2016

An Integrative Review Focusing on Accuracy and Reliability of Clinical Thermometers

Julie Black Black
Walden University

Walden University

College of Health Sciences

This is to certify that the doctoral study by

Julie Black

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Eileen Fowles, Committee Chairperson, Health Services Faculty
Dr. Catherine Harris, Committee Member, Health Services Faculty
Dr. Jonas Nguh, University Reviewer, Health Services Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2016

Abstract

An Integrative Review Focusing on Accuracy and Reliability of Clinical Thermometers

by

Julie A. Black

MS, University of Maryland, 1999

BS, Azusa Pacific University, 1986

Project Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Nursing Practice

Walden University

May 2016

Abstract

Technological advances in clinical thermometers have resulted in a variety of minimally invasive devices that give rapid results but may not have the accuracy necessary for use in acutely ill adults. Inaccurate temperatures can result in missed opportunities for the early identification and treatment of infection and sepsis. Following the methodology outlined by Whittemore and Knafl, the purpose of this project was to conduct an integrative review of the research on the accuracy of clinical thermometers used for acutely ill adults. The evidence was categorized using the Hierarchy of Evidence for Interventional Studies, and the quality of the studies was appraised using the indicators described by Hooper and Andrews. Forty-seven studies met the inclusion criteria; the findings on device accuracy were contradictory. Device accuracy was found in 10 (n = 27) studies on the tympanic (TM), 2 (n = 8) on the chemical dot (CH), 7 (n = 19) on the temporal artery (TAT), and 3 (n = 13) on the axillary (AX) thermometers. Two of 2 studies found the no-touch (NT) device clinically inaccurate. Diagnostic accuracy was found in 3 (n = 8) and 0 (n = 5) studies on the TM and TAT, respectively. Only 22 studies had an acceptable quality grade of A or B, limiting the validity of the evidence. The evidence did not support the use of the NT and TAT thermometers or the AX route for acutely ill adults. The CH device should be use with caution, and abnormal temperatures should be validated with a more reliable device. For thermometers in use, appropriate training and technique are essential for the most accurate results. Closing the knowledge-to-practice gap on clinical thermometers can change the culture of nursing practice, improve early sepsis identification, and increase the quality of patient care.

An Integrative Review Focusing on Accuracy and Reliability of Clinical Thermometers

by

Julie A. Black

MS, University of Maryland, 1999

BS, Azusa Pacific University, 1986

Project Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Nursing Practice

Walden University

May 2016

Dedication

Above all, this project is dedicated to my heavenly Father and to my savior, Jesus Christ. Hebrews 12:1-2 says "…let us run with endurance the race set before us, fixing our eyes on Jesus…" and we are assured in Philippians 4:13 that "I can do all things through Christ who strengthens me." It is through God's grace that I have finished this race and am ready to embark on the next leg of this journey

This project is also dedicated to my husband and best friend, Byron Black. No one accomplishes these monumental journeys without the love and support of family and friends. Without Byron's love and endless encouragement, I would not have persevered to the end.

Acknowledgments

I would like to acknowledge, with the deepest appreciation, the time, effort, and counsel of my committee chair, Dr. Eileen Fowles. The road to the completion of this project has been filled with many passages, by-ways, and roadblocks. Dr. Fowles helped me navigate each of these challenges with the calm assurance that I would be successful in the end. I would also like to acknowledge my committee member, Dr. Catherine Harris, and my URR, Dr. Jonas Nguh, for their feedback and help through this project.

A special thank you to my friends and colleagues that have prayed for me as I struggled my way from class to class and paper to paper. Without your love, support, and prayers, I would not have reached this goal.

Table of Contents

List of Tables

List of Figures

Section 1: Nature of the Project

Introduction

The assessment of temperature as a marker for illness has been identified in the literature as early as 1592 (Pearce, 2002). Boerhaave (1668–1738) pioneered the use of the clinical thermometer at the bedside and correlated temperature with illness progression (Pearce, 2002). In 1871, Wunderlich wrote, "a knowledge of the course of the temperature in disease is highly important to the medical practitioner, and, indeed, indispensable" (p. vi.). Wunderlich (1871) documented observations about temperature that still have implications today. One observation was that an abnormal temperature was something that could not be contrived or faked. Because temperatures could not be faked, one could conclude that there was a physiologic disturbance simply from the change in the temperature. Lastly, Wunderlich also identified that the observation of changes in temperature could provide information about the course of some diseases. Using the thermometer quite extensively in medical practice, Wunderlich performed over 1 million temperatures readings in 25,000 patients (Pearce, 2002). According to Pearce (2002), Wunderlich was the first person to identify the normal range of temperature as 36º C to 37º C.

In the early 1800s, the foot-long size of the thermometer remained a significant barrier for use (Pearce, 2002). By 1852, further advances in the thermometer were accomplished by adding a bulb reservoir for the mercury and narrowing the column (Pearce, 2002). Allbutt (1836–1925) reduced the size of the mercury thermometer to one that was six inches, and with it, the advent of routine temperature assessment in clinical

practice had begun (Pearce, 2002). The mercury thermometer has been the gold standard for routine temperature assessments until medical and environmental concerns related to the mercury pushed the development of various electronic, digital, and infrared clinical thermometers (Davie & Amoore, 2010).

Currently, there are a wide range of clinical thermometers used for hospitalized adult patients. Thermometers that can be used in any clinical area include oral (O), rectal (R), tympanic membrane (TM), axillary (AX), temporal artery (TAT), and no-touch (infrared; NT). In critical care or the operating room, more invasive devices may be used such as the esophageal (ES), bladder (BL), and pulmonary artery catheter (PAC) thermometers. Clinical thermometers may be chosen for their novelty, convenience, rapidity, and lack of invasiveness for the patient often without knowledge of differences in accuracy (Dunleavy, 2010; Ostrowsky, Ober, Wenze, & Edmond, 2003). Factors which can impact accuracy or reliability include (a) device characteristics and configuration, (b) patient characteristics and physiology, (c) user technique, and (d) calibration and maintenance (Davie & Amoore, 2010).

Problem Statement

Temperature assessment is integral to the care of all hospitalized adult patients. Imprecise temperature measurements may lead to unrecognized infection, increased morbidity and mortality, and increased health care costs (Dellinger et al., 2012; Hall, Williams, DeFrances, & Golosinskiy, 2011). In the clinical environment, nurses may choose to use a thermometer because of convenience (rapid results, noninvasive), the ease of operating the device, or because no other thermometers are available (bulk

purchase by the organization). When staff nurses at the former practicum site were asked to clarify why they chose a specific thermometer or route, they were unable to specify any related evidence to support their practice. These practices demonstrated a knowledge-to-practice gap related to the use of clinical thermometers used on adult hospitalized patients.

Additionally, the organization had identified the early identification and treatment of infection and sepsis a system-wide priority. The importance of accurate temperature assessment in early sepsis identification, as described by Dellinger et al. (2012), provided the foundation for a clinical inquiry related to clinical thermometers. The evidence-based practice (EBP) model used by the organization is the Iowa model for evidence-based practice (hereafter referred to as the Iowa model; Titler et al., 2001). The Iowa model was used to provide the structure for the steps in this clinical inquiry, specifically the comprehensive review of the literature. The comprehensive review of the literature was developed as an integrative review (IR) for the DNP project.

The IR can have a significant impact on the field of nursing practice. An IR can provide a synthesis of past research on a topic of interest and a summary on the recommendations (Russell, 2005). An IR allows for the inclusion of both experimental and nonexperimental evidence in order to obtain the best understanding of the problem or the clinical question (Whittemore & Knafl, 2005). The evidence in this body of literature may provide clarity related to which devices are the most accurate and reliable for clinical use.

Purpose

The purpose of this project was to conduct an IR of the body of research related to the accuracy and reliability of clinical thermometers. The IR review provides EBP information to narrow the knowledge-to-practice gap identified in the clinical environment. Also, a synthesis of the evidence will facilitate organizational decision-making regarding which devices are best for early sepsis identification. The guiding clinical practice question for this IR was: For adult patients in acute care hospitals, which clinical thermometer or thermometers provide the most accurate and reliable temperature readings?

Although there is a large body of research on the accuracy and reliability of many devices, nurses may not be knowledgeable about the thermometers they are using in their environment. The knowledge gap related to temperature devices may result in inaccurate temperature measurements leading to missed opportunities to identify an early infection. An IR of the pertinent body of literature may help to narrow this significant knowledge-to-practice gap.

Nature of the Doctoral Project

The DNP project consisted of an IR of the existing research on the accuracy and reliability of clinical thermometers. Russell (2005) defined an IR as "one in which past research is summarized by drawing overall conclusions from many studies" (p. 8). A systematic and comprehensive review of the research was conducted by accessing computerized databases such as CINAHL, MEDLINE, and ProQuest. Search methodology, search terms, and results are discussed in Section 3.

The methodology for IR, described by Whittemore and Knafl (2005), was used as a framework for the review. Additionally, the available research was categorized and analyzed using the levels of evidence described by Fineout-Overholt, Melnyk, Stillwell, and Williamson (2010). The IR provides a resource for nurses and leaders to help narrow the knowledge-to-practice gap observed within the organization. In addition, the results from the IR will also support leadership decision-making related to clinical thermometers used within the organization.

Significance

Temperature assessment is a standard of care in all areas of nursing practice. My former practicum site (part of a five-hospital system) identified early recognition and treatment of sepsis as a system-wide organizational priority. Sepsis is a significant health concern that can occur in any hospitalized patient; without early identification and targeted interventions, the mortality rate can be as high as 20% (Dellinger et al., 2012). Abnormalities in temperature (< 37ºC or > 37ºC), together with other clinical indicators, have been identified as a potential marker for infection or sepsis (Davie & Amoore, 2010). The annual cost associated with treating sepsis (as of 2008) was estimated to be approximately $14.6 billion (Hall et al., 2011). Kumar et al. (2006) determined that for every hour in which there is a delay in treatment, patient mortality increases 7.6%. Therefore, the accuracy of the assessed temperature is key in the early identification and treatment of sepsis and is critical to survival.

Given the importance of temperature as part of recurring vital sign assessments, the devices used in one's organization should provide the most accurate and reliable

results (Flynn-Makic, VonRueden, Rauen, & Chadwick, 2011). In many institutions, changes in temperature may result in a cascade of diagnostic studies in order to identify potential infection; these can be costly to both the patient and the organization (Flynn-Makic et al., 2011). The IR of the research on clinical thermometers provides organizational leaders with critical information related to decision-making about any potential changes in the devices used within the system. Additionally, this IR supports nursing practice and clinical decision-making in other acute care hospitals concerned with questions about accuracy and reliability of the clinical thermometers.

Implications for Social Change

Closing the knowledge-to-practice gap concerning temperature assessment devices and their accuracy and reliability has significant implications for changing the culture of nursing practice. At the organizational level, effective temperature assessments provide data that can reduce morbidity and improve patient care. Safe, quality patient care is a fundamental tenet in healthcare, as is our mandate to sustain these processes while striving to mitigate increasing healthcare costs (Zaccagnini & White, 2011). Without the appropriate use of EBP in the area of temperature assessment, devices are often chosen for the novelty, the convenience, or the noninvasive nature of the device (Manian & Griesenaur, 1998). Furthermore, many nurses presume a device is accurate and reliable simply because it is adopted by an organization (Ostrowsky et al., 2003). Often it is the nursing staff of healthcare organizations that raise safety concerns about a device and are change agents and advocates for their patients (Bahr, Senica, Gringas, & Ryan, 2010; Dunleavy, 2010; Ostrowsky et al., 2003).

Summary

The relevance of temperature assessment in all areas of nursing is apparent. In the early identification and treatment of sepsis, temperature accuracy is even more important (Birriel, 2013; Dellinger et al., 2012). An IR of this body of research provides nurses and nurse leaders with an evidence-base resource on clinical thermometers and helps to narrow the research-to-practice gap for this organization. As organizational leaders consider the implications of temperature inaccuracy and missed opportunities to identify infection, an IR provides additional evidence to support any recommended device changes.

In Section 2, I provide a thorough description of the models and methods used to inform this project. Additionally, I discuss the relevance of this problem to nursing practice. Finally, I describe the local context for the project and my role as a DNP student in the development of this project.

Section 2: Background and Context

Introduction

The accuracy of temperature assessment in the adult hospitalized patient is an important factor in the early identification and treatment of sepsis. The purpose of this project was to conduct an IR of the body of research related to accuracy and reliability of clinical thermometers. The clinical practice question was: For adult patients in acute care hospitals, which clinical thermometer or thermometers provide the most accurate and reliable temperature readings?

In the following section of this study, I examine the concepts and models that were used to guide the project. The models include a discussion on the IR and the methodology described by Whittemore and Knafl (2005) and the levels of evidence proposed by Fineout-Overholt et al. (2010). Additionally, I used the quality indicators and quality score described by Hooper and Andrews (2006) and the Iowa model for EBP (Titler et al., 2001) to develop this review. Also included in the following section is a discussion on the relevance of this project to nursing practice and my role as a DNP student in conducting the project.

Concepts, Models, and Theories

Methodology

This project was an IR of the existing literature on the accuracy and reliability of clinical thermometers. The IR methodology developed by Whittemore and Knafl (2005) was used to provide the framework for this project. Whittemore and Knafl stated that the IR is the broadest type of research review and can incorporate both experimental and

nonexperimental design. The inclusion of different types of research can lead to a more robust understanding of the project question (Whittemore & Knafl).

The strategy for the IR described by Whittemore and Knafl (2005) consisted of five stages:

1. Problem identification – includes clear problem identification and identification of the variables of interest.
2. Literature search – specific search strategies, search terms, computerized data bases and the means to identify literature not found with computerized search. These methods include reference reviews and research registries.
3. Data evaluation – determination of the quality of each study.
4. Data analysis – a synthesis of the evidence using (a) data reduction, (b) data display, (c) data comparison, (d) conclusion drawing, and (e) verification.
5. Presentation – findings of the review are presented; conclusions are supported by the evidence.

Hierarchy of Evidence

I used the Hierarchy of Evidence for Interventional Studies (see Figure 1) described by Fineout-Overholt et al. (2010) to describe the levels of evidence for this body of research. The seven levels describe the strength of the research. The categories identify the strength of the evidence from the highest, level I evidence, to the lowest, level VII evidence.

Hierarchy of Evidence for Intervention Studies

Type of evidence	Level of evidence	Description
Systematic review or meta-analysis	I	A synthesis of evidence from all relevant randomized controlled trials.
Randomized controlled trial	II	An experiment in which subjects are randomized to a treatment group or control group.
Controlled trial without randomization	III	An experiment in which subjects are nonrandomly assigned to a treatment group or control group.
Case-control or cohort study	IV	Case-control study: a comparison of subjects with a condition (case) with those who don't have the condition (control) to determine characteristics that might predict the condition. Cohort study: an observation of a group(s) (cohort[s]) to determine the development of an outcome(s) such as a disease.
Systematic review of qualitative or descriptive studies	V	A synthesis of evidence from qualitative or descriptive studies to answer a clinical question.
Qualitative or descriptive study	VI	Qualitative study: gathers data on human behavior to understand *why* and *how* decisions are made. Descriptive study: provides background information on the *what, where,* and *when* of a topic of interest.
Expert opinion or consensus	VII	Authoritative opinion of expert committee.

