically and comprehensively reduce physiologic monitor alarm rates in neuroscience intensive care units (NICU). This study is innovative in that it will be the first to assess the effectiveness of daily nursing interventions in reducing alarm rates specific to SpO_2, technical, and arrhythmia alarms in a NICU. Although existing studies have merit, they also possess limitations. For example,

these studies do not provide essential alarm data, such as information regarding lower severity physiologic monitor alarm levels (e.g., “message” alarms that trigger only visual alerts) and all monitor alarms that may occur simultaneously. Moreover, existing studies rarely report patient outcome data, nor do they provide in-depth analysis and annotation of arrhythmia alarms to determine the proportion of false-positive critical arrhythmia alarms, assess the frequency of audible alarms generated by physiologic monitor alarms, and the frequency of total physiologic monitor alarm (audible and inaudible). Furthermore, existing studies have not reported physiologic monitor alarm rates on the basis of patients’ monitoring hours; that is, past studies have used physiologic monitor alarms as the unit of analysis, which fails to account for variations in the frequency of alarms each unique patient contributes to the aggregate unit alarm burden—both true alarms and false-positive alarms. This study addresses these shortcomings and provides a more comprehensive understanding of physiologic monitor alarm rates, builds on prior research, and contributes to the formulation of alarm practice standards.

Statement of the Problem

Physiologic monitors play a vital role in saving patients’ lives—yet expose clinicians to an overwhelming number of alarms. For clinicians, excessive physiologic monitor alarm frequency may lead to alarm fatigue, which is associated with decreased responsiveness to alarms and consequent medical error. In neuroscience intensive care units (NICUs), physiologic monitors enable surveillance of electrocardiographic (ECG) rhythms and other physiologic waveforms and parameters (e.g., heart rate, blood pressure, oxygen saturation, intracranial pressure) on a continuous "real-time" basis. Each of these measurements has associated auditory and visual alarms that alert clinicians to unanticipated changes in a patient’s condition. The ECG

alone can be the source of 20–30 types of alarms that are triggered by changes in cardiac rhythm (e.g., asystole, ventricular fibrillation) and rate (e.g., too fast, too slow).

In addition, poor clinical practice associated with use of medical equipment and patient monitoring supplies can generate numerous technical alarms and hence contribute to physiologic monitor alarm burden. Although strategies to minimize alarm burden associated with technical alarms are unsophisticated and relatively simple, such strategies are often understudied and underappreciated. Published intervention studies, including the single study that has investigated the effect of daily ECG electrode change on reductions in technical alarms, have not reported whether interventions were effective in reducing false–positive arrhythmia alarms or whether they had an effect on patient care or outcomes (Cvach, Biggs, Rothwell, & Charles-Hudson, 2012). The present study will address these shortcomings and will include annotations of critical arrhythmia alarms and report patient outcomes such as the incidence of cardiopulmonary respiratory arrests and acute respiratory compromise in which a hospital-wide resuscitation response was activated. Furthermore, this study strengthens the design of previous clinical research and enables a better understanding of physiologic monitor alarm rates. In addition, this investigation comprehensively studies all parameter, technical, and arrhythmia alarms and their individual contributions to physiologic monitor alarm rates in a NICU. Because alarm fatigue related to physiologic monitor alarm rates is multifaceted and multilayered, a variety of approaches to reduce alarms must be employed to reduce the frequency of nuisance alarms. The most important and perhaps most difficult aspect of this effort is fostering clinicians' recognition and acceptance that past practices, default alarm settings, and organization's policies and procedures may actually contribute to the risk of alarm hazards. Furthermore, changes in clinical practice require open-mindedness, creativity, and support from nurses, physicians, clinical

engineers, and the executive leadership team—in order to facilitate procurement of resources aimed at reducing clinical alarm burden and fatigue.

Purpose of the Study

This study describes a randomized clinical trial study that assesses the impact of select nursing interventions such as the modification of existing oxygen saturation (SpO_2) default alarm settings, the introduction of a new monitoring feature (i.e., SpO_2 alarm delay), and a *daily* skin preparation using ECG skin preparation paper and the application of *daily* Ag/AgCl-foam, pre-gelled *wet* ECG electrodes on reducing physiologic monitor alarm rates in a NICU. Our research study will answer the question: "Is the mean hourly alarm rate difference between Assessment 1 and Assessment 2 in the experimental unit different from that of the control unit?"

The primary aims of this study are

1. To determine whether (a) the mean hourly rate for SpO_2 low-threshold alarms in the experimental unit (whose default SpO_2 alarm setting is adjusted to a low threshold of less than or equal to 88% with a 15-s alarm delay) between Assessment 1 and Assessment 2 is different from that of (b) the mean hourly rate of SpO_2 low-threshold violation alarms in a control unit that utilizes traditional SpO_2 alarm settings (i.e., low threshold of less than or equal to 90% and a 5-s SpO_2 alarm delay) between Assessment 1 and Assessment 2.

   Hypothesis 1: The difference between Assessment 1 and Assessment 2 in the experimental unit is not the same as the difference between Assessment 1 and Assessment 2 in the control unit. We expect the mean hourly rate for SpO_2 low-threshold alarms to decrease in the experimental unit over time and remain the same in the control unit.

2. To determine whether a) the mean hourly rate for technical alarms (i.e., ECG lead fail, artifact, and arrhythmia suspend) in the experimental unit, whose patients receive a *daily* skin preparation using ECG skin preparation paper and the application of *daily* Ag/AgCl-foam, pre-gelled *wet* ECG electrodes between Assessment 1 and Assessment 2 is different from (b) the mean hourly rate for technical alarms in the control unit, whose patients may or may not receive the usual skin preparation (i.e., use of soap and water, dry with a dry gauze) and application of Ag/AgCl-foam *solid hydrogel* ECG electrodes every *two days* between Assessment 1 and Assessment 2.

   Hypothesis 2: The difference between Assessment 1 and Assessment 2 in the experimental unit is not the same as the difference between Assessment 1 and Assessment 2 in the control unit. We expect the mean hourly rate for technical alarms (i.e., ECG lead fail, artifact, and arrhythmia suspend) to decrease in the experimental unit over time and remain the same in the control unit.

3. To determine whether (a) the mean hourly rate for arrhythmia alarms (i.e., 6 critical arrhythmia and all arrhythmias alarms) in the experimental unit—whose patients receive a *daily* skin preparation using ECG skin preparation paper and the application of *daily* Ag/AgCl-foam, pre-gelled *wet* ECG electrodes between Assessment 1 and Assessment 2—is different from (b) the mean hourly rate for arrhythmia alarms in the control unit, whose patients may or may not receive the usual skin preparation (i.e., use of soap and water, dry with a dry gauze) and application of Ag/AgCl-foam *solid hydrogel* ECG electrodes every *two days* between Assessment 1 and Assessment 2.

   Hypothesis 3: The difference between Assessment 1 and Assessment 2 in the experimental unit is not the same as the difference between Assessment 1 and

Assessment 2 in the control unit. We expect the mean hourly rate for arrhythmia alarms to decrease in the experimental unit over time and remain the same in the control unit.

4. To determine whether the mean hourly audible alarm rate in the experimental unit—whose patients have a new oxygen saturation (SpO_2) default alarm setting and receive a *daily* skin preparation using ECG skin preparation paper and the application of *daily* Ag/AgCl-foam, pre-gelled *wet* ECG electrodes between Assessment 1 and Assessment 2—is different from (b) the mean hourly rate for audible alarms in the control unit, whose patients have a traditional SpO_2 default alarm setting and may or may not receive the usual skin preparation (i.e., use of soap and water, dry with a dry gauze) and application of Ag/AgCl-foam *solid hydrogel* ECG electrodes every *two days* between Assessment 1 and Assessment 2.

   Hypothesis 4: The difference between Assessment 1 and Assessment 2 in the experimental unit is not the same as the difference between Assessment 1 and Assessment 2 in the control unit. We expect that, over time, the experimental unit's mean hourly audible alarm rate will decrease, and the control unit's rate will remain the same.

5. To determine whether the mean hourly physiologic monitor alarm rate in the experimental unit—whose patients have a new oxygen saturation (SpO_2) default alarm setting and receive a *daily* skin preparation using ECG skin preparation paper and the application of *daily* Ag/AgCl-foam, pre-gelled *wet* ECG electrodes between Assessment 1 and Assessment 2—is different from (b) the mean hourly physiologic monitor alarm rate in the control unit, whose patients have a traditional SpO_2 default alarm setting and may or may not receive the usual skin preparation (i.e., use of soap and water, dry with a

dry gauze) and application of Ag/AgCl-foam *solid hydrogel* ECG electrodes every *two days* between Assessment 1 and Assessment 2.

Hypothesis 5: The difference between Assessment 1 and Assessment 2 in the experimental unit is not the same as the difference between Assessment 1 and Assessment 2 in the control unit. We expect that, over time, the experimental unit's mean hourly physiologic monitor alarm rate will decrease, and the control unit's rate will remain the same.

6. To determine whether (a) the mean percentage of false-positive cardiac arrhythmia alarms (e.g., asystole, accelerated ventricular, pause, ventricular bradycardia, ventricular tachycardia, ventricular tachycardia/ventricular fibrillation) in the experimental unit—whose patients receive a *daily* skin preparation using ECG skin preparation paper and the application of *daily* Ag/AgCl-foam, pre-gelled *wet* ECG electrodes between Assessment 1 and Assessment 2—is different from (b) the mean percent of false-positive cardiac arrhythmia alarms in the control unit, whose patients may receive the usual skin preparation (use of soap and water, dry with a dry gauze) and application of Ag/AgCl-foam *solid hydrogel* ECG electrodes every *two days* between Assessment 1 and Assessment 2.

   Hypothesis 6: The difference between Assessment 1 and Assessment 2 in the experimental unit is not the same as the difference between Assessment 1 and Assessment 2 in the control unit. We expect the mean percentage of false-positive cardiac arrhythmia alarms to decrease in the experimental unit over time and remain the same in the control unit.

Dissertation Chapters

Chapter 1 provides an introduction to the dissertation, presents the research question, and discusses the need for nursing research to study the effectiveness of nursing interventions in reducing physiologic monitor alarm rates in intensive care units.

Chapter 2 Part 1 reviews literature that focuses on what is known about the topic of interest, what remain unclear, disagreements between studies, and current issues.

Chapter 2 Part 2 offers a theoretical framework for studies on alarm burden in the context of physiologic monitor alarms.

Chapter 3 provides a rationale for the study's methods (including study design and assessment instruments).

Chapter 4 discusses the methods utilized in this study to measure the effect of nursing interventions on reducing physiologic monitor alarm rates in an adult neuroscience intensive care unit, including inclusion–exclusion criteria, methods of measurement, data analysis, and human subject protection.

Chapter 5 presents the study's findings. This discussion uses the patient as the unit of analysis, thus enabling calculation of per-patient physiologic monitor alarm rates on the basis of individual patients' monitoring hours. Use of these calculations in turn enables determination of each individual patient's unique contribution to the overall unit alarm burden.

Chapter 6 summarizes the findings of the research study on physiologic monitor alarm rates and concludes with discussion of implications for clinical practice and recommendations for future nursing research.

References

Cvach, M., Biggs, M., Rothwell, K. J., Charles-Hudson, C. (2012). Daily electrode change and effect on cardiac monitor alarms: An evidence based practice approach. *Journal of Nursing Care Quality. 28*(3), 265-271.

Emergency Care Research Institute. (2013). Top 10 health technology hazards for 2014. *Health Devices, 42*(11), 354–380.

Chapter 2 Part 1

Literature Review

Over the past two decades, medical devices have played a vital role in protecting and saving patients' lives, yet the very devices that have improved patient safety now also present a threat. Medical equipment that generates alarms includes not only physiologic monitors for measuring and monitoring vital signs but also therapeutic devices, such as intravenous pumps, enteral pumps, and mechanical ventilators, which support bodily functions. Clinical settings typically contain multiple alarm-equipped devices; these devices present a diverse array of auditory and visual alarm signals that independently alert clinicians to changes in a patient's condition or therapy. The continuous operation of such devices—and, often, the simultaneous presentation of multiple, discordant alarm signals—impose an extraordinary sensory and cognitive challenge on clinicians' attentional capacity.

Medical devices are designed to improve patient safety; however, they also contribute to clinicians' exposure to an overwhelming number of alarms. Estimates of alarm burden reported in recent adult studies range from one alarm every 1.5 minutes to one alarm every 10 minutes (Chambrin et al., 1999; Graham & Cvach, 2010). Moreover, research has reported that, in adult intensive care units (ICUs), more than 85% of all physiologic monitor alarm events are technically false or occur because of some form of manipulation by clinicians (Biot, Carry, Perdix, Eberhard, & Baconnier, 2000; Chambrin et al., 1999; Koski, Mäkivirta, Sukuvaara, & Kari, 1990; Siebig et al., 2010a); in other words, fewer than 15% of alarm events are clinically relevant. Similar observations have been reported in pediatric critical care settings (Lawless, 1994; Talley et al., 2011; Tsien & Fackler, 1997).

In aggregate, the large frequency of physiologic monitor alarms and the simultaneous presentation of the diverse array of alarm types may lead to medical errors because clinicians are

either distracted by alarms or simply choose to ignore alarms. To reduce the cacophony of alarms, clinicians have silenced, suspended, or deactivated lifesaving medical devices in misguided attempts to reduce alarms. The purpose of alarms is to call the attention of clinicians to the patient or conditions that deviate from a pre-determined "normal" status; yet, the sheer abundance of alarms now challenges their original intention.

The proliferation of technology in the clinical environment has an unintended consequence for clinicians: alarm-related sensory overload that can result in "alarm fatigue" (ECRI, 2012a). The ECRI defines *alarm fatigue* as a condition that occurs when clinicians are exposed to an excessive number of alarms. Sensory overload causes staff to become desensitized to medical device alarms and results in a range of adverse effects on clinicians' responses to alarms: (a) delay in initiating response; (b) slower execution of response; (c) failure to respond altogether; and (d) error in identifying correct response (ECRI, 2012b). The most serious outcome of sensory overload is the failure to recognize and respond to true alarms (TA) that require intervention as a result of the high occurrence of alarms. Moreover, the detrimental consequences of excessive alarm presentation are not limited to effects on clinicians; the profusion of alarms in care environments also prevents patients from receiving rest (ECRI, 2012b).

In 2013, the ECRI identified *alarm hazard* as a key safety issue and a health care technology hazard for 2014. Excessive physiologic monitoring alarms and their byproduct, alarm fatigue, have become a serious threat to public health. Internationally, adverse alarm events have resulted in patient deaths across the continuum of care—from home health care to state-of-the-art tertiary and quaternary care hospitals. In Germany from 2007 to 2009, 75 adverse alarm events resulted in patient outcomes ranging from no harm to brain damage and death (Borowski et al.,

2011). In the United States from 2005 to 2008, 566 alarm-related deaths were recorded in the U.S. Food and Drug Administration (FDA) Manufacturer and User Facility Device Experience (MAUDE) database; examination of these records revealed that, most often, users were not familiar with the monitoring equipment, had not checked the alarm parameters, or had not checked the status of the alarm (Weil, 2009). Regrettably, between 2005 and 2010, more than 200 hospital patient deaths were linked to problems with alarms on physiologic monitors (Kowalczyk, 2011a). A series of untimely patient deaths in premier academic medical centers has generated media attention regarding alarm failures (Kowalczyk, 2010, 2011b, 2011c; McKinney, 2010). This heightened public awareness has given rise to a national call to action directed at the academic and health care community, medical device industry, and governmental agencies to acknowledge alarm fatigue and seek inter-professional solutions to mitigate this phenomenon (Association for the Advancement of Medical Instrumentation [AAMI], 2011).

Although organizations are recognizing that implementing a comprehensive alarm management system is a complex undertaking, exploring current clinical practices and the use of existing technologies that may reduce alarm burden and improve alarm efficacy is imperative (AAMI, 2012; ECRI, 2007; Phillips, 2006). This chapter provides a comprehensive review of published research on physiologic monitor alarm burden—with a particular focus on research pertaining to monitor interventions and clinical practice. This review also identifies significant deficits in current knowledge and areas for future research. Ultimately, knowledge gained through this review of the literature will guide research pertaining to physiologic monitor alarm burden and contribute to the existing body of knowledge.

Methods

Search Strategies

For the present literature review, a comprehensive search utilized three databases: Medline PubMed, Cumulative Index Nursing Allied Health Literature Plus (CINAHL), and Embase. A search was performed using the key terms *alarms, fatigue, ECG monitoring,* and *nurses;* this search identified a small number of research studies. Subsequently, the PubMed Medical Subject Heading (MeSH) database was used to find key terms and build searches. Multiple iterative searches using the MeSH terms *monitoring, physiologic, telemetry, arrhythmias, clinical alarm/standards, equipment failure, fatigue,* and *intensive care unit* identified a limited number of studies. A search involving *monitoring, physiological,* and *clinical alarms* yielded 58 studies. Although these articles were published in prominent journals and were of general interest, many of these studies were not relevant to the research question "What interventions are effective for minimizing alarm fatigue related to physiologic monitor alarms"? Inclusion of the term *fatigue* in the search reduced the number of articles that focused on monitor alarm burden. The reference lists of these research articles contained additional relevant studies and other types of cited sources in a variety of disciplines. An iterative process was used repeatedly to perform a comprehensive literature search; the final review included 38 research articles published in peer-reviewed journals. A detailed summary of the research articles identified in this search is provided in Table 2.1.

Numerous observational studies have examined physiologic monitor alarm burden; however, few studies have examined interventions to reduce alarms. The present search located no seminal scientific studies or leading researcher in this field of study. Of late, alarm research has attracted growing attention from various nurse scientists, physicians, biomedical engineers,

and industry leaders; as a result, large or otherwise important studies involving technological advancements and nursing practice regarding alarm system management are expected.

Review of Current Research

Physiologic Monitor Alarms

In order to promote understanding and standardization in the nascent field of alarm burden research, the ECRI developed a glossary of common alarm terminology:

- *Nuisance alarms* are clinical alarms that (a) are perceived by clinicians to be annoying, (b) typically do not indicate an adverse patient condition, and (c) may interfere with patient care activities. These types of alarms are a cause for concern because they may interrupt staff performance of necessary tasks—even when no condition exists that would warrant attention or action.
- *False alarms* (FA) are alarms that indicate a need for a clinical action in response to a physiologic event when no true event has actually occurred (ECRI, 2012a).

Excessively frequent FAs have numerous safety ramifications that stem from clinicians' becoming desensitized to—and ultimately, habituated to—the physiologic monitor alarms in their environment. Furthermore, FAs reduce clinicians' trust in patient monitoring systems. That is, nurses tend to ascribe less importance to alarms that are preceded by FAs than to alarms not preceded by FAs (Breznitz, 1984). The reduced importance that nurses accord to alarms preceded by FAs manifests as suboptimal response to alarms (e.g., slower responses, less frequent responses, less accurate assessments of responses). Of course, many factors can influence the effect of an FA on nurses' evaluation of subsequent alarms—for example, FA frequency, length of time between FA and TA, the relative importance of alarms had they not been FAs, and a nurse's level of fatigue (in general and at various points during a shift).

Nuisance alarms and FAs have been identified as major culprits in the etiology of alarm fatigue, and their occurrence must be reduced to promote patient safety.

Causes of Alarms

The causes of physiologic monitor alarms vary across patient populations and are influenced by phase of care, severity of illness, and variables of patient monitoring. Many descriptive studies have identified and quantified causes of physiologic monitor alarms among diverse patient groups. Although some of the causes of alarms are common to all patient subpopulations, causes can also differ across subpopulations. This diversity of causative factors points to a need for multiple approaches to diminish staff and patient exposure to non-actionable alarms. Descriptive physiologic monitor alarm studies report frequencies, types, and causes of alarms, and provide the basis on which to initiate concerted efforts to alleviate alarm fatigue.

Pediatric Studies

Few alarm studies have been conducted in pediatric clinical settings, and these studies have had design limitations. Only one neonatal study and three pediatric studies have identified and quantified sources of physiologic monitor alarms. A thorough literature search found no studies on interventions to reduce alarm burden in pediatric units. Several investigators have reported that respiratory-related events—primarily hypoxemia and apnea—comprised the majority of monitor alarm events among neonates and critically ill children (Bitan, Meyer, Shinar, & Zmora, 2004; Lawless, 1994; Talley et al., 2011; Tsien & Fackler, 1997). In two pediatric intensive care unit (PICU) studies, oxygen saturation alarms generated over 43% of the recorded alarms and were the largest contributor of total alarms (Lawless, 1994; Tsien & Fackler, 1997). Moreover, these studies reported that more than 92% of monitor alarms were false-positive alarms. Comparable results were seen in the neonatal intensive care unit study

(Bitan et al., 2004).

The pediatric and neonatal observation studies were relatively small and were conducted in single units. To measure monitor alarm burden, all researchers utilized a cumbersome methodology: direct observation and recordings obtained from physiologic monitors. This approach introduced potential bias, given that the RNs were aware of being observed and, in addition, some participated in data collection. Notably, inclusion and exclusion criteria were not well defined in these studies. Bitan et al. (2004) excluded monitor alarms that were triggered while RNs were providing care; as a result, not all alarm data were captured, and the alarm frequency calculation was probably inaccurate.

While these studies provide a pediatric perspective, they possess significant limitations and therefore should be interpreted with caution. While cardiac arrhythmias are uncommon in infants and children (Young & Seidel, 1999), a major limitation of these pediatric studies is that researchers did not investigate or report the prevalence of cardiac arrhythmia alarms. Taken collectively, the studies' findings indicate that most pediatric physiologic monitor alarms are related to respiratory issues and that future research should examine factors that contribute to the numerous oxygen saturation alarms (such as inappropriate parameter settings and equipment failures).

Adult Studies

Descriptive studies—prospective and retrospective—performed in the adult inpatient setting reveal that predictable sources of physiologic monitor alarms are associated with types and levels of care. To date, only four observational studies—Burgess, Herdman, Berg, Feaster, and Hebsur (2009), Graham and Cvach (2010), Gross, Dahl, and Nielsen (2011), and Whalen et al. (2013)—have assessed the effects of alarm burden on acute and progressive care units. Alarm

research has been conducted in the context of critical care; the majority of these studies have significant design limitations. In four adult studies that systematically analyzed physiologic monitor alarms, the majority of alarms were classified (in order of decreasing frequency) as parameter threshold violations, technical events, or critical arrhythmia alarms (Biot et al., 2000; Blum, Kruger, Sanders, Gutierrez, & Rosenberg, 2009; Siebig et al., 2010a, 2010b).

Parameter alarms. Parameter alarm limits are for monitored parameters (i.e., heart rate, blood pressure, and pulse oximetry) and their high- and low-threshold limits. Unit default alarm limits are determined by the type of patient population of a particular care area. Specific alarm limits (e.g., nondefault) are chosen and adjusted on the basis of the individual patient's physiologic condition (Pennsylvania Patient Safety Authority, 2008).

Non-ICU studies of parameter alarms. The four non-ICU studies that quantify physiologic monitor alarms had dissimilar aims—and, as a result, establishing alarm ascendancy relative to these studies is difficult. Burgess et al. (2009) performed an analysis utilizing a secondary data set from an earlier study that evaluated an automated, non-invasive patient vigilance system. This study's results revealed that low heart rate (HR) and low and high respiratory rate (RR) threshold alarms produced the largest number of physiologic monitor alarms. Unfortunately, data from continuous pulse oximetry (SpO_2) were not reported.

Subsequently, Gross and colleagues (2011) examined the impact of physiologic monitor alarms on an adjudicated sample of medical–surgical patients (n = 30). This study found that the majority of physiologic monitor alarms were, collectively, SpO_2 alarms, RR low-threshold alarms, and HR high-threshold alarms. Furthermore, the authors reported that more than 40% of their high-priority alarms were FAs; these FAs predominately included high or low HR alarms, high or low RR alarms, and low SpO_2 alarms. This Gross et al. study has a methodological

strength unseen in other investigations: it is the first to adjust the numbers of critical and high-priority alarms/patient/day to the duration of cardiac monitoring. This approach may yield a more accurate representation of alarm burden, given that some non-ICU patients may not be monitored for a complete 24-hour period.

A frequently cited study conducted at The Johns Hopkins Hospital by Graham and Cvach (2010) in a medical progressive care unit (MPCU) found that HR threshold alarms, both high and low, and SpO_2 low-threshold alarms accounted for the largest number of alarm parameter violations—mostly, non-actionable alarms. In progressive care patients, the high occurrence of non-actionable alarms related to continuous pulse oximetry monitoring is similar to findings from pediatric studies (Lawless, 1994; Tsien & Fackler, 1997). Although the Graham and Cvach study was described as being a quality initiative, this investigation was the first to implement interventions to improve overall alarm management—including modification of default settings, nurse education, and use of software to allow remote views of monitored patients.

Lastly, a recent study by Whalen et al. (2013) conducted on a general medical–surgical unit at Boston Medical Center substantiated a finding by Graham and Cvach (2010)—that HR threshold alarms, bradycardia, and tachycardia alarms are the largest contributors of alarms on telemetry units. This investigation is valuable in that it engaged an interdisciplinary participation; furthermore, unlike in previous studies, in the study by Whalen et al., the alarm setting modifications were formalized in the cardiac monitoring order sets in the electronic medical record. The study's limitations—the intervention serving as its own control unit, lack of operational definitions for alarms, potential for observer bias, and the lack of precise and reliable instruments—exemplify the challenges of performing alarm research in patient care areas.

ICU studies of parameter alarms. Because of the prolific use of advanced life-saving medical devices and the hemodynamic instability of patients, the number of clinical alarms in most ICUs is excessive. Frequently, the configuration of these units is compact; this compact design characteristic may exacerbate the adverse effects of unnecessary physiologic monitor alarms on staff and further compromise patient care. For these reasons, reducing alarm burden is vital.

In three single-site descriptive studies that comprehensively reported threshold alarm settings, the physiologic parameters most commonly associated with threshold violations were systemic arterial blood pressure (systemic ABP), oxygen saturation, and heart rate—in decreasing order of contribution (Blum et al., 2009; Siebig et al., 2010a, 2010b). In a single multicenter study, Chambrin et al. (1999) reported that, in order of decreasing contribution, ABP, HR violation alarms, and SpO_2 parameter alarms were the primary sources of threshold violation alarms. In contrast, Biot et al. (2000) reported the physiologic parameters that created the most frequent true positive alarms, rather than incidence of parameter alarm. Regardless of the method used to report parameter alarms, ABP threshold alarms generated the largest number of monitor alarms. Notably, the frequently cited study by Görges, Markewitz, and Westenskow (2009)—designed to identify means of reducing the number of FAs—recorded HR alarms and arrhythmia alarms as a single set of data. This non-differentiation of alarm data made evaluation of the frequency of parameter alarm events difficult. Despite this unusual approach, after excluding ventilator alarms, continuous SpO_2 monitoring generated more alarms than did ABP and non-invasive blood pressure (NIBP) combined.

In 2004, Zong, Moody, and Mark performed an off-line analysis to demonstrate that false ABP alarms could be reduced with an algorithm that used signal quality assessment and

relationships between ECG and ABP. The investigators reported that through the use of a test algorithm, the false ABP alarm rate was reduced by 27%—to less than 0.5%—while the algorithm accepted 99% of true ABP alarms. Sensitivity was reported at 99.8%, and the algorithm's positive predictive value was 99.3%. Although the algorithms used by commercial monitors are proprietary and publicly unavailable, reducing false ABP alarms through the biomedical signal processing and pattern recognition beneficially reduces nuisance alarms. For example, the use of "smart BP" alarms with certain patient monitoring systems (i.e., GE Healthcare) is an arterial artifact rejection approach that reduces the frequency of occurrence of needless alarms by preventing most of the alarms associated with zeroing the transducer, fast flushing the system, or drawing blood. These ICU studies show that ABP alarms are the most common type of parameter alarms; ABP alarms can be cause by threshold violations, incorrect staff manipulation, or poor clinical technique.

Furthermore, in the first diagnoses-related study to examine the reliability and frequency of alarms in coronary artery bypass and cardiac valve replacement patients (N = 10), HR alarms accounted for the largest percentage of clinically significant parameter alarms (Koski, Mäkivirta, Sukuvaara, & Kari, 1990). Koski et al. (1990) defined *significant alarm* as an alarm event in which a clinician should (a) examine a patient's condition or (b) take therapeutic action during all post-operative monitoring periods. This result, albeit different from those of the above studies, may be attributed to the patients' underlying cardiac disease and resulting surgery. Similar to the approach undertaken by Biot et al. (2000), Koski et al. did not report the incidence of physiologic monitor alarms but, rather, reported the distribution of alarms according to their clinical significance. The investigators found that most parameterized alarms occurred during the rehabilitation period; in the Koski et al. study, *rehabilitation period* was defined as being the

period from extubation until removal of chest drains. This research finding suggests that the majority of parameter threshold limit alarms are triggered as patients stabilize and recover during hospitalization. Continuous pulse oximetry was not a variable of interest in this study, and the omission of this measurement is a limitation.

In the body of research on physiologic monitor alarms, the majority of parameter alarms that occur in adult ICUs are primarily attributed to ABP, pulse oximetry, and heart rate limit threshold violations. These alarms are triggered when physiologic parameter limits (low and high) are inappropriately adjusted by clinicians—that is, when the threshold limits are set too narrowly or too broadly relative to patients' actual hemodynamic values. Setting and changing alarm parameter thresholds based on a patient's physiologic condition is a dynamic process requiring staff technical knowledgeability, clinical judgment, and participation. More important, clinicians must examine the trends of a patient's physiologic parameters—and not just a single physiologic value at the time that an alarm event occurs. This examination requires the knowledge and ability to adjust alarm threshold settings in order to suppress clinically irrelevant parameter alarms. If adjustment is not done, the alarm will repeat incessantly and become a nuisance.

Despite the limitations of individual studies, the results of studies performed in ICUs are similar to the results of studies performed in non-ICU patient settings. This similarity strengthens the general findings and points to the importance of examining parameter threshold violation alarms—because such alarms are the primary contributors of monitor alarm events. Parameter alarms supply little actionable information and minimally influence the course of treatment. Furthermore, parameter alarms play a significant role in the development of alarm fatigue and in

clinicians' failures to respond to clinically significant alarms (Scalzo & Hu, 2013; Solet & Barach, 2012).

Specialty departments studies of parameter alarms. Few physiologic monitor alarm studies have been conducted in operating room (OR) or post-anesthesia care unit (PACU) settings. A single intra-operative study was conducted by researchers who introduced "fixed" alarm threshold settings for patients ($N = 25$) undergoing elective cardiac surgery (Schmid et al., 2011). Notably, clinicians were instructed to not modify patients' alarm settings after extracorporeal circulation in order to characterize patterns of alarms with the use of a specific physiologic monitor and an anesthesia workstation. Although the investigators aimed to identify false-positive alarms, the rationale for the instructions regarding not changing the settings is unclear, given that this prohibition is contrary to guidelines that recommend adjusting parameter threshold settings on the basis of the individual patient's physiologic condition (The Joint Commission [TJC], 2004; Pennsylvania Patient Safety Authority, 2008). Schmid et al. found that ABP (mean, systolic, and diastolic) accounted for the majority of alarms (more than HR and SpO_2)—a finding similar to those of the adult ICU studies that investigated alarm threshold violations (Blum et al., 2009; Chambrin et al. 1999; Siebig et al., 2010a, 2010b). While the OR and PACU are unique care environments, they also appear to be subject to the hazards of excessive monitor alarms, and future studies assessing alarm burden associated with clinical anesthesia would be beneficial. The literature search identified one study that investigated intracranial pressure alarms in a neurosurgical–neurological ICU (Scalzo & Hu, 2013). No Emergency Department (ED) studies were found that focused on physiologic monitor alarm parameters. Furthermore, no studies examining QT interval or ST-segment alarms on alarm burden have been identified.

Technical alarms. Technical alarms are the second most common type of physiologic monitor alarms. The U.S. FDA recognizes the International Electrotechnical Commission's (2006) definition of *technical alarm event*: an alarm event caused by a monitored equipment-related or alarm system-related variable. Types of technical alarm failures include electrical, mechanical, sensor, component, and "other" supply failures. These types of failures may result in an *ECG leads fail* alarm or a *SpO_2 probe off* alarm. Equipment failure can be caused by an unsafe voltage, high impedance, signal impedance, artifact, noisy signal, disconnection, calibration error, or tubing obstruction. Also, technical alarm conditions can arise from algorithms that cannot classify or resolve the available data. In some instances, these alarm conditions can manifest as *arrhythmia suspend* alarms—which indicate that the monitor algorithm is no longer analyzing—a significant threat to patient safety. With hospitalized patients, technical alarms generally result from the accidental detachment of disposable electrocardiographic (ECG) monitoring electrodes or lead wire, a malfunctioning pulse oximetry sensor, or an incorrectly sized or incorrectly positioned blood pressure cuff. Also, these alarms can be a caused by an inadequate signal detection related to motion artifact.

A small number of studies have identified causes of technical alarms and measured the frequency of these causes. In these studies, intentional or unintentional detachment of the ECG electrode, lead wire, or other sensor was repeatedly identified as being a contributing factor. Technical alarms due to ECG lead failures or sensor disconnections are a common type of nuisance alarm.

ECG electrodes. In one of the first physiologic monitor alarm studies conducted in a combined adult–pediatric ICU, O'Carroll (1986) observed that most of the physiologic monitor alarms signaled equipment malfunction rather than critical patient conditions. A total of 1,455

monitor alarms were recorded; of these alarms, over 75% were FAs activated by the removal of the ECG electrodes or lead wires. This investigation is an older study and has recognizable weaknesses; however, despite decades of technological advancement in the field of medical monitoring, the generation of FAs via inadvertent removal of ECG electrodes or lead wires remains a challenge.

Pulse oximetry sensors. In one of the two PICU studies referred to earlier, 64% of false-positive technical alarms problems were caused by inadequate pulse oximetry connection or contact; 10% of such alarms were caused by ECG lead wire movement (Tsien & Fackler, 1997). This finding is in accordance with a PACU study by Wiklund, Hök, Ståhl, and Jordeby-Jönsson (1994), who reported that more than 75% of SpO_2 alarms were false; these SpO_2 alarms were caused by sensor displacement, motion artifact, poor perfusion, or other related factors.

ECG lead fail alarm. Equipment challenges are not restricted to pediatric and post-anesthesia care units. In the Johns Hopkins study by Graham and Cvach (2010), the ECG leads fail technical alarm comprised over 7% of total alarms in the MPCU. When the total number of physiologic monitor alarms was reduced through application of an intervention focused on modifying alarm default settings, the proportion of ECG-lead fail alarms more than doubled (from 7% to 16%). Likewise, pre-intervention arrhythmia-suspend alarms constituted 4% of total alarms; post-intervention, the proportion tripled (12%). The increased post-intervention proportions of ECG-lead fail alarms and of arrhythmia suspend alarms may have been due to the fact that the intervention targeted parameter alarms rather than ECG electrode problems. Graham and Cvach's finding are in alignment with the findings of ICU studies by Siebig et al. (2010a) and Chambrin et al. (1999), who reported that an estimated 7% of all physiologic monitor alarms were attributed to technical alarms. Furthermore, Chambrin et al. found that 22% of the sensor-

related technical alarms required an electrode change, and 50% of these alarms resulted in pulse oximetry sensor repositioning.

Arterial disconnect alarm. The technology for reducing alarms related to staff manipulation remains an opportunity for improvement, as an ICU alarm study by Blum et al. (2009) has reported that the discontinuation of invasive monitor lines triggers numerous technical alarms. Blum et al. (2009) have reported that the act of removing invasive lines from hemodynamic monitors accounted for an estimated 6% of total physiologic monitor alarms. These technical alarms typically display as an *ART disconnect* alarm signal and often occur when patients depart from the critical care setting. The investigators recommend that ART disconnect alarms can be avoided through use of a standardized protocol of first terminating the transduction of the monitored line and then removing the patient's invasive catheter or device. The high frequency of technical alarms raises important considerations: the effects of (a) staff manipulation of equipment and (b) providing routine patient care on the generation of monitor alarms and the development of alarm fatigue.

Few studies have investigated technical alarm conditions. However, the existing research, albeit limited, reveals opportunities to reduce alarm burden by addressing equipment and system-related issues. An intervention that can be implemented quickly and easily as part of a wider range of solutions may include (a) addressing challenges associated with ECG electrodes and lead wires, SpO_2 sensors, and clinical practice surrounding use, and (b) nursing practice related to discontinuation of invasive pressure monitoring.

Arrhythmia alarms. All types of physiologic monitor alarms are susceptible to motion artifact and noise (Hu et al., 2012; Imhoff & Kuhls, 2006; Wiklund et al., 1994). While arrhythmia alarms constitute a small proportion of all physiologic monitor alarms—the failure to

recognize and respond to cardiac arrhythmias can lead to adverse patient outcomes (Drew et al., 2004). Comparatively, these alarms may be few but their potential consequences are considerable. Arrhythmia alarms have significant implications for patient safety—given the potential consequence of arrhythmia and the need for immediate intervention if an arrhythmia alarm is in fact a true alarm. In a large proportion of studies, the investigators have not been able to fully assess, categorize, or evaluate the appropriateness of arrhythmia alarms. Investigators' reasons for their not conducting detailed analysis of arrhythmia alarms include resource limitations and lack of access to the ECG waveform data. Recently, innovative device integration applications are commercially available to assist researchers annotate cardiac arrhythmia alarms—and the use of this clinical software will augment the existing body of physiologic monitor alarm research.

Non-ICU studies of arrhythmia alarms. While three non-ICU studies reported arrhythmia alarms, only one non-ICU study, by Gross et al. (2011), analyzed arrhythmia alarms in terms of whether they were true or false, and the internal validity of this study is weakened by the study's small patient sample size (N = 30 patients). The investigators counted 13.1 critical alarms/patient/day (SD = 21.4, median = 6.0), with over 34% of these alarms identified as being true (on the basis of alarm adjudication). This alarm load has a large standard deviation, which implies that (a) the distribution was skewed, and (b) most of the critical alarms analyzed were associated with a small number of patients. Given these limitations, critical alarms with highest frequency were associated with apnea (33% true), desaturation (39% true), and tachycardia and ventricular tachycardia–ventricular fibrillation combined (Tachy and VTach/VFib; 38% true). The investigators reported that most of the tachycardia alarms for HR exceeding 160 beats per minute (bpm) were false. Gross et al. (2011) identified one bradycardia alarm and one asystole

alarm; both were false (i.e., 0% of these two alarms was true).

Secondly, Graham and Cvach (2010) reported that baseline arrhythmia alarms comprised 40% (6,875) of total alarms (N = 16, 953). The triggering conditions that contributed the largest number of arrhythmia alarms in medical patients were (a) bradycardia, (b) tachycardia, (c) ventricular tachycardia greater than 2 (i.e., a series of ventricular beats that has fewer than six beats but more than two beats and an average HR greater than or equal to 100 bpm), (d) asystole, and (e) ventricular tachycardia. While their performance improvement interventions did not specifically aim to reduce arrhythmia alarms, the number of cardiac arrhythmia alarms was reduced to 33% (3,185) post-intervention. This study, which involved a variety of professional disciplines, is feasible, reproducible, and represents real-world events. However, the study would have been strengthened by (a) an in-depth analysis and annotation of arrhythmia alarms to determine the rate of false-positive arrhythmia alarms and (b) a report of patient outcome data.

The studies by Gross et al. (2011), Graham and Cvach (2010), and Whalen et al. (2013) are the first to report the frequency and prominence of physiologic monitor alarms outside of the ICU setting. Although Gross et al. reported alarms based on priority messaging level and Graham and Cvach and Whalen et al. reported alarms based on arrhythmia classifications, the findings of these three studies were in some ways similar. In particular, in all above studies, bradycardia and tachycardia were the most common arrhythmia alarms in the non-ICU setting. It is important to highlight that Gross et al.'s study utilizes a more robust methodology for reporting the numbers of critical and high-priority; average number of alarms per patient adjusted to the duration of monitoring (which accounts for dispersions in alarm data) rather than describing simple proportions.

ICU studies of arrhythmia alarms. As discussed earlier, arrhythmia alarms occur less frequently than parameter and technical alarms; however, arrhythmia alarms are the most serious. Arrhythmia condition alarms alert clinicians of patients experiencing complex and/or life threatening cardiac arrhythmia—an abnormal heart rate or rhythm. Physiologic monitors require reliable arrhythmia detection algorithms to promote early recognition of true alarms while minimizing FAs.

Among the ICU studies that explored the frequency of arrhythmia alarms utilizing vendor supplied software with visual records (i.e., video monitoring) for annotation, of the total alarms (N = 5,820), it was determined that cardiac arrhythmia alarms occurred on average less than 1.7% (n = 104) of the time, with over 85% of the arrhythmia alarms being classified as technically true. Similarly, Seibig et al.'s (2010b) pilot study found that arrhythmia alarms constituted a small quantity, 2.7%, of total alarms. Biot et al. (2000) utilized a MD observer in the ICU setting to validate cardiac arrhythmia alarms. Investigators reported that cardiac arrhythmias alarms contributed a small proportion of total positive physiologic monitor alarms—fewer than 5%—with an estimated 3% of cardiac arrhythmias alarms identified as false positive, and 2% identified as true positive (TP). Both Biot et al. and the two studies by Seibig et al. reported a low incidence of arrhythmia alarms and an even lower proportion of true arrhythmia alarms.

Critical arrhythmia alarms. The objective of a retrospective study by Aboukhalil et al. (2008) was to determine the frequencies of true and of false arrhythmia alarms. In this study, the investigators utilized a large multi-parameter ICU database to test an algorithm designed to suppress false arrhythmia alarms. The study conducted an expert review of 5,386 arrhythmia alarms associated with simultaneous ECG and ABP waveform analysis. Aboukhalil et al. found

that an average of 42% of the arrhythmia alarms were false. The investigators also reported low rates of true arrhythmia alarms; these rates were similar to the rates reported by Siebig et al. (2010a, 2010b) and Biot et al. (2000). Notably, five cardiac arrhythmias had FA rates ranging from 23% and 90%. Asystole alarms (n = 579) comprised 91% (n = 525) of arrhythmia FAs and 10% of total FAs. Next, was VTach/VFib (n = 313) with an estimated 80% (n = 249) of arrhythmia FAs and 5% of total FAs. Lastly, VTach contributed the highest proportion of alarms (n = 1900) with over 47% (n = 885) of arrhythmia FAs and 16% of total FAs.

The results of the Aboukhalil et al. (2008) study were derived from an off-line analysis of the Physionet's MIMIC II database; given this approach, the study has certain limitations. Mainly, the ICU staff from whom the data were collected chose to standardize arrhythmia analysis on only one selected ECG lead (despite the monitors' having the functionality to perform multi-lead analysis). Still, the use of (a) a large heterogeneous sample of patient alarms from over 48 medical, surgical, and cardiac ICUs in tertiary care hospitals and (b) "gold standard" annotation and adjudication of arrhythmia alarms by experts strengthen the findings. The study's findings regarding the proportion of false-positive arrhythmia alarms are generalizable to adult ICU settings and are similar of FA rates cited in previous descriptive studies. The proportion of cardiac arrhythmia FAs is alarming, and cross-disciplinary efforts are needed to reduce alarm burden by optimizing accurate rhythm recognition.

PVC alarms. Few published research have examined the impact of premature ventricular complexes (PVCs) on physiologic monitor alarm burden. In some areas, PVCs are monitored and are a major contributor to either visual or audible physiologic monitor alarms. In 2011, Fidler, Pickham, and Drew conducted an analysis to quantify PVC alarm frequencies in an academic medical center; specifically this study quantified paired, multiform, R-on-T, bigeminy,

and trigeminy PVC alarm frequencies. The investigators found that during the 2-month study on six patient care units, a total of 318,009 alarms occurred at an estimated frequency of 883 alarms/unit/day. Remarkably, over 38% (120,732) of alarms were due to PVCs. In a comparison of two similar ICUs, the unit with activated PVC alarms had 23,761 PVC alarms, while no PVC alarm was reported in the unit that had de-activated the alarm. All units combined had a total of 19 code blue events (17 patients) and seven resulting deaths (Fidler et al., 2011).

In the 14-bed cardiothoracic ICU study by Blum et al. (2009), investigators did not record the number of arrhythmia alarms in their physiologic alarm notification study, however, their figures reveal that frequent arrhythmia alarms occurred. Despite lack of raw numbers and proportions, predominant cardiac arrhythmias were determined to be PVCs, couplets, VTach greater than 2, tachycardia, and bradycardia. With the Fidler et al. and Blum et al. studies, it is unclear why PVCs alarms were activated, given that treating patients who are PVC-asymptomatic or mildly PVC-symptomatic with antiarrhythmic drugs is no longer recommended or common medical practice—because these drugs may provoke life-threatening ventricular arrhythmias (The Cardiac Arrhythmia Suppression Trial [CAST], 1989). It was not the primary intent of these investigators to provide complete data on the quantity and type of arrhythmias observed, nor was it the investigators' primary purpose to validate these alarms or to report individual unit outcomes. Nevertheless, the learning gained from these studies underscores the importance that future research includes arrhythmia alarms as a key variable of interest.

Unfortunately, no physiologic monitor alarm studies have investigated patient outcomes associated with (a) configuring PVC, to personalized threshold limits; (b) using visual PVC alarms (rather than audible alarms); or (c) deactivating audible PVC alarms.

Sensitivity and Specificity

Patient monitoring systems include intelligent functions that detect subtle changes in patients' condition. While these functions are highly valued, they have a downside: increased potential to contribute to alarm overload. A delicate balance exists between *sensitivity*—which refers to the likelihood that an alarm will correctly signal a true-positive event, and *specificity*—which refers to the likelihood that an alarm will appropriately remain silent during a true-negative event period (Burgess et al., 2009). Commonly, patient monitoring systems have been configured to respond with high sensitivity and specificity—in order to minimize the possibility of missing a clinically true event. However, this configuration results in an excessive frequency of false-positive alarms and consequent alarm fatigue.

Three studies—by Lawless (1994), Chambrin et al. (1999), and Biot et al. (2000)—have reported the sensitivity and specificity of physiologic monitor alarms in adult and pediatric ICUs. In an early work with critically ill pediatric patients, total alarm sensitivity and specificity were reported to be 31% and 82%, respectively, with a positive predictive value (PPV) of 5% and a negative predictive value (NPV) of 98% (Lawless, 1994). Subsequently, a descriptive analysis by Chambrin et al. (1999), involving five adult ICUs in two university and three general hospitals, found an alarm sensitivity of 97% and a specificity of 58%. Moreover, the PPV and NPV were 27% and 99%, respectively (Chambrin et al., 1999).

An alarm study involving 25 patients with acute respiratory distress syndrome conducted in France by Biot et al. (2000) reported sensitivities, specificities, and PPV for each monitoring parameter. The focus of this study was the low sensitivity and PPV assigned to arrhythmia and HR alarms. Specifically, the investigators reported that cardiac arrhythmia sensitivity was 59%, 95% CI [51, 67]; specificity was 99%, 95% CI [99, 100]; and PPV was 42%, 95% CI [36, 49].

Heart rate sensitivity was 81%, 95% CI [76, 85], specificity was 99%, 95% CI [99, 99], and PPV was 69%, 95% CI [64, 73]. Moreover, Biot et al. reported that false-negative alarms (n = 175) were detected as a result of their methodology (direct MD observations). The investigators stated that 50% of the false-negative alarms were related to alarm conditions that were not being detected by the monitor but that were seen by the MD observer. Fifty percent of the false negative alarms were related to cardiac arrhythmia alarms (because the arrhythmia detection software was not in use) and to aspiration alarms. False-negative alarms were attributed to rapid fluctuations in parameter thresholds (25%), and to excessively wide alarm parameters settings relative to the patients' condition (25%). Because the arrhythmia detection algorithm software was not used consistently (i.e., the feature was deactivated), the reported sensitivity and specificity findings should be interpreted with caution. However, the studies by Lawless (1994), Chambrin et al. (1999), and Biot et al. (2000) provide evidence that additional research examining the sensitivity, specificity, and PPV of physiologic monitor alarms is warranted.

During the past decade, medical monitoring has been enhanced by technological developments such as improvements in FA suppression algorithms and rhythm recognition, yet during this period no new studies have reported sensitivity and specificity values. New studies investigating alarm sensitivities and specificities are long overdue—especially with regard to cardiac arrhythmia alarms.

Strategies for Optimizing Physiologic Monitor Alarms

Given the complexity of the alarm systems problem, the review of the literature reveals that a variety of approaches is essential to combat alarm fatigue (Borowski et al., 2011; Cvach, 2012). A range of strategies—some simple, some sophisticated—have been proposed to manage this pressing problem (Brown & Anglin-Regal, 2008; ECRI, 2012b; Edworthy, 2011). Efforts to

minimize hospital alarm fatigue often begin with re-examining current alarm management practices regarding default alarm thresholds and delays, alarm notification, and severity levels (ECRI, 2012b; Konkani, Oakley, & Bauld, 2012, Welch, 2011). In addition, focused assessment of patient monitoring supplies is needed to reduce nuisance alarms and promote safe patient care.

Alarm threshold settings. Parameter threshold alarms have been identified as the biggest source of physiologic monitor alarms—of which, the ABP, SpO_2, and HR threshold alarms collectively contribute a significant proportion. The majority of alarm reduction studies have focused on reducing alarm frequencies through modifying any of a variety of features, including alarm threshold and delay settings. However, most published studies have failed to investigate the impact of modifying parameter thresholds on patient outcomes.

Oxygen saturation threshold alarms. Pulse oximetry is one of the most commonly prescribed patient monitoring parameters across the continuum of care. SpO_2 alarm signals account for large proportions of alarms, have a low positive predictive value, and, consequently, account for few true-positive alarms (Biot et al., 2000; Gross et al., 2011; Rheineck-Leyssius & Kalkman, 1998; Tsien & Fackler, 1997; Wiklund, et al., 1994). Furthermore, a large majority of SpO_2 alarms have been found to be brief self-correcting desaturations; in conjunction with conservative threshold limits, these desaturations generate many FAs (Pan & Gravenstein, 1994). Accordingly, several recent studies have investigated approaches to reducing the incidence of non-actionable SpO_2 alarms.

A review of the literature identified several industry-sponsored medical studies that examined relationships between low-threshold limits and SpO_2 alarm frequency. Most physiologic monitors have a default SpO_2 low-threshold limit of 90%, which is often used as a reference point, and predictably is a common low-threshold limit in health care settings. The

intervention study by Graham and Cvach (2010) found that oxygen saturation alarms could be reduced by 63% by simply lowering the SpO_2 threshold from 90% to 88%. With the lowering of a threshold alarm limit, a decrease in alarm frequency was not unexpected. The focus of Graham and Cvach's study was alarms as the unit of analysis—and not patients; however, the lack of attention to patient outcome variables post-intervention was a limitation in this study's design.

Similar results regarding the effect of modifying SpO_2 settings on alarm frequency have been obtained from off-line analysis. A retrospective study by Rheineck-Leyssius and Kalkman (1998; N = 200 postoperative patients) found that lowering the low-threshold alarm limit significantly reduced SpO_2 alarm frequency. For example, lowering this setting from the default of 90% to 85% reduced the frequency of SpO_2 alarms by 82%. Notably, this off-line study used patient information obtained from a PACU, where the SpO_2 monitoring time is relatively short (median duration, 46 min; range, 13–200 min).

In a pulse oximetry study with findings similar to those of Rheineck-Leyssius and Kalkman (1998), Welch (2011) reported that lowering the low-threshold settings significantly reduced alarm frequency. Welch investigated over 32 million SpO_2 data points and reported that lowering SpO_2 low-threshold limits from 90% to 85% reduced the frequency of pulse oximetry alarms by over 75%. The studies described above were conducted in a variety of patient care environments, yet the studies reported similar results with a low SpO_2 setting of 85%; accordingly, these findings appear to be generalizable.

In contrast to Rheineck-Leyssius and Kalkman's findings and to Welch's findings, Gross et al. (2011) reported less change in SpO_2 alarm frequency—a mere 36% reduction—after reducing the low-threshold setting from 90% to 85%. Gross et al. also reported that a 65% reduction of SpO_2 alarm load resulted from further reducing the low-threshold limit from 85% to

80%. While a threshold limit of 80% would appear unsafe, evidence to the contrary has been reported. Taenzer, Pyke, McGrath, and Blike (2010) performed a landmark patient surveillance study in a 36-bed orthopedic unit and reported significant reductions of SpO_2 alarms achieved through use of unconventional threshold limits and alarm delay settings (low SpO_2 alarm threshold of 80% and 15-s alarm delay)—on average, less than four alarms/patient/day—while reporting satisfactory patient outcomes.

Heart rate threshold alarms. Pulse oximetry is a customary monitoring parameter and so too is HR and RR; but, few studies have investigated HR or RR threshold violation alarms. As noted earlier, only three non-ICU studies—by Gross et al. (2011), Graham and Cvach (2010), and Whalen et al. (2013)—have explored heart rate parameter variables and their contribution to nuisance alarms. Through the use of an alarm history analysis, Gross et al. demonstrated that HR alarm burden could be reduced by greater than 50% with an adjustment of the high HR alarm from 120 bpm to 130 bpm. Graham and Cvach (2010) reported that raising the high HR threshold alarm limit from 120 bpm to 150 bpm reduced alarm frequency by 84%; correspondingly, lowering the low HR threshold from 60 bpm to 50 bpm reduced HR parameter alarm frequency by 88%. However, the studies by Gross et al. and Graham and Cvach studies do not provide information on the potential effects of these HR alarm threshold changes on in-hospital mortality, code blue events, or other outcome variables, such as delays in care or hospital length of stay. In contrast, the more recent study by Whalen et al. evaluated raising the high HR threshold alarm limit from 120 bpm to 130 bpm and lowering the low HR threshold from 50 bpm to 45 bpm reduced parameter alarms by 91%. In comparing the alarm data, in the 2-week period before and after implementation of the alarm modifications, the authors reported an 89% reduction in audible alarm burden per week on the experimental unit ($t = 8.84$; $p <$

.0001). When combining the bradycardia, tachycardia, and HR parameter limit alarms, a 93% (t = 6.34: $p < .0001$) reduction in alarms was observed. Furthermore, the distribution of alarms is reported in three broad categories; arrhythmia alarms, HR limit alarms, and system alarms. It is reported that weekly arrhythmia alarms reduced by 91%, and parameter limit alarms and technical alarms—limited to no telemetry alarm, lead fail alarm, probe off, and arrhythmia suspend alarms—decreased by 94% and 36%, respectively. Differing from previous studies, Whalen et al. reported that there were no changes in the frequency of response team (RRT) activations, whereas the incidence of code blues decreased by 50% (from 6 to 3) on the experimental unit in the 6 months preceding and after implementation of the alarm changes.

Although new HR default alarm thresholds were implemented in the Graham and Cvach and Whalen et al. studies, the studies did not report how often the nurses may have re-adjusted the low and high alarm parameter thresholds to other values. In addition to the fact that Whalen et al. widened the HR threshold alarm default setting which consequentially affects the bradycardia and tachycardia alarms because these two arrhythmia alarms are triggered based upon the average of the most recent eight R-to-R intervals at a heart rate less than the set low or high heart rate limits—the investigators also modified the alarm severity levels. The accelerated ventricular, bradycardia, and tachycardia alarms were elevated to crisis levels (from warning), atrial fibrillation was raised from a visual alarm (i.e., message) to an audible advisory alarm, and the premature ventricular complex alarm was reduced to a message from an advisory level notification. Although the intervention was focused on modifying heart rate related alarms, the investigators did not explain why they did not report the frequencies of the other arrhythmia alarms such as accelerated ventricular, pause, and PVCs, given that the alarm severity levels for these arrhythmia alarms were also modified as part of the intervention.

These study findings should be interpreted with caution because the methodology for collecting alarms and capturing alarm audibility data is not reported, and arrhythmia alarms were not annotated to determine whether they were true or false pre- and post-alarm management intervention. Whalen et al. reported numerous outcome variables, including noise levels, nursing staff perception of noise, nursing staff satisfaction, and patient satisfaction scores; however, the study did not report validity and reliability of its instruments and did not acknowledge that the results observed can be related to other extraneous and confounding variables. Lastly, it is not surprising that an alarm reduction was achieved, given that the unit's alarm settings pre-intervention were very conservative—which undoubtedly contributed to the excessive frequency of alarms during the pre-intervention during the baseline period.

In an older study conducted in a PACU at a university hospital, Wiklund et al. (1994) reported that a mere 3% of alarm frequencies and failures were associated with ECG alarms related to HR alarm threshold violations (i.e., instances of HR less than 40 bpm or greater than 160 bpm), with 14% of the alarms categorized as true. It is worth mentioning that this PACU had wide-ranging alarm threshold settings, and perhaps this contributed to their low HR alarm frequencies.

A human factors study by Solsona et al. (2001) revealed that nursing staff were more apt to set and periodically adjust physiologic alarm limits if they were obligated to document alarm limit values in patients' medical records. That is, this documentation requirement resulted in alarm limits being more closely and more consistently configured to the patients' actual physiologic values and promoted the use of appropriate alarm settings. While proponents might advocate that documenting alarm limits may encourage the customization of alarm threshold settings, this approach creates additional work and potentially introduces compliance concerns.

Alarm delay settings. New software in physiologic monitors have an alarm condition delay feature that adds a time delay from the onset of a triggering event to the point in time at which the monitor actually triggers an audible or visual alarm (IEC, 2006). The effectiveness of such software features is typically studied off-line (utilizing stored patient databases), because these features are not ordinarily amenable to clinical experimentation. Theoretical effects of alarm delays have been investigated by many researchers, and, with the rapidity of technological advancements, recent studies are likely forecast more accurate estimates.

***Oxygen saturation alarm delays*.** Nearly two decades ago, the hypothetical effects of SpO_2 alarm delay conditions were studied by Pan and Gravenstein (1994), who reported that FA-triggered data discrepancies could be reduced by delaying alarm signals. Specifically, a12-s delay prevented 63% of such discrepancies, and a 30-s delay prevented 93% of such discrepancies. Soon after, Rheineck-Leyssius and Kalkman (1998) found that instituting a more conservative SpO_2 alarm delay of 6 s reduced the frequency of nuisance alarms by 50%. Reductions in the incidence of FAs have also been reported in off-line analysis examining the theoretical effects of an alarm delay on monitored parameters (e.g., HR, SAP) in post-operative cardiac patients (Mäkivirta & Koski, 1994; Mäkivirta, Koski, Kari, & Sukuvaara, 1991). Despite this research and for a variety of reasons, many monitoring systems in use today do not use clinical monitoring software with extended alarm delay features.

A decade later, Görges, Markewitz, and Westenskow's (2009) study of clinical alarms in a medical ICU found that the median alarm duration was 17 s, and that, for certain alarms, introducing a 14-s delay or a 19-s delay could reduce alarms by 50% and 67%, respectively. Independent of other monitoring parameters, a 19-s alarm delay would reduce SpO_2 alarms by 52%. Görges et al. acknowledged that introducing alarm delays is not recommended for asystole

and apnea alarms; however, they did not comment on the applicability of alarm delays in the context of other parameters or rhythms (such as tachycardia and bradycardia). While this study is often cited in the literature and has merit, obvious shortcomings weaken the validity and reliability of the results. To increase the strength of alarm delay research, it may be advantageous to focus on a single type of alarm—for example, physiologic monitor alarms—rather than on all clinical alarms seen in the ICU setting.

A dedicated pulse oximetry study by Welch (2011) verified that increasing SpO_2 alarm delays from 5 s to 10 s could decrease alarm frequency by 57%, and increasing SpO_2 alarm delays from 10 s to 15 s could decrease the frequency of alarms by 70%. Moreover, the investigator analyzed reductions in SpO_2 alarm frequencies when a lowered threshold alarm and an alarm delay were combined. That is, a low SpO_2 alarm threshold of 80% with an alarm delay of 15 s produced a 98% reduction in alarm frequency. The results of this industry-sponsored research are compelling; however, prospective studies to determine the possible effect of these changes on patient outcomes are required before recommendations for practice can be issued.

A single study by Taenzer et al. (2010), has explored the clinical effects of simultaneous application of an unconventional SpO_2 low and HR high threshold setting and a prolonged alarm delay setting The investigators recommended that for all patients admitted to the study unit, the SpO_2 low-threshold limit should be set to 80% with an extended alarm delay of 15 s and an HR high threshold of 140 bpm. With these novel alarm settings, the investigators reported a low number of monitor alarms (average, four alarms/patient/day) and highly satisfactory patient outcomes. Rescue events declined from 3.4 to 1.2 per 1,000 patient days, and patient transfers to the ICU decreased from 5.6 to 2.9 per 1,000 patient days. Observed deaths declined from four deaths pre-implementation to two deaths post-implementation, and there were no significant

changes in length of stay. In addition, the study unit outperformed two comparison units on the patient outcome measures. This research suggests that simultaneous application of lowered SpO_2 low threshold setting and extended alarm delay setting produces the greatest reductions in nuisance alarms while effectively maintaining patient safety; however, for future research, a randomized clinical trial design would be better than this pre–post study design. More research and replication studies examining the clinical effects of alarm thresholds, and delays are clearly necessary; the above clinical study can serve as a foundation for further investigation.

Patient monitoring supplies and clinical practice. Technical alarms are the second most common source of monitor alarms. The studies in this literature review indicate that a majority of these alarm signals are related to components such as ECG electrodes and lead wires, pulse oximetry sensors, and related cables. The quality of the electrical signal received from the ECG electrodes is a direct result of good skin preparation, use of high-quality electrodes, and proper electrode application (Turkmen & Pantiska, 2011). In a survey sponsored by the AAMI, the investigators reported that biomedical equipment technicians found that the most common ECG cable and wire problems in the clinical environment were bad electrode placement, dry or old electrodes, improper skin preparation, and broken lead wires, clips, and connector pins; many of these problems cause failures that clinicians may not be aware of (Oster, 2000).

Skin preparation. It is widely accepted that the structural components of skin contribute to a poor electrical conductivity between skin and electrode—known as *electrical impedance*—and that electrical impedance is reliant on the quality of the conductive gel between the skin and the electrode and the extent to which the outer epidermal layer is connected by the conductive gel (Smith, 1984).

ECG skin preparation paper, which has an abrasive fine sandpaper finish, removes part

of the *stratum corneum* (outer layer of epidermis) and scratches the *stratum granulosum* (middle layer of epidermis), thereby enabling transmission of electrical signals to the electrode; this scratching of the stratum granulosum reduces motion artifact during cardiac monitoring (see Figure 2.2; Oster, 2000; Philips Healthcare, 2008; Smith, 1984).

Well over three decades ago, Patterson (1978) performed a study investigating the electrical characteristics of a variety of Ag/AgCl ECG electrodes and skin preparation methods—namely, concerning the benefits of light skin abrasion. The investigator reported the difference between the use of a light sanding preparation and the alcohol or acetone skin preparation was significant in reducing electrical impedance from 100 K-ohms to 2 K-ohms in men and 200 K-ohms to 1.6 K-ohms in women ($p < .01$). This result indicates that use of disposable commercial Ag/AgCl electrodes without prior light skin abrasion may not yield satisfactory ECG tracings. Second, the investigator compared the effects of various skin preparation methods—using fine sandpaper, alcohol, or acetone—on median electrode offset voltage and found that the fine sandpaper preparation method resulted in an improvement in the voltage offset potential ($p < .05$). Comparable results were observed in subsequent studies that investigated the use of fine sandpaper strokes (i.e., 1–5 strokes) for skin abrasion in reducing skin potential and motion artifact (Clochesy, Cifani, & Howe, 1991; Medina, Clochesy, & Omery, 1989; Oster, 2000; Tam & Webster, 1977). Moreover, although the use of alcohol is known to defat the skin of skin oils, skin preparation techniques that involve mild rubbing with an alcohol swab have not been shown to reduce motion artifact in continuous cardiac monitoring (Hanish, Neustein, Van Cott, & Sanders, 1977; Oster, 2000; Patterson, 1978); however, this technique remains in practice—35 years after its efficacy has been called into question.

Performing satisfactory skin preparation prior to the application of ECG electrodes is

often forgotten or dismissed. Reasons for the lapse in practice can include but are not limited to (a) supplies being unavailable or inaccessible, (b) time pressures and competing clinical demands, (c) confusion regarding who is responsible for completing this task, and (d) staff members being unaware of the value of performing this clinical task. However, this skin preparation is beneficial: proper skin preparation may improve ECG signal quality and reduce false physiologic monitor alarms (Adams-Hamoda, Caldwell, Stotts, & Drew, 2003).

ECG electrodes. Although a variety of ECG electrodes are available for clinical use, the Ag/AgCl, foam, pre-gelled *wet* ECG electrode has been assessed as high-quality for continuous cardiac monitoring (Tronstad, Johnsen, Grimnes, & Martinsen, 2010). Tronstad et al. (2010) have reported that sweating may contribute to negative skin conductance responses for wet-gels, and sweating may be related to the thickness of the epidermal stratum corneum layer. However; this type of electrode is known for its clinical advantage in adhering very well to patients' skin (Chi, Jung, & Cauwenberghs, 2010).

A recent study by Cvach, Biggs, Rothwell, and Charles-Hudson (2012) investigated the effect of traditional skin preparation (use of soap and water, rubbing skin with gauze) and daily ECG electrode change on alarms utilizing Ag/AgCl-foam, pre-gelled *wet* ECG electrodes. The daily electrode change intervention affected two types of technical alarm: the ECG leads fail alarm and the arrhythmia-suspend alarm. Oddly, post-intervention, a 13% increase in ECG leads fail alarms occurred in the medical unit, but a 15% decrease was observed in the cardiology unit. A 60% decrease in arrhythmia-suspend alarms was seen in the medical unit, and an even greater reduction in these alarms, 74%, was observed in the cardiology unit (Cvach et al., 2012). The slight increase in ECG lead fail alarms in the remaining unit was unexplained. Potentially, this finding may have been related to the physical action of changing the ECG electrodes more

frequently. However, the investigators reported moderate reductions (33% to 51% change) in technical alarms (which includes *all* medical equipment failure related alarms not simply ECG or RR leads fail alarms) on both study units. These results may provide some support for daily ECG electrode change and the use of wet-gel electrodes. However, it is unknown whether using ECG preparation paper to lightly abrade the skin prior to ECG electrode application would have further reduced the frequency of alarms.