Figure 1. Hierarchy of Evidence for Interventional Studies. Adapted from "Evidence-based practice, step by step: Critical appraisal of the evidence Part III," by E. Fineout-Overholt, B. M. Melnyk, S. B. Stillwell, and K. M. Williamson, 2010, American Journal of Nursing, 110, p. 48. Copyright 2010 by Wolters Kluwer Health, Inc. Used with permission (obtained 1/7/2016)

The Iowa model. According to Taylor-Piliae (1999), "evidence-based practice aims to establish clinical practice based on scientific findings…and has the potential to influence practice and education" (p. 357). Medicine and physician training has been grounded in EBP and the synthesis of available evidence to guide practice (Taylor-Piliae, 1999). While there are large bodies of research evidence available to inform the practice of nursing, the use of or ability of nurses to utilize this research has been limited (Hicks & Hennessy, 1997). EBP models, such as the Iowa model, were developed to narrow this

research to practice gap (Taylor-Piliae, 1999; Zaccagnini & White, 2011). The Iowa model (see Figure 2) was originally developed in 1994 and was designed to guide nurses and other health care professionals to facilitate the use of research to improve patient care (Titler et al., 2001).

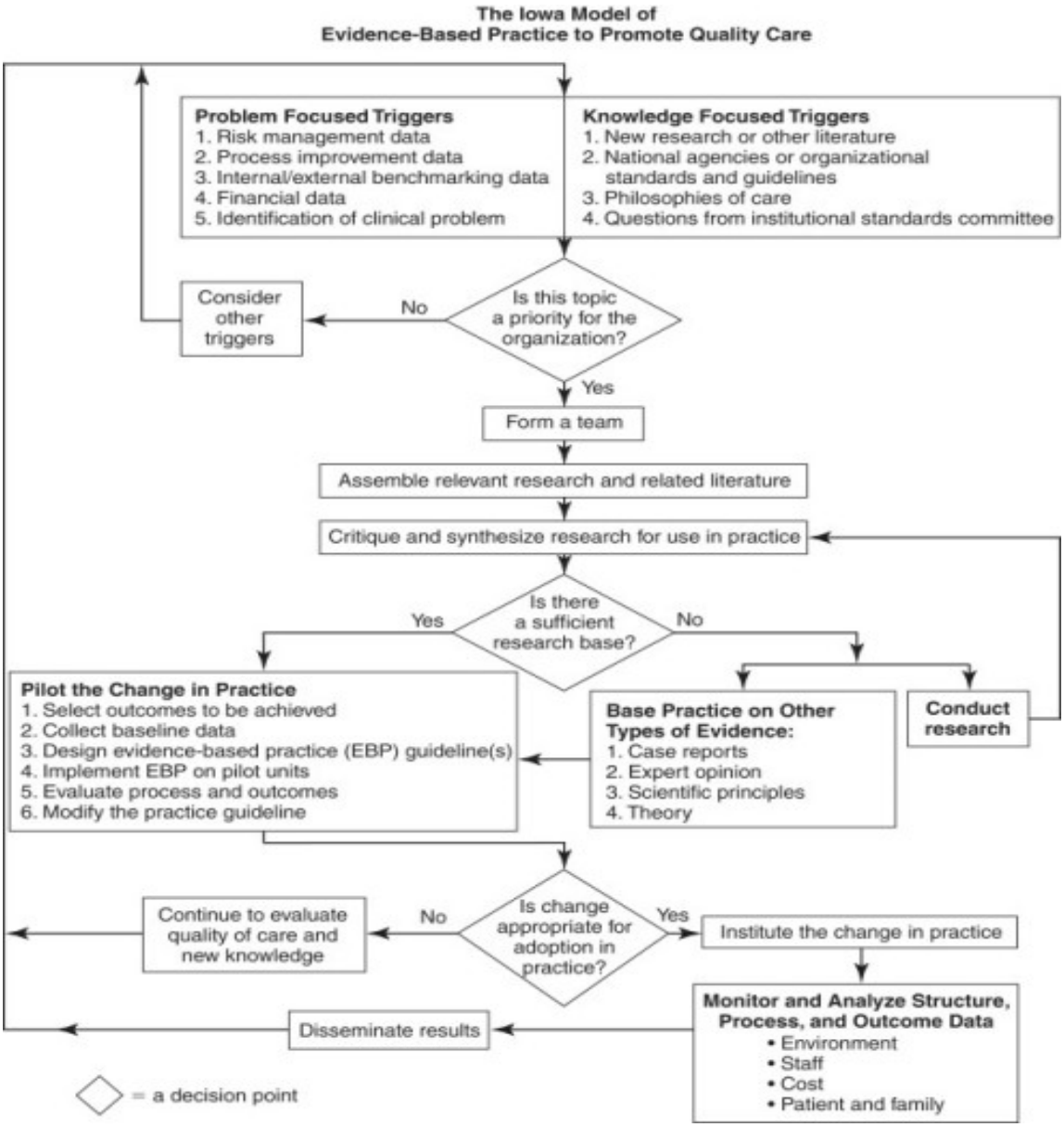

Figure 2. Iowa model for evidence-based practice. Adapted from "The Iowa model of evidence-based practice to promote quality care," by M. G., Titler, C. Kleiber, V. J. Steelman, B. A. Rakel, G. Budreau, L. Q. Everett,…C. J. Goode, 2001*, Critical Care Clinics of North America, 13*, p. 499. Copyright 1998 by the University of Iowa Hospitals and Clinics and Marita G. Titler, PhD, RN, FAAN. Used/Reprinted with permission (obtained 6/28/14). For permission to use or reproduce the model, please contact the University of Iowa Hospitals and Clinics at *(319)384-9098.*

A review of the literature demonstrated the use of the Iowa model in a wide range of disciplines including neonatal (Haxton, Doering, Bringas, & Kelly, 2012), oncology (Brown, 2014), critical care (Kowal, 2010), and nursing administration (Johnson, Gardner, Kelly, Maas, & McCloskey, 1991). In addition, the model has been used in nursing literature from Europe (C. Doody & O. Doody, 2011) and Asia (Chan, Lee, Poh, Ng, & Prabhakaran, 2011; Taylor-Piliae, 1999). Zaccagnini and White (2011) stated "selecting and defining the problem is the earliest and most critical step in an evidenced-based intervention" (p. 104). The Iowa model identifies these problems as a problem-focused trigger or a knowledge-focused trigger (Titler et al., 2001). The problem for this project was a knowledge-focused trigger related to nurses' lack of knowledge regarding the accuracy and reliability of different temperature devices. In addition, there was also a problem-focused trigger related to the organizational priority for early sepsis identification.

The second step of the Iowa model requires one to consider if the problem is an important issue to the organization (Titler et al., 2001). A topic or problem that is consistent with organizational priorities and targets a high-risk or a high-cost issue, has greater potential to be supported by key leaders (Titler et al., 2001). According to N. Tauzon (personal communication, June 13, 2014), the organization under study targeted the early identification and treatment of sepsis as a key process improvement issue. Since a change in patient temperature is one of the early indicators of infection, the accuracy of temperatures obtained within the organization was an important question to consider. Once organizational support has been determined, the next major step to undertake is an

assessment and critical review of the relevant literature to determine if there is sufficient evidence to address a practice change (Titler et al., 2001). An IR of the literature was conducted and that analysis will be discussed in Section 4 of this study.

Quality evaluation. Finally, the IR included an evaluation of the quality of the research. The quality indicators and quality grade described by Hooper and Andrews (2006) was used for this IR (see Table 1). According to Hooper and Andrews, an A or B is considered to be an acceptable grade.

Table 1

Quality Indicators

Indicator	Quality Score and Grade
Number of temperature measurements	A: > 8 B: 5–7 C: 3–4 D: 0–2
One data collector or interrater reliability of multiple data collectors addressed	
Data collector training	
Temperature measurement technique	
Water-bath calibration of instruments	
Core setting used for tympanic thermometers	
Accuracy standard established	
Results reported using instrument bias statistics	
Temperature linearity addressed	

Note. From "Accuracy of Noninvasive Core Temperature Measurement in Acutely Ill Adults: The State of Science" by V. D. Hooper and J. O. Andrews, 2006, *Biological Research for Nursing, 8*(24), p. 28. Copyright 2006 by Sage Publications. Used with permission (obtained 3/8/2016).

Relevance to Nursing Practice

Existing Scholarship

The body of existing scholarship on the question of accuracy and reliability of clinical thermometers has grown along with the advances in technology. Moreover, change was also driven by rising concern related to the mercury used in thermometers. Mercury thermometers had the advantage of long-term stability, little maintenance or training needed, and device failure was readily apparent (National Institute of Standards and Technology, 2011). However, environmental and health concerns related to the mercury resulted in a ban throughout most of the United States and in some European countries (Environment Protection Agency, 2015). An alternative glass thermometer was developed containing gallium (gallium-in-glass thermometer) and is also widely reported on in the literature (Lefrant et al., 2003; Rubia-Rubia, Arias, Sierra, & Guirre-Jaime, 2011; Smith, 2003). While not considered safe to use in the United States, mercury thermometers are still in use in many countries around the world and are included in the current body of research (B. Jensen, F. Jensen, Madsen, & Lossl, 2000; Leon, Rodríguez, Fernández, & Flores, 2005; Prentice & Moreland, 1999; Rajee & Sultana, 2006).

Evaluation of clinical thermometers for accuracy and reliability in different patient populations is important as results from one population, site, or route may not be generalizable to other populations (Grove, Burns, & Gray, 2013). The available research covers a wide range of clinical specialty areas and explores conditions specific to those areas. In the critical care area, a variety of factors can influence the assessed temperature to include vasopressors and physiologic condition (Giuliano, K., Scott, Elliot, &

Giuliano, A., 1999; Lawson et al., 2007; Moran et al., 2007). In the pediatric population, clinical thermometers that have been used for adult populations, were found to be clinically inaccurate in this population (Nimah, Bshesh, Callahan, & Jacobs, 2006; Siberry, Diener-West, Schappell, & Karron, 2002). In the perioperative arena, Barringer et al. (2011) found that some thermometers were not appropriately sensitive for assessing hypothermia. Finally, oncology patients have unique needs and have a lower threshold for fever, so the sensitivity of thermometers used for this population is also important to evaluate (Dzarr, Kamal, & Baba, 2009).

Nursing Practice and Impact of Temperature Assessment

According to Zaccagnini and White (2011), one's practice should be based on research evidence and not on historical processes. There are many variables to consider when assessing temperature to include the device, the technique, and the patient (Davie & Amoore, 2010). Although there is a large body of research evidence on the accuracy and reliability of different temperature devices, a persistent evidence-to-practice gap remains.

Temperature assessment is a standard of care in all areas of nursing practice including peri-anesthesia (pre-operative, intra-operative, and post-operative care), ED, medical and surgical wards, pediatrics, and the critical care environment (Barringer et al., 2011; Hooker & Houston, 1996; Lawson et al., 2007; Nimah et al., 2006). In the peri-anesthesia environment, temperature assessment and ensuring normothermia throughout the perioperative timeframe has been linked to decreased surgical wound infections (Kurz, Sessler, & Lenhardt, 1996). Failure to mitigate hypothermia in the perioperative

arena has been associated with an increase in adverse outcomes which also results in increased healthcare costs (Brown-Mahoney & Odom, 1999).

In the ED, all patients are routinely screened for abnormalities in temperature (Hooker & Houston, 1996; Singler et al., 2013). Data from the National Hospital Ambulatory Medical Care Survey (Centers for Disease Control and Prevention, 2010) showed that of the 129,843 ED patient records reviewed, 123,888 patients had their temperatures screened upon presentation. Abnormal body temperature is one of several factors including respiratory failure, vasopressor use, and bandemia which were identified as early predictors of bacteremia in patients presenting to the ED (Chase et al., 2012).

In the pediatric population, fever is one of the most common reasons parents seek care for their children (Siberry et al., 2002). According to Browne, Currow, and Rainbow (2001), between 20% and 30% of ED visits for children are related to episodes of fever temperature. According to the American Academy of Pediatrics (2007), many temperatures do not require treatment; however, immediate temperature assessment and treatment can be critical for a small subset of pediatric patients. In the critically ill pediatric patient, sepsis criteria includes a core temperature of > 38.3ºC or < 36ºC as well as other physiologic indicators (Goldstein, Giroir, & Randolph, 2005).

In the ICU, early identification of infection has been identified as critical to the successful treatment of sepsis (Dellinger et al., 2012). Sepsis is identified using several physiologic components including evidence of a potential new infection and temperature > 38.3ºC or < 36ºC (Birriel, 2013; Dellinger et al., 2012). Early identification of infection

and administration of antibiotics within the first hour (grade 1B) has been demonstrated to improved mortality (Dellinger et al., 2012).

Given the importance of temperature as part of regular vital sign assessments, thermometers used on adult hospitalized patients should provide accurate and reliable measurements (Flynn-Makic et al., 2011). In many institutions, changes in temperature may result in a cascade of diagnostic studies in order to identify potential infection; these can be costly to both the patient and the organization (Flynn-Makic et al., 2011). Additionally, temperature inaccuracies can lead to missed opportunities for the identification of infection or sepsis, which also increase morbidity and mortality and health care costs (Dellinger et al., 2012).

Previous Gap-in-Practice Strategies

There were two quality improvement (QI) projects related to concerns about the accuracy of clinical thermometers in the literature. Bahr, Senica, Gingras, and Ryan (2010) and Ostrowsky, Ober, Wenzel, and Edmond (2003) identified a clinical practice issue with a new temporal artery thermometer (TAT). In both reports, concerns related to the devices (accuracy and reliability) were raised which prompted the QI projects. User technique, cleaning, and maintenance were also identified in both reports as leading possibilities for inaccurate measurements. Although hospital-wide retraining was accomplished in both facilities, neither group of authors reported a favorable outcome with the TAT and these devices were removed from their respective hospitals.

Operator technique is often cited as a cause for variation in assessed temperatures with new devices. As described by Bahr et al. (2010) and Ostrowsky et al. (2003)

reteaching and retraining may be the first steps when concerns about new devices are raised. Ideally, a thorough review of the available research should be evaluated prior to the bulk purchase of any new device. Ostrowsky et al. reported that the only information on accuracy and reliable of the TAT was from the manufacturer; no supporting data were found in their literature review. Even with low-tech devices such as mercury-in-glass and gallium-in-glass thermometers, technique variability using the axillary and rectal route as accurate placement and dwell time are important (Sund-Levander & Grodzinsky, 2013).