Research regarding the optimal frequency of ECG electrode application (i.e., whether every 12hr, 24 hr, 48 hr, or 72 hr) is scant. The majority of recommendations are predominately from review articles and guidelines; authors primarily advise clinicians to verify that the electrode gel has not dried out (ECRI, 2007; Patel & Souter, 2008; Pennsylvania Patient Safety Authority, 2008; Turkmen & Pantiskas, 2011; Smith, 1984). Moreover, most manufacturers provide different recommendations on their electrode packaging and in their operators' manuals (Covidien, 2013; GE Healthcare, 2005; Philips Healthcare, 2008). Other than Cvach et al.'s study, no research has examined the effect of monitoring equipment on physiologic monitor alarm burden. Replication studies and new studies that investigate clinical practice regarding skin preparation and patient monitoring components such as electrodes, pulse oximetry sensors, NIBP cuffs, cables, and invasive pressure monitoring are needed to confirm the findings reported in Cvach's research.

Impact of Alarm Fatigue

Patient monitoring systems are designed to optimize patient safety through the use of advanced technologies; however, excessive monitor alarms have become a hazard that impacts key stakeholders in varying ways and degrees. In particular, clinicians have an obligation to

protect patients from harm by practicing safely, adhering to established policies and procedures, and participating in efforts to resolve alarm problems that might be harmful to patients.

Impact of alarm on clinicians. The knowledge and opinions of clinical staff regarding the usefulness of monitor alarms may help identify problems with physiologic monitoring systems and priorities for the future. To date, five cross-sectional surveys of clinicians' perceptions of clinical alarms have been conducted. Of these studies, one survey focused on anesthesiologists' and ICU RNs' perceptions of the cardiovascular monitoring system, and two surveys predominately sought input from MDs on monitor alarm limits. The remaining two surveys involved a variety of health care professionals and investigated perceptions regarding all clinical alarms in the health care environment. Because a re-survey was performed utilizing the same instrument, a comparison over time is possible. In addition, a single qualitative study on monitor alarms and their effects on pediatric acute care RNs was published.

In a prominent national survey sponsored by the Healthcare Technology Foundation (HTF) of the American College of Clinical Engineers (2006), investigators explored the effectiveness of clinical alarms. Over 1,300 health care professionals participated, and a majority (81%) of respondents identified frequent nuisance alarms as problematic; 77% respondents indicated that nuisance alarms disrupt patient care, and 78% of respondents indicated that nuisance alarms lead to distrust of alarms and disabling of devices (Korniewicz, Clark, & Yadin, 2008). A re-survey conducted 5 years later found that, unfortunately, these perceptions had changed little during the intervening period. Specifically, 76% of respondents reported that nuisance alarms occurred frequently—a decline (i.e., improvement) of 5% since 2006—with the highest estimates coming from RNs (84%) and respiratory therapists (71%). In the 2011 HTF survey, 71% of respondents felt that nuisance alarms disrupt care—a decline of 6% from the

2006 survey; in the 2011 survey, the percentage of respondents who felt that nuisance alarms reduce trust and cause caregivers to turn alarms off remained at 78% (i.e., unchanged from the 2006 survey). Notably, the 2011 survey revealed that participants' perceptions and recollections regarding compliance with alarm policies, procedures, and documentation that alarms were set and appropriate for each patient was less than 5% (HTF, 2011).

The investigators recognized that the demographics of the clinicians who responded to the 2011 survey differed from those of the 2006 survey respondents. In particular, in the more recent survey, increased survey participation was seen among acute care hospital staff, ICU departments, respiratory therapists, and, in the overall years of health care experience, among the respondents. Although in the second survey the 81% survey response rate for clinical groups as a whole was satisfactory, the RN response rate declined significantly—from 51% to 33%. Also, the 2011 survey (N = 3,454) had more participants from all clinical areas than did the 2006 survey (N = 1,327), but fewer RNs participated in the 2011 study than in the 2006 survey; the reasons for this decrease is unclear. In addition, physician participation in the 2011 survey was minimal. Two factors may have affected the on-line survey completion rates of nurses and of physicians: the reduced survey timeframe (5 months in 2006 versus 1 month in 2011), and the fact that the 2006 survey was also made available in paper format for participating hospitals. Neither study report provided detailed descriptions of sampling design, survey administration, and eligibility criteria; these deficits introduced bias and limit the generalizability of the findings.

A study conducted by Block, Nuutinen, and Ballast (1999) on physician practices in the United States, Finland, and the Netherlands reported that the leading reason for physicians' turning off monitor alarms was excessive FA frequency. The study corroborated the findings of an earlier study by Koski, Mäkivirta, Sukuvaara, and Kari (1995), in which physicians reported

that they routinely ignored many of the monitors' parameter limit alarms because of poor specificity and, in the respondents' view, an unacceptable length of time required to configure individual parameter threshold limits. User feedback clearly indicated that adjustment of parameter alarms for individualized threshold limits must be less time consuming and more automated (Koski et al., 1995). Establishing alarm safety in hospitals requires interdisciplinary input in order to (a) determine appropriate default alarm limits for various diagnostic categories and (b) develop configuration criteria and procedures for responding to parameter threshold alarms in specific care settings.

In a large survey study involving over 180 ICUs, Siebig et al. (2009) sought RNs' and MDs' opinions regarding parameter threshold alarms. Regarding clinician control of alarm limits, 52% of the RNs viewed this function as solely a nurse responsibility, whereas 90% of MDs viewed this function as a shared responsibility. This difference of opinion underscores the importance of defining and delimiting staff responsibilities regarding initiation, adjustment, and discontinuation of parameter alarms. Optimizing congruence between physiologically appropriate alarm limits and MD notification orders of vital sign abnormalities (often specified on standardized order sets) also adds complexity to the determination of alarm limits and settings. Each of the three studies that focused specifically on physiologic monitor alarms (Block et al., 1999; Koski et al., 1995; Siebig et al., 2009) used a unique survey instrument—rendering comparison of results difficult. Unfortunately, a valid and reliable instrument that assesses clinicians' opinions regarding physiologic monitor alarms or all clinical alarms has not yet been developed.

An investigation by Varpio, Kuziemsky, MacDonald, and King (2012) was the first qualitative examination of the effects of monitor alarms on nurses' work flow. The findings of

this study were in accord with those of earlier alarm burden studies and substantiated concerns regarding the occurrence and detrimental consequences of alarm fatigue. Analyzing interviews from 14 nurses assigned to acute pediatric patients, the investigators found that RNs are burdened by FAs, which interrupt workflow. Participants stated that FAs are problematic and that it was standard practice to dismiss or ignore alarms altogether. The practice of ignoring alarms was corroborated when the investigators observed 446 monitor alarms—of which only 3.8% generated a response from the RNs (i.e., attending to patients). The investigators also found that nurses responded to alarms in different ways in order to manage frequent interruptions in their clinical duties. Remarkably, over 70% of investigator-observed RN responses to monitor alarms were delayed, and 41% of these delayed responses were to critical alarms that warranted prompt action. This observation is alarming, and the problem is probably underreported in this study—because the nurses most likely were on their "best behavior," knowing that an investigator was observing their practice. Future qualitative research is required to explore the multiple impacts of alarm fatigue on nurses and on their professional practice. This research will generate findings that will assist nurse leaders in developing strategies to minimize hospital alarm burden and enhance patient safety.

The harsh reality of alarm burden is brought into focus by studies that have reported clinicians' inability to identify audible alarms. An influential study by Cropp et al. (1994) revealed that, on average, RNs, MDs, and RTs were able to correctly identify only 50% of critical alarms (n = 10) and 40% of non-critical alarms (n = 23). In particular, RNs were able to identify only an estimated 50% of critical alarms. These findings are in accordance with other hospital and laboratory studies in which alarms were correctly identified less than 40% of the time by both non-clinical and clinical staff (Momtahan, Translet, & Hetu, 1993; Sanderson, Wee,

& Lacherez, 2006). Ergonomics and human-factors engineering data have shown that individuals are limited in terms of the audible information they can identify and remember. Miller (1956) reported that humans could identify up to seven sounds with reliable accuracy and validity. Human factors studies underscore the importance of research to minimize audible alarm overload—alarm frequency, alarm diversity, and alarm sound volume.

Lastly, in a seminal environmental medicine study, Rhyerd and Persson-Waye (2008) examined the RNs' opinions towards medical device alarms. Many RNs (43%) stated that noises from alarms influenced their ability to perform their job tasks, and almost half of the RNs divulged that they sometimes adjusted the alarm levels so that they would not hear them. Nonetheless, a large proportion of Swedish RNs agreed that adjustments to the alarms should be documented in the medical record. Ninety-one percent of the RNs reported that noise negatively affected their daily work environment; the nurses' reported detriments included irritation (66%), fatigue (66%), concentration problems (43%), and tension headaches (40%). The investigators report that neurological intensive care noise—including noise from monitor alarms—warrants further research to evaluate the impact of hospital noise on staff and on patients who may develop ICU syndrome related to their care environment.

Impact of alarms on patients. It is beyond the scope of this review to examine the effect of noise levels on patients; however, it is important to understand that a comprehensive alarm management system may reduce hospital noise pollution. Alarm research is gaining momentum and the health care community is recognizing that physiologic monitor alarms impact patient welfare. The World Health Organization (1999) guidelines for community noise recommend that (a) during the night, to promote good quality sleep, an individual's exposure to noise should be less than 30 dB(A), and (b) during daytime hours, an individual's exposure to noise should be

less than 35 dB(A). Moreover, the Occupational Safety and Health Administration mandates that employees should not be subjected to sound levels exceeding 85 dB over a period of 8 hr, and employees should never be subjected to a sound level exceeding 115 dB (Occupational Safety and Health Administration [OSHA], 2012).

The study by Rhyerd and Persson-Waye (2008) aimed to identify acoustic descriptors relevant for the sound environment and RNs' perceptual, psychological, and physiological reactions to noise in a neurological intensive care unit. The investigators discovered that the sound pressure level near the patients was typically 53–58 dB. This sound pressure level—which significantly exceeds the level stipulated in the WHO guidelines for community noise—approaches a hazardous noise level. Additional studies characterizing the clinical noise problem (Elliott, McKinley, & Eager, 2010; Johansson, Bergbom, Persson-Waye, Ryherd, & Lindahl, 2012; Monsén & Edéll-Gustafsson, 2005) confirm the neurological ICU findings reported by Rhyerd and Persson-Waye (2008). Elevated noise levels have also been described in neonatal and pediatric studies (Darcy, Hancock, & Ware 2008; Williams, van Drongelen, & Lasky, 2007). As with adult research on alarm-related sound levels, questionnaire surveys and staff interviews consistently identified monitors or alarms as being key contributors to increased sound levels (Darcy et al., 2008; Williams et al., 2007).

With nurse wellbeing and patient safety in mind, Daly and Wilson (1983) were able to determine that fatigue decreased overall ECG detection rates and increased latency on the vigilance of nurses. Although Daly and Wilson's study was conducted in a controlled setting, the study's findings suggest that fatigue significantly diminishes RN performance as measured by both response rate (i.e., time from alarm onset to nursing response) and arrhythmia detection rate.

Impact of alarm on health care organizations. In 2002, the National Quality Forum—an influential non-profit organization that aims to improve the quality of American health care—created and endorsed a list of "serious reportable events" (SREs) of great concern to both public and health care providers. Subsequently updated in 2006, SRE classification uses six categories: surgical, product-of-device, patient protection, care management, environment, and criminal. Product-of-device events include adverse patient events (i.e., death or serious disability) associated with the use or function of a device in patient care—specifically, instances in which a device is used or functions in a manner other than as intended (National Quality Forum [NQF], 2011). Notably, the recent 2011 HTF national survey on clinical alarms reported that almost one in five institutions experienced adverse patient events during the 2-year period prior to implementation of the survey. Survey respondents from slightly more than 20% of participating health care institutions reported that clinical alarm improvement initiatives were implemented in their institutions within the past 2 years; fewer than half of the respondents (47.5%) were unsure if action had been taken. These survey results are startling and re-affirm the need for future research to test interventions to reduce nuisance alarms, alarm burden, and consequent adverse events.

TJC first promoted the need to address the challenges associated with clinical alarms as the sixth goal of the 2003 National Patient Safety Goals (NPSG): "to improve the effectiveness of clinical alarms systems." This goal primarily focused on regular preventive maintenance, testing of alarm systems, assuring that alarms were activated with appropriate settings, and assuring that alarms were sufficiently audible with regard to clinician–alarm distances. In 2005, following achievement of a high degree of compliance, the goal was retired and incorporated into TJC standards (Catalano, 2005). Nearly a decade since the 2003 TJC patient safety goal was

set; researchers have compiled compelling documentation that alarm mismanagement is frequently a factor in sentinel events. In late 2011, TJC announced its investigation of including alarm management to improve the safety of clinical alarm systems in the 2013 NPSG for critical access hospitals (CAH) and hospital accreditation programs (HAP; Pujols-McKee, 2011). However, the proposal was not accepted. In 2013, TJC conducted a standards field review seeking input on a proposed 2014 NPSG for medical device alarm management. Months later, TJC approved and announced a new National Patient Safety Goal .06.01.01 on clinical alarm safety, which impacts all CAH and accredited hospitals (TJC, 2014). In Phase I (beginning January 2014), NPSG .06.01.01 encourages hospitals to establish alarm safety as an organizational priority and identify the most important alarms to manage on the basis of an organizational assessment. In Phase II (beginning January 2016), hospitals must develop and implement policies related to alarm management. In addition, the element of performance for NPSG .06.01.01 includes staff education and endorses practical, targeted, readily implementable interventions, such as leveraging existing technology (e.g., clinically appropriate settings for alarm signals) and promoting best practices (e.g., monitoring and responding to alarm signals) as an immediate strategy to minimize clinician alarm fatigue in health care organizations seeking accreditation.

Similar to the NQF and TJC, The Centers for Medicare and Medicaid Services (CMS; 2008) announced that Medicare would no longer reimburse hospitals for additional costs of treating preventable errors or conditions that could reasonably have been prevented—including errors or conditions considered 29 “Never Events” (i.e., events that should never occur; Centers for Medicare and Medicaid Services, 2011). By the same token, many states and private insurers have adopted similar reimbursement policies. In the context of a challenging economic climate,

rising consumer expectations, and the increasingly regulated and competitive nature of healthcare, stakeholders must develop collaborative academic and industry partnerships that will focus on effective and efficient alarm management approaches to protect patients.

Limitations of Current Research

Monitor alarm research is in its early stages and is fraught with challenges. Studies on physiologic monitor alarm burden is not readily amendable to experimentation; this difficult is due to well-entrenched clinical practices that are difficult to change, challenges with access, and lack of practical expertise and clinical experience among investigators in the hospital setting. The majority of alarm studies have been observational. To date, no multicenter randomized clinical trials have tested the effect of a comprehensive alarm intervention on nursing practice and patient outcomes. Existing foundational studies have primarily explored certain elements of monitoring, including alarm sources (i.e., parameter, technical, and arrhythmia alarms and their corresponding frequencies) and predictive value. Research has been performed in a variety of settings; however, most of this research on physiological monitor alarms has involved single sites and small convenience samples of patients (with related physiologic monitor alarms).

Another research limitation is that investigations have often examined the theoretical effects of various alarms settings and algorithms using hypothetical data —and such studies do not represent the “real world” of clinical practice. A chief reason for this approach is the technical difficulty of acquiring the necessary alarm data in an accurate and reliable manner. Obtaining data on recorded alarms and alarm adjustments requires direct observation, which introduces bias. An alternative approach to obtaining such data entails video monitoring, which is labor- and technology intensive. Alarm research has also been somewhat constrained by the lack of development of clinical informatics, which is yet a nascent discipline. In summary, the

limited research and the numerous deficiencies identified in the current review make multi-study comparisons and syntheses of research difficult.

Suggestions for Future Research

Future studies with more rigorous designs are sorely needed to understand alarm burden and to resolve problems associated with alarm fatigue and related patient safety issues. Both qualitative and integrated quantitative-qualitative designs would augment the existing body of literature. The majority of quantitative studies have used descriptive–observational designs and have primarily focused on measuring the incidence and frequency of alarms. Although these studies have established the significance of the alarm problem, more intervention studies are needed to identify and develop effective clinical strategies to reduce alarm burden and improve patient and organizational outcomes. Well-designed multicenter randomized clinical trial studies are required to determine optimal default alarm limits and provide evidence for practice guidelines. This work would greatly benefit patients and the health care community.

Future research should also comprehensively study arrhythmia alarms. For example, the increasing use of proarrhythmic pharmacologic agents warrants the routine use of QT–interval monitoring. Given that recent technology enables this computerized surveillance, it is regrettable that these features are often underutilized (Drew et al., 2004). Likewise, the ST-segment monitoring feature of physiologic monitors is currently underutilized, even in hospital units caring for patients with acute coronary syndromes (Funk et al., 2010; Patton & Funk, 2001; Sangkachand, Cluff, & Funk, 2012). This underutilization of available monitor features stems primarily from clinicians' concerns regarding adding to an already unmanageable number of non-actionable alarms and FAs. Studies evaluating the impact of activating these monitoring features on alarm frequencies may reveal the addition of relatively few alarms with the added

patient safety benefits of QT interval and ST-segment monitoring. Furthermore, additional studies that investigate reducing the notification of PVCs to a visual alert (rather than using an audible alarm) and the widening of the PVC alarm parameter thresholds on reported alarm frequencies and patient outcomes may also be beneficial.

Although the existing research has value, the failure to report physiologic monitor alarms based on actual patient cardiac monitoring times (rather than on unit proportions or on average daily census) and the lack of rigor in reporting patient outcome data in the research makes the assessment of intervention safety difficult. To test alarm fatigue interventions, it is imperative (a) to obtain Institutional Review Board approval for future prospective randomized clinical trials—many of which the reviewed studies failed to report and (b) to have a Data Safety Monitoring Board composed of clinical experts who can prospectively review patient outcomes and apply pre-determined stopping rules to halt a study if patient harm is observed.

Lastly, additional research should investigate the effects of the environmental noise that is generated by physiologic monitor alarms. Specifically, research should rigorously and reliably assess the unintended effects of this noise on patients, families, and care providers. These studies would provide a more holistic understanding of physiologic monitor alarm research.

Conclusion

Throughout the modern health care system, initiatives to promote efficient bed utilization, use of cutting-edge treatments, and implementation of clinical guidelines have led to increased use of medical devices across the continuum of care. Continuous physiological monitoring, once coveted and restricted to critical care, is now ubiquitous in multiple types of acuity units. Unfortunately, the proliferation of patient monitoring systems has not been accompanied by industry innovations or by a robust alarm management infrastructure among

organizations. Given the known hazards, physiologic monitor alarm complexity and burden have been underestimated and understudied.

Leading professional societies, academic and industry leaders, and accrediting agencies have begun to recognize the growing importance of alarm research. Among all licensed professional groups, ICU nurses have the greatest vested interest in studying alarm phenomena, given that, with their constant presence at the patient's bedside, these nurses are at greatest risk for developing alarm fatigue. Notably, patients remain the primary benefactors of this research and an organization's efficacious alarm management policies and procedures, because patients' well-being is dependent on the attention and actions of others.

While medical device industries and leading scientists are collaborating to improve the intelligence of physiological alarms and integrated alarm systems, health care organizations can adopt a "What can we do now?" position to reduce alarm burden, and consequently, risk of harm to patients. Academicians in partnership with clinicians must begin to research straightforward efforts to improve alarm efficacy by first systematically identifying, quantifying, and understanding physiologic monitor alarms. Once understood, responding appropriately by implementing interventions, considering specific patient populations, and evaluating their clinical effectiveness with meaningful outcomes is critical.

Above all, dissemination of physiologic monitor alarm research is essential, as alarm fatigue is a widespread problem that has detrimental effects for nurses, patients and their families, and health care organizations. Innovation and dissemination of knowledge regarding alarm safety initiatives is vitally important because this program of research, though highly specialized, is beneficial for all. Knowledge gained from future studies will build on prior research and can potentially minimize clinical risks associated with excessive alarms and alarm

fatigue. Furthermore, research findings will provide evidence for development of alarm practice standards and can serve as a model for other medical device alarm management and research.

Table 2.1. Physiologic Monitor Alarms: Summary

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Aboukhalil et al., 2008 USA	▪ A subset of stored patient records selected from the MIMIC II database ▪ 447 patients ▪ 41,301 hr ECG and ABP waveforms containing 5386 alarms	▪ Secondary analysis of a convenience sample of stored patient records from the Physionet MIMIC II database ▪ Data from adult patients in 48 medical surgical and coronary ICUs at an urban hospital	▪ Descriptive observational study ▪ Retrospective off-line analysis ▪ To test an algorithm to suppress false critical ECG arrhythmia alarms using morphological and timing information derived from the ABP waveform ▪ 5 categories of false critical ECG arrhythmia alarms include: asystole, extreme bradycardia and tachycardia, V-Tach, V- Tach/Fib	▪ An average of 43% of the critical ECG arrhythmia alarms were found to be false, with each of the five alarm categories having FA rates between 23% and 91% ▪ The FA suppression algorithm was able to suppress 60% of the false alarms, with FA reduction rates as high as 94% for asystole and 81% for extreme bradycardia ▪ FA reduction rates were lowest for extreme tachycardia (64%) and ventricular-related alarms (58%) for ventricular fibrillation/tachycardia and 33% for ventricular tachycardia) ▪ True alarm (TA) reduction rates were all 0%, except for VTach alarms (9%) ▪ The FA suppression algorithm reduced the incidence of false critical ECG arrhythmia alarms from 43% to 17%, where simultaneous ECG and ABP data were available	▪ Large heterogeneous patient sample ▪ Alarm definitions and thresholds well defined ▪ Detailed and strong critical arrhythmia annotation process ▪ Complex methodology and intensive study with seasoned investigators	▪ Non critical arrhythmia (parameter and technical) alarms were not included in this study given specific aims ▪ Clinical staff at the data collection site selected to use single lead arrhythmia analysis
ACCE- HTF, 2011 USA	▪ 81% response rate (3,454) ▪ Included RNs, RTs, MDs, Clinical Manager, Admin, etc.	▪ National on-line survey ▪ Convenience sample of acute care hospitals	▪ Cross-sectional study ▪ To re-survey the health care field to determine changes in the profession's perception of clinical alarm issues, improvements made at their facilities, and priorities for future action	▪ 76% strongly agree/agree that nuisance alarms occur frequently ▪ 71% strongly agree/agree that nuisance alarms disrupt patient care ▪ 78% strongly agree/agree that nuisance alarms reduce trust in alarms and cause caregivers to turn alarms off ▪ 78% strongly agree/agree that smart alarms would be effective to use for reducing FA and would be effective to use for improving clinical response ▪ Almost 1 in 5 institutions experienced an AE over the last 2 year with alarm initiatives ▪ Technological solutions to improve alarm safety were reported by < 20% of the responders	▪ Large heterogeneous sample (RN, RT, clinical engineers, biomechanical engineers, etc.) ▪ Strong response rate ▪ On-line re-survey utilizing same instrument ▪ Allowed comments Face validity ▪ Findings are consistent with initial survey in 2006 ▪ Participants received no monetary incentives	▪ Short survey (1 month) period which occurred from 08/8/2011 to 09/10/2011 ▪ No information on the validity or reliability of the instrument's ▪ Shift in demographics between surveys (less RN responses) ▪ Inclusion and exclusion criteria is vague ▪ Risk for selection bias as it is unclear who received the survey and if only an on-line survey was available (paper forms of the survey were distributed in 2006 study

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Biot et al., 2000 France	▪ 25 adult patients with acute respire-tory distress syndrome (ARDS) ▪ 3665 alarms ▪ 54 study periods (250 hr)	▪ Intensive Care Unit ▪ Convenienc e sample ▪ Lyon, France ▪ Marquette 7010™ (Marquette Electronics Inc) ▪ Cliptiper™, Ohmeda for pulse oximetry	▪ Descriptive observational study ▪ To evaluate the efficiency of a hemodynamic and respiratory monitoring system by a clinical analysis of the alarms	▪ 1 alarm every 6 minutes with 92% of audible alarms being FP alarms ▪ 1 ventilator alarm every 24 minutes with 8% of alarms being FP ventilator alarms ▪ Sources of alarms originated from: 44% of alarms from ABP, 17% of alarms from and 12% from P_{aw} ▪ The PPV for ABP was 51%, Spo_2 18% ▪ HR sensitivity was 81% and specificity was 99%, PPV 69% ▪ Arrhythmia sensitivity was 59%, specificity 99% and PPV 42% ▪ Only 5% of true positive alarms had clinical significance for patients care ▪ 9% of ABP alarms, 4% of SpO_2 and 4% of HR alarms were categorized as TP and clinically significant alarms warranting therapeutic action ▪ 175 false negative alarms reported with no clinical consequences as a TP alarm did not occur after event ▪ 50% of FN alarms were due to arrhythmia alarms or aspiration alarms	▪ Early study and findings support other studies with more rigorous methodology ▪ First study to report false negative alarms and report sensitivity, specificity, PPV and NPV for their 9 monitoring parameters ▪ Surrogate consent obtained ▪ Studied 25 patients with the same number of monitoring parameters (9) which has not been performed in other studies	▪ Small sample size focused on patients with ARDS (on mechanical ventilation with invasive pressure monitoring) which limits external validity ▪ MD in-room performing direct observations introduced potential for bias (Hawthorne effect) as data collection occurred for 5 hours during peak care times (7:99-12:00, 12:00-17:00, 17:00-22:00) ▪ All alarm parameters were set by treating physician (per standard of care) yet, it appears these could be modified ▪ Poor sensitivity and PPV for HR and arrhythmia alarms ▪ Potential for instrument bias /measurement error as 3 devices were utilized ▪ Investigators collected data on cardiac arrhythmias but do not provide a detailed report on validity of alarms (FP and TP) ▪ Should be translated into English

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Bitan et al., 2004 Israel	▪ 6-beds in a 22 bed NICU ▪ ADC: 6 patients ▪ 35 RNs ▪ 47 observation periods with total observation time of 30 hr	▪ Neonatal Intensive Care Unit ▪ Convenience sample of patients and RNs ▪ The Soroka Medical Center ▪ Beer Sheva, Israel ▪ Mennen Medical Horizon 1000 ™	▪ Descriptive observational study ▪ Measure the occurrence of alarms from different causes in a Neonatal ICU, record the nurses' reactions and analyze the relationship between the alarms and reactions	▪ A monitor alerted on average 16.74 times per hr and the average alarm duration was 15 s ▪ The percentage of time and duration of alarms differed between shifts ▪ The main cause of alarms were: ▪ Spo_2: 13.25 alarms/hr ▪ HR: 3.18 alarms/hr ▪ RR: .75 alarms/hr ▪ > 90% of alarms, the RNs did not attend to the infant during the minute following the alarm ▪ RNs were more likely to intervene after a relative rare alarm then after frequent alarms (oxygen saturation) ▪ RNs responded to longer alarms which may indicate that nurses effectively "filter out " short alarms while being aware of alarms and monitoring their duration ▪ Results demonstrate that RNs often do not respond directly to alarms, but rather use them as additional sources of information	▪ Research design is appropriate ▪ The introduction explains the central concept and objectives of the study ▪ Although not in depth, the literature review cites several published summaries of research and is sufficient to support their hypothesis ▪ Computed probabilities for determining the relationship between the nurses' actions and the alarms ▪ Acknowledged limitations	▪ Convenience sample of with undefined inclusion/exclusion criteria ▪ Homogeneous sample limits external validity ▪ No mention of conceptual and operational definitions, theoretical basis or institutional review board approval ▪ No discussion of inter-rater reliability ▪ Possible Hawthorne effect ▪ The recorded alarms consisted of only three monitored alarm parameters (HR, RR & Spo_2) ▪ Excluded monitor alarms that were triggered while the RN was treating the infant
Block et al., 2009 USA	▪ 1054 MDs from the USA ▪ 449 MDs from Finland ▪ 700 MDs from the Nederland ▪ 30 medical device companies	▪ Convenience sample of anesthesiolo-gists ▪ USA, Finland, and the Netherlands	▪ Cross-sectional survey ▪ Study the selection of alarm limits, reasons for turning off alarms, and the perceived incidence of false alarms	▪ Leading reason for turning off alarms was that there were too many false alarms ▪ Dutch MDs used lower settings for their high HR alarms ▪ Response rate: USA 12%, Finland 37%, Netherlands 24%	▪ Two different survey forms of equal number distributed ▪ Survey A provided as an example ▪ Study confirms the findings of a previous study regarding FAs ▪ Novel and early study that sought MD opinions on alarm practices	▪ No operational definitions or theoretical framework provided ▪ Mail-in survey ▪ Risk for selection bias as survey for monitor manufacturers was distributed at an exhibit ▪ Poor response rate ▪ Researchers developed the survey instrument and no reliability or validity information ▪ Acknowledged confusing questions on the questionnaire ▪ Unclear how MDs were selected to participate in the survey which introduce selection bias

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Blum et al., 2009 USA	▪ 14-bed Cardio-thoracic Intensive Care Unit ▪ 115 patients	▪ University of Michigan Medical Center ▪ Convenience sample ▪ Ann Arbor, MI, USA ▪ GE Solar 9000™ monitors, General Electric	▪ Descriptive observational study ▪ To address false positive alarms, a computerized architecture based on simple reactive intelligent agent (IA) technology was developed and implemented in a critical care unit to facilitate the investigation of deterministic algorithms for the improvement of sensitivity and specificity of physiologic alarms	▪ During the 28 day study, 1012 alarm events were triggered and captured by the IA system with system availability >99% ▪ In comparison, during this period a total of 293,049 alarms were generated using the default settings ▪ Specificity was predictably higher for extreme physiologic disturbances. As the cardiac index decreased from <2.0 l/min to <1.8 l/min, the specificity improved from 88% to 95% ▪ Specificity also improved as the physiologic disturbance persisted for longer periods of time ▪ There were 327 events with SBP <80 mmHg for more than 15 min and 282 periods with SBP <80 mmHg for more than 30 min. These alerts had specificities of 76% and 85% respectively. ▪ When SBP was low for 30 min compared to 15 min, the specificity improved by 11% ▪ Median filtered alerts for blood pressure drops were associated with generally high specificity, CVP trends were not. Clinically important rises in CVP above 20 mmHg were only found to have 33% specificity ▪ A high predisposition for events occurring during admission/discharge to the ICU ▪ 9% of alerts occurred within 30 min of patients entering the ICU, 84% occurred during patient stays, and 6% occurred as invasive lines were removed for departure from the ICU ▪ Data analyzed to determine the specificity of the alarm system using a standard specificity equation	▪ Practicable and efficient study employing a complex biomedical methodology ▪ True events were generally classified as those where the IA system generated some alarm in conjunction with the patient monitor that was classified by a clinician chart review, as a situation requiring clinician intervention ▪ Findings are credible and a statistically significant relationship between the variables exists	▪ Single site with homogeneous sample ▪ Short data collection period (4 weeks) ▪ Study limited to 4 alarm parameters (SBP, MAP, CVP & CI) ▪ Weak operational definitions and no theoretical framework cited ▪ Cumbersome methodology for the classification of the data set into false positive and false negative ▪ Potential for recall bias as MD and RNs caring for patients the prior day were consulted about the alarm to determine if the alarm occurred in response to a valid condition ▪ Authors acknowledged alerts generated by the system were based on median values over a minimum of 15 min meaning, physiologic instability would need to be occur for a minimum of 8 min in order to generate an alert for any 15 min period ▪ Setting had 24/7 MD coverage which may have contributed to the number of reduced alarms by rapid MD intervention and treatment

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Burgess et al., 2009 USA	▪317 patients on Medical-Surgical Units (MSU) ▪781 days ▪18,737 hr of monitoring	▪ Medical - surgical units at 2 hospital sites ▪ Stratified sampling strategy ▪ Secondary analysis from previous studies datasets and obtained from July 2003 to July 2006 ▪ Lifebed™ Patient Vigilance System	▪ Descriptive observational study ▪ Retrospective off-line analysis ▪ To provide objective data to assist with setting alarm limits for early warning system	▪ Study demonstrates that as alarms limits became more sensitive, more false positive alarms are generated and as settings become less sensitive, fewer false positive alarms are generated ▪ The following alarm limits appeared consistent with an appropriate level of false positive alarm rates: HR high 130-135 bpm, low 40-45 pm; RR high 30-35, low 7-8 breath/min	▪ The introduction conveys the importance of the study and is amendable to the discipline of nursing. ▪ The background and review of the literature is extensive ▪ IRB approval obtained in previous studies and dataset was de-identified and made anonymous ▪ Multisite- 2 hospitals with participation of several MSUs ▪ Large sample size ▪ Detailed descriptive statistics provided to assist quantify each alarm limit level to study the effect on RN workflow ▪ 1st study that makes reference regarding how to perform basic calculation of alarm frequency per individual RN (the number of patients a nurse cared for would be multiplied by percentages of patients with alarms per shift) but does not actually perform the analysis	▪ Potential for selection bias as the inclusion criteria excluded patients with < 12 hr of monitoring and those who experienced a CPA, unanticipated ICU transfer or sudden death during monitoring ▪ Study has limited generalizability ▪ Parameters investigated were limited to HR & RR ▪ Unclear oxygen saturation alarm limits not studied ▪ Analysis was limited to alarm parameter limits and did not integrate functional or technical device issues such as lead detachment or poor sensor contact ▪ The sample alarm limits are system specific with alarm delays of HR = 30 s and RR = 60 s and findings may not be generalizable to different systems using other algorithms ▪ Authors state a power analysis was not performed because the study was non-experimental yet, authors reported they aimed for a sample size >100 patients ▪ Study did not allow for subgroups or multivariate analysis ▪ Authors report risk for co-intervention and proficiency bias

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Chambrin et al., 1999 France	▪ 5 Medical Intensive Care Units ▪ 131 patients ▪ 3188 monitor alarms	▪ 2 university hospitals and 3 general hospital Intensive Care Units (ICUs) ▪ Convenience sample ▪ Northern France ▪ Make of monitors not specified	▪ Descriptive observational study ▪ To assess the relevance of current monitoring alarms as a warning system in the Adult ICU	▪ There were 3188 alarms, an average of one alarm every 37 minutes /patient ▪ The study confirms that the level of monitoring in the ICUs generates a great number of false positive alarms ▪ Alarm sensitivity is 97%, specificity 58%, PPV 27% and NPV 99% ▪ 69% of alarms occurred in non-sedated patients ▪ The distribution of alarms according to RN shifts was significantly different with days shift have the most alarms ▪ 23.7% of alarms were due to staff manipulation, 17.5% to technical problems and 58.8 % to the patients ▪ Alarms originated from ventilators (37.8%), cardiovascular monitors (32.7%), pulse oximeter (14.9%) and capnography (13.5%) ▪ Only 5.9% of the alarms led to MD notification (among which 51.3% were from the cardiac monitor, 23% from pulse oximeter, 20.9% from ventilator and 4.8% from capnography) ▪ Triggering the alarm was known for only 1982 (62.2%) alarms: 68.8% were classified as alerts and 31.8% as vital alarms and among the vital alarms 28.5% were induced by staff manipulation, 12.3% by technical problems and 59.2% by the patients	▪ Multi-centric study which occurred over a 16-month period between November 1996 to February 1998 ▪ Diagram for conceptual reasons and consequences of excessive alarms ▪ Reports alarm sensitivity specificity, PPV and NPV ▪ Detailed operational definitions for true-and false-positives, true and false negatives provided ▪ Voluntary experienced nurses were paid to record observations outside their normal job duties ▪ Data gathered on 3 shifts ▪ Appropriate statistical analysis; descriptive and hypothesis testing methods ▪ Estimated probability of treatment during the 1 minute following an alarm for all alarms	▪ No mention of theoretical foundation ▪ Relatively low quantity of alarms given study involved 5 ICUs over 16 months ▪ Unclear which model of physiologic monitors was utilized in the 5 ICUs which introduces variability ▪ Observer bias: Experienced RNs were asked to record the patient's characteristics and, for each alarm event, the reason type and consequence was completed on a form ▪ Hawthorne effect: Bedside RNs were observed by their peers (also RNs) which may have influenced their behaviors and actions ▪ The authors fail to state if inter-rater reliability (for RN observers) was performed to promote accuracy of alarm classification reporting ▪ While authors quantified parameter alarms, they did not quantify or study cardiac arrhythmias alarms ▪ Recorded some numerical measurements every 5 s and others every second with alarm and alarm settings entered manually into programs as part of the annotation

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Cropp et al., 1994 USA	▪ 100 subject: ▪ 25 MDs ▪ 41 RNs ▪ 34 RTs ▪ 33 alarm sounds	▪ Medical Intensive Care Unit ▪ Convenience sample of patients and staff in a 11 bed unit with a occupancy rate > 95% ▪ Ohio, USA ▪ Seimens™ monitor	▪ Quasi-experimental study ▪ Experimental model of ICU audio signal recognition study ▪ To determine whether clinical alarms (cardiac monitor and ventilator) are identifiable by sound alone to the ICU staff	▪ 43% of the alarm were identified ▪ For the critical alarms, the mean correct response rate was 50% and 40% for noncritical sounds ▪ Compared by occupation, RTs identified more sounds than RNs, nonregistered therapists or physicians ▪ Registered Therapist Technicians (RTT) were able to identify 51% of all sounds, 62% of the critical alarms and 46% of the noncritical ▪ RN scores were lower than the RRTs scores but higher than the MD scores and not significantly different than the nonregistered therapists' score. ▪ For the critical alarms, 26% of the variance in scores (r =.26) were attributable to occupation	▪ Heterogeneous sample of health care professionals ▪ Attained IRB approval for waived consent and anonymity provided ▪ Studied a variety of alarm sounds (23 non-critical sounds, 10 critical alarms). 12 of the 33 alarm sounds were directly related to the ventilation ▪ Reasonable reliability for the instrument (audiotape and testing procedure) o	▪ Convenience sample from single site ▪ Inclusion and exclusion criteria not specified ▪ No power analysis ▪ This testing was done in rapid succession and more representative of an auditory task ▪ No description of noncritical and critical alarm sounds was provided ▪ Study performed in laboratory like setting under controlled conditions which does not reflect the "real world" clinical environment with competing noise ▪ No visual alarm signals included in the study
Cvach et al., 2012 USA	▪ 2 adult units ▪ 15-bed Medical Progressive Care Unit ▪ 25-bed Cardiology Care Unit	▪ Non- ICU units at John Hopkins Hospital, Baltimore, MD ▪ Convenience sample ▪ GE Solar 8000i™ monitors, General Electric	▪ Quasi-experimental study ▪ Pre and Post Intervention study ▪ Investigate the impact of daily ECG electrode changes on technical cardiac monitor alarms	▪ The average alarms/bed/day was reduced by 47% and 56% on a medical progressive and cardiology care unit ▪ The overall technical alarms decreased by 34% and 45% on the medical and cardiology units, respectively. ▪ Post-intervention, a 10% increase in ECG leads fail alarms occurred in the medical unit while a 14% decline was observed in the cardiology care unit. ▪ In the medical unit a 60% decrease in arrhythmia suspend alarms was seen and a 74% reduction was witnessed in the cardiology unit	▪ Feasible and realistic intervention study ▪ First study to perform an intervention regarding technical alarms ▪ Involvement of 2 inpatient care units ▪ Nurse driven ▪ Demonstrated improved results (decreased frequency in total alarms)	▪ Short study and sample size with 8 days of alarm data pre and post-intervention ▪ Small sample of technical alarms ▪ Dedicated staff member utilized to perform this daily task during study ▪ No operational definitions or theoretical framework ▪ Unclear if IRB approval obtained ▪ It is unexplained why ECG lead fail alarms increase post-intervention on the medical unit ▪ Authors do not report associated costs, efforts to promote sustainability and impact to clinical workflow ▪ No mention of the effect of daily electrode changes on patient experience ▪ No annotation of arrhythmia alarms

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Daly et al.,1983 USA	▪ 96 RNs participated with 8 RNs lost to attrition (n=88)	▪ Urban hospital in the Midwest ▪ Convenience sample ▪ Make of monitors not specified	▪ Quasi-experimental study ▪ To determine the effects of fatigue on the vigilance of nurses observing continuous ECGs	▪ Age, gender and duration of work experience were studied and no significant differences between the groups were observed. ▪ Performance varied according to the duration of time over which the RNs observed the oscilloscope ▪ Fatigue affected the performance significantly as measured by both latency and correct detection rate, when fatigue was defined by comparing actual hours of sleep to reported usual hours of sleep ▪ The percent of correct detection was affected when fatigue was defined according to whether or not the RN had just completed 8 hours of work. The overall performance of the "rested" RNs was better than the fatigued" nurses ▪ "Rested" RNs had less variation in performance over time than had the fatigued RNS, whose performance declined from the 1st quarter to the second ▪ The performance of the rested compared to fatigue individuals was most similar during the first 20 min. In several grouping, the fatigued nurses initially performed slightly better to be followed by major deterioration in performance	▪ Early study appropriate for the time period and available devices and technology ▪ Operational definitions satisfactorily defined (vigilance, decrement latency, fatigue and well rested) ▪ Reference the Learning Theory as a theoretical foundation for the study and study based on classic vigilance, research ▪ To measure the effects of time and fatigue, ANOVA was utilized ▪ Study consisted of an 80-min test period (divided into 4 – 20 min periods) during which the RNs monitored an ECG on an oscilloscope and was instructed to identify PVCs	▪ Moderate sample size in a single site ▪ Unclear which model of physiologic monitors was utilized ▪ Dated study with older monitoring devices therefore difficult to promote the generalization of the study's findings ▪ Limited information provided on the sample and whether the RNs worked in a single unit or in a variety of departments (ICUs, telemetry units or the Emergency Department) which might affect the RNs' competency in PVC identification ▪ Author does not specify whether Institutional Review Board and informed consent was obtained from RNs ▪ Each RN was provided with a 2 minute training period which appears limited given the foreign environment and importance of the task ▪ Study performed in laboratory like setting under controlled conditions which does not reflect the "real world" clinical environment ▪ No replication studies found

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Görges et al., 2009 USA	▪ 12-bed Medical Intensive Care Unit ▪ 22 patients ▪ 1,214 monitor alarms ▪ 200 hr	▪ Medical Intensive Care Unit ▪ University of Utah Health Sciences ▪ Convenience sample ▪ Hewlett Packard (HP M1094B™) Philips Medical System	▪ Descriptive observational study ▪ To observe alarms in a Medical Intensive Care Unit to identify methods for reducing the number of false alarms by using time delays and the correlations between alarms and their clinical context	▪ Alarms occurred 6.07 times/hr , 23% were effective, 36% were ineffective, 41% were ignored ▪ The median alarm duration was 17s ▪ A 14-s delay before alarm presentation would remove 50% of the ignored and ineffective alarms and a 19-s delay would remove 67% ▪ Suctioning and ADLs caused 152 ignored or ineffective ventilator alarms ▪ 57% of the alarms occurred when a staff was in the room. If a 2 min alarm silence had been activated before suctioning, 74 alarms would have been prevented and the rate would decrease from 2.8 to 2.5 alarms/hr ▪ Infusion pump alarms had a high effective alarm rate (83%) because an alarm, once triggered, does not stop until the technical problem is resolved	▪ IRB obtained from the University of Utah Health Sciences Center's and informed consent obtained from 22 team members ▪ Medical devices well defined ▪ Random selection of patient room at the beginning of each observation day (a different patient was selected each day, except 1 patient was selected twice) ▪ Length of study sufficient (24-day period)	▪ Single site and convenience sample of patients which limits external validity ▪ Cumbersome methodology ▪ All monitoring was performed on weekdays from 09:00 to 17: 30 ▪ Single observer was not a clinician ▪ Failed to mention whether inter-rater reliability was established ▪ Reports on all clinical alarms (cardiac monitor and therapeutic devices) which is confusing ▪ Did not classify alarms as false and true; however, they classified them as effective, ineffective or ignored which adds complexity ▪ Did not annotate cardiac arrhythmias alarms
Graham et al., 2010 USA	▪ 15-bed Medical ▪ Progres-sive Care Unit ▪ ADC: 13 patients ▪ 30 RNs ▪ 16,953 monitor alarms	▪ Medical Progressive Care Unit ▪ The John Hopkins Hospital Baltimore, MD, USA ▪ Convenienc e sample ▪ GE Solar 8000i™ monitors, General Electric	▪ Intervention study to analyze monitor alarms to improve overall alarm management ▪ Determine RN knowledge of monitor alarms and the perceived response of the RNs to alarms	▪ Baseline alarm data: 16,953 alarms in 18 day period (942 alarms/24 hours) or 1 critical alarm every 92 s ▪ Post-intervention : 9,647 alarms ▪ Alarms were reduced by 43% ▪ Pretest: 83% of RNs changed alarm parameters when vital signs changed, 78% of RNs changed parameters at the beginning of their shift, and 41% of RNs reported compliance with changing alarm parameters after patients returned from temporary sites ▪ Post-test: 94% of RNs changed alarm parameters when a vital signs changed and changed parameters at shift star, 56% of RNs reported compliance with changing alarm parameters after patients returned from temporary sites	▪ Feasible, realistic and replicable study ▪ Multidisciplinary team approach ▪ Inclusion and exclusion criteria were not cited ▪ Good representation of real-world events ▪ Innovative approach to describing causes of nuisance alarms and appropriate nursing interventions to decrease their occurrence	▪ Small, homogeneous and convenience sample ▪ Short data collection ▪ No mention of operational definitions, theoretical basis, and approval from IRB for study ▪ Do not specify length of time post-intervention data collection period lasted (18 day?) ▪ Did not annotate cardiac arrhythmia alarms ▪ Advisory alarms not included in the analysis ▪ Developed a RN survey but they do not address instrument 's validity and reliability criteria ▪ Low RN survey response rate

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Gross et al., 2011 USA	▪79 medical-surgical sub-acute care beds community hospital ▪4104 patients with 4393 monitor alarms in 1040.5 hr ▪Subset of 30 patients with 2218 alarms in 529.5 hr	▪ Sub-acute medical and surgical units ▪ Community hospital Arizona, USA ▪ Convenience sample ▪ Intellivue MP5, and Telemetry System™, Philips Healthcare	▪ Descriptive observational study ▪ To assess alarm limits and workflow impact using 2-way audio/video system and population management system (eICU)	▪ For all patients, the average hrof monitoring per patient was 16.5 hr ▪ (*s*=8.3, median 22 hours) ▪ The average number of alarms per patient was 69.7 alarms (*s*=90.3, median = 28) ▪ When adjusted for the duration of monitoring, the average alarms/patient/day rate was 95.6 (*s*= 124.2, median = 34.2) alarms Adjudicated sample: ▪ There was a total of 13.1 critical alarms /pt/day 1 (*s*= 21.4, median 6.0) ▪ 34% of critical alarms were true and 63% of the high priority alarms were also true based on adjudication ▪ Alarm with greatest frequency was apnea, followed by desaturation, and tachycardia or ventricular tachycardia Off-line: ▪ Analysis of alarm history resulted in ability to reduce HR alarm burden by >50% with an adjustment of high HR from 120 to 130 bpm ▪ A 36% or 65% reduction in Spo_2 alarm burden by reducing Spo_2 limit from 90% to 85% or 80%, respectively	▪Multi-unit study in single institution ▪IRB approval obtained ▪Occurred over 9 months from April 2009 to January 2010 ▪First study to adjust the number of alarms to the duration of monitoring ▪Alarms reviewed by two independent researchers for validity and third researcher if necessary ▪Result of the study are legitimate and possess clinical and statistical significance ▪Sound study design (large sample, specific inclusion and exclusion criteria) ▪Reported alarms for study units then performed an in-depth analysis of a smaller sample of patients ▪Investigators report that a separate study was conducted to look at workflows, and satisfaction	▪ No mention of theoretical foundation or definition of concepts studied ▪ Study focused on alarms related to the frequency of alarms over specific physiologic thresholds ▪ Reported on large sample results and then adjudicated sample - unclear at times ▪ Reported information based on classification of alarm (critical alarms)by institution ▪ Unclear why Tach and VT/VF alarms were combined in the analysis ▪ Conducted adjudication of all alarms based on a smaller sampling (n=30). Uncertain why authors selected such a small sample and inclusion and exclusion criteria for this subset analysis vague and there is a potential for selection bias ▪ For the average number of alarms per patient per day there is a large standard deviation (*s*= 90.3) which indicates a skewed distribution and that most alarms were perhaps a result of a small number of patients

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Jacobs et al., 2007 USA	▪293 acute care medical-surgical patients	▪Tertiary care hospital ▪Honolulu, HI, USA ▪Convenience sample ▪Sham device	▪ Descriptive observational study ▪ To investigate the effect of an early prototype of an automated noninvasive no contact early warning system on patient outcomes and nursing satisfaction	▪ For patients being actively monitored, the system notified staff an average of 1.45 times per 24-hr patient day ▪ 53% of alerts were triggered by potential unauthorized bed exit or abnormal physiological conditions ▪ 12 alerts notified the RN staff of conditions with potential clinical consequence: 5 attempted unauthorized bed exits, and 7 HR or RR abnormalities ▪ The remainder of the alerts were not clinically significant 91.9% of RNs agreed or strongly agreed they liked the system	▪ Novel study utilizing a sham-device which investigated bed exit alarms along with cardiac monitor alarms ▪ Study conducted over sufficient time period from May 2005 to January 2006 ▪ Explored RN satisfaction	▪ Very weak study and findings published in low-impact journal ▪ Methodological flaws threatens the study's internal and external validity ▪ Results should be interpreted with caution
Korniewicz et al., 2008 USA	▪ 1327 partici-pants ▪Included RNs, RTs, MDs, Clinical Manager, Admin, etc.	▪Convenience sample of health care professionals in the USA	▪ Cross-sectional survey ▪ To determine the problems associated with alarms in hospitals and to evaluate reasons healthcare workers do not respond to clinical alarms	▪ Participants (>90%) agreed or strongly agreed regarding the need for prioritized and easily differentiated audible/ visual alarms ▪ Most respondents identified frequent false alarms as problems (81%), and most agreed or strongly agreed that nuisance alarms disrupt patient care (77%) and that such alarms can cause healthcare workers to "distrust" the alarms and disable the devices (78%) ▪ 49% of respondents believed that having a "monitor watcher" is helpful and 34% were neutral ▪ 54% of participants see usefulness in integrating alarm information with communication systems (e.g., pagers) and 30% were neutral	▪ Survey developed by a national taskforce – satisfactory face validity ▪ Occurred over 5 months from August 15, 2005 to January 16, 2006 ▪ Strong response rate ▪ Findings are consistent with other research performed on monitor alarms ▪ Participants received no monetary incentives ▪ Large heterogeneous sample (RN, RT, clinical engineers, biomechanical engineers, etc.) ▪ On-line survey	▪ Unclear how the participants were recruited to participate in the survey ▪ Authors report survey was administered on-line and was also made available in a paper version yet, tables demonstrate only on-line survey results ▪ Demographic data lacks specificity (academic medical centers vs. community hospitals, hospital departments other than critical care, years of experience interval too broad) ▪ No operational definitions or theoretical frameworks noted ▪ Authors developed a survey but do not report the instrument 's validity and reliability criteria

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Koski et al.,1990 Finland	10 adult cardiac post-operative patients ▪ 2322 monitor alarms	▪ Adult Intensive Care Unit ▪ Kuopio University Central Hospital ▪ Kuopio, Finland ▪ Convenience sample ▪ Kone 565, Kone Monitoring System™	▪ Descriptive observational study ▪ Determine both the frequency and the significance of audible limit alarms during the postoperative hemodynamic monitoring of cardiac patients	▪ Total of 400 hours approximating 26 hours per patient ▪ The total number of individual parameter alarms was 2322 and the number of alarm events, occasions when the alarms were activated by one or more parameters, was 1307 ▪ 87% of alarms were insignificant (related to artifact or undue alarms) ▪ 12% of alarms were deemed significant and required a patient's condition to be checked or a treatment provided ▪ The total number of alarms recorded per period was high during the rewarming, recovery and rehabilitation period ▪ A lower number of alarms were observed during the weaning period and the lowest during the follow-up period	▪ This is an early study and demonstrates excessive monitor alarms was recognized as an patient safety issue over 30 years ago and remains a significant patient safety issue ▪ A valiant attempt to quantify monitor alarms related to parameters in a specific patient population. ▪ The lack of technology is representative of the historical period of when the study was conducted	▪ Very small convenience sample of 10 post-operative cardiac patients with undefined inclusion and exclusion criteria monitored over 5 sub-periods ▪ No mention of conceptual and operational definitions, theoretical basis and approval from IRB for study ▪ Limited parameter alarms studied (HR, SAP, SAPm, SAPd, PAPs, PAPm, PAPd, and CVPm). ▪ SpO_2 alarms were not reported ▪ Instrument developed to capture /record alarm events but authors do not state the tool 's validity and reliability criteria ▪ Response bias: RNs recorded the time and parameters of every alarm activated and the significance of the alarm per alarm classification. RNs classified all alarms during their normal working time ▪ Short study period: from admission and ending at the 1st post-operative day, 6 hrafter removal of chest drains ▪ Study is weak due to the relatively short duration, lack of control for threats to internal validity and external validity ▪ Results should be regarded with uncertainty

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Koski et al., 1995 Finland	▪23 respond-ents ▪21 MDs ▪ 2 cardiac surgeons	▪ Convenience sample of medical staff in 5 Finnish hospitals	▪ Cross-sectional survey ▪ To assess the MDs' present use of limit alarms and their decisions when therapeutic responses are required ▪ To survey the routine use of monitors in post-operative cardiac patients	▪ No correlations were found between the limit values and how long these values were tolerated without therapeutic response ▪ The MDs clinical experience had no statistical effect either on the chosen limit or on how long the MDs tolerated the suboptimal values ▪ Physicians routinely used few of the monitors' parameter limit alarms resulting from poor specificity and from the time-consuming process of configuring individual parameter threshold alarm limits	▪ Good attempt to identify MD alarm practices ▪ Used a standardized questionnaire	▪ Small homogeneous sample in a single service ▪ 2 anesthesiologist responses were excluded due to misinterpretation of the instructions ▪ No information on instruments' validity or reliability ▪ Investigated 5 alarm parameters alarm limit threshold (HR, SAP, CVP, PAP and $ETCO_2$) ▪ No mention of incentives for participants
Lawless,1994 USA	▪16-bed PICU ▪ADC: 11 patients ▪2176 monitor alarms ▪ 2000 hours over a 7-day period	▪ Pediatric Intensive Care Unit ▪ Alfred I. DuPont Institute, Wilmington DE, USA ▪ Convenience sample ▪ Hewlett Packard monitors 78532A and 78534C™ ▪ Nellcor™ monitors for pulse oximetry	▪ Descriptive observational study ▪ Determine the predictive value of commonly used alarm systems in a PICU in identifying clinically important events	▪ Observations performed over 928 of a possible 1680 hours of patient care (55% response rate) ▪ There were 2,176 alarms observed: 1,481 (68%) false, 576 (26.5%) induced, and 119 (5.5%) significant ▪ There was one significant alarm for every 17.3 false or induced alarm ▪ There averaged one alarm every 26 min but one significant alarm each 7.8 hours ▪ More alarms during the day shift ▪ 44% of alarm related to pulse oximeter, 31% from the ventilator, 24% from the monitor and 1% from the capnograph ▪ Number of FAs and alarm soundings /patient/shift was greater during the day as opposed to evening/night shift ▪ The percentage of significant alarms as compared with FAs was greater during evening and night shift as opposed to day shift	▪ Concepts an operational definitions were well defined for: false, induced and significant alarms and for false negatives, false positives, true positives and true negative ▪ Performed 3-day trial period prior to study ▪ IRB approval obtained ▪ Study included data on ventilator and capnography alarms ▪ Used a protocol running with defined alarm limits which were set through fixed percentages of the initial value observed during a stable period	▪ Pediatric patient population affects external validity (generalizability) ▪ Small sample size; study conducted in a single unit over 7 days ▪ No mention of theoretical foundation ▪ Study utilized self-reported data from RNs, RTs, MDs to capture alarm information ▪ RNs classified all alarms during their normal working time ▪ Authors failed to provide specifics and cite inter-rater reliability of reporters (other than they offered a series of in-services) ▪ Authors developed an instrument to capture /record alarm events but do not specify the instrument 's validity and reliability criteria ▪ Potential for response bias

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Mäkivirta et al., 1991 Finland	▪ 10 post-cardiac surgery patients	▪ Adult Intensive Care Unit ▪ Kuopio University Central Hospital ▪ Convenience sample ▪ Kone 565, Kone Monitoring System™	▪ Descriptive observational study ▪ To study the effects of removing brief variations in the monitoring data on the quality of limit alarms during the postoperative hemodynamic monitoring of cardiac patients	Selected median filters: ▪ Increasing the length of the median filter decreased the incidence of false alarms ▪ The FA frequency was reduced by more than two-thirds ▪ The ratio of correct alarms to total alarms increased with increasing filter length from 24% (reference patient monitor) to 56% (1 min delay) to 73% (2.5 min delay) Dual-limit alarm simulation: ▪ The greatest number of alarms was triggered by deviations in the SAP (40%) ▪ Deviations in PAP accounted for 40%, mean CVP 22% and HR 7% of alarms ▪ The frequency of correct alarms increase from .29 to 4.3 alarms/patient hr ▪ The frequency of FA decreased from 12 to 4.5/pt monitoring hour	▪ Preliminary study and satisfactory attempt to analyze use of median filters and novel dual-limit alarm simulation ▪ Authors designed and clinically tested a novel dual-limit alarm system with two median filters which demonstrates benefits of alarm delays	▪ Research design is appropriate yet the study has a small sample size ▪ Homogenous sample of patients ▪ Limited monitoring period (from the time of admission to the ICU and ending 1 hour after the patient's peripheral temperature measured on the first toe had reached 31^{o} C) ▪ Limited parameter alarms studied (HR, SAP, SAPm, SAPd, PAPs, PAPm, PAPd, and CVPm)
Mäkivirta et al., 1994 Finland	▪ 10 post-cardiac surgery patients ▪ Total monitoring time 52.6 hr	▪ Adult Intensive Care Unit ▪ Kuopio University Central Hospital ▪ Convenience sample ▪ Kone 565, Kone Monitoring System™	▪ Descriptive observational study ▪ To determine the distribution of invasively measured hemodynamic data to enhance the reliability of monitor alarm systems	▪ 15% of all registered values fell outside of commonly applied alarm limits ▪ Doubling the pre-alarm delay from 5 to 10 s reduced the mean alarm rate by 26% ▪ A further decrease of 8% in the alarm rate was observed when a multidimensional vector median filter was used to remove the variable value interdependencies ▪ Preprocessing can decrease the alarm rate effectively ▪ Multidimensional preprocessing may produce more reliable alarms than one-dimensional processing	▪ Research builds on earlier study (1991) ▪ Authors provided brief operational definitions	▪ Research design is appropriate yet the study has an insufficient sample size ▪ Very small sample of patients and monitoring hours. ▪ Limited monitoring period (from the time of admission to the ICU and ending 1 hour after the patient's peripheral temperature measured on the first toe had reached 31^{o} C ▪ Limited parameter alarms studied (HR, SAP, SAPm, SAPd, PAPs, PAPm, PAPd, and CVPm) ▪ Wide confidence intervals (90%) related to small sample ▪ Research should be regarded with uncertainty

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
O'Carroll, 1986 USA	▪ 26 patients ▪ 1455 alarms ▪ Mean ADC: 4.5 patients	▪ Intensive Care unit ▪ Leicester Royal Infirmary ▪ Leicester, England ▪ Convenienc e sample ▪ Make of monitors not specified	▪ Descriptive observational study ▪ To provide data on the frequency and origins of alarms in a general intensive care unit	▪ 8 of the 1455 clinical alarms indicated potentially life threatening problems ▪ Only 25% of the monitor alarms indicated a dysrhythmia while 75% were false alarms triggered by the removal of ECG leads ▪ Ventilator alarms were mostly triggered by manipulation ▪ Infusion pump alarms were triggered by incorrect alignment or for no observable reason ▪ Little variation in the overall frequency of alarms by day or night ▪ The number of alarms per patient varied considerably ▪ Majority of alarms signaled minor malfunctions rather than events which threatened patient safety	▪ Early study appropriate for the time period and available devices and technology ▪ Sufficient study period 3 week period ▪ Good attempt to quantify all clinical alarms in ICU (cardiac monitor, ventilator and infusion pumps)	▪ Dated study with older monitoring devices therefore difficult to establish external validity ▪ Unclear which model of physiologic monitor was utilized although investigator reports manufacturer for ventilators and pumps ▪ Small sample size in a single site therefore difficult to generalize findings ▪ Limited information provided on the sample, setting and inclusion and exclusion criteria ▪ Conceptual and operational definition not described ▪ No reference to a theoretical framework ▪ The nursing staff was responsible for manually recording all alarm calls in the ICU during a 3 week period. No dedicated observers which introduces bias (observer, social desirability, novelty, etc.) ▪ Monitor alarms captured limited to 3 alarms (leads removed, tachycardia and bradycardia) ▪ No information regarding instrument utilized to capture frequency and type of alarm was provided ▪ Study lacks comprehensive of types of monitor and clinical alarms yet, appropriate for the time period

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Pan et al., 1994 USA	▪ 46 patients under-going elective surgery	▪ University medical center ▪ FL, USA ▪ Convenienc e sample of patients ▪ Pulse oximetry monitor Model 3700™, Ohmeda	▪ Descriptive observational study ▪ To determine when and how often discrepant data occur during intra-operative pulse oximetry to devise a way to prevent discrepant data from leading to false alarms	▪ The number of episodes per hour of discrepant oximeter data and the duration of the episodes were recorded by phase of anesthesia: induction, maintenance, and emergence ▪ Discrepant data occurred most frequently and lasted longest during emergence ($p < 0.05$); the majority of episodes of discrepant data during emergence lasted less than 12 s ▪ Excluding discrepant data that lasted less than 12 s decreased the frequency of discrepant data by 63% and excluding those that lasted less than 30 s decreased the frequency of discrepant data by 93% ▪ An alarm delay triggered by discrepant data and lasting 12 - 30 s would keep most discrepant data from becoming FA.	▪ Satisfactory description of discrepant data ▪ Use of single type of device and non-disposable finger pulse oximeter probe secured to finger ▪ Authors state the accuracy of the algorithm was validated by manually observing the oximeter while it screened a set of previously recorded oximeter data that contained discrepant data ▪ Appropriate statistical analysis (chi-square, analysis of variance and Tukey test)	▪ Small sample size in single site ▪ No theoretical framework specified ▪ Authors failed to describe care team involved in application of pulse oximeter probe and Spo_2 monitoring during 3 phases ▪ Sample of pre and postoperative patients in 3 phase of anesthesia which limits generalizability of findings ▪ Cumbersome recording of simultaneous ECG on paper

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Rheineck-Leyssius et al., 1998 Netherlands	▪ Data from 200 post-operative patients	▪ Post Anesthesia Care Unit (PACU) ▪ Twenteborg Hospital ▪ The Netherland ▪ Convenience sample ▪ Pulse oximetry monitor Criticare 504™ Criticare Systems Inc,	▪ Descriptive observational study ▪ Retrospective off-line analysis ▪ To better understand the potential effects of different settings on alarm generation and identification of hypoxemia	▪ Baseline limits: built in signal averaging epoch was 3 s and low Spo_2 limit was 85% ▪ Median duration of Spo_2 monitoring was 46 min ▪ During 1.5% of the time Spo_2 values were artifact and during 2% hypoxemic ▪ 830 episodes of hypoxemia ($Spo_2 \leq 90\%$) and 73 episodes of severe hypoxemia ($Spo_2 \leq 85\%$) occurred ▪ With an alarm Spo_2 limit of 90%, the alarm was triggered 1535 times (830 true, 705 false) ▪ With a Spo_2 alarm limit of 85%, the alarm was triggered 282 times (73 true, 209 false) ▪ Artifact rejection reduced alarms by almost 50% ▪ An alarm delay of 6 s or an averaging or median filtering epoch of 10 s resulted in an alarm reduction of almost 50% ▪ No differences in the reduction of alarms between averaging and median filtering ▪ Changing the alarm limit to 85% reduced alarms by 82% ▪ A similar reduction of alarms was obtained with either an alarm delay of 18 s or an averaging or median filtering epoch of 42 s ▪ An alarm limit of 85% reduced the false alarms less than the other three algorithms ($p<0.01$) ▪ Statistical analysis: Two-way ANOVA was used to compare differences in the frequency of alarms between averaging and median filtering. Chi square was used to test for differences in the ratio of artifacts to hypoxemia and Kruskall-Wallis tests for non-parametric data	▪ Early study focused on Spo_2 alarm parameter setting ▪ Feasible, realistic and replicable study ▪ Operational definition for artifact and hypoxemia provided ▪ Researchers calculated the effects of 5 different methods (artifact rejection, alarm delay, averaging median filtering and decreasing the lower alarm limit) on the number of alarms generated and detection of hypoxemia	▪ Single site and convenience sample ▪ Homogenous sample of post-operative patients in a PACU which may limit generalizability of findings ▪ No inclusion and exclusion criteria provided ▪ No conceptual model or theoretical framework specified ▪ Unclear if sample of patients include adult and pediatric patients ▪ Unclear if PACU staff were able to individualize customize Spo_2 alarm parameters ▪ Potential for higher nurse to patient staffing ratios in the PACU environment which ▪ Short duration of Spo_2 monitoring time (46 min) with wide range (13-200 min) ▪ Data was processed off-line and the effect of delayed alarm generation on clinical course is unknown as study does not report clinical outcomes ▪ Study does not address human factors with regards to the PACU RN and Spo_2 monitoring devices ▪ No mention of multidisciplinary team involvement