IR Impacts Gap-in-Practice

Some of the challenges with the use of EBP identified by nurses at the bedside are that the research is inaccessible, they are unable to understand the findings, and they do not have time to search for current research evidence (Hicks & Hennessy, 1997; Krom, Batten, & Bautista, 2010). The benefit of an IR review on the clinical thermometers for the adult hospitalized patient is a synthesis of the evidence in one report. In addition, the IR review can be used as a resource for clinical nurses and nurse leaders as an early source of information when considering the bulk purchase of any new device.

Local Background and Context

The former practicum site was part of a five-hospital, for-profit system with 1,673 licensed beds. The relevance for the clinical question regarding accuracy of temperature assessment was first identified through observation of the clinical nursing staff. Temperature measurements were observed being accomplished with many clinical areas with different devices and via different routes. Further, the nurses were unable to clarify the rationale for their choice of device or route. In addition to the observed knowledge-to-

practice gap, the organizational priority for early sepsis identification also supports the need to answer the clinical question.

Federal Context

The hospital system studied is eligible for Medicare and Medicaid reimbursement through the Centers for Medicare and Medicaid (CMS). The CMS is the federal agency responsible for the management of Medicare and works in cooperation with state governments to administer Medicaid (CMS, n.d.). The Hospital Value-Based Purchasing (VPB) Program, a CMS initiative, is designed to reward acute care hospitals for the quality of the care they provide (CMS, 2012). The hospital is located in the downtown area of a large urban city and provides services to a large number of low-income patients. Ensuring full CMS reimbursement for all eligible patients is consistent with the organization's fiscal and quality goals. There are a wide variety of measures incorporated into the CMS reimbursement base including infection. Surgical site infection, catheter-associated urinary tract infection, and vascular catheter-associated infections have been identified by the CMS as a preventable healthcare-acquired condition (HAC; CMS, 2012). According to Mattie and Webster (2008), HAC resulted in 2.4 million additional hospital days and cost between $17 billion and $29 billion. Under the VBP program, acute care hospitals can lose up to 1% of the diagnostic-related group payments (this number will rise to 2% in 2017; CMS, 2012).

Role of the DNP Student

According to the American Association of Colleges of Nursing (AACN; 2006), the emphasis of the practice-focused doctoral degree is the focus on the translation of

research into practice rather than the generation of new evidence. As a critical care nurse and as an educator, the practice-focused degree fit well with my current practice and will help prepare me for additional educational or clinical roles. As a critical care course instructor, the importance of EBP is at the forefront of our didactic content and in daily discussions at the clinical bedside.

My doctoral project, an IR, was developed in response to clinical questions that surfaced from observations in the clinical environment. The IR is one method of facilitating the translation of the research on clinical thermometers into practice at the former practicum site. The five-hospital system associated with my practicum site identified early sepsis identification as an organizational priority. Given the importance of temperature assessment with this organizational priority, the topic of thermometer accuracy merged well. The chief nursing officer at my practicum site, who was also my preceptor, supported an in-depth literature review in order to evaluate the evidence on which devices are best for the populations they serve.

The practicum experiences provided an opportunity to observe a broad variety of clinical areas in a number of hospitals within the system. The questions surrounding temperature accuracy were discussed with nurse leaders but more importantly with the clinical nurses. It was the responses from nurses at the bedside, and the certified nursing assistants, that led me towards the development of the clinical question.

The inspiration for an evaluation of temperature devices came from early observations in the critical care environment, where a number of devices and routes were used interchangeably. The rationales for use or route might include the ability to obtain

rapid results or to improve patient comfort, which is consistent with what is described in the literature (Davie & Amoore, 2010). The question of clinical accuracy of thermometers was easily translatable to my practicum site and to the organizational priority for early sepsis identification.

An important area of potential bias in this project included the preconceptions I had on the accuracy and reliability of specific thermometers. Additionally, I had preconceptions about the best routes for temperature assessment. These preconceptions had the potential to lead me to discount valid and reliable data in favor of evidence which supported my own preconceptions.

Mitigating bias in this project was important as the goal was to conduct an accurate evaluation of the body of research, not an evaluation determined to support my own personal opinions. One means of mitigating bias was to critically evaluate the data from the literature, using a wide variety of resources to help me understand the statistical analyses. Understanding how the data were reported was critical in my ability to determine if the findings supported the authors conclusions or not. However, I think the most important means of addressing potential bias in this project was to be aware of my preconceptions.

Summary

Whitmore and Knafl (2005) stated that the IR can be used for a variety of purposes including context definition, a review of theories, or to answer a specific practice problem. The IR provides a resource for clinical nurses on the current research evidence on clinical thermometers for the adult hospitalized patient. The availability of

this resource will help to narrow the identified knowledge-to-practice gap. Further, the IR serves as the comprehensive literature review identified in the Iowa model (Titler et al., 2001), which may facilitate decision-making for any potential changes in thermometers for the hospital.

In the next section of this paper, I describe the collection and analysis of the evidence. A thorough discussion of the sources of evidence will be provided along with the specific databases, search terms, and inclusion and exclusion criteria used. Finally, I present an analysis of the early findings.

Section 3: Collection and Analysis of Evidence

Introduction

The accuracy of temperature assessment in the adult hospitalized patient is important for the early identification and treatment of sepsis. The purpose of this project was to conduct an IR of the body of research related to accuracy and reliability of clinical thermometers. The IR merges well with observed knowledge-to-practice gaps in the clinical environment as well as supporting the organizational priority of the hospital study site for early sepsis identification. The synthesis from this review can be used by clinical nursing staff, nurse educators, and organizational leaders when considering alternative devices for their hospitals.

In the following section, I will describe the practice-focused question, sources of evidence, and search methodology. Also, a description of the methodology for tracking, organizing, and synthesizing the research will also be provided. Finally, I will present an early analysis of the findings from the literature search.

Practice-Focused Question

The accuracy of temperature assessments is essential for the early identification and treatment of sepsis in the adult, hospitalized patient (Dellinger et al., 2012). According to Oermann and Hays (2011), the purpose of an IR is to advance one's understanding of a specific topic or clinical question. Organizational leaders can rely on the synthesis of the literature in the IR to facilitate decision-making about their current devices and needs. Additionally, the information from the IR guides clinical nurses and helps to narrow the observed knowledge-to-practice gap at the former practicum site.

The project question was: For adult patients in acute care hospitals, which clinical thermometer or thermometers provide the most accurate and reliable temperature readings? The research question should be developed using four components, the population, issue of interest, the comparison being made, and the desired outcome (Terry, 2012). The project question in this format was:

- Population – adult patients in the acute care hospital;
- Issue of interest – accuracy/reliability of clinical thermometers for the most accurate temperature assessments;
- Comparison – device comparisons were reviewed from the available research;
- Outcome – use the best evidence in selecting a clinical thermometer or thermometers; improved accuracy of assessed temperature may improve early recognition and treatment of sepsis.

Sources of Evidence

The databases that I queried for literature from January 1999 to December 2015 included CINHAL & MEDLINE Simultaneous Search, ProQuest Nursing & Allied Health Source, the Cochrane Database of Systematic Reviews and PubMed. A search of these databases was undertaken using the following keywords: *temperature assessment, temperature assessment AND methods, body temperature determination, body temperature determination AND methods, thermometry,* and *thermometry AND methods AND comparison*. The inclusion and exclusion criteria for this review is described in Table 2.

Table 2

Inclusion and Exclusion Criteria

Inclusion	Exclusion
Peer-reviewed journals	Ambulatory settings
English language	Outpatient settings
Human studies	Prototype experimental studies
Critical care or intensive care unit	Intraoperative TM thermometer
Perioperative	Exercise related studies
Emergency Department	Healthy volunteers
Inpatient	Pediatrics (< 19 years old)
Adult (19+)	

The inclusion of multiple clinical areas in this review was specific to the population identified in the clinical practice question of "hospitalized adult patient." While the ED may be considered an outpatient treatment area, it is also a significant source (high volume, high risk) for patients being screened for infection or sepsis (Singler et al., 2012; Varney et al., 2002). Lastly, the review did not include research on children, as their physiologic differences limit the generalizability of findings to adult populations. The level of evidence described by Fineout-Overholt et al. (2010) was used to categorize the body of evidence for this project. The quality of the evidence was evaluated using the quality indicators and quality score (see Table 1) described by Hooper and Andrews (2006). The evidence will be described by the level of evidence and grouped by device.

Protection of Human Rights

The protection of human rights in research is paramount and is governed by the U. S. Department of Health and Human Services (2009). This capstone project was an IR of the literature related to clinical thermometers used for adult hospitalized patients; no human subjects were used. Nevertheless, the project purpose and methods were reviewed

and approved by the Walden University Institutional Review Board (IRB approval number 03-11-16-0450734).

Analysis and Synthesis

Search results for database and keyword searches were documented on a separate spreadsheet to ensure continuity in search procedures. Documentation included the number of hits (for each search), number of relevant articles, number of repeated relevant articles, and number of articles selected for early review. Additional search methods included hand searches through reference lists to ensure all relevant research was included in this review.

The search results yielded 2,643 papers and the abstracts were reviewed for their relevance to the clinical question. The initial abstract review resulted in the selection of 85 papers. Further review resulted in the exclusion of 38 papers described in Table 3.

Table 3

Table of Article Exclusion

Author, Year	Article Title	Rationale for Exclusion
Abolnik et al., 1999	Comparison of oral and tympanic temperatures in a Veterans Administration outpatient clinic	Outpatient sample
Ahmadnia et al., 2010	A comparison between urinary bladder temperature and rectal, axillary, and oral temperatures following kidney transplantation	Letter to the editor, no copy of research available
Arslan et al., 2011	Analysis of the effect of lying on the ear on body temperature measurement using a tympanic thermometer	Outpatient sample
Bock et al., 2005	The accuracy of a new infrared ear thermometer in patients undergoing cardiac surgery	Prototype temperature device
Camboni et al., 2008	Accuracy of core temperature measurement in deep hypothermic circulatory arrest	Brain temperature as reference, not generalizable for review purposes
Cronin and Wallis, 2000	Temperature taking in the ICU: Which route is best?	Quality improvement project
Dowding et al., 2002	An investigation into the accuracy of different types of thermometers	Sample included healthy volunteers

(table continues)

Author, Year	Article Title	Rationale for Exclusion
Dzarr et al., 2009	A comparison between infrared tympanic thermometry, oral, and axilla with rectal thermometry in neutropenic patients	Sample included children
Fallis 2005	The effect of urine flow rate on urinary bladder temperatures in critically ill adults	No comparison related to the purpose of the review
Gasim et al., 2013	Accuracy of tympanic measurement using an infrared tympanic membrane thermometer	Sample included children
Giantin et al., 2008	Reliability of body temperature measurements in hospitalized older patients	Comparisons were related to functional assessments
Gobolos et al, 2014	Reliability of different body temperature measurement sites during aortic surgery	Intraoperative tympanic thermometer is not comparable to device used for intermitted temperature assessments
Hamilton et al., 2013	Clinical performance of infrared consumer-grade thermometers	Sample included children
Harioka et al., 2000	"Deep-forehead" temperature correlates well with blood temperature	Device not available for general population
Hausfater et al., 2008	Cutaneous infrared thermometry for detecting febrile patients	Sample included children
Huang & Kurz, 2001	Body warmer and upper extremities positions affect accuracy of cutaneous thermometers during anesthesia	Additional variables of skin temperature and body position not related to purpose of review
Hocker et al., 2012	Correlation, accuracy, precision, and practicality of perioperative measurement of sublingual temperature in comparison with tympanic membrane temperature in awake and anaesthetised patients	Intraoperative tympanic thermometer is not comparable to device used for intermitted temperature assessments
Hutton et al., 2008	Accuracy of different temperature devices in the postpartum population	Sample included newborns
Khorshid et al., 2005	Comparing mercury-in-glass, tympanic, and disposable thermometers in measuring body temperature in healthy young people	Outpatient sample
Kimberger et al., 2009	Accuracy and precision of a novel noninvasive core thermometer	Investigational device
Kistemaker et al., 2006	Reliability of an infrared forehead skin thermometer for core temperature measurements	Sample included outpatients and exercise
Lu et al., 2009	The effects of measurement site and ambient temperature on body temperature values in healthy older adults: A cross-sectional comparative study	Outpatient sample
Masamune et al., 2011	The usefulness of an earphone-type infrared tympanic thermometer during cardiac surgery with cardiopulmonary bypass	Intraoperative tympanic thermometer is not comparable to device used for intermitted temperature assessments
Modell et al., 1999	Hope for the infrared tympanic thermometer: One model outperforms the others	Letter to the editor, no copy of research found

(table continues)

Author, Year	Article Title	Rationale for Exclusion
Nguyen et al., 2010	Comparison of three infrared thermal detection systems and self-report for mass fever screening	Devices were not comparable to other infrared devices
Nuckton et al., 2001	A comparison of two methods of measuring rectal temperatures with digital thermometers	Outpatient sample
O'Brien et al., 2000	The accuracy of oral predictive and infrared emission detection tympanic thermometers in an Emergency Department setting	Sample included children
Onur et al., 2008	Oral, axillary, and tympanic temperature measurements in older and younger adults with or without fever	Sample included children
Rabbani et al., 2010	Tympanic temperature comparison with oral mercury thermometer readings in an OPD setting	Outpatient sample
Schey et al., 2009	Skin temperature as a noninvasive marker of haemodynamic and perfusion status in adult cardiac surgical patients: An observational study	Skin temperature used as comparison
Schmal et al., 2006	Effect of status after ear surgery and ear pathology on the results of infrared ear thermometry	Sample included healthy volunteers
Sehgal et al., 2002	Comparison of tympanic and rectal temperature in febrile patients	Sample includes children
Sener et al., 2012	Agreement between axillary, tympanic, and mid-forehead body temperature measurements in adult emergency department patients	Sample included children 16+
Singh et al., 2000	Variation of axillary temperature and its correlation with oral temperature	Sample included children
Smith, L.S. 2003	Using low-tech thermometers to measure body temperatures in older adults: A pilot study	Pilot study
Smith, L. S. 2004	Temperature measurement in critical care adults: A comparison of thermometry and measurement routes	Experimental device
Sund-Levander et al., 2002	Normal oral, rectal, tympanic, and axillary body temperatures in adult men and women: A systematic literature review	Topic was normal body temperature
Washington & Matney, 2008	Comparison of temperature measurement devices in post anesthesia patients	Sample included children

Summary

The assessment of temperatures is considered a routine activity in all areas of nursing. The accuracy of assessed temperatures is important as abnormalities in temperature may be an early indication of infection or sepsis (Dellinger et al., 2012). There is a significant body of research available comparing different clinical thermometers: however, many nurses are not aware of this evidence. An IR of this body of evidence provides nurses with a resource to narrow this knowledge-to-practice gap. Advancing nursing knowledge in this area can also lead to early identification and treatment of sepsis, improved quality of care, and reduced health care costs. The IR will also provide organizational leaders with the comprehensive literature review necessary to make decisions about the clinical thermometers used in their hospitals.