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Ryherda et al., 2008 Sweden	▪ Neuro ICU	▪ Neuro ICU at a major research hospital ▪ Göteborg, Sweden ▪ Convenience sample of patients and RNs	▪ Descriptive observational with cross-sectional survey study of RNs ▪ To perform sound measurements and occupant evaluations	▪ Measurements near the patients showed average L(Aeq) values of 53-58 dB ▪ The mean length of restorative periods (L(Aeq) below 50 dB for more than 5 min) was 9 and 13 min for day and night, respectively ▪ 90% of the time, the L(AFMax) levels exceeded 50 dB and L(CPeak) exceeded 70 dB ▪ Dosimeters worn by the staff revealed higher noise levels ▪ Personnel perceived the noise as contributing to stress symptoms ▪ RNs (38%) thought that an intensive work day with many alarms affected their sleep ▪ RNs (43%) thought that the alarm noises influenced their ability to perform their job tasks ▪ Almost half of the RNs surveyed (49%) said they sometimes adjusted the alarm levels so that they would not hear them ▪ 70% of RNs conceded that adjustments to the alarms should be documented ▪ 91% of RNs felt that noise may negatively affect them in their daily work environment ▪ 66% RNs reported experiencing irritation, fatigue (66%), concentration problems (43%), and tension headaches due to the sound environment (40%) ▪ 49% of RNs reported they often discussed the sound environment, ▪ 96% of RNs felt that noise contributed to the development of ICU syndrome in patients	▪ Feasible and realistic study with team approach ▪ Study performed in a dedicated Neurological Intensive Care Unit ▪ Research explores implications of unit noise levels on nursing practice surrounding alarm management strategies ▪ Study explored perception of noise in relationship to clinical alarms and provides valuable insight from the nursing perspective ▪ Authors acknowledge future work is warranted	▪ Short study (5 days) with 3 patients in monitored room ▪ During the noise measurement period, two nurses were always on staff in the measurement room, with one being lower and one higher rank and it is unclear how many other clinicians interacted in the space ▪ Patients on a variety of medical devices generating sounds and alarms ▪ Acuity of patients (and number of therapeutic medical devices) may have impacted noise levels yet authors did not cite the use of standardized patient acuity scale

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Schmid et al., 2011 Germany	▪ 25 consecutive patients schedules for elective cardiac surgery ▪ 124 hr of monitoring ▪ 8975 monitor alarms: ▪ 7556 hemodynamic alarms ▪ 1419 ventilator alarms	▪ Operating Room ▪ University Medical Center, Hamburg-Eppendorf ▪ Germany ▪ Kappa XLT™; Dråger Medical	▪ Descriptive observational study ▪ To characterize the patterns of alarms of a current patient monitor and an anesthesia workstation during elective cardiac surgery with the use of extracorporeal circulation	▪ A mean density of alarms of 1.2/minute ▪ For each procedure 359 +/- 158 alarms were recorded ▪ 30% of alarms were false ▪ Of the FAs, 40% were caused by artifacts or manipulations of patient monitoring system or its sensors ▪ In 28% , the reasons for the artifacts could be identified as blood draws or flushing of the arterial line ▪ Overall reaction time to the alarms amounted to 4 +/- 43.67 seconds ▪ 6386 life-threatening alarms ▪ 2589 alarms were technical ▪ 96% of life threatening alarms were caused by threshold ▪ Of the serious alarms, 70% were valid and 30% were artifact ▪ Of the valid alarms (39%), 39% were relevant and 61% were not	▪ Well-designed study ▪ Strong methodological approach ▪ Investigators included ventilator alarms in analysis ▪ First intra-operative study involving > 5 patients	▪ Small homogeneous sample in a controlled environment ▪ Short intra-operative surgery (4.95 hours +/-.96 hours) ▪ Single site ▪ Unclear how many anesthesiologists were involved in the study ▪ No explanation provided for utilizing fixed alarm settings for all patients ▪ No information on arrhythmia alarms ▪ Difficult to compare results with other studies as most performed in ICUs ▪ Examined a combination of a single patient monitor and an anesthesia workstation
Siebig et al., 2009 Germany	▪ 185 ICUs ▪ Response rate: 30% (274 of 915) ▪ 160 MDs (58.4%) ▪ 114 RNs (41.6%)	▪ Randomized sample of ICUs across Germany ▪ Utilized a representative stratified sampling scheme to account for type of hospital	▪ Cross-sectional survey ▪ To learn more about staff reception of monitoring alarms and their clinical judgment of techniques designed to reduce the number of alarms in order to reduce acoustic stress and maintain patient safety	▪ Most respondents (87%) estimated that less than 50% of alarms result in clinical consequences ▪ The number of alarms was judged as too frequent in 76% and much too frequent in 13% of respondents ▪ 93% of respondents used monitors from 3 different manufacturers and satisfaction did not differ devices ▪ 52.2% of the RNs considered themselves as the only ones regularly controlling alarm limits, ▪ 47.8% of RNs thought modifying alarm limits was also done by MDs and 90% of MDs regarded RNs and MDs controlling alarms together ▪ 54% and 75% of respondents agreed that smoothing of a signal to prevent artifact is beneficial ▪ 60% of subjects confirmed new alarms combining HR & BP alarms was helpful	▪ Representative survey ▪ 1 year study from May 2006 to June 2007 ▪ Randomized sample of ICUs ▪ Content validity established ▪ Findings confirm previous opinion survey on cardiac monitor alarms ▪ Self-administered questionnaire with 24 partly closed-ended partly open-ended questions	▪ Low response rate despite long study period ▪ No operational or definitions provided ▪ No mention of conceptual model or theoretical foundation ▪ Potential for selection and social desirability response bias as each medical director received 5 questionnaires for distribution to 5 MDs or RNs working in the ICU ▪ No inclusion or exclusion criteria regarding who could participate in the survey from each ICU (years of experience, AM or PM shift, Residents, Fellow or attending MDs) ▪ Unclear if a power analysis was performed

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Siebig et al., 2010 Germany	▪ 12- bed Medical Intensive Care Unit ▪ 5934 monitor alarms ▪ 68 patients, with 982 hours of monitoring	▪ Medical Intensive Care Unit ▪ Hospital of University of Regensburg ▪ Germany ▪ Convenience sample ▪ Infinity Patient Monitoring System™, Dråger Medical	▪ Descriptive observational study ▪ To validate cardiovascular alarms in patients by generating a data base of physiologic data and clinical alarm annotations and report the current rate of alarms and their clinical validity using video recordings	▪ 5934 alarms corresponding to an average of 6 alarms per hour were recorded ▪ The video annotation of the clinical validity of alarms revealed only 885 (15%) of all alarms indicated a clinically relevant situation ▪ Most of the generated alarms were threshold alarms (70%) and were related to arterial blood pressure (45%) ▪ 40% of all alarms did not correctly describe the patient condition and were classified as technically false; 68% of those were caused by manipulation	▪ Satisfactory definitions of key concepts and study received IRB approval ▪ General statement regarding technical validity and clinical relevance of alarms was provided ▪ Moderate amount of alarms compared to other studies ▪ The video assisted annotations offered the possibility to reduce bias ▪ Results are credible and strengthen previous studies	▪ Single site and unit with 68 patient cases with 982 hr of monitoring ▪ No description of inclusion and exclusion criteria ▪ No clear description of “other serious alarms” ▪ Authors state that MDs with at least 3 years’ experience in intensive care medicine annotated the data on the basis of video but they do not specify the quantity of MD observers or why other professionals (RNs) were not utilized to perform the annotation ▪ The quantity of arrhythmia alarms that were annotated for clinical validity is very small (n=104) and were classified as alarm-relevant, not alarm relevant or alerting –not TP or FP ▪ False negative alarms with no alarm occurring during an alarm relevant situation were not discovered due to the study design (fast-forwarding of video to the next alarm)

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Siebig et al., 2010 Germany	▪ 12- bed Medical Intensive Care Unit ▪ 3682 cardiac monitor alarms ▪ 38 patients with 515 hr of annotated monitoring	▪ Hospital of University of Regensburg ▪ Germany ▪ Convenience sample ▪ Infinity Patient Monitoring System™, Dråger Medical	▪ Descriptive observational study ▪ Pilot study to validate alarm procedure ▪ To analyze the alarm frequency and reliability of the alarm system ▪ To establish a reference database based on extracted data and video recordings	▪ Preliminary study results for larger study by research team. ▪ 2512 (68.2%) of alarms were threshold alarms, 535 (14.5%) other serious alarms, 535 (14.5%) technical alarms and 100 (2.7%) arrhythmia alarms. ▪ The vital parameters that caused the most alarms were systolic arterial blood pressure (45.4%), oxygen saturation (29.5%), mean arterial blood pressure (7.3%), heart rate (7.1%) and RR (4.8%). ▪ 54.5% of alarms were judged as technically true, 43.6% as technically false, and 1.9% could not be assessed. ▪ High percentage of alarms (44.2%) were caused by manipulation ▪ 16.6% of all alarms was classified as alarm relevant ▪ In 46.5% of the cases, the situation was non alarm relevant and in 35% of all alarms, the alarms were judged as alerting	▪ Definitions of alarms used for annotations are given as well as examples for technical true and false alarms ▪ Authors provided the taxonomy used for alarm annotation ▪ Study offers a method for collecting real-time monitoring data in conjunction with annotations of clinically relevant events by means of video monitoring ▪ Intra and inter variability obtained and satisfactory ▪ Observer bias is reduced through the use of video recordings ▪ Alarms were annotated by an experienced MD	▪ Single site and small homogeneous sample ▪ No theoretical framework or conceptual map provided ▪ Unclear if Institutional Review Board approval or informed consent from patients for videotaping was obtained ▪ The authors' state inter-observer variability was investigated with 2 intensivists annotating the data on the basis of video but they do not specify why other RNs were not involved in the annotation ▪ With the use of video annotations, potential opportunity for missed clarification by staff ▪ Potential for selection bias as the 38 patients in the study, only 4 patients fell in the cardiac disease category which may limit external validity. Subjects included patients with gastroenterological disease (n=11), respiratory disease (n=10) and others diseases such as stroke and sepsis, (n=13)

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Solsona et al., 2001 Spain	▪ 12 bed Medical-Surgical ICU in a teaching hospital ▪ 100 patients	▪ Medical-Surgical ICU ▪ Barcelona, Spain ▪ Convenience sample ▪ Make of monitors not specified	▪ Quasi-experimental study ▪ To determine whether documenting the alarm parameters in the patient's record is an effective intervention for improving alarm adjustments	▪ A total of 60 and 40 patients were included in Phase 1 and Phase 3, respectively. ▪ No statistically significant differences between the degree of severity of the cases (SAPS II) included in the study ▪ In phase 1, certain alarm values were set a long way off from the real values of the patient ▪ When upper and lower values of alarms in the 1st and 3rd phase were compared, there were statistically significant improvements in expired volume, hear rate, and systolic arterial pressure, ▪ Having to note alarm limit values on the patients' chart changed the way parameters were employed and resulted in their improved use ▪ Alarm limits were set closer to the corresponding real values but were adjusted more frequently ▪ No differences were found between the values recorded in the patients' medical records and real values	▪ Early and novel study investigating nursing practice and effect of documentation on setting alarm parameter settings	▪ Single center study with convenience sample which limits the generalizability of findings to other ICU settings ▪ Unclear which model of physiologic monitors was utilized ▪ No mention of conceptual or operational definitions ▪ No theoretical framework is specified ▪ Inclusion criteria were limited to patients with stable hemodynamic and respiratory parameters which limits external validity ▪ Limited the study of alarm parameters (high/low) to: respiratory rate, expired volume per minute, airway pressure, oxygen saturation, arterial blood pressure and heart rate ▪ Potential for observer bias as 5 RNs participated and recorded the data on each shift ▪ To prevent response bias, the physicians and RNs were not informed of the reason for recording alarm settings ▪ Did not address if the adjustment of alarm limit values closer to the patient's real values produced more nuisance alarms

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Taenzer et al., 2010 USA	▪ 36-bed Ortho unit ▪ Average of 200 patient days and 53 patient discharges per week	▪ Dartmouth Hitchcock Medical Center ▪ Lebanon, NH ▪ Convenience sample ▪ Masimo™ Patient Safety Net Monitoring System	▪ Quasi-experimental study ▪ A before and after concurrence study ▪ To determine the impact of routine patient monitoring post-operatively on patient outcomes	▪ Alarms average 4 /patient/day or 2 /12 hr nursing shift ▪ Observed deaths after implementation were 2 as opposed to 4 in the previous time frame ▪ Rescue events decreased from 3.4 to 1.2/1000 patient discharges after implementation while the comparison units changed from 2 to 1.3 and 2.7 to 3.4/1000 patient discharges ▪ Transfers to ICU declined from 5.6/1000 patient days to 2.9 whereas the comparison units changed from 5.7 to 5.2 and 15 to 12.7 /1000 patient days	▪ Well-conceived study ▪ 1st study to monitor all non-ICU patients during hospitalization ▪ IRB approval obtained ▪ Data collected 11 month before and 10 months after implementation ▪ Robust patient outcome data on rescue events, ICU patient transfers and mortality rates ▪ 2 comparison units ▪ Reported on return on investment on monitoring system in terms of annual savings ▪ Reports alarms in terms of alarm densities/ shift	▪ No operational definitions or theoretical framework provided ▪ Alarm densities pre implementation of Masimo ™ PatientSafetyNet System is not disclosed ▪ Unclear what physiologic monitoring system was in place prior to the implementation of the Masimo ™ PatientSafetyNet System ▪ It is unknown if all patients on the unit were monitored prior to the implementation of the new monitoring system ▪ Unclear if patient populations on comparison units are similar (Urology/Gyn/Vascular surgery and general surgery) ▪ No reference to patient acuity provided ▪ No discussion on technical alarms ▪ Does not appear to have RN participation

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Talley et al., 2011 USA	▪ 24-bed Level 1 Pediatric Intensive Care Unit ▪ ADC: 20 patients ▪ 2245 monitor alarms (3 star only)	▪ Pediatric Intensive Care Unit (PICU) ▪ Children's National Medical Center ▪ Washington DC, USA ▪ Convenience sample ▪ Philips MP70™ and Philips Patient Information Center ™	▪ Descriptive observational study ▪ To compare monitor alarms to clinically significant events and to estimate sensitivity and specificity of alarms	▪ Cardiopulmonary monitor (CPM) and clinically significant events (CSE) were evaluated in 98 critically ill children ▪ During the observation period, 2,245 three-star alarms (cardiac arrhythmia, apnea or oxygen desaturation) were recorded with 68 clinically significant events ▪ Respiratory related clinically significant events (hypoxia and apnea) comprised the majority alarm events	▪ The literature review was brief and assisted the reader by laying a foundation for the study ▪ Authors cite the study was deemed exempt from Institutional Review Board approval ▪ Feasible and realistic study with team approach ▪ Researchers acknowledged the study's significant limitations and they were unaware that all the alarms were not saved ▪ The authors state this oversight has improved their awareness of their equipment capabilities and settings	▪ Single site and unit ▪ Poor study: unable to answer research questions ▪ No theoretical foundation or definition of concepts ▪ The authors exclusion criteria may have left out valuable patient and alarm information (patients pending organ donation, admitted for < 12hr, or length of stay < 24 hr.) ▪ As a result of the inability to capture all relevant monitor alarms (methodological flaw), the researchers were unable to study the relationship between monitor alarms and clinically significant events. The alarm file only recorded three-star alarm data and did not record the one-star (equipment alerts) and two-star alarms (vital sign that exceed high/low parameter settings) secondary to set-up ▪ The data stored did not definitively identify all alarms that occurred with each study patient and missing data points related to patient whose medical record numbers was not recorded ▪ Study failed to determine the relationship between monitor alarms and AE due to the many methodological flaws inherent in the study design

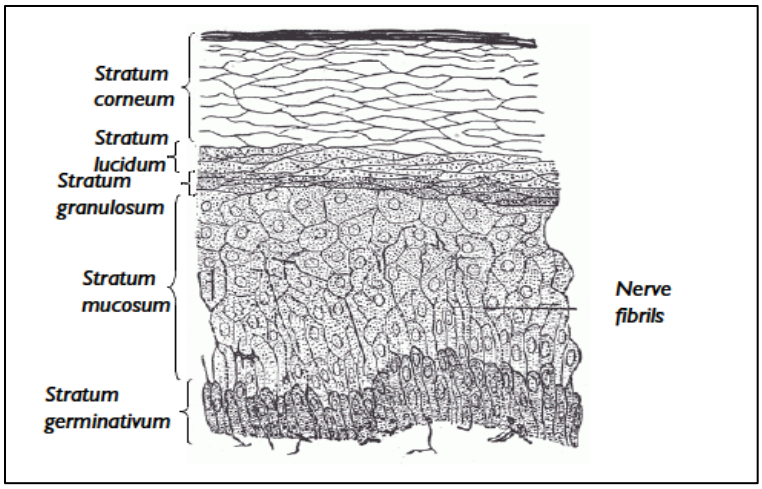
Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Tsien et al.,1997 USA	▪ Pediatric Medical Intensive Care Unit (ICU) ▪ 2942 monitor alarms ▪ 298 hr of over a 10-wk period	▪ Multidisciplinary ICU in a university affiliated children's hospital ▪ MA, USA ▪ Convenience sample ▪ SpaceLabs ™	▪ Descriptive observational study ▪ To identify the areas requiring the most improvement in the ICU and to accurately determine the positive predictive value of routine critical care patient monitoring alarms, as well as common causes for false positive alarms	▪ A total of 2942 total alarms were recorded during 298 hr of monitoring ▪ 86% of alarms were found to be FP alarms, 6% were classified as TP-I, 8% were classified as tTP-R ▪ All monitored signals had FP alarm rates exceeding 90%; the two exceptions were the RR signal (75%) and arterial mean BP signal (46%) ▪ The arterial mean BP signal had the highest rate (38%) of clinically significant TP alarms while remaining signals had clinically significant true positive rates ≤5% ▪ The largest contributor to the total alarms was the SpO_2 (43%) ▪ The oximeter caused FP alarms most frequently (45%) ▪ 18% of alarms were associated with patient interventions while 74% occurred with no interventions	▪ Concepts and operational definitions well defined for TP-R, TP-I and FP " were provided ▪ Alarms annotated by trained observers at the bedside ▪ Accommodations made since not all devices were tracked for an equal number of hours, alarm occurrence were normalized to frequency of alarms/100 hr. of device monitoring ▪ Helpful tables with occurrences of different alarm types and positive predictive values of alarms	▪ Authors did not cite the average daily census, number or diagnoses patient studied ▪ No details on type of SpaceLab™ monitor utilized in the medical PICU ▪ Methodology cumbersome and introduces bias ▪ Although validated by a bedside RN, a single trained observer situated at the bedside recorded all alarm occurrences for devices being tracked within a single bed space at a time in the ICU ▪ All monitoring was performed on weekdays from 09:00 to 17: 30 – missing PM shift and weekend alarm data ▪ The presence of an observer may have produced different behaviors from staff members especially on the night shift and for non-sedated patients ▪ Extracted values every 5 to 6 s and alarm settings entered manually into programs as part of the annotation

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Varpio et al., 2012 Canada	▪14 RNs ▪49 hours of non-participant observa-tion ▪2446 observed alarms	▪ Convenience sample of pediatric acute care RNs ▪ 31-bed medical surgical acute care unit ▪ Tertiary care pediatric teaching hospital ▪ Patient monitoring System, Dråger Medical™	▪ Qualitative study ▪ To explore the qualitative impact of monitor generated alarms on nurses tasked with responding to alarms in an inpatient pediatric unit ▪ To explore strategies nursing staff develop to handle alarm-driven interruptions and the effects those alarms have on nursing work practices	▪ Interpreting alarms and using workaround strategies is a learned skill ▪ Interruptions generated by patient monitors can be viewed as nurse "check-ins" where RNs review specific physiologic information then decide son required care actions or non-actions ▪ Observations revealed that RNs respond to alarms in 7 possible ways ▪ Of 446 alarms, RNs responded 3.8% of the time to the alarms by going to the patients' room	▪ First qualitative study on alarm fatigue among hospital nurses ▪ Feasible study ▪ Obtained rich descriptions on RN workflow and responses to alarms ▪ 2 cycles of individual semi-structured interviews ▪ Confirmability and dependability is reported ▪ Two forms of triangulation was utilized to support trustworthiness	▪ Single inpatient pediatric acute care unit ▪ Direct observations conducted only the day shift RNs (night shift excluded) – selection bias ▪ No clearly and broadly accepted definition of what constitutes "immediate response" – investigators determine 60 s was acceptable ▪ Research assistant was not a licensed health care professional which may have limited the contextual data collected ▪ Limited data collected on many variables of interest including sources of cardiac monitor alarms (e.g. HR, RR, Spo2)
Welch, 2011 USA	10 million SpO_2 data points	▪ 10 hospitals ▪ General post-surgical care units ▪ Convenience sample Masimo Patient SafetyNet™ Monitoring System	▪ Descriptive observational study ▪ Retrospective off-line analysis ▪ To determine the incidence of alarms at various alarm threshold and delay settings	▪ Decreasing alarm thresholds from 90% to 88% decreased alarms by 45%; decreasing alarm thresholds from 90% to 85% decreased alarms by 75% ▪ Increasing alarm delays from 5 to 10 s can decreased alarms by 57%; increasing alarm delays from 5 to 15 s can decreased alarms by 70% ▪ Combining a low SpO_2 alarm threshold of 88% and a 10 s alarm delay reduced alarm frequency by 78%; combining a low SpO_2 alarm threshold of 88% and a 15 s alarm delay reduced alarm frequency by 85%; combining a low SpO_2 alarm threshold of 85% and a 10 s alarm delay reduced alarm frequency by 87%; combining a low SpO_2 alarm threshold of 85% and a 15 s delay reduced alarm frequency by 90%	▪ Large sample of SpO_2 data points ▪ Pulse oximetry data from multiple sites ▪ Acute care study ▪ Theoretical effects of various low SpO_2 alarm parameter thresholds alarms and alarm delays provided ▪ Performed analysis of combined low alarm thresholds and alarm delays on frequency of alarms ▪ Clear and simple figures and tables	▪ No patient outcome data provided or predicted given proposed modification in parameter thresholds and alarm delays ▪ No data on pulse oximetry technical alarms provided ▪ No mention of theoretical foundation or definition of concepts ▪ No inclusion or exclusion criteria provided ▪ No mention that IRB approval was obtained from 10 participating institutions ▪ No nursing participation in the study

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Whalen et al, 2013 USA	▪ 87,823 alarms/ week (pre-pilot) ▪ 9,967/ week alarms (pilot)	▪ Medical Cardiology Acute Care Unit ▪ Convenience sample	▪ Quality improvement study to determine variables that would safely reduce alarms on a general medical-surgical unit where standard manufacturer defaults were used. ▪ Intervention included modifying the accelerated ventricular alarm, tachycardia, and bradycardia alarm to crisis (from warning), increasing the atrial fibrillation alarm severity level to advisory (from message), and reducing PVC alarms to message (from advisory). HR thresholds limits modified to a <45, >130 and crisis level (from <50, >120, warning) ▪ Direct observation to observe staff's response to alarms	▪ A 89% reduction in total mean weekly audible alarms was achieve (t=8.84; P <.0001) ▪ Staff satisfaction improved ▪ Rapid response team (RRT) activation and the incidence of code-blue on the pilot unit in the 6 months preceding and after implementation of the alarm changes demonstrated the number of RRT activations remained constant at 11, whereas the incidence of code blues decreased by 50%, from 6 to 3.	▪ 6 week study ▪ Unit served as its own control unit ▪ Inter-professional engagement ▪ Electronic cardiac monitoring order sets were re-designed ▪ Utilized a 2 sample t-test ▪ Evaluated patient outcome variables (response team activation, code blue)	▪ Methodology for collecting alarm and capturing alarm audibility data is not specified ▪ Observer bias ▪ Focused solely on HR limits, bradycardia, and tachycardia alarm ▪ No annotation of tachycardia or bradycardia alarms to determine FP or TP alarm rates ▪ Technical alarm specific to telemetry – and included 4 alarms ▪ Authors fail to report validity and reliability of survey tool utilized to measure nursing staff satisfaction (Used a home-grown tool). ▪ Noise levels reduction observed can be related to other extraneous variables – unclear how the authors controlled for other sources of noise ▪ Alarm settings pre-pilot were conservation (manufacturer defaults) therefore it is not surprising that an alarm reduction was observed after modification of the default alarm settings

Authors Year/Country	N	Sample Setting	Design Study Aims	Major Findings	Strengths	Weaknesses
Wiklund et al., 1994 Sweden	▪ 123 adult and pediatric patients	▪ Post-Anesthesia Care Unit (PACU) ▪ Uppsala University Hospital ▪ Uppsala, Sweden ▪ Convenience sample ▪ Cardiocap CH IIS™, Datex Instrument-arium	▪ Descriptive observational study ▪ To determine the frequency of true and false alarms and to determine the frequency of alarm failures for various parameters ▪ To determine the occurrence of events in which the equipment failed to alarm	▪ During a mean observation period of 101 min per patient, the average frequency of Spo_2 alarms was once every 8 min ▪ 77% of the alarms were false caused by sensor displacement, motion artifacts, poor perfusion, or a combination of factors ▪ Apnea alarms occurred on average once every 37 min with the 'false" fraction being 28% and 27% for impedance and flow detection, respectively. ▪ The impedance sensor failed to detect apnea on at least 6 occasions; the flow sensor failed on 1 occasion. ▪ The coincidence of Spo_2 and apnea alarms was small and ECG exhibited a low alarm rate but a high fraction of false alarms ▪ Patients receiving opioids /neuromuscular relaxants had a higher frequency of "true" apneas	▪ A standard form was used to record patient and equipment data, alarm categories in abbreviated form, time of alarm, and measurements taken ▪ Dedicated observer relieved of normal duties to observe ≤ 3 patients at once ▪ Occurred over 4 months from January to April 1992 ▪ Default alarm parameters provided (bradycardia <40 bpm tachycardia >160 bpm; Spo_2 <90%; apnea > than 30 s)	▪ Adult/pediatric patients ▪ Very limited inclusion criteria ▪ Observer bias ▪ Only those parameters being continuously monitored (ECG, RR & Spo_2) were included thereby excluding NIBP ▪ In 8 of the 10 monitors, RR was derived from transthoracic impedance measurements between the ECG whereas in 2 units, RR was obtained from a gas flow sensor developed by authors – potential for measurement error ▪ New equipment was installed 6 months before the study which may have contributed to the frequency of alarms Risk for personal bias as authors are originators of one of the sensors being tested
Zong et al., 2004 USA	▪ Stored patient records from Physionet MIMIC II database ▪ 46 patient records	▪ Secondary analysis of a subset of stored patient records from the Physionet MIMIC II database ▪ Convenience sample of patient records	▪ Descriptive observational study ▪ To investigate an algorithm for reducing false ABP alarms by assessing the signal quality of the ABP waveform and by fusing information from simultaneous ECG and ABP signals ▪ simultaneous ECG and ABP signals	▪ The monitors produced 445 ABP alarms of which 126 (28.3%) were false based on visual review ▪ The algorithm rejected 117 of 126 FAs (93%) and only 2 true alarm ▪ The algorithm reduced the false alarm rate in the development data set from 28% to 3% at a cost of rejecting .4% of 319 true alarms ▪ Experiment repeated using the test data set and the monitor produced 604 ABP alarms of which 162 (27%) were false based on review ▪ The algorithm rejected 159 (98%) of these false alarms, reducing the false alarm rate from 28% to .4% while rejecting only 1 (.2%) of 442 true alarms	▪ Homogenous sample ▪ Detailed information provided on instruments, techniques, and complex methodology	▪ Weak description of introduction, background and clinical significance ▪ Small sample size and secondary analysis from longstanding database ▪ Lacks "real-world" clinical perspective ▪ Non critical arrhythmia alarms were excluded ▪ No alarms related to RR, Spo_2 or other cardiac monitor alarms included given study aims ▪ Researchers did not acknowledge any limitations

Figure 2.1. Epidermis Cross Section
(Improving ECG Quality, by Philips Healthcare, 2008.
Adapted with permission)



Abbreviations

ABP	Arterial blood pressure
ADC	Average daily census
ADL	Activities of daily living
AE	Adverse event
Bpm	Beats per minute
BP	Blood pressure
CEM	Continuous electrocardiographic monitoring
CI	Cardiac index
CPM	Cardiopulmonary monitor
CSE	Clinically significant events
CVP	Central venous pressure
CPA	Cardiopulmonary arrest
dB	Decibels
ECG	Electrocardiography
ED	Emergency department
EEC	Extracorporeal circulation
eICU	Electronic intensive care unit
EtCO2	End tidal carbon monoxide
FA	False alarm
FN	False negative
FP	False positive
Hr	Hour
HR	Heart rate
IA	Intelligent agent
ICU	Intensive care unit
IRB	Institutional review board
L_{Aeq}	Weighted average sound level
L_{AFMax}	Maximum a-weighted noise level –fast time weighting
L_{CPeak}	C-weighted peak measurement
MAP	Mean arterial pressure
MD	Medical doctor
Min	Minutes
mmHg	Millimeter of mercury
NPV	Negative predictive value
P_{aw}	Airway pressure
PACU	Post anesthesia care unit
PAPd	Pulmonary arterial pressure, diastolic
PAPm	Pulmonary arterial pressure, mean
PAPs	Pulmonary arterial pressure, systolic
PPV	Positive predictive value
RN	Registered nurse
RR	Respiratory rate
RRT	Rapid response team
RT	Respiratory therapist
RTT	Respiratory therapist technicians
s	Seconds
S	Standard deviation
SAPd	Systemic arterial pressure, diastolic
SAPm	Systemic arterial pressure, mean
SAPs	Systemic arterial pressure, systolic
SAPS II	Simplified Acute Physiological Score
SBP	Systolic blood pressure
SpO_2	Pulse oximeter oxygen saturation
TA	True alarm
TP-I	True positive, irrelevant
TP-R	True positive, relevant
V-Tach	Ventricular tachycardia
V-Tach/V-Fib	Ventricular tachycardia/ ventricular fibrillation

References

Aboukhalil, A., Nielsen L., Saeed, M., Mark, R. G., & Clifford, G. D. (2008). Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform. *Journal of Biomedical Informatics, 41,* 442–451.

Adams-Hamoda, M. G., Caldwell, M. A., Stotts, N. A., & Drew, B. J., (2003). Factors to consider when analyzing 12-lead electrocardiograms for evidence of acute myocardial ischemia. *American Journal of Critical Care, 12*(19), 9–16.

Association for the Advancement of Medical Instrumentation (2011). *Clinical Alarms; 2011 Summit*. Retrieved from http://www.aami.org/htsi/alarms /pdfs/2011 _Alarms_Summit_publication.pdf

Association for the Advancement of Medical Instrumentation (2012). *Safety innovations: Using data to drive alarm system improvement efforts. The Johns Hopkins Hospital experience.* Retrieved from http://www.aami.org/htsi/SI_Series/ Johns_Hopkins_White_Paper.pdf

Biot, L., Carry, P. Y., Perdrix, J. P., Eberhard, A., & Baconnier, P. (2000). Évaluation clinique de la pertinence des alarmes en reanimation. *Annales Françaises d'Anesthésie et de Réanimation, 19,* 459–466.

Bitan, Y., Meyer J., Shinar, D., & Zmora, E. (2004). Nurses' reactions to alarms in a neonatal intensive care unit. *Cognition, Technology & Work, 6,* 239–246.

Block, F. E., Jr., Nuutinen, L., & Ballast, B. (1999). Optimization of alarms: A study of alarm limits, alarm sounds, and false alarms, intended to reduce annoyance. *Journal of Clinical Monitoring and Computing, 15,* 75–83.

Blum, J. M., Kruger, G. H., Sanders, K. L., Gutierrez, J., & Rosenberg, A. L. (2009). Specificity improvement for network distributed physiologic alarms based on a simple deterministic

reactive intelligent agent in the critical care environment. *Journal of Clinical Monitoring and Computing, 23,* 21–30.

Borowski, M., Görges, M., Fried, F., Such, O., Wrede, C., & Imhoff, M. (2011). Medical device alarms. *Biomedical Engineering/ Biomedizinische Technik, 56,* 73–83.

Breznitz, S. (1984). *Cry-wolf: The psychology of false alarms.* Lawrence Erlbaum Associates, Hillsdale, NJ.

Brown, J. C., & Anglin-Regal, P. (2008). Clinical alarm management: A team effort. *Biomedical Instrumentation & Technology, 42*(2), 142–144.

Burgess, L. P. A., Herdman, T. H., Berg, B.W., Feaster, W. W., & Hebsur, S. (2009). Alarm limit settings for early warning systems to identify at-risk patients. *Journal of Advanced Nursing, 65*(9), 1844–1852.

Catalano, K. (2005). Clinical Alarms Management and the JCAHO Standards. *Biomedical Instrumentation & Technology, 39*(6), 4-5.

Centers for Medicare and Medicaid Services (2011). Department of Health & Human Services, *State Medicaid Director Letter.* Retrieved from https://www.cms.gov/SMDL/downloads/SMD073108.pdf

Chambrin, M. -C., Ravaux, P., Calvelo-Aros, D., Jaborska, A., Chopin, C., & Boniface, B. (1999). Multicentric study of monitoring alarms in the adult intensive care unit (ICU): A descriptive analysis. *Intensive Care Medicine, 25,* 1360–1366.

Chi, Y. M., Jung, T-P., & Cauwenberghs, G. (2010). Dry-contact and noncontact biopotential electrodes: Methodological review. *IEEE Reviews in Biomedical Engineering, 3,*106–118.

Clochesy, J. M., Cifani, L., & Howe, K. (1991). Electrode site preparation techniques: A follow-up study. *Heart & Lung, 20*(1), 27–30.

Covidien (2013). *Kendall 530 series foam electrodes: Market leading electrode* Retrieved from http://www.covidien.com/criticalcare/pageBuilder.aspx?topicID =143786&breadcrumbs=0:121623,143512:0,143513:0,157490:0,157494:0

Cropp, A. J., Woods, L. A., Raney, D., & Bredle, D. L. (1994). Name that tone: The proliferation of alarms in the intensive care unit. *Chest, 105*(4), 1217–1220.

Cvach, M. (2012). Monitor alarm fatigue: An integrative review. *Biomedical Instrumentation & Technology, 46*(4), 268–277.

Cvach, M., Biggs, M., Rothwell, K. J., Charles-Hudson, C. (2012). Daily electrode change and effect on cardiac monitor alarms: An evidence based practice approach. *Journal of Nursing Care Quality. 28*(3), 265-271.

Daly, B. J. & Wilson, C. A. (1983). The effect of fatigue on the vigilance of nurses monitoring electrocardiograms. *Heart and Lung, 12(*4), 384–388.

Darcy, A. E., Hancock, L. E., & Ware, E. J. (2008). A descriptive study of noise in the neonatal intensive care unit. Ambient levels and perceptions of contributing factors. *Advances in Neonatal Care, 8*(3), 165–175.

Drew, B. J., Califf, R. M., Funk, M., Kaufman, E. S., Krucoff, M. W., Laks, M. M., . . . Van Hare, G. F. (2004). Practice standards for electrocardiographic monitoring in hospital settings. *Circulation, 110*, 2721–2746.

Edworthy, J. (2011). Designing effective alarm sounds. *Biomedical Instrumentation & Technology, 45*(4), 290–294.

Elliott, R. M., McKinley, S. M., & Eager, D. (2010) A pilot study of sound levels in an Australian adult general intensive care unit. *Noise & Health, 12*(46), 26–36.

Emergency Care Research Institute (2007). The hazards of alarm overload. Keeping excessive physiologic monitoring alarms from impeding care. *Health Devices, 36*(3), 73–83.

Emergency Care Research Institute. (2012a). *Medical device alarm related terminology*. Retrieved from http://www.aami.org/meetings/summits/alarms.html

Emergency Care Research Institute (2012b). *Top 10 technology hazards for 2013*.Retrieved from https://www.ecri.org/Documents/Secure/Health_Devices_Top_10_Hazards_2013.pdf

Emergency Care Research Institute (2013). Top 10 health technology hazards for 2014. *Health Devices, 42*(11), 354–380.

Fidler R., Pickham, D., & Drew, B. J. (2011). Quantification of hospital cardiac monitoring alarms. *Journal of Electrocardiology, 44*, 744.

Funk, M., Winkler, C.G., May, J. L., Stephens, K., Fennie, K. P., Rose, L.L., . . . Drew, B. J. (2010). Unnecessary arrhythmia monitoring and underutilization of ischemia and QT interval monitoring in current clinical practice: baseline results of the Practical Use of the Latest Standards for Electrocardiography (PULSE) Trial. *Journal of Electrocardiology, 43*, 542-547.

General Electric Healthcare (2005). *Critical care monitoring clinical reference and troubleshooting guide.* Waukesha, WI: GE Healthcare.

Görges, M., Markewitz, B. A., & Westenskow, D. R. (2009). Improving alarm performance in the medical intensive care unit using delays and clinical context. *Anesthesia & Analgesia, 108*(5), 1546–1552.

Graham, K., & Cvach, M. (2010). Monitor alarm fatigue: Standardizing use of physiological monitors and decreasing nuisance alarms. *American Journal of Critical Care, 19*(1), 28–34.

Gross, B., Dahl, D., & Nielsen, L., (2011). Physiologic monitoring alarm load on medical/surgical floors of a community hospital. *Biomedical Instrumentation & Technology, 45*(1), 29–36.

Hanish, H. M., Neustein, R. A., Van Cott, C. C., & Sanders, R. T. (1971). Technical aspects of monitoring the heart rate of active persons. *The American Journal of Clinical Nutrition, 24*(9), 1155–1163.

Healthcare Technology Foundation (2011). *National Clinical Alarm Survey: Perceptions, issues, improvements, and priorities of healthcare professionals*. Retrieved from http://www.thehtf.org

Hu, X., Sapo, M., Nenov, V., Barry, T., Kim, S., Do, D. H., . . . Martin, N. (2012). Predictive combinations of monitor alarms preceding in-hospital code blue events. *Journal of Biomedical Informatics, 45*(1), 913–921.

Imhoff, M., & Kuhl, S. (2006). Alarm algorithms in critical care monitoring. *Anesthesia & Analgesia, 102,* 1525–1537.

International Electrotechnical Commission (2006). Medical electrical equipment- Part 1-8: General requirements for basic safety and essential performance—Collateral Standard: *General requirements, tests and guidance for alarm systems in medical electrical equipment and medical electrical systems* IEC 60606-1-8. (2nd ed.). Geneva: IEC.

Johansson, L., Bergbom, I., Persson-Waye, K., Ryherd, E., & Lindahl, B. (2012). The sound environment in an ICU patient room: A content analysis of sound levels and patient experiences. *Intensive and Critical Care Nursing, 28*(5), 269–279.

Konkani, A., Oakley, B., & Bauld, T. J. (2012). Reducing hospital noise: A review of medical device alarm management. *Biomedical Instrumentation & Technology*, *46*(6), 478–487.

Korniewicz, D. M., Clark, T., & Yadin, D. (2008). A national online survey on the effectiveness of clinical alarms. *American Journal of Critical Care, 17*(1), 36–41.

Koski, E., M., Mäkivirta, A., Sukuvaara, T., & Kari, A. (1990). Frequency and reliability of alarms in the monitoring of cardiac postoperative patients. *International Journal of Clinical Monitoring and Computing, 7*(2), 129–133.

Koski, E. M., Mäkivirta, A., Sukuvaara, T., & Kari, A. (1995). Clinicians' opinions on alarm limits and urgency of therapeutic responses. *International Journal of Clinical Monitoring and Computing*, *12*, 85–88.

Kowalczyk, L. (2010, April 3). "Alarm fatigue" linked to patient's death. *The Boston Globe.* Retrieved from http://www.boston.com/news/health/articles /2010/04/03/alarm_fatigue_linked_to_heart_patients_death_at_mass_general/

Kowalczyk, L. (2011a, February 13). Patient alarms often unheard, unheeded. *The Boston Globe*. Retrieved from http://www.boston.com/lifestyle/health/articles /2011/02/13/patient_alarms_often _unheard _unheeded/

Kowalczyk, L. (2011b, February 14). No easy solution for alarm fatigue. *The Boston Globe*. Retrieved from http://www.boston.com/news/local/massachusetts /articles/2011/02/14/no_easy_solutions_for_alarm_fatigue/

Kowalczyk, L. (2011c, April 18). Groups target alarm fatigue at hospitals. *The Boston Globe*. Retrieved from http://www.boston.com/lifestyle/health/articles /2011/04/18/groups_target_alarm_fatigue_at_hospitals/

Lawless, S.T. (1994). Crying wolf: False alarms in a pediatric intensive care unit. *Critical Care Medicine, 22,* 981–985.

Mäkivirta, A., & Koski, E. M. J. (1994). Alarm-inducing variability in cardiac postoperative data and the effects of prealarm delay. *Journal of Clinical Monitoring, 10*(3), 153–162.

Mäkivirta, A., Koski, E. M. J., Kari, A., & Sukuvaara, T. (1991). The median filter as a preprocessor for a patient monitor limit alarm system in intensive care. *Computer Methods and Programs in Biomedicine, 34,* 139–144.

McKinney, M. (2010). Alarm fatigue sets off bells: Mass. Incident highlights need for protocols check. *Modern Healthcare, 40*(15), 14.

Medina, V., Clochesy, J. M., & Omery, A. (1989). Comparison of electrode site preparation techniques. *Heart & Lung, 18*(5), 456–460.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Momtahan, K. L., Tanslet, B. W., & Hetu, R. (1993). Audibility and identification of auditory alarms in the operating room and intensive care unit. *Ergonomics, 36,* 1159–1176.

Monsén, M. G., & Edéll-Gustafsson. U. M. (2005). Noise and sleep disturbance factors before and after implementation of a behavioural modification programme. *Intensive and Critical Care Nursing, 21,* 208–219.

National Quality Forum (2011). *Serious reportable events in healthcare—2011 update: A consensus report.* Retrieved from http://www.qualityforum.org/Topics/SREs /Serious_Reportable_Events.aspx

O'Carroll, T. M. (1986). Survey of alarms in an intensive therapy unit. *Anaesthesia, 41*, 742–744.

Occupational Safety and Health Administration (2012). *Occupational Health and Environmental Control, Standard 1910.95. Occupational Noise Exposure*. Retrieved from http://www.osha.gov/pls/oshaweb/owadisp.show _document?p_table=STANDARDS&p_id=9735

Oster, C. D. (2000). Improving ECG trace quality. *Biomedical Instrumentation & Technology, 34*(3), 219–222.

Pan, P., & Gravenstein, N. (1994). Intraoperative Pulse Oximetry: Frequency and distribution of discrepant data. *Journal of Clinical Anesthesia, 6*, 491–495.

Patel, S. I., & Souter, M. J. (2008). Equipment-related electrocardiographic artifacts: casuses, characteristics, consequences, and correction. *Anesthesiology, 108,* 138-148.

Patterson, R. P. (1978). The electrical characteristics of some commercial ECG electrodes. *Journal of Electrocardiology, 11*(1), 23–26.

Patton, J. A., & Funk, M. (2001). Survey of use of ST-segment monitoring in patients with acute coronary syndromes. *American Journal of Critical Care, 10*, 23–34.

Pennsylvania Patient Safety Authority (2008). Alarm interventions during medical telemetry monitoring [Supplemental Material] *Pennsylvania Patient Safety Advisory*, March, 1–50.

Philips, J. (2006). Clinical alarms: Complexity and common sense. *Critical Care Nursing Clinics of North America, 18*(2), 145–156.

Philips Healthcare (2008). *Improving ECG quality: Application note.* Retrieved from http://incenter.medical.philips.com/doclib/enc/fetch/2000/4504/577242/577243/577245/577817/577869/Improving_ECG_Quality.pdf%3fnodeid%3d8616008%26vernum%3d1

Pujols-McKee, A. (2011, October). Alarms Management Summit. In M. Logan (Chair), *Medical Device Alarm Summit*. Symposium conducted at the Association for the Advancement of Medical Instrumentation, Herndon, VA.

Rheineck-Leyssius, A. T., & Kalkman, C. J. (1998). Influence of pulse oximeter settings on the frequency of alarms and detection of hypoxemia. *Journal of Clinical Monitoring and Computing*, *14*(3), 151–156.

Ryherd, E. E., Persson-Waye, K., & Ljungkvist, L. (2008). Characterizing noise and perceived work environment in a neurological intensive care unit. *Journal of the Acoustical Society of America, 123*(2), 747–756.

Sanderson, P. M., Wee, A., & Lacherez, P. (2006). Learnability and discriminability of melodic medical equipment alarms. *Anesthesia, 61*, 142–147.

Sangkachand, P., Cluff, M., & Funk M. (2012). Detecting myocardial ischemia with continuous ST-segment monitoring: Two case studies. *Heart & Lung, 41*, 284–289.

Scalzo, F., & Hu, F. (2013). Semi-supervised detection of intracranial pressure alarms using waveform dynamics. *Physiological Measurement, 34*, 465–478.

Schmid, F., Goepfert, M. S., Kuhnt, D., Eichhorn, V., Diedrichs, S., Reichenspurner, H., Goetz, A. E., & Reuter, D. A. (2011). The wolf is crying in the operating room: Patient monitor and anesthesia workstation alarming patterns during cardiac surgery. *Anesthesia and Analgesia, 12*(1), 78–83.

Siebig, S., Kuhls, S., Imhoff, M., Gather, U., Schölmerich, J., & Wrede, C. (2010a). Intensive care unit alarms- How many do we need? *Critical Care Medicine, 38*(2), 451–456.

Siebig, S., Kuhls, S., Imhoff, M., Langgartner, J., Reng, M., Schölmerich, J., Gather, U., & Wrede, C. (2010b). Collection of annotated data in a clinical validation study of alarm algorithms in intensive care: A methodological framework. *Journal of Critical Care, 25*, 128–135.

Siebig, S., Sieben, W., Kollmann, F., Imhoff, M., Bruennler, T., Rockmann, F., Gather, U., & Wrede, C. (2009). Users' opinions on intensive care unit alarms: A survey of German intensive care units. *Anaesthesia and Intensive Care, 37*(1), 112–116.

Smith, M. (1984). Rx for ECG monitoring artifact. *Critical Care Nurse, 4*(1), 64–66.

Solet, J. M., & Barach, P. R. (2012). Managing alarm fatigue in cardiac care. *Progress in Pediatric Cardiology, 33*, 85–90.

Solsona, J. F., Altaba, C., Maúll, E., Rodríguez, L., Bosqué, C., & Mulero, A. (2001). Are auditory warnings in the intensive care unit properly adjusted? *Journal of Advanced Nursing, 35*(3), 402–406.

Taenzer, A. H., Pyke, J. B., McGrath, S. P., & Blike, G. T. (2010). Impact of pulse oximetry surveillance on rescue events and intensive care unit transfers. *Anesthesiology, 112*(2), 282–287.

Talley, L. B., Hooper, J., Jacobs, B., Guzzetta, C., McCarter, R., Sill, A., Cain, S., & Wilson, S. L. (2011). Cardiopulmonary monitors and clinically significant events in critically ill children. *Biomedical Instrumentation & Technology, 45*(1), 38–45.

Tam, H. W., & Webster, J. G. (1977). Minimizing electrode motion artifact by skin abrasion. *IEEE Transactions on Biomedical Engineering, 24*(2), 134–139.

The Cardiac Arrhythmia Suppression Trial (CAST) Investigators (1989). Preliminary Report: Effect of Encainide and Flecainide on Mortality in a Randomized Trial of Arrhythmia Suppression after Myocardial Infarction. *New England Journal of Medicine, 321,* 406–412.

The Joint Commission (2004). Case Study: Improving the effectiveness of clinical alarm systems. *Joint Commission Perspectives on Patient Safety, 4*(6), 9–10.

The Joint Commission (2014). *2014 Hospital National Patient Safety Goals.* Retrieved from http://www.jointcommission.org/assets/1/6/2014_HAP_NPSG_E.pdf

Tronstad, C., Krogh Johnsen, G., Grimnes, S., & Martinsen, Ø. G. (2010). A study on electrode gels for skin conductance measurements. *Physiological Measurement, 31,* 1395–1410.

Tsien, C. L., & Fackler J. C. (1997). Poor prognosis for existing monitors in the intensive care unit. *Critical Care Medicine, 25* (4), 614–619.

Turkmen, A., & Pantiskas, C. (2011). Management of ECG cables and leadwires. *Biomedical Instrumentation & Technology, 45*(2), 130–133.

Varpio, L., Kuziemsky, C., MacDonald, C. King, J. W. (2012). The helpful or hindering effects of in-hospital patient monitor alarms on nurses. *Computers, Informatics, Nursing, 30*(4), 210–217.

Welch, J. (2011). An evidence-based approach to reduce nuisance alarms and alarm fatigue. *Biomedical Instrumentation & Technology, 45*(1), 46–52.

Weil, K. M. (2009). Alarming monitor problems. *Nursing, 9*, 58.

Whalen, A. A., Covelle, P.M., Piepenbrink. J. C., Villanova, K. L., Cuneo, C. L., & Awtry, E. H. (2013). Novel approach to cardiac alarm management on telemetry units. *Journal of Cardiovascular Nursing*, 1-10.

Wiklund, L., Hök, B., Ståhl, K., & Jordeby-Jönsson, A. (1994). Postanesthesia monitoring revisited: Frequency of true and false alarms from different monitoring devices. *Journal of Clinical Anesthesia*, *6*, 182–188.

Williams, A. L., van Drongelen, W., & Lasky, R. E. (2007). Noise in contemporary neonatal intensive care. *The Journal of the Acoustic Society of America, 121*(5 Pt1), 2681–2690.

World Health Organization (1999). *Guidelines for community noise*. Retrieved from http://www.euro.who.int/en/what-we-do/health-topics/environment-and-health/noise/facts-and-figures

Young, K. D., & Seidel, J. S. (1999). Pediatric cardiopulmonary resuscitation: A collective review. *Annals of Emergency Medicine, 33*(2), 195–205.

Zong, W., Moody, G. B., & Mark, R. G. (2004). Reduction of false arterial blood pressure alarms using signal quality assessment and relationships between electrocardiogram and arterial blood pressure. *Medical and Biological Engineering and Computing, 42*, 698–706.

Chapter 2 Part 2

Theoretical Considerations

Registered nurses protect critically ill patients by practicing intensive observation (sometimes referred to as "watchful vigilance") and by attending to patients according to their acuity or physiologic conditions (Fairman, 1992). Ultimately, this practice and approach to patient care provided the foundation for the development in the 1950s of the first intensive care units (ICUs; Fairman, 1992). Within the next decade, coronary care units (CCUs) were developed, and soon thereafter electrocardiographic (ECG) monitoring was introduced (Daly, 1968). In the 1980s, computers were rapidly introduced in the care environment, along with complex medical devices and increasingly sophisticated monitoring techniques. By the 1990s, clinicians were encountering a burgeoning volume of clinical data from computer analyses and other sources; as one consequence, clinical decision making now required that ICU nurses assimilate information from over 230 variable categories while concurrently making good clinical decisions (East, 1992; East, Wallace, Morris, Gardner, & Westenskow, 1995).

Clinicians have a penchant for state-of-the-art computers and medical devices, and, as a result, applications of technology have proliferated throughout health care settings (Hagenouw, 2007). In a typical ICU setting, monitoring systems and other equipment can supply 40–100 alarm sources; in some ICUs, the number may be even higher. Each of these sources has alarm conditions that must be configured, activated, or manipulated (Block, 2011; Chambrin, 2001). Moreover, the prolific use of patient monitoring systems outside of the critical care environment sponsored by national guidelines, regulations, and organizational commitments has resulted in clinicians' being subjected to an astounding number of medical device alarms (Keller, Diefes, Graham, Meyers, & Pelczarski, 2011).

Recent studies report that, on a variety of patient care units, staff may be exposed to over 800 physiologic monitor alarms every day (Fidler, Pickham, & Drew, 2011; Graham & Cvach, 2010). Moreover, estimates of physiologic monitor alarm burden differ, depending on the clinical setting and on the research methodology utilized to collect the sources and types of alarms. Despite varying approaches, researchers have reported that up to 90% of alarms in all critical care units were false-positive alarms (Imhoff & Kuhls, 2006). Because of the excessive frequency of false-positive alarms, most clinicians mistrust alarms; this excessive frequency has led clinicians to exhibit new alarm response behaviors characterized by reacting more slowly, ignoring alarm notifications, or worse, deactivating alarm signals (Bliss & Dunn, 2000; ECRI, 2007). Not surprisingly, the high frequency of clinical alarms has also resulted in the development of alarm fatigue among health care professionals. This condition undermines efforts to (a) accomplish the objectives of medical monitoring, (b) improve patient safety, and (c) reduce medical errors (Kohn, Corrigan, & Donaldson, 1999).

Health care errors related to clinical alarms are not new. Adverse events related to alarms can be traced back to when medical devices were first introduced into hospitals (Fairman & Kagan, 1999). Alarm-related errors are devastating in terms of the human toll and economic consequence. Of the 26 alarm-related injury claims made between 1970 and 2002, 88% involved permanent brain damage or death, with a median legal award of $450,000 (Olympio, 2004). While these figures may appear negligible, alarm-related errors are undoubtedly under-reported by clinicians and health care facilities. In fact, the actual number of alarm system-related deaths may be tenfold (or more) the number reported (Keller, 2011). In a recent HTF (2011) study, investigators reported that 18% of surveyed participants acknowledged that their hospital had had an adverse patient event related to clinical alarms within the 2-year period preceding the

survey. These results affirm that alarm fatigue is undeniably a health care technology hazard—a potential source of danger based on the possibility of an alarm system-related adverse patient event (ECRI, 2009, 2010, 2011, 2012a, 2013).

The focus of this chapter is to present a theoretical framework for understanding physiologic monitor alarm fatigue and its effects on clinician behavior in the hospital environment. To illuminate this discussion, the well-known *habituation: a dual-process theory of response plasticity* will be reviewed. In particular, an evaluation will be made of the tenets of this theoretical perspective as it pertains to alarm signals associated with physiologic monitors. Moreover, it is important to note that the habituation: a dual-process theory of response plasticity will be used in a novel way, given that it has not been previously adapted or applied to nursing research related to alarm fatigue.

The ECRI—a leading nonprofit organization that researches and identifies the best approaches to improve the safety, quality, and cost-effectiveness of patient care—provides the most satisfactory conceptual definition of alarm fatigue. ECRI describes alarm fatigue as a condition that occurs when staff is exposed to an excessive number of clinical alarms. Alarm fatigue, which results from sensory overload, may cause clinicians to become desensitized to medical device alarms; this desensitization in turn may result in delayed alarm response or missed alarms (ECRI, 2012b). Other professional societies have proposed conceptual definitions; however, most descriptions fail to provide an adequate explanation and tend to focus on patient and staff outcomes. Most investigators use the word “desensitized” to describe the change in clinicians’ behavior in response to redundant or needless alarms. The word *desensitize* is defined as “to make (a sensitized or hypersensitive individual) insensitive or nonreactive to a sensitizing agent; . . . to extinguish an emotional response” (“desensitize,” *Merriam-Webster.com*,

2012a). In contrast, *habituate* is defined as "to become used to something; . . . to undergo habituation, a decrease in responsiveness upon repeated exposure to a stimulus" ("habituate," *Merriam-Webster.com*, 2012b). Over time, the novice nurse, like the expert nurse, can habituate to repeated alarms as she or he gains experience and develops reasoning and clinical judgment. Because of the high frequency and repetitive nature of alarms, the term *habituate* better reflects the basis for clinicians' behavioral response, or lack thereof, to physiologic monitor alarms.

A comprehensive review of the literature reveals that extant research regarding alarm fatigue and staffs' behavior toward clinical alarms is predominately atheoretical. That is, there are few referenced theoretical frameworks or models to guide alarm fatigue research. Theories utilized in domains such as health information technology, the aviation industry, and nuclear power plant research are derived from a variety of disciplines (such as psychology, sociology, and ergonomics) and can lend themselves to this program of research. The application of certain theoretical foundations provides guidance for understanding alarms and subsequent behavioral responses; yet, such theories are not readily adaptable to health care researchers (Holden & Karsh, 2009). Nevertheless, the field of vigilance research and related theories can provide a theoretical underpinning to better appreciate physiologic monitor alarm fatigue.

Vigilance Research

In the early 20th century, Sir Henry Head first employed the word "vigilance" to describe a state of maximum physiological and psychological readiness to react (Warms, Matthews, & Finomore, 2008). Years later, this definition was further refined, and today *vigilance* (also referred to as "sustained attention") is described as the ability of observers to maintain their focus of attention and to remain responsive to stimuli over prolonged periods of time (Ballard, 1996; Davies & Parasuraman, 1982; Warm, 1977, 1984, 1993). The foundation of vigilance

research lies in world history and provides valuable insight regarding the nature of human performance in high-stress occupational contexts (such as industrial inspection, military target detection, commercial transport aviation, and medical monitoring).

Classic vigilance research was first conducted during World War II. The concept of vigilance stemmed from the seminal work of Norman Mackworth, a neurologist and cognitive scientist. In 1943, Mackworth was commissioned by the British Royal Air Force to investigate factors affecting vigilance behaviors—specifically, factors that contributed to airborne observers' failure to identify radar signals (Warm, 1993). To study vigilance, Mackworth developed the "Clock Test," which simulated the essential functions of the radar operator's job. The Clock Test was used in an experiment that evaluated the operator's ability to detect critical signals for prolonged continuous periods of time. Mackworth's formative research found that, in individual subjects, initial levels of signal detection were high, but over the monitoring period, radar operators' performance markedly declined. That is, the quality of sustained attention was found to be fragile and diminish quickly over time (Helton & Warm, 2008; Warm, Matthews, & Finomore, 2008). The progressive decline in sustained attention—known as *vigilance decrement*—is the most common finding in vigilance research (Davies & Parasuraman, 1982; Matthews, 2000; Warm, 1984).

Mackworth's pioneering work on the psychology of vigilance was later advanced by Jerison (1959) and modified by Warm and Berch (1985) who described the components of vigilance. Researchers now posit that performance efficiency is a function with five dimensions: sense modality, signal salience, stimulus uncertainty, the background event context, and stimulus complexity. Three of these dimensions—sense modality, signal salience, and stimulus complexity—are important in the study of clinical alarms and are reviewed later in this paper. Of

lesser influence, the background event context refers to critical signals that may be available for recognition as they are present within a matrix of repeated background events (Warm, 1993). For instance, individuals may be asked to identify a distinct audible tone amidst lower auditory sounds emanating from the same display. While this dimension is critical in certain lines of work, such as target detection, it is of minor significance in relation to medical monitoring. Similarly, according to Warm (1993), stimulus uncertainty is focused on challenges in knowing when signals appear (temporal) or difficulties in identifying where signals appear (spatial). Therefore, this dimension is less applicable to study of physiological monitor alarm fatigue in health care.

Performance efficiency, a measure of vigilance by the consideration of the above dimensions, can be objectively measured. Procedural steps include (a) examining an individual's ability to correctly identify signals (referred to as "detection rate"), (b) assessing the time one takes to detect and respond to a signal (referred to as "detection latency"), and (c) evaluating the number of occasions for which a signal is reported when in fact there is none (i.e., a "false alarm"; Davies & Parasuraman, 1982; Matthews, 2000). Together, these indicators can contribute to the assessment of an individual's performance in medical monitoring and consequently, enhance the understanding of alarm burden and fatigue.

With regard to vigilance, the parallels between the task of radar operators (who use monitors to observe aviation events) and the task of RNs (who use monitors to observe patients' physiological status) are striking. Both job tasks require sustained attention, sound judgment, and a quick response from observers to ensure the safety of individuals whose lives depend on the vigilance of others. Understanding elements that contribute to vigilance decrement among clinicians—that is, elements that are associated with medical monitoring—is instrumental in

addressing alarm fatigue, a modern-day threat to patient safety. Several researchers in psychology have incorporated factors known to affect human performance into their theoretical constructs. Such theories can be useful for identifying interventions to improve alarm conditions and optimize behavioral responses—and through these interventions, save patient lives.

Theoretical Framework

The performance of medical monitoring in the hospital environment can be viewed from various perspectives. One theoretical perspective useful in conducting research related to clinicians' use and responsiveness to physiologic monitor alarms is reviewed in this chapter. Notably, this theory was developed decades ago when physiological monitors were first being manufactured and adopted in CCUs and ICUs. The application of a behavioral theory seems most appropriate to the study of physiologic monitor alarm burden and resultant clinician response.

Habituation: A Dual-Process Theory

The Groves and Thompson (1970) habituation: a dual-process theory of response plasticity to repeated stimulation model can be adapted to the study of alarm fatigue related to physiologic monitor alarm burden. The theory is based on neurophysiological studies of habituation and sensitization, yet it is considered a behavioral theory that explains how changes in behavior can occur as a result of experience (Domjan, 1996). Habituation and sensitization *processes* are neural mechanisms that are responsible for an individual's behavior towards a stimulus. The habituation process is described as the weakening of a response to an incitement; this weakening occurs when the stimulus is repeatedly provoked (Domjan, 1996). Habituation has been most studied in relation to a recurrent auditory stimulus (Watts, 1979). In contrast, the sensitization process is described as an increase in responsivity to a stimulus. In the context of

response magnitude, habituation is decremental, and sensitization is incremental. Moreover, in the dual-process theory of response plasticity to repeated stimulation, these two processes are believed to develop independently in the central nervous system. Dual process theory also holds that habituation and sensitization processes interact on the level of neurophysiological function in producing behavioral outcomes (Groves & Thompson, 1970).

Domjan (1996) emphasizes that habituation and sensitization processes are underlying neural mechanisms and that the "effects" of these processes are observable human behaviors. Groves and Thompson (1970) postulate that repeated exposure to a stimulus can have two independent behavioral responses: a habituation effect and a sensitization effect. The *habituation effect* is characterized as a decrease in the vigor of the initial behavioral response that is elicited by the stimulus, and the *sensitization effect* is an enhancement of the initial behavioral response (Domjan, 1996). Habituation and sensitization can occur concurrently, and the resultant change in behavior is most influenced by whichever effect is most influential (Thompson, 2009). Habituation and sensitization effects are a component of conditioning and learning. Operating in mutual opposition, these processes regulate an individual's response to environmental factors. An illustration of the different behavioral responses to repeated stimulations is provided in Figure 2.2.

The processes of habituation and sensitization play a vital role in influencing living organisms' responses to environmental stimuli—that is, responses to events that may be either significant and warranting response or inconsequential and ignorable (Domjan, 1996). If the habituation effect is dominant, the behavioral response decreases. Meanwhile, if the sensitization effect is dominant, the behavioral response increases. Thorpe (1956) described habituation as the simplest form of learning. Notably, habituation is distinguished from other learning practices in

that the decrease in response strength is effortless and is itself unlearned (Thompson & Spencer, 1966). In essence, habituation is a precursor process for more complex forms of learning as it permits the "filtering out" of irrelevant stimuli and facilitates selective focus on important stimuli—a prerequisite for learning (Rankin et al., 2009).

A variety of factors influence the habituation and sensitization effects and are noticeably similar to the dimensions of vigilance developed by Mackworth (1948) and later revised by Warm and Berch (1985). Habituation (and not sensitization) will be the focus of discussion in this paper. In particular, factors contributing to the habituation effect will be described in the context of physiologic monitor alarms. The following section will concentrate four factors that influence the habituation effect (i.e., *effects of stimulus frequency, intensity, change,* and *time*) as they pertain to habituation to alarms. (The two other factors in habituation—*exposure to a second stimulus* and *time after a dishabituating stimulus*—are less studied and not well understood.) Classic vigilance research has also identified three additional factors that influence response to alarms—*signal salience, modality*, and *complexity;* these factors are also examined in the section that follows.

Effect of stimulus frequency. The habituation effect is intensified with greater stimulus frequency and shorter time periods between stimulations or duration of respite between repetitions of the stimulus. A corollary to this observation is that behaviors do not diminish in quality as quickly if the interval between repeated stimulations is greater (Davis, 1970). Recently, investigations in the critical care setting have reported a high frequency of physiologic monitor alarms, ranging from one alarm occurring every 1.8 to one alarm every 10 minutes (Chambrin, 1999; Fidler et al., 2011; Siebig et al., 2010). On the basis of research findings, we can infer that a high frequency of monitor alarms observed in the ICU adversely affects vigilance

because the time interval between alarm occurrences is relatively short. Given the relative constancy of alarm signals and the relative brevity of the intervals between alarms, it is not surprising that clinicians habituate to alarms. Alarm fatigue leads to reduced vigilance and, as Drew (2011) has noted, clinicians' can ignore or deactivate alarms to minimize annoyance. Furthermore, clinicians may hesitate to initiate additional monitoring (ST-segment) from fear of increasing the number of clinical alarms— which may lead to underutilizing monitor features, and missing critical patient events (Drew, 2011).

The effect of stimulus frequency helps explain the assertion that concurrent alarms as if compete for clinicians' limited attention (i.e., when presented with multiple near-simultaneous alarms, clinicians can attend only to a very limited number of the alarms and must ignore or postpone response to the rest) span. Given that the overwhelming proportion (over 85%) of ICU physiologic monitor alarms are false, it is not surprising that clinicians attach little importance to them (Biot, Carry, Perdix, Eberhard, & Baconnier, 2000; Chambrin et al., 1999; Koski, Mäkivirta, Sukuvaara, & Kari, 1990; Siebig et al., 2010). Instead, clinician response to these alarms slows as clinicians habituate to them. The high frequency of FAs lessens the perceived importance of physiologic monitor alarms; this reduction in perceived importance undoubtedly negatively influences clinician response. This concept is in accord with findings from a recent survey in which 33% of respondents identified frequent FAs as the most important issue contributing to alarm fatigue, because such alarms are known to reduce attention and delay clinicians' response (HTF, 2011).

Over time, the constant onslaught of alarms results in clinicians' becoming habituated to alarms and disregarding alarm signals assumed to be relatively unimportant in the context of substantial workload demands. These cognitive processes account for clinicians' behavior (i.e.,

deferred response or no response to clinical alarms) in a variety of settings (Bitan, Meyer, Shinar, & Zmora, 2004; Varpio, Kuziemsky, MacDonald, & King, 2012). Laboratory studies also describe the detrimental effects of frequent FAs on attention and behavioral response. Notably, Bliss, Gilson, and Deaton (1995) conducted an influential study that investigated the alarm response-related "cry-wolf" effect on performance. Specifically, the investigators explored the influence of alarm reliability (a measure inversely related to the frequency of FAs) on alarm response frequency, speed, and accuracy. The investigators found that alarm *reliability* did not significantly affect response speed, but alarm *urgency* did affect alarm response frequency: high-urgency alarms were responded to more often than were low-urgency alarms. This finding is congruent with study results reported by Bitan et al. (2004).

Furthermore, Bliss et al. reported that greater than 90% of participants (N = 138) did not respond to all auditory alarms, but rather, tended to match their response rate to the estimated probability that the alarm was true. For example, if these clinicians believed an alarm system to be 90% reliable, they responded 90% of the time; if the clinicians believed an alarm system to be 10% reliable; the clinicians responded only 10% of the time.) Bliss and colleagues' research findings identified that when alarm reliability is low, ICU staff attaches relatively low importance to physiologic monitor alarm signals and, as a result, respond to only a small fraction of alarm signals. If the ultimate purpose of medical alarms is to improve the speed, accuracy of nurses' alarm responses, and alarm reliability— reducing the frequency of FAs is critical (Bliss et al., 1995). These findings add to the evidence that clinical understanding of alarm response can be elucidated by theory.

Effect of stimulus change. A key factor known to influence the habituation effect is the specificity of the repeated stimulus. During the course of habituation, the behavioral response

diminishes in quality with repeated stimulations; this reduction of quality is due to fatigue (i.e., the individual's weariness from repeated performance of identical elicited responses). McSweeney and Murphy (2009) hypothesized that the habituation effect can be disrupted if the stimulus is modified. This hypothesis implies that stimulus modification may produce a recovery in the response; the degree of recovery is inversely contingent upon the degree of similarity between the new stimulus and the previous one (Domjan, 1996). McSweeney and Murphy's hypothesis was substantiated in a neonatal ICU study in which investigators reported that RNs were more likely to respond to rare alarms (e.g., life threatening cardiac arrhythmia alarms) than they were to common alarms—oxygen saturation alarms (Bitan et al., 2004). Respiratory distress will generate monitor alarms (e.g., low oxygen saturation, excessively high/low respiration rate) which are ordinary in a neonatal ICU because the most common clinical emergency in the pediatric population are respiratory—and not cardiac arrest in origin (Young & Seidel, 1999). Therefore, it was not surprising that neonatal nurses were more prompt in responding to rare alarms (e.g., arrhythmia alarms) than to the ordinary ones (e.g., pulse oximetry alarms). This clinical observation suggests that a stimulus change can indeed disrupt habituation related to physiologic monitor alarms. This alarm-response is an intriguing result; however, other studies have not reported similar findings. For example, Varpio et al. (2012) reported little change in level of response among pediatric RNs for infrequent life threatening alarms (i.e., alarms with highest clinical priority). Of 446 physiologic monitor alarms, 13% (n = 34) were life-threatening; only 41% of those life-threatening alarms prompted the RNs to respond within 60-s of alarm sounding. The hypothesis that staff responds more appropriately to infrequent alarms than to frequent alarms merits further inquiry, given the potential implications of this hypothesis and the fact that current findings are inconclusive.

Effect of stimulus intensity. The habituation effect is influenced by the intensity of the stimulus. Specifically, behavioral responses deteriorate more slowly if the stimulus is more intense, and vice versa: if the signal is weak, the behavioral response declines more rapidly (Groves, Lee, & Thompson, 1969). The intensity of monitor alarm signals can be amplified in a variety of ways. The simplest approach is to increase the auditory volume of alarms.

Although increasing alarm volume has traditionally been used to heighten stimulus intensity, recent studies suggest that this practice is problematic. Amplifying auditory alarm volume is not the best strategy for capturing clinicians' attention because alarms that are too loud simply annoy patients and staff. In addition, if the alarm signal is too loud, staff may easily silence or pause the alarm before the clinical situation is attended to or properly resolved (Patterson, 1989). Making alarm sounds more shrill (i.e., higher in pitch) or otherwise more noxious would involve consideration of effects other than the effect of increasing signal amplitude or volume (Edworthy, 1994; Hedley-White, 1988; Schmidt & Baysinger, 1986). An alarm's excessive loudness can inhibit clinicians' ability to hear adjacent medical device alarms or to perform clinical tasks (Edworthy & Hellier, 2006; Momtahan, Tanslet, & Hetu, 1993; Patterson, 1990). Lastly, alarms that are too piercing are counterproductive. That is, if an alarm is too intense or otherwise aversive, staff may be provoked to deactivate the offending alarm, and subsequently, staff may not re-activate the alarm. Unfortunately, these harmful alarm setting practices are well documented both in the aviation and in the medical environment (Kerr & Hayes, 1983; Rood, Chillery, & Collister, 1985; Thorning & Ablett, 1985).

Groves, Lee, and Thompson (1969) affirm that the factors most influential in determining the degree and speed of habituation are signal frequency and intensity. To increase alarm intensity and thereby attenuate habituation, a more direct approach entails simply reducing the

frequency of alarm signals by individualizing and customizing alarm settings for monitored patients. Lawless (1994) famously described FA noise and data pollution by invoking Aesop's fable of the boy who cried "Wolf!" As noted earlier, when confronted with a high frequency of FAs, clinicians habituate to the alarm signals; this habituation results in potential delays in response to a significant alarm sounding, which can lead to tragic outcomes. By implementing approaches to reduce the frequency of nuisance alarms, "true" alarms—which are actionable and significant—would then be perceived by clinicians as being more intense. Researchers believe that the above alarm management strategy will ultimately improve outcomes—clinician attentiveness and response to alarms (Konkani, Oakley, & Bauld, 2012; Sendelbach, 2012).

Effect of time. The effects of habituation are temporary—that is, adaptation to a particular signal diminishes as time passes absent of the eliciting signal. This attenuation of adaptation, known as *spontaneous recovery*, may be short or long in duration; spontaneous recovery is closely related to the duration of the rest period (Domjan, 1996). To date, no studies have evaluated how staffs' time away from the clinical setting (e.g., due to vacation or some other cause of absence) might affect response to clinical alarms. This is an important area for future investigation. In a related study, Daly and Wilson (1983) sought to determine the effects of fatigue on the vigilance of RNs who were observing continuous ECGs. This laboratory study compared the performance of fatigued nurses with that of rested nurses. Nurses were characterized as being "fatigued" if they had just completed an 8-hour shift or if they had obtained less sleep than expected; nurses were characterized being "rested" if, during the previous night, they had had their normal amount of sleep or had not worked during the previous 8-hour period. The nurses' performance was measured in terms of *latency* (defined as elapsed time from initiation of alarm signal to RN response) and *accuracy* of arrhythmia identification;

these measures were assessed during four consecutive 20-min periods. The investigators reported that fatigue affected performance significantly: in general, the overall performance of fatigued RNs was diminished in comparison with that of their rested colleagues. Surprisingly, however, during the first 20-min assessment period, the fatigued nurses' performance was similar to that of the rested RNs; in fact, some fatigued RNs outperformed their rested counterparts. However, in the three subsequent assessment periods, the performance of the fatigued nurses deteriorated significantly; (during these periods, performance was assessed in terms of absence of omission errors and latency). Notably, Daly and Wilson's study is almost 30 years old and the study's results are consistent with the earlier findings from classic vigilance research. Additional studies, preferably those conducted in the "real-world" inpatient setting, are necessary to understand how a reprieve from clinical duties can influence physiologic monitor alarm response in our current care environment.

Signal salience. Mackworth's study of vigilance considers signal salience as a condition that influences human performance. Signal salience is described in terms of the stimulus' amplitude, gain on sensory channels, and stimulus duration. Early studies reported that improving signal salience occurred following incremental increase in acoustic amplitude on sensory channels being monitored (Corcoran, Mullin, Rainey, & Frith, 1977; Guralnick, 1972). In addition, Warm, Loeb, and Alluisi (1970) hypothesized that signals would be interpreted as being more salient by assessing their duration; short signals might be overlooked; longer signals were less likely to be missed.

Few studies have investigated the saliency of auditory alarms or visual alerts. An experimental study by Bliss, Fallon, and Nica (2007) reported that participants (i.e., psychology students in a laboratory setting) perceived long-duration signals as more representative of true

alarms; the students determined their behavioral responses based on alarm duration. The cognitive association of signal duration with alarm validity may explain reactivity, or lack thereof, among clinicians. That is, this association explains the finding that if an alarm's signal gradually attenuates, it is likely to be associated with an FA or alarm system unreliability (Bliss et al., 2007). Furthermore, in a neonatal ICU, Bitan et al. (2004) observed that over 90% of the time RNs did not respond to patients within a minute of the initiation of a monitor alarm signal. The investigators reported that although RNs responded to longer alarm signals, they monitored alarm durations and simply overlooked or ignored shorter alarms—on the assumption that shorter alarms were self-correcting or nuisance alarms. The results of Bitan et al. (2004) are in agreement with those of a pediatric study by Varpio et al. (2012), who reported similar coping strategies in use by RNs. The qualitative research suggests that (a) clinicians monitor alarm duration as a strategy for managing numerous alarms in the care environment and (b) this strategy accounts for delayed alarm responses (Varpio et al., 2012). Furthermore, the above clinical and laboratory findings are in accordance with a previous aviation study (Bliss, 2003).

The mechanism of signal salience illuminates our understanding of how clinicians selectively attend to some alarms—those that signify importance—and disregard others. Without this innate regulating mechanism, it would be difficult, if not impossible, to cope with simultaneous work demands and to accomplish tasks. Research confirms that alarms, whether they are actionable or not, disrupt patient care activities (Kerr & Hayes, 1983; Korniewicz, Clark, & David, 2008). Notably, Bitan et al. (2004) found that RNs in neonatal intensive care units hear monitor alarms an estimated 7% of the time for each assigned patient during a shift. Because RNs are usually assigned more than one patient, the exposure to physiologic monitor alarms arithmetically increases with the assignment of more patients. Moreover, Bitan et al.

concluded that the probability of a RN's responding to an alarm increases the longer the duration of the alarm—which relates to signal salience, a dimension known to influence human performance. That is, the likelihood of a RN's responding to an alarm within 15 s is 5%; within 30 s, 7%; and within 60 s, 10%. By delaying responses, alarm self-correction precludes RN response (Konkani et al., 2012). This finding supports the idea that clinicians perform some form of filtering, because attending to every monitor alarm, would render provision of care unmanageable.

The 2011 HTF cross-sectional survey revealed an interesting finding: 22% of responding clinicians ranked noise from other non-clinical alarms and sources (e.g., pages) in the top ten of most important contributors to alarm fatigue. Nearly half (42%) of respondents agreed that environmental background noise interferes with alarm recognition. Also, only 66% of respondents agreed that clinicians are sensitive to alarms and respond quickly. The survey's findings also suggest that the remaining respondents (34%) may have believed that the clinicians were desensitized to alarms and, as a consequence, did not efficiently respond. These findings support the view that clinicians are besieged with simultaneous inputs (e.g., alarms and other sources of noise) that must be simultaneously synthesized, assigned a meaning, and further analyzed in order to respond most effectively. Given constraints on the capacity of the human mind to accomplish this feat, it makes sense that all messages undergo a form of prioritization in order to determine which signal receives attention and a corresponding behavioral response. Furthermore, these results indicate that excessive clinical alarms and environmental noise can overburden staff, contribute to alarm fatigue, and, ultimately, decrease staff vigilance. Sabar and Zmora (1997) found that nurses respond to physiologic monitor alarms according to a set of assigned priorities and meanings that are associated with the severity of a patient's illness and

with the alarm parameter. On the basis of their clinical observation, the investigators reported that RN response time was significantly shorter for alarms associated with critically ill infants than for alarms associated with infants in less serious condition. This finding is in accord with the concept of assigning meaning to certain inputs (i.e., alarms) to determine a behavioral response. Notably, the mean RN response time of evening-shift RNs was twice that of day-shift RNs and night-shift RNs. This study is valuable in that it broadly examined RN alarm response times in relation to patient acuities and nursing shift; however, a more robust methodology and statistical analysis would have strengthened the study.

Signal modality. A second component thought to influence human performance is *signal modality*—the methods in which alarm signals are presented. Signal modality can be visual, auditory, or a combination of both. The majority of medical devices that generate alarm signals produce auditory signals; a smaller proportion of medical devices produce visual alerts. Signal modality affects the intensity of the alarm signal. In 2006, The International Electrotechnical Commission (IEC) issued to medical equipment manufacturers a set of alarm-system design standards (for both visual and auditory systems). These standards, collectively titled "Internal Organization for Standardization (IEC–ISO) standard 60601-1-8," provide guidance for alarm standardization by the medical device industry; however, these standards are not legally mandated. As a result, functionally equivalent medical devices manufactured by different companies use a variety of signal modalities (rather than a single modality; Wilcox, 2011). The diversity of signal modalities used by functionally identical devices significantly complicates clinicians' efforts to distinguish between and properly identify alarm types presented in clinical units. Vendor-dependent and even unit-specific alarm settings add heterogeneity and complexity to alarm systems management as clinicians attempt to distinguish a variety of alarms among the

profusion of devices used for patient care.

Alarm urgency is yet another factor that complicates accurate alarm recognition. *Alarm urgency* refers to the relative importance of responding to a given alarm. As with signal modality, alarm urgency schema are not standardized across devices produced by different manufacturers; in addition, medical facilities or even individual patient care units within facilities may use unique alarm urgency default settings. Researchers have found that low-priority events do not require audible alarms because there is usually sufficient time to notify clinicians of such events through visual methods (i.e., inspecting monitor displays for on-screen messages or observing patients; Edworthy, 1994; Edworthy & Hellier, 2005). Furthermore, lower-priority events typically represent non-urgent conditions that do not require immediate clinical interventions or attention from clinicians. One type of low-priority alarm—the *message* alarm—can use visual alerts to minimize alarm fatigue and reduce hospital noise pollution (Edworthy & Hellier, 2005). Years ago, Posner, Nissen, and Klein (1976) observed that clinicians responded more quickly to auditory alarms than to visual alarms; however, in ensuing years, clinical work conditions (i.e., staffing ratios) and monitoring technology have advanced significantly. In addition, past research has found that a traditional dual-mode approach (i.e., using both visual and auditory signals) can enhance monitoring efficiency (Craig, Colquhoun, & Corcoran, 1976; Doll & Hanna, 1989). While the method of communicating alarm conditions in relation to improving clinician response has varied over time, more recent research on clinical alarms underscores the importance of reducing the quantity of non-actionable alarms and their related signals in order to arrive at a condition that is both manageable and safe.

Stimulus complexity. Most patient monitoring systems produce a diverse repertoire of alarm tones (beeps, foghorns, etc.); these tones are pre-determined by manufacturers with regard

to alarm type and recommended sound volume level. This profusion of diverse alarm signals makes it difficult for clinicians to differentiate between and recognize specific signals (Edworthy & Hellier, 2006). A seminal study by Momtahan et al. (1993) compared the alarm identification accuracy of two groups of clinicians, (a) RNs and (b) anesthesiologists and anesthesiology technicians. The RN group and anesthesiologist–anesthesiology technician group were able to correctly identify only 39% and 40% of alarms, respectively. The investigators also noted that the clinicians were unable to recognize that the alarms' acoustic urgency was tied to the clinical urgency of the condition being signaled (Momtahan et al., 1993). Similarly, Momtahan et al. found that when several medical devices were used simultaneously on a patient, 50% of the clinicians reported being confused as to which device was producing an alarm. Similar findings have been reported in many other alarm studies (Cropp, Woods, Raney, & Bredle, 1994; HTF, 2011; Loeb, Jones, Behrman, & Leonard, 1990).

Patient monitoring research has reported that ICU staff members are prone to the effects of habituation. As a result of habituation, ICU staff differentiation of concurrent auditory alarm signals is often difficult—even when the alarms vary significantly in loudness, pitch, tone, and other modality features. The degree of perceived similarity among different alarm signals explains (a) clinicians' difficulties in discriminating between auditory alarms and (b) deficiencies in performance related to a lengthy response to an alarm. Among researchers, recommendations regarding the maximum number of signal types that can be processed in any given setting vary widely—from four auditory alarm types, maximum, to nine alarm types (Miller; 1956; Patterson, 1982; Sanders & McCormick, 1987). This recommendation originates from human information-processing ability and presents a dilemma: how to reconcile the need for signal specificity (in

order to minimize habituation) and limitations in clinicians' abilities to differentiate between types of auditory signals.

The psychology of vigilance and the habituation effect derived from a behavioral theory provides a framework that is relevant in understanding how a high alarm frequency can adversely affect clinicians' response to alarms and lead to alarm fatigue. As noted by Breznitz (1984) decades ago, both habituation and excessive FAs lead to inadequate responsiveness to TAs. This basic behavioral theory provides an overarching view for understanding how clinicians naturally habituate to physiologic monitor alarms as a consequence of imperfect technology and prevailing alarm settings.

Through understanding of the complex nature of alarms and the response demand that alarms enforce on clinicians' attention, appropriate alarm management interventions (e.g., involving optimization of alarm settings) can be implemented to promote vigilance and minimize habituation, and its adverse effect, alarm fatigue. See Figure 2.3 for an illustration of anticipated changes in alarm characteristics following an alarm management intervention in a hospital setting.

Conclusion

Few theoretical frameworks or conceptual models have informed the advance of alarm research. This lack of theory application to research may be a consequence of the fact that alarm research has largely been performed by physicians, biomedical engineers, and industry leaders and that this research is published in journals that ordinarily do not include a study's conceptual framework. Of the five existing nurse-led studies, four have been atheoretical; the third study, by Daly and Wilson (1983), refers only briefly to vigilance research. The general absence of conceptual definitions, coupled with the scarcity of applicable theories, contributes to the

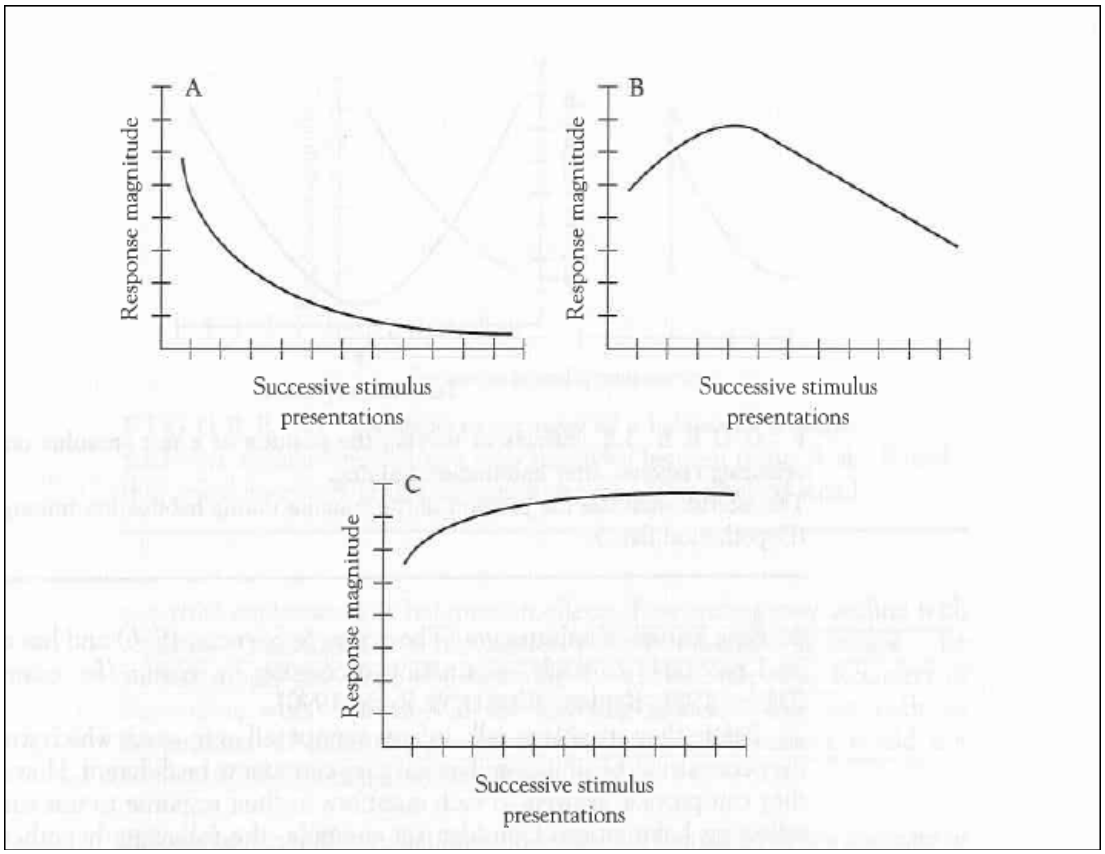
increasingly complex nature of alarm fatigue research. This lack of theoretical framework underpinning in the development of clinical alarm research science is a shortcoming that must be rectified.

Meleis (2007) has observed that, by facilitating outcome prediction and identification, theoretical frameworks render research more efficient, more useful, and more influential. Physiologic monitor alarm research is complex. Given the incidence and ramifications of inadequate response to alarms, the application of a theoretical framework to physiologic monitor alarm research is paramount. The application of Groves and Thompson's dual-process theory of response plasticity to research findings on repeated stimulation constitutes an adequate framework for conceptualizing the impact of physiologic monitor alarms on clinicians, their practice, and patient care. This framework has obvious shortcomings: it is based on neurophysiology studies and cognitive models, it is challenged by alternative theories and models, and it has not been applied to clinical research. However, despite these limitations, this theory has merit. Habituation: a dual-process theory of response plasticity illuminates the study of physiologic monitor alarms by providing insight into nursing practice situations and relationships between alarm signals and clinician response.

Habituation: a dual-process theory of response plasticity to repeated stimulation provides a basis for understanding the factors that contribute to habituation as related to managing multiple alarm signals and determining approaches for attracting and optimizing clinician attention and response to alarms. Also, this behavioral theory presents an explanation of humans' behavioral abilities to filter out unwanted messages (such as alarms), in order to effectively respond to the constant barrage of sensory information. These abilities enable humans to function—by selecting how and what we use our limited attentional capacity for (Wilcox, 2011).

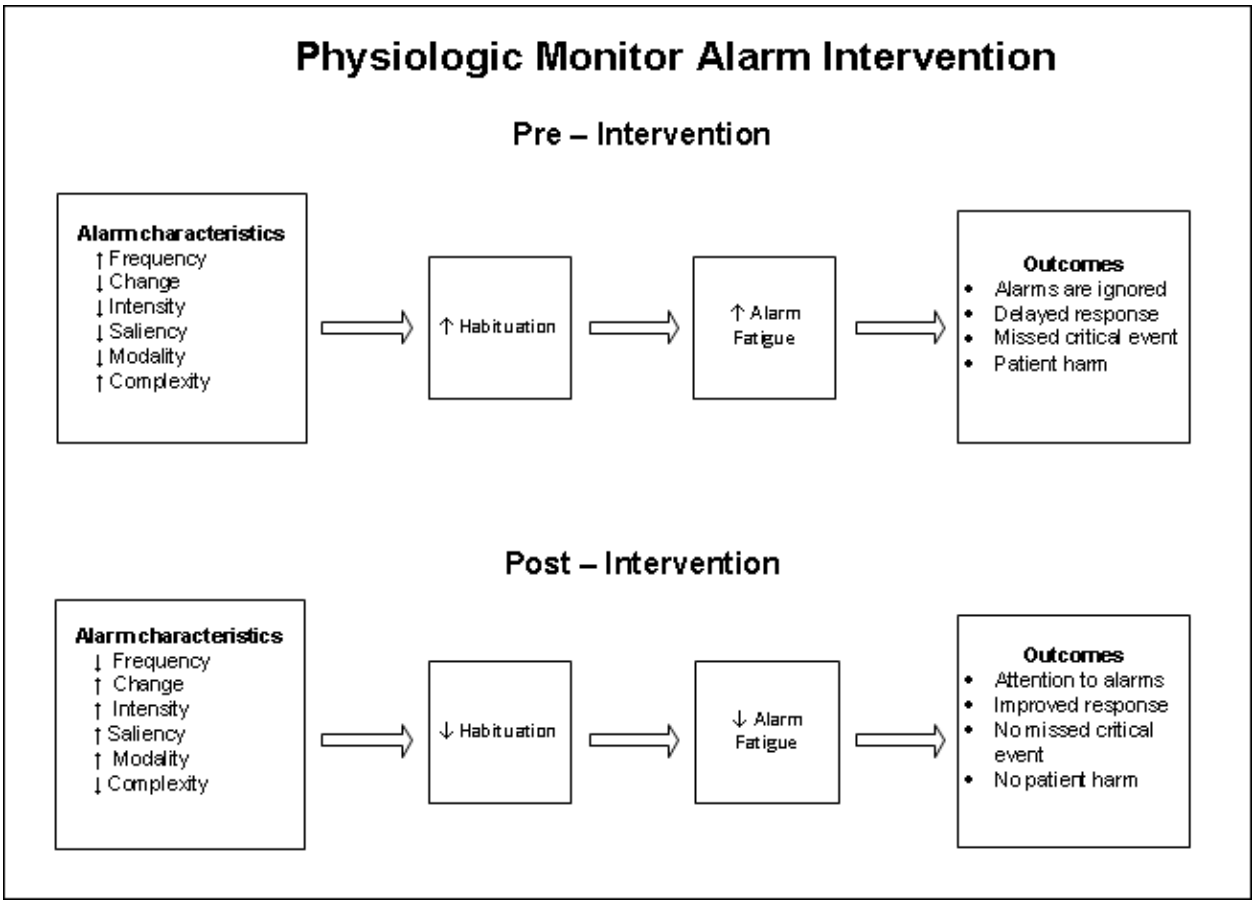
With this in mind, in-depth nursing research is necessary to study the impact of technology on nurses' selective attention and behaviors towards patient monitoring and clinical alarms. One important outcome of this research will be the attenuation and prevention of actual and potential alarm hazards. Groves and Thompson's theory presents opportunities to not only understand alarms, but also to improve clinician performance and, possibly, monitoring technology.

Figure 2.2. Schematic Illustration of Different Possible Outcomes of Repeated Presentations of a Stimulus on Elicited Behavior. ("The Essentials of Conditioning and Learning" by M. Domjan, 1996, Brooks/Cole Publishing).



Panel A illustrates the phenomenon of habituation. Panel B illustrates a transient sensitization effect followed by a habituation effect. Panel C illustrates the phenomenon of sensitization. (Hypothetical data.)

Figure 2.3 Schematic Illustration of Anticipated Changes in Physiologic Monitor Alarm Characteristics Following an Alarm Management Intervention (Mammone, 2013, with permission).



References

Ballard, J. C. (1996). Computerized assessment of sustained attention: A review of factors affecting vigilance performance. *Journal of Clinical and Experimental Neuropsychology, 18*, 843–863.

Biot, L., Carry, P. Y., Perdrix, J. P., Eberhard, A., & Baconnier, P. (2000). Évaluation clinique de la pertinence des alarmes en reanimation. *Annales Françaises d'Anesthésie et de Réanimation, 19*, 459–466.

Bitan, Y., Meyer J., Shinar, D., & Zmora, E. (2004). Nurses' reactions to alarms in a neonatal intensive care unit. *Cognition, Technology & Work, 6*, 239–246.

Bliss, J. P. (2003). An investigation of alarm related incidents in aviation. *International Journal of Aviation Psychology, 13*(3), 249–268.

Bliss, J. P., & Dunn, M. C. (2000). Behavioural implications of alarm mistrust as a function of task workload. *Ergonomics, 43*(9), 1283–1300.

Bliss, J. P., Fallon, C. K., & Nica, N. (2007). The role of alarm signal duration as a cue for alarm validity. *Applied Ergonomics, 38*, 191–199.

Bliss, J. P., Gilson, R. D., & Deaton, J. E. (1995). Human probability matching behavior in response to alarms in varying reliability. *Ergonomics, 38*(11), 2300–2312.

Block, F. (2011). Expert Perspective: If I ran the alarms. *Clinical Alarms 2011 Summit*. Retrieved from http://www.aami.org/htsi/alarms/pdfs/2011_Alarms_Summit_publication.pdf

Breznitz, S. (1984). *Cry–wolf: The psychology of false alarms.* Hillsdale, NJ: Lawrence Erlbaum.

Chambrin, M. –C., Ravaux, P., Calvelo–Aros, D., Jaborska, A., Chopin, C., & Boniface, B. (1999). Multicentric study of monitoring alarms in the adult intensive care unit (ICU): A descriptive analysis. *Intensive Care Medicine, 25*, 1360–1366.

Chambrin, M.–C. (2001). Alarms in the intensive care unit: How can the number of false alarms be reduced? *Critical Care, 5*(4), 184–188.

Corcoran, D.W. J., Mullin, J., Rainey, M. T., & Frith, G. (1977). The effects of raised signal and noise amplitude during the course of vigilance tasks. In R.R. Mackie, ed. *Vigilance: Theory, Operational Performance and Physiological Correlates* (pp. 645–664). New York, NY: Plenum.

Craig, A., Colquhoun, W. P., & Corcoran, D. W. J. (1976). Combining evidence presented simultaneously to the eye and the ear: A comparison of some predictive models. *Perception and Psychophysics, 19*, 473–484.

Cropp, A. J., Woods, L. A., Raney, D., & Bredle, D. L. (1994). Name that tone: The proliferation of alarms in the intensive care unit. *Chest, 105*(4), 1217–1220.

Daly. H. W. (1968). Acute coronary care—A five year report. *The American Journal of Cardiology, 21*(2), 252–257.

Daly, B. J., & Wilson, C. A. (1983). The effect of fatigue on the vigilance of nurses monitoring electrocardiograms. *Heart & Lung, 12*(2), 384–388.

Davis, M. (1970). Effects of interstimulus interval length and variability on startle–response habituation in the rat. *Journal of Comparative and Physiological Psychology, 72*, 177–192.

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance.* London, UK: Academic Press.

Doll, T. J., & Hanna, T. E. (1989). Enhanced detection with sonar display. *Human Factors, 31,* 539–550.

Domjan, M. (1996). Habituation and sensitization. In M. Domjan (Eds.), *The essentials of conditioning and learning* (pp. 23–36). Pacific Grove, CA: Brooks/Cole Publishing.

Drew, B. J. (2011, October). Clarion Theme 1: Deepen all stakeholders' understanding of use environments. *Clinical Alarms 2011 Summit.* Retrieved from http://www.aami.org/htsi/alarms/pdfs/2011_Alarms_Summit_publication.pdf

East, T. D. (1992). Computers in the ICU; Panacea or Plague? *Respiratory Care, 37,* 170–180.

East, T. D., Wallace, J., Morris, A. H., Gardner, R. M., & Westenskow, D. R. (1995). Computers in critical care. *Critical Care Nursing Clinics of North America, 7*(2), 203–217.

Edworthy, J. (1994). The design and implementation of nonverbal auditory warnings. *Applied Ergonomics, 25*, 202–210.

Edworthy, J., & Hellier, E. (2005). Fewer but better auditory alarms will improve patient safety. *Quality & Safety in Health Care, 14,* 212–215.

Edworthy, J., & Hellier, E. (2006). Alarms and human behavior: Implications for medical alarms. *British Journal of Anaesthesia, 97*(1) 12–17.

Emergency Care Research Institute (2007). The hazards of alarm overload: Keeping excessive physiologic monitoring alarms from impeding care. *Health Devices, 36*(3), 73–83.

Emergency Care Research Institute (2009). Top 10 technology hazards. *Health Devices, 38*(11), 364–373.

Emergency Care Research Institute (2010). Top 10 technology hazards for 2011. *Health Devices, 39*(11), 404–416.

Emergency Care Research Institute (2011). *Top 10 technology hazards for 2012*. Retrieved from https://www.ecri.org/Documents/Secure/Health_Devices _Top_10_Hazards_2012.pdf

Emergency Care Research Institute (2012a). *Top 10 technology hazards for 2013*.Retrieved from https://www.ecri.org/Documents/Secure/Health_Devices _Top_10_Hazards_2013.pdf

Emergency Care Research Institute. (2012b). *Medical device alarm related terminology*. Retrieved from http://www.aami.org/meetings/summits/alarms.html

Emergency Care Research Institute (2013). Top 10 health technology hazards for 2014. *Health Devices, 42*(11), 354–380.

Fairman, J. (1992). Watchful vigilance: Nursing care, technology, and the development of intensive care units. *Nursing Research, 41*(1), 56–60.

Fairman, J., & Kagan, S. (1999). Creating critical care: The case of the hospital of the University of Pennsylvania, 1950–1965. *Advances in Nursing Science, 22*(1), 63–77.

Fidler, R., Pickham, D., & Drew, B. J. (2011). Quantification of hospital cardiac monitoring alarms. *Journal of Electrocardiology, 44*, 744.

Graham, K., & Cvach, M. (2010). Monitor alarm fatigue: Standardizing use of physiological monitors and decreasing nuisance alarms. *American Journal of Critical Care*, *19*(1), 28–34.

Groves, P. M., Lee, D., & Thompson, R. F. (1969). Effects of stimulus frequency and intensity on habituation and sensitization in acute spinal cat. *Physiology and Behavior, 4,* 383–388.

Groves, P. M., & Thompson, R. F. (1970). Habituation: A dual process theory. *Psychological Review, 77,* 419–459.

Guralnick, M. J. (1972). Observing responses and decision processes in vigilance. *Journal of Experimental Psychology, 93*, 239–244.

Hagenouw, R. (2007). Should we be alarmed by our alarms? *Current Opinion in Anesthesiology, 20,* 590–594.

Healthcare Technology Foundation (2011). *National Clinical Alarm Survey: Perceptions, issues, improvements, and priorities of healthcare professionals*. Retrieved from http://www.thehtf.org

Hedley–White, J. (1988). Monitoring and alarm philosophy and practice. *Baillière's Clinical Anaesthesiology, 2*, 379–389.

Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica, 129,* 18–25.

Holden, R. J., & Karsh, B.–T. (2009). A theoretical model of health information technology usage behavior with implications for patient safety. *Behaviour & Information Technology, 28*(1), 21–23.

Imhoff, M., & Kuhls, S. (2006). Alarm algorithms in critical care monitoring. *Anesthesia & Analgesia, 102,* 1525–1537.

International Electrotechnical Commission (2006). Medical electrical equipment– Part 1–8: General requirements for basic safety and essential performance – Collateral Standard: *General requirements, tests and guidance for alarm systems in medical electrical equipment and medical electrical systems* IEC 60606–1–8.

Jerison, H. J. (1959). Effects of noise on human performance. *Journal of Applied Psychology, 43* 96–101.

Keller, J. (2011, October). *Perspectives on problem reporting, RCAs, and alerts and recalls.* Presentation at the AAMI/FDA Medical Device Alarm Summit. Herdon, VA. Retrieved

from http://www.aami.org/meetings/summits/alarms/Materials /PDFs/keller-Tues-1045.pdf

Keller, J. P., Diefes, R., Graham, K., Meyers, M., & Pelczarski, K. (2011). Why clinical alarms are a “top ten” hazard: How you can help reduce the risk. *Biomedical Instrumentation & Technology*, 45(1), 17–23.

Kerr, J. H., & Hayes, B. H. (1983). An “alarming” situation in the intensive therapy unit. *Intensive Care Medicine, 9,* 103–104.

Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (1999). *To err is human: Building a safer health system*. Washington, DC: Institute of Medicine.

Konkani, A., Oakley, B., & Bauld, T. J. (2012). Reducing hospital noise: A review of medical device alarm management. *Biomedical Instrumentation & Technology*, *46*(6), 478–487.

Korniewicz, D. M., Clark, T., & David, Y. (2008). A national online survey on the effectiveness of clinical alarms. *American Journal of Critical Care, 17*(1), 36–41.

Koski, E., M., Mäkivirta, A., Sukuvaara, T., & Kari, A. (1990). Frequency and reliability of alarms in the monitoring of cardiac postoperative patients. *International Journal of Clinical Monitoring and Computing, 7*(2), 129–133.

Lawless, S. T. (1994). Crying wolf: False alarms in a pediatric intensive care unit. *Critical Care Medicine*, *22,* 981– 985.

Loeb, R. G., Jones, B. R., Behrman, K. H., & Leonard, R. J. (1990). Anesthetists cannot identify audible alarms. *Anesthesiology, 72,* A538.

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology, 1*, 6–21.

Mammone, T . (2013). *Physiologic monitor alarm burden* (Unpublished doctoral dissertation). University of California, San Francisco.

Matthews, G. (2000). Vigilance and sustained attention. In Matthews, G., Davies, D. R., Westerman, S. J., & Stammers, R.B. (2000). *Human Performance: cognition, stress, and individual differences* (pp. 107–124). East Sussex, UK: Psychology Press.

McSweeney, F. K., & Murphy, E. S. (2009). Sensitization and habituation regulate reinforce effectiveness. *Neurobiology of Learning and Memory, 92,* 189–198.

Meleis, A. I. (2007). *Theoretical nursing: Development and progress* (4th ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Desensitized. (2012). In *Merriam–Webster.com.* Retrieved from http://www.merriam–webster.com/dictionary/desensitized

Habituate. (2012). In *Merriam–Webster.com* Retrieved from http://www.merriam–webster.com/dictionary/habituate

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Momtahan, K. L., Tanslet, B.W., & Hetu, R. (1993). Audibility and identification of auditory alarms in the operating room and intensive care units. *Ergonomics, 36,* 1159–1176.

Olympio, M. (2004). APSF workshop recommends new standards. *Anesthesia Patient Safety Foundation Newsletter, 19*(41), 52–53.

Patterson, R. D. (1982). Guidelines for auditory warning systems on civil aircraft. *Civil Aviation Authority Report* 8201.

Patterson, R. D. (1989). Guidelines for the design of auditory warning sounds. *Proceeding of the Institute of Acoustic, 11*(5), 17–24.

Patterson, R. D. (1990). Auditory warning sounds in the work environment. *Philosophical Transactions of the Royal Society of London, B, 327,* 485–492.

Posner, M. I., Nissen, M. J., & Klein, R. (1976). Visual dominance: An information processing account of its origins and significance. *Psychological Review, 83*, 157–171.

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D., Colombo, J., . . . Thompson, R. F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory, 92*(2), 135–138.

Rood, G. M., Chillery, J. A., & Collister, J. B. (1985). Requirements and applications of auditory warnings to military helicopters. *Ergonomics International 85,* (pp. 169–171). London, UK: Taylor & Francis.

Sabar, R., & Zmora, E. (1997). Nurses response to alarms from monitoring systems in NICU. [Abstract]. *Pediatric Research, 41*, 174.

Sanders, M. S., & McCormick, E. J. (1987). *Human Factors Engineering and Design* (6th ed.). Montreal, Canada: McGraw–Hill.

Schmidt, S. I., & Baysinger, C. L. (1986). Alarms help or hindrance? *Anaesthesiology, 64*, 65.

Sendelbach, S. (2012). Alarm fatigue. *Nursing Clinics of North America*, *47*(3) 375–382.

Siebig, S., Kuhls, S., Imhoff, M., Gather, U., Schölmerich, J., & Wrede, C. (2010). Intensive care unit alarms—How many do we need? *Critical Care Medicine*, *38*(2), 451–456.

Thompson, R. F. (2009). Habituation: A History. *Neurobiology of Learning and Memory, 92*(2), 127–134.

Thompson, R. F., & Spencer, W. A. (1966). Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychological Review, 73*(1), 16–43.

Thorning, A. G., & Ablett, R. M. (1985). Auditory warning systems on commercial transport. *Ergonomics International 85,* (pp. 166–168). London, UK: Taylor & Francis.

Thorpe, W. H. (1956). *Learning and instinct in animals.* London, UK: Methuen.

Varpio, L., Kuziemsky, C., MacDonald, C., & King, J. W. (2012). The helpful or hindering effects of in–hospital patient monitor alarms on nurses. *Computers, Informatics, Nursing, 30*(4), 210–217.

Warm, J. S. (1977). Psychological processes in sustained attention. In R. R. Mackie (Eds.), *Vigilance: Theory operational performance and physiological correlates* (pp. 623–644). New York, NY: Plenum Publishing.

Warm, J. S. (1984). *An introduction to vigilance in sustained attention to human performance* (pp.1–14). Chichester, UK: Wiley.

Warm, J. S. (1993). Vigilance and target detection performance. In B. Messick Huey., & C. D. Wickens (Eds.), *Workload transitions: Implications for individual and team performance* (pp.139–170). Washington, DC: National Academy Press.

Warm, J. S., & Berch, D. B. (1985). Sustained attention in the mentally retarded: The vigilance paradigm. In N. R Ellis & N. W. Bray (Eds.), *International review of research in mental retardation.* Vol 13. Orlando, FL: Academic Press.

Warm, J. S., Loeb, M., & Alluisi, E. A. (1970). Variations in watchkeeping performance as a function of the rate and duration of visual signals. *Perception and Psychophysics, 7*, 97–99.

Warm, J. S., Matthews, G., & Finomore, V. S. (2008). Vigilance, Workload, and Stress. In P. A. Hancock, & J. L. Szalma (Eds.), *Performance under stress* (pp. 115–141). Abington, Oxon: Ashgate Publishing Group.

Watts, F. N. (1979). Habituation model of systemic desensitization. *Psychological Bulletin, 86*(3), 627–637.

Wilcox, S. B. (2011). A human factors perspective: Auditory alarm signals. *Biomedical Instrumentation and Technology, 45*(4), 284–289.

Young, K. D., & Seidel, J. S. (1999). Pediatric cardiopulmonary resuscitation: A collective review. *Annals of Emergency Medicine, 33*(2), 195–205.

Chapter 3

Design and Measurement Issues

Over 150 years ago, Florence Nightingale famously said "It may be a strange principle to enunciate as the very first requirement in a hospital that it should do the sick no harm" (Nightingale, 1859/2010). Although this tenet was primarily intended for RNs, Nightingale's classic, fundamental principle of patient safety applies to all clinicians. Indeed, maintaining patients' safety is of paramount importance for clinicians. Consequently, as administration of care has become more complex, so too have efforts to protect patients from iatrogenic injury.

Medical devices can save patient lives; however, they can also cause harm if poorly designed or ill-advisedly used. For three consecutive years, the Emergency Care Research Institute (ECRI) has ranked alarm hazards as foremost on its annual list of top ten medical technology hazards (ECRI, 2010, 2011, 2012). Fortunately, most alarm-related errors are preventable, and known risks can be minimized. Alarm-related error prevention and risk minimization begins with a thorough critique of physiologic monitor alarm research and identification of hazards. One important type of alarm hazard is *alarm fatigue*, as experienced by both clinicians and patients alike.

The body of knowledge on alarm fatigue related to physiologic monitors has been described in Chapter 2. This chapter discusses many deficiencies, including limitations in study design, sampling methodology, and instruments used to identify, measure, and annotate physiologic monitor alarms. In addition, the review discusses limitations in studies that have assessed clinicians' perception and behavior towards clinical alarms. The following section presents critiques of the methodology used in previous research on alarm fatigue associated with physiologic monitor alarms. In particular, this chapter focuses on methods of data collection and

examines the psychometric properties of selected measures that are most suitable for measuring the prevalence of physiologic monitor alarms and related concepts regarding alarm fatigue.

Study Designs

Thirty-eight studies on physiologic monitor alarms were reviewed in the literature review chapter. One study used qualitative research methods; of the remaining 37 studies, 31 studies used an observational design, and six studies used quasi-experimental designs. Of the observational design studies focused on physiologic monitor alarms, 20 were prospective cohort studies (including one pilot study), five were retrospective cohort studies, and five were cross-sectional studies. One observational investigation in a neurological intensive care unit examined noise related to physiologic monitor alarms and employed a prospective cohort study design.

Qualitative Research

Alarm fatigue associated with physiologic monitor alarms is a complex and understudied phenomenon. To gain a holistic understanding of alarm fatigue, a qualitative approach that involves examination of the clinician's subjective experience is essential. Use of qualitative methodology is appropriate for exploration of the behaviors, perspectives, feelings, and experiences of individuals, groups, and cultures in the context of their work (Holloway & Wheeler, 2002). A review of the literature revealed that a single qualitative study assessed RNs' response to alarms and reported (a) their explanations of the impact of physiologic monitor alarms on workflow and (b) their strategies for responding to alarms (Varpio, Kuziemsky, McDonald, & King, 2012). A qualitative approach enabled the investigators to explore pediatric RNs' subjective experiences and to describe RNs' lived experience in the context of their work with physiologic monitors in an acute care setting. Given the limitations of this lone published

study, its findings should be interpreted judiciously. From critical analysis of this study, it is apparent that future qualitative studies would strengthen the existing body of knowledge.

Quantitative Research

Quantitative studies primarily focus on the relationship between predictor and outcome variables; accordingly, researchers first identify variables of interest, then develop operational definitions of those variables, and finally proceed to collect relevant data (Polit & Beck, 2004). Three observational designs are commonly used in quantitative clinical research: cohort, cross-sectional, and case control. Also, researchers can use a quasi-experimental design that includes an intervention but lacks randomization or a randomized clinical trial (RCT) design. Hulley et al. (2001) assert that no single approach is better than any other; rather, each research question requires a judgment about which design is most likely to yield satisfactory answers.

Observational study designs. Among the reviewed studies, the majority of physiologic monitor alarm studies that have measured alarm incidence) have primarily used observational designs (i.e., cohort and cross-sectional). Fewer studies have used quasi-experimental designs, and no studies have used RCT designs. This distribution of study design types is expected, because the study of a new subject commonly begins with observational studies (Hulley et al., 2001).

Cohort Studies. A *cohort study* is a study that examines a group of subjects (i.e., a "cohort") over time; a *cohort* is a group of individuals who share a common characteristic. Cohort designs may be either descriptive or analytic. *Descriptive studies* describe the incidence of certain outcomes over time; *analytic studies* analyze associations between predictor and outcome variables (Hulley et al., 2001). Cohort studies can have either a *control group* or a *comparison group*; these other groups are also investigated throughout the study period

(Crombie, 1996). A control group ordinarily receives no new intervention and is often described as receiving the "standard of care." Participants in a comparison group can receive a treatment that is different from but similar to the treatment that they would typically receive. Comparison groups are highly useful, because they provide context for understanding the study's findings (Polit & Beck, 2004). In cohort studies, the inclusion of either a control group or a comparison group is not necessary, and many studies do not have either of these types of groups.

Cohort studies may also be either prospective or retrospective, according to when subjects are identified (Carlson & Morrison, 2009). In *prospective studies* (also called *concurrent studies*), the investigator selects or recruits a sample and measures predictor variables before the outcomes occur; prospective studies may be either longitudinal or cross-sectional. In contrast, in *retrospective studies*, the outcomes have already occurred (Hulley et al., 2001; Mann, 2003).

Prospective cohort alarm studies. The literature search identified 20 alarm studies that used a prospective cohort design. All of the studies were relatively brief; in none of the investigations did the study period exceed 1,971 hours or a 28-day period (Blum, Kruger, Sanders, Gutierrez, & Rosenberg, 2009; Chambrin et al., 1999). The primary aim of most of the prospective studies was to describe the incidence and frequency of physiologic monitor alarms and in some cases, other clinical alarms; the studies were conducted in a variety of settings (Biot, Carry, Perdrix, Eberhard, & Baconnier, 2000; Bitan, Meyer, Shinar, & Zmora, 2004; Blum et al., 2009; Chambrin et al., 1999; Görges, Markewitz, & Westenskow, 2009; Gross, Dahl, & Nielsen, 2011; Jacobs & Eron, 2007; Koski, Mäkivirta, Sukuvaara, & Kari, 1990; Lawless, 1994; Mäkivirta & Koski, 1994; Mäkivirta, Koski, Kari, & Sukuvaara, 1991, O'Carroll, 1986; Pan & Gravenstein, 1994; Schmid et al., 2011; Siebig et al., 2010a, 2010b; Taenzer, Pyke, McGrath, & Blike, 2010; Talley et al., 2011; Tsien & Fackler, 1997; Whalen, Covelle, Piepenbrink,

Villanova, Cuneo, & Awtry, 2013; Whalen, Covelle, Piepenbrink, Villanova, Cuneo, & Awtry, 2013; Wiklund, Hök, Ståhl, & Jordeby-Jönsson, 1994).

The strengths of prospective studies lie in their thorough measurement of variables, definition of prevalence, and identification of potential causes of conditions (Hulley et al., 2001). The majority of the prospective alarm studies were able to achieve these objectives; the studies' reported alarm incidence ranged from 1,214 alarms (in a 200-hr study) to 293,049 alarms (in a 28-day study; Blum et al., 2009; Görges et al., 2009). In addition, several prospective studies reported potential causes of physiologic monitor alarms: motion artifacts, equipment failures, staff manipulations, and nursing interventions such as suctioning and turning patients (Chambrin et al., 1999; Görges et al., 2009; Schmid et al., 2011; Siebig et al., 2010a, 2010b; Tsien & Fackler, 1997). Prospective cohort designs also have weaknesses—specifically, that they are expensive and resource-intensive. Accordingly, these factors must be carefully considered, given that they are significantly stronger than retrospective studies (Mann, 2003).

In the present review, only one prospective study—an investigation by Taenzer et al. (2010)—used an experimental group–comparison group design (specifically, the study had an experimental group and two comparison groups). Taenzer et al. performed a before-and-after concurrent study of the implementation of a patient surveillance system that used pulse oximetry surveillance to facilitate early recognition and detection of patients' clinical deterioration. This study compared the orthopedic units' outcomes (i.e., rates of rescue events and transfers to higher levels of care) with the outcomes of two comparable acute surgical units. The investigators reported that their use of unconventional alarm settings, a standardized alarm adjustment schema, and an alarm notification delay resulted in a low incidence of oxygen saturation (SpO_2) alarms.

Retrospective cohort alarm studies. Retrospective cohort studies offer other important features: they are less costly and time consuming than are prospective studies, because in the former type of study, the data have already been collected (Hulley et al., 2001). Physiologic monitor alarm studies using a retrospective cohort study design have been feasible because researchers have contributed to and maintained large collections of physiologic signal and waveform data. Five such studies have included retrospective analyses that utilized stored patient records from large databases, including the free web-based PhysioNet MIMIC II Database and manufacturer-specific databases (Aboukhalil, Nielsen, Saeed, Mark, & Clifford, 2008; Burgess, Herdman, Berg, Feaster, & Hebsur, 2009; Rheineck-Leyssius & Kalkman, 1998; Welch, 2011; Zong, Moody, & Mark, 2004). These retrospective studies were predominately conducted to test arrhythmia algorithms, to assist with setting alarm limits for early warning systems, and to determine the incidence of alarms at various alarm thresholds and delay settings. Drawbacks of retrospective cohort design are that (a) investigators have less control over the selection of subjects and (b) the data may be incomplete or may have been measured or otherwise obtained in suboptimal conditions (Hulley et al., 2001). These limitations only minimally constrain the feasibility of retrospective physiologic monitor alarm studies, because the quantity of objective physiologic data available in reputable databases from contributors worldwide is substantial.

Cross-sectional studies. Cross-sectional studies attempt to determine the prevalence of various conditions, diseases, risk factors, or other outcomes, as well as the factors associated with such outcomes (Carlson & Morrison, 2009; Hulley et al., 2001; Polit & Beck, 2004). A major advantage of cross-sectional studies is that no time is lost in waiting for the outcome to be observed, because all measurements are made on a single occasion; hence, cross-sectional studies do no entail lengthy follow-up periods (Hulley et al., 2001). This benefit enables a shorter

time commitment for researchers and a shorter time commitment and less inconvenience for participants. Furthermore, cross-sectional study designs enable efficient evaluation of a large sample of subjects, which makes this design practical and economical (Carlson & Morrison, 2009).

Serial surveys. Hulley et al. (2001) have noted that a series of cross-sectional studies involving a single population observed at different points in time can be useful for drawing inferences about patterns in behavior, perceptions, and experiences that change over time. A potential limitation of serial surveys that examine the same cohort of participants is that the initial survey can produce a learning effect, thereby influencing participants' responses to subsequent surveys. To mitigate this limitation, serial surveys can study different participants over time. A serial cross-sectional approach was utilized in two multisite clinical alarm studies sponsored by the HTF (HTF, 2011; Korniewicz, Clark, & David, 2008). The investigators conducted two cross-sectional survey studies 5 years apart, to determine (a) changes in clinicians' perception of clinical alarm issues, (b) alarm management improvements made at their hospitals, and (c) priorities for future efforts to reduce alarm hazards. The investigators' findings summary included a comparative analysis of the 2006 and 2011 surveys; this methodology was important for evaluating progress and identifying areas in need of further action to minimize clinical alarm fatigue.

Surveys. Three other cross-sectional alarm studies used a straightforward survey design. These cross-sectional studies offered insight into RNs' and physicians' behaviors concerning the use of parameter alarm limits, their opinions on the frequency of FAs, and their impression of techniques designed to reduce alarms in order to minimize acoustic stress and improve patient safety (Block, Nuutinen, & Ballast, 1999; Koski, Mäkivirta, Sukuvaara, & Kari, 1995; Siebig et

al., 2009). In all of the cross-sectional studies, use of a cross-sectional survey design enabled investigators to obtain answers to the research questions and to examine the distribution of clinician responses in the samples. Cross-sectional studies have inherent limitations; for example, such studies do not establish sequences of events and, as a result, determination of causal relationships is difficult. In addition, cross-sectional studies measure prevalence rather than incidence (Hulley et al., 2001; Polit & Beck, 2004). For instance, the HTF 2011 study reported that nuisance alarms occurred frequently (i.e., 76%) and that hospitals had adverse patient events related to alarm problems (i.e., 18%); however, these findings did not conclusively indicate that nuisance alarms caused the adverse events. Furthermore, cross-sectional study designs typically require use of extensive analysis (rather than intensive analysis), and often have low internal validity and high external validity (Carlson & Morrison, 2009; Polit & Beck, 2004). Extensive analysis was noted in the three multisite cross-sectional survey studies (HTF, 2011; Korniewicz et al., 2008; Siebig et al., 2009) and to a lesser degree, in the physician studies (Block et al., 1999; Koski et al., 1995).

Despite their inherent limitations, cross-sectional study designs are convenient for examining networks of causal links and for generating screening hypotheses (Carlson & Morrison, 2009; Hulley et al., 2001). The five cross-sectional studies (Block et al.; HTF; Korniewicz et al.; Koski et al., Siebig et al.) provide estimates of alarm fatigue prevalence; more important, these studies also provide reasons why clinicians do not respond to alarms and suggest remedies for improving clinicians' recognition of and response to alarms. Technological and practice solutions often originate from front-line clinicians, whose input can be readily obtained using self-administered questionnaires (SAQs) that present open-ended questions. A web-based SAQ was used in the HTF (2011) survey of clinical alarms. The survey's

questionnaire, which contained seven open-ended questions, generated 3,192 comments from front-line clinicians. Unfortunately, although this survey was conducted in 2011, as of April 2013, analysis of the survey comments has not been reported. Use of open-ended questions in SAQs enables participants to respond in their own words; however, this flexibility in responding to questions makes data analysis time-consuming and challenging, because responses must be classified according to fixed categories (Polit & Beck, 2004).

Quasi-experimental study design. Campbell and Stanley (1963) popularized a class of studies called "quasi-experiments," in which the predictor variable is manipulated, but participants are not randomly assigned to groups. Quasi-experiments are used to test descriptive hypotheses about manipulable causes; the inclusion of a control group or comparison group can be used to support an opposing conclusion in a quasi-experiment about the outcome in the absence of the intervention (Shadish, Cook, & Campbell, 2002). In addition, quasi-experimental studies use self-selection or administrator selection rather than randomization—a design feature that is useful for investigations in which pre-selection or randomization is difficult or impossible (Polit & Beck, 2004).

The benefits of the quasi-experiment approach to sample selection—reduced expenditures of resources and time—must be weighed against limitations imposed on statistical analysis (Polit & Beck, 2004; Shadish et al., 2002). The advantages of this design are derived from the forgoing of randomization to group and consequent minimization of time and resource expenditures. On the other hand, the limitations of the quasi-experiment design are inherent difficulties with statistical analysis because of the lack of certain controls (Polit & Beck, 2004; Shadish et al., 2002). In the review, six alarm studies used a quasi-experimental design (Cropp, Woods, Raney, & Bredle, 1994; Cvach, Biggs, Rothwell, & Charles-Hudson, 2012; Daly, 1983;

Graham & Cvach, 2010; Solsona et al., 2001; Whalen et al., 2013). Three of these studies were conducted in adult progressive care or telemetry units. The intervention in one study focused on daily electrocardiographic (ECG) electrode changes (Cvach et al., 2012); the intervention in the other two studies modified alarm limit thresholds and changed priority messaging levels for cardiac arrhythmias and parameter alarm thresholds (Graham & Cvach, 2010; Whalen et al., 2013). Neither study used a control group (i.e., intervention groups served as their own controls), and study units were assigned to the intervention group by a nurse administrator. The primary aim of both studies was to reduce the high frequency of nuisance (non-actionable) alarms. Following intervention, each study reported moderate reductions in the incidence of physiologic monitor alarms—unadjusted for cardiac monitoring time.

The remaining three quasi-experimental studies are uniquely important in that each investigated a different yet important aspect of physiologic monitor alarm research. The studies investigated audio alarm recognition by staff (Cropp et al., 1994), effects of fatigue on the vigilance of RNs observing continuous ECGs (Daly, 1983), and whether documenting alarm parameters in patients' records is an effective strategy for improving alarm limit adjustments (Solsona et al., 2001). For all six studies, the use of a quasi-experimental research design was appropriate and practical.

Randomized clinical trial (RCT) study design. Clinical trials are designed to assess the effectiveness of clinical interventions. Accordingly, in clinical trials, an observed difference between two or more randomly assigned groups is likely to be due to the intervention and not to any pre-existing differences between groups (Shadish et al., 2002). The primary objective of RCT studies is to determine whether an intervention is more effective than the standard of care. Unlike the observational study design, the RCT design has the ability to determine causality

(Hulley et al., 2001; Polit & Beck, 2004). In health science research, randomized experiments are highly valued and are often referred to as the "gold standard" for treatment outcome research (Shadish et al., 2002). To date, no published RCT studies have investigated physiologic monitor alarms. This lack of RCT studies may be related to the core elements that define this design: random assignment to group; large, heterogeneous samples; and, ideally, use of multiple geographically dispersed sites to ensure that findings are not unique to a particular unit or geographical location. These requirements, albeit costly, are undertaken in effort to increase samples size (and, consequently, increase the power of the statistical analysis) and to promote validity.

Summary of Study Designs

Every design has strengths and limitations, and the investigators' research question ultimately determines the selection of the design (Hulley et al., 2001). The review of the literature has revealed that the various designs used in physiologic monitor alarm research have been useful for elucidating the nature of alarm fatigue, quantifying alarm burden (e.g., unit proportion, alarm rates), and for assessing the effectiveness of relevant interventions. For example, observational studies have been necessary in order to gain understanding of the sources, types, and frequencies of physiologic monitor alarms and thereby provide a scientific basis for designing alarm management interventions. In contrast, quasi-experimental studies have been necessary for performing hospital-based alarm management interventions—given the difficulty of conducting true experiments in "real-world" contexts (Polit & Beck, 2004). Also, use of a quasi-experimental design in studies conducted in patient care areas enables investigators to have a degree of research control when full experimental rigor is not possible (Polit & Beck, 2004).

Accordingly, although no physiologic monitor alarm RCTs have been performed to date, it would behoove the health care and scientific communities to conduct multicenter physiologic monitor alarm RCTs. These studies should be led by experienced nurse scientists who can oversee the studies' execution—in order that research quality and integrity are maintained; the nurse scientists can also help to ensure that effective systems of inter-professional communication are used among multidisciplinary research teams and other study stakeholders at the multiple research sites.

Dimensions of Data Collection Approaches

Numerous data collection methods are used in nursing research, including observation, self-report, interview, and physiologic measurements. Data collection methods for quantitative work are defined in terms of four characteristics: *structure, quantifiability, researcher obtrusiveness*, and *objectivity* (Polit & Beck, 2004). Each dimension is considered, yet the unit of analysis (i.e., monitor alarms) and the research question primarily determine the appropriateness of the approach.

Structure

In qualitative research, data collection most often uses unstructured or semi-structured interviews; in rare instances, qualitative research may use a structured method of data collection. Qualitative research that uses unstructured or semi-structured interviews provides an opportunity for the investigator to discover the informants' feelings, perceptions, and behaviors regarding the study topic (Holloway & Wheeler, 2002). The single qualitative alarm study that used nonparticipant observation and semi-structured interviews for data collection, by Varpio et al. (2012), reported that alarms prompt RNs to regularly consider and interpret patient information—a useful contribution to the literature concerning physiologic monitor alarms.

In contrast, quantitative research relies primarily on a structured plan to collect data. Structured methods of data collection are beneficial in that data is collected with a standardized instrument; this data collection method facilitates subsequent data analysis. The drawback of structured methods is that they provide little opportunity for study participants to offer input or contribute explanations of their responses (Polit & Beck, 2004). The majority of quantitative physiologic monitor alarm studies have been appropriately structured—using variety data collection methodologies—to enable satisfactory data compilation. However, in one study, investigators acknowledged that their research had a substantial methodological flaw that prevented identification of all physiologic monitor alarms and data tabulation (Talley et al., 2011). In addition, this flaw precluded identification of possible relationships between alarms and clinically significant events in a pediatric critical care setting.

Quantifiability

Qualitative data are typically collected in a narrative form; in contrast, quantitative data are gathered using highly structured procedures that enable quantification of variables. For example, Varpio et al. (2012) was able to study the work experience of the pediatric RNs by conducting two cycles of individual, semi-structured interviews. In this investigation, inductive thematic content analysis was used to identify emerging patterns and themes. Quantitative research is used to determine the quantity of an attribute that is present in an object (Nunnally & Bernstein, 1994). This approach facilitates complex statistical analysis (i.e., for descriptive statistics and inferential statistics, Polit & Beck, 2004). In the literature review, 36 studies collected quantitative data on physiologic monitor alarms; these data were described and summarized in statistical analyses.

Researcher Obtrusiveness

Polit and Beck (2004) have suggested that participants' detection of a researcher in the experimental site can cause the participants to modify their behaviors—thereby compromising the quality of collected data. In both qualitative and quantitative studies, the presence of researchers may be problematic when an observed participant is engaged in socially unacceptable behaviors; is non-compliant with medical and nursing standards of care, policies, or procedures; or is striving to perform well in front of others (Polit & Beck, 2004). In the reviewed studies, the researchers did not indicate whether the observers adopted a completely passive role to become unobtrusive bystanders in the patients' room or care setting. Furthermore, in alarm studies that utilized observers for data collection, the staff RNs may have modified their behaviors because of observer presence. The role of participant observation in physiologic monitor alarm studies may have influenced the data collection process and, ultimately, the value of the research.

Objectivity

Polit and Beck (2004) describe *objectivity* in terms of the degree to which two independent observers of a variable of interest obtain similar quantitative findings or perform similar observations—such that the findings or observations are free from the influences of bias and personal emotions. Variability between observers (i.e., deficits in inter-observer objectivity) can occur among multiple members of a research team, and a variety of tactics can be employed to reduce error in order to enhance accuracy and precision. Tactics can include training and certifying the observers, standardizing the measurement methods, automating the instrument, and making measurements unobstrusively (Hulley et al., 2001). Such tactics are usually specified by the principal investigator, who has primary responsibility for overseeing the study (Polit & Beck,

2004). Furthermore, Carthey (2003) emphasizes that whether the observer is a medical professional or a non-medical professional, the observer should have competence in domain knowledge and observational skills in order to make reliable clinical observations.

Data Collection Methods

In physiologic monitor alarm research, data collection has been performed by observers (e.g., RNs, MDs) who have utilized structured instruments or through the use of automated data collection systems—such as manufacturer proprietary software and middleware software in association with video recordings obtained from network surveillance cameras—in order to quantify and to annotate alarms. Physiological monitoring alarm data can be collected using several methods; numerous factors influence the selection of the data collection strategy. Typically, feasibility and cost are major determinants. Other factors that might influence decisions about data collection methods include time pressures, staff availability and expertise, and anticipated methodological burden on both the research team and the participants (Polit & Beck, 2004). Moreover, the availability of various automated data collection strategies (i.e., strategies that use advanced software), support from the hospital, and assistance from experienced clinicians and biomedical engineers are important considerations in selecting a data collection method that will optimize clinical research.

Observation

Observational methods—which are performed directly through the human senses or with the aid of technical equipment—are versatile techniques for collecting and recording data about phenomena in their natural setting (Polit & Beck, 2004). Observational research has been used to gain information about a variety of issues and factors related to physiologic monitor alarms, including incidence, prevalence, frequency, sources, contributing factors, RN response, and

clinicians' behavioral responses to alarms.

Observational methods are well suited to nursing research because evidence of nursing effectiveness can often be obtained through observation (Polit & Beck, 2004). However, observational data collection methods have significant inherent limitations, such as observer bias, the Hawthorne effect, and ethical issues. In alarm research to date, ten studies used direct observation to collect physiologic monitor alarm data; all of these studies reported methodological limitations (Biot et al., 2000; Bitan et al., 2004; Chambrin et al., 1999; Görges et al., 2009; Koski et al., 1990; Lawless, 1994; O'Carroll, 1986; Talley et al., 2011; Tsien & Fackler, 1997; Wiklund et al., 1994).

Observer bias. Hulley et al. (2001) describe *observer bias* as a conscious or unconscious distortion in the perception or reporting of the measurement. A variety of factors can contribute to observer bias: (a) personal interest or commitment may cause observers to see what they want to see—and overlook what they do not want to see; (b) anticipation of what is observed may affect what is observed; and (c) the observer's emotions, values, or prejudices may produce erroneous conclusions (Polit & Beck, 2004). In the data collection plan, researchers make an important decision regarding selecting research personnel who will actually collect the data. Candidate qualification factors include prior research experience, compatibility with the sample characteristics, and professional attributes (Carthey, 2003; Polit & Beck, 2004).