The next section of this paper reports on the findings of this IR and includes implications for clinical practice. Also, I will include recommendations for the organizational leaders regarding device accuracy and the potential to impact early recognition of sepsis within their organization.

Section 4: Findings and Implications

Introduction

The accuracy of temperature assessment in the adult hospitalized patient is important; changes in temperature (< 37ºC or > 37ºC) can be an early indicator of infection (Dellinger et al., 2012). At my practicum site, an organizational priority for early sepsis identification, together with an observed knowledge-to-practice gap related to clinical thermometers, provided the foundation for this IR. The practice-focused question that guided this inquiry was: For adult patients in acute care hospitals, which clinical thermometer or thermometers provide the most accurate and reliable temperature readings? An IR review of the literature provides nursing leaders with a resource to narrow the knowledge-to-practice gap identified in the clinical environment. The IR also provides organizational leaders with an evidence-based review of the literature to facilitate decision-making about thermometers used in their hospitals.

A search of the literature was undertaken using a variety of search terms (previously described in Section 3) in four databases: CINHAL & MEDLINE Simultaneous Search, ProQuest Nursing & Allied Health Source, the Cochrane Database of Systematic Reviews, and PubMed. The review methodology described by Whittemore and Knafl (2005) guided this review. The Hierarchy of Evidence (see Figure 1) described by Fineout-Overholt et al. (2010) was used to categorize the body of evidence. Lastly, the quality of the evidence was evaluated using the quality indicators and quality grade described by Hooper and Andrews (2006; see Table 1). The review of the literature will first be described by the level of evidence (highest to lowest strength), followed by

device, and reference site. The articles were grouped into those considered clinically unacceptable, clinically acceptable, or inconclusive. Additionally, 11 studies included an evaluation of the diagnostic accuracy of some thermometers; therefore, those findings are also delineated.

Findings and Implications

There were 47 articles which met the inclusion criteria for this review. A summary of these articles can be found in Table 4. Most of the populations were from the intensive care unit (ICU; 21), followed by medical/surgical wards (9), perioperative patients (6), the ED (4), and oncology (3). Several articles had combined populations from more than one area such as the ICU and medical/surgical wards (2), ICU and the ED (1), and the ICU and perioperative patients (1). Many of the studies provided comparison data on several devices, and the results are presented in the device specific section.

Table 4

Summary of Articles Included in the Integrative Review

Author/ Year	Study Design	Population Sample	Purpose / A priori acceptability	Findings/Conclusion
			Level I Evidence	
Jefferies et al., 2011	Systematic review	ICU *N* = 3	- Determine accuracy of peripheral thermometers in detecting fever (> 37.5ºC) TM, O, R compared to PAC (reference) - Acceptability: mean difference ±0.2ºC of PAC	- 5 of 7 TM's and the O were within pre-defined criteria while R was outside this limit in all three studies. -TM and O provide accurate measure of core temp within the febrile range. R is clinically inaccurate.
			Level IV Evidence	
Amoateng-Adjepong et al., 1999	Prospective observational cohort study	ICU *N* = 51 918 paired readings; 153 observations	- Determine accuracy of TM compared to PAC (reference) - Evaluate intra-observer variability between ICU RN educator (1), ICU RNs, floor RNs, medical assistants (MA) - Acceptability: within 0.5ºF of PAC	- TM to PAC correlation range 0.83 to 0.89 - Accuracy and correlation coefficient differed depending on operator skill ICU RN educator – 98% accurate; *r* = 0.98; ICU RNs – 80% accurate; *r* = 0.90 Floor RNs/MA – 61% accurate *r* = 0.82 - TM is accurate, but accuracy is dependent on operator skill

(table continues)

Author/ Year	Study Design	Population Sample	Purpose / A priori acceptability	Findings/Conclusion
Barringer et al., 2011	Repeated measures comparison design	Peri-Op *N* = 86 258 paired readings	- Evaluate equivalence between TAT, AX to O (reference) - Acceptability: not defined	- Preoperative – TAT bias were -0.27ºF (95% LOA -1.46, 0.91); AX bias 0.5ºF (95% LOA -0.9, 1.8) - Postoperative – TAT -0.12 (95% LOA -1.49, 1.24) AX bias -0.2 (95% LOA -2.1, 1.7) - TAT is acceptable replacement for oral
Bodkin et al., 2014	Prospective observational cohort design	ED *N* = 100 (febrile = 47; afebrile = 53) 200 readings	- Compare TAT measurements to O (reference) - Evaluate accuracy of TAT to detect fever (38ºC) - Acceptability: difference of ± 0.5ºC	- Bias 0.48ºC (SD ± 0.8) *P* < .0001 - 49% had clinically significant different temperatures between TAT and O - 57% (*n* = 27) of fever detected by O, were not measured by the TAT - TAT should be used with caution; screen for other clinical indicators of infection
Calonder et al., 2010	Repeated measures comparison design	Peri-Op *N* = 23 46 measures per site	- Evaluate the difference between core measured by O and TAT compared to ES (reference) - Acceptability: temperature difference of > 0.4ºC	- O biased high (to ES) by 0.12ºC (*P* = .0008; 95% CI 0.061, 0.187); TAT biased high 0.075ºC (*P* = .03; 95% CI 0.010, 0.133). - Statistically significant differences between O, TAT and ES, but within clinically acceptable criteria
Counts et al., 2014	Method comparison, cohort design	ICU *N* = 48 144 paired readings	- Determine differences in temperature obtained with CH and TAT compared to O (reference) - Acceptability: bias ≤ ± 0.3ºC; precision ≤ ± 0.5ºC	- CH within acceptable bias, precision slightly outside acceptable value (0.56ºC) - TAT – bias and precision exceeded recommendations; not recommended for routine use
Duncan et al., 2008	Prospective comparison design	ED/ICU *N* = 93 Paired readings ED – 148 ICU – 38	- Assess reliability/validity of NT compared to O (reference) and BL (reference) - Acceptability: ± 0.3ºC	- NT reliability – strong correlation between NT readings (*r* = 0.94) bias between readings 0.00ºC (*SD* 0.15) - NT and O – poor correlation/poor agreement (*r* = 0.26); bias 0.87ºC (*SD* 0.58) - NT and BL – highly correlated/poor agreement (*r* = 0.83); bias 1.17ºC (*SD* 0.67) - NT is reliable, but does not agree with O or BL; NT not recommended for use
Dunleavy, 2008	Comparative descriptive design	ICU *N* = 10 241 paired readings	- Determine which device is most accurate TM to O (reference); TM to BL (reference); BL to PAC (reference) - Acceptability: not defined	- TM to O: variance of ≥ 0.8ºC in 58% - TM to BL: variance of ≥ 0.8ºC in 35%; of the 35%, 38% had variance of ≥ 1.5ºC. - Based on variance of TM to O and BL; TM not recommended for ICU patients. BL is an acceptable alternative for PAC
Fallis, et al., 2006	Prospective observational comparison design	Peri-Op Obstetrics *N* = 62 212 paired readings	- Assess agreement between CH and O (reference) - Acceptability: difference of 0.3ºC	- Bias 0.35ºC ± 0.32°C (*p* <.0001, 95% CI 0.31, 0.40) - LCCC poor (0.443) - CH underestimated O in 81%; overestimated O in 10% - CH significantly under-measures O and is not a reliable indicator for temperature evaluation
Farnell et al., 2005	Prospective observational comparison design	ICU *N* = 25 160 paired readings	- Compare accuracy and reliability of CH and TM to PAC (reference) - Clinical significance (determined by medical staff) – would inaccuracy cause a delayed intervention or result in an unnecessary intervention	- CH and TM to PAC bias 0.2ºC (*SD* 0.34; *P* < 0.0001) and 0.0ºC (*SD* 0.59; *P* = 0.39), respectively - Clinical significance: 15.3% (*n* = 26) CH and 21.1% (n = 35) TM might have had delayed interventions; while 28.8% (*n* = 44) CH and 37.8% (*n* = 58) TM might have received unnecessary interventions - CH was more accurate/reliable than TM. TM not recommended for use.

(table continues)

Author/ Year	Study Design	Population Sample	Purpose / A priori acceptability	Findings/Conclusion
Fetzer et al., 2008	Prospective descriptive correlational design	Peri-Op *N* = 222 444 paired readings	- Evaluate agreement between TAT and TM (reference) - Acceptability: within 1.0ºC of TM; 95% CI within 1.0ºC	- TAT to TM bias - 0.04ºC (*SD* 0.64) 95% CI -1.29, 1.21 - Pearson's *r* = 0.42; *P* = .000 - TAT and TM cannot be considered equivalent
Fountain et al., 2008	Method comparison cohort design	Oncology *N* = 60 240 readings	- Evaluate agreement between CH, TM, and TAT to O (reference) - Acceptability: not defined	- CH, TM, TAT to O bias 0.00ºF (*SD* 0.92); 0.39 (SD 1.01); 0.68 (*SD* 0.99), respectively - Significant difference between O, TM, and TAT ($F_{3,\ 171} = 12.51$, $p < 0.0001$) - Significant difference between TM and TAT ($p = 0.003$ and $p < 0.0001$, respectively) - TM / TAT not recommended for use. - CH – good agreement with O; authors recommend limited use
Frommelt et al., 2008	Prospective, method-comparison design	Surgical ward *N* = 84 333 readings	- Compare TM, TAT, and CH to O (reference) - Acceptability: not defined	- TM to O bias -1.21ºF (*SD* 0.79); $t = 14.09$, $p < 0.0001$; - TAT to O bias 0.37ºF (*SD* 0.67); $t = -5.11$, $p < 0.0001$ - CH to O bias -0.28ºF (*SD* 0.69); $t = 3.78$, $p = 0.0003$ - TM and TAT had greatest variability, not recommended for use; CH had less variability, use with caution.
Gilbert et al., 2002	Repeated measures design	Surgical ward *N* = 257 514 paired readings	Examine reproducibility of TM and O; Acceptability: difference of 0.2ºC	-Bias between TM1 and TM2 0.28°C, 46% were ≥ ± 0.2°C Bias between O1 and O2 was 0.19°C, 63% were ≥ ± 0.2°C - TM to O bias 0.36ºC, 34% were clinically significant - Strong negative correlation between TM and O ($r = -0.96$, $p < .001$) - No correlation between devices; important to use the same thermometer for serial temperature measurement
Giuliano et al., 1999	Prospective descriptive comparative design	ICU *N* = 102 393 readings	- Determine reliability and accuracy of O and TM when compared to PAC (reference) - Acceptability: bias ±0.3ºC and a SD 0.3ºC	- O to PAC bias -0.15ºC (*SD* 0.36); $p = .0001$ - TM (core mode) to PAC bias -0.11ºC (*SD* 0.57); $p = .0795$ TM (oral mode) bias -0.52ºC (*SD* 0.53); $p = .0001$. - TM demonstrated greatest variability, not recommended. O is acceptable alternative for PAC
Giuliano et al., 2000	Prospective descriptive cohort study	ICU *N* = 72 812 readings	- Determine accuracy and variability of O and TM, in febrile (>38ºC) patients, compared to PAC (reference) - Acceptability: accuracy tolerance zone of ±0.5ºC	- O to PAC bias 0.18ºC, *SD* in afebrile = 0.50ºC; febrile = 0.47ºC - 47 data points outside tolerance; bias in febrile patients TM to PAC bias -0.17ºC, *SD* in afebrile = 0.64ºC; febrile 0.65ºC - 75 data points outside tolerance - In febrile patients, wide variability with both TM devices, even with expert operators. O temperatures had the best agreement with PAC
Hasper et al., 2011	Prospective correlational cohort study	ICU *N* = 10 558 readings	- Compare BL and TM to ES (reference) during therapeutic hypothermia (32-34ºC) - Acceptability: not defined	- BL, TM to ES bias, LOA (± 2SD) 0.019ºC, ± 0.61 and 0.021ºC, and ± 0.80ºC, respectively - Strong positive correlation TM to ES and TM to BL $r = 0.95$, $p < 0.0001$, 95% CI 0.93, 0.96; $r = 0.96$, $p < 0.0001$, 95% CI 0.95, 0.97, respectively - TM temperature is an accurate representation of ES and BL in hypothermic range (32-34ºC)

(*table continues*)

Author/ Year	Study Design	Population Sample	Purpose / A priori acceptability	Findings/Conclusion
Haugan et al., 2012	Prospective correlational agreement study	Surgical ward and ICU *N* = 200 406 readings per method ICU – 252 readings	- Explore precision between two new TM (right to left ear) - Ward - compare TM to R (reference) - ICU - compare TM to PAC (reference) - Acceptability: difference of 0.25ºC	- No statistically significant differences found for left vs right ear for either brand. - Agreement between TM devices; bias Braun 0.04ºC; Genius -0.01ºC - Both brands measured consistently lower temps than R (Braun 0.36ºC, $p < 0.001$) Genius 0.85ºC, $p < 0.001$) - Authors concluded TM devices are acceptable
Irvin, 1999	Comparison study	Medical Surgical Ward *N* = 160	- Evaluate reliability, validity and variability of TM compared to O (reference) - Acceptability: not defined	- Reliability – no sign differences between nurses. -Validity – significant difference between TM and O $F(1;156) = 41.8$, $p < 0.001$). -Wide variability – 58% of O readings were 1ºF higher than TM - TM may not be as accurate as O
Jensen et al., 2000	Prospective comparison design	Medical Surgical *N* = 200 7 per subject	- Determine accuracy of electronic thermometry. Compare R, O, AX (electronic) and TM to R (mercury) (reference) - Acceptability: ± 0.5ºC	- R, O, AX (electronic), TM to R (mercury) bias (SD) -0.05ºC (0.12); 0.53ºC (0.53); 0.62ºC (0.49); 0.54ºC (0.41), respectively - R (electronic) significantly more accurate than TM, O and AX $p < 0.001$, $p < 0.001$, $p < 0.001$, respectively - In febrile patients (T > 37.5ºC), R more accurate than TM, O and AX $p > 0.001$, $p < 0.001$, $p < 0.001$, respectively - TM is as inaccurate as O and AX, especially in febrile patients. Electronic O, AX, and TM not recommended
Khan et al., 2006	Prospective comparison design	ICU *N* = 49 629 readings	- In post-cardiac surgery patients, does TM (right and left) and AX correlate with BL (reference) - Acceptability: not defined	- Left TM, Right TM, AX to BL bias 0.65ºC (95% CI -0.24, 1.58) 0.57ºC (95% CI -0.48, 1.63) 0.55ºC (95% CI -0.27, 1.36), respectively - AX and TM are unreliable for post-cardiac surgery patients
Kimberger et al., 2007	Prospective comparison design	Neurological operative; Neuro ICU *N* = 70 280 readings	- Determine agreement between TAT and BL (reference) - Evaluate TAT sensitivity and specificity for hypothermia (35.5ºC) and hyperthermia (37.8 ºC) - Acceptability: LOA < ± 0.5ºC	- TAT to BL (normothermic) bias 0.1ºC (*SD* 0.07); > 37.8ºC bias 0.4ºC (*SD* 0.7); <35.5ºC bias -0.7ºC (SD 1.1) - TAT sensitivity and specificity for detecting fever 0.72 and 0.97; for hypothermia 0.29 and 0.95. - TAT not recommended for perioperative temperature monitoring
Langham et al., 2009	Prospective comparison observational design	Peri-Op *N* = 50 200 readings	- Evaluate accuracy and precision of TAT, TM (right and left), O, AX compared to BL (reference) - Acceptability: within 0.5°C of BL and 95% proportion of measurements within 0.5°C of BL	- TAT to BL bias 0.23°C (*SD* 0.50); proportion within 0.5ºC 0.70; LCCC 0.53 (95% CI 0.41, 0.64) -Right TM to BL bias -1.04°C (*SD* 0.51); proportion 0.13; LCCC 0.34 (95% CI 0.22, 0.44) -Left TM to BL bias -1.06°C (*SD* 0.51); proportion 0.13; LCCC 0.34 (95% CI 0.22, 0.44) - O to BL bias -0.25°C (*SD* 0.38); proportion 0.81; LCCC 0.79 (95% CI 0.69, 0.86) - AX to BL bias -0.50°C (*SD* 0.42); proportion 0.61; LCCC 0.64 (95% CI 0.49, 0.75) - None fully met acceptability criteria, however, O, TAT agreed best with BL, 70–80% all pairs differing by no more than 0.5°C. Accuracy "probably acceptable"; TM and AX not acceptable for clinical practice