In the structured observational physiologic monitor alarm research examined in this literature review, the data collectors' experiential backgrounds varied significantly. Four studies recruited RNs to serve as dedicated observers (Chambrin et al., 1999; Koski et al., 1990; Talley et al., 2011; Wiklund et al., 1994). Three studies employed observers but did not specify their clinical work experience (Bitan et al., 2004; Görges et al., 2009; Tsien & Fackler; 1997; Whalen

et al., 2013). One other study recruited a physician to be the observer in the patients' room (Biot et al., 2000). Remarkably, the two remaining studies enlisted patient care RNs themselves to assess and record alarm data during performance of their regular clinical duties (Lawless, 1994; O'Carroll, 1986). In this instance, the use of patient care RNs as observers further elevated risk for observer bias—especially because the RNs were providing care to the patients whose condition generated the physiologic monitor alarms. Moreover, given that the RNs were performing data collection during their shifts, the potential for incomplete or missing data is significant.

Observer bias cannot be completely eliminated; however, researchers can implement preventive measures to minimize its impact can be undertaken. For example, to enhance accuracy, investigators can use a variety of strategies such as standardizing and structuring the measurement method, meticulous training (initial and periodic) of observers, and endorsing procedures to establish intra- and inter-reliability measurement (Carthey, 2003; Hulley et al., 2001; Polit & Beck, 2004). In the ten studies that used observers to determine the incidence of physiologic monitor alarms, most investigators omitted or only minimally described the procedures they undertook throughout the study periods to minimize bias and maintain objectivity among the observers. That is, information regarding strategies for increasing accuracy and precision (such as performing intra-rater and inter-rater reliability measures to verify the consistency of observations within and between observers) were minimally provided—a reporting deficit that threatens the validity of the findings.

Hawthorne effect. Just as observer expectations can influence outcomes, so also can a subject's awareness of being observed by observers or by other research personnel. In field experiments, this phenomenon is often referred to as the "Hawthorne effect"; in laboratory

research, this phenomenon is referred to as the "guinea pig effect." Specifically, this effect refers to behavior changes resulting from participants' awareness that they are participating in an experiment and that they are being experimentally observed (Roethlisberger & Dickson, 1939). This behavioral effect reduces the validity of observational research because both patients and health care professionals often behave non-authentically when an observer is present in the experimental setting (Johnson, 1992).

Observation studies that primarily rely on the use of observers for data collection may be susceptible to these expectation- and observation-related biases and hence may be subject to limitations in validity. In the 37 studies reviewed, 10 research teams chose to use the direct observation approach to collect alarm data in accordance with the studies' aims. With the exception of two studies by Talley et al. and Bitan et al., which used a combination of data collection methods (i.e., observers and a manufacturer' propriety software), the remaining studies can be viewed as being early investigations, given their selected methodology.

Ethical issues. Ethical problems related to participant and patient observation may arise in observational health care research (Johnson, 1992). In clinical research situations, moral dilemmas can arise for observers who witness a deviation from proper intervention protocol; such events are problematic both ethically and methodologically (Polit & Beck, 2004). Alarm studies that utilized clinicians (i.e., RNs, MDs) as observers did not describe the role of observers in instances in which a patient's clinical deterioration triggered a life-threatening physiologic monitor alarm. That is, the study reports did not indicate whether in a clinical event warranting immediate attention—and in which an observer was the first clinician on hand—the observer should attend to the patient, or, instead, seek assistance from a member of the patient's health care team.

One might consider whether the need for observer intervention could be precluded by a study's use of personnel who do not have medical backgrounds. According to Slagle, Weinger, Dinh, Brumer, and Williams (2002), this approach could obviate certain ethical considerations because, in Slagle's view, assessments by non-medical observers do not significantly differ from those of medical observers. However, Slagle et al. concede that medical observers are superior to non-medical observers in assessing content-specific characteristics; in contrast Schaefer, Helmreich, and Scheidegger, (1994) report that non-medical observers are superior to medical observers in assessing interpersonal factors in the environment.

Information Technology

Data collection can include the use of an automated instrument in health care research. Because human observers unintentionally differ in the way they perform measurements, the use of information technology (e.g., software) in data collection can reduce variation and, hence, increase precision (Hulley et al., 2001). In physiologic monitor alarm research, investigators have used three information technology (IT) approaches to collect alarm data from physiologic monitors: (a) researcher and vendor-supplied physiological monitoring software, (b) vendor-supplied physiological monitoring software with visual records, and (c) middleware integration software that monitors, captures, and manages alarms from various and different systems.

Researcher and vendor-supplied software. The review of the literature identified 10 quantitative studies that used vendor-supplied software (e.g., GE Healthcare, Masimo, Philips Healthcare) to collect alarm data—such as alarm threshold violations, arrhythmia, and technical alarms. This data collection approach is typically time-consuming, involves manual record keeping (due to limited storage capabilities), and requires the assistance of biomedical engineers. Moreover, some physiologic waveform alarm data and clinician-configured settings (i.e.,

parameter threshold limits) may not be available due to inherent technological constraints. Additional limitations include lack of flexibility in reviewing and managing patient information during transitions in care and over excessive lengths of time required for retrospective review.

Investigators who conduct alarm research often rely on data from a single brand of physiologic monitors that happens to be in use in their clinical research setting. Although a study' exclusive use of software from a single manufacturer provides some uniformity regarding alarm data, some alarm characteristics and other key factors—such as priority levels, behavior, conditions required to generate alarms, and alarm terminology—vary among monitors produced by different manufacturers; these variations can hamper multi-study comparisons. However, combining the alarm data collection approach (that uses vendor-supplied physiological monitoring software) with complementary methods (i.e., observation, video-monitoring, third-party software component) can strengthen the acquisition of alarm data and validation, and, hence, the quality of the research.

Vendor-supplied software with visual records. Observational research can include video recording (via audio-visual equipment such as cameras) in order to enhance physiological observations and minimize bias that otherwise could be introduced by use of human observers. Advantages of video records include (a) capture of details of complex or simultaneous events that might otherwise elude human observers, (b) enhancement of physiological observations, (c) elimination of human observer bias, (d) potential use as s permanent record, and (e) use for review and verification of the accuracy of assessments performed by annotators; in addition, cameras can be easily and unobtrusively ensconced in the environment (Polit & Beck, 2004). On the other hand, video recording is expensive and technically complex. Also, participants who

are being filmed may behave differently than they would otherwise—a manifestation of the Hawthorne effect.

Two prospective observational studies and one pilot study utilized bedside video monitoring–recording (via network surveillance cameras) and manufacturer proprietary software to collect and evaluate physiologic monitor alarm burden (Gross et al., 2011; Siebig et al., 2010a; 2010b). To measure alarm burden on 79 medical–surgical beds on acute care units in a community hospital, Gross et al. (2011) utilized a two-way audio–video telepresence and population management system (eICU VISICU Philips Healthcare) and a bi-directional WMTS telemetry system (IntelliVue Telemetry System, Philips Healthcare). The physiologic data consisted of all monitored waveforms; most of the patient records included data from (a) a single lead of ECG, (b) respiration via the impedance method, and (c) a photoplethysmogram from the oxygen saturation (SpO_2) monitor. In addition, data from all alarms and event conditions were collected and saved to a secure external server.

In contrast, a pilot study and a subsequent research study by Siebig et al. (2010a; 2010b) used a simpler approach. All patients in the 12-bed medical ICU at a university hospital received physiological monitoring (Infinity Patient Monitoring System, Dräger Medical, Lübeck, Germany) and were videotaped using a network camera (Mobotix M10D Dualnight, Kaiserslautern, Germany). The positioning of the camera enabled recording of the patient, surrounding monitoring devices, and actions performed by the nurse at the patient's bedside. Data acquisition from the monitoring system included numerical measurements, physiologic waveforms, occurrences of alarms, alarm settings, and technical messages; these data were stored in dedicated, full-disclosure files via dedicated software (eData, Dräger Medical, Lübeck, Germany).

Gross et al. (2011) and Siebig et al. (2010a; 2010b) used similar methods to collect data on the incidence of physiologic monitor alarms, but the two research teams used different approaches to in classifying alarms. In the Gross et al. study, alarm adjudication was performed by two independent clinical researchers; when the two researchers differed in opinion, a third reviewer's opinion was sought. The Gross et al. article described the alarm classification process but did not discuss the reliability of the classification process. Also, the article did not describe the professional experience of the researchers or any efforts to establish intra-observer and inter-observer reliability. To perform annotation, Siebig et al. recruited a single experienced physician who had approximately 3 years of experience in intensive care medicine. The investigators reported that with over 75% of 200 alarms situations, both physicians classified the alarms identically—indicating a high degree of inter-observer reliability. Both technical validity and clinical relevance were congruent (i.e., at 95% and 85%, respectively; Siebig et al., 2010b). After 12 weeks, intra-observer reliability was assessed in 100 alarm conditions, and an intra-observer difference of 7% was reported; technical validity and clinical relevance were 99% and 97%, respectively.

Hulley et al. (2001) have suggested that researchers can enhance measurement precision by minimizing random variation through training and certifying observers. Providing extensive training to observers confers several benefits: (a) enhanced measurement consistency among multiple observers, (b) the ability to evaluate data plan procedures, and (c) ensured qualification of observers for participation. Although costly and time-consuming, observer training is a necessary in order to standardize data collection, reduce methodological flaws, and enhance overall research quality. In the quantitative alarm studies that utilized clinical observers to collect or annotate physiologic monitor alarms, the researchers provided only minimal descriptions of

how they educated and trained observers to perform their important tasks; this deficit is a limitation because absent use of a high-quality data collection and alarm classification methods, the validity of the studies' findings is compromised.

Integration software. Another method to collect physiologic monitor alarm data involves use of a software developer who is independent of the device manufacturer. The literature review identified two quasi-experimental studies (conducted in the same clinical setting) that used an event-notification management solution (GlobeStar Systems, Connexall) to facilitate export of physiologic monitor alarm data via middleware technology (Cvach et al., 2012; Graham & Cvach, 2010). This approach to agnostic software alarm data collection requires a substantial organizational investment in order to enable information exchange though an interoperability engine designed to integrate communication among medical devices—a methodology not readily available to most researchers and hospitals. Notably, while the breadth of information that can be collected using this platform is substantial, the comprehensiveness of data regarding physiologic monitor alarms is uncertain, because thus far researchers have not provided detailed information regarding how alarm data are retrieved, displayed, or reported for the various types of physiologic monitor alarms. For example, researchers have not reported whether access to associated physiologic waveforms is available for annotation (validation of arrhythmia alarms). The studies that used this event-notification data collection methodology did not report essential alarm data, such as information regarding manipulation of alarm parameter threshold settings, RN responsiveness to alarms, and lower severity physiologic monitor alarm levels (e.g., "message" alarms that only trigger visual alerts).

Most important, studies to date have not indicated whether physiologic monitor alarm burden was defined as the incidence of audible alarms or as the incidence of all physiologic

monitor alarms. A variety of alarms can occur in a single alarm session, but the physiologic monitor will annunciate only the alarm that has the highest priority level (i.e., crisis). Moreover, the two studies that utilized the GlobeStar Systems applications to collect physiologic monitor alarm data did not annotate the arrhythmia alarms to determine the proportion of false-positive cardiac arrhythmia alarms; lack of annotation may have been due to intrinsic limitations of the middleware technology.

BedMasterEx. BedMasterEx (Excel-Medical Electronics) is another vendor-agnostic software application that can be used to collect physiologic monitor alarm data from patient monitoring systems and ancillary medical devices. This Microsoft Windows-based software application offers high-resolution video capability for viewing patients with their time-aligned physiological data and streaming waveforms. BedMasterEx can generate shift reports that have a 7-s ECG rhythm strip of the top two waveforms channels, alarms in a selected time interval, and ECG measurements that can be integrated into the patient's electronic medical record. Furthermore, the use of BedMasterEx's paperless shift strip reports reduces costs (through reduced waste and employee time), increases accessibility and security of ECG rhythm strips, and reduces the occurrence of missing or fading rhythm strips. The BedMasterEx software has powerful features: it collects and records (a) an unlimited number of vital sign measurements, (b) alarms, and (c) physiologic waveform data (125 samples per second for Philips, 240 samples per second for General Electric) from monitored patients in a networked environment; these data can be displayed in near real-time and stored indefinitely.

Vital signs. Vitals sign data are collected in a tabular format and can be gathered at a variety of collection intervals: 5 s, 15 s, 30 s, 1 min, 5 min, 30 min, and 1 hr; also, these patient data can be collected until monitoring is discontinued (Excel-Medical Electronics, 2012).

Patients' vital sign trends can also be calculated using a graphical representation with vital sign data (i.e., heart rate, respiration rate) plotted on various time scales; all graphs auto-update when new data are acquired. In addition, all vital signs and graphic trends can be exported to software applications such as Microsoft Excel to XML for further analysis and/or review by hospital administrators and researchers.

Alarms. Data on every physiologic monitor alarm (arrhythmia, parameter threshold violation, or technical alarm) are acquired and stored along with the related waveform strip for indefinite review and export. Alarms can be sorted by priority notification level (i.e., advisory, warning, crisis, system warning) or chronologically, and calipers are provided for interval measurements. When the physiologic monitor generates an alarm, the original alarm is stored, and the database is updated at each 2-s interval. Alarm data are stored in the structured query language (SQL) database—a special-purpose programming language designed for managing large data in relational databases; storage of data for a single alarm requires approximately 100 bytes. Reports generated from these data can include every alarm generated; these reports can be produced even after the patient is discharged from the monitor or health care organization for indefinite review and analysis**.**

Physiological waveforms. BedMasterEx's unique feature is its ability to collect, store, and navigate waveform data from physiologic monitors—capabilities that sets this automated data collection instrument apart from other software applications. When annotating cardiac arrhythmia alarms, inclusion of physiologic waveform data is necessary in order to determine whether an alarm is true-positive or false-positive. Arrhythmia alarms (i.e., asystole, VTach/VFib) cannot be validated without the ability to review the patient's physiological waveforms (including waveforms from multiple leads —i.e., I, II, III and V), invasive pressures,

respiration, and oxygen saturation in a time-synchronized data series (Excel-Medical Electronics, 2012).

BedMasterEx is designed for research, hospital quality assurance, and administrative oversight purposes—specifically, for examining physiological conditions leading to sentinel events and for performing root-cause analysis (Excel-Medical Electronics, 2012). This application is useful for physiologic monitor alarm research because it has a robust analytics platform, an IBM Streams and Matlab Interface, and SQL database query tools. In addition, all of the functions of BedMasterEx can be accessed remotely from any networked personal computer or via a "thin client" (i.e., a computer or computer program that relies on another computer—i.e., the server) at a clinician's home or office—including near real-time, historical trends and historical full-disclosure data.

BedMasterEx contains the required security specifications to protect the privacy of a patient's health information according to the Health Insurance Portability and Accountability Act (HIPAA). In this regard, the application includes features that enable administrative oversight of users through system settings (such as logons determined by user privileges, group, and profile), server disk space, and database usage. Given the flexibility, granularity, and amount of data provided by this automated instrument, it is not surprising that over 70% of the 2010–2011 *US News & World Report* (USNWR) Best Hospitals and 75% of the USNWR Best Children's Hospitals utilize BedMasterEx in their institutions to collect physiologic monitor alarm data (Excel-Medial Electronics, 2012).

Methodological Limitations

Currently, acquiring comprehensive alarm information—or even partial detailed alarm information—from bedside physiologic monitors or central monitoring stations is impracticable.

As discussed above, these limitations have motivated researchers to adopt a variety of methodological approaches, including direct observation, visual records, and use of information technology—for obtaining information regarding unit alarm histories beyond the applications of the operating system. However, most of these traditional data collection approaches do not provide alarm data regarding the manipulation of alarm parameter threshold settings, clinicians' responsiveness to alarms, and information on lower severity alarm levels—(such as message alarms and other low-priority alarms) that may only present visual alerts—and are necessary in order to fully understand physiologic monitor alarm fatigue.

Researcher efforts have also been stymied by a lack of operational definitions of measurement concepts pertaining to physiologic monitor alarms. This lack has arisen from manufacturers' use of diverse proprietary alarm labels and definitions. In the studies considered in this review, few investigators provided descriptions of the physiologic monitor alarms being assessed and evaluated. With the profusion of patient monitoring systems utilized by health care providers, it is unclear whether studies are classifying arrhythmia alarms in a standardized manner or in reference to internationally accepted societal guidelines. For example, the confounding of data on different types of alarms was observed in two studies: Görges et al.'s study (2009), which merged data on cardiac arrhythmia and heart rate alarms, and Gross et al.'s study (2011), which combined data on tachycardia and VTach/VFib arrhythmia alarms. This lack of standardization casts doubt on the validity of the findings. Such practices affect estimates of the incidence of physiologic monitor alarms and of alarm burden, thereby undermining the integrity of the results, are important weaknesses in the studies.

Inconsistent definition, calculations, and reports of alarm burden in published studies have resulted in erroneous results and conclusions. For example, a number of investigators have

utilized the unit of measurement *mean alarms/24 hr* to define alarm burden; however, this convention does not take into account the unit size or patient census, or patient factors (e.g., acuity, monitoring time) that affects the generation of physiologic monitor alarms. Moreover, some studies use the unit of measurement *mean alarms/patient/day*—which is better than mean alarms/24 hr, because "mean alarms/patient/day" considers patients; however, some investigators fail to indicate whether, in reporting their findings, they are utilizing the average daily census (ADC) or all patient activity within the unit. Utilizing the ADC can exaggerate alarm burden, because typically the number of patients hospitalized in a unit is greater than the unit's officially stated bed capacity. (This difference between ADC and bed capacity is due to patient admission, transfer, or discharge activity that is not recorded in the midnight census.)

Recent studies better report alarm burden by adjusting for the duration of physiologic monitoring per patient or by normalizing the data to a frequency of alarms/100 hr of device monitoring. Both of these approaches are much more complex than is the unit of measurement "mean alarms/patient/day" or "mean alarms/24 hr" and require detailed information and bioengineering support. Often, these data and expertise are not readily available to researchers, and the reporting of alarm burden post-implementation of alarm reduction efforts are over-simplified, which can lead to misunderstanding of the study findings. The diversity of units of measurement used in alarm burden assessments presents challenges when comparatively evaluating multiple research studies.

In the future, investigators might consider adjusting for the number of patients and actual monitored hours in a unit during a study. Also, because the majority of alarms are associated with a minority of patients, the alarm data is skewed by outliers (Gross et al., 2011) or, given that many patients do not trigger certain critical arrhythmia alarms (e.g., ventricular

fibrillation/tachycardia), the type and quantity of alarms generated by each patient vary significantly.

Furthermore, some descriptive observational studies have included data on both alarm signals originating from physiologic monitors and alarm signals emanating from other clinical devices (of which ventilators are the main contributors of such signals). The inclusion of alerts and notifications from a variety of clinical devices is at times confusing. To avoid confusion, physiologic monitor alarm data and clinical alarm data should be reported separately or, better yet, individual studies may be warranted, given the importance, complexity, and hazards of the various types of medical device alarms.

The cross-sectional studies also have substantial limitations. The five studies that explored clinicians' perspectives on alarms each developed their individual self-administered questionnaires. Investigators' analyses of their data on clinicians' perceptions of and practices regarding alarms are important contributions to the nascent study of alarm fatigue; however, as noted earlier; a valid and reliable survey instrument has yet to be developed. Among published alarm studies, the only study that has mentioned validity is the 2011 HTF re-survey; (in this regard, the HTF report stated that the survey questionnaire's content was validated by an expert panel). The lack of a valid and reliable survey instrument is a significant shortcoming that cannot be overlooked. Development of a credible instrument would enable comparative studies in various populations and settings.

Few intervention studies have investigated the clinical effectiveness of alarm strategies for reducing physiologic monitor alarm burden. Notably, no studies conducted in an ICU setting have reported patient outcomes related to implementation of strategies aimed at reducing alarm fatigue. Except for the work by Taenzer et al. (2010) and Whalen et al. (2013), most intervention

studies have not reported patient or organizational outcomes. The omission of patient data related to possible delays in care or to negative patient experience as a result of modifications in unit alarm settings is a significant deficit in the research literature.

Design and Instrument Selection

Current studies that evaluate the effectiveness of select interventions—clinical practice or modification of alarm features—utilize a quasi-experimental study design. This design is preferred by some in light of the numerous challenges inherent in performing true experiments in the clinical setting. To answer the over-arching research question, "Is the difference in mean hourly alarm rates between Assessment 1 and Assessment 2 different for the experimental unit compared with the control unit?" a randomized clinical trial design was determined to be the best methodological approach for assessing and evaluating an alarm management intervention focused on reducing physiologic monitor alarms in an intensive care unit. This method allows for the random allocation of one unit to the experimental group, while the second unit will serve as a control—whose performance data will be useful for the evaluation of outcomes in the group of primary interest, the intervention group (Polit & Beck, 2004).

In addition, a variety of data collection instruments have been used to determine the prevalence and validate the accuracy of physiologic monitor alarms, which makes the interpretation of results among studies difficult. Each of these instruments has proven useful; however, technology evolves at a rapid pace—and a device-integration software application, BedMasterEx, has been assessed as being superior and was selected as the instrument for data collection. Specifically, BedMasterEx was considered the instrument of choice to determine and annotate the prevalence of total and audible physiologic monitor alarm data because of the strengths of this automated data collection application. The dimensions of this application are

highly structured, quantifiable, and objective, and entail minimal researcher obtrusiveness. Furthermore, the use of BedMasterEx software application precluded challenges associated with observer bias, Hawthorne effect, and ethical issues. This clinical software application has a myriad of technological and innovative features that simplify the comprehensive task of physiologic monitor alarm collection and annotation. In this regard, also, BedMasterEx is unmatched by other methodologies—including other physiologic monitor commercial software applications and video monitoring.

Contextual Issues

Even with the use of an automated instrument such as BedMasterEx, important contextual issues must be considered when preparing to perform a randomized clinical trial to investigate a physiologic monitor alarm management intervention in an intensive care unit. One such issue is the multiplicity of internal departmental processes that are beyond the control of the investigator and that may influence the study's methodological merits.

Compensatory equalization and rivalry. Schumacher et al. (1994) described *compensatory equalization* as an effect in which equalization may occur among study groups—and involves one group procuring resources from the intervention or treatment group. This threat to validity can negatively affect the planned contrast among the two groups—the intervention group and the control (usual care) group. Compensation equalization is a risk that must be considered when performing an alarm management intervention in one of two neuroscience intensive care units—especially because RNs and ancillary staff members can be assigned to work on either the experimental unit or the control unit.

When innovations such as novel patient monitoring supplies (i.e., skin prep paper and new ECG electrodes) are introduced in a clinical environment, they can naturally generate excitement and motivation among staff. These subjective reactions that can contribute to a study's success, especially if the area of research has previously received little attention—or, conversely, it can promote rivalry among units (Shadish et al., 2002). If a member of one group perceives that she or he is being mistreated or disregarded, this perception may give rise to an inherent urge among individuals to somehow equalize the performance among the study units by modifying clinical conditions so that they are "fairer"—which may involve the displacement of supplies (dedicated to the experimental unit) and the application of knowledge beyond the standardized education received in order for both units to have an equal chance to succeed.

Saretsky (1972) posits that the public knowledge and assignment of groups to the intervention and control or comparison group can instigate competition. This rivalry can manifest as an attempt on the part of the comparison group to outperform the intervention group by working harder to overcome the disadvantage of being in the comparison group—and deprived of the benefits and resources accorded to the intervention group.

This reactive human behavior is prevalent in sociological research and has been called the "John Henry effect," named after the legendary (and possibly mythical) American steel worker who worked so hard to outperform a steam powered hammer that he eventually died of over-exertion. The John Henry effect, a particular form of Hawthorne effect, occurs when the control group participants alter their behavior out of awareness that they are in the control group.

Acknowledging that compensatory equalization and rivalry can occur among participants in study units is important, because, if these issues do manifest in the context of a study, they can

threaten construct validity. In such an event, it may be necessary to address perceived inequities through thoughtful discussions with participants.

Diverse participants and patient placement. Another consideration that is outside the investigator's sphere of influence is the allocation of RNs and ancillary staff members (i.e., patient care assistants) assigned to work on the study units. This issue is of a concern regardless of whether the intervention or comparison unit is their "home" unit or whether they are temporarily assigned (i.e., "float") to one of the two units for a shift. Given the dynamic nature of staffing in intensive care units, nurses from other specialty units are often assigned to work in both the intervention and comparison units each day—and sometimes within a shift—in order to ensure appropriate coverage due to unanticipated absences (i.e., sick calls), increased patient acuity, and fluctuations in unit census. Having a variety of staff participate in a nursing intervention—albeit once a day with no advanced preparation—will require daily oversight to ensure adherence to the intervention and related procedures.

Furthermore, because of operational needs, patients can be placed on different units throughout their hospitalization (including at times of high patient census) in order to efficiently manage patient flow and ensure patient safety. Accordingly, patients may be transferred to several patient care units (including ICUs) throughout their stay. Such transfers may be necessary because of any of a variety of circumstances, such as (a) when specialty patient volume exceeds bed capacity on a unit; (b) when a patient requires subspecialty care or clinical expertise that is best provided on another nursing unit (i.e., continuous renal replacement therapy); or (c) when the patient's specific and/or immediate needs (e.g., negative pressure isolation, hemodialysis) warrant transfer to a room that meets patient care requirements. Such transfers may involve patient relocation to an altogether different unit.

Conclusion

Physiologic monitor alarm fatigue is a complex, inadequately studied phenomenon that affects many clinicians who care for patients in the technology-rich intensive care environment. The majority of published research studies that have focused on physiologic monitor alarms have been observational, and many of these studies have been fraught with methodological limitations that affect the validity and reliability of the studies' findings. Evaluation of the clinical effectiveness of nursing interventions aimed at reducing the frequency and rates of non-actionable physiologic monitor alarms is a burgeoning area of research. In designing such research, investigators have begun to recognize the value of quasi-experimental studies. While the quasi-experimental design has limitations, the advantages of this approach are considerable. Chiefly, quasi-experiments have enabled researchers to investigate and measure outcomes associated with various clinical interventions; quasi-experiments also permit flexibility in selection and sampling that may be necessary for conducting scientific research in challenging health care settings. Until certain clinical barriers can be rectified or surmounted by researchers, this design can be considered satisfactory alternative.

However, it is without question that an RCT study design is generally accepted as being the most powerful and rigorous approach. Our study will be the first to apply this design in the physiologic monitor alarm research in the health care setting and report alarm burden as the mean hourly alarm rates for alarms based on monitoring times. Furthermore, quantitative physiologic monitor alarm studies can produce rigorous research if the quality of the instrument selected for data collection is valid and reliable. Assessment of the BedMasterEx software application for the collection of physiologic monitor alarms indicates that this software enables

accurate, objective, and precise measurements; at the same time, use of this application obviates challenges associated with direct observation and vendor-supplied software.

Of late, advancements in information technology such as the introduction of specialized software now provide researchers with the tools necessary for comprehensive analysis of physiologic monitor alarms. Studies that adopt these innovations will be able to capture vast amounts of information—not limited to simply data on the incidence of all alarms. Notably, BedMasterEx facilitates the review of cardiac arrhythmias and their ensuing alarms; more important, this software application facilitates annotation of physiologic waveforms and determination of alarm accuracy. The use of information technology in health care research can provide a wealth of alarm information that has previously been unavailable to researchers. As a result, future studies that leverage the advantages of applied technology will elevate the quality of published alarm research.

References

Aboukhalil, A., Nielsen L., Saeed, M., Mark, R. G., & Clifford, G. D. (2008). Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform. *Journal of Biomedical Informatics, 41,* 442–451.

Biot, L., Carry, P. Y., Perdrix, J. P., Eberhard, A., & Baconnier, P. (2000). Évaluation clinique de la pertinence des alarmes en reanimation. *Annales Françaises d'Anesthésie et de Réanimation, 19,* 459–466.

Bitan, Y., Meyer J., Shinar, D., & Zmora, E. (2004). Nurses' reactions to alarms in a neonatal intensive care unit. *Cognition, Technology & Work, 6,* 239–246.

Block, F. E., Jr., Nuutinen, L., & Ballast, B. (1999). Optimization of alarms: A study of alarm limits, alarm sounds, and false alarms, intended to reduce annoyance. *Journal of Clinical Monitoring and Computing, 15,* 75–83.

Blum, J. M., Kruger, G. H., Sanders, K. L., Gutierrez, J., & Rosenberg, A. L. (2009). Specificity improvement for network distributed physiologic alarms based on a simple deterministic reactive intelligent agent in the critical care environment. *Journal of Clinical Monitoring and Computing, 23,* 21–30.

Burgess, L. P. A., Herdman, T. H., Berg, B.W., Feaster, W. W., & Hebsur, S. (2009). Alarm limit settings for early warning systems to identify at-risk patients. *Journal of Advanced Nursing, 65*(9), 1844–1852.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: RandMcNally.

Carlson, M. D. A., & Morrison, R. S. (2009). Study design, precision, and validity in observational studies. *Journal of Palliative Medicine, 12*(1), 77–82.

Carthey, J. (2003). The role of structured observational research in health care. *Quality and Safety in Health Care, 12,* Suppl II, ii13–ii16.

Chambrin, M.-C., Ravaux, P., Calvelo-Aros, D., Jaborska, A., Chopin, C., & Boniface, B. (1999). Multicentric study of monitoring alarms in the adult intensive care unit (ICU): A descriptive analysis. *Intensive Care Medicine, 25,* 1360–1366.

Crombie, I. K. (1996). *The pocket guide to critical appraisal.* London, UK: BMG Publishing Group.

Cropp, A. J., Woods, L. A., Raney, D., & Bredle, D. L. (1994). Name that tone: The proliferation of alarms in the intensive care unit. *Chest, 105*(4), 1217–1220.

Cvach, M., Biggs, M., Rothwell, K. J., Charles-Hudson, C. (2012). Daily electrode change and effect on cardiac monitor alarms: An evidence based practice approach. *Journal of Nursing Care Quality, 28*(3), 265–271.

Daly, B. J. (1983). The effect of fatigue on the vigilance of nurses monitoring electrocardiograms. *Heart and Lung*, *12(*4), 384–388.

Emergency Care Research Institute (2010). Top 10 technology hazards for 2011. *Health Devices, 39*(11), 404–416.

Emergency Care Research Institute (2011). Top 10 technology hazards for 2012.Retrieved from https://www.ecri.org/Documents/Secure/Health_Devices _Top_10_Hazards_2012.pdf

Emergency Care Research Institute (2012). Top 10 technology hazards for 2013. Retrieved from https://www.ecri.org/Documents/Secure/Health_Devices _Top_10_Hazards_2013.pdf

Excel-Medical Electronics (2012). BedMasterEx Operators Manual (Version 4.1) p.1-131.

Graham, K., & Cvach, M. (2010). Monitor alarm fatigue: Standardizing use of physiological monitors and decreasing nuisance alarms. *American Journal of Critical Care, 19*(1), 28–34.

Görges, M., Markewitz, B. A., & Westenskow, D. R. (2009). Improving alarm performance in the medical intensive care unit using delays and clinical context. *Anesthesia & Analgesia, 108*(5), 1546–1552.

Gross, B., Dahl, D., & Nielsen, L., (2011). Physiologic monitoring alarm load on medical/surgical floors of a community hospital. *Biomedical Instrumentation & Technology, 45*(1), 29–36.

Healthcare Technology Foundation, (2011). National Clinical Alarm Survey: Perceptions, issues, improvements, and priorities of healthcare professionals. Retrieved from http://www.thehtf.org

Holloway, I., & Wheeler, S. (2002). *Qualitative research in nursing* (2nd ed.).Oxford, UK: Blackwell Science.

Hulley, S. B., Cummings, S. R., Browner, W. S., Grady, D., Hearst, N., & Newman, T. B. (2001). *Designing clinical research* (2nd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Jacobs, J. L., & Eron, L. (2007). Automated medical vigilance on the medical/surgical ward of an acute care hospital. *Journal of Hospital Medicine, 2*(3), 196–197.

Johnson, M. (1992). A silent conspiracy?: Some ethical issues of participant observation in nursing research. *International Journal of Nursing Studies, 29*(2), 213–223.

Korniewicz, D. M., Clark, T., & David, Y. (2008). A national online survey on the effectiveness of clinical alarms. *American Journal of Critical Care, 17*(1), 36–44.

Koski, E. M. J., Mäkivirta, A., Sukuvaara, T., & Kari, A. (1990). Frequency and reliability of alarms in the monitoring of cardiac postoperative patients. *International Journal of Clinical Monitoring and Computing, 7*(2), 129–133.

Koski, E. M. J., Mäkivirta, A., Sukuvaara, T., & Kari, A. (1995). Clinicians' opinions on alarm limits and urgency of therapeutic responses. *International Journal of Clinical Monitoring and Computing, 12*, 85–88.

Lawless, S. T. (1994). Crying wolf: False alarms in a pediatric intensive care unit. *Critical Care Medicine, 22*, 981–985.

Mäkivirta, A., & Koski, E. M. J. (1994). Alarm-inducing variability in cardiac postoperative data and the effects of prealarm delay. *Journal of Clinical Monitoring, 10*(3), 153–162.

Mäkivirta, A., Koski, E. M. J., Kari, A., & Sukuvaara, T. (1991). The median filter as a preprocessor for a patient monitor limit alarm system in intensive care. *Computer Methods and Programs in Biomedicine, 34*, 139–144.

Mammone, T. (2013). *Physiologic monitor alarm burden* (Unpublished doctoral dissertation). University of California, San Francisco.

Mann, C. J. (2003). Observational research methods. Research design II: cohort, cross sectional, and case-control studies. *Emergency Medicine Journal, 20*, 54–60.

Nightingale, F. (2010). *Notes on hospitals.* Charleston, SC: Nabu Press.

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

O'Carroll, T. M. (1986). Survey of alarms in an intensive therapy unit. *Anaesthesia, 41*, 742–744.

Pan, P., & Gravenstein, N. (1994). Intraoperative Pulse Oximetry: Frequency and distribution of discrepant data. *Journal of Clinical Anesthesia, 6,* 491–495..

Polit, D. F., & Beck, C. (2004). *Nursing research: Principles and methods* (7th ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Rheineck-Leyssius, A. T., & Kalkman, C. J. (1998). Influence of pulse oximeter settings on the frequency of alarms and detection of hypoxemia. *Journal of Clinical Monitoring and Computing, 14*(3), 151–156.

Roethlisberger, F. S., & Dickson, W. J. (1939). *Management and the worker.* Cambridge, MA: Havard University Press.

Saretsky, G. (1972). The OEO PC experiment and the John Henry effect. *Phi Delta Kappan, 53,* 579-581.

Schaefer, H. G., Helmreich, R. L., & Scheidegger, D. (1994). Human factors and safety in emergency medicine. *Resuscitation, 8*(3), 221–225.

Schmid, F., Goepfert, M. S., Kuhnt, D., Eichhorn, V., Diedrichs, S., Reichenspurner, H., . . . Reuter, D. A. (2011). The wolf is crying in the operating room: Patient monitor and anesthesia workstation alarming patterns during cardiac surgery. *Anesthesia and Analgesia, 12*(1), 78–83.

Schumacher, J.E., Milby, J. B., Raczynski, J. M., Engle, M., Caldwell, E. S., & Carr, J. A. (1994). Demoralization and threats to validity in Birmingham's homeless project. In K. J. Conrad (Ed.), *Critically Evaluating the Role of Experiments* (pp. 41-44). San Francisco, CA: Jossey-Bass.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* New York, NY: Houghton Mifflin Company.

Siebig, S., Kuhls, S., Imhoff, M., Gather, U., Schölmerich, J., & Wrede, C. (2010a). Intensive care unit alarms—How many do we need? *Critical Care Medicine, 38*(2), 451–456.

Siebig, S., Kuhls, S., Imhoff, M., Langgartner, J., Reng, M., Schölmerich, J., . . . Wrede, C. (2010b). Collection of annotated data in a clinical validation study of alarm algorithms in intensive care: A methodological framework. *Journal of Critical Care, 25*, 128–135.

Siebig, S., Sieben, W., Kollmann, F., Imhoff, M., Bruennler, T., Rockmann, F., . . . Wrede, C. (2009). Users' opinions on intensive care unit alarms: A survey of German intensive care units. *Anaesthesia and Intensive Care, 37*(1), 112–116.

Slagle, J., Weinger, M. B., Dinh, M. T., Brumer, V. V., & Williams, K. (2002). Assessment of the intrarater and interrater reliability of an established clinical task analysis methodology. *Anesthesiology, 96,* 1129–1139.

Solsona, J. F., Altaba, C., Maúll, E., Rodríguez, L., Bosqué, C., & Mulero, A. (2001). Are auditory warnings in the intensive care unit properly adjusted? *Journal of Advanced Nursing*, *35*(3), 402–406.

Taenzer, A. H., Pyke, J. B., McGrath, S. P., & Blike, G. T. (2010). Impact of pulse oximetry surveillance on rescue events and intensive care unit transfers. *Anesthesiology*, *112*(2) 282–287.

Talley, L. B., Hooper, J., Jacobs, B., Guzzetta, C., McCarter, R., Sill, A., . . . Wilson, S. L. (2011). Cardiopulmonary monitors and clinically significant events in critically ill children. *Biomedical Instrumentation & Technology, 45*, 38–45.

Tsien, C. L., & Fackler J. C. (1997). Poor prognosis for existing monitors in the intensive care unit. *Critical Care Medicine, 25* (4), 614–619.

Varpio, L., Kuziemsky, C., MacDonald, C., & King, W. J. (2012). The helpful or hindering effects of in-hospital patient monitor alarms on nurses. *Computers, Informatics, Nursing, 30*(4), 210–217.

Welch, J. (2011). An evidence-based approach to reduce nuisance alarms and alarm fatigue. [Supplemental Material]. *Biomedical Instrumentation & Technology*, Spring, 46–52.

Whalen, A. A., Covelle, P.M., Piepenbrink. J. C., Villanova, K.L., Cuneo, C. L., & Awtry, E.H. (2013). Novel approach to cardiac alarm management on telemetry units. *Journal of Cardiovascular Nursing*, 1-10.

Wiklund, L., Hök, B., Ståhl, K., & Jordeby-Jönsson, A. (1994). Postanesthesia monitoring revisited: Frequency of true and false alarms from different monitoring devices. *Journal of Clinical Anesthesia*, *6*, 182–188.

Zong, W., Moody, G. B., & Mark, R. G. (2004). Reduction of false arterial blood pressure alarms using signal quality assessment and relationships between electrocardiogram and arterial blood pressure. *Medical and Biological Engineering and Computing, 42*, 698–706.

Chapter 4

Methods

This study used a randomized clinical trial study design, and, for data collection and analysis, a quantitative approach. The intervention for this study had two components: modifying unit default SpO_2 alarm settings and using a novel skin preparation technique with the daily application of high-quality ECG electrodes.

Setting. The study was conducted at the University of California, San Francisco (UCSF) Medical Center in San Francisco, California. UCSF Medical Center is a large tertiary–quaternary care center that provides both acute care and ambulatory care services to a multicultural patient population. The study was performed in two neuroscience ICUs—"11 NICU" (16 beds) and "8NICU" (13 beds)—that were similar in size, patient acuity, and staffing. Participating patients were randomly assigned to either an experimental group or to a control group. For both groups, alarm data were collected for two 31-day periods (i.e., Assessment 1[baseline] and Assessment 2).

Sample. The 8 NICU and 11 NICU patient population is comprised of a diverse adult patient population ranging in age from 18 years to 100 years; this population represents diverse cultural and ethnic backgrounds. Patients who require intensive medical–surgical care or nursing interventions are admitted to 8 NICU and 11 NICU in accordance with written admission criteria. Such patients include those with specific invasive line (including arterial, intracranial pressure [ICP] monitoring) and patients who require tracheal intubation or continuous assessment and management.

Patients. Neuro-critically ill patients cared for in the 8 NICU and 11 NICU include but are not limited to patients with respiratory failure, hemodynamic instability, vasospasm

requiring aggressive treatment, acute stroke, Guillain-Barre, craniotomies for tumors, aneurysms, arteriovenous malformations, subarachnoid hemorrhages, spinal disorders, endovascular treatment of aneurysms, carotid and vertebral stenosis; patients requiring external ventricular drainage (EVD); and patients who require frequent assessment of neurological status. Patients with *Do Not Resuscitate* (DNR) orders may be admitted to the 8 NICU and 11 NICU on the basis of their clinical needs and plan of care. Treatments and nursing procedures include wound and skin care, diabetic management, intravenous therapy, anticoagulation, care of chest tubes, pain management, care of invasive lines, and management of patients with an EVD.

All physiological data (i.e., alarms and waveforms) from patients admitted to the 8 NICU and 11 NICU were obtained from the Solar 8000i physiologic monitors during the assessment periods and were included in the analysis; the samples had no exclusion criteria. The study could not report the true prevalence of NICU physiologic monitor alarm rates without a waiver of patient consent because obtaining every NICU patients' consent was not practical (i.e. either patients are critically ill, sedated, and/or comatose or surrogate consent is difficult to obtain on admission).

Registered nurses. The 8 NICU and 11 NICU are staffed for nurse-to-patient ratios of 1:1 and 1:2. The 8 NICU and 11 NICU charge nurses assess staffing for each 12-hr shift based on an established acuity system (i.e., Clairvia®) that is specific to unit-based patient populations and patient care requirements. Other ICU RNs and PCAs are assigned to 8 NICU and 11 NICU on an ad hoc basis (i.e., in "float" assignment) to meet departmental staffing needs. Also, PCAs assist in provision of care on all shifts.

Timeline. The timeframe from study initiation to completion was greater than 1 year. A preliminary analyses was performed in June 2012 for 2 weeks to gain a better understanding of

the prevalence of physiologic monitor alarm in the study units (see Table 4.1). Nine months later, alarm data for both the experimental unit and the control unit during Assessment 1 (pre-intervention baseline) were collected in March 2013 (i.e., for 31 days). The collection of alarm data for Assessment 2 in both units occurred in August 2013 (i.e., for 31 days; see Table 4.2). Performing the intervention in August provided time for (a) acquiring the Solar 8000i software upgrades (Version 5.5); (b) coordinating the upgrade process with unit nursing leadership and biomedical engineering; this process required that each physiologic monitor be shut down for a period of 5 min; this monitor shut down necessitated patients' being placed on a transport monitor during that time period, (c) modifying the experimental unit's default SpO_2 alarm low-limit threshold setting; (d) training the RNs on the new SpO_2 low-limit threshold alarm setting (i.e., less than 88%) with a SpO_2 alarm delay (i.e., 15 s); (e) training RNs and patient care assistants (PCAs) regarding the novel skin preparation technique and daily application of Ag/AgCl-foam, pre-gelled *wet* ECG electrode, and (f) educating the nursing staff regarding proper storage of electrodes (i.e., electrodes should be kept in their sealed packages until ready for patient use).

Table 4.1. Physiologic Monitor Alarms by Alarm Type (Pilot Data)

Physiologic Monitor Alarms	8 NICU (ADC 9) June 1-14, 2012 (*N* = 66,597)	11 NICU (ADC 14) June 1-14, 2012 (*N*= 80,430)
Parameters	63,562 (95%)	76,339 (95%)
Technical	1,078 (2%)	1,908 (2%)
Arrhythmia	1,957 (3%)	2,183 (3%)

Table 4.2. Alarm Study Timeline

Month/Year	Study
June 2012	1. 8 and 11 NICU Pilot (2 weeks)
March/April 2013	1. Annotate alarms collected continuously from all patient monitors (8 NICU; 13 monitors; 11 NICU;16 monitors) for 31 days using BedMasterEx (Assessment 1) 2. Data collection/entry of patient variables from Epic 3. Data collection/entry for Code Blue events and RN assignments
May/June/July 2013	1. Prepare for intervention
August 2013	1. Intervention on 11 NICU begins for 31 days
September/October/November 2013	1. Annotate alarms collected continuously from all patient monitors (8NICU; 13 monitors: 11 NICU:16 monitors) for 31 days using BedMasterEx (Assessment 2) 2. Data collection/entry of patient variables from Epic 3. Data collection/entry for Code Blue events and RN assignments
December 2013	1. Final data entry 2. Final annotation 3. Data cleaning (entry of all missing data)
January/February/March 2014	1. Statistical analysis
April/May/June/July 2014	1. Prepare for publications to selected journals

Human Subjects Assurance

A study application was submitted to the Human Research Protection Program Committee on Human Research, and a notification of expedited review approval (IRB# 12-09927, Reference # 056096) was received on November 21, 2012. The committee of record was the San Francisco General Hospital panel, which determined that the study risk assignment would be minimal.

Confidentiality. Preservation of data confidentiality was the investigator's major concern. All information was collected in accordance with provisions of the Health Information Portability and Accountability Act (HIPPA). Data from the GE patient monitoring system were downloaded to a dedicated encrypted server. Access to both BedMasterEx and the EMR remains

password protected and auditable. Analysis of alarm data using the BedMasterEx software was done on an encrypted UCSF Medical Center computer. Since all the data included in this study were routinely collected in the course of patient care activities during hospitalization, the analysis of these data for research purposes presented similar risks for the patients admitted to UCSF Medical Center.

Data Safety Monitoring Board. A Data Safety Monitoring Board was in place for the length of the study. The Board was comprised clinical experts who could prospectively review patient outcomes and apply pre-determined stopping rules to halt the study if patient harm was observed. The principal investigator (PI) was responsible for monitoring adverse events for every patient admitted to the study units throughout the course of the project. The principal investigator collaborated closely with senior nurse scientists and advisors and ensured that all safety measures were in effect during the study period. No patient safety issues were identified during the study.

Risks to Human Subjects. The risk to patients was determined to be minimal, given that modification of monitor alarms and use of patient monitoring supplies is part of routine care for the hospitalized adult receiving continuous physiological monitoring. Moreover, the modifications of the experimental units' default SpO_2 alarms settings (i.e., low-limit threshold of 88% and 15-s alarm delay) were already implemented in all non-ICU units utilizing the Masimo Patient SafetyNet Surveillance System® throughout UCSF Medical Center; to date, no adverse patient outcomes have been reported or documented associated with the above default alarm settings.

Intervention: SpO_2 Low-limit Threshold Alarm and SpO_2 Alarm Delay

For this study, the intervention for reducing non-actionable SpO_2 alarms involved modification of the experimental unit's default SpO_2 low-limit threshold setting and the activation of a SpO_2 alarm delay feature.

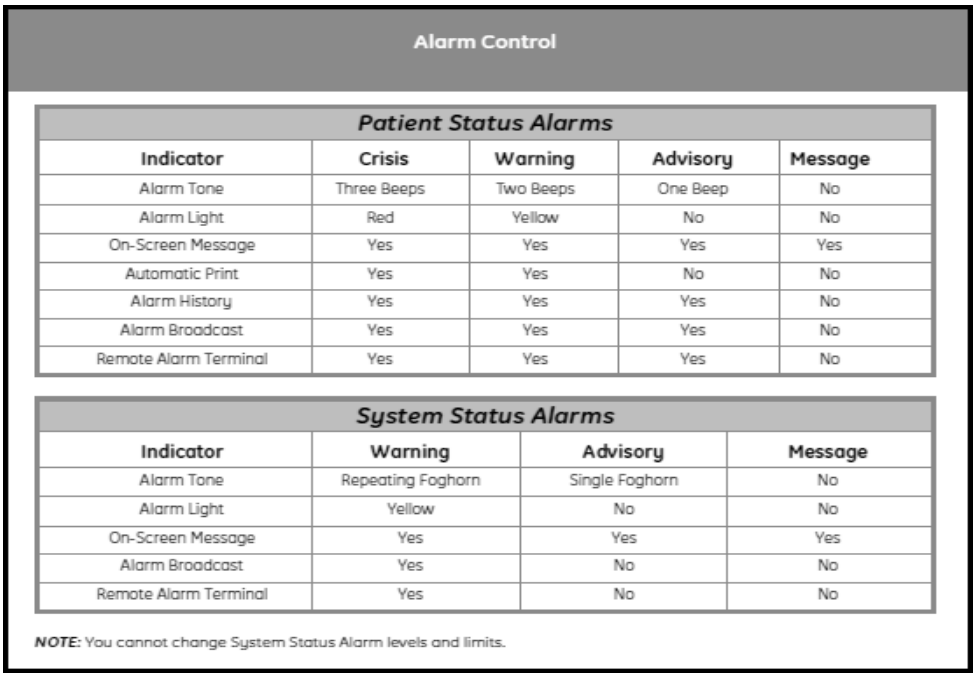
SpO_2 Low-limit Threshold Alarm

During the preliminary data collection period (June 1–14, 2012), the 8 NICU and 11 NICU default alarm setting for the SpO_2 low-threshold limit alarm was set to generate an audible alarm if a patient's SpO_2 was less than or equal to 90%. On 8 NICU, the SpO_2 low-threshold alarm generated 7% of all parameter alarms, resulting in approximately 40 alarms/patient/day. Similarly, on 11 NICU, the SpO_2 low-threshold alarm generated 9% of all parameters alarms, resulting in approximately 35 SpO_2 alarms/patient/day (see Table 4.3). The severity level of the SpO_2 alarm defaults to an "advisory" (i.e., low importance) alarm—which indicated an event that required monitoring but was not serious. However, an advisory alarm sounds continuously until the alarm condition self-corrects or is fixed by a clinician. Different types of physiologic monitor alarms—patient status and system status (technical) alarms—generate different audible alarms and visual alerts that are associated with different alarm severity levels (see Figure 4.1).

Table 4.3. Parameter Alarms (Pilot Data)

Physiologic Monitor Alarms Parameter Alarms	8 NICU (ADC 9) June 1–14, 2012 (*n* = 63,562)	11 NICU (ADC 14) June 1–14, 2012 (*n* = 76,339)
RR High	17,051 (26%)	17,811 (23%)
Art BP Systolic High	12,734 (19%)	16,805 (22%)
Art BP Systolic Low	8,624 (13%)	5,712 (7%)
Art BP Mean High	6,066 (9%)	8,393 (11%)
Art BP Mean Low	779 (1%)	1,277 (2%)
Art BP Diastolic High	5,500 (8%)	7,339 (10%)
Art BP Diastolic Low	2,820 (4%)	3,221 (4%)
SpO2 Low	4, 987 (7%)	7,146 (9%)
No Breath (Apnea)	350 (<1%)	985 (1%)

Figure 4.1. Alarm Notification and Severity Levels
(Solar 8000i: Quick reference guide by GE Healthcare, 2008a.Adapted with permission)



Alarm Control

Patient Status Alarms

Indicator	Crisis	Warning	Advisory	Message
Alarm Tone	Three Beeps	Two Beeps	One Beep	No
Alarm Light	Red	Yellow	No	No
On-Screen Message	Yes	Yes	Yes	Yes
Automatic Print	Yes	Yes	No	No
Alarm History	Yes	Yes	Yes	No
Alarm Broadcast	Yes	Yes	Yes	No
Remote Alarm Terminal	Yes	Yes	Yes	No

System Status Alarms

Indicator	Warning	Advisory	Message
Alarm Tone	Repeating Foghorn	Single Foghorn	No
Alarm Light	Yellow	No	No
On-Screen Message	Yes	Yes	Yes
Alarm Broadcast	Yes	No	No
Remote Alarm Terminal	Yes	No	No

NOTE: You cannot change System Status Alarm levels and limits.

Ordinarily, the SpO_2 low-limit threshold alarm is set at a conservative setting (less than or equal to 90%) with an advisory level alarm (one beep and on–screen message) that has neither been adjusted nor evaluated since the monitors were purchased in 2007 (see Table 4.4 and 4.5). The SpO_2 alarm threshold settings can be customized by RNs according to the individual patient's needs. It is important to note that adjustments can be made to the SpO_2 threshold alarm (low and high limit) and to the alarm severity level (message-to-crisis), but these adjustments are temporary, and the threshold limits and associated alarm severity level will revert back to the unit default settings when the RN resets the default alarms (at the beginning of each new shift or when patient monitoring is discontinued). However, in light of our preliminary analysis and the known contributions of SpO_2 alarms associated with a low-threshold limit of less than or equal to 90% toward alarm burden, we aimed to reduce the frequency of non-actionable SpO_2 alarms by lowering the unit default SpO_2 low-limit threshold alarm to less than or equal to 88%. Recently, investigators have reported moderate reductions in the frequency of non-actionable and clinically insignificant SpO_2 alarms by using lower SpO_2 low-limit threshold alarms (Graham & Cvach, 2010; Taenzer, Pyke, McGrath, & Blike, 2010; Welch, 2011).

Table 4.4. 8 and 11 NICU Default Parameter Alarm Threshold Limits (shortened)

Parameter Limits		
	Low	High
HR	50	130
PVC/min	…	10
ST-I	-2.0	2.0
ST-II	-2.0	2.0
ST-III	-2.0	2.0
ST-V1	-2.0	2.0
ST-AVL	-2.0	2.0
ST-AVF	-2.0	2.0
ST-AVR	-2.0	2.0
ST-V2	-2.0	2.0
ST-V3	-2.0	2.0
ST-V4	-2.0	2.0
ST-V5	-2.0	2.0
ST-V6	-2.0	2.0
NBP-S	90	160
NBP-D	50	90

Parameter Limits		
	Low	High
CO2-Insp	…	5
CO2-Resp	5	30
No Breath	…	20
SpO2	90	105
SpO2-R	50	130
BT	32 C	40 C
SvO2	60	80
RR	5	30
RR-Apnea	…	20
TEMP 1	32 C	40 C
TEMP 2	32 C	40 C
O2-Insp	18	102
O2-Exp	-1.0	102.0
N2O-Insp	-1.0	80.0
N20-Exp	-1.0	80.0
N2-Insp	-1.0	85.0

Table 4.5. 8 NICU and 11 NICU Default Arrhythmia and Parameter Alarm Severity Levels (shortened)

Arrhythmia Alarm Levels				
	Crisis	Warning	Advisory	Message
Asystole	x			
Vfib/Vtac	x			
V Tach	x			
VT >2			X	
V Brady		x		
Couplet				X
Bigeminy				X
Acc Vent		x		
Pause		x		
Trigeminy				X
R on T				X
PVC				X
Tachy			X	
Brady			X	
Atrial Fib for Patient data modules			X	

Parameter Alarm Levels				
	Crisis	Warning	Advisory	Message
HR		X		
CO2 No Breath		X		
RM No Resp		X		
PVC				x
ST			X	
ART		X		
PA			X	
CVP				X
CO2			X	
NBP			X	
NBP M Only			X	
SPO2			X	
FEM		X		
UAC			X	
UVC			X	
LA			X	
ICP		X		
SP				X
SVO2			X	
TC				X
BIS				X
ART Rate				X
SPO2 Rate				X
BT				X

SpO_2 Alarm Delay

In addition to the new SpO_2 low-limit threshold of 88%, the experimental unit received a Solar 8000i monitor software upgrade (Version 5.5) in order to activate the SpO_2 15-s alarm delay. With the GE Solar 8000i bedside monitoring system (Version 5.4), there is no SpO_2 alarm delay feature to assist with reduction of nuisance alarms; instead, there is only a 5-s delay from the time a SpO_2 alarm low-threshold alarm is triggered to the time this alarm becomes audible. The combined use of a lower SpO_2 low-limit threshold alarm and a 15-s alarm delay feature—was investigated to reduce the frequency of brief, self-correcting SpO_2 alarms—which are clinically insignificant and simply annoy patients and staff (see Table 4.6). Combining lower SpO_2 alarms limits and an alarm delay is most often studied in patient care areas that have implemented patient surveillance systems (Gross, Dahl, & Nielsen, 2011; Taenzer, Pyke, McGrath, & Blike, 2010; Welch, 2011). Although researchers advocate use of the above combination of SpO_2 alarm settings, the literature search found no study that measured the impact of these settings on the incidence of SpO_2 alarms (pre- and post-setting modification) in an ICU setting—which is what our study measured.

Table 4.6. Summary of Intervention: Modification of SpO_2 Alarm Settings

SpO_2 Alarm	Comparison Unit	Experimental Unit
SpO_2 low-limit threshold	≤ 90%	≤ 88%
Alarm delay	5 s	15 s

Intervention: Electrode-site Skin Preparation and Daily Use of Ag/AgCl–Foam, Pre-gelled *Wet* ECG Electrodes

For this study, the intervention for reducing technical, arrhythmia alarms , and false positive arrhythmia alarms involved performing a novel electrode-site skin preparation—using fine abrasive ECG skin preparation paper—prior to the daily application of Ag/AgCl-foam, pre-gelled *wet* ECG electrodes in the experimental unit. The June 2012 preliminary analysis found that the most frequent technical alarms were the *lead fail* (ECG and RR) and *arrhythmia suspend* (artifact level 2) alarms; the suspected cause of these technical alarms was poor ECG electrode signal related to unsatisfactory conductivity or motion artifact. Notably, the data did not include the frequency of artifact level 1 alarms because these alarms were classified as "message" level alarm (i.e., least critical) with no audible alarm; therefore, the frequency of artifact level 1 alarms could not be retrieved from the CIC Pro Clinical Information Center.

Technical Alarms

ECG lead fail alarm. The *lead fail alarm* is an alarm condition wherein no ECG waveforms are displayed on the physiologic monitor because the *smart lead fail feature* is ineffective and has failed. To support continuous monitoring, the smart lead fail feature continuously checks the integrity of the ECG electrodes. If the quality of an electrode signal deteriorates to an unsatisfactory level, a lead fail message is displayed on the physiologic monitor. It is important to note, one single ECG lead fail alarm will only display a "message" alarm and a V-lead and a single ECG limb lead failure will not cause an ECG Leads Fail alarm. However, a combination of two single ECG limb lead failures will cause an "ECG Leads Fail" alarm and this alarm will be measured. If the failing lead negatively influences the ECG waveform being monitored in the top position on the display, the ECG monitoring automatically

switches to another lead (see Table 4.7). On 8 NICU, the lead fail (ECG and RR) alarms accounted for 62% of all technical alarms (76 alarms/day; 5 alarms/patient/day). Similarly, in 11 NICU, the report identified that the lead fail alarms accounted for 52% of all technical alarms (69 alarms/day; 5 alarms/patient/day; see Table 4.8 for the frequency of common technical alarms that occurred in June 2012).

Table 4.7. Smart Lead Fail Feature

Message	New Lead Monitored
RA FAIL	Lead III
RL FAIL (Patient Data Module only)	The lead selected to display in the top trace position
LL FAIL	Lead I
LA FAIL	Lead II
V FAIL	Lead II
LEADS FAIL	No waveform displayed

Artifact alarm. Artifact alarms are due to a transient condition resulting from intermittent noise and artifact. Artifact alarms begin at level 1 and progress to level 2 if the ECG noise lasts for 20 s of the last 30 s (GE Healthcare, 2007). It is important to note that when *artifact level 1* alarms are triggered, full arrhythmia processing is suspended, yet the lethal arrhythmias detection software (EK-Pro, Version 11) remains active for two life threatening arrhythmia alarms—ventricular fibrillation/tachycardia and asystole (GE Healthcare, 2007).

Arrhythmia suspend alarm. The arrhythmia suspend alarm condition occurs as a result of ECG artifact In contrast, the *artifact level 2* alarm displays as an *arrhythmia suspend* alarm, and arrhythmia interpretation is completely suspended. The arrhythmia suspend alarm generates a continuous foghorn alarm until the quality of the ECG signal improves. To resume arrhythmia processing and alarms, the alarm condition must be resolved by checking lead placement, performing skin preparation and/or replacing the ECG electrodes, or adjusting electrode

placement (GE Healthcare, 2007).

The June alarms report showed that 129 arrhythmia suspend (artifact level 2) alarm events occurred over 2 weeks, which represented 12% (15 alarms/day) of all technical alarms on 8 NICU. Conversely, on 11 NICU, arrhythmia suspend alarm events represented 16% of all technical alarms (21 alarms/day) in a 14-day period (see Table 4.8). Although arrhythmia suspend alarms are infrequent in comparison with other types of monitor alarms, their potential effect on patient safety is considerable and potentially devastating. That is, during an arrhythmia suspend alarm event, arrhythmia analysis is not conducted—including analysis for the most lethal arrhythmias.

Table 4.8. Common Technical Alarms (Pilot Data)

Physiologic Monitor Alarms Technical Alarms	8 NICU (ADC 9) June 1–14, 2012 (*n* = 1,078)	11 NICU (ADC 14) June 1–14, 2012 (*n* = 1,908)
Lead fail (ECG)	564 (52%)	830 (44%)
Lead fail (RR)	104 (10%)	163 (9%)
Arrhythmia suspend	129 (12%)	303 (16%)
NBP max time	57 (5%)	215 (11%)
SpO2 sensor	42 (4%)	130 (7%)
No ECG	76 (7%)	99 (5%)
Art disconnect	63 (6%)	99 (5%)
Invasive pressure sensor	38 (4%)	54 (3%)

Arrhythmia alarms. In June 2012, the incidence of arrhythmia alarms represented less than 3% and 1.6% of all monitor alarms on 8 NICU and 11 NICU, respectively (see Table 4.1). This finding is in accordance with reported incidences of arrhythmia alarm in adult ICUs (Biot et al., 2000; Seibig et al., 2010a, 2010b). However, while few, false-positive arrhythmia alarms set at a severity level of crisis play a significant role in the development of physiologic monitor

alarm fatigue because the alarm continuously sounds until it is silenced by the user (GE Healthcare, 2008b). The preliminary 8 NICU and 11 NICU alarm data revealed that the most common cardiac arrhythmia alarms were those related to an abnormal heart rate, with tachycardia being the most common alarm type and bradycardia the next most common. The most prevalent arrhythmia alarms related to an abnormal heart rhythm (in ascending order) were ventricular fibrillation/tachycardia, ventricular bradycardia, accelerated ventricular, asystole, pause, atrial fibrillation, ventricular tachycardia, and ventricular tachycardia, greater than 2 (see Table 4.9). Some cardiac arrhythmias alarms are classified with an alarm notification and severity level of crisis (e.g., ventricular tachycardia and ventricular tachycardia/ventricular fibrillation) and are considered "latching alarms." The International Electrotechnical Commission's (IEC) standard 60601-1-8 defines a *latching alarm signal* as an alarm that continues to be generated after its triggering event has ceased. This implies that the alarm signal (audio alarm) continues to annunciate—and requires that the RN acknowledge the alarm by pressing the silence button. In contrast, a "non-latching" alarm signal automatically ceases being generated when its triggering event no longer exists. Physiologic monitor are configured with a mixture of both latching and non-latching alarm signals (IEC, 2006).

Table 4.9. Arrhythmia Alarms (Pilot Data)

Physiologic Monitor Alarms Arrhythmia Alarms	8 NICU (ADC9) June 1–14, 2012 (*n* = 1,957)	11 NICU (ADC 14) June 1–14, 2012 (*n* = 2,183)
Tachycardia	862 (44%)	469 (21%)
Bradycardia	658 (34%)	415 (19%)
Ventricular Tachycardia ≥ 2	169 (9%)	884 (40%)
Atrial Fibrillation	165 (8%)	0
Ventricular Tachycardia	37 (2%)	415 (19%)
Pause	21 (1%)	0
Asystole	19 (1%)	0
Accelerated Ventricular	16 (<1%)	0
Ventricular Bradycardia	8 (<1%)	0
Ventricular Fibrillation/Tachycardia	2 (<1%)	0

Electrode-site Skin Preparation. The UCSF Medical Center Department of Nursing's procedure on cardiac monitoring (2013) specifies cleaning the skin with alcohol swabs every 48 hr and briskly drying the area with gauze prior to the application of ECG electrodes; however, it is known that RNs find this task time-consuming and that this task is not consistently performed; accordingly, in common practice, skin preparation may or may not be done prior to application of ECG electrodes. During Assessment 2, the experimental unit utilized single-use Philips ECG Skin Preparation Paper (M4606A) composed of fine sandpaper strips to stroke the skin (i.e., 1–5 strokes per electrode site) for the purpose of gently abrading the skin to reduce skin potential and motion artifact. This ECG preparation paper is commercially available and sold for adult and pediatric use.

The professed advantages of using ECG skin preparation paper over other skin preparation methods is that the former method removes dead skin cells that impede conduction;

at the same time, use of ECG skin preparation paper causes little skin damage or irritation (Mirvis et al., 1989; Smith, 1984). In their training, registered nurses were informed that the only contraindication for using this preparation paper was that it was not to be used on skin sites with established erythema, lesions, or injuries of any kind (Philips Healthcare, 2008).

ECG electrodes. The Kendall 530 series Ag/AgCl-foam *solid hydrogel* ECG electrodes were both used on the study units during Assessment 1 and on the control unit during Assessment 2. The manufacturer's recommendation for the Kendall 530 series Ag/AgCl-foam *solid hydrogel* ECG electrodes is that they are to be used primarily for Holter monitoring, stress tests, and on very diaphoretic patients (Covidien, 2013a). This electrode utilizes an "aggressive adhesive" and the foam electrode adheres strongly to the skin. Because this type of electrode uses solid hydrogel for conductivity (and not a superior wet-gel), the manufacturer claims this electrode can remain out of its packaging for up to 1 month (Covidien, 2013a). It is unclear why the Kendall 530 series Ag/AgCl-foam *solid hydrogel* ECG electrode was introduced and selected as the primary ECG electrode for use in both the experimental and control unit many years ago—and consequently used during Assessment 1 in both units and during Assessment 2 in the control unit.

Although a variety of ECG electrodes are available for clinical use, the Ag/AgCl, foam, pre-gelled *wet* ECG electrode has been assessed as high-quality for continuous cardiac monitoring and is well regarded for its clinical advantage in adhering very well to patients' skin (Chi, Jung, & Cauwenberghs, 2010; Tronstad, Johnsen, Grimnes, & Martinsen, 2010). During Assessment 2, the experimental unit used the Philips Ag/AgCl-foam, pre-gelled *wet* ECG electrode in individually packaged packets of 5 electrodes.

Application. Electrode application is also an important and often over-looked step in the site preparation process. Specifically, the lead wire should be attached to the electrode before applying the ECG electrode on the patient, the electrode should be placed on flat, fleshy parts while avoiding bony prominences and major muscles. When the electrode is applied to the patient, gentle pressure should be applied to the outer edges of the electrode and not directly on the center of the electrode, in order to minimize the creation of air pockets and gel leakage (Hanish, Neustein, Van Cott, & Sanders, 1971, Mirvis et al., 1989; Philips Healthcare, 2008). Nurses received training on the correct electrode application technique, and all patients were monitored with Solar 8000i devices (GE Healthcare, WI) using Mason-Likar limb leads and one precordial lead.

Storage. Electrode gel influences the transmission of signals from the patient's skin to the electrode. Insufficient conductive gel due to evaporation from improper storage can cause high electrode impedance and unstable ECG traces (Smith, 1984). Preferably, electrodes should be stored in sealed metal foil packets that are resistant to moisture in single packets of 5 electrodes (see Table 4.10). Electrodes should not be stored in open bins, bags, or attached to lead wires when not is use (Melendez & Pino, 2012; Smith, 1984; Turkmen & Pantiskas, 2011). See Table 4.11 for a summary of the intervention regarding electrode-site skin preparation and ECG electrode application. In our study, we expected that skin preparation and the daily application of high-quality ECG electrodes would improve the quality of the ECG and other waveform signals; this improvement would be evidenced by fewer technical alarms (i.e., ECG lead fail, artifact, and arrhythmia suspend alarms) and false positive arrhythmia alarms during Assessment 2 in the experimental unit.

Table 4.10. Use and Storage of Patient Monitoring Supplies

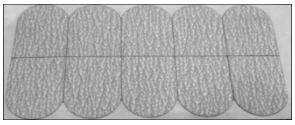
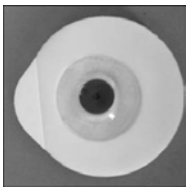
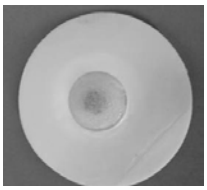
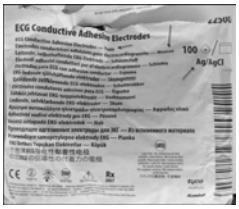
Elements	Control Unit	Experimental Unit
Skin Preparation	Usual care	 ECG Skin Preparation Paper
Electrode	 Ag/AgCl–hydrogel	 Ag/AgCl–wet gel
Package	 Bulk electrodes in open packages, bins/shelves/attached to lead wires	 Single packets of 5 electrodes kept at each bedside

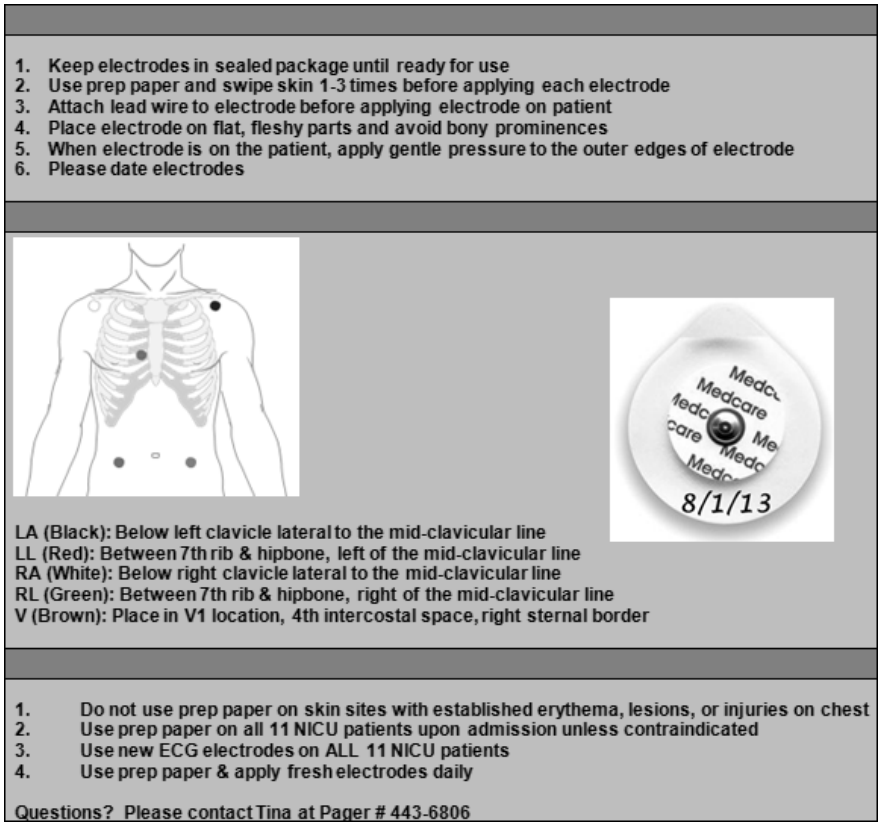
Table 4.11. Summary of Intervention: Skin Preparation and Electrode Application

Elements	Control Unit *(Usual care*)	Experimental Unit
Skin preparation	a. Cutting or shaving hair (if necessary) b. Cleaning skin with alcohol pad, and briskly drying with a regular gauze pad	a. Cutting or shaving hair (if necessary) b. 1–3 strokes of the ECG skin preparation paper (Philips M4606A)
Type of electrode	Ag/AgCl, foam, *solid hydrogel* ECG electrodes (Kendall, 530 series # 22500)	Ag/AgCl, foam, pre-gelled, *wet* ECG electrodes (Philips, 40493D)
Application	Every 48 hours (and as needed)	Every 24 hours (and as needed)
Storage	100 disposable ECG electrodes stored in bulk package and electrodes are loosely distributed at each patients' bedside	5 disposable ECG electrodes stored in an individual package at each patients' bedside

Availability of new patient monitoring supplies. The new supplies were provided to the experimental unit in kits; the kits contained (a) ECG skin preparation paper (5 tabs), (b)

Ag/AgCl, foam, pre-gelled *wet* ECG electrodes in sealed packets of 5 electrodes, (c) black pen marker, and (d) simple skin preparation and electrode application instructions (see Table 4.10 and Figure 4.2). Supplies were delivered on a daily basis and as needed and were kept at the patient's bedside. Staff received instructions that all patients admitted to the experimental unit were to receive the new skin preparation and daily application of Ag/AgCl foam, pre-gelled *wet* ECG electrodes—even new patients who were being admitted or transferred with other brands of electrodes in place. For patients with established erythema, lesions, or skin injuries of any kind, the nursing staff was instructed to not use the skin preparation paper and simply apply the high-quality ECG electrodes. Furthermore, the Ag/AgCl-foam *solid hydrogel* ECG electrodes were removed from the experimental unit during Assessment 2 to promote the use of the novel patient monitoring supplies.

Figure 4.2 Instructions for Use of New Patient Monitoring Supplies



Training. Prior to the intervention (Assessment 2), the RNs were trained on the new SpO_2 alarm settings, the SpO_2 low-limit threshold of less than or equal to 88%, and the 15-s SpO_2 alarm delay. Furthermore, the RNs and PCAs received training on the daily use of skin preparation paper (Philips M4606A), dating and application of Ag/AgCl, foam, pre-gelled *wet* ECG electrodes (# 40493D), and proper ECG electrode placement to improve electrical signal and reduce motion artifact. The investigators trained the nursing staff during unit-based staff meetings and education was reinforced with staff during daily rounds on the experimental unit during the 31-day intervention period (Assessment 2). In addition, the staff was informed that either the investigator or a research assistant would perform daily rounds to confirm that the new skin preparation has been performed and that the new ECG electrodes had been applied.

Data Collection

Instruments. For this study, the CareScape Gateway (CSG, GE Healthcare) research Version 1.1 was used to transfer physiological monitoring data from the monitoring system to an external server (see Figure 4.3). Data were collected and processed using BedMasterEx software (Excel-Medical Electronics). The BedMasterEx software was installed on the external server and had client licenses that enabled investigators to analyze the data obtained from the physiologic monitors. The Solar 8000i software Version 5.4 and ECG arrhythmia detection software (EK-Pro, Version 11) were used on both units during Assessment 1 (baseline). At Assessment 2, the experimental unit used Solar 8000i software Version 5.5. The only difference between Version 5.4 and 5.5 is that Version 5.5 has the SpO_2 alarm delay feature. The BedMasterEx software program was used to export full physiologic monitor waveforms and alarm data to a dedicated server. Use of the BedMasterEx viewer enabled the investigators to analyze each alarm and the accompanying waveforms at the time the alarm was triggered (see Figure 4.4).

Figure 4.3. 8 and 11 NICU Alarm Study Connectivity Diagram (Adapted with permission by B. J. Drew, 2013)

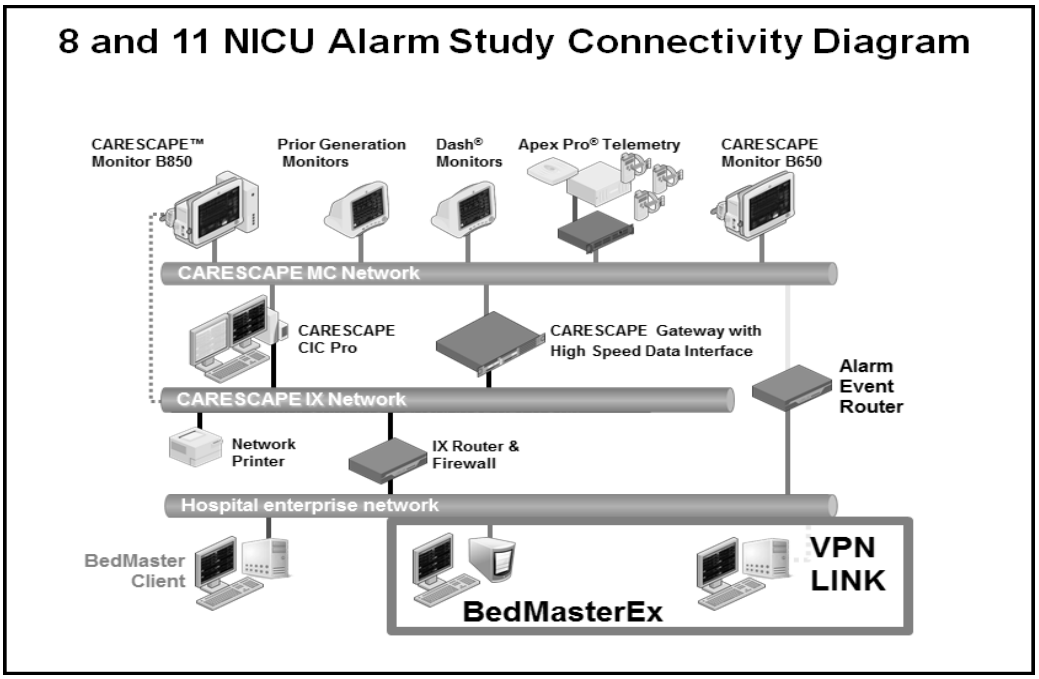
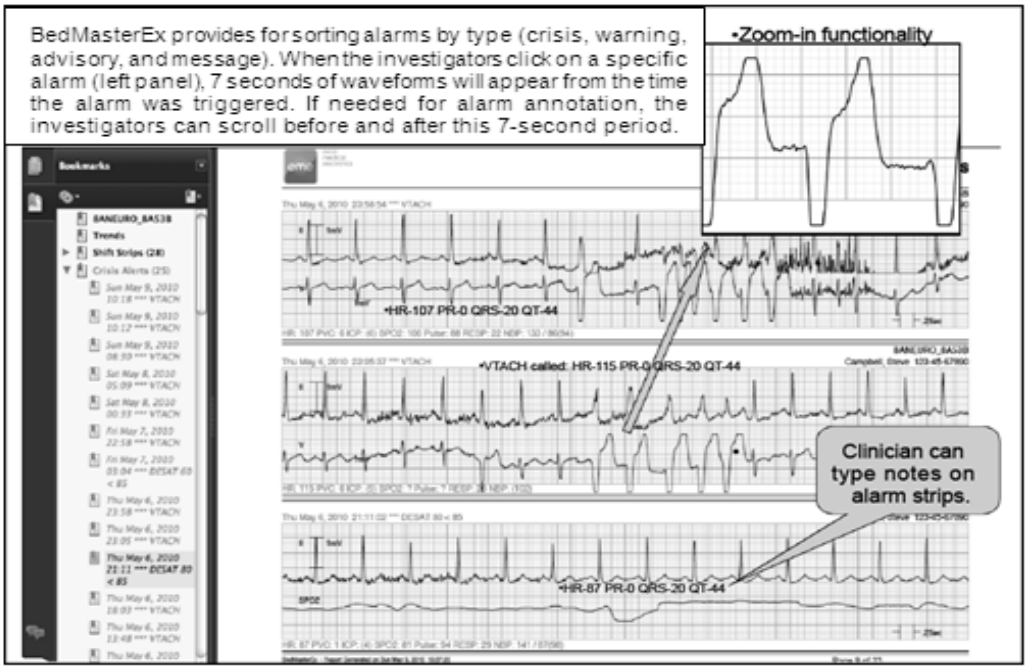


Figure 4.4. Illustration of Process for Alarm Annotation Using BedMasterEx Software (With permission by Excel-Medical Electronics, 2013)



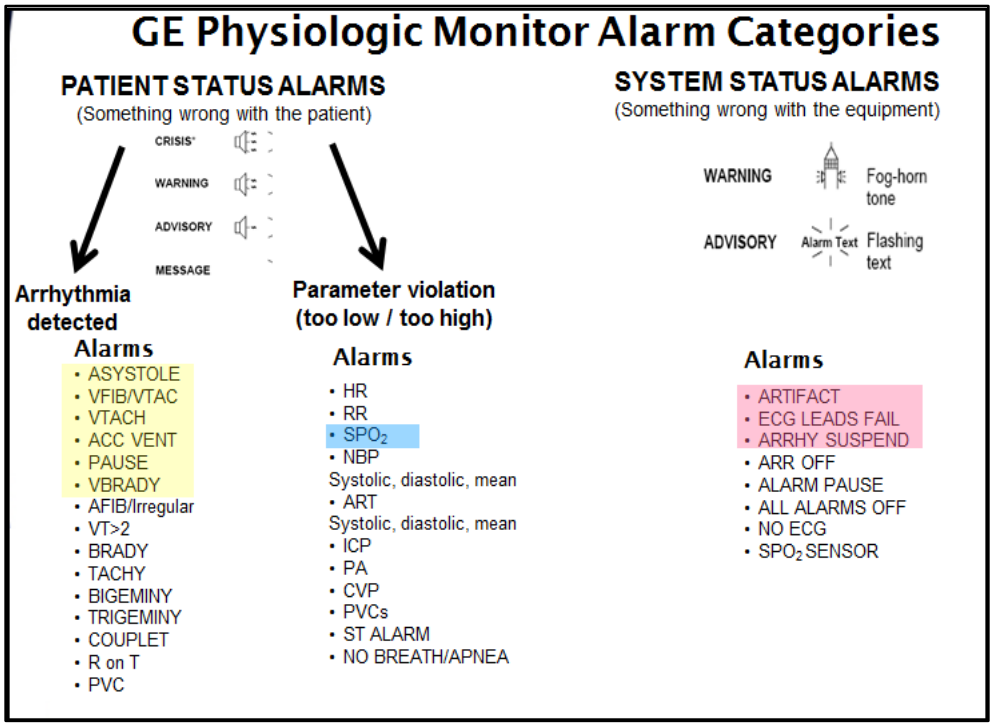
Alarm data. The unit of analysis in this study was patients. The analysis used retrospective subject-case physiologic monitor data sets from patients admitted to the 8 and 11 NICU during the two 31-day assessment periods. Subjects were not be enrolled in this study, and bias in subject selection was be precluded. We anticipated that the study's sample size would consist of 317 patients. The number of study patients was calculated as follows: the ADC for 8 NICU is 9 patients, and the ADC for 11 NICU is 14 patients, for a total of 23 patients. The total ADC was multiplied by 62 days (i.e., two 31-day study periods), and the resulting product was divided by an average length of stay of 4.5 days. The resulting calculations were 124 patients for 8 NICU and 193 patients for 11 NICU.

The data were divided into records that began when the physiologic monitor was activated for a patient admission and ended when the monitor was deactivated for a patient transfer or discharge. Based on our preliminary analysis in June 2012, the quantity of physiologic monitor alarms was predicted to be high—an estimated 646,935 alarms, total, for both units throughout the study. The investigator estimated that the patients in 8 NICU would generate 4,757 alarms/day (i.e., 147,467 alarms for Assessment 1 [baseline] and a similar frequency for Assessment 2). The 11 NICU patients would generate 5,745 alarms/day (i.e., 178,095 alarms for Assessment 1 [baseline] and a similar frequency for Assessment 2). We predicted that, in total, 651,124 physiologic monitor alarms would be collected, and 18,352 arrhythmia alarms (including atrial fibrillation and ventricular tachycardia greater than or equal to 2) would potentially require annotation to determine the frequency of false-positive cardiac arrhythmia alarms. The alarm estimates were conservative and were expected to fluctuate (i.e., higher or lower) based on patient census and acuity during the two assessment periods. These arrhythmia alarm predictions did not account for a reduction in cardiac arrhythmia alarms that were

anticipated to occur on the experimental unit during Assessment 2. All waveforms and alarm data were analyzed using BedMasterEx software developed by Excel-Medical Electronics. Data from each unique patient's alarm were collected along with their total physiologic monitoring time (which was computed by an algorithm that excluded interruptions).

Arrhythmia annotation. All physiologic monitor alarms were collected to determine the alarm rates (i.e., mean hourly rate of physiologic monitor alarms per patient); alarms highlighted in yellow were annotated by the investigator and were used to determine the impact of the new skin preparation method and use of high quality ECG electrode on mean hourly arrhythmia alarm rates (see Figure 4.5). Arrhythmia alarms were analyzed and annotated as true-positive or false-positive alarms according to GE arrhythmia alarm definitions (see Appendix A). We did not investigate false-negative arrhythmia alarms. Multiple ECG waveforms and other physiological measurements were examined for each patient's arrhythmia alarms. As part of a larger UCSF alarm study, documentation of alarm annotation for both study units occurred in a secure, password protected database (i.e., Medidata Rave) that is accessible only to members of the research team who have successfully completed the required on-line training modules.

Figure 4.5. Physiologic Monitor Alarm Categories for 8 and 11 NICU Alarm Study (Adapted with permission by B.J. Drew, 2012)



Specific SpO_2 alarms, highlighted in blue, were used to identify reductions in mean hourly SpO_2 alarm rates associated with a lower SpO_2 low-limit alarm threshold and a 15-s alarm delay. Alarms highlighted in pink were used to determine the effect of the new skin preparation and use of high-quality ECG electrodes on the mean hourly ECG lead fail, artifact, and arrhythmia suspend alarm rates (see Figure 4.5). The prevalence of all physiologic monitor alarms was counted according to a standardized protocol (see Appendix B).

Patient demographic variables. In addition to obtaining information from administrative nursing records, basic patient information was collected via the EMR for select demographic variables (see Table 4.12). It is important to note that the BedMasterEx application is bed-centric and not patient-centric. Accordingly, the BedMasterEx database did not contain

patient health information other than two patient identifiers (i.e., patient name and medical record number) if entered correctly by RNs.

Table 4.12. Alarm Study: Patient Variables

Patient Variables	Additional data
1. ICU admit time and date 2. Monitoring admit time and date 3. Age 4. Gender 5. Race and ethnicity 6. Primary admitting service 7. Discharge status	1. Code blue event 2. In-hospital death

Patient outcome variables. To obtain information on patient outcomes on the experimental unit and control unit, the incidence of "code-blue" events was examined during 6 months preceding and after implementation of alarm changes. Data on code blue events were collected and classified as cardiopulmonary arrest (chest compressions were administered and/or the patient was defibrillated) and acute respiratory compromise (neither chest compressions nor defibrillation occurred, but the patient required assisted ventilation). These events were documented on a code blue records and this information was attainable through the medical center's quality department. Furthermore, a log of overhead code-blue announcements was obtained from Security Services. The log of code-blue announcements was reconciled with the code-blue records to ensure that all adverse patient outcomes were collected during the assessment periods for both units. It was expected that adverse patient events would be few and the intervention would not contribute to an increase incidence of events in the experimental unit.

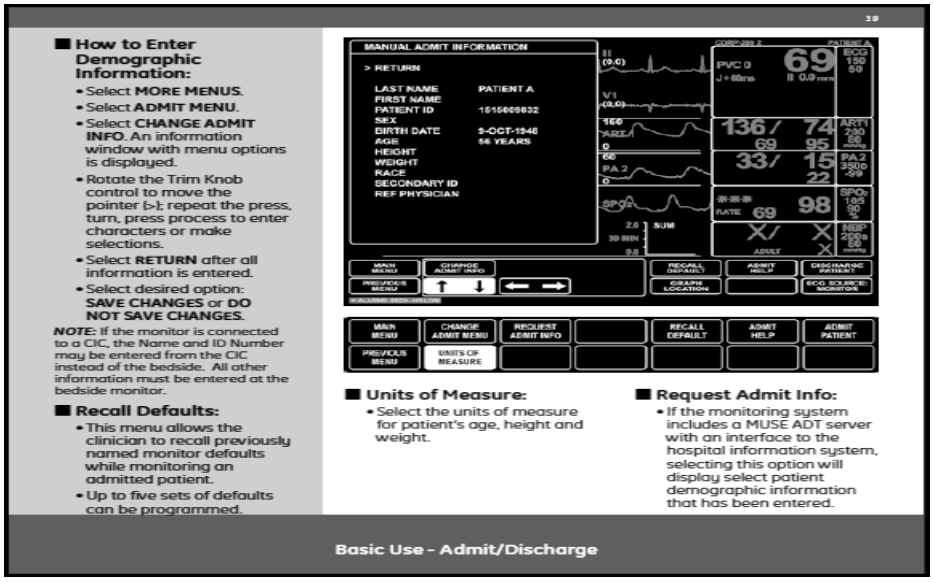
Administrative information. Existing nursing records were reviewed and data including but not limited to the study units' bed occupancy and acuity were collected through administrative, quality assurance, and financial reports for both study units.

Training plan. Prior to the start of data collection, the principal investigator and all research personnel received clinical software training on the EMR (i.e., Epic) and BedMasterEx applications. BedMasterEx training was provided to the principal investigator on February 15 and on March 13, 2013. On March 14 and April 8, 2013, the principal investigator received additional Epic training specifically customized to facilitate the collection of patient variables by an Epic credentialed trainer and the principal educator. The investigator also attended an educational session, "How to Annotate GE Monitor Arrhythmia Alarms" that was presented by Barbara Drew, PhD, on February 25, 2013 and has previously completed the N225 Cardiac Rhythm and Analysis course at the University of California, San Francisco, School of Nursing (also taught by Dr. Barbara Drew). Furthermore, the principal investigator has over 20 years of clinical nursing experience, including analyzing continuous ECG and physiological monitoring data in adult ICU and emergency department settings. In addition, the investigator fulfilled the requirements for Medidata Rave® 5.6 certified clinical research coordinator training in order to enter study variables and alarm annotations.

The principal investigator and members of the research team provided visual proof of all true- and false-positive arrhythmia alarms, to allow for further examination and review by a member of the ECG Monitoring Research Lab in the UCSF School of Nursing. The use of the BedMasterEx software application permitted the review of cardiac arrhythmia alarms and associated physiological waveforms as often as needed for escalation and final determination by Dr. Barbara Drew, Professor of Nursing Science and Clinical Professor of Medicine, Cardiology.

Missing data. Patient information such as acuity, diagnoses, Admit– Discharge–Transfer (ADT) information is documented in the electronic medical record (i.e., Epic); therefore, we anticipated that few data would be missing—provided that the clinicians entered the information according to the medical center's documentation standards. We anticipated that some physiologic monitor alarm data would not be captured or would be "lost" (e.g., due to periods of system "down-time" when either the CSG or BedMasterEx application would be serviced or to periods when data collection would inadvertently be interrupted). Service interruptions were documented in an event log on a secure server. Even if alarm data were not missed, we anticipated that if RNs failed to correctly admit or discharge a patient (either at the bedside physiologic monitor or at the CIC Pro Clinical Information Center), the patient's physiologic waveform data and alarms might become integrated into the previous patient's record, thus leading to potential errors in determining the prevalence of alarms per patient and individual patient monitoring time (see Figure 4.6). A process was implemented whereby patient census obtained from the EMR system was reconciled with the names and medical records of patients assigned to physiologic monitors in the BedMasterEx application software three times a day (i.e., at around 8:00, 16:00, and 23:00). When patient identification discrepancies were identified, a comment with the correct information was entered in the BedMasterEx (under the waveform tab) to aid the investigators during the annotation and data analysis phase of the study; this procedure ensured that, to the greatest extent possible, alarms were correctly attributed to patients.

Figure 4.6. Patient Admit/Discharge Screen (Solar 8000 I: Quick reference guide by GE Healthcare, 2008a.Adapted with permission)



Data Analysis and Statistics

Stata V. 13 Data Analysis and Statistical Software and the IBM Statistical Package for the Social Sciences (SPSS) for Windows Graduate Student V.22 were used for the statistical analyses. The alpha level for statistical significance was preset at p value less than .05. In addition to parametric and non-parametric models, descriptive statistical analyses were applied as appropriate according to the study variables. Data were reported as frequencies and percentages for categorical variables; means ± standard deviations (*SD*) were reported for continuous variables.

Testing for an interaction

For Assessment 1 and Assessment 2, a two-way ANOVA was used to obtain the mean hourly unique alarm rates (number of unique alarms per patient per monitoring hour) for the experimental and control unit 2; however, a corresponding negative binomial regression was

used to obtain *P* values. A negative binomial regression was used to establish statistical significance because research on life-threatening arrhythmia alarms and infrequently occurring alarm events, involves counts in a fixed period of time; these counts are non-normally distributed with highly skewed distributions. Therefore, using traditional statistical methods such as ordinary least square (OLS) regression was deemed inadequate (Hutchinson & Holtman, 2005). The negative binomial regression was favored over the Poisson regression because the former does not assume a normal distribution of the errors terms and dependent variables (e.g., alarm counts). In essence, the negative binomial regression makes no assumptions regarding equidispersion and does not require special adjustments when overdispersion in counts are present (Allison, 1999).

The negative binomial regression is appropriate for the study of infrequently occurring count data, such as alarms, because typically rare alarm counts for life-threatening arrhythmia alarms (e.g., ventricular fibrillation, ventricular bradycardia, and asystole) have a propensity to gather around discrete values (e.g., 0, 1, 2); these counts are somewhat analogous to other discrete health–related events such as pregnancy or hospitalizations (Hutchinson & Holtman, 2005). This positively skewed distribution is typically truncated at 0 and progressively trails off towards higher values. In this type of distribution the mean is generally low but greater than the median because of the effect of a few relatively large observations (Hutchinson & Holtman).

This statistical model was chosen because overdispersion (i.e., too much variability) of alarm counts was anticipated in our samples—meaning some patients would not trigger certain types of alarms during their ICU hospitalization and some patient (i.e., outliers) would generate numerous alarms—thereby creating an excess variation between alarm counts. As with the Gross et al. (2011) alarm study on medical–surgical units, in the present study, overdispersion was

expected and was further supported by the examination of the preliminary June 2012 analysis, which found that high per patient rates of physiologic monitor alarms were predominately generated by a minority of patients. A statistical model was needed to accommodate this phenomenon because overdispersion whose variance exceeds the mean can cause standard errors of the estimates to be underestimated—that is, a variable may appear to be falsely significant (Hilbe, 2011). The combined statistical approach, two-way ANOVA and binomial negative regression, enabled the researcher to test for interactions among units and assessments using methods that provided the best fit for the data.

Subsequently, if the interaction was determined to be significant, a Student's *t*-test was performed to test the simple effects. The first of these effects was differences between the two independent group means; these differences were tested in order to evaluate the experimental unit's change in mean hourly alarm rates between Assessment 1 and Assessment 2. The second effect was the control unit's change over time in mean hourly alarm rates between Assessment 1 and Assessment 2. For the above reasons and to have the ability to handle the challenge of infrequently occurring repeatable alarms counts at the patient level, the above methods were used to specifically answer research questions 1 through 5:

Research Question 1. Is the mean hourly SpO_2 low-limit alarm rates difference between Assessment 1 and Assessment 2 for the experimental unit ($SpO_2 \leq 88\%$ with a 15-s alarm delay) different compared with the control unit ($SpO_2 \leq 90\%$ with a 5-s alarm delay)?

Research Question 2. Is the mean hourly technical alarm rates (e.g., ECG, artifact, arrhythmia suspend) difference between Assessment 1 and Assessment 2 for the experimental unit (who received a novel ECG skin site preparation technique and daily applications of high-quality ECG electrodes) different compared with the control unit (who received usual care)?

Research Question 3. Is the mean hourly arrhythmia alarm rates difference between Assessment 1 and Assessment 2 for the experimental unit (who received a novel ECG skin site preparation technique and daily applications of high-quality ECG electrodes) different compared with the control unit (who received usual care)?

Research Question 4. Is the in mean hourly audible alarm rates difference between Assessment 1 and Assessment 2 for the experimental unit (who received a novel ECG skin site preparation technique and daily applications of high-quality ECG electrodes) different compared with the control unit (who received usual care)?

Research Question 5. Is the mean hourly physiologic monitor alarm rates difference between Assessment 1 and Assessment 2 for the experimental unit (who received a novel ECG skin site preparation technique and daily applications of high-quality ECG electrodes) different compared with the control unit (who received usual care)?

Given that the patient is the unit of analysis in our methodology, a two-way ANOVA was performed to determine whether the experimental unit and the control unit differed in between-assessment changes (i.e., from Assessment 1 to Assessment 2) in mean percent of false-positive cardiac arrhythmia alarms (i.e., accelerated ventricular, asystole, pause, ventricular bradycardia, ventricular tachycardia, ventricular fibrillation/ventricular tachycardia). Of note, it was expected that the sample size would be small because only a certain proportion of patients would generate one or more of each of the six types of arrhythmia alarms that were annotated as either *true-positive* or *false-positive arrhythmia alarms*. The analysis met the assumptions of equal variance and *a P* less than .05 was evaluated for statistical significance. The above approach allowed the investigators to answer the last research question:

Research Question 6. Is the mean percentage of false-positive cardiac arrhythmia alarms (e.g., asystole, accelerated ventricular, pause, ventricular bradycardia, ventricular tachycardia, ventricular tachycardia/ventricular fibrillation) difference between Assessment 1 and Assessment 2 for the experimental unit different compared with the control unit?

Monitoring time and Age

A one-way ANOVA was used to determine whether there were differences in total patient monitoring hours and age (i.e., continuous variables) during Assessment 1 and Assessment 2 in the experimental unit and the control unit. The analysis met the assumptions of equal variance and an *F*-ratio statistic was evaluated for significance.

Demographic variables

A nonparametric procedure was utilized to complete the data analysis. The chi-squared (χ^2) test is reported to be the most appropriate test to test hypotheses when frequency data between proportions have been obtained for two or more exclusive categories that contain all the data (Shott, 1990). Examples of such data include patient demographic (i.e., categorical data) information for the experimental and control unit during the two assessment periods. Four main assumptions about the data were met in order to use the chi-squared test of hypothesized proportions: (a) random sampling, (b) independent observations, (c) mutually exclusive categories that include all observations, and (d) adequately large expected frequencies (Shott, 1990). All assumptions have been satisfied for use of the chi-square test for the analysis of race, gender, primary service, and patient discharge status. The Pearson Chi-square test was used to determine whether between-group differences in these characteristics.

Patient outcome variables

Descriptive statistics were obtained from the medical center's quality department to

report the incidence of (a) cardiopulmonary arrest (CPA; chest compressions were administered and/or the patient was defibrillated) and (b) acute respiratory compromise (ARC; neither chest compressions nor defibrillation occurred, but the patient required assisted ventilation). Data were obtained for both the experimental and control unit for 6 months preceding and after the intervention; these data are reported as rates: the number of CPAs/1,000 patient discharges and the number of ARCs /1,000 patient discharges.

Administrative information

Likewise, descriptive statistics were also utilized to report aggregate patient acuity data during Assessment 1 and Assessment 2 for both the experimental and control unit. A unit acuity summary report was obtained from the medical center's clinical software application (i.e., Clairvia®) to collect RN assessments of patient acuity during the study. Lastly, average daily census (ADC) data were obtained from the medical center's financial administrative database. The units' ADC was identified as the number of patients at midnight in both the experimental unit and control unit during Assessment 1 and Assessment 2.

Funding

This study was funded by the American Association of Critical Care Nurses (AACN), 2013 Philips Outcome Grant. In addition, investigators received computer and technical support from Excel-Medical Electronics (Jupiter, FL) and GE Healthcare (Waukesha, WI).

Summary

Physiologic monitor alarm fatigue is an understudied phenomenon, albeit one with great significance; as the proliferation of existing medical devices and the introduction of new technologies continues to increase the care environment. The review of the literature identified

measurement issues related to the variation in instruments used to collect alarm data and the descriptive or statistical tests utilized to analyzed the data and generate findings. Our study differentiates itself from previous research because the patient is the unit of analysis—and not alarms. This approach identifies unique alarm rates each patient and their individual contributions to the alarm rates based on their individual monitoring time.

Furthermore, our study has a robust data analysis plan: (a) Means and standard deviations of the hourly alarm rates and percentage of false-positive arrhythmia alarms were determined for the experimental unit and control unit at Assessment 1 and 2; (b) a negative binomial regression was performed to test the main effect of unit, the main effect of assessment, and the unit by assessment interaction and (c) when unit by assessment interactions were significant, tests of simple effects were performed to examine the differences between the two assessments within each unit separately. Lastly, data analysis was performed in collaboration with an experienced statistician and lecturer in the School of Nursing. Dr. Steven Paul, who provided guidance for the data analysis, possessed the requisite expertise to supervise all of the statistical analyses in this study.

References

Chi, Y. M., Jung, T-P., & Cauwenberghs, G. (2010). Dry-contact and noncontact biopotential electrodes: Methodological review. *IEEE Reviews in Biomedical Engineering, 3*,106–118.

Covidien (2013a). *ECG monitoring false alarm "leads-off" clinical practice bundle*. Retrieved from http://www.kendallhq.com/imageServer.aspx?contentID=35771 &contenttype=application/pdf

General Electric Healthcare (2007) *Solar 8000 i/M: Patient monitor operator's manual software Version 5*. Waukesha, WI: GE Healthcare.

General Electric Healthcare (2008). *Solar 8000i v5: Participant guide.* Waukesha, WI: GE Healthcare.

Graham, K., & Cvach, M. (2010). Monitor alarm fatigue: Standardizing use of physiological monitors and decreasing nuisance alarms. *American Journal of Critical Care*, *19*(1), 28–34.

Gross, B., Dahl, D., & Nielsen, L., (2011). Physiologic monitoring alarm load on medical/surgical floors of a community hospital. *Biomedical Instrumentation & Technology, 45*(1), 29–36.

Hanish, H. M., Neustein, R. A., Van Cott, C. C., & Sanders, R. T. (1971). Technical aspects of monitoring the heart rate of active persons. *The American Journal of Clinical Nutrition, 24*(9), 1155–1163.

Hilbe, J. M. (2011). *Negative Binomial Regression*. Cambridge, UK: Cambridge University Press.

International Electrotechnical Commission (2006). Medical electrical equipment—Part 1–8: General requirements for basic safety and essential performance – Collateral Standard: *General requirements, tests and guidance for alarm systems in medical electrical equipment and medical electrical systems* IEC 60606–1–8.

Melendez, L. A., & Pino, R. M. (2012). Electrocardiogram interference: A thing of the past? *Biomedical instrumentation & Technology, 46*(6), 470–477.

Mirvis, D. M., Berson, A. S., Goldbergeer, A. L., Green, L. S., Heger, J. J., Hinohara, T., . . . Wagner, G. S. (1989). Instrumentation and practice standards for electrocardiographic monitoring in special care units. *Circulation, 79,* 464–471.

Philips Healthcare (2008). *Improving ECG quality: Application note.* Retrieved from http://incenter.medical.philips.com/doclib/enc/fetch/2000/4504/577242/577243/577245/577817/577869/Improving_ECG_Quality.pdf%3fnodeid%3d8616008%26vernum%3d1

Polit, D. F., & Beck, C. (2004). *Nursing research: Principles and methods* (7th ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Shott, S. (1990). *Statistics for health professionals*, Philadelphia, PA: W. B. Saunders.

Smith, M. (1984). Rx for ECG monitoring artifact. *Critical Care Nurse, 4*(1), 64–66.

Taenzer, A. H., Pyke, J. B., McGrath, S. P., & Blike, G. T. (2010). Impact of pulse oximetry surveillance on rescue events and intensive care unit transfers. *Anesthesiology, 112*(2), 282–287.

Tronstad, C., Krogh Johnsen, G., Grimnes, S., & Martinsen, Ø. G. (2010). A study on electrode gels for skin conductance measurements. *Physiological Measurement, 31*, 1395–1410.

Turkmen, A., & Pantiskas, C. (2011). Management of ECG cables and leadwires. *Biomedical Instrumentation & Technology, 45*(2), 130–133.

University of California, San Francisco (UCSF) Medical Center. (2013). *Cardiac monitoring (adults), Nursing Procedure*. Retrieved from http://manuals.ucsfmedicalcenter.org/NursingDept/AdultProcedures/PDFsafter12-29-2003/CardiacMonitoringAdult.pdf

Welch, J. (2011). An evidence-based approach to reduce nuisance alarms and alarm fatigue. *Biomedical Instrumentation & Technology*, *45*(1), 46–52.

Chapter 5

Results

Patient Demographics

A total of 429 patients were treated in two NICUs over the two 31-day assessment periods. In the experimental unit (11 NICU), 124 patients were treated during Assessment 1, and 120 patients were treated during Assessment 2. In the control unit (8 NICU), 89 patients were treated during Assessment 1, and 96 patients were treated during Assessment 2 (see Table 5.1). The sample size was larger than originally predicted (429 vs. 317); in the study units, the larger size was related to higher patient volume during both assessment periods. Mean age did not differ between the experimental group (M = 58 years; n = 244) and the control group (M = 58 years, n = 185). The one-way ANOVA found no differences in mean age among the four groups, F = .603, df = 3, p < .613. The overall sample consisted of slightly more women than men (51% vs. 49%, respectively); in both units, more women were hospitalized during Assessment 1 than during Assessment 2. However, the chi-square test found no meaningful differences in gender among the units over time, χ^2 = 3.30, df = 4, p < .348.

During Assessment 1 and Assessment 2, the study units did not differ in regards to patient ethnicity. A large proportion of patients reported being not Hispanic or Latino (n = 389; 90%), a smaller number of patients identified themselves as Hispanic or Latino (39; 9%), and a single patient's ethnicity is unknown. Similarly, there was no difference among the study units in terms of racial demographics. The majority of patients were Caucasian (n = 284; 66%), with other race–ethnicities reported as Asian (n = 62; 14%), Black/African American (n = 20; 5%), Native Hawaiian/Pacific Islander (n = 8; 2%), and American Indian/Alaska Native (n = 1; < 1%). The remaining patients declined to state their race or it was unknown (n = 54; 12%). The racial composition was representative of the San Francisco Bay Area community. The chi-square test

found no meaningful differences in race among the 4 groups over time, χ^2 =8.35, *df* = 12, *p* < .758.

More than half of the patients were hospitalized for neurosurgical disease (*n* = 286; 68%), with other reported primary medical services treating neurovascular patients (*n* = 56, 13%), neurological patients (*n* = 23; 5%), and "off-service" patients (e.g., medicine, general surgery, orthopedic; *n* = 64; 15%). Notably, during assessment periods, the patients' primary medical services differed; a difference was found in the patients' primary medical services over assessment periods, χ^2 =17.55, *df* = 9, *p* < 0.41. Analysis of this difference found that during Assessment 2, the proportion of neurosurgical patients in the experimental unit and the control unit increased equally; neurosurgical patient volume increased in August 2013.

With regard to patient discharge status, the majority of patients from both the experimental unit and the control unit were discharged home after hospitalization (*n* = 246; 57%). Other discharge status changes included transfer to rehabilitation–skilled nursing facility (*n* = 117; 27%); transfer to another acute care facility (*n* = 39; 9%); in-hospital death (*n* = 24; 6%); and transfer to hospice (*n* = 3; < 1%). Again, among the four groups over assessment periods, no differences were detected in the patients' discharge disposition, χ^2 =13.291, *df* = 12, *p* < .348. Because no significant differences were found in the units over time, control for the above demographic variables was not needed.

Of note, three patients were not included in the sample during Assessment 1. Specifically, in the experimental unit, one patient was excluded because the providers requested that physiologic monitoring be discontinued; in the control unit, two patients did not have any recorded physiological monitoring related to a service interruption with the BedMasterEx

clinical application. In the above instances, no alarm or physiologic monitoring data (i.e., alarms or waveforms) were available for analysis.

Table 5.1. Patient Demographics

Variables	Control Unit		Experimental Unit		Total	*p*
	Assessment 1	Assessment 2	Assessment 1	Assessment 2		
	ADC: 9.6	ADC: 10.5	ADC: 14.2	ADC: 13.7		
Patients (Pts)	89	96	124	120	429	
Gender						.348
Females	51 (57%)	47 (49%)	67 (54%)	55 (46%)	220 (51%)	
Males	38 (43%)	49 (51%)	57(46%)	65 (54%)	209 (49%)	
Mean age (years)	59 (21–94 yrs) *SD* = 17	57 (19–87 yrs) *SD* = 16	59 (23–94 yrs) *SD* = 17	57 (18–91 yrs) *SD* = 18	58 (18–94 yrs) *SD* = 17	.613
Ethnicity						.072
Not Hispanic or Latino	78 (88%)	84 (87%)	111 (90%)	116 (97%)		
Hispanic or Latino	10 (11%)	12 (13%)	13 (10%)	4 (3%)		
Unknown/Not Reported	1 (1%)	0	0	0		
Race						.758
American Indian /Alaska Native	0	0	0	1 (<1%)	1 (<1%)	
Asian	12 (13%)	11 (11%)	17 (14%)	22 (18%)	62 (14%)	
Black or African American	4 (4%)	3 (3%)	8 (6%)	5 (4%)	20 (5%)	
Native Hawaiian/Pacific Islander	1 (1%)	1 (1%)	4 (3%)	2 (2%)	8 (2%)	
Caucasian	62 (70%)	67 (70%)	79 (64%)	76 (63%)	284 (66%)	
Unknown/decline to state	10 (11%)	14 (15%)	16 (13%)	14 (12%)	54 (12%)	
Primary Service						.041
Neurology	2 (2%)	5 (5%)	9 (7%)	7 (6%)	23 (5%)	
Neurosurgery	55 (62%)	72 (75%)	72 (58%)	87 (73%)	286(67%)	
Neurovascular	14 (16%)	11 (11%)	16 (13%)	15 (12%)	56 (13%)	
Other	18 (20%)	8 (9%)	27 (22%)	11 (9%)	64 (15%)	
Discharge Status						.348
Home	57 (64%)	59 (62%)	65 (52%)	65 (54%)	246 (57%)	
Rehab/SNF	22 (25%)	27 (28%)	30 (24%)	38 (31%)	117 (27%)	
Transfer to acute care	6 (7%)	7 (7%)	18 (15%)	8 (7%)	39 (9%)	
In-hospital death	3(3%)	3 (3%)	10 (8%)	8 (7%)	24 (6%)	
Hospice	1 (1%)	0	1 (<1%)	1 (< 1%)	3 (<1%)	

Cardiac Monitoring Time

During Assessment 1 and Assessment 2, an average of 79.73 hrs of continuous monitoring data was collected from each patient (*SD* = 100.65, range: 1.93–694.60 hrs). In total, 34,206.10 hrs of continuous cardiac monitoring data were collected (see Table 5.2). Between-unit mean monitoring time did not differ, $F = .243$, $p < .866$ (see Table 5.3).

Table 5.2. Total Cardiac Monitoring Time in Hours for Patients in the Experimental and Control Unit

	Control Unit Assessment 1	Control Unit Assessment 2	Experimental Unit Assessment 1	Experimental Unit Assessment 2	Total
Hrs	6,816.30	7,223.05	9,856.46	10,310.20	34,206.10

Table 5.3. Mean Cardiac Monitoring Time in Hours for Patients in the Experimental and Control Unit

Unit	*N* *Pts*	*M*	*SD*	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower bound	Upper bound		
Control Unit Assessment 1	89	76.58	96.44	10.22	56.27	96.90	2.42	516.58
Experimental Unit Assessment 1	124	79.48	97.88	8.79	62.08	96.88	2.58	608.55
Control Unit Assessment 2	96	75.24	85.24	8.69	57.96	92.51	4.53	519.48
Experimental Unit Assessment 2	120	85.91	117.43	10.72	64.69	107.14	1.93	694.60
Total	429	79.73	100.65	4.85	70.18	89.28	1.93	694.60

Compliance with New Default SpO_2 Low-limit Alarm Settings

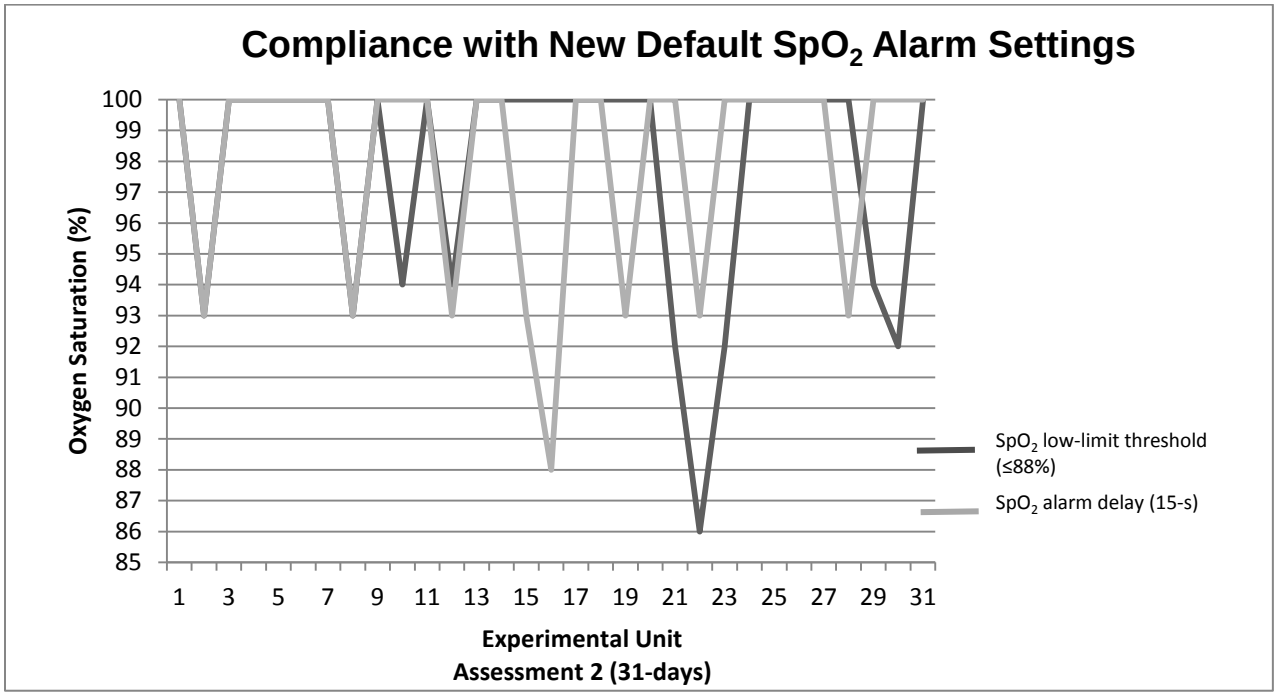
During Assessment 2, the RN investigators performed daily audits in the experimental unit to verify the nurses' compliance with the new default SpO_2 low-limit alarm threshold settings—SpO_2 low-limit threshold alarm to less than or equal to 88% and SpO_2 15-s alarm

delay. The patients' default SpO_2 alarms settings were recorded from the CIC Pro Clinical Information Center and the GE Solar 8000i physiologic monitors. The analysis revealed that 98% of the time, patients in the experimental unit had their SpO_2 low-limit threshold alarm set to less than or equal to 88% and a SpO_2 15-s alarm delay (see Table 5.4; Figure 5.1). In addition, 99% of the time, the SpO_2 alarm was set to the default alarm severity level (i.e., advisory).

Table 5.4. Compliance with New SpO_2 Default Alarm Settings

Date	Patients (*n*)	Compliance with default SpO_2 low-limit 88% (n)	Compliance with default SpO_2 low-limit 88% (%)	Compliance with default notification level (*n*)	Compliance with default notification level	Compliance with default alarm delay 15–s (*n*)	Compliance with default alarm delay 15–s (%)
8/1/13	15	15/15	100	15/15	100	14/14	100
8/2/13	15	14/15	93	14/15	93	14/15	93
8/3/13	16	16/16	100	16/16	100	16/16	100
8/4/13	14	14/14	100	14/14	100	13/13	100
8/5/13	14	14/14	100	14/14	100	13/13	100
8/6/13	16	16/16	100	16/16	100	15/15	100
8/7/13	16	14/14	100	14/14	100	14/14	100
8/8/13	16	14/15	93	15/15	100	13/14	93
8/9/13	16	16/16	100	16/16	100	15/15	100
8/10/13	16	15/16	94	16/16	100	16/16	100
8/11/13	16	16/16	100	16/16	100	15/15	100
8/12/13	16	15/16	94	16/16	100	13/14	93
8/13/13	14	13/13	100	13/13	100	13/13	100
8/14/13	13	13/13	100	13/13	100	13/13	100
8/15/13	16	15/15	100	15/15	100	14/15	93
8/16/13	16	16/16	100	16/16	100	14/16	88
8/17/13	15	15/15	100	15/15	100	15/15	100
8/18/13	14	14/14	100	14/14	100	13/13	100
8/19/13	14	14/14	100	14/14	100	13/14	93
8/20/13	16	16/16	100	16/16	100	16/16	100
8/21/13	12	11/12	92	12/12	100	12/12	100
8/22/13	14	12/14	86	14/14	100	13/14	93
8/23/13	13	12/13	92	13/13	100	12/12	100
8/24/13	9	8/8	100	8/8	100	8/8	100
8/25/13	9	9/9	100	9/9	100	9/9	100
8/26/13	9	9/9	100	9/9	100	9/9	100
8/27/13	16	16/16	100	16/16	100	15/15	100
8/28/13	16	15/15	100	15/15	100	14/15	93
8/29/13	16	15/16	94	16/16	100	16/16	100
8/30/13	14	12/13	92	13/13	100	13/13	100
8/31/13	16	16/16	100	16/16	100	16/16	100
Total	448	430/440	98	439/440	100	419/428	98

Figure 5.1. Compliance with New Default SpO_2 Alarm Settings



Compliance with Novel Electrode-site Skin Preparation and Daily Use of Ag/AgCl–Foam, Pre-gelled—Wet ECG Electrodes

During Assessment 2, RN investigators performed daily observations in the experimental unit to assess the nurses' compliance using the ECG skin preparation paper, the application and dating of Ag/AgCl, foam, pre-gelled *wet* ECG electrodes, and proper electrode placement to improve electrical signal and reduce motion artifact. The analysis found that on average, RNs complied with the new ECG electrode regimen 85% of the time; see Table 5.5.

Of note, the investigators observed that 100% compliance with the electrode-site skin preparation and daily use of Ag/AgCl–foam, pre-gelled *wet* ECG electrodes was achieved on only 4 calendar days during the 31-day intervention—specifically, three times during the first week of the intervention (perhaps when staff was more attentive due to the novelty of the practice), and once during the third week. On one occasion, the day after the start of the

intervention, compliance fell to 56%. Furthermore, investigators observed that compliance declined to 60%–70% compliance on three occasions: once during the third week and twice during the last 2 days of the study (see Figure 5.2). Although unit RNs did not comply with the prescribed ECG electrode application regimen (i.e., when a patient's ECG electrodes were observed to not have been changed daily and had the prior day's date), upon observation, the investigators or the staff nurse (if available) immediately performed the skin preparation and dated and applied new ECG electrodes, using the study supplies.

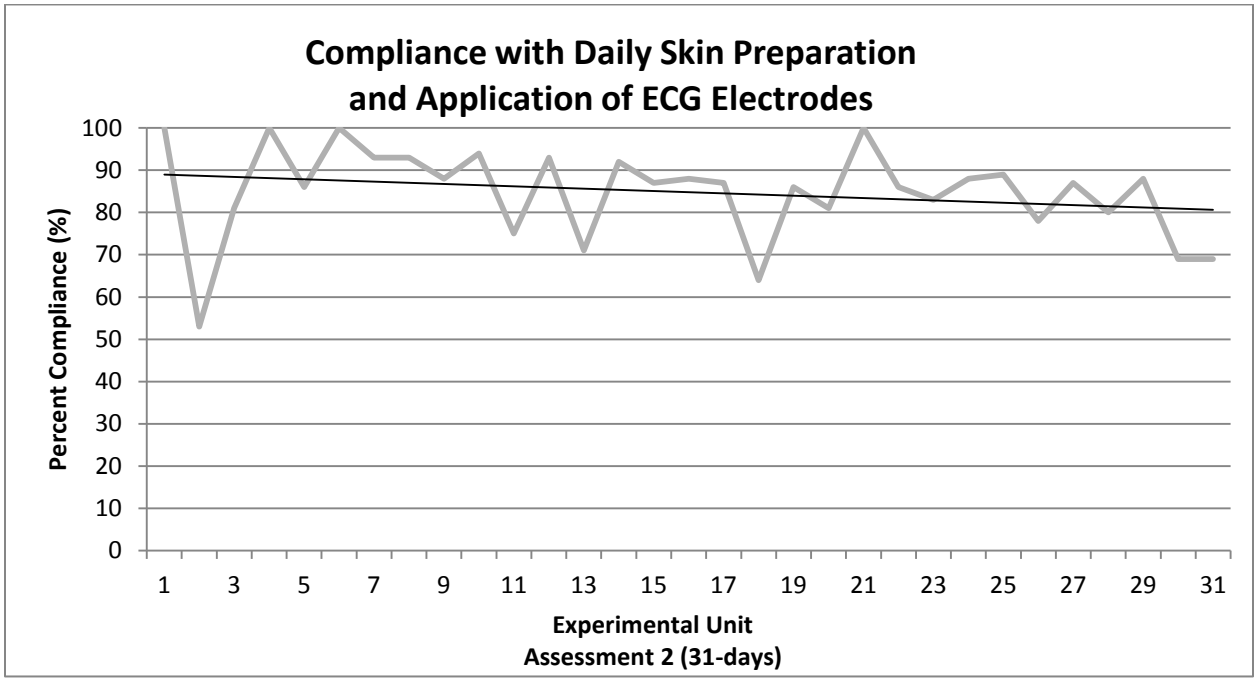
Skin issues were found in two patients during the intervention (see Table 5.5). One patient developed a minor reaction to the electrode adhesive (not to the conductive gel); therefore, the use of the fine abrasive skin preparation paper and Ag/AgCl–foam, pre-gelled *wet* ECG electrodes was stopped for this patient. However, investigators continued to observe that this particular patient's usual ECG electrodes were changed and dated daily. The second patient had naturally fragile skin, so the use of the ECG skin preparation paper was discontinued; nevertheless, the use of the Ag/AgCl–foam, pre-gelled *wet* ECG electrodes continued during her ICU stay without incident.

Table 5.5. Compliance with Novel Skin Preparation and Daily Application of ECG Electrodes

Date	Patients (*n*)	Compliance with dated ECG electrodes (*n*)	Compliance with dated ECG electrodes (%)	Patient days with skin issues (*n*)
8/1/13	15	15/15	100	0/15
8/2/13	15	8/15	53	0/15
8/3/13	16	13/16	81	0/16
8/4/13	14	13/13	100	0/13
8/5/13	14	12/14	86	0/14
8/6/13	16	16/16	100	0/16
8/7/13	16	14/15	93	0/15
8/8/13	16	14/15	93	1/16
8/9/13	16	14/16	88	1/16
8/10/13	16	15/16	94	0/16
8/11/13	16	12/16	75	0/16
8/12/13	16	13/14	93	0/15
8/13/13	14	10/14	71	0/14
8/14/13	13	12/13	92	0/13
8/15/13	16	13/15	87	0/15
8/16/13	16	14/16	88	0/16
8/17/13	15	13/15	87	0/15
8/18/13	14	9/14	64	0/14
8/19/13	14	12/14	86	0/14
8/20/13	16	13/16	81	0/16
8/21/13	12	12/12	100	0/12
8/22/13	14	12/14	86	1/14
8/23/13	13	10/12	83	1/12
8/24/13	9	7/8	88	1/8
8/25/13	9	8/9	89	1/8
8/26/13	9	7/9	78	1/9
8/27/13	16	13/15	87	1/16
8/28/13	16	12/15	80	1/15
8/29/13	16	14/16	88	1/16
8/30/13	14	9/13	69	1/13
8/31/13	16	11/16	69	0/16
Total	448	370/437	85	11/439

*2 patients developed skin issues (1 patient for 2 days and 1 patient for 9 days)

Figure 5.2. Compliance with Novel Skin Preparation and Daily Application of ECG Electrodes

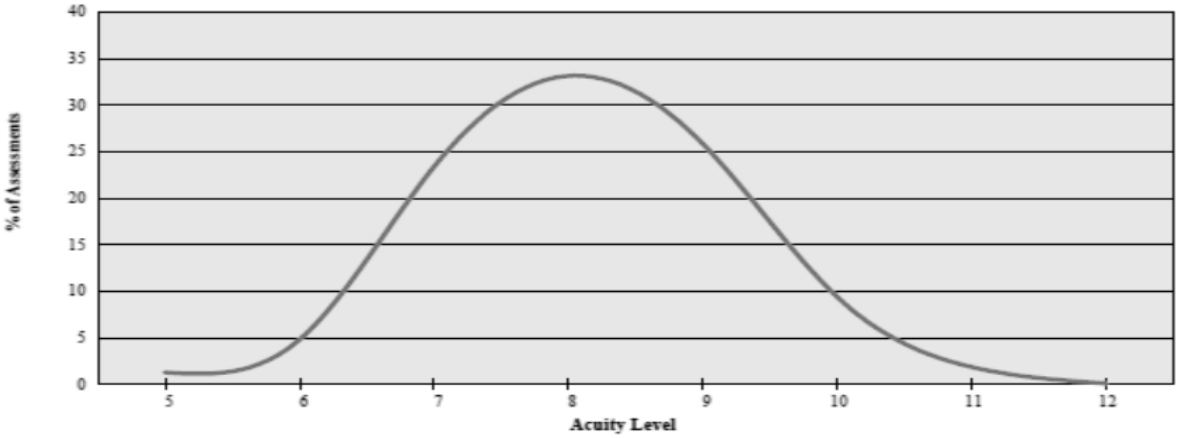


Patient Acuity

UCSF Medical Center uses a patient classification system called Clairvia®. This system is an outcome-driven acuity tool that measures variable nursing care required by acute and critical care patients. Patient acuity levels range from 1 (lowest) to 12 (highest). During our study, RNs assessed patients on both the experimental unit and control unit once every shift and upon observing any change in patient condition; assessments were documented in the electronic medical record. In the experimental unit, RNs assessed patient acuity during Assessment 1 and Assessment 2 as 8.14 (n = 850 assessments) and 8.42 (n = 842 assessments), respectively (see Figures 5.3 and 5.4). Patient acuity was determined to be higher in the control unit than in the experimental unit; during Assessment 1 and Assessment 2. In the control unit, RNs assessed patient acuity during Assessment 1 and Assessment 2 as 10.07 (n = 589 assessments) and 10.06 (n = 610 assessments), respectively (see Figures 5.5 and 5.6). Despite the difference in average

acuity level scores, unit acuity did not change in the experimental unit and control unit among assessment periods and remained relatively stable.

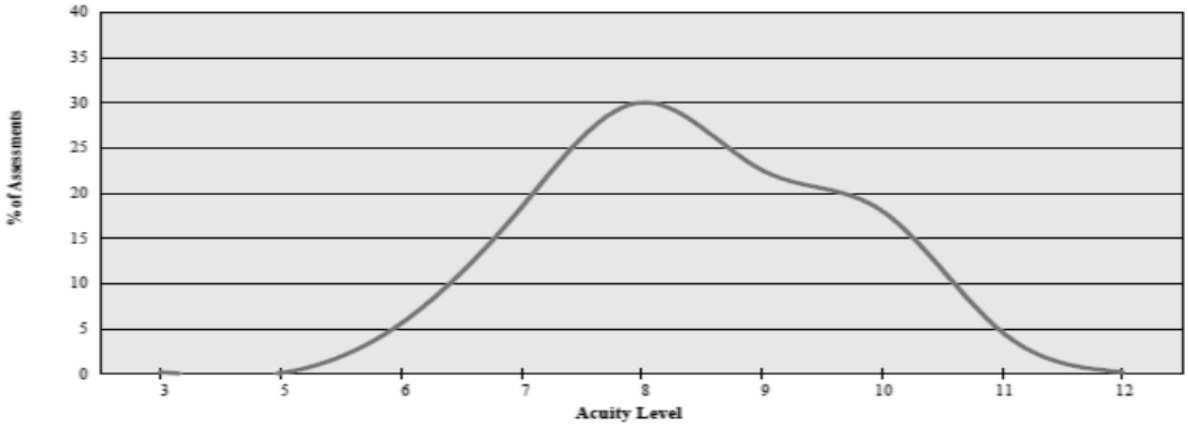
Figure 5.3. Acuity Summary Report for Experimental Unit during Assessment 1



Number of Assessments:	850
Average Acuity Level:	8.14

Acuity	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Assessments	0	0	0	0	11	41	199	282	220	80	16	1
% of Total	0%	0%	0%	0%	1%	5%	23%	33%	26%	9%	2%	0%

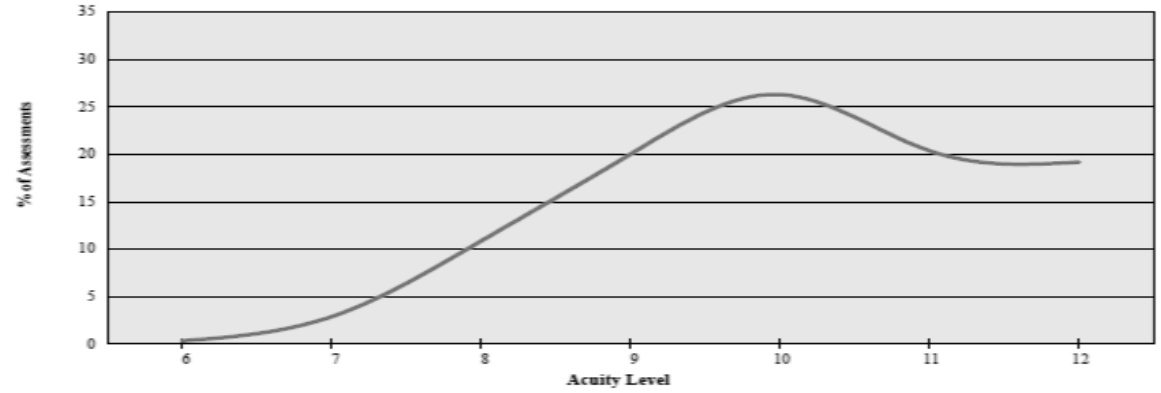
Figure 5.4. Acuity Summary Report for Experimental Unit during Assessment 2



Number of Assessments:	841
Average Acuity Level:	8.42

Acuity	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Assessments	0	0	2	0	1	47	155	253	190	152	39	2
% of Total	0%	0%	0%	0%	0%	6%	18%	30%	23%	18%	5%	0%

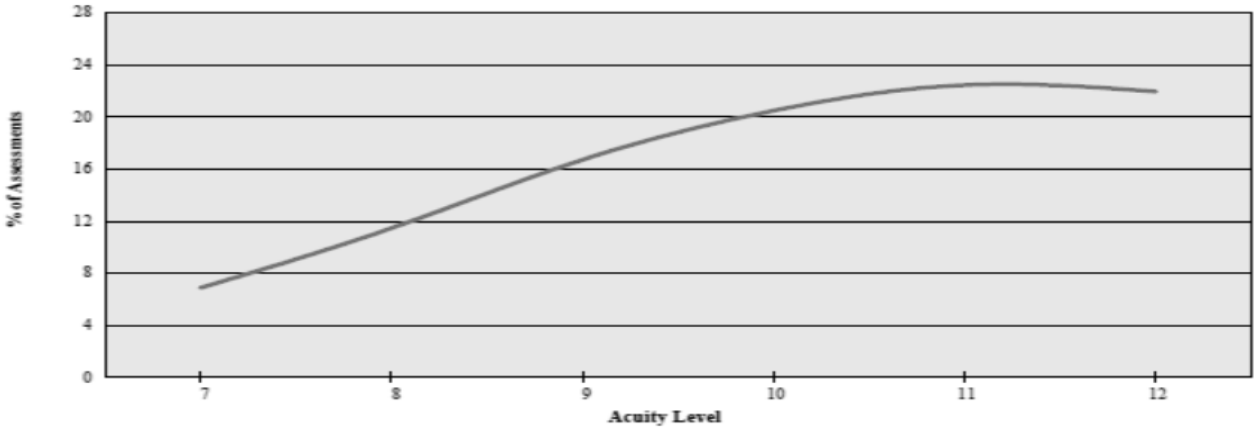
Figure 5.5. Acuity Summary Report for Control Unit during Assessment 1



Number of Assessments:	589
Average Acuity Level:	10.07

Acuity	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Assessments	0	0	0	0	0	2	17	64	118	155	120	113
% of Total	0%	0%	0%	0%	0%	0%	3%	11%	20%	26%	20%	19%

Figure 5.6. Acuity Summary Report for Control Unit during Assessment 2



Number of Assessments:	610
Average Acuity Level:	10.06

Acuity	1	2	3	4	5	6	7	8	9	10	11	12
Assessments	0	0	0	0	0	0	42	70	102	125	137	134
% of Total	0%	0%	0%	0%	0%	0%	7%	11%	17%	20%	22%	22%

Patient Outcome Data

As predicted, the experimental unit reported few adverse patient events during 6 months preceding and 6 months following implementation of the new default SpO_2 low-limit alarm threshold settings (i.e., SpO_2 low-limit threshold alarm to less than or equal to 88% and SpO_2 15-s alarm delay). Similar findings were observed in the control unit. Review of baseline events (February 2013–July 2013) revealed that the experimental unit reported four instances of cardiopulmonary arrest (CPA; CPA rate = 0.27/1,000 patient discharges) and two events of acute respiratory compromise (ARC: ARC rate = 0.13/1,000 patient discharges); the control unit had three instances of CPAs (CPA rate = 0.20/1,000 patient discharges) and zero events of ARC; see Table 5.6 and 5.7). Following implementation of the new default SpO_2 low-limit alarm threshold

settings (August 2013–January 2014), the experimental unit had two instances of CPAs (CPA rate of 0.13/1,000 patient discharges and one event related to ARC (ARC rate = 0.06/1,000 patient discharges). The control unit had one instance of CPA (CPA rate = 0.06/1,000 patient discharges) and zero events of ARC. In review of the reported events, no significant differences were observed between Assessment 1 and Assessment 2 in the experimental unit and control unit.