(table continues)

Author/ Year	Study Design	Population Sample	Purpose / A priori acceptability	Findings/Conclusion
Lawson et al., 2007	Prospective repeated measures design	ICU *N* = 60 180 readings per site	- Determine accuracy (bias) and precision (SD) of O, TM, TAT and AX compared to PAC (reference) - Acceptability: > ±0.5ºC from PAC and identify number of data points outside range	- O to PAC bias 0.09 (*SD* 0.42ºC); 95% CL -0.75, 0.93; 34 of 180 readings (19%) outside 0.5ºC - TM to PAC bias -0.36ºC (*SD* 0.56ºC); 95% CL -1.46, 0.74; 88 of 180 (48%) readings outside 0.5ºC - TAT to PAC bias -0.02ºC (*SD* 0.47ºC); 95% CL -0.92, 0.88; 36 of 180 (20%) readings outside 0.5ºC - AX to PAC bias 0.23ºC (*SD* 0.44ºC); 95% CL -0.64, 1.12; 49 of 180 (27%) readings outside 0.5ºC - O and TAT were most accurate and precise. AX underestimates PAC; TM, least accurate or precise
Lefrant et al., 2003	Prospective comparison cohort study	ICU *N* = 42 529 readings	- Compare ES, BL, R; IN, AX (both measured with Gallium-in-glass) to PAC (reference) - Acceptability: not defined	- Bias between PAC and ES 0.11 (*SD* ± 0.30), R -0.07 (*SD* ± 0.40), AX 0.27 (*SD* ± 0.45), IN 0.17 (S*D* ± 0.48), BL -0.21 (*SD* ± 0.20) - BL and ES can be used as alternatives to PAC; BL and ES are more reliable than R, which was better than IN and AX
Leon et al., 2005	Prospective comparison descriptive design	ICU *N* = 50 429 readings	- Determine the accuracy of TM compared to AX (mercury) (reference) - Determine sensitivity and specificity for different temperatures 37ºC, 38ºC, 39ºC - Acceptability: not defined	- TM to AX bias 0.006ºC, 95% LOA -1.09 and 1.102ºC TM strongly correlated with AX ($r = 0.813$, $P < .0005$) TM sensitivity, specificity, PPV, NPV at 37 ºC was 74%, 85%; 81%, 78% at 38ºC was 70%, 95%; 70%, 95% and at 39ºC was 25%, 99.8%; 50%, 99% - TM device highly reliable for use in ICU
Marable et al., 2009	Prospective comparative design	ICU *N* = 69 215 readings	- Determine if TAT (three techniques – forehead and ear, forehead only, ear only) or AX are acceptable alternative to O (reference) - Evaluate influence of body mass index (BMI ≥ 30 or BMI<30) on TAT and AX results - Acceptability: difference of 0.5ºF from O and number of readings >0.5ºF	- TAT (forehead and ear) bias 0.27ºF (95% CI -2.13, 2.66); 60.6% of readings were > 0.5ºF; sensitivity 90.4% - TAT (forehead only) bias -0.56ºF (95% CI -2.65, 1.54); 60.9% of readings were > 0.5ºF; sensitivity 94.6% - TAT (ear only) to O: bias -0.26ºF (95% CI -2.79, 2.26); 65.6% of readings were > 0.5ºF; sensitivity 94% - AX to O: bias 0.03 (95% CI -1.97, 2.03); 55.4 % of readings were > 0.5ºF; sensitivity 89.5% - TAT lower than O with BMI ≥ 30 compared with BMI ≤ 30 ($P = .0313$ and $P = .0065$, respectively) - TAT not recommended
Mason et al., 2015	Repeated measures equivalence design	Oncology *N* = 33 40 readings	- Determine equivalence of TAT, AX to O (reference) - Acceptability: difference of 0.2ºF from O	- TAT-O difference was 0.14°F, - AX-O difference was 0.25°F, which exceeded the criterion - TAT device is acceptable; AX should not be used or limited use.
McConnell et al., 2013	Method comparison design	Med/Surg ward *N* = 34 68 readings	- Evaluate intra- and inter-rater reliability of TAT to O (reference) - Determine bias / precision of TAT to O (reference) - Acceptability: intra-, interrater reliability SD ≤0.6ºF; between devices: bias ≤ 1.0ºF and precision (SD) ≤ 0.6ºF	- Intra-rater reliability (two investigators) differences 0.14ºF (±0.43ºF) and 0.13ºF (±0.4ºF) - Inter-rater reliability difference -0.19ºF (±0.48ºF) - TAT to O bias (two investigators) 0.48ºF (SD 0.88) and 0.47ºF (SD 0.57ºF) - TAT is reliable method for temperature measurement

(table continues)

Author/ Year	Study Design	Population Sample	Purpose / A priori acceptability	Findings/Conclusion
Moran et al., 2007	Prospective observational cohort study	ICU *N* = 110 6,703 readings	- Compare accuracy of TM, AX, BL to PAC (reference) Acceptability: not defined	- LCCC TM, BL and AX was 0.77, 0.92, 0.83 respectively - TM to PAC bias 0.36ºC (LOA -0.56, 1.28) - AX to PAC bias 0.30ºC (LOA -0.42, 1.01) - BL to PAC bias -0.05ºC (LOA -0.69, 0.59) - Agreement between TM and PAC was inferior to BL, which was overall more likely to reflect PAC
Myny et al., 2005	Prospective descriptive comparison design	ICU *N* = 57 318 readings	- Determine accuracy and variability of TAT, AX, compared to PAC (reference) - Acceptability: ±0.3ºC from the PAC; SD 0.3ºC to 0.5ºC	- TAT to PAC bias 0.14ºC (*SD* 0.51); 95% CI 0.04, 0.23; $p = 0.33$ - AX to PAC bias: 0.46ºC (*SD* 0.39); 95% CI 0.39, 0.54; $p < 0.001$ - TAT is acceptably accurate in normo-thermic patients
Nonose et al., 2012	Prospective observational comparison design	ICU *N* = 73 1,793	- Compare accuracy and precision of TM, AX to BL (reference) and PAC (reference) - Acceptability: not defined	- BL, TM, AX to PAC bias 0.02ºC (*SD* 0.21); -1.03 (*SD* 1.23); -0.60 (*SD* 0.53), respectively - TM, AX to BL bias 0.51ºC (*SD* 1.02) (95% LOA -2,51, 1.48); -0.33 (*SD* 0.55) (95% LOA -1.42, 0.75), respectively - Correlation TM, AX to BL $R^2 = 0.64$; $R^2 = 0.23$, respectively - BL agreed with PAC; AX agreed more with BL than TM. AX is acceptable alternative to BL and PAC
Potter et al., 2003	Prospective descriptive design	ICU *N* = 85 170 readings	- For isolation patients, is the CH an acceptable alternative to O (reference) Acceptability: difference from O 0.3ºC	- Bias 0.001ºC (SD 0.18ºC; t 84 = 0.34, *P* =. 97; 95% CI -0.061, 0.070) - Correlation was high ($r = 0.937$) - 25% of all CH were overestimates (11.8%) or underestimates (10.8%) by 0.4 ºC - CH useful as a screening tool; consider validation with electronic O for abnormal findings
Prentice & Moreland, 1999	Prospective comparison design	Geriatric chronic care *N* = 30 180 readings	- Evaluate test/retest reliability of TM, O (electric) and O (mercury) - Evaluate accuracy of TM, O (electric) to O (mercury) (reference) - Evaluate sensitivity and specificity of TM, O (electric) to detect fever (37.5 ºC) Acceptability: not defined	- O (merc) findings were more consistent between times - - O (electric) sensitivity and specificity for fever were 60% (95% CI 17%, 100%) amd 84% (95% CI 70%, 98%) - TM sensitivity and specificity for fever was 60% (95% CI 17%, 100%) and 92% (95% CI 81%, 100%) - Oral (electric) more accurate and reliable than TM. Poor sensitivity for detecting fever
Rajee & Sultana, 2006	Prospective comparison design	ED *N* = 200 1200 readings	- Evaluate repeatability of TM, CH - Evaluate agreement of CH, TM to R (mercury - reference). - Evaluate sensitivity and specificity of TM and CH to detect fever (≥38ºC) - Acceptability: repeatability ±0.3ºC; agreement within ±0.5ºC	- TM repeatability significant bias for second reading -0.8ºC (95% CI -0.9, 0.7) to 0.5ºC (95% CI 0.5, 0.6) - CH repeatability nonsignificant bias −0.3ºC (95% CI -0.4, -0.3ºC) to 0.4°C (95% CI 0.4, 0.5) - TM, CH to R (mercury) bias 0ºC (95% CI -0.1, 0.1) and -0.1ºC (95% CI -0.1, 0), respectively - CH can be used interchangeable with TM and R (mercury)

(table continues)

Author/ Year	Study Design	Population Sample	Purpose / A priori acceptability	Findings/Conclusion
Rubia-Rubia et al., 2011	Comparative descriptive design	ICU *N* = 201 3015 readings	- Evaluate inter- and intra-rater reliability - Evaluate concordance of TM, TAT, CH (axillary) AX (gallium) to PAC (reference) - Evaluate sensitivity and specificity of TM, TAT, CH (axillary), AX (gallium) for fever (>38.5ºC) - Acceptability: ±0.2°C	- TM - lowest inter-/intra-rater reliability (76% and 85% respectively) (none reported for TAT, CH or AX) - Bias from PAC (range) TM -0.1°C (-0.7; 0.5); $p = 0.003$ TAT 1.0°C (-0.4; 2.4); $p < 0.001$ CH (axillary) 0.2°C (-0.6; 1.0); $p < 0.001$ AX (gallium) 0.4ºC (-.04; 1.2); $p < 0.001$ - Sensitivity and specificity TM - 98.3%, 93% TAT - 81.6%, 88% CH (axillary) – 96.7%, 91% AX (gallium) 97.3%, 94% - AX (gallium) with 12-minute dwell time was the most accurate and reliable
Shin et al., 2013	Prospective observational cohort design	ICU *N* = 21 1479 readings	Evaluate agreement of BL, R to PAC (reference) during three phases of therapeutic hypothermia (TH) Acceptability: not defined	Bias to PAC and correlation Induction phase, BL (−0.24 ± 1.30ºC; $r = 0.827$) RE (−0.52 ± 1.40ºC; $r = 0.834$) Maintenance phase BL (0.06 ± 0.79ºC; $r = 0.812$) RE (−0.30 ± 1.16ºC; $r = 0.600$) Rewarming phase: BL (0.08 ± 0.86ºC; $r = 0.915$) RE (−0.03 ± 1.71ºC; $r = 0.684$) - Bias between BL and PA temperatures is lower than those in other sites during TH. Use of R only may result in overcooling.
Singler et al., 2013	Prospective quality measurement design with retrospective analysis	ED *N* = 427 3 readings per patient	- Evaluate diagnostic accuracy for infection of TM, TAT compared to R (reference) - Compare reliability of TM, TAT compared to R (reference) Acceptability: adjudicated final diagnosis of infection by two independent physicians after review of all clinical data	- In patients with confirmed infection (n = 105), 22.8%, 35.5% and 43.8% had temperature > 38ºC using TM, TAT, and R, respectively. - TM to R bias 0.54ºC (95% LOA -0.14, 1.21) - TAT to R bias 0.03ºC (95% LOA -0.94, 1.01) - Diagnostic accuracy (AUC) comparable R AUC: 0.72 (95% CI 0.65, 0.80) and TM AUC: 0.73 (95% CI 0.66, 0.81). TAT significantly lower AUC: 0.65 (95% CI 0.57, 0.73; $P < 0.001$). - R and TM have sufficient diagnostic accuracy;
Smith, 2003	Descriptive correlational design	Medical Surgical *N* = 120 960 readings	- Compare clinical accuracy of Gallium-in-glass– O, AX, IN, R to Mercury – O, AX, IN, R (reference) Acceptability: not defined	- Correlation mercury to gallium O $r = 0.929$ ($p < .001$); AX $r = 0.886$ ($p < .001$); IN $r = 0.701$ ($p < .001$); R $r = 0.927$ ($p < .001$) - Bias and 95% CI by site (°F): O 0.20 (0.142; 0.265), AX 0.25 (0.167; 0.339), IN 0.18 (0.037; 0.321), and R 0.06 (-0.111; 0.111). - Gallium-in-glass is an appropriate replacement for mercury

(table continues)