Table 5.6. Patient Outcomes for Experimental Unit (February 2013 to January 2014)

	Feb 13	Mar 13	Apr 13	May 13	Jun 13	Jul 13	Aug 13	Sep 13	Oct 14	Nov 14	Dec 14	Jan 14
Total hospital patient discharges	2,389	2,601	2,485	2,441	2,552	2,534	2,734	2,509	2,658	2,463	2,509	2,611
Cardiopulmonary Arrests (CPA)	2	1	1	0	0	0	0	1	0	1	0	0
Acute Respiratory Compromise (ARC)	0	0	2	0	0	0	0	0	0	0	1	0
CPA /1,000 patient discharge	0.84	0.38	0.40	0.00	0.00	0.00	0.00	0.40	0.00	0.40	0.00	0.00
ARC/1,000 patient discharges	0.00	0.00	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.00

Table 5.7. Patient Outcomes for Control Unit (February 2013 to January 2014)

	Feb 13	Mar 13	Apr 13	May 13	Jun 13	Jul 13	Aug 13	Sep 13	Oct 14	Nov 14	Dec 14	Jan 14
Total hospital patient discharges	2,389	2,601	2,485	2,441	2,552	2,534	2,734	2,509	2,658	2,463	2,509	2,611
Cardiopulmonary Arrests (CPA)	0	0	0	2	0	1	0	0	1	0	0	0
Acute Respiratory Compromise (ARC)	0	0	0	0	0	0	0	0	0	0	0	0
CPA /1,000 patient discharge	0.00	0.00	0.00	0.82	0.00	0.39	0.00	0.00	0.38	0.00	0.00	0.00
ARC/1,000 patient discharges	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

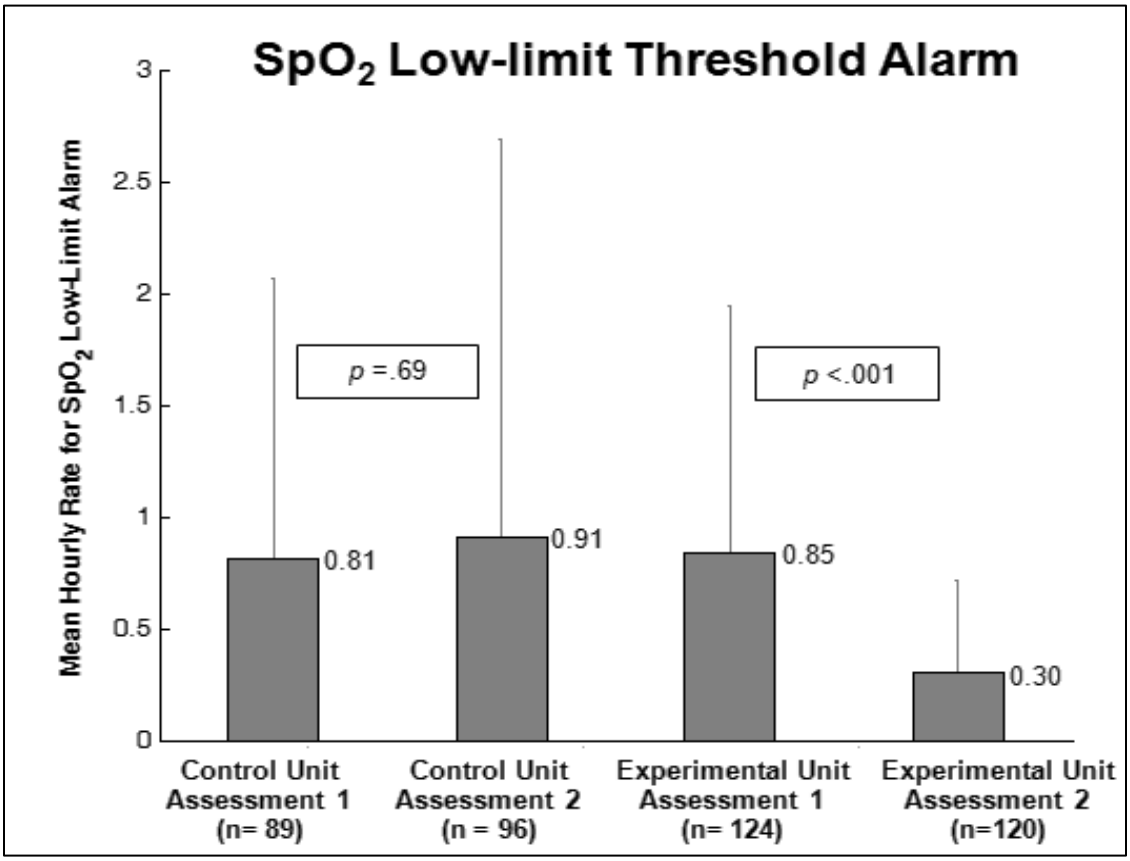
Research Question 1. Is the difference in mean hourly SpO_2 low-limit threshold alarm rate rates between Assessment 1 and Assessment 2 different for the experimental unit (SpO_2 low-limit ≤ 88% with a 15-s alarm delay) compared with the control unit (SpO_2 low-limit ≤ 90% with a 5-s alarm delay)?

For mean hourly rate of SpO_2 low-limit threshold alarms, the unit-by-assessment interaction was statistically significant $p < .001$(see Table 5.8). The difference between Assessment 1 and Assessment 2 for the experimental unit was not the same as the difference between Assessment 1 and Assessment 2 for the control unit. Because the interaction was significant, the simple effects were tested; the test found that at Assessment 1, the experimental unit had a mean hourly rate of SpO_2 low-limit threshold alarms of .85; during Assessment 2, the mean hourly SpO_2 low-limit threshold alarm rate decreased significantly to .30 $p < .001$ (see Figure 5.7). For the control unit, at Assessment 1 and Assessment 2, the mean hourly SpO_2 low-limit threshold alarm rate was .81, and then increased slightly to .90. However, this change was not significant ($p = .69$). The prevalence of SpO_2 low-limit threshold alarms for patients hospitalized in the experimental unit and control unit during Assessment 1 and Assessment 2 is provided in Table 5.9.

Table 5.8. Mean Hourly SpO_2 Low-Limit Threshold Alarm Rates for Patients in the Experimental and Control Unit

Alarm	Control Unit Assessment 1 *N*= 89 patients	Control Unit Assessment 2 *N*= 96 patients	Experimental Unit Assessment 1 *N*= 124 patients	Experimental Unit Assessment 2 *N*= 120 patients	Unit-by-assessment interaction
SpO_2 low-limit threshold	*M* (*SD)*	*M* (*SD)*	*M (SD)*	*M (SD)*	*p*
	0.81(± 1.30)	0.90 (± 1.78)	0.85 (± 1.10)	0.30 (±0.41)	.001

Figure 5.7. Mean Hourly Rates for SpO_2 Low-Limit Alarms



Unit-by-assessment interaction $p < .001$

Table 5.9. Prevalence of SpO_2 Low-Limit Threshold Alarms for the Experimental and Control Unit

Alarm	Control Unit Assessment 1	Control Unit Assessment 2	Experimental Unit Assessment 1	Experimental Unit Assessment 2	Total
SpO_2 low-limit threshold	5,536	4,029	7,627	2,970	20,162

Research Question 2. Is the difference in mean hourly technical alarm rates (e.g., ECG, artifact, arrhythmia suspend) between Assessment 1 and Assessment 2 different for the experimental unit (in which patients received a novel ECG skin site preparation technique and daily applications of high-quality ECG electrodes) compared with the control unit (whose patients received usual care)?

ECG lead fail alarm

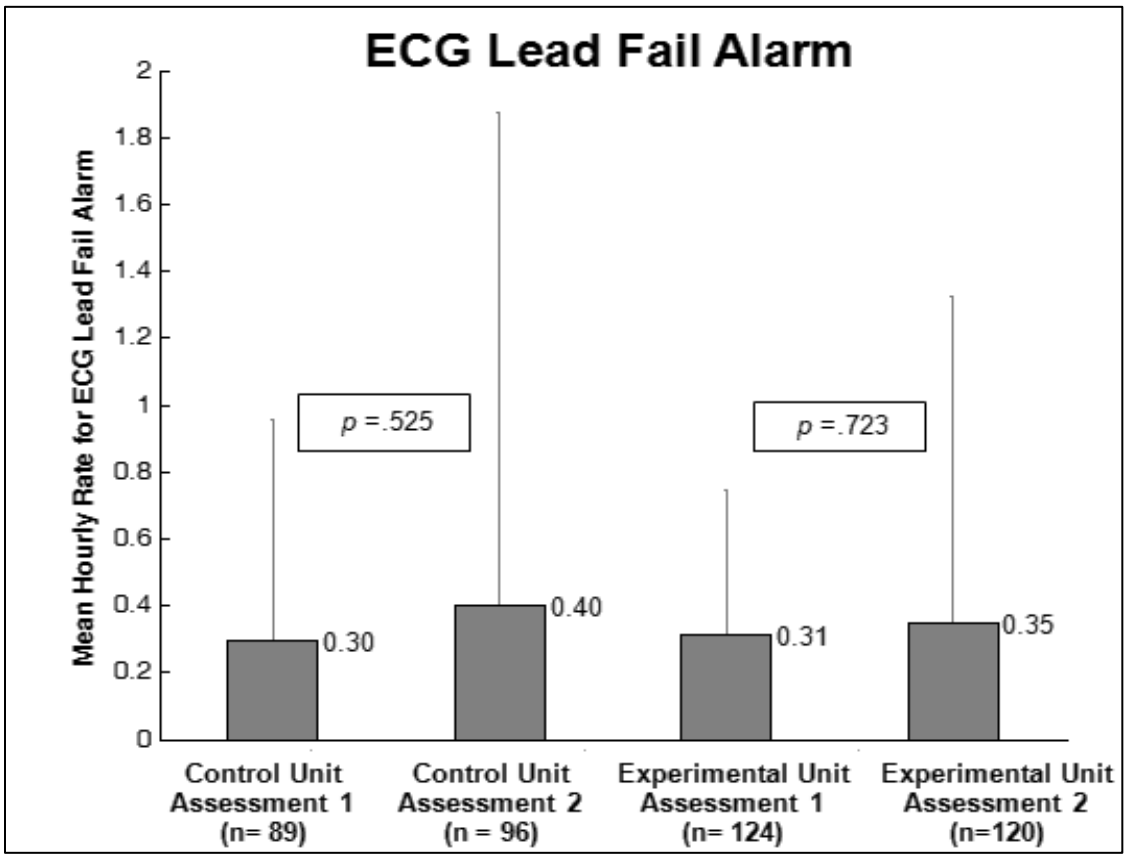
For mean hourly rate of ECG lead fail alarms, analysis found no statistically significant unit-by-assessment interaction (p = .741). Although there was no interaction, simple effects were tested; the tests found that during Assessment 1 and Assessment 2, the experimental unit had a mean hourly rate of ECG lead fail alarms of .31 and .35, respectively (p = .723; see Table 5.10). Similarly, during Assessment 1 and Assessment 2, the control unit had a mean hourly ECG lead fail alarm rate of .30; during Assessment 2, this rate increased slightly to .40 (p = .525). Notably, the mean hourly ECG lead fail alarm rate in both units increased slightly from Assessment 1 to Assessment 2 (see Figure 5.8). The prevalence of ECG technical alarms for both the experimental unit and control unit during both assessment periods is reported in Table 5.11. Of note, during Assessment 2, the frequency of ECG leads fails alarms in the experimental unit increased by 8% (see Table 5.11).

Table 5.10. Mean Hourly Technical Alarm Rates for Patients in the Experimental Unit and Control Unit

Alarms	Control Unit Assessment 1 *N* = 89 patients	Control Unit Assessment 2 *N* = 96 patients	Experimental Unit Assessment 1 *N* = 124 patients	Experimental Unit Assessment 2 *N* = 120 patients	Unit-by-assessment interaction
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*p*
ECG lead fail	0.29 (± 0.66)	0.40 (± 1.47)	0.31 (± 0.43)	0.35 (± 0.97)	.741
Artifact*	13.40(± 19.40)	9.41 (± 11.09)	11.64 (± 13.51)	8.48 (± 12.59)	.891
Arrhythmia suspend	0.08 (± 0.20)	0.03 (± 0.06)	0.06 (± 0.14)	0.05 (± 0.09)	.130

*Visual

Figure 5.8. Mean Hourly Rates for ECG Lead Fail Alarms

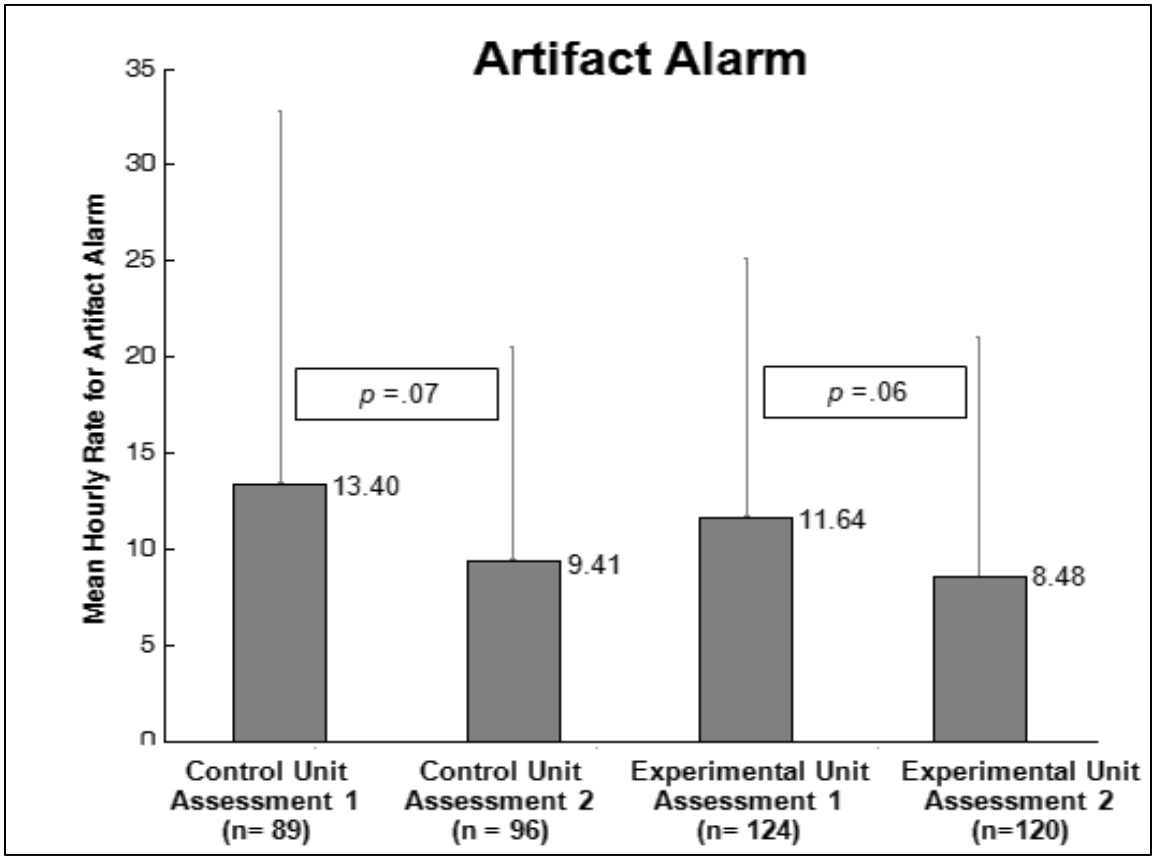


Unit-by-assessment interaction $p = .741$

Artifact alarm

No statistically significant unit-by-assessment interaction for mean hourly artifact alarms, (p = .891) was found. Examination of simple effects revealed that the experimental unit had a mean hourly rate of artifact alarms of 11.64 and 8.48 (p = .061) during Assessment 1 and Assessment 2, respectively (see Table 5.10). During Assessment 1, the control unit had a mean hourly artifact alarm rate of 13.40, which decreased to 9.41 during Assessment 2 (p = .070). Both units had a reduction in the mean hourly artifact alarm rate between Assessment 1 and Assessment 2; however, this reduction was not significant (see Figure 5.9).

Figure 5.9. Mean Hourly Rates for Artifact Alarms

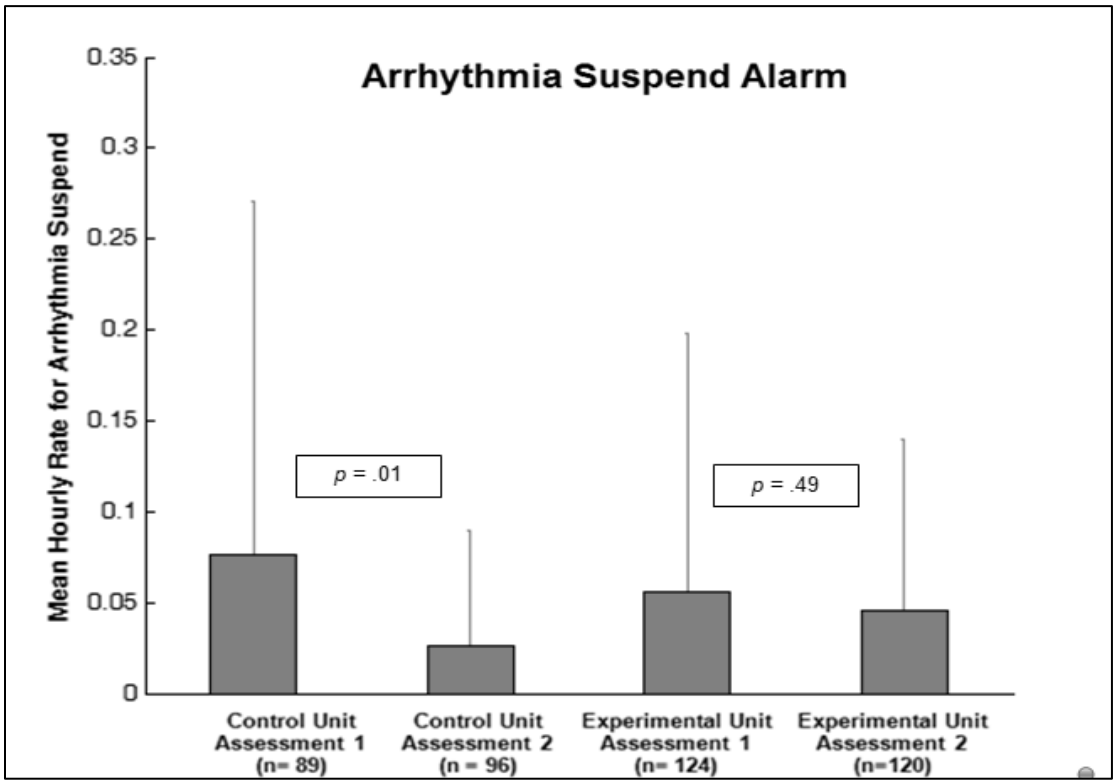


Unit-by-assessment interaction p = .891

Arrhythmia suspend alarm

No unit-by-assessment interaction was found with regard to the mean hourly rate of arrhythmia suspend alarms (p =.130). In the experimental unit, the mean hourly arrhythmia suspend alarm rate was 0.06 during Assessment 1 and decreased to 0.05 during Assessment 2 (p = .49); likewise in the control unit, the mean hourly arrhythmia suspend alarm rate decreased from 0.08 to 0.03 during Assessment 1 and Assessment 2 (p = .01; see Figure 5.10). Although the decrease in the control unit was greater than that in the experimental unit and was significant, no significant unit-by-assessment interaction was observed. Mean hourly arrhythmia suspend alarm rates were small, which indicates that these technical alarms occurred infrequently (i.e., relative to patients' total monitoring hours; see Table 5.10 and Table 5.11).

Figure 5.10. Mean Hourly Rates for Arrhythmia Suspend Alarms



Unit-by-assessment interaction p = .130

Table 5.11. Prevalence of Technical Alarms for Patients in the Experimental Unit and Control Unit

Alarms	Control Unit Assessment 1	Control Unit Assessment 2	Experimental Unit Assessment 1	Experimental Unit Assessment 2	Total
ECG lead fail	2,081	2,067	2,510	2,713	9,371
Artifact*	87,107	63,273	128,002	79,337	357,719
Arrhythmia suspend	383	223	745	675	2,026

*Visual

Research Question 3. Is the difference in mean hourly arrhythmia alarm rates between Assessment 1 and Assessment 2 different for the experimental unit (whose patients received a novel ECG skin site preparation technique and daily applications of high-quality ECG electrodes) compared with the control unit (whose patients received usual care)?

For mean hourly rate of arrhythmia alarms—which includes all types of patient status arrhythmia alarms—the unit-by-assessment interaction was statistically significant (p =.05; see Table 5.12). The Assessment 1–Assessment 2 change in the experimental unit differed from that of the control unit. The list of all patient status arrhythmia alarms is presented in Table 5.13. The interaction was significant: analysis of the simple effects revealed that at Assessment 1, the experimental unit had a mean hourly rate of total arrhythmia alarms of 14.10; during Assessment 2, the mean hourly arrhythmia alarm rate decreased significantly to 5.42 (p < .002; see Table 5.12). During Assessment 1, the control unit had a mean hourly arrhythmia alarm rate of 10.33, which increased moderately to 12.49 during Assessment 2 (p = .707; see Figure 5.11). During

Assessment 2, the reduction in frequency of arrhythmia alarms in the experimental group was significantly greater than that of the control unit (see Table 5.14).

Interestingly, for critical arrhythmia alarms (i.e., accelerated ventricular, asystole, pause, ventricular bradycardia, ventricular tachycardia, and ventricular tachycardia/ventricular fibrillation), the mean hourly rate for the six individual critical arrhythmia alarms—there was no statistically significant unit-by-assessment interaction (see Table 5.15). Likewise, there was no unit-by-assessment interaction for the combined six arrhythmia alarms (p = .273); no reduction in critical arrhythmia alarm rates was found in the experimental unit (see Table 5.15.).

Table 5.12. Mean Hourly Arrhythmia Alarm Rates for Patients in the Experimental Unit and Control Unit

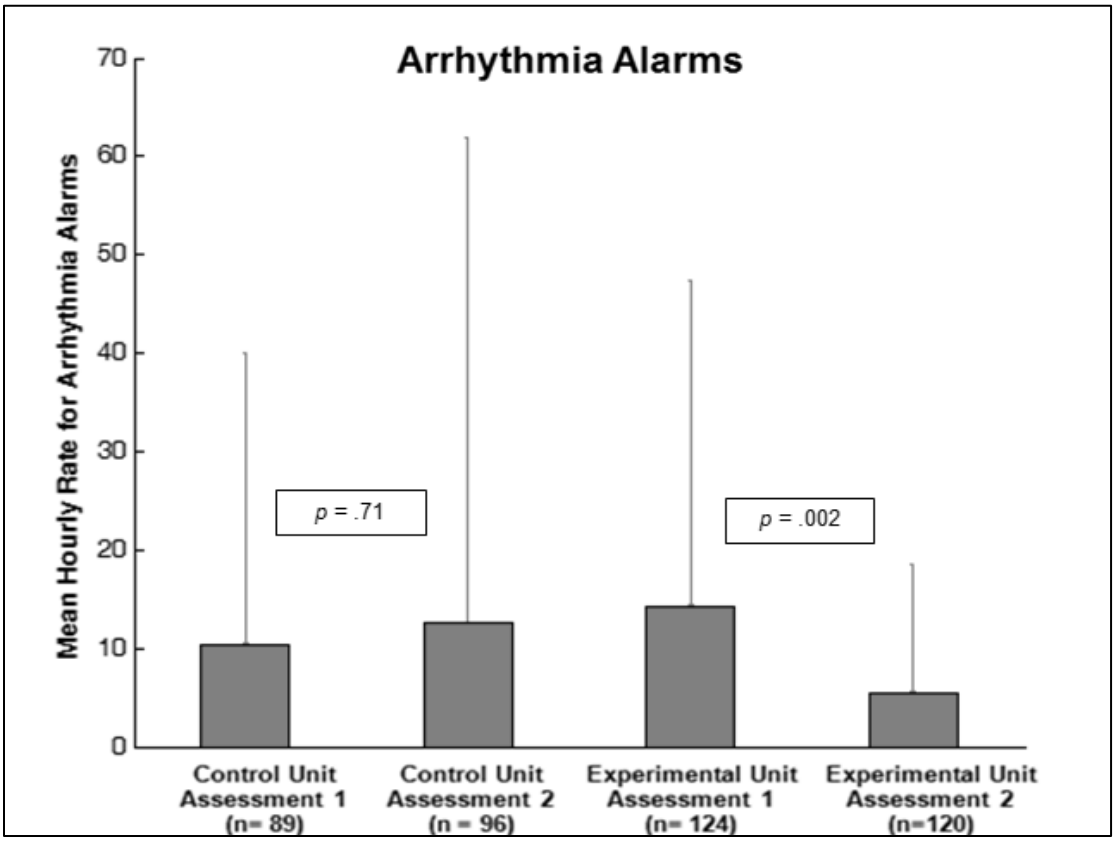
Alarms	Control Unit Assessment 1 N = 89 patients	Control Unit Assessment 2 N = 96 patients	Experimental Unit Assessment 1 N = 124 patients	Experimental Unit Assessment 2 N = 120 patients	Unit-by-assessment interaction
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*p*
Arrhythmia*	10.33 (± 29.62)	12.50 (± 49.49)	14.10 (± 33.30)	5.42 (± 13.07)	.05
All audible**	6.71 (± 3.91)	7.07 (± 5.38)	6.73 (± 4.13)	6.25 (± 4.37)	.316
All alarms**	38.43 (± 39.03)	38.29 (± 57.32)	40.77 (± 38.99)	28.23 (±25.97)	.100

*Includes all types of arrhythmia alarms
** Includes all types of patient and system status alarms

Table 5.13. Types of Patient Status Arrhythmia Alarms and Default Notification Levels

Patient Status Arrhythmia Alarm				
	Crisis	Warning	Advisory	Message
Asystole	x			
Ventricular Fibrillation/Tachycardia	x			
Ventricular Tachycardia	x			
Ventricular Tachycardia ≥ 2			x	
Ventricular Bradycardia		x		
Couplet				x
Bigeminy				x
Accelerated Ventricular		x		
Pause		x		
Trigeminy				x
R on T				x
PVC				x
Tachycardia			x	
Bradycardia			x	
Atrial Fibrillation/Irregular			x	

Figure 5.11. Mean Hourly Rates for Arrhythmia Alarms



Unit-by-assessment interaction $p < .05$

Table 5.14. Prevalence of Alarms for Patients in the Experimental Unit and Control Unit

Alarms	Control Unit Assessment 1	Control Unit Assessment 2	Experimental Unit Assessment 1	Experimental Unit Assessment 2	Total
Arrhythmia*	79,226	68,455	122,233	68,647	338,561
Audible**	42,027	46,580	62,159	56,490	207,256
All Alarms	258,776	227,692	401,758	272,610	1,160,836

*Includes all types of arrhythmia alarms
** Includes all types of patient and system status alarms

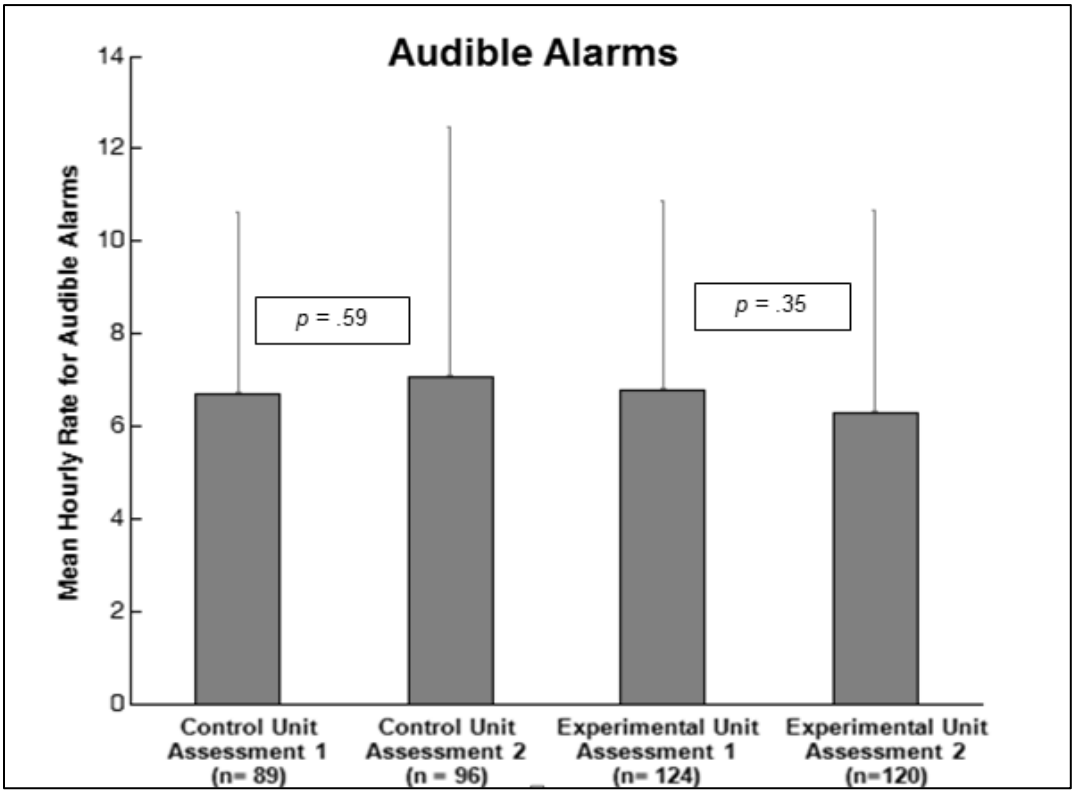
Table 5.15. Mean Hourly Rates for Critical Arrhythmia Alarms for Patients in the Experimental Unit and Control Unit

Alarms	Control Unit Assessment 1 *N* = 89 patients	Control Unit Assessment 2 *N* = 96 patients	Experimental Unit Assessment 1 *N* = 124 patients	Experimental Unit Assessment 2 *N* = 120 patients	Unit–by-assessment interaction
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*p*
Accelerated ventricular	0.018 (± 0.080)	0.004 (± 0.014)	0.005 (± 0.025)	0.016 (± 0.123)	.577
Asystole	0.020 (± 0.132)	0.006 (± 0.019)	0.019 (± 0.145)	0.011 (± 0.034)	.398
Pause	0.012 (± 0.046)	0.008 (± 0.033)	0.027 (± 0.125)	0.031 (± 0.128)	.840
Ventricular Bradycardia	0.006 (± 0.044)	0.002 (± 0.010)	0.001 (± 0.011)	0.002 (± 0.019)	.954
Ventricular tachycardia	0.014 (± 0.032)	0.011 (± 0.033)	0.017 (± 0.057)	0.022 (± 0.196)	.557
Ventricular Fib /Tach	0.000 (± 0.000)	0.000 (± 0.000)	0.000 (± 0.001)	0.000 (± 0.000)	.958
Combined	0.070 (± 0.237)	0.033 (± 0.067)	0.070 (± 0.280)	0.083 (± 0.376)	.273

Research Question 4. Is the difference in mean hourly audible alarm rates between assessment one and assessment two different for the experimental unit (received the alarm management interventions) compared with the control unit (received usual care)?

For the mean hourly rate of audible alarms, the unit-by-assessment interaction was not statistically significant (p = .316). Although no interaction was found, simple effects were examined; at Assessment 1 and Assessment 2, the experimental unit's mean hourly audible alarm rate was 6.73; this rate then slightly decreased to 6.25 (p = .357). However, the control unit's Assessment 1 mean hourly audible alarm rate, 6.71, increased to an Assessment 2 rate of 7.07 (p =. 595; see Table 5.12). The experimental unit had a reduction in the mean hourly rate of audible alarms; however, the reduction was insufficient to substantiate an interaction.

Figure 5.12. Mean Hourly Rates for Audible Alarms



Unit-by-assessment interaction p = .316

Research Question 5. Is the difference in mean hourly rate of all physiologic monitor alarms between Assessment 1 and Assessment 2 different for the experimental unit (received the alarm management interventions) compared with the control unit (received usual care)?

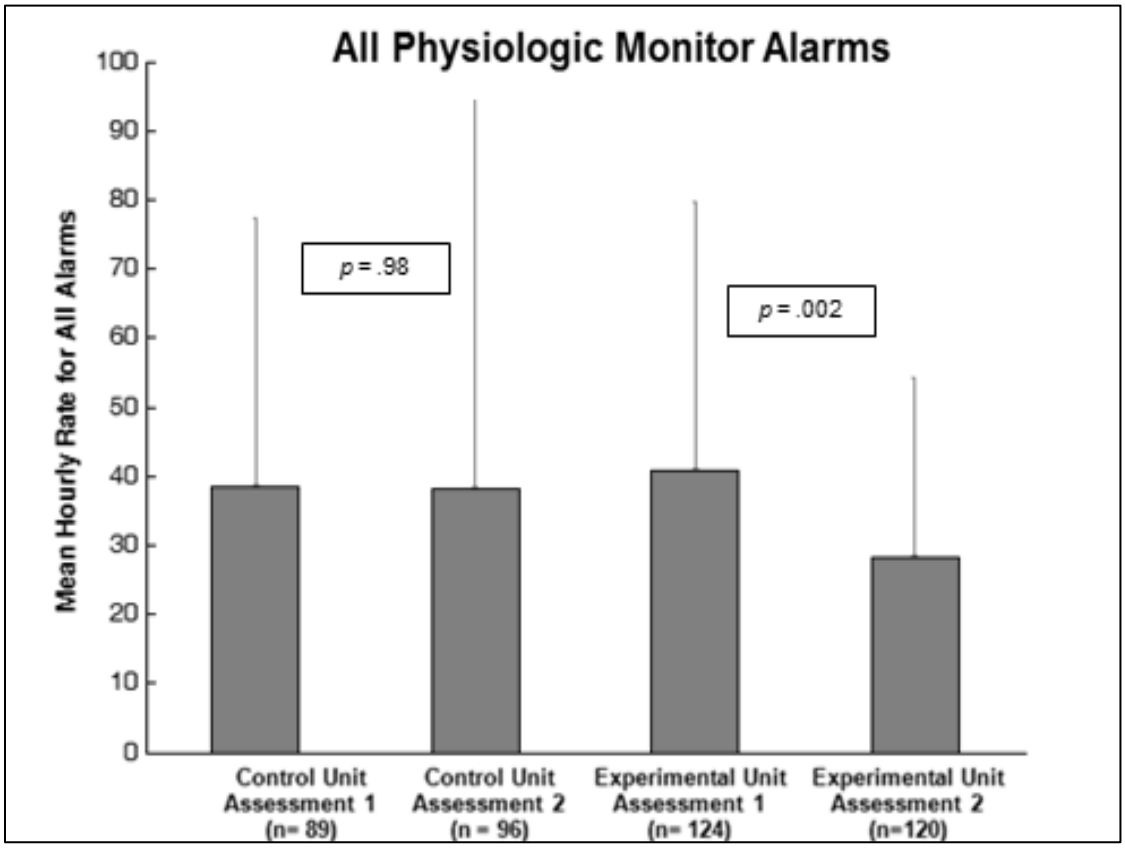
There was no significant unit-by-assessment interaction in the mean hourly rate of all physiologic alarms, which includes patient status (arrhythmia and parameter) alarms and system status (technical) alarms (see Table 5.16). Nonetheless, between Assessment 1 and Assessment 2, the mean hourly rate of all alarms in the experimental unit diminished significantly, from 40.78 to 28.22 (p = .002). Conversely, the control unit's mean hourly rate of all physiologic monitor alarms remained relatively unchanged from Assessment 1 (38.42) to Assessment 2 (38.28; p = .983; see Table 5.12). Although in the experimental unit the mean hourly alarm rate declined significantly, this reduction was insufficient to substantiate a statistically significant interaction (see Figure 5.13). Physiologic monitor alarm counts in both units were reduced over time; however, the reduction was greater in the experimental unit (-129,410) than in the control unit (-31, 084; see Table 5.14).

Table 5.16. Types of GE Physiologic Monitor Alarms

Physiologic Monitor Alarms (GE Solar 8000i)		
Patient Status Alarms		System Status Alarms
Arrhythmia	Parameter (≥ and ≤)	Technical
Asystole	HR	Artifact
Ventricular fibrillation/tachycardia	CO2 No Breath	Arrhythmia Suspend
Ventricular tachycardia	RM No Resp	Arrhythmia Off
Ventricular tachycardia ≥2	PVC	ECG Leads Fail
Ventricular bradycardia	ST	Resp Leads Fail
Couplet	ART	No ECG
Bigeminy	PA	ART Sensor Fail
Accelerated ventricular	CVP	ICP Sensor Fail
Pause	CO2	FEM Sensor Fail
Trigeminy	NBP	RAP Sensor Fail
R on T	NBP M Only	SP Sensor Fail
PVC	SPO2	LAP Sensor Fail
Tachycardia	FEM	CVP Sensor Fail
Bradycardia	UAC	PA Sensor Fail
Atrial fibrillation/ Irregular	RA	UVC Sensor Fail
	NICO	UAC Sensor Fail
	LA	SpO2 Connect Probe
	ICP	SpO2 Probe Off
	SVO2	SpO2 Probe Fail
	TC	SpO2 Low Sig
	ICG	SpO2 Incompatible Cable
	ART Rate	SpO2 Interf Def
	SPO2 Rate	Nbp Invalid Command
	BT	Nbp Excessive Pressure 200
	FEM Rate	Nbp Exceeded 3 min
	TMP	Nbp Deflation Failure
	POC	Nbp Inflation Time Exceeded
	CCO	PA Art Line Disconnect
	Resp no breath	CVP Art Line Disconnect
	RR	ART Art Line Disconnect
		FEM Art Line Disconnect

ART: arterial, CO_2: carbon dioxide, BT: blood temperature, CCO: continuous cardiac output, CVP: central venous pressure; ECG: electrocardiograph, Fem: femoral, HR: heart rate, ICG: impedance cardiography, ICP: intracranial pressure: LA: left atrial, LAP: left atrial pressure: NBP: non-invasive blood pressure: NBP M: non-invasive blood pressure mean: NICO: non-invasive cardiac output, PA: pulmonary artery: PVC: premature ventricular complex, RA: right atrial, RAP: right arterial pressure: RESP: respiration: RM: respiratory mechanics, RR: respiration rate: POC: point of care, SP: special, SpO_2: arterial oxygen saturation from pulse oximetry: ST: interval of ventricular repolarization, S_VO_2: mixed venous oxygen saturation, TC: transcutaneous, TMP: temperature, UAC: umbilical artery catheter, UVC: umbilical venous catheter.

Figure 5.13. Mean Hourly Rates for Physiologic Monitor Alarms



Unit-by-assessment interaction $p = .100$

Research Question 6. Is the difference in the mean percentage of false-positive cardiac arrhythmia alarms (e.g., accelerated ventricular, asystole, pause, ventricular bradycardia, ventricular tachycardia, ventricular tachycardia/ventricular fibrillation) between Assessment 1 and Assessment 2 different for the experimental unit compared with the control unit?

A 2-way ANOVA was performed to determine whether, between Assessment 1 and Assessment 2, the mean percentage of false-positive cardiac arrhythmia alarms (e.g., accelerated ventricular, asystole, pause, ventricular bradycardia, ventricular tachycardia, ventricular tachycardia/ventricular fibrillation) in the experimental unit was different from that in the control unit. As expected, the alarm samples were small because the analysis included only patients who

generated a minimum of one of the six possible types of critical arrhythmia alarms that were analyzed and determined to be either true or false-positive alarms according to the annotation plan (see Appendix A). For the analysis of mean percentage of false-positive arrhythmia alarms, no statistically significant unit-by-assessment interaction was found for each critical arrhythmia alarm (see Table 5.17). That is, the daily skin preparation and application of ECG electrodes did not reduce the mean percentage of false-positive critical arrhythmia alarms in the experimental unit during Assessment 2.

Accelerated ventricular alarms

More than ten patients had accelerated ventricular alarms in both the control unit and experimental unit during each assessment period. Notably, the mean percentage of false-positive accelerated ventricular alarms in the experimental unit increased from 88.88 to 90.00 (p = .931; see Table 5.18). In contrast, the control unit's mean percentage of false-positive accelerated ventricular alarms declined significantly, from 87.07 to 63.63 between Assessment 1 and Assessment 2 (p = .170; See Table 5.19). However, this difference in changes was not sufficient to substantiate an interaction (p = .465).

Asystole alarms

During the two assessment periods, over 20% of all patients (n = 94) had a minimum of one asystole alarm in the study units (experimental units, 56 patients; control units, 36 patients). Asystole alarms were the second most frequently occurring critical arrhythmia alarms (19%) during the study (see Table 5.20). Analysis of the mean percentage of false-positive asystole alarms found no unit-by-assessment interaction (p = .246; see Table 5.17). Tests of simple effects revealed that the experimental unit had a significant increase in mean percentage of false-positive asystole alarms, from 76.53 to 94.11 (p = .046; see Table 5.18). Although this result was

significant, it should be interpreted with caution because the increase in the mean percentage of asystole alarms between Assessment 1 and 2 may be due factors un-related to the intervention (e.g., increased patient disconnection from the physiologic monitor upon discontinuation of continuous monitoring). In contrast, between Assessment 1 and Assessment 2, the control unit had a slight reduction in mean percentage of false-positive asystole alarms, from 94.11 to 89.47; however this result was not significant (p = .627; see Table 5.19).

Pause alarms

During the Assessment 1 and Assessment 2, a similar number of patients generated at least one pause alarm in each study unit. In the experimental unit, 31 patients generated at least one pause alarm for the duration of both assessment periods (see Table 5.17). For the control unit, during Assessment 1 and 2, 21 and 14 patients generated pause alarms, respectively. In our study, pause alarms were the most frequently occurring critical arrhythmia alarm (n = 918; 42% see Table 5.20). The mean percentage of false-positive pause alarm analysis revealed no unit-by-assessment interaction (p = .547; see Table 5.17). Tests of simple effects found that, between Assessment 1 and Assessment 2, the mean percentage of false-positive pause alarms in the experimental unit increased from 78.49 to 87.63; however, this result was also insignificant (p = .315; see Table 5.18). In contrast, the mean percentage of false-positive pause alarms in the control unit decreased from 86.45 to 85.71 however, this result was also insignificant (p = .952; see Table 5.19).

Ventricular bradycardia alarms

During the study, ventricular bradycardia alarms were the second least frequently occurring arrhythmia alarms (67 alarms from 23 patients; experimental unit, 12 patients; control unit, 11 patients) at Assessment 1 and Assessment 2 (see Table 5.17). Analysis of the mean

percentage of false-positive ventricular bradycardia alarms found no unit-by-assessment interaction (p = .277; see Table 5.17). Tests of simple effects found that the mean percentage of false-positive ventricular bradycardia alarms decreased in the experimental unit, from 83.33 to 80.00 (p = .891); still, this result was irrelevant. Between Assessment 1 and Assessment 2, the mean percentage of false-positive ventricular bradycardia alarms increased moderately in the control unit, from 25.00 to 64.29; however, this result was not significant (p = .227; see Table 5.19).

Ventricular tachycardia alarms

Regarding the six critical arrhythmia alarms that were annotated to be either true or false, over 27% of all study patients (n = 114; experimental 60 patients; control 54 patients) had at least one ventricular tachycardia alarms during their ICU hospitalization. Although frequent, no unit-by-assessment interaction was found (p = .813; see Table 5.17 and 5.20). Tests of simple effects found that the mean percentage of false-positive ventricular tachycardia alarms in the experimental unit decreased from 89.08 to 77.27 (p = .197; see Table 5.18); even so, this result was insignificant. Between Assessment 1 and Assessment 2, the mean percentage of false-positive ventricular tachycardia alarms in the control unit decreased slightly, from 82.80 to 79.56; however this result was also unimportant (p = .540; see Table 5.19).

Ventricular fibrillation/tachycardia alarms

Ventricular fibrillation/tachycardia alarms were the least frequently occurring arrhythmia alarms. During Assessment 1 and Assessment 2, in the control unit, no patients generated this alarm; in the experimental unit, two patients generated this alarm (see Tables 5.18 and 5.19). For this reason, a unit-by-assessment interaction was not feasible. During both assessment periods, in the experimental unit, the mean percentage of false-positive ventricular fibrillation/tachycardia

alarms was 100%; however, only two alarms were generated during Assessment 1, and only two alarms were generated during Assessment 2 (see Tables 5.17 and 5.18).

Table 5.17. Mean Percent False-positive Critical Arrhythmia Alarms for Patients in the Experimental Unit and Control Unit

Alarms	Control Unit Assessment 1		Control Unit Assessment 2		Experimental Unit Assessment 1		Experimental Unit Assessment 2		Unit–by–assessment interaction
	N Pts	*M (SD)*	*N Pts*	*M (SD)*	*N Pts*	*M (SD)*	*N Pts*	*M (SD)*	*p*
Accelerated ventricular	15	87.07 (± 34.13)	11	63.63 (± 50.45)	18	88.88 (± 32.33)	10	90.00 (± 31.62)	.465
Asystole	17	94.11 (± 24.25)	19	89.47 (± 31.53)	22	76.53 (± 40.67)	34	94.11 (± 23.88)	.246
Pause	21	86.45 (± 34.13)	14	85.71 (± 36.31)	31	78.49 (± 39.44)	31	87.63 (± 31.02)	.547
Ventricular bradycardia	4	25.00 (± 50.00)	7	64.29 (± 47.56)	7	83.33 (± 37.27)	5	80.00 (± 44.72)	.277
Ventricular tachycardia	29	82.80 (± 34.50)	25	79.56 (± 40.66)	38	89.08 (± 27.30)	22	77.27 (± 42.89)	.813
Ventricular fibrillation/ tachycardia	0	0	0	0	2	100.0 (±.00)	2	100.00 (±.00)	-

Table 5.18. Comparison of Mean Percent False-positive Critical Arrhythmia Alarms for Patients in the Experimental Unit during Assessment 1 and 2

Alarms	Experimental Unit Assessment 1		Experimental Unit Assessment 2		
	N Pts	*M (SD)*	*N Pts*	*M (SD)*	*p*
Accelerated ventricular	18	88.88 (±32.33)	10	90.00 (±31.62)	.931
Asystole	22	76.53 (±40.67)	34	94.11 (±23.88)	.046
Pause	31	78.49 (±39.44)	31	87.63 (±31.02)	.315
Ventricular bradycardia	7	83.33 (±37.27)	5	80.00 (±44.72)	.891
Ventricular tachycardia	38	89.08 (±27.30)	22	77.27 (±42.89)	.197
Ventricular fibrillation/tachycardia	2	100.0 (±.00)	2	100.00 (±.00)	.540
Combined	53	84.14 (±30.77)	61	86.20 (±33.22)	.734

Table 5.19. Comparison of Mean Percent False-positive Critical Arrhythmia Alarms for Patients in the Control Unit during Assessment 1 and 2

Alarms	Control Unit Assessment 1		Control Unit Assessment 2		
	N Pts	*M (SD)*	*N Pts*	*M (SD)*	*p*
Accelerated ventricular	15	87.07 (± 34.13)	11	63.63 (± 50.45)	.170
Asystole	17	94.11 (± 24.25)	19	89.47 (± 31.53)	.627
Pause	21	86.45 (± 34.13)	14	85.71 (± 36.31)	.952
Ventricular bradycardia	4	25.00 (± 50.00)	7	64.29 (± 47.56)	.227
Ventricular tachycardia	29	82.80 (± 34.50)	25	79.56 (± 40.66)	.540
Ventricular fibrillation/tachycardia	0	0	0	0	-
Combined	45	86.80 (± 28.81)	42	76.92 (± 39.88)	.187

Table 5.20. Prevalence of Critical Arrhythmia Alarms for Patients in the Experimental Unit and Control Unit

Types of Alarms	Control Unit Assessment *1*	Control Unit Assessment *2*	Experimental Unit Assessment 1	Experimental Unit Assessment 2	Total Alarms
Accelerated ventricular	199 (34%)	34 (17%)	87 (9%)	26 (5%)	346 (16%)
Asystole	46 (8%)	43 (22%)	251 (27%)	84(18%)	424 (19%)
Pause	194 (33%)	58 (29%)	361 (39%)	305 (65%)	918 (42%)
Ventricular bradycardia	34 (5%)	9(4%)	15 (2%)	9 (2%)	67 (3%)
Ventricular tachycardia	117 (20%)	55 (28%)	197 (22%)	42 (9%)	411 (19%)
Ventricular fibrillation /tachycardia	0	0	3 (<1%)	7 (1%)	10 (<1%)
Total Alarms	590	199	914	473	2176

Power Analysis

The primary analysis strategy to address aims 1 through 6 was a design with two between subject's factors (i.e., unit and assessment). This design allowed for testing the main effect of unit, the main effect of assessment, and the unit-by-assessment interaction. It is the test of the interaction that answers the primary question of whether the difference between Assessment 1 and Assessment 2 in experimental unit is the same as the difference between Assessment 1 and Assessment 2 in the control unit for any of the study alarm outcome variables. The sample size for the study was 244 in the experimental unit and 185 patients in the control unit. The total sample of 429 patients provided power of at least 80%, at a 2-tailed alpha of .05 to detect a small effect size, $f = .0971$. Cohen (1988) gives rough guides for what can be considered small, medium, or large effects for effects measured on the *f* scale. A small effect is $f = .10$, a medium effect is $f = .25$, and a large effect is $f = .40$.

For example, in the control unit the mean hourly SpO_2 alarm rate increased between Assessment 1 and Assessment 2 but in the experimental unit it decreased between Assessment 1 and Assessment 2. The effect size for this interaction was $f = .1323$, which was close to a small effect size and was found significant. As another example, the mean hourly arrhythmia alarm rate increased between Assessment 1 and Assessment 2 in the control unit but in the experimental unit it decreased between Assessment 1 and Assessment 2. The effect size for this interaction was $f = .0817$, which was also close to a small effect size and just made the cutoff for statistical significance. For the mean hourly ECG lead fail alarm rate, the test of the interaction was not significant. However, the effect size of the interaction was only $f = .012$, a very small and perhaps not clinically meaningful effect.

So, the study total sample size was adequate to detect what Cohen would consider small effect sizes. For some of the critical arrhythmia alarm outcome variables (e.g., accelerated ventricular), not all of the 429 patients generated such an arrhythmia alarm. There were in fact only a total of 54 patients who had an accelerated ventricular alarm. The much smaller sample does not provide power to detect a very small effect size. The sample 54 patients would provide power of at least 80%, at a 2 tailed alpha of .05, to detect a medium effect of f =.275. The effect size in the study for the mean hourly accelerated ventricular alarm interaction was only f = .193, and therefore it is not surprising that it wasn't statistically significant.

Reference

Cohen, J. (1988). *Statistical power analysis for behavioral sciences.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Chapter 6

Discussion

This study is the first physiologic monitor alarm randomized clinical trial to be conducted in an adult intensive care unit. The investigation discussed in this dissertation was conducted at a renowned academic medical center that provides tertiary and quaternary care to some of the sickest patients in northern California. Six months prior to the start of our study, the American Nurses Credentialing Center awarded the medical center with the prestigious Magnet designation. The study was generously supported by the University of California, San Francisco, School of Nursing, the institutions' Department of Nursing, and the University of California San Francisco (UCSF) Medical Center and exemplifies the organizations' values and commitment to safe patient care.

Objective of Dissertation; Purpose of Intervention

The purpose of the study presented in this dissertation was to assess the effectiveness of an alarm management nursing intervention for reducing select mean hourly physiologic monitor alarm rates and the mean percentage of false-positive arrhythmia alarms in a NICU. The intervention was bimodal, involving (a) modification of default oxygen saturation (SpO_2) alarm setting (i.e., lowering the low-limit threshold violation alarm and increasing the alarm delay) and (b) use of ECG skin preparation paper and the application of daily Ag/AgCl-foam, pre-gelled, *wet* ECG electrodes. The study's method entailed pre-/post-intervention assessments of an experimental group and a control group. The study's specific objective was to determine whether changes in the mean hourly alarm rates and mean percentage of false-positive arrhythmia alarms between Assessment 1 and Assessment 2 were different in the experimental unit than in the control unit.

Technology–based Nursing Intervention

As described in Chapter 2, parameter alarms resulting from inappropriate alarm threshold settings contribute to a substantial proportion of physiologic monitor alarms. In particular, oxygen saturation low-limit alarms result in few relevant alarms that require intervention (Biot, Carry, Perdrix, Eberhard, & Baconnier, 2000; Gross, Dahl, & Nielsen, 2011; Rheineck-Leyssius & Kalkman, 1998; Tsien & Fackler, 1997; Wiklund, Hök, Ståhl, & Jordeby-Jönsson, 1994). Analyses of our pilot alarm data identified that SpO_2 low-limit alarms contributed to an excessive amount of alarms (7-9% of all monitor alarms) as a result of the units' conservative default alarm settings (see Table 4.3, 4.4, and 4.5). This discovery provided the basis for our technology-based nursing intervention—modification and activation of new SpO_2 alarm features.

Because continuous SpO_2 monitoring has become ubiquitous on acute care medical–surgical units, *lowering* the SpO_2 low-limit threshold alarm setting and *activating* a SpO_2 alarm delay have been the two most frequently studied strategies for reducing non-actionable SpO_2 alarms; these investigations have reported substantial reductions in the frequency of unnecessary SpO_2 alarms (Graham & Cvach, 2010; Gross et al. 2011; Taenzer, Pyke, McGrath, & Blike, 2010). We aimed to investigate the merits of the above alarm reduction strategy in a clinical setting with a higher patient acuity —an intensive care unit.

Research Question 1. Is the difference in mean hourly SpO_2 low-limit threshold alarm rate rates between Assessment 1 and Assessment 2 different for the experimental unit (SpO_2 low-limit ≤ 88% with a 15-s alarm delay) compared with the control unit (SpO_2 low-limit ≤ 90% with a 5-s alarm delay)?

During Assessment 2, daily observations in the experimental unit, in combination with information obtained from the CIC Pro Clinical Information Center and the physiologic monitors, found that the SpO_2 low-limit threshold alarm and the SpO_2 alarm delay setting were parameter settings that were rarely adjusted by RNs; compliance with both a SpO_2 low-limit threshold alarm set to less than or equal to 88% and a SpO_2 15-s alarm delay was 98% (see Figure 5.1). Furthermore, our observations revealed that RNs rarely customized the alarm notification level either higher (e.g., warning) or lower (e.g., message) beyond the pre-determined unit default setting, which was defaulted to an advisory level. During Assessment 2, we observed that 99% of the time the RNs did not modify the SpO_2 alarm notification level; the alarm remained programmed at an advisory level (see Table 5.4).

An advisory level alarm is characterized by a single beep and displays a flashing white-on-red visual on-screen alert message. All SpO_2 advisory level alarms are also retained in the patient's alarm history; these alarms annunciate both at the patient's bedside and at the central station (CIC Pro Clinical Information Center). Assessment of unit noise level was not measured and was not a study aim. However, during Assessment 2, anecdotal staff reports indicated that the reduction in mean hourly SpO_2 alarm rates contributed to a reduction of noise both at the bedside and throughout the experimental unit. What could not be confirmed was (a) whether the RNs adhered to medical center policy in reviewing assigned patients' parameter alarms settings and (b) whether the RNs adjusted patients' alarm settings appropriately for patients' condition upon admission, subsequent to each shift, and as necessary. Regrettably, our clinical research software (BedMasterEx) and existing technology (GE Carescape Gateway) lacked the technological capability for acquiring this detailed information for all parameter alarms; such

information would have provided insight regarding the end-user interface with the physiologic monitors.

Our study's findings are congruent with those of prior clinical research that has promoted the innovative *combination* of a lowered SpO_2 alarm low-limit alarm setting and extending the alarm delay. By examining the frequency of SpO_2 alarms through use of unsophisticated calculations such as simple reductions in alarm proportions, our study found that SpO_2 alarm frequency in the experimental unit decreased by more than 60% (7,662 vs. 2,970) post-intervention (see Table 5.9). Although this result does not consider the patient as the unit of analysis, this reduction is consistent with findings of past studies that used simple calculations to measure alarm reductions following modifications in SpO_2 default alarm settings.

Our study results substantiate the clinical findings of a study by Taenzer and colleagues (2010), who reported reductions in the frequency of SpO_2 alarms that resulted from lowering SpO_2 alarm low limits *and* increasing alarm delays. In addition, Taenzer et al. found that these modifications were safe (i.e., did not result in increased adverse patient events) and contributed to fewer rescue events and unanticipated patient transfers to a higher level of care (i.e., ICU). Furthermore, with the modest lowering of the SpO_2 low-limit alarm threshold setting and the activation of an alarm delay to reduce the impact of transient events or artifact on the frequency of SpO_2 alarms, our results closely met or exceeded both theoretical predictions and past clinical research performed in progressive care units or on telemetry units in community hospitals (Graham & Cvach, 2010; Gross et al., 2011; Pan & Gravenstein, 1994; Rheineck-Leyssius & Kalkman, 1998).

We would be remiss to not mention that our 60% reduction in frequency of SpO_2 alarms fell short of the manufacturer's (i.e., Masimo®) assertion that a combined lower SpO_2 limit

setting (i.e., less than or equal to 88%) with an alarm delay (15 s) would result in a 85% reduction in SpO_2 alarms (Welch, 2011). However, it is important to note that Welch's premise regarding the effect of various SpO_2 alarm settings (thresholds and delays) on alarm frequency is primarily based on patient data obtained from 10 hospitals using the Masimo Patient SafetyNet Remote Monitoring and Clinician Notification System®—devices that are *not* typically used in intensive care units.

While our study's results are promising, a majority of bedside physiologic monitoring systems and most telemetry systems typically do not give clinicians the option of selecting the amount of time the SpO_2 value can fall outside of the pre-determined SpO_2 limits before an alarm sounds (e.g., GE Healthcare, Philips). For older physiologic monitors, the SpO_2 alarm delay feature may be available via a complex software upgrade; for most new physiologic monitors, this feature is a standard component. This feature constraint is a technological barrier that is apparent in recent studies that have explored optimizing parameter alarm limits. Whalen, Covelle, Piepenbrink, Villanova, Cuneo, & Awtry (2013) from Boston Medical Center widened the parameter alarm limits for heart rate (low/high limits) to minimize nuisance alarms; however, notably, the investigators neglected to mention modifications to the SpO_2 alarm (i.e., low-limit threshold and alarm delay). This oversight may have been related to the researchers' inability to make SpO_2 alarm adjustments (because of telemetry monitoring equipment limitations).

In our study, we were fortunate to be able to perform a software upgrade that enabled activation of a SpO_2 alarm delay (i.e., greater than or equal to 5 s) on the physiologic monitors in the experimental unit. While obtaining numerous administrative approvals is somewhat challenging, activation of an SpO_2 alarm delay was demonstrated to be safe and effective for minimizing nuisance alarms related to brief and self-correcting SpO_2 alarms in our experimental

unit. One way in which our study differed from prior studies is that we were able to satisfactorily demonstrate that clinically irrelevant SpO_2 alarms can be safely minimized by decreasing the SpO_2 low-limit threshold alarm to less than or equal to 88% and increasing the SpO_2 alarm delay to 15 s in an *adult ICU* without incurring an increase in adverse patient events. For the experimental unit, examination of baseline events (February 2013– July 2013) found a cardiopulmonary arrest (CPA) rate of 0.27/1,000 patient discharges and acute respiratory compromise (ARC) rate of 0.13/1,000 patient discharges. After implementation of the new default SpO_2 low-limit alarm threshold settings (August 2013–January 2014), the experimental unit had a CPA rate of 0.13/1,000 patient discharges and an ARC rate of 0.06/1,000 patient discharge—a considerable reduction (50%) in adverse patient events (see Table 5.6). To the author's knowledge, this is the first time that the effect of modifying the SpO_2 alarm setting has been studied in an adult ICU.

This alarm reduction strategy provides one promising technological approach that an interprofessional team of nurse leaders, academicians, biomedical engineers, leaders from device manufacturers, and others can put into practice to mitigate alarm fatigue. As a result of this study, our technology-based alarm management intervention (lowering the SpO_2 low-limit threshold alarm and activation of an alarm delay) is being adopted in the remaining four adult ICUs (and will become the default SpO_2 alarm setting for future adult ICUs) at the study site.

Our study's findings related to reducing the mean hourly SpO_2 alarm rate underscores the importance of ensuring that the unit default alarms are thoughtfully established and reassessed regularly as there may be changes in patient populations within units. Default alarm settings should not be determined merely on the basis of conservative physiological values or a traditional understanding of physiologic monitor alarms. Default physiologic monitor alarm

settings need to be established based upon research in conjunction with consideration of the harmful and repetitive effect that alarms can have on clinicians in the modern health care environment—habituation and indifference.

Practice–based Nursing Intervention

Many researchers have suggested that excessive technical alarms are due to poor signal quality from electrode management and that this alarm frequency can be minimized by satisfactory skin preparation and frequent ECG electrode applications to ensure good conductivity (Cvach, Biggs, Rothwell, Charles-Hudson, 2012; ECRI, 2007, 2013a; 2013b; Graham & Cvach, 2010; Meziane, Webster, Attari, & Nimbukar, 2013; O'Carroll, 1986; Oster, 2000; Patel & Souter, 2008; Turkmen & Pantiskas, 2011). Our research aimed to investigate this claim by studying the effect of a novel skin preparation and ECG electrode management regimen on technical (ECG lead fail, artifact, and arrhythmia suspend) and arrhythmia alarms.

Research Question 2. Is the difference in mean hourly technical alarm rates (e.g., ECG, artifact, arrhythmia suspend) between Assessment 1 and Assessment 2 different for the experimental unit (whose patients received a novel ECG skin site preparation technique and daily applications of high-quality ECG electrodes) compared with the control unit (whose patients received usual care)?

Our daily regimen specified using fine abrasive skin preparation paper to swipe the skin 1–3 times before applying each electrode (kept in sealed packages), attaching the lead wires to the dated electrodes before applying the electrodes, placing the electrodes on flat, fleshy parts and being careful to avoid bony prominences, and, once the electrodes were applied, applying gentle pressure to the electrode's outer edges. The RNs were instructed to change the patients'

electrodes around the "usual" time they bathe unconscious patients (for the most part, 4:00 a.m.–7:00 a.m.) and for patients that were conscious, during the daytime. RNs were instructed to press "pause" twice on the physiologic monitors to obtain 5 min of silence (to prevent an audible alarm) while they performed this task. On the basis of researchers' observations, RN compliance with the novel skin preparation and ECG electrode management regimen was 85% (see Table 5.5). Notably, researchers performed the skin preparation and applied ECG electrodes the remaining 15% of the time; accordingly, overall compliance was 100%. A 56% daily compliance occurred on the 2nd day and 15th day following the start of the intervention, and 100% compliance was achieved by the staff nurses on four days (see Figure 5.2). RN compliance was satisfactory; however, it is important to note that the RNs were aware that their compliance with the electrode regimen was being assessed daily; with little to no oversight, this level of performance would be difficult to sustain over time.

Furthermore, to encourage and promote compliance with the new patient monitoring supplies, the kits containing the ECG skin preparation paper and Ag/AgCl-foam, pre-gelled *wet* ECG electrodes were available in the experimental unit only at the point of care (i.e., at each patient's bedside). The investigators delivered a sufficient amount of bedside kits on a daily basis (rather than in bulk). The control unit staff maintained their standard inventory of ECG electrode supplies, and the appropriate storage of electrodes was inspected each day. These procedures prevented the inadvertent sharing of the novel patient monitoring supplies and ensured that only the study supplies (i.e., ECG skin preparation paper and high-quality electrodes) were used in the experimental unit.

During Assessment 2, of the 120 patients hospitalized in the experimental unit who received the novel skin preparation and ECG electrode regimen, two patients developed minor

skin breakdown. One breakdown incident lasted for 9 days and was related to the adhesive on the study electrodes; the other incident lasted for 2 days and was related to an elderly patient's naturally fragile skin (see Table 5.5). It is unknown what the impact of *daily* skin preparation and application of ECG electrodes would have on patients' skin integument during prolonged physiologic monitoring and hospitalization. Given the heightened awareness and national concerns to reduce hospital-acquired pressure ulcers (including skin breakdown related to medical devices) due to high treatment cost and reimbursement issues—recommendations to perform daily skin preparation and apply new ECG electrodes on monitored patients must be regarded with caution given the lack of published research to endorse this clinical practice.

ECG lead fail alarm

The ECG lead fail alarm, an alarm condition wherein no ECG waveforms are displayed on the physiologic monitor, indicates interruption of cardiac monitoring due to an equipment failure. To aid continuous cardiac monitoring, if the quality of an ECG electrode signal deteriorates to an unsatisfactory level, a lead fail message for the affected lead is displayed on the physiologic monitor; each ECG leads fail alarm will be preceded by alarm events for each individual lead (I, II, III, and V; GE Healthcare, 2007). These interruptions can be intentional when discontinuing monitoring or unintentional in several circumstances: (a) detachment of the ECG lead wire from the electrode; (b) detachment of the ECG electrode from the patient's skin; and even (c) poor sensor placement (ECRI, 2007; Medina, Clochesy, Omery, 1989; Oster, 2000; Paparella, 2014; Sendelbach & Funk, 2013). Failure to respond quickly to ECG lead fail alarms—an alarm that signals discontinuation of cardiac monitoring and arrhythmia analysis—can lead to adverse events, including patient death (ECRI, 2007). Because of this potential for a sentinel event, researchers have focused on reducing ECG lead fail alarms

primarily in progressive units and acute care telemetry units (rather than in ICUs), where direct view of either the patient or the monitor display may not be possible.

Although our study found no statistically significant unit-by-assessment interaction ($p <$ .741; see Table 5.10), tests of simple effects found that during the assessment periods, the experimental unit's mean hourly ECG lead fail alarm rate increased over time; during Assessment 1 and 2, these rates were .31 and .35, respectively ($p = .723$). Similarly, during Assessment 1 and 2, the control unit's mean hourly ECG lead fail alarm rate was .30 and .40, respectively ($p = .525$; see Figure 5.8). This negative finding (i.e., an increase in the mean hourly ECG lead fail alarm rate in the experimental unit) was not the result we hypothesized would occur during Assessment 2. We had anticipated that the mean hourly ECG lead fail alarm rate would decline between Assessment 1 and Assessment 2 in the experimental unit and that the mean hourly ECG lead fail alarm rate would remain unchanged in the control unit, thereby producing a significant unit-by-assessment interaction.

Given that we achieved an overall 100% compliance with the electrode regimen, we suspect that the slight increase in the mean hourly ECG lead fail alarm rate was the consequence of the more frequent ECG electrode applications, which negated any reductions in this alarm outcome variable. Interestingly, our study results pertaining to the ECG lead fail alarm were somewhat similar to results reported in a study performed at Johns Hopkins Hospital, which measured the effect of a traditional skin preparation (use of soap and water, rubbing skin with gauze) and daily ECG electrode change. In this study, Cvach et al. (2012) reported that a 13% increase in ECG leads fail alarms was observed in the 15-bed medical progressive care unit, while a 15% reduction in ECG lead fail alarms was observed in the 25-bed cardiology care unit.

The investigators proposed that the cause of this increase may have been related to possible failure on the part of the dedicated technician (who performed this task during the 8-day quality improvement project) to pause the alarm on the bedside monitors (i.e., GE Solar 8000i) before changing a patient's electrodes.

From collaboration with device manufacturer engineers, we have learned that in our study, an ECG lead fail alarm was "broadcasted" and recorded as an alarm event regardless of whether the ECG lead fail alarm was audible or silent on the GE Solar 8000i physiological monitors (GE Healthcare, Waukesha, WI) that were used in both the experimental unit and control unit during Assessment 1 and Assessment 2. This technological element may explain our findings of increased mean hourly ECG lead fail alarm rates in the experimental unit.

Although a remarkable amount of discussion and effort has focused on reducing the ECG lead fail alarm, it would seem that clinicians and researchers should anticipate that a reasonable number of ECG lead fail alarms will occur (e.g., 1 ECG lead fail alarm/patient/day)—that is, if the nursing staff are indeed changing ECG electrodes regularly per their institutions' policies and procedures. On the contrary, having an unusually low frequency of ECG lead fail alarms would be a matter of concern—as a possible indication that staff are not regularly assessing the ECG electrode skin site or changing electrodes as specified.

In this analysis, the duration of the ECG lead fail alarms was not measured because without direct observation or video recordings, ascertaining the accuracy of the ECG lead fail alarm duration is difficult. That is, ECG lead fail alarms can occur when a patient is disconnected from patient monitoring or can persist when a clinician is in the midst of providing patient care (i.e., resolving the ECG lead fail alarm) and cannot acknowledge the alarm by pressing "pause" once or twice to silence the alarm—thus leading to prolonged alarm duration.

To mitigate the generation of audible alarms that are related to patient care (e.g., skin care) and that are due to the manipulation of monitoring equipment (e.g., application of ECG electrodes), nurse education can include the anticipation of alarms and pausing or silencing alarms in advance of performing tasks. In addition, education can emphasize the importance of silencing physiologic monitor alarms only when the problem has been addressed or the patient has been assessed (Lipton et al., 2009; Scott Allen, Hileman, & Ward, 2013; Siebig, Kuhls, Imhoff, Gather, Schölmerich, & Wrede, 2010; Way, Beer, & Wilson, 2014).

Artifact alarm

Artifact alarms indicate a transient condition resulting from intermittent noise and artifacts, which often generate false-positive arrhythmia alarms and nuisance alarms. Artifact alarms begin at “level 1” and progress to an arrhythmia suspend alarm “level 2” if the ECG noise lasts for 20 s of the last 30 s. As previously discussed, it is important to note that when artifact level 1 alarms are triggered, full arrhythmia processing is suspended, yet the lethal arrhythmias detection software (EK-Pro, Version 11) remains active for only the ventricular fibrillation/ tachycardia and asystole alarms (GE Healthcare, 2007).

To the best of the author’s knowledge, our investigation is the first to report on mean hourly rates of artifact alarms. This is a unique aspect of our study, given that artifact alarms are typically visual “message” notifications, and data pertaining to these alarm events are not easily collected. In our study, collection of these data was feasible via the use of BedMasterEx (Excel-Medical Electronics, Jupiter, FL) and the GE Research Carescape Gateway (GE Healthcare, Waukesha, WI). Our study found no statistically significant unit-by-assessment interaction for mean hourly artifact alarms (p = .891; see Table 5.10). Although our study found no interaction, examination of simple effects revealed that during Assessment 1 and 2, the experimental unit had

a mean hourly rate of artifact alarms of 11.64 and 8.48, respectively (p = .061). During Assessment 1, the control unit had a mean hourly artifact alarm rate of 13.40; this rate decreased to 9.41 at Assessment 2 (p = .069). Between Assessment 1 and Assessment 2, mean hourly artifact alarm rate in both units decreased slightly; this decrease was not statistically significant (see Figure 5.9). The above results indicate that our intervention (novel skin preparation and a daily electrode regimen) did not affect the mean hourly artifact alarm rate. We had expected that between Assessment 1 and Assessment 2, the mean hourly rate for artifact alarms in the experimental unit would decrease substantially and remain relatively unchanged in the control unit during the two assessment periods. Because this is the first physiologic monitor alarm study to report artifact alarms, it is unknown how our mean hourly artifact alarm rates compare with those of similar units and institutions pre- and post-intervention.

Artifact alarms occur frequently. During Assessment 1 and Assessment 2, patients in the experimental unit triggered 128,002 and 79,337 artifact alarms, respectively, and the control unit had fewer artifact alarms—87,107 and 62,273, respectively (see Table 5.11). In the author's opinion, artifact alarms are often underappreciated—perhaps because this alarm is typically assigned as a low-priority alarm condition (i.e., visual) and are difficult to collect. It is the author's belief that most clinicians are unaware that artifact alarms (level 1), if unresolved, progress to an arrhythmia suspend (level 2) alarm—and arrhythmia interpretation is completely suspended, including interpretation of lethal arrhythmias. If artifact alarms were more widely understood, perhaps greater attention would be devoted to the prevention of artifact alarms—given the potential ramifications for patient safety.

Arrhythmia suspend alarm

As discussed above, with the arrhythmia suspend alarm, all arrhythmia interpretation is

completely suspended (including ventricular fibrillation/tachycardia and asystole). The arrhythmia suspend alarm generates a continuous foghorn alarm (i.e., system warning) until the quality of the ECG signal improves. To resume arrhythmia processing and alarms, the alarm condition must be resolved by clinicians verifying lead placement, performing skin preparation, and/or subsequently applying new ECG electrodes (GE Healthcare, 2007).

Similar to the artifact alarm outcome variable, we had hypothesized that the practice-based intervention would influence the arrhythmia suspend alarm. However, we found no unit-by-assessment interaction with regard to the mean hourly rate of arrhythmia suspend alarms (p =.130; see Table 5.10). For the experimental unit, the mean hourly arrhythmia suspend alarm rate decreased from 0.06 and 0.05, respectively (p = .490; see Figure 5.10). Although no significant interaction was discovered, the control unit did experience a statistically significant reduction. During Assessment 1 and Assessment 2, the control unit had a reduction in mean hourly arrhythmia suspend alarm rate from 0.08 to 0.03 (p = .01), respectively. However, it is important to note that the mean hourly rates for this particular alarm outcome variable are very small; this result must be interpreted with caution given the rarity of this alarm. Although the mean hourly arrhythmia suspend alarm rates decreased in both units over time (and greater in the control unit which cannot be explained), the reduction was insufficient to generate an interaction; therefore, this change cannot be associated with our practice-based intervention.

Arrhythmia suspend alarms occur infrequently and the prevalence of the arrhythmia suspend alarm is low. During Assessment 1 and Assessment 2, the experimental unit had 745 and 675 alarms, respectively (see Table 5.11), and patients in the control unit also generated few arrhythmia suspend alarm (i.e., 383 and 223, respectively). Only two quality improvement projects have reported proportions of arrhythmia suspend alarms following modifications in unit

default setting or following daily ECG electrode change (Graham & Cvach, 2010; Cvach et al., 2012). Both studies were conducted by the same investigator in the same institution, and one unit (i.e., the medical progressive care unit) was involved in both quality improvement projects. The percent reduction in arrhythmia suspend alarm frequency ranged from 8% in the 2010 project (which focused on alarm setting modifications) to as high as 56% in the 2012 project (which examined the effect of daily ECG electrode change on alarm proportions). Because our study used a different methodology to account for overdispersion of alarm data, comparing our results with previous results is difficult. Nonetheless, statistical analysis showed that our practice-based intervention did not have a significant effect on the arrhythmia suspend alarm outcome variable.

Research Question 3. Is the difference in mean hourly arrhythmia alarm rates between Assessment 1 and Assessment 2 different for the experimental unit (whose patients received a novel ECG skin site preparation technique and daily applications of high-quality ECG electrodes) compared with the control unit (whose patients received usual care)?

Our study found that for the alarm outcome variable, arrhythmia alarms, the unit-by-assessment interaction was statistically significant ($p = .05$; see Table 5.12). The Assessment 1–Assessment 2 change in the experimental unit differed from that of the control unit. Simple effects revealed that during Assessment 1, the mean hourly rate of total arrhythmia alarms in the experimental unit was 14.10; during Assessment 2, the mean hourly arrhythmia alarm rate decreased significantly to 5.42 ($p < .002$; see Figure 5.11). Meanwhile, in the control unit, the mean hourly arrhythmia alarm rate increased between Assessment 1 and Assessment 2, from 10.33 to 12.49 ($p = .707$). Our results indicate that our novel ECG electrode regimen affected all arrhythmia alarms—comprising15 different types of cardiac arrhythmia alarms (see Table 5.13).

Our hypothesis that the experimental unit's mean hourly Assessment 1–Assessment 2 arrhythmia alarm rate change is different from that of the control unit was correct. The study's findings confirmed that the mean hourly rate of arrhythmia alarms decreased between Assessment 1 and Assessment 2 in the experimental unit, whose patients received a daily skin preparation using ECG skin preparation paper and the application of daily Ag/AgCl-foam, pre-gelled, *wet* ECG electrodes and remained the same in the control unit. This finding must be regarded with caution because excess variations in arrhythmia alarm counts (e.g., atrial fibrillation, ventricular tachycardia ≥ 2, PVCs) generated by relatively few patients can influence the mean hourly arrhythmia alarm rates during assessment periods; arrhythmia alarm rates fluctuate considerably because of variations in hospitalized patients' cardiovascular conditions. For example, during Assessment 1, one patient in the experimental unit generated 6,256 atrial fibrillation alarms and one patient in the control unit triggered 1,626 atrial fibrillation arrhythmia alarms. Similarly, during Assessment 2, one patient in the experimental unit and one patient in the control unit generated 9,105 and 6,119 atrial fibrillation arrhythmia alarms, respectively.

Notably, our results indicate that six critical arrhythmia alarms (asystole, accelerated ventricular, pause, ventricular bradycardia, ventricular tachycardia, and ventricular fibrillation/tachycardia) had no unit-by-assessment interaction either for individual arrhythmia alarms or for these alarms in combination ($p < .273$; see Table 5.15). This finding indicates that perhaps the remaining 9 arrhythmia alarms (atrial fibrillation/irregular, ventricular tachycardia ≥ 2, bradycardia, tachycardia, bigeminy, trigeminy, couplet, R on T, and PVC) were more sensitive to the intervention; however additional analyses are required to confirm this alternative explanation and the impact of patient outliers. It is important to recognize that, with the inclusion of the 10 non-lethal arrhythmia alarms, several of these lower priority alarms are not audible.

According to the units' alarm default, five of the non-lethal arrhythmia alarms (couplet, bigeminy, trigeminy, R on T, and PVC) are set to a notification level of "message," which is simply a visual alert (not an audible alarm). The remaining four arrhythmia alarms (ventricular tachycardia ≥2, tachycardia, bradycardia, and atrial fibrillation/irregular) default to an advisory level alarm; this defaulting has implications for alarm audibility.

Research Question 4. Is the difference in mean hourly audible alarm rates between assessment one and assessment two different for the experimental unit (whose patients received a novel ECG skin site preparation technique and daily applications of high-quality ECG electrodes) compared with the control unit (whose patients received usual care)?

Because our technology and practice-based alarm management interventions were specifically aimed at reducing mean hourly alarm rates for select audible alarm outcome variables (i.e., SpO_2 low-limit alarms, technical alarms, and cardiac arrhythmia alarms) along with reducing the mean percentage of false-positive arrhythmia alarms, in hindsight, it was not surprising that we did not observe a statistically significant unit-by-assessment interaction for mean hourly audible alarm rates (p <.316; see Table 5.12). Examination of simple effects revealed that during Assessment 1 and 2, the experimental unit had a mean hourly rate of audible alarms of 6.73 and 6.25, respectively (p = .357; see Figure 5.12). At Assessment 1, the control unit had a mean hourly audible alarm rate of 6.71; this rate increased to 7.07 at Assessment 2 (p = .595). Although the mean hourly audible alarm rates decreased in the experimental unit and increased in the control unit during Assessment 2, this result was insignificant and was incapable of generating a unit-by-assessment interaction that could be attributed to our technology-based and practice-based intervention.

This finding suggests that research must investigate alternative strategies for reducing mean hourly audible alarm rates. Such strategies can include (and are not limited to) adjusting the unit default alarm settings for both parameter (i.e., high-and-low limits) and arrhythmia alarms (i.e., visual notification vs. audible alarms) and using research to validate safe alarm settings—not simply the manufacturer's recommendations. Parameter high- and low-limit alarms must be appropriate for the patient population (e.g., in terms of patient age and unit specialization) and must be based on scientific research such as the reviewed studies that investigated a less conventional SpO_2 low-threshold limit (Gross et al., 2011; Taenzer et al., 2010; Welch 2011).

Another strategy, albeit less studied, entails enabling actionable alarms only and re-prioritizing select arrhythmia alarms from auditory to visual alarms (i.e., lower priority). The American College of Cardiology (ACC), American Heart Association (AHA), and the European Society of Cardiology (ESC; 2006) practice guidelines recommend that for the acute management of in-hospital patients, neither accelerated ventricular rhythm nor non-sustained ventricular tachycardia (greater than 30 s) warrant antiarrhythmic therapy. In addition, the guidelines recommend that only sustained or hemodynamically compromising ventricular tachycardia requires treatment (ACC/AHA/ESC, 2006). It appears that some device manufacturers and some clinicians may not be comprehensively familiar with these published guidelines. For example, some manufacturer defaults for ventricular tachycardia ≥ 2 is a crisis level alarm (3 beeps); however, according to the 2006 practice guidelines, the ventricular tachycardia ≥ 2 alarm could be set to a message alarm (visual alarm; GE Healthcare, 2007). Similarly, some hospitals (Johns Hopkins Hospital and Boston Medical Center) report assigning a higher priority level to arrhythmia alarms than what is clinically recommended for early

intervention and treatment by practice guidelines in the United States and in Europe.

For instance, Johns Hopkins Hospital currently sets the ventricular tachycardia ≥ 2 alarm to a warning level (2 beeps), whereas alternatively, this arrhythmia alarm could also be changed to an inaudible alert (Graham & Cvach, 2010). In addition, Graham and Cvach (2010) and Fidler, Pickham, and Drew (2011) report that the PVC alarm were set as an audible alarm in some units at both Johns Hopkins Hospital and Stanford Hospital and Clinics, respectively. These are yet additional examples in which improper setting of a default arrhythmia alarm's notification level (i.e., in excess of recommended treatment guidelines) results in unwarranted annunciation of clinical urgency (It is no longer recommended or common clinical practice to treat PVCs because the drugs used for PVC treatment may provoke life-threatening ventricular arrhythmias; The Cardiac Arrhythmia Suppression Trial [CAST], 1989).

Another renowned academic medical center, Boston Medical Center, has also recently reported elevating certain arrhythmia notification levels beyond what is clinically necessary—for example, increasing the accelerated ventricular alarm from a warning level to a crisis level and maintaining the ventricular tachycardia ≥ 2 at a crisis level alarm (Whalen et al., 2013). The justification for giving these arrhythmia alarms such a high severity notification level is unclear, given that these cardiac arrhythmias do not require clinical intervention. These modifications are not in accordance with the principle that alarm fatigue can be minimized by reducing non-actionable alarms, and hence, only higher, audible alarms should be actionable—that is, linking the clinical urgency with the alarm urgency.

In addition to optimizing unit default alarm settings, institutions must invest in and support initial and ongoing training to ensure that clinicians understand patient monitoring equipment and alarm features so that they (clinicians) can adjust alarm parameter limits and

arrhythmia alarm notification levels on the basis of published research and their clinical judgment regarding the patient's physiological condition.

Research Question 5. Is the difference in mean hourly rate of all physiologic monitor alarms between Assessment 1 and Assessment 2 different for the experimental unit (whose patients received a novel ECG skin site preparation technique and daily applications of high-quality ECG electrodes) compared with the control unit (whose patients received usual care)?

Our hypothesis was that the experimental unit's mean hourly Assessment 1–Assessment 2 physiologic monitor alarm rate change is different (i.e., following our intervention), from that of the control unit was unsubstantiated ($p < .100$; see Table 5.12). We expected the mean hourly rate of physiologic monitor alarms to decline in the experimental unit over time and remain the same in the control unit. Although the mean hourly rate of physiologic monitor alarms remained mostly unaffected in the control unit, the alarm rate significantly declined over assessment periods in the experimental unit from 41 alarms/ hr to 28 alarms/hr, respectively ($p = .002$; see Table 5.12 and Figure 5.13). However, the decrease in physiologic monitor alarm rates (subsequent to significant reductions to the SpO_2 low-limit alarm rates and arrhythmia alarm rates) was inadequate to impact the overall physiological monitor alarm outcome variable and to be attributed to our interventions.

Summary of Mean Hourly Alarm Outcome Variables

The first intervention—lowering the unit default SpO_2 low-limit threshold alarm to less than or equal to 88% and increasing the SpO_2 alarm delay to 15 s—was found to be safe and effective in reducing mean hourly SpO_2 alarm rates ($p = .001$). Our second intervention—the use of a daily skin preparation using ECG skin preparation paper and the application of daily

Ag/AgCl-foam, pre-gelled, *wet* ECG electrodes—was effective in reducing only the mean hourly alarm rates for all patient status arrhythmia alarms (i.e., for all 15 alarms collectively; $p = .05$).

Our clinical practice-based intervention did not have a statistically significant unit-by-assessment interaction for all technical alarm studies (i.e., ECG lead fail, artifact, arrhythmia suspend) or for individual critical arrhythmia alarms. The intervention's ineffectiveness in this regard indicates that other strategies are required to minimize mean hourly alarm rates for particular alarm outcome variables, such as technical alarms, critical arrhythmia alarms, and even other parameter alarms. Our findings reinforce the view that effective alarm management requires a multifaceted approach to reduce non-actionable alarms—and that a variety of unique interventions or alarm setting modification are required to minimize the mean hourly alarm rates for the large subset of physiologic monitor alarms. This proposition is in alignment with recommendations by many respected experts, professional societies, and agencies dedicated to reducing clinical alarm fatigue (AAMI, 2011; ECRI, 2007, 2012, 2013a, 2013b; TJC, 2013, 2014).

Research Question 6. Is the difference in the mean percentage of false-positive cardiac arrhythmia alarms (e.g., accelerated ventricular, asystole, pause, ventricular bradycardia, ventricular tachycardia, ventricular tachycardia/ventricular fibrillation) between Assessment 1 and Assessment 2 different for the experimental unit compared with the control unit?

We hypothesized that the impact of a daily skin preparation using ECG skin preparation paper and the application of daily Ag/AgCl-foam, pre-gelled, *wet* ECG electrodes would substantially reduce the mean percentage of false-positive arrhythmia alarms. We had posited that many false-positive arrhythmia alarms were related to poor electrode management as a result

of (a) electrode signal degradation to an inadequate level, (b) low-voltage signals, (c) intermittent noise, and (d) artifacts that could be mitigated by an effective electrode regimen. It appears that we overrated the utility and value of this recommended practice, given that the analysis of mean percentage of false-positive critical arrhythmia alarms revealed no unit-by-assessment interaction for any of the six critical arrhythmia alarms that were annotated (see Table 5.17). Meaning, RNs performing the novel ECG electrode regimen on patients daily did not reduce the mean percentage of false-positive critical arrhythmia alarms.

Our investigation is the first to study the impact of a novel ECG skin site preparation and electrode application regimen on the mean percentage of false-positive critical arrhythmia alarms—not only for assessing the accuracy of the arrhythmia alarms in terms of true-positive/false-positive but also for performing the analysis at the patient level (i.e., not using alarms as the unit of analysis). Given these design elements and the absence of similar research, comparison of our findings with previous research is difficult. To a degree, we can compare our study's finding of mean percentage of false-positive cardiac arrhythmia alarms with findings from a retrospective, off-line analysis conducted by Aboukhalil, Nielsen, Saeed, Mark, and Clifford (2008). Aboukhalil et al. studied the proportion of false-positive arrhythmia alarms using both simultaneous ECG morphological and arterial blood pressure waveform information. The two studies differ in their use of operational definitions for life-threatening ECG arrhythmia alarms outcome variables and in methodology (e.g., Aboukhalil et al. calculated the frequency of true alarms and false alarms on a per-alarm basis). Accordingly the studies' results must be compared with caution.

Our prospective study and the study by Aboukhalil et al. had three critical cardiac arrhythmia alarms in common: asystole, ventricular tachycardia, and ventricular

tachycardia/ventricular fibrillation alarms. Aboukhalil et al. reported the following proportion of false-positive alarms: asystole, 91%; ventricular tachycardia, 47%; and ventricular tachycardia/ventricular fibrillation, 80%. In our study, the mean percentages of asystole false-positive alarms during Assessments 1 and 2 in the experimental unit were 77% and 94%, respectively, and in the control unit, 94% and 89%, respectively (see Table 5.17.). Despite methodological differences, the accuracy of the asystole alarms in the studies was somewhat similar (i.e., for the control unit during both assessments and during Assessment 2 for the experimental unit). In our study, the mean percentage of false-positive ventricular tachycardia alarms during Assessment 1 and Assessment 2 in the experimental unit was 89% and 77%, respectively, and in the control unit, 83% and 80%, respectively. In comparison with Aboukhalil et al.'s results for ventricular tachycardia, our study's mean percentage of false-positive ventricular tachycardia alarms was higher (i.e., more false-positive ventricular tachycardia alarms) than that previously reported by investigators (i.e., 47%). Lastly, our study found that in the experimental unit, the mean percentage of false-positive ventricular tachycardia/ventricular fibrillation alarms during both assessment periods was 100%. In the control unit, no ventricular fibrillation/tachycardia alarms were triggered during either assessment period. In contrast, Aboukhalil et al. reported a better result; 80% frequency of false-positive ventricular fibrillation/tachycardia alarms on a per-alarm basis. It is worth mentioning that the number of arrhythmia alarms investigated in the study by Aboukhalil et al. was much larger than the number investigated in our study, and alarms were not obtained from patients in dedicated neuroscience intensive care units.

One factor that may have contributed to our failure to detect a statistically significant reduction in the mean percentage of false-positive arrhythmia alarms is that the arrhythmia

detection software (EK-Pro V.11) requires a minimum QRS size of 0.5 mV (5 mm) in all four analyzed leads. In light of this current technological constraint, reduction of false-positive critical arrhythmia alarms may largely depend on improving the arrhythmia detection software and shifting research focus away from clinical practice. Given the variation in methodological approaches and paucity of studies that include arrhythmia alarm annotation, additional research (including replication studies) is required to assess the effectiveness of interventions for reducing the frequency of false-positive arrhythmia alarms.

Lastly, other physiologic monitor alarm studies that have annotated arrhythmia alarms used diverse approaches to determine technical validity and clinical validity. (In this regard, a technically false alarm is an alarm erroneously annunciated on the basis of measurements that do not correctly reflect the patient's condition. The term *not alarm relevant* refers to an alarm that is not followed by a diagnostic or therapeutic decision; Siebig et al., 2010). Other investigators have used descriptors such as "effective" and "ineffective" and "ignored" as a result of many observers not being clinicians (Görges, Markewitz, & Westenskow, 2009). The dissimilarity in terminologies and approaches are so substantial that it is impossible to compare the above studies results regarding the relative frequency or percentage of false-positive cardiac arrhythmia alarms with our findings.

Review of Study Findings

Our study had two major statistically significant findings, among other interesting results. First, our modification of the default SpO_2 alarm settings had a statistically significant effect on reducing mean hourly SpO_2 alarm rates ($p < .001$; see Table 5.8 and Figure 5.7). In this regard, our technologically based nursing intervention aimed at minimizing brief and clinically insignificant SpO_2 alarms by (a) lowering the unit default SpO_2 low-limit threshold alarm to less

than or equal to 88% and (b) increasing the SpO_2 alarm delay to 15 s (from less than or equal to 90% with a 5-s alarm delay) was an effective alarm management strategy. Modifications of the unit default SpO_2 alarm settings were implemented safely with no reported increase in adverse patient events (i.e., cardiopulmonary arrests and/or acute respiratory compromise over assessment periods (see Tables 5.6 and 5.7).

Second, the study's novel clinical practice regarding an ECG electrode regimen had a statistically significant effect on only one alarm outcome variable—all arrhythmia alarms, which included 16 different unique arrhythmia alarms (see Tables 5.12 and 5.13). The unit-by-assessment interaction for mean hourly arrhythmia alarm rates was statistically significant (p = .05). In this regard, our nursing intervention focused on clinical practice; the use of a daily skin preparation using fine abrasive ECG skin preparation paper and the application of Ag/AgCl-foam, pre-gelled, *wet* ECG electrodes was effective in minimizing the mean hourly rates of arrhythmia alarms. However, this result must be interpreted with caution, given the over-dispersion of arrhythmia alarm counts and known patient outliers.

Surprisingly, no statistically significant unit-by-assessment interactions were observed for the remaining physiologic monitor alarm outcome variables: (a) technical alarms (i.e., ECG leads fail, artifact, arrhythmia suspend); (b) six critical cardiac arrhythmia alarms; (c) audible physiologic monitor alarms (i.e., patient and system status); and (d) all physiologic monitor alarms (see Table 5.11, 5.12, and 5.15). In addition to examining hourly alarm rates, examination of changes from Assessment 1 to Assessment 2—in terms of mean percentage of false-positive cardiac arrhythmia alarms (i.e., accelerated ventricular, asystole, pause, ventricular bradycardia, ventricular tachycardia, ventricular tachycardia/ventricular fibrillation) following a novel electrode management regimen was evaluated and found to be insignificant (see Table 5.17).

Notably, comparison of our study's findings with those of previous research is difficult, for two reasons. First, our study is the first to consider the patient as the unit of analysis; previous studies have considered alarms as the unit of analysis. Second, our study is the first to analyze alarm outcome variables expressed as a mean hourly rate based on total cardiac monitoring time; computation of this rate used an algorithm that excluded interruptions in patient monitoring. However, our conclusions can be compared with conclusions of previous studies and with results from quality improvement projects.

Limitations

The investigator recognized that the study design has limitations. Although the data were collected at a single hospital site, the two NICUs are located on separate geographical units (i.e., three floors apart). Also, due to the dynamic nature of ICU staffing in a large academic medical center, the nursing staff (RNs and PCAs) can be assigned to either unit. This limitation was recognized and because the data collection period and intervention is short (31 days for both assessment periods), the degree of staff ad hoc assignment between the units was minimized. In addition, for continuity of patient care and to promote work satisfaction, RNs and PCAs were typically assigned to one unit for the 4-week scheduling period.

It is important to note that there are limitations with the acuity tool—primarily patient acuity scores are based on RNs' clinical judgment and subjective assessment of patient care requirements—and often time, scores can be overestimated or inflated for certain shifts and days to drive staffing levels. However; for simple comparison purposes, this data does provide information on overall unit acuity during Assessment 1 and 2 as reported by the staff nurses assigned to provide direct patient care to the very patients who generated the physiologic monitor alarms under study. While the experimental and control units' acuity did not change among

assessment periods, a better methodology for reporting patient acuity would have been preferable. The use of a highly valid and reliable instrument such as the Acute Physiology and Chronic Health Evaluation (APACHE II/III) or the Simplified Acute Physiology Score (SAPS II) which reliably measures severity of disease based primarily on objective physiological measurements for adult patients admitted to intensive care units would have been preferred—rather than an instrument designed to measure patient care requirements.

Lastly, although this study was performed in a large academic medical center and included physiologic monitor alarms from 2 ICUs over two-31-day assessment periods, we obtained a relatively small sample of critical arrhythmia alarms defined as; accelerated ventricular, asystole, pause, ventricular bradycardia, ventricular tachycardia, and ventricular tachycardia/ventricular fibrillation alarms. While collecting more arrhythmia alarms may have reduced the challenges associated with overdispersion of data, it would have been particularly concerning to have obtained many more arrhythmia alarms—especially true alarms—as it could be interpreted that the collection of more arrhythmia alarms could indicates a failure to adequately anticipate or treat patients in the study units.

In addition, it is acknowledged that during Assessment 2 study, RNs working on the control unit could decrease the default SpO_2 low-threshold alarm to 88% (with a 5-s alarm delay), and conversely, the RNs on the experimental unit could increase the default SpO_2 low-threshold alarm to 90% or higher (with a 15-s alarm delay). To verify the adoption of the new SpO_2 low-threshold alarm setting in the experimental unit, daily verification was conducted at the CIC Pro Clinical Information Center and physiologic monitors with minimal staff disturbance and awareness. Lastly, although the study involved a single site and consisted of alarm data from a relatively homogeneous patient population, we believe that the knowledge

gained from our study can be applied to other clinical areas that conduct similar physiological monitoring.

Implications for Nursing

The implications of our study's findings for nursing practice and the care of monitored patients are broad. Our foremost finding is that our technology-based nursing intervention—modification of the unit default SpO_2 alarm (i.e., lowering the low-limit threshold violation alarm and increasing the alarm delay)—successfully reduced the mean hourly SpO_2 alarm rate. Not only was our hypothesis substantiated but also the reduction in SpO_2 alarm rates was achieved with minimal expenditures (cost and workload). In addition, this improvement was in conjunction with a highly satisfactory RN compliance (98%) over the 31-day assessment period. These factors suggest that this alarm reduction strategy is sustainable.

However, our practice-based intervention was not as effective as we had predicted. Moreover, this intervention was costly (largely because of required purchases of high-quality patient monitoring supplies). This analysis revealed that although the novel skin preparation and daily application of ECG electrodes reduced mean hourly arrhythmia alarm rates, this practice did not have statistically significant benefit for other alarm outcome variables—namely, technical alarms (ECG lead fail, artifact, and arrhythmia suspend). In addition, the intervention did not effectively reduce audible alarms, physiologic monitor alarms, or the mean percentage of false-positive critical arrhythmia alarms. For these reasons, the optimal interval for performing skin preparation and applying ECG electrodes is undetermined. Accordingly, the merits of this practice must be thoughtfully considered given the practice's expense—in terms of both clinician workload and supply costs, which are so often transferred to the health care consumer or payer. Moreover, given that there is very little science to substantiate changing ECG electrodes daily,

we may be at risk for increasing skin breakdown among monitored patients.

Finally, in considering the dynamic health care environment, it is important to keep in mind the numerous challenges associated with sustaining staff compliance with best practice recommendations in fast-paced, high-acuity nursing units, given the great variation in clinicians' skills and performance. In this regard, it is the author's professional opinion that efforts to reduce alarm fatigue must be largely directed at developing and implementing technology-based strategies (e.g., improved algorithms and use of appropriate alarm default settings and alarm delays for select parameters) rather than at improving clinical practice dependent upon human capabilities.

Implications for Research

Little is known about the characteristics of the study patients, including those who generated many physiologic monitor alarms and those who generated few alarms. This study's findings raise questions about individual's clinical characteristics (e.g., seizures, fever, confusion, smoking cessation, alcohol use) and about other factors (e.g., mobilization, mechanical ventilation, post-operative recovery phase) that may influence alarm generation. Another related research approach would be to investigate whether following an alarm management intervention, if a patient's hourly alarms rates would decrease over time (i.e., for the same patient). This approach might facilitate evaluation of select nursing and technological interventions (because these strategies would be studied on the same patients during their hospitalizations, thereby eliminating some potential confounding variables).

Furthermore, investigators might consider performing analyses to examine ECG waveform quality (i.e., tracings) between patients in the experimental unit and control unit to assess differences in signal quality at baseline (Assessment 1) and following the novel skin

preparation and daily application of ECG electrodes intervention (Assessment 2). An outcome variable such as signal-to-noise ratio is a measure that can be used to compare the desired signal with the level of background noise (i.e., artifact) in assessing the effectiveness of an intervention on signal quality for patient monitoring. Another approach to analyze the quality of waveforms is signal abnormality index. This waveform quality index can quantify the quality of recorded signals and has been previously used to analyze arterial blood pressure waveforms and compare the vital sign documentation with automated acquired values using waveform indices (Sapo et al., 2009; Sun, Reisner, & Mark, 2006).

Given that our study demonstrated that a novel electrode management intervention does not reduce technical nor false-positive arrhythmia alarms, clinicians cannot rely on improvement in clinical practice to be the cure-all for alarm fatigue; in addition to improving clinical practice, we must improve algorithms for accurate arrhythmia detection. Future studies are required to determine whether improved algorithms will reduce the high percentage of false-positive arrhythmia alarms. Efforts to reduce false-positive arrhythmia alarms and ensuing alarm fatigue require an interprofessional approach, innovation, and a commitment to improve the care of the monitored patient. Time and resource investments from numerous stakeholders, especially device manufactures, are required for investigations of new arrhythmia detection and motion artifact reduction algorithms.

Also, as discussed in the review of the literature and as suggested by Görges et al. (2009), additional alarm delays for monitored physiological parameters (e.g., heart rate, respiratory rate) besides oxygen saturation must be explored. For example, the hypothetical effects of the application of an alarm delay for the heart rate low-and-high limit alarm in the context of arrhythmia alarms (bradycardia and tachycardia) should be tested for safety on a rich,

comprehensive database prior to manufacturers' submitting premarket notifications and to seeking approvals from the U.S. Food and Drug Administration.

It is this author's opinion that, when measuring intervention effectiveness, future alarm studies should use more statistical analysis that are more robust (rather than descriptive statistics such as percent change) in order to statistically accommodate the overdispersion of alarm counts and the large variation in alarm frequencies over time—factors that are predominately affected by individual patients (i.e., outliers) and that result in highly skewed alarm distributions. Gardner, Mulvey, and Shaw (1995), endorse the application of a negative binomial regression as a model that best fits skewed distributions, because this statistical technique facilitates understanding of particular situations in which many patients may experience no events (e.g., alarms) and few patients experience many events. Furthermore, the use of a negative binomial regression model provides optimum results when analyzing data containing excess zeros that cannot be excluded from the sampling plan (Coxe, West, & Aiken, 2008).

Furthermore, alarm analysis should be based on alarm rates (i.e., normalized per monitoring hours), because the alarm counts per patient are highly skewed and suggests that a relatively small proportion of patients trigger the majority of alarms (Gross et al., 2011). Moreover, research should focus on investigating patients and examine metrics such as adverse patient events (e.g., cardiopulmonary arrests, acute respiratory compromise), unanticipated transfers to a higher level of care, rapid response rescue events, and delays and/or complications in patient care—and not exclusively on the percent change of specific alarm outcome variables.

The growing body of research on alarm fatigue can be enriched by nurse scientists who conduct robust mixed-method studies involving a complementary blend of quantitative and qualitative data. This type of research can provide insight into the nature of nurses' work

experience in technologically intensive patient care environments, the impact of alarms on workflows, the approaches used by nurses to handle interruptions caused by physiologic monitor alarms, and strategies for responding to alarms. Furthermore, further research must be conducted to develop appropriate alarm management strategies for reducing clinically irrelevant physiologic monitor alarms for age-specific patient populations—especially, for pediatric populations, because most studies to date have been conducted in adult settings.

Conclusion

The study presented in this dissertation was the first prospective, randomized clinical trial to generate the highest level of evidence substantiating the effectiveness of a technology-and practice-based intervention for reducing physiologic monitor alarms in a neuroscience intensive care unit. The study demonstrated that these technology-based interventions (modification of SpO_2 unit default alarm setting) can effectively reduce non-actionable SpO_2 low-limit alarms that are short, that typically autocorrect after a few seconds, and that contribute to clinical alarm fatigue.

However, our investigation also yielded some surprising findings. Clinicians have long assumed that better ECG electrode management could lead to reduced alarm fatigue—indeed, this strategy has been viewed as yielding easily generable “low-hanging fruit”; however, our study’s results challenge the recommended practice of changing ECG electrodes on a daily basis. Existing practice guidelines recommending application of “fresh” electrodes daily are primarily based on quality improvement projects that used inappropriate instruments for collecting alarm outcome variable data and unsuitable statistics for subsequent analysis. Without rigorous research to measure the effectiveness of novel skin preparation and ECG electrode regimens,

clinicians' view of the effectiveness of such a basic nursing intervention may be exaggerated, resulting in clinicians' overvaluing or overlooking other potential interventions strategies to minimize nuisance alarms.

Reputable professional societies, industry leaders, and accrediting agencies are recognizing the importance of robust study designs for alarm research. In particular, ICU nurses have a strong vested interest in alarm research, given their constant presence at the patient's bedside. The need for additional nursing research to assess strategies for reducing the frequency of non-actionable alarms is clear. Greater attention must be given to identifying technology-based solutions for minimizing and preventing clinically irrelevant (i.e., nuisance) alarms that are associated with inappropriate or overly conservative threshold settings or false-positive arrhythmia alarms related to imperfect arrhythmia detection algorithms. In addition, scientists, device manufacturers, and clinicians should explore the use of alarm delays for other rate-related parameter alarms. Thus, research should examine all known physiologic monitor alarms associated with the development of alarm fatigue. This effort will require greater awareness on the part of the health care community, regulatory bodies, and accrediting agencies regarding the detrimental effects of nuisance alarms on the health and well-being of both clinicians and monitored patients.

Despite well-intentioned efforts to combat alarm fatigue, increased knowledge, preparation, and training are warranted for researchers, clinicians, biomedical engineers, and other members of the interprofessional team involved in alarm research and hospital-based quality assurance projects. Adequate preparation and expertise is essential to ensure that studies are rigorously executed and that findings and project outcomes are valid prior to the dissemination of results. Use of inappropriate statistical analytic techniques or tests can lead to

misinterpretation of study results, to subsequent unwarranted generalization of erroneous findings, and to formulation of incorrect and detrimental guidelines. As our study has revealed, enhancement of RN clinical practice is not a panacea for reducing excessive alarm rates and ameliorating alarm fatigue. Substantial investments in new alarm technologies such as the development of “smarter” alarms and improved arrhythmia detection software are overdue; however, given the complexity of alarm data, these complex efforts require time, human resources, and funding.

Efforts to reduce the frequency of physiologic monitor alarms and consequent alarm fatigue among clinicians requires a collaborative approach, innovation, and a commitment to improving the care and experience of the monitored patient. Only through the generation and dissemination of nursing research can the quality of physiological monitoring improve.

References

Aboukhalil, A., Nielsen L., Saeed, M., Mark, R. G., & Clifford, G. D. (2008). Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform. *Journal of Biomedical Informatics, 41*, 442–451.

American College of Cardiology., American Heart Association. , & European Society of Cardiology.(2006). Guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. *Journal of the American College of Cardiology, 48*(5), e247–e346.

Association for the Advancement of Medical Instrumentation (2011). *Clinical Alarms; 2011 Summit*. Retrieved from http://www.aami.org/htsi/alarms /pdfs/2011 _Alarms_Summit_publication.pdf

Biot, L., Carry, P. Y., Perdrix, J. P., Eberhard, A., & Baconnier, P. (2000). Évaluation clinique de la pertinence des alarmes en reanimation. *Annales Françaises d'Anesthésie et de Réanimation, 19*, 459–466.

Coxe, S., West, S. G., & Aiken, L. S (2008). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. *Journal of Personality Assessment, 9*(12), 121–136.

Cvach, M., Biggs, M., Rothwell, K. J., Charles-Hudson, C. (2012). Daily electrode change and effect on cardiac monitor alarms: An evidence based practice approach. *Journal of Nursing Care Quality. 28*(3), 265–271.

Emergency Care Research Institute (2007). The hazards of alarm overload. Keeping excessive physiologic monitoring alarms from impeding care. *Health Devices, 36*(3), 73–83.

Emergency Care Research Institute (2012). *Top 10 technology hazards for 2013*. Retrieved from https://www.ecri.org/Documents/Secure/Health _Devices_Top_10_Hazards_2013.pdf

Emergency Care Research Institute. (2013a). The life cycle of an alarm. *Health Devices, 42*(9), 294–300.

Emergency Care Research Institute. (2013b). Top 10 health technology hazards for 2014. *Health Devices, 42*(11), 354–380.

Fidler, R., Pickham, D., & Drew, B. J. (2011). Quantification of hospital cardiac monitoring alarms. *Journal of Electrocardiology, 44*, 744.

Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological Bulletin, 118*(3), 392–404.

General Electric Healthcare (2007) *Solar 8000 i/M: Patient monitor operator's manual software Version 5*. Waukesha, WI: GE Healthcare.

Görges, M., Markewitz, B. A., & Westenskow, D. R. (2009). Improving alarm performance in the medical intensive care unit using delays and clinical context. *Anesthesia & Analgesia, 108*(5), 1546–1552.

Graham, K., & Cvach, M. (2010). Monitor alarm fatigue: Standardizing use of physiological monitors and decreasing nuisance alarms. *American Journal of Critical Care, 19*(1), 28–34.

Gross, B., Dahl, D., & Nielsen, L. (2011). Physiologic monitoring alarm load on medical/surgical floors of a community hospital. *Biomedical Instrumentation & Technology, 45*(1), 29–36.

Lipton, J. A., van Ettinger, M. J. B., Barendse, R. J., van Dam, T. B., van der Putten, N. H. J. J., & Newlan, S. P. (2009). Alarms on the intensive cardiac care unit. *Computers in Cardiology, 36*, 253–256.

Medina, V., Clochesy, J. M., & Omery, A. (1989). Comparison of electrode site preparation techniques. *Heart & Lung, 18*(5), 456–460.

Meziane, N., Webster, J. G., Attari, M., & Nimunkar, A. J. (2013). Dry electrodes for electrocardiology. *Physiological Measurement, 34*(9), R47–R69.

O'Carroll, T. M. (1986). Survey of alarms in an intensive therapy unit. *Anaesthesia, 41*, 742–744.

Oster, C. D. (2000). Improving ECG trace quality. *Biomedical Instrumentation & Technology, 34*(3), 219–222.

Pan, P., & Gravenstein, N. (1994). Intraoperative Pulse Oximetry: Frequency and distribution of discrepant data. *Journal of Clinical Anesthesia, 6,* 491–495.

Paparella, S. F. (2014). Signals or noise: Maximizing the use of clinical alarm systems for patient safety. *Journal of Emergency Nursing, 40*(2), 169–171.

Patel, S. I., & Souter, M. J. (2008). Equipment-related electrocardiographic artifacts. *Anesthesiology, 108,* 138–148.

Rheineck-Leyssius, A. T., & Kalkman, C. J. (1998). Influence of pulse oximeter settings on the frequency of alarms and detection of hypoxemia. *Journal of Clinical Monitoring and Computing, 14*(3), 151–156.

Sapo, M., Wu, S., Asgari, S., McNair, N., Buxey, F., Martin, N., & Hu, X. (2009). A comparison of vital signs charted by nurses with automated acquired values using waveform quality indices. *Journal of Clinical Monitoring and Computing, 23,* 263–271.

Scott Allen, J., Hileman, K., & Ward, A. (2013). Simple solutions for improving patient safety in cardiac monitoring—eight critical elements to monitor alarm competency. *Safety*

Innovations. Retrieved from http://www.aami.org/htsi/SI_Series/Alarm_Competency %20_White_Paper.pdf

Sendelbach, S., & Funk, M. (2013). Alarm fatigue. *AACN Advanced Critical Care, 24*(4), 378–386.

Siebig, S., Kuhls, S., Imhoff, M., Gather, U., Schölmerich, J., & Wrede, C. (2010). Intensive care unit alarms—How many do we need? *Critical Care Medicine, 38*(2), 451–456.

Sun, J. X., Reisner, A. T., & Mark, R. G. (2006). A signal abnormality index for arterial blood pressure waveforms. *Computers in Cardiology, 33*, 13–33.

Taenzer, A. H., Pyke, J. B., McGrath, S. P., & Blike, G. T. (2010). Impact of pulse oximetry surveillance on rescue events and intensive care unit transfers. *Anesthesiology, 112*(2), 282–287.

The Cardiac Arrhythmia Suppression Trial (CAST) Investigators. (1989). Preliminary Report: Effect of Encainide and Flecainide on Mortality in a Randomized Trial of Arrhythmia Suppression after Myocardial Infarction. *New England Journal of Medicine, 321,* 406–412.

The Joint Commission. (2013). Alarm system safety. Retrieved from http://www.jointcommission.org/assets/1/18/R3_Report_Issue_5_12_2_13_Final.pdf

The Joint Commission. (2014). *National Patient Safety Goal on Alarm Management.* Retrieved from http://www.jointcommission.org/assets/1/18 /JCP0713 _Announce_New_NSPG.pdf

Tsien, C. L., & Fackler J. C. (1997). Poor prognosis for existing monitors in the intensive care unit. *Critical Care Medicine, 25* (4), 614–619.

Turkmen, A., & Pantiskas, C. (2011). Management of ECG cables and leadwires. *Biomedical Instrumentation & Technology, 45*(2), 130–133.

Way, R. B., Beer, A. A., & Wilson, S. J. (2014). What's that noise? Bedside monitoring in the emergency department. *International Emergency Nursing*. Retrieved from http://dx.doi.org/10.1016/j.ienj.2014.01.001

Welch, J. (2011). An evidence-based approach to reduce nuisance alarms and alarm fatigue. *Biomedical Instrumentation & Technology*, *45*(1), 46–52.

Wiklund, L., Hök, B., Ståhl, K., & Jordeby-Jönsson, A. (1994). Postanesthesia Monitoring Revisited: Frequency of True and False Alarms from Different Monitoring Devices. *Journal of Clinical Anesthesia*, *6*, 182–188.

Whalen, A. A., Covelle, P. M., Piepenbrink. J. C., Villanova, K. L., Cuneo, C. L., & Awtry, E.H. (2013). Novel approach to cardiac alarm management on telemetry units. *Journal of Cardiovascular Nursing*, 1–10.

Glossary of Alarm Terminology

Alarm burden: A term to quantify the number of alarms of a predetermined period of time (e.g., over a 24-hour period).

Alarm condition: State of the alarm system when it has determined that a potential or actual hazardous situation exists for which operator notification is required. An alarm condition can be invalid (i.e., a false positive alarm condition). An alarm condition can be missed (i.e., a false negative alarm condition).

Alarm condition delay: Time from the occurrence of a triggering event either in the patient, for physiological alarm conditions, or in the equipment, for technical alarm conditions, to when the alarm system determines that an alarm condition exists.

Alarm fatigue: When staff is exposed to an excessive number of alarms, this can result in sensory overload, causing staff to become desensitized to the alarms. Desensitization may result in delayed alarm response or missed alarms.

Alarm management: Orchestration of the culture, staff responsibilities, technology, policies and procedures, practices, and other factors, tasks, and processes that are required to support prompt and efficacious alarm verification, notification, response, and documentation.

Alarm prioritization: Visual and audible differentiation of alarms (e.g., life-threatening vs. other types of less serious events) in which the visual and auditory alarm prominence connotes the level of urgency with which clinicians should respond.

Alarm session: Defined as cluster of physiologic monitor alarms that occur simultaneously yet, only the alarm event with the highest alarm severity level is audible.

Alarm signal: Type of signal generated by the alarm system to indicate the presence (or occurrence) of an alarm condition.

Arrhythmia suspend: The arrhythmia suspend alarm condition occurs as a result of ECG artifact. Artifact alarms begin at level 1 and progress to level 2 if the ECG noise lasts for 20 of the last 30 seconds. Artifact level 2 alarm displays as an arrhythmia suspend alarm, and arrhythmia interpretation is completely suspended.

Artifact Alarm: The artifact alarm is a transient condition resulting from intermittent noise and artifact. When artifact level 1 alarms are triggered, full arrhythmia processing is suspended, yet the lethal arrhythmias detection software remains active.

False alarms: An alarm detected by a medical device or system that indicates a need for a response to a physiologic event when a not true event has occurred.

False negative alarm condition: Absence of an alarm condition when a valid triggering event has occurred in the patient, the equipment or the alarm system. Note an alarm condition can be

rejected or missed because of spurious information produced by the patient, the patient-equipment interface, other equipment or the equipment itself.

False positive alarm condition: Presence of an alarm condition when no valid triggering event has occurred in the patient, the equipment or the alarm system. A false positive alarm condition can be caused by spurious information produced by the patient, the patient-equipment interface, other equipment or the alarm system itself.

Latching alarms: An alarm signal that continues to be generated after its triggering event no longer exists. This implies that the alarm signal (audio alarm) continues to annunciate—and requires the clinician to acknowledge the alarm by pressing the silence button.

Non-actionable alarms: Alarms that correctly sound, but for an event that has no clinical relevance.

Non-latching alarms: An alarm signal that automatically ceases being generated when its triggering event no longer exists.

Nuisance alarms: Alarms, perceived by staff to be annoying, that may interfere with patient care, and typically do not result from adverse or potential adverse patient conditions. Nuisance alarms become a problem because alarm signals can distract caregivers from other tasks despite there not being any real patient condition requiring attention and can contribute to alarm fatigue.

Patient status alarms: Patient status alarms are triggered by a patient condition that exceeds parameter limits or by an arrhythmia condition. Patient status alarms provide the user with the highest priority information.

Sensitivity: Refers to the likelihood that an alarm will correctly signal a true-positive event.

Specificity: Refers to the likelihood that an alarm will appropriately remain silent during a true-negative event.

System status alarms: System status alarms are triggered by mechanical or electrical problems and are of lesser priority than patient status alarms.

Technical alarms: An alarm event caused by a monitored equipment-related or alarm system-related variable.

References

Association for the Advancement of Medical Instrumentation (2011). *Clinical Alarms: 2011 Summit*. Retrieved from http://www.aami.org/htsi/alarms/pdfs/2011_Alarms_Summit_publication.pdf

Association for the Advancement of Medical Instrumentation (2011). *Clinical Alarms: 2011 Summit*. Retrieved from http://www.aami.org/meetings/summits/alarms/Materials/PDFs/Drew_Tues.pdf

Emergency Care Research Institute. (2012a). *Medical device alarm related terminology*. Retrieved from http://www.aami.org/meetings/summits/alarms.html

APPENDICES

Appendix A

Physiologic Monitor Alarms: Annotation Plan (With permission of B. J. Drew, 2012)

CRISIS ECG ALARMS			
Alarm Condition	**Arrhythmia Algorithm**	**Potential Cause of False Alarm**	**Proof of True/False Alarm by Investigator**
1. ASYSTOLE	Displayed heart rate drops to zero. No QRS detected for 5-6 seconds	a. Monitor is not detecting sufficient QRS amplitude in analysis leads (I, II, III, & V). b. Noisy signal	Asystole True Alarm Proof: (either condition would confirm true alarm) 1. Simultaneous drop in ***invasive*** arterial or PA pressure to near zero (abrupt decrease in pressure waveform amplitude to near isoelectric line); cannot use non-invasive BP 2. Code Blue documentation of asystolic or PEA arrest at same time (<5 sec asystole would not be expected to cause loss of consciousness/Code Blue so asystole must persist) 3. Confirm that asystole lasts at least 5 seconds with e-calipers 4. If rhythm is determined to be low amplitude VF, count asystole alarm as true Asystole False Alarm Proof: (either condition would confirm FA) 1. There is no simultaneous drop in ***invasive*** arterial or PA pressure (abrupt decrease in pressure waveform amplitude) 2. There is a visible QRS ***in at least one lead*** (may be low amplitude and barely visible; must examine all available [7] leads) 3. ASYSTOLE episode lasts >60 seconds but no Code Blue or other documentation that it was recognized clinically (syncope, seizure, LOC) 4. ASYSTOLE episode lasts >60 seconds but no decrease in SpO_2 waveform amplitude
2. V-FIB/ V-TACH	Course flutter waves without QRS complexes	a. Noisy signal (motion, electrical interference, or other artifact)	VFIB/VTAC True Alarm Proof: (either condition would confirm true alarm) 1. Simultaneous drop in ***invasive*** arterial or PA pressure to near zero (abrupt decrease in pressure waveform amplitude to near isoelectric line) 2. Code Blue documentation of VF or VT arrest at same time VFIB/VTAC False Alarm Proof: (any of the following conditions would confirm FA) 1. There is no simultaneous drop in ***invasive*** arterial or PA pressure (abrupt decrease in pressure waveform amplitude) 2. There are QRS complexes at the same rate as the patient's normal rhythm visible throughout a noisy signal ***in any lead*** (check RR intervals before, during, after event to see if they "march through") 3. VFIB/VTAC episode lasts >60 seconds but no Code Blue or other documentation that it was recognized clinically (syncope, seizure, LOC) 4. VFIB/VTAC episode lasts >60 seconds but no decrease in SpO_2 waveform amplitude
3. V-TACH	≥6 consecutive PVCs with rate ≥100 bpm	a. Motion or other artifact (rapid repetitive motion by patient as in brushing teeth or scratching an electrode) b. Event is due to a supra-ventricular tachycardia (SVT) in a patient with a pre-existing right or left bundle branch block (BBB)	VTACH True Alarm Proof: (any condition would confirm true alarm) 1. Simultaneous drop in ***invasive*** arterial or PA pressure 2. Code Blue documentation of VT 3. Stat 12-lead ECG documentation of VT read by cardiologist 4. AV dissociation (sinus P waves can be seen "marching through" VTACH) 5. QRS morphology that is different than patient's underlying rhythm with BBB VTACH False Alarm Proof: (any condition would confirm FA) 1. There is no simultaneous drop in invasive arterial or PA pressure (abrupt decrease in pressure waveform amplitude; if it is "slow" VT with rate 100-150, pressure may not drop to near zero; however, there will be a visible decrease in pressure waveform amplitude) 2. There are QRS complexes at the same rate as the patient's normal rhythm visible throughout a noisy signal in any lead (check RR intervals before, during, after event to see if they "march through") 3. VTACH episode lasts >60 seconds but no decrease in SpO_2 waveform 4. Event is due to SVT in a patient with pre-existing BBB (the patient's dominant rhythm has a RBBB or LBBB; the same BBB morphology must be identified during the event ***in all 7 leads***; if one lead differs, it is VT, not SVT). Additional confirmation is seeing a premature P wave initiating the tachycardia.

Appendix A. *(continued)*

WARNING ECG ALARMS			
Alarm Condition	**Arrhythmia Algorithm (EK-Pro definition)**	**Potential Cause of False Alarm**	**Proof of True/False Alarm by Investigator**
4. ACC VENT	≥6 ventricular beats with HR 50-100 bpm	a. Patient has ventricular paced beats; however, pacemaker mode has not been activated b. Event is due to a supra-ventricular rhythm in a patient with a pre-existing right or left BBB	ACC VENT False Alarm Proof: (either condition would confirm FA) 1. Event is ventricular pacing (DETECT PACE may be off so no pacer spikes are visible) a. Patient is known to have ventricular pacemaker ***and*** b. Event QRS morphology is identical to paced rhythm on 12-lead ECG or prior monitoring rhythm in all available (7) ECG leads b. Event is due to sinus rhythm with intermittent or new onset BBB (P waves prior to each wide beat with normal PR interval)
5. PAUSE	3-second interval without a QRS complex	a. Monitor is not detecting sufficient QRS amplitude in analysis leads (I, II, III, & V)	PAUSE False Alarm Proof: (either condition would confirm FA) 1. No simultaneous drop in ***invasive*** arterial or PA pressure 2. During the pause, there is a visible QRS (may be low amplitude and barely visible) ***in any of the 7 available leads***
6. VBRADY	≥3 consecutive ventricular beats with HR ≤50 bpm	a. Patient has ventricular paced beats; however, pacemaker mode has not been activated b. Patient has BBB with rate ≤50 bpm due to sinus brady or atrial fibrillation with slow ventricular rate	VBRADY False Alarm Proof: (either condition would confirm FA) 1. Event is ventricular pacing (DETECT PACE may be off so no pacer spikes are visible) a. Patient is known to have ventricular pacemaker set at this slow rate (e.g., 50 may be the low rate limit set on a DDD pacemaker) ***and*** b. Event QRS morphology is identical to paced rhythm on 12-lead ECG or prior monitoring rhythm in all available (7) ECG leads 2. Event is due to sinus rhythm with intermittent or new onset BBB (P waves prior to each wide beat with normal PR interval)

Appendix B

Alarm Study Plan for Counting, Analyzing and Reporting GE Physiologic Monitor Alarms (by permission of B.J. Drew, 2014)

Arrhythmia (Patient Status) Alarms		
Our Label	**BMEX Label** Include Pace Mode 1 & 2	**Definition**
Acc Vent	Acc Vent	≥6 ventricular beats with an average heart rate for the ventricular beats between 50-100 bpm (adult HR)
Afib	1. Afib 2. Irregular	Irregular timing of QRS complexes and absence of P waves preceding the QRS complex. Note: Afib and Irregular are mutually exclusive so alarms will not be double counted (when Afib is enabled, irregular does not occur)
Asystole	Asystole	The displayed HR drops to zero; alarm delay depends upon HR (higher HR takes longer to drop to zero); typically no QRS for 5-6 seconds
Brady	Brady	HR below user-defined low limit setting; average of the most recent 8 R-to-R intervals that fall below the low limit setting.
Pause	Pause	No QRS for a 3-second interval
Tachy	Tachy	HR above user-defined high limit setting; average of the most recent 8 R-to-R intervals that fall above the high limit setting.
V Brady	V Brady	≥3 consecutive ventricular beats at an average rate ≤50 bpm (adult HR)
Vfib/Vtac	Vfib/Vtac	Course flutter waves without QRS complexes
Vtach	Vtach	≥6 consecutive ventricular beats at rate ≥100 bpm (adult HR)
All PVCs	Combine all the alarms below to create a Drew Lab new variable	
	VT>2	3-5 consecutive ventricular beats at rate ≥100 bpm (adult HR)
	PVC	Isolated PVCs: a ventricular beat has non-ventricular beats before and after
	R on T	A ventricular beat (PVC) falls on the ST or T wave portion of the previous non-ventricular beat (normal QRS beat)
	Couplet	Two consecutive PVCs with rate >100 (coupling interval <600 milliseconds)
	Bigeminy	A ventricular beat is followed by a non-ventricular beat for ≥3 cycles; e.g., PVC-normal QRS; PVC-normal QRS; PVC-normal QRS
	Trigeminy	A ventricular beat is followed by 2 non-ventricular beats for ≥3 cycles

Parameter (Patient Status) Alarms		
Our Label	**BMEX Label** Include PaceMode1-2	**Definition**
All HR	Combine all the alarms below to create a Drew Lab new variable that represents heart rate is outside high or low limit settings	
	HR>X	
	HR<X	
	HR=X	
	SpO2 Rate>X	
	SpO2 Rate<X	
	SpO2 Rate=X	
	ART Rate = X	
	ART Rate > X	
	ART Rate < X	
	FEM Rate = X	
	FEM Rate > X	
	FEM Rate < X	
All RR	Combine all the alarms below to create a Drew Lab new variable; accept all RR alarms regardless of respiratory sensitivity setting (e.g., 40%, 20%, etc.)	
	RESP >X	Respiratory rate per minute is above user-defined high limit setting
	RESP <X	Respiratory rate per minute is below user-defined low limit setting
	RESP = X	
	APNEA >X	No breaths have been detected for a period of seconds that the user defines; e.g., APNEA >20 means no breaths detected for >20 seconds
All SpO2	Combine all the alarms below to create a Drew Lab new variable	
	SpO2<X	
	SpO2>X	
	SpO2=X	
All ST	Combine all the alarms below to create a Drew Lab new variable	
	ST-I >X	ST amplitude measured at user-defined position (e.g., J+60 ms; J+80 ms) is greater than the reference point (PR segment level) by a user-defined setting; e.g., ST high limit of 2 mm in this single ECG lead
	ST-I <X	ST amplitude measured at user-defined position (e.g., J+60 ms; J+80 ms) is less than the reference point (PR segment level) by a user-defined setting; e.g., ST low limit of minus 2 mm in this single ECG lead
	ST-II >X	Same as above
	ST-II <X	Same as above
	ST-III >X	Same as above
	ST-III <X	Same as above
	ST-aVR >X	Same as above
	ST-aVR <X	Same as above
	ST-aVL >X	Same as above
	ST-aVL <X	Same as above
	ST-aVF >X	Same as above
	ST-aVF <X	Same as above
	ST-V >X	Same as above
	ST-V <X	Same as above
	ST-V1 >X	Same as above

Parameter (Patient Status) Alarms		
Our Label	**BMEX Label** Include PaceMode1-2	**Definition**
	ST-V1 <X	Same as above
	ST-V2 >X	Same as above
	ST-V2 <X	Same as above
	ST-V3 >X	Same as above
	ST-V3 <X	Same as above
	ST-V4 >X	Same as above
	ST-V4 <X	Same as above
	ST-V5 >X	Same as above
	ST-V5 <X	Same as above
	ST-V6 >X	Same as above
	ST-V6 < X	Same as above
PVC Rate	PVC = X	PVC count is equal to user-defined limit; e.g., if count limit setting is 20, this alarm indicates the patient had 20 PVCs over the past minute
	PVC > X	PVC count exceeds the user-defined limit; e.g., if count limit setting is 20, this alarm indicates the patient had >20 PVCs per minute
All ART	Combine all the alarms below to create a Drew Lab new variable for **invasive** arterial blood pressure	
	ART Sys = X	
	ART Sys > X	
	ART Sys < X	
	FEM Sys = X	
	FEM Sys > X	
	FEM Sys < X	
	ART Dia = X	
	ART Dia > X	
	ART Dia < X	
	FEM Dia = X	
	FEM Dia > X	
	FEM Dia < X	
	ART Mean = X	
	ART Mean > X	
	ART Mean < X	
	FEM Mean = X	
	FEM Mean > X	
	FEM Mean < X	
All NIBP	Combine all the alarms below to create a Drew Lab new variable for **non-invasive** arterial blood pressure	
	Npb Sys = X	
	Npb Sys > X	
	Npb Sys < X	
	Npb Dia = X	
	Npb Dia > X	
	Npb Dia < X	

Parameter (Patient Status) Alarms		
Our Label	**BMEX Label** Include PaceMode1-2	**Definition**
	Npb Mean = X	
	Npb Mean > X	
	Npb Mean < X	
All Heart pressures	Combine all the alarms below to create a Drew Lab new variable for central, intra-cardiac and pulmonary artery pressures (CVP=central venous pressure; RAP = right atrial pressure; LAP = left atrial pressure; PAP = pulmonary artery pressure)	
	CVP Sys = X	
	CVP Sys > X	
	CVP Sys < X	
	CVP Dia = X	
	CVP Dia > X	
	CVP Dia < X	
	CVP Mean = X	
	CVP Mean > X	
	CVP Mean < X	
	PA Sys = X	
	PA Sys > X	
	PA Sys < X	
	PA Dia = X	
	PA Dia > X	
	PA Dia < X	
	PA Mean = X	
	PA Mean > X	
	PA Mean < X	
	LAP Mean = X	
	LAP Mean > X	
	LAP Mean < X	
	RAP Mean = X	
	RAP Mean > X	
	RAP Mean < X	
All ICP	Combine all the alarms below to create a Drew Lab new variable for **invasive** intracranial pressure	
	ICP Mean = X	
	ICP Mean > X	
	ICP Mean < X	

Technical (System Status) Alarms		
Our Label	**BMEX Label**	**Definition**
All Technical	Combine all alarms below to create a Drew Lab new variable that indicates a problem with electrodes or sensors (e.g., pulse ox probe).	
	Artifact	ECG artifact (noisy signal) is detected
Arrhythmia suspend	Arrh Suspend	When artifact lasts for 20 of the last 30 seconds, ARRHY SUSPEND occurs
	Arr Off	
	ECG Leads Fail	
RR Leads Fail	Resp Leads Fail	
	No ECG	
	ART Sensor Fail	
	ICP Sensor Fail	
	FEM Sensor Fail	
	RAP Sensor Fail	
	SP Sensor Fail	
	LAP Sensor Fail	
	CVP Sensor Fail	
	PA Sensor Fail	
	UVC Sensor Fail	
	UAC Sensor Fail	
	SpO2 Connect Probe	
	SpO2 Probe Off	
	SpO2 Probe Fail	
	SpO2 Low Sig	
	SpO2 Incompatible Cable	
	SpO2 Interf Def	
	Nbp Invalid Command	
	Nbp Excessive Pressure 200	
	Nbp Exceeded 3 min	
	Nbp Deflation Failure	
	Nbp Inflation Time Exceeded	
Nursing Intervention	Combine all system status alarms below to create a Drew Lab new variable that indicates interruption of monitoring due to nursing intervention (e.g., drawing arterial blood gas sample from the arterial line). These alarms are preventable because the RN could silence the alarm first).	
	PA Art Line Disconnect	
	CVP Art Line Disconnect	
	ART Art Line Disconnect	
	FEM Art Line Disconnect	

UCSF Library Release

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my theses, dissertation, or manuscripts to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Tina Hammond
Author Signature

June 11, 2014
Date