Author/ Year	Study Design	Population Sample	Purpose / A priori acceptability	Findings/Conclusion
Smitz et al., 2000	Prospective comparison sequential measures	Geriatric unit *N* = 45 34 sets of readings	- Evaluate agreement between TM and R (mercury – reference) - Evaluate validity of TM in detecting R fever (≥37.6ºC) - Acceptability: not defined	- Significant positive correlation (95% CI 0.52, 0.86, *P* < .01; *r* = 0.78) - Bias 0.50ºC ± 0.37ºC (95% CI 0.41, 0.59) - Sensitivity and specificity of TM to detect R fever was 86% and 89%, respectively -Acceptable agreement between TM and R
Smitz et al., 2009	Prospective comparison design	Geriatric unit *N* = 100 800 readings	- Evaluate accuracy of TM (2 different models) to predict R (reference) fever (≥37.8ºC) - Acceptability: not defined	- Bias TM (Thermoscan) 0.20ºC (SD 0.32) 95% LOA -0.83, 0.42; fever predictability max error 0.7ºC (mean error 0.3ºC) - Bias TM (Genius) -0.56ºC (SD 0.39 ºC) 95% LOA -1.32, 0.20; fever predictability max error 1.6ºC (mean error 0.4ºC) - Strong positive correlation TM to R *R* = 0.91; 95% CI 0.75, 0.89; *p* < 0.001) - TM can predict R rectal temperature in normothermic and in febrile inpatients. However, the predictive accuracy depends on both operator technique and quality of instrumentation.
Spitzer 2008	Prospective comparison design	ICU *N* = 66 198 readings	- Evaluate agreement between TM (R ear), TM (L ear) and O (reference) - Acceptability: not defined	Bland-Altman data (bias and LOA not reported) - Right TM mean 98.7ºF (*SD* 1.4); correlation *r* = 0.70; higher in 29% (n = 19) Left TM mean 98.6ºF (*SD* 1.5); correlation *r* = 0.44; higher in 44% (n = 29) Versus O- higher 33% (*n* = 22) -No significant difference between three measures (*p* = .6428)
Stelfox et al., 2010	Descriptive comparison design	ICU *N* = 14 736 readings	- Evaluate agreement between TAT and BL (reference). - Determine accuracy (sensitivity/specificity) of TAT to detect fever and hypothermia - Acceptability: ± 0.5ºC	- Agreement greatest for normothermia (bias -0.35ºC, 95% CI -0.37, -0.33) - Hypothermia (<36ºC) TAT measured higher temperatures (bias 0.66ºC, 95% CI 0.53, 0.79) - Hyperthermia (≥38.3°C) TAT measured lower temperatures (bias -0.90°C; 95% CI, -0.99, -0.81). - Sensitivity and specificity for fever 0.26 (95% CI 0.20, 0.33) and 0.99 (95% CI 0.98, 0.99), respectively - TAT should not to be used in situations where body temperature needs to be measured with accuracy
Varney et al., 2002	Cross-sectional design	ED *N* = 95 275 readings	- Evaluate correlation of O, TM measurements to identify R fever (38 ºC) in patients presenting with symptoms of infection. - Acceptability: discordance defined as any R temp over 38ºC and 0.5ºC over O or TM	- O, TM to R correlation *r* = 0.621 and *r* = 0.764, respectively - R identified fevers missed by O 14.7% (*n* = 14) and TM 12.2% (*n* = 11); 5.6% (*n* = 5) had R fever but were afebrile O and TM - In 19 episodes of R fever (afebrile O and TM), 68% (*n* = 13) required admission - Identification of fever, in addition to other clinical signs and symptoms, may be an important determination in the search for evidence of infection.

(table continues)

Author/ Year	Study Design	Population Sample	Purpose / A priori acceptability	Findings/Conclusion
Winslow et al., 2012	Prospective descriptive	Peri-Op *N* = 64 447 readings	- Evaluate agreement of TAT to O – preoperative (reference) and BL – postoperative (reference) to identify hypothermia (<36 ºC) - Acceptability: LOA ± 0.5 ºC	- Preoperative TAT to O bias 0.43ºC (*SD* 0.52; LOA -1.46, 0.61) - Postoperative TAT to BL mean bias -0.76 ºC (*SD* 1.14; LOA -3.04, 1.52) - TAT failed to detect any hypothermic (< 36 ºC) temperatures - BL hypothermic readings 33 (52%) intraoperative; 27 (42%) postoperative - - Lack of agreement between TAT and O, BL. TAT not recommended for the perioperative population
Wolfson et al., 2013	Method comparison design	Oncology *N* = 34 68 readings	- Evaluate agreement, in febrile patients, between TAT and O (reference) - Acceptability: bias ± 0.6ºF, precision between -1.0 ºF and +1.0 ºF	- Bias and precision 0.80ºF (*SD* 1.2) - Number of temperature differences >± 1.0ºF n = 13 (43%); > ± 2.0ºF n = 5 (17%) - TAT not recommended for febrile patients
Woodrow et al., 2006	Quantitative comparison design	Medical-Surgical; ICU *N* = 178 178 readings	- Evaluate agreement between NT to TM (reference), in oral equivalent and core equivalent modes -Acceptability: maximum difference of 1.0 ºC	- NT to TM (oral equivalent) bias 0.47ºC (*SD* 0.69; 95% CL -0.883, 1.83; $p < 0.001$); $t = 7.038$ - NT to TM (core equivalent) bias -0.59ºC (*SD* 0.75; 95% CL -0.88, 2.08; $p < 0.001$); $t = -6.73$ - Devices are not comparable; accuracy is undetermined

Note. LOE = Level of Evidence: I = systematic review or metanalysis; IV = cohort studies; RN = registered nurse; Peri-Op = perioperative; ES = esophageal temperature; O = oral temperature; TAT = temporal artery temperature; TM = tympanic membrane temperature; AX = axillary temperature; R = rectal temperature; CH = chemical/disposable dot thermometers; PAC = pulmonary artery catheter; BL = bladder temperature; NT = no touch forehead thermometer; IN = inguinal; LCCC = Lin's concordance correlation coefficient

The body of evidence was categorized using the Hierarchy of Evidence (see Figure 1) described by Fineout-Overholt et al. (2010). There was one level I study, a systematic review (SR), which met the inclusion criteria for this review. The rest of the included studies were cohort studies, categorized as level IV evidence.

The quality of the research was evaluated using the quality indicators described by Hooper and Andrews (2006; see Table 1). Of the 46 level IV studies, two were determined to have a quality grade of A, while there were 20 with a grade of B, and 18 with a grade of C. The lowest quality grade, D, was assessed for five studies, as they were found to have less than two quality indicators. The quality grade for each article is specified in Table 5.

Table 5

Quality Indicator Grades

Author/ Year	Grade	Author/Year	Grade
Level I Evidence			
Jefferies et al., 2011	A		
Level IV Evidence			
Amoateng-Adjepong et al., 1999	B	Lefrant et al., 2003	D
Barringer et al., 2011	C	Leon et al., 2005	B
Bodkin et al., 2014	C	Marable et al., 2009	B
Calonder et al., 2010	B	Mason et al., 2015	C
Counts et al., 2014	B	McConnell et al., 2013	B
Duncan et al., 2008	C	Moran et al., 2007	B
Dunleavy, 2008	D	Myny et al., 2005	B
Fallis, et al., 2006	B	Nonose et al., 2012	C
Farnell et al., 2005	C	Potter et al., 2003	C
Fetzer et al., 2008	C	Prentice & Moreland, 1999	C
Fountain et al., 2008	C	Rajee & Sultana, 2006	C
Frommelt et al., 2008	C	Rubia-Rubia et al., 2011	B
Gilbert et al., 2002	B	Shin et al., 2013	C
Giuliano et al., 1999	A	Singler et al., 2013	B
Giuliano et al., 2000	B	Smith, 2003	B
Hasper et al., 2011	D	Smitz et al., 2000	C
Haugan et al., 2012	B	Smitz et al., 2009	C
Irvin, 1999	C	Spitzer 2008	D
Jensen et al., 2000	B	Stelfox et al., 2010	B
Khan et al., 2006	D	Varney et al, 2002	C
Kimberger et al., 2007	B	Winslow et al., 2012	B
Langham et al., 2009	C	Wolfson et al., 2013	B
Lawson et al., 2007	A	Woodrow et al., 2006	C

Level I Evidence: Systematic Review

Jefferies, Weatherall, Young, and Beasley (2011) conducted a SR to evaluate the accuracy of peripheral thermometers in the detection of fever (> 37.5°C) in critically ill patients. While only three studies met the inclusion criteria, data were evaluated on the seven TM thermometers (including different brands and models), O (digital), and R (digital). Five of the TM thermometers and the O thermometer were within ±0.2ºC of the PAC. The bias of the R to PAC was outside the acceptable criterion. The authors concluded that the TM and O devices provided accurate temperature readings on febrile patients; the R device was not recommended (Jefferies, Weatherall, Young, & Beasley, 2011).

Level IV Evidence: Cohort Studies

Tympanic membrane (TM) thermometer. The TM thermometer is one of the earliest noninvasive devices to be developed and used in hospitalized patients (Gallimore, 2004). The ease of use, rapidity of results, and noninvasive nature of the TM device created an opportunity for the rapid diffusion of this technology into patient care (Gallimore, 2004; Manian & Griesenauer, 1998). TM readings can be affected by ambient temperature, operator technique, a narrow ear canal, and can vary from side to side (Sund-Levander & Grodzinsky, 2013). Ear wax and otitis media may affect TM readings though there is conflicting data on both of these variables (Sund-Levander & Grodzinsky, 2013). Although there have been significant advances in technology, the evidence on the accuracy and reliability of this device continues to be inconsistent.

There were 27 studies which included an evaluation of the TM thermometer included in this review. There were 17 studies in which had findings indicated that the TM thermometer was clinically unacceptable for use. Alternatively, the findings from 10 others studies led the authors to conclude that the TM thermometer is clinically acceptable for use in hospitalized adult patients. Eight studies addressed diagnostic accuracy and those results are described separately.

Clinically unacceptable. In the ICU or in the perioperative environment, most authors using the PAC (Farnell et al., 2005; Giuliano et al., 1999; Giuliano et al., 2000; Lawson et al., 2007; Moran et al., 2007; Nonose et al., 2012) or BL (Dunleavy, 2010; Khan, Vohra, Paul, Rosin, & Patel, 2006; Langham et al., 2009; Nonose et al., 2012) as the reference temperatures, concluded that the TM device had unacceptably wide variability and did not recommend the device for use. Clinically acceptable differences (from the reference standard) ranged from ± 0.2ºC to ± 0.5ºC or ± 0.2ºF to ± 0.5ºF. Farnell et al. (2005), using an expert panel, defined clinical acceptability by determining if the inaccuracy would result in a delay in care or unnecessary interventions. A priori clinically acceptable differences were not established in all studies (Dunleavy, 2010; Khan et al., 2006; Moran et al., 2007; Nonose et al., 2012).

Most of the studies using the PAC as the reference temperature scored high (A or B) on the quality of the research (Giuliano et al., 1999; Giuliano et al., 2000; Lawson et al., 2007; Moran et al., 2007; Rubia-Rubia et al., 2011). The studies by Farnell et al. (2005), Langham et al. (2009), and Nonose et al. (2012) had three to four quality indicators and received a quality grade of C. The research by Dunleavy (2008) and Khan

et al. (2006) were evaluated and found to have less than two quality indicators and received a grade of D.

Outside of the ICU or the perioperative environments, researchers used alternative reference sites including the O (digital or mercury; Dunleavy, 2010; Fountain et al., 2008; Frommelt et al., 2008; Gilbert et al., 2002; Irvin, 1999; Prentice & Moreland, 1999). The R (digital or mercury) was also used the reference temperature by Jensen et al. (2000) and Varney et al. (2002). In these studies, half of the authors identified clinically acceptable differences which ranged from ± 0.3ºC to ± 0.5ºC. Four groups failed to specify clinically acceptable differences (Dunleavy, 2010; Fountain et al., 2008; Frommelt et al., 2008; Irvin, 1999). Of the eight studies using the O or R as the reference temperature, two had a quality grade of B (Gilbert et al., 2002; Jensen et al., 2000). Five of these studies received a quality grade of C (Fountain et al., 2008; Frommelt et al., 2008; Irvin, 1999; Prentice & Moreland, 1999; Varney et al., 2002).

Clinically acceptable. There were 10 studies in which the findings indicated that the TM thermometer was accurate and was acceptable as an alternative temperature device. Three studies used the PAC as the reference temperature (Amatoeng et al., 1999; Haugan et al., 2012; Rubia-Rubia et al., 2011) and one used the ES as a reference temperature (Hasper et al., 2011). Haugan et al. (2012) also used R as the reference temperature as well as Rajee and Sultana (2006), Singler et al. (2013), Smitz, Giagoultsis, Dewe, and Albert (2000), and Smitz, Van de Winckel, and Smitz (2008). Spitzer (2008) used the O thermometer as the reference while Leon et al. (2005) used the AX (mercury) for the reference temperature. Clinically acceptable parameters, described in four studies,

ranged from ± 0.5ºF and ± 0.2ºC to ± 0.5ºC (Amatoeng et al., 1999; Haugan et al., 2012; Rajee & Sultana, 2006; Rubia-Rubia et al., 2011).

When considering the quality grade for this group of studies, five had enough quality indicators to be graded a B, although there was variation in the indicators scored (Amatoeng et al., 1999; Haugan et al., 2012; Leon et al., 2005; Rubio-Rubio et al., 2011; Singler et al., 2013). There were three with a quality grade of C (Smitz et al., 2000; Smitz et al., 2008; Rajee & Sultana, 2006). The studies by Hasper et al. (2011) and Singler et al. (2013) had less than two quality indicators and received a grade of D.

Diagnostic Accuracy

In addition to the evaluation of the TM thermometer for accuracy and reliability, one SR and eight cohort studies included an evaluation of the diagnostic accuracy of the TM device to detect fever (Jefferies et al., 2011; Leon et al., 2005; Prentice & Moreland, 1999; Rajee & Sultana, 2006; Rubia-Rubia et al., 2011; Singler et al., 2013; Smitz et al., 2000; Smitz et al., 2009; Varney et al., 2002). Fever was defined in seven studies and ranged from 37.5ºC to 39ºC. Singler et al. (2013) used an adjudicated diagnosis of infection, determined by a panel of two independent physicians.

Unacceptable diagnostic accuracy. The TM thermometer was found to have a low diagnostic accuracy for fever in five of the eight studies. In the SR, Jefferies et al. (2011) determined that two of seven TM thermometers did not meet accuracy criteria. When sensitivity to predict fever was evaluated, the values ranged from 51% to 73% (Leon et al., 2005; Prentice & Moreland, 1999; Rajee & Sultana, 2006). When evaluating the diagnostic accuracy for infection, Singler et al. (2013) found the TM area under the

curve (AUC) to be 0.73 (95% CI 0.66-0.81). Out of those patients with the adjudicated diagnosis of infection (*n* = 105), 22.8% had TM fevers >38.ºC whereas R fever was present in 43%. Varney et al. (2002), using R temperature as the reference, found 12.2% (*n* = 11) with R fever that were afebrile using the TM. Of 19 occurrences of R fever that were afebrile by other routes (O and TM), 68% (*n* = 13) required admission.

Acceptable diagnostic accuracy. There were findings from three studies which concluded that the TM could provide acceptable diagnostic accuracy for fever. In their SR, Jefferies et al. (2011) found that five of seven TM thermometers were accurate for diagnosing fever. Rubia-Rubia et al. (2011) found the sensitivity of the TM for fever (PAC fever range 38.5ºC to 38.9ºC) was 0.987 ± 0.007. In 2000, Smitz et al. concluded sensitivity of the TM thermometer, compared to R (mercury) was 86%. In a later study, Smitz et al. (2009) evaluated the ability of two different TM thermometers to predict R fever. While both models had comparable sensitivities (94%), the maximal errors (0.7ºC and 1.6ºC) were pointedly different. Therefore, the predictive accuracy of the TM thermometer was dependent upon operator technique and the quality of the equipment (Smitz et al., 2009).

Disposable Chemical Thermometers. Another alternative for temperature assessment in the adult hospitalized patient is the disposable chemical (CH) dot thermometer (such as TempaDot ™ or NexTemp™). Often considered a device used for patients in isolation, there is a growing body of literature evaluating the accuracy and reliability of the CH thermometer in other clinical areas. The CH thermometer has a series of dots on the strip which change from white (or green) to black when exposed to

heat (mouth or axilla); each black dot represents an incremental change in temperature (Potter, Schallom, Davis, Sona, & McSweeny, 2003). The assessed temperature is determined by reading the corresponding temperature of the last black dot. Reading accuracy for the CH thermometers has been the most significant issue identified (Creagh-Brown, James, & Jackson, 2005; Frommelt, Ott, & Hays, 2008; Potter et al., 2003). In a study on reading accuracy, Creagh-Brown, James, and Jackson (2005) reported that of 78 nurses, only 23% gave the correct temperature reading.

There were eight papers comparing the CH thermometer to various reference temperatures in this review. As with other devices, the findings for use of the CH thermometer are mixed. Limited use or cautionary use was a recommendation from half of these papers, so their findings are delineated separately.

Clinically unacceptable. The findings from two studies did not support the use of CH thermometer in adult, acute care patients. Counts et al. (2014), and Fallis et al. (2006) each compared the CH to the O thermometer. Clinically acceptable differences of ± 0.3ºC were defined by both groups. In addition to reporting mean differences, Counts et al. reported 21% (n = 10) differences > ± 0.5ºC and 13% (n = 6) differences ≥ ± 1.0ºC. Fallis et al. found that 91% of CH readings either overestimated (81%) or underestimated (10%) the O temperature. Additionally, Fallis et al. reported 65 instances where the CH thermometers failed to demonstrate any color change. The quality of both studies were evaluated using the quality indicators described by Hooper and Andrews (2006). Both studies had sufficient quality indicators to receive a grade of B.

Clinically acceptable. In this review, findings from two of eight studies demonstrated that the CH thermometer was an acceptable thermometer for use in adult hospitalized patients. Clinical acceptability was defined as ±0.2ºC by Rubia-Rubia et al. (2011); Rajee and Sultana (2006) did not define this parameter. Rubia-Rubia et al. used the CH thermometer in the axilla and found a narrow mean difference of 0.2ºC with the PAC temperature. Rajee and Sultana using the R (mercury) as the reference site, found a nonsignificant bias for both agreement and repeatability between the R and CH thermometers. When evaluating the studies for quality indicators, Rubia-Rubia et al. had a quality grade of B, while Rajee and Sultana had a quality grade of C.

Cautionary or limited use. Authors of four additional studies concluded that the CH thermometer should be used with caution or have limited use (Farnell et al., 2005 Fountain et al., 2008; Frommelt et al., 2008; Potter et al., 2003). Clinically acceptable differences were described by Farnell et al. (2005) and defined by Potter et al. (2003) to be ± 0.3ºC. Farnell et al. compared the PAC to axillary CH thermometers. Although the mean difference from the PAC was 0.2ºC, they also reported a large percentage of readings, 88.6%, that overestimated or underestimated the PAC. Farnell et al. added that, based on the measured CH temperatures, 70 patients would have had delayed interventions (15.3%) or would have received unnecessary interventions (28.8%).

Fountain et al. (2008), Frommelt et al. (2008), and Potter et al. (2003) used the O (digital or mercury) as the reference temperature. These authors identified either a narrow bias (Fountain et al., 2008; Frommelt et al., 2008) or a strong correlation (Potter et al., 2003) in the evaluation of the CH thermometer. However, each group also reported a

number of measurements that differed by more than 1.0ºF. Fountain et al. noted that 30% (*n* = 18) differed from the O temperature by 1.0 ºF and 3% (*n* = 2) differed by 2.0 ºF. Frommelt et al. reported 2% (*n* = 2) measured ≥ 2.0ºF. Lastly, Potter et al. noted that 25% of the CH temperatures measured either overestimated (11.8%) or underestimated (10.8%) body temperature by 0.4ºC or more. The authors of these four studies concluded that there were sufficient readings which were ≥ 1.0º (F or C) to cause concern for use in the clinical environment. The CH thermometer should be used for screening and abnormal findings should be validated with a more accurate thermometer. All four studies in this group received a quality grade of C.

Temporal Artery Thermometer (TAT). The TAT device, marketed for its rapid results and noninvasiveness, was developed for patient care in 1999 (Healthcare Improvement Scotland, 2012; Ostrowsky et al., 2003). According to Sund-Levander and Grodzinsky (2013), the TAT provides an indirect measure of patient temperature and can be influenced by operator technique, skin thickness, local blood flow, and ambient temperature. There were 19 studies included in this review that evaluated the TAT against various reference temperature sites. As with other clinical thermometers, the findings are inconsistent. Five studies addressed the diagnostic accuracy of the TAT and these results are described separately.

Clinically unacceptable. Findings from 13 studies comparing the TAT to a variety of reference temperatures, did not support the use of this device. Four studies were conducted using PAC or BL as the reference temperature (Kimberger et al., 2007; Rubia-Rubia et al., 2011; Stelfox et al., 2010; Winslow et al., 2012). Rubia-Rubia et al.

(2011) found the TAT device had the lowest overall valuation score and the poorest validity. In their comparison studies Kimberger et al. (2007), Stelfox et al. (2010), and Winslow et al. (2012) found the TAT to lack agreement, particularly at the hypo- and hyperthermic ranges. In perioperative patients, Winslow et al. found the TAT did not register any hypothermic (<96.8 ºF) temperatures. In contrast, 52% (n = 33) of the BL readings indicated intra-operative hypothermia and 52% (n = 42) were hypothermic in the post-anesthesia care unit. All of these studies earned a quality grade of B.

When using O as the reference temperature, authors from seven studies concluded the TAT was not an acceptable alternative for clinical (Bodkin et al., 2104; Counts et al., 2014; Fountain et al., 2008; Frommelt et al., 2008; Marable et al., 2009; Winslow et al., 2012; Wolfson et al., 2013). Clinically acceptable differences were defined in five studies and ranged from ± 0.3ºC to ± 0.5ºC and ± 0.5ºF to ± 0.6ºF. As previously stated, neither Fountain et al. (2008) and Frommelt et al. (2008) defined clinically acceptable criteria.

Additional findings of interest, Fountain et al. (2008) found significant temperature differences of > 1.0ºF in 43% (n = 26) and >2.0ºF in 8% (n = 5) while Frommelt et al. (2008) noted differences of > 2.0ºF in 6% (n = 5). Only one study (Marable et al., 2009) described the evaluation of the TAT device using three different methods (forehead to ear; forehead only; behind the ear only). The authors found that two of the three methods exceeded their pre-defined clinically acceptable differences. Marable et al. (2009) also studied the influence body mass index (BMI; < 30 or ≥ 30) might have on TAT readings and found that TAT readings were lower than O in obese patients (BMI ≥ 30; p = 0.313).

The quality indicators and grade described by Hooper and Andrews (2006) were used to evaluate these studies. When considering the quality of these studies, four were determined to have a quality grade of B (Counts et al., 2014; Marable et al., 2009; Winslow et al., 2012; Wolfson et al., 2013). The research by Bodkin et al. (2014), Fountain et al. (2008), and Frommelt et al. (2008) received a quality grade of C.

One group of authors compared the TAT device to the TM as the reference site (Fetzer et al., 2008). The authors found that, although the mean was within the clinically acceptable range, the confidence intervals were significantly wider than a priori criterion. Fetzer et al. (2008) also reported a moderate correlation (r = .421; p = .000), but a low coefficient of determination (17.7%). The quality grade assigned to the research done by Fetzer et al. was a C.

Clinically acceptable. The results of seven studies supported the use of the TAT thermometer in hospitalized adults. When compared to the PAC, Lawson et al. (2007) and Myny et al. (2005) had comparable findings with the TAT of -0.02 ºC and 0.14 ºC, respectively. Calonder et al. (2010) found a statistically significant difference between the TAT and ES; however, the differences did not meet the clinically significant threshold. Langham et al. (2009), using the BL thermometer as the reference, determined that the TAT device, while not meeting their whole criteria for acceptability, performed well enough for use in the perioperative patient. Clinically acceptable criteria were defined in all four studies and ranged from ± 0.3ºC to ± 0.5ºC. In evaluating the quality of the research, one study earned an A (Lawson et al., 2007), while the two earned a quality

grade of B (Calonder et al., 2010; Myny et al., 2005); Langham et al., received a quality grade of C.

Three comparison studies of the TAT used the O temperature as the reference site (Barringer et al., 2011; Mason et al., 2015; McConnell et al., 2013). In pre- and postoperative patients Barringer et al. (2011) found adequate agreement between the TAT and O temperatures; however, they did not define the criterion for clinical acceptability. Mason et al. (2015) and McConnell et al. (2013) defined clinically acceptable differences to be 0.2ºF and ≤1.0ºF, respectively. McConnell et al. also evaluated the intrarater reliability and found the mean differences between investigators was within acceptable standards. With regards to the quality of the research, the study by McConnell et al. was evaluated as a B, while Barringer et al. and Mason et al. earned a grade of C.

Diagnostic Accuracy

The diagnostic accuracy of the TAT thermometer was evaluated in five studies. One group used the PAC as the reference site (Rubia-Rubia et al. 2011) while two others used the BL (Kimberger et al. 2007; Stelfox et al., 2010). Lastly, the O (digital) and R (digital) were used as reference sites by Bodkin et al. (2014) and Singler et al. (2013), respectively. The conclusion reached by all groups was that the TAT thermometer had significant limitations in detecting hypo- or hyperthermia. Rubia-Rubia et al. (2011) found that the TAT had the lowest AUC, specificity and positive predictive value (PPV) for PAC fever of 38.5ºC (0.853; 83%; PPV 47%). Concerning to the ability of the TAT to detect fever (BL >37.8 ºC), Kimberger et al. (2007) found a sensitivity and specificity of 0.72 and 0.97, respectively. The sensitivity of the TAT to detect fever was found to be

much lower (0.26) by Stelfox et al. (2010). In their study, Bodkin et al. found the mean difference in afebrile patients was 0.12ºC while the mean difference in the febrile group was much greater (0.87ºC). Additionally, 57% of the fevers recorded by the O device, were not measured by the TAT (Bodkin et al., 2014).

The accuracy of a thermometer to detect hypothermia is also an important consideration, particularly for the perioperative environment. Kimberger et al. (2007) and Stelfox et al. (2010) used the BL for comparison. Kimberger et al. reported the TAT sensitivity, specificity, and PPV for hypothermia (BL <35.5 ºC) was 0.29, 0.95, and 0.31. Stelfox et al. found that TAT recorded higher temperatures (mean 0.66ºC) at hypothermic ranges (<36ºC). Finally, Singler et al. (2013) evaluated the accuracy of the TAT in predicting infection (determined by the AUC). The authors used the R thermometer as the reference site and found the diagnostic accuracy for the TAT was significantly lower (AUC 0.65 [95% CI 0.57-0.73] $p < 0.001$).

No-Touch (NT) Infrared Forehead Thermometers. The employment of NT thermometers has grown as a public health tool to screen large numbers of patients for fever, as seen with the recent Ebola virus epidemic and the 2003 severe acute respiratory syndrome epidemic (Liu, R. Chang, & W. Chang, 2004). Along with the public health use, the rapid results and non-contact nature of the NT thermometers have found favor in acute care hospitals. Although the NT devices are widely used, there is a paucity of data on the accuracy and reliability of these devices. Only two studies (Duncan, Bell, Chu, & Greenslade, 2008; Woodrow et al., 2006) on NT thermometers met the inclusion criteria for this review.

Clinically unacceptable. Neither group of researchers found the device valid for use in acutely ill hospitalized adults. Duncan, Bell, Chu, and Greenslade (2008) compared the NT to O temperature in ED and BL temperature in the ICU while Woodrow et al. (2006) used the TM as the reference thermometer. Clinically acceptable differences were established in both studies (0.3ºC and 1.0ºC, respectively). Duncan et al. found the NT device to be reliable (between NT readings, $r = 0.94$). However, agreement with the O and BL was poor. Woodrow et al. found the agreement between the NT and TM to be acceptable. However, the analysis demonstrated a number of temperature differences over 1.0ºC, TAT to TM oral mode (24.7%; $n = 26$) and TAT to TM core mode (43.8%; $n = 32$). Additionally, the *t*-test for both comparison groups (TAT to TM oral and TAT to TM core) were statistically and significantly greater ($t = 7.038$; $p < 0.001$ and $t = -6.736$; $p < 0.001$).

Both of the studies in this category had four of nine quality indicators and received a quality grade of C. Duncan et al. (2008) provided information on temperature measurement techniques while Woodrow et al. (2006) provided information on data collector training. Otherwise, the quality indicators were the same in both studies.

Axillary (AX) Thermometry. As with other devices and routes, the AX site is favored for its noninvasiveness and accessibility. However, the accuracy and reliability of routine temperatures assessed via the axilla remain a debatable topic. According to Sund-Levander and Grodzinsky (2013), the axillary site can be affected by ambient temperature, local blood supply, sweat, placement of the probe and dwell time. Further, temperatures can vary by as much as 1.4ºC between the right and left axilla (Sund-

Levander & Grodzinsky, 2013). Lastly, axillary temperatures lag significantly far behind other sites, especially during rapid temperature change (Sund-Levander & Grodzinsky, 2013).

Clinically unacceptable. Of the 13 studies which included a comparison to axillary temperature, 10 concluded that axillary temperatures should not be used as a source for routine temperature assessment. When compared to PAC (Lawson et al., 2007; LeFrant et al., 2003; Moran et al, 2007; Myny, De Waele, Defloor, Blot, & Colardyn, 2005), BL (Langham et al., 2009; Khan et al., 2006; Moran et al., 2007), O (Barringer et al., 2011; Marable et al., 2009; Mason et al., 2015), and R (mercury; Jensen et al., 2000), axillary temperatures were not recommended as a source for temperature assessment. Clinically acceptable differences were described by six groups and ranged from 0.3ºC to 0.5ºC and 0.2ºF to 0.5ºF. Lefrant et al. (2003), Moran et al. (2007), Khan et al. (2006) and Barringer et al. (2011) did not report clinically acceptable differences in their studies.

The quality of research in this group was variable. Only one study (Lawson et al., 2007) was found to have all nine quality indicators and received a quality grade of A. Four studies (Jensen et al., 2000; Marable et al., 2009; Moran et al., 2007; Myny et al., 2005) were found to have five to seven quality indicators and earned a quality grade of B. A quality grade of C was given to studies by Barringer et al. (2011), Langham et al. (2009), and Mason et al. (2015). There were two studies (Khan et al., 2006; Lefrant et al., 2003) that received a quality grade of D, as each study had only one or two quality indicators.

Clinically acceptable. Three studies included in this review had findings which supported the use of the axillary thermometry as an acceptable alternative for temperature assessment. Nonose et al. (2012) compared AX to BL and PAC reference temperatures while Rubia-Rubia et al. (2011) used the PAC alone. Smith (2003) compared the gallium-in-glass to R (mercury). Only Rubia-Rubia et al. specified a pre-defined clinically acceptable range of 0.2ºC.

Nonose et al. (2012) determined that the AX was an acceptable alternative for temperature assessment AX as it had a better correlation (r = 0.64;) and narrower limits of agreement with the BL than the TM. Smith (2003) determined the AX (gallium-in-glass) was acceptable based on correlation (r = 0.886) and mean difference (0.25; 0.167; 0.339). Rubia-Rubia et al. (2011) evaluated accuracy, reliability and validity of the AX along with other variables (ease of use, cost, speed, durability of the instrument, and patient comfort). Based on the valuation score, the authors concluded that the gallium-in-glass AX route demonstrated the strongest results. Of note, the authors noted the required dwell time for this device was 12 minutes, a significant limitation given the time constraints for nurses.

The quality of the studies by these three groups was also variable. The studies by Rubia-Rubia et al. (2011) and Smith (2003) were evaluated for the quality indicators, and a quality grade of B was assessed. The research by Nonose et al. (2012) had a quality grade of C.

Unanticipated Limitations

There were several unanticipated limitations in this IR which may impact the findings. First, because the DNP project was designed as an IR of the literature, no actual research or comparison of thermometers was conducted. The findings from the IR are the result of conclusions drawn by the authors of the research. Another limitation of this IR review is that this author was unable to access the most recently completed studies. Therefore, the most recent findings analyzed were at least 2 years old (time from completing a project to publication). Finally, new advances in clinical thermometers occur rapidly, and I was unable to evaluate the accuracy of the most recent technologies.

Implication of the Findings

The implication of the findings from this IR can be viewed from the perspective of the individual (nurses and patients), the community, the institution and the system. In 2001, the Institute of Medicine (IOM) described a health care system that was unable to provide quality healthcare in the face of rapid technological and medical advances. The IOM (2001) recommended that "patients should receive care based on the best available scientific knowledge…care should not vary from clinician to clinician" (p. 8). The statement from the IOM is consistent with some of the challenges associated with advances in thermometer technology.

Nurses and assistant nursing staff are the health care providers directly responsible for the assessment of vital signs, including temperature. Patients are the recipients of our nursing care, and their outcomes can be directly related to the accuracy or inaccuracy of the assessed temperatures. The findings from the IR can serve as a

resource for nursing personnel, providing a synthesis of the body of evidence related to thermometers. Nurses can use the recommendations from the IR to guide their practice.

The implications of the findings from the IR as it relates to the community are broader, but also important to consider. The community in this context was the patients, families, and area served by the hospital. As nurses and nurse leaders utilize the IR as a resource to help nurses obtain more accurate temperatures, the quality of the patient care is improved. Improving the quality and safety of patient care and patient outcomes leads to strengthened patient and family engagement (Agency for Healthcare Research and Quality, 2013). Stronger patient engagement leads to improved patient satisfaction, which increases patient loyalty and the reputation of the community (Hall, 2008).

The findings and recommendations from the IR may also impact the hospital and hospital system. According to the IOM (2001) report "a health care system frequently falls short in its ability to translate knowledge into practice, and apply new technology safely and appropriately" (pp. 2–3). The knowledge-to-practice gap which drove the need for the IR was the inability of nurses to specify the evidence to support their practice related to in temperature assessment. Inaccurate devices or those that rely on particular technique or training can result in faulty low or high patient temperatures. The IR provides a synthesis of the current body of evidence on clinical thermometers and can serve as a resource to nursing leaders and system leaders as they consider current practice. Additionally, the IR would be an excellent resource for the healthcare system in considering the purchase of new or alternative thermometers. Lastly, implementation of the recommendations from the IR may improve the accuracy of temperature assessment

which is directly tied to the system-wide priority for early sepsis identification. Early sepsis identification has been demonstrated to decrease health care costs and morbidity and mortality (Dellinger et al., 2012; Hall et al., 2011).

Implication for Positive Social Change

Closing the knowledge-practice gap of temperature assessment devices, their accuracy, and reliability has significant implications for changing the culture of nursing practice. Safe, quality patient care is a fundamental tenet in healthcare, as is our mandate to sustain these processes while striving to mitigate increasing healthcare costs. Without the appropriate use of EBP in the area of temperature assessment, devices are often chosen for the novelty, the convenience or the noninvasive nature of the device (Manian & Griesenaur, 1998). Many nurses assume a device is accurate and reliable just because it is adopted by an organization (Ostrowsky et al., 2003). It is critical to have knowledge of the devices used in an organization or on one's patient population to ensure high quality and safe patient care. Often it is the nursing staff of healthcare organizations that raise safety concerns about a device and are change agents and advocates for their patients (Bahr et al., 2010; Dunleavy, 2010; Ostrowsky et al., 2003).

Recommendations

The use of clinical thermometers has been considered a routine procedure with little potential for error. However, there are many factors which can impact the accuracy of temperature assessment including the thermometer specifications, operator technique, and patient characteristics (Davie & Amoore, 2010). The wide variety of clinical

thermometers currently available, each with different specifications, has increased the complexity of the issues related to temperature accuracy.

The review of the literature did not reveal a clinical thermometer with better accuracy and reliability than any other. There were contradictory findings for all of the clinical thermometers evaluated, except the NT thermometer. The quality of this body of research was limited, with 22 of 47 studies having an acceptable quality grade of A or B. Given the lack of evidence supporting any one thermometer, the recommendations will address each device specifically. Finally, additional suggestions are presented for improving the accuracy of clinical thermometers already in use.

Tympanic Membrane (TM) Thermometers

There was one systematic review on the TM thermometer, the strongest level of evidence, which concluded that the device was clinically accurate (Jefferies et al., 2011). It is important to note that the accuracy was specified to different TM models. Descriptions of differences by model was also consistent with other researchers (Giuliano et al., 2000; Haugan et al., 2012; Smitz et al., 2009).

Another factor associated with the variability and accuracy of the TM thermometer is technique. In both studies by Giuliano et al. (1999, 2000), they included a discussion about the challenges of training and technique with the TM thermometer. Amoeteng et al. (1999) unexpectedly found that the accuracy of the TM temperatures was lower with staff that used the device routinely. Gilbert et al. (2002) observed staff taking TM temperatures by reaching over the patient to the opposite ear (9.44%).

For organizations where the TM thermometer is currently in use, ensure adequate initial training and consider annual revalidation of this skill. Appropriate technique is also critical and should be emphasized for nurses and nursing personnel using the TM thermometer routinely. The recommendation for organization leaders considering the purchase of the TM thermometer should consider evidence which specifies a superior model.

Disposable Chemical (CH) Thermometers

While CH thermometers are favored for their convenience, portability, and ease of use, they should be limited to use in specific circumstances (such as isolation). Reading accuracy was identified as a potential limiting factor with this thermometer (Creagh-Brown et al., 2005; Frommelt, Ott, & Hays, 2008; Potter et al., 2003). Although most of the authors found the CH thermometer to be clinically accurate, they also reported a significant number of CH measurements that differed by more than 0.5ºC (F). These findings create concern for the routine use the CH thermometer. When the patient's condition requires the use of this thermometer, any abnormal temperatures should be validated by another, more reliable, thermometer.

Temporal Artery Thermometer (TAT)

Most of the data support the conclusion that the TAT is accurate in the normothermic range, but is less accurate in the hypo- and hyperthermic ranges. A significant limitation when the screening for fever or hypothermia in perioperative patients. The accuracy of the TAT device may also rely on appropriate technique; however, only one study (Marable et al., 2009) evaluated this variable. Given the

concerning results of this device in measuring temperatures outside of anything clinically normal, the TAT thermometer may not be accurate enough for use in acutely ill adults.

No-Touch (NT) (Infrared) Forehead Thermometers

There were limited data on the accuracy and reliability of these devices. There were only two studies which met the inclusion criteria for this review and neither found the device to be accurate enough for clinical use. More research is needed on the NT thermometer before use in acutely ill adults can be recommended.

Axillary (AX) Thermometry

The AX site as a source for routine clinical temperature assessment in adult hospitalized patients were not recommended. Only three of 13 studies determined the AX site to be acceptable. Two were specific to the gallium-in-glass thermometer, which is not as widely use as the digital thermometer. The recommendation by Rubia-Rubia et al. (2011) to use the gallium-in-glass with a 12-minute dwell time is unrealistic for clinical nurses today.

When clinical situations preclude the use of an oral thermometer, such as with combative or confused patients, clinicians may consider the AX site as a safe alternative. However, the data on the use of AX thermometry do not support the use of this site for clinical use. Instead, an alternative noninvasive thermometer should be used (Sund-Levander & Grodzinsky, 2013).

Additional considerations. Although the conclusions from this review did not pinpoint a clearly superior clinical thermometer, there are some basic factors which can help improve the clinical accuracy of thermometers already in use. The first is the

importance of correct anatomical placement. Anatomical placement for many temperature sites is based on proximity to arterial flow, allowing the opportunity to evaluate changes in blood temperature (Sund-Levander & Grodzinsky, 2013). Therefore, even if thermometers with validated accuracy are not placed appropriately, this can result in anomalous readings.

The second basic factor to improve the accuracy of clinical thermometers is appropriate training and technique. The importance of appropriate training and technique when using clinical thermometers cannot be overstated. The accuracy of TM and TAT thermometers has been described as technique dependent (Amoeteng et al., 1999; Bahr et al., 2010; Gilbert et al., 2002; Ostrowsky et al., 2003).

An anecdotal experience by I had highlights the issues of new thermometers, training and technique. During a medical appointment, a nursing assistant obtained vital signs including a temperature. The device he used was a noncontact infrared thermometer and which he aimed it at the carotid artery. When asked about this new thermometer and technique (aiming at the carotid artery), the nursing assistant relayed that the thermometers were new, sent out for use only recently. No training or user manuals were available, as the device was "self-explanatory." Given my interest in the topic, a picture of the device was obtained with the goal of adding to my knowledge and to the EBP project.

After several internet searches, I verified that the device was a NT temporal artery thermometer. The company representative provided the user manual, which specified that the correct technique was to aim the thermometer at the forehead. The representative also

specified that temperature accuracy could not be guaranteed if the device was not used to its specifications.

Strength and Limitations of the Project

Strengths

There are several strengths that can be identified with this project. First, the topic was tied directly to an organizational priority for early sepsis identification at the former practicum site. Because this was a system-wide priority, the project also had the support of chief nursing officers and nursing leaders within the organization. Additionally, the inclusion of the EBP model adopted by the organization, the Iowa model, supported the identification of both a practice- and knowledge-focused problem. The integrative review also fits well with the Iowa model, as meets the need for the analysis and synthesis of current literature in provide an opportunity make an informed decision. Finally, I would describe my own interest in the project topic as a strength and a limitation. My personal interest in the topic is a strength because I have been using the available research to help guide practice with new ICU nurses and in the critical care environment.

Limitations

My personal knowledge and interest in the topic of clinical thermometry is also a limitation. My experience and personal bias related to different devices might have led to bias in this review. Another limitation is that the IR was limited to the body of literature specific to adult hospitalized patients, so the findings may not be translated to other populations.

Recommendations for Future Research

Given that this IR review was specific to adult hospitalized patients, one potential project would include an IR on accuracy and reliability of different clinical thermometers in children. Additionally, an IR on the literature specific to the geriatric population may also be of value as they are the fastest growing population and have age-related physiologic changes affect thermoregulation (Norman, 2000). While research on the geriatric population was included in this review, discussion about factors important to this population were not within the scope of the paper.

Section 5: Dissemination Plan

Executive Summary

Introduction

The assessment of temperature as a marker for illness has been identified in the literature as early as 1592 (Pearce, 2002). In 1871, Wunderlich wrote, "a knowledge of the course of the temperature in disease is highly important to the medical practitioner, and, indeed, indispensable" (p. vi.). The evolution of the clinical thermometer has advanced significantly since the first crude device developed by Galileo in 1592 (Pearce, 2002). In 1866, Allbutt reduced the size of the mercury thermometer from 12 inches to six inches in length (Pearce, 2002). The smaller, more portable mercury thermometer led to the advent of routine temperature assessment in clinical practice (Pearce, 2002). The mercury thermometer was the gold standard for routine temperature assessment until medical and environmental concerns related to the mercury pushed the development of various electronic, digital, and infrared clinical thermometers (Davie & Amoore, 2010).

Currently, there is a wide range of clinical thermometers used for hospitalized adult patients. Reusable gallium-in-glass thermometers replaced the mercury thermometers and are still in use (Lefrant et al., 2003; Rubia-Rubia et al., 2011; Smith, 2003). Digital and infrared thermometers include O, R, TM, AX, TAT, and NT. More invasive devices, such as the ES, BL, and PAC thermometers, may be used in the critical care or perioperative areas (Giuliano et al., 2000; Khan et al., 2006; Lawson et al., 2007; Winslow et al., 2012). Factors which can impact the accuracy or reliability of thermometers include (a) device characteristics and configuration, (b) patient

characteristics and physiology, (c) operator technique, and (d) calibration and maintenance (Davie and Amoore, 2010).

The accuracy of temperature assessment is a critical factor in the early identification and treatment of infection or sepsis (Dellinger et al., 2012). Significantly, inaccuracy in temperature assessment can lead to missed opportunities for early sepsis identification, which was identified as a system-wide organizational priority at the hospital study site. Clinical thermometers are chosen for their novelty, convenience, rapidity, and lack of invasiveness for the patient, often without the knowledge of differences in accuracy (Dunleavy, 2010; Ostrowsky, Ober, Wenze, & Edmond, 2003).

In 2001, the IOM described a healthcare system that was limited in its ability to "translate knowledge into practice, and apply new technology safely and appropriately" (pp. 2–3). Although there is a large body of research on the accuracy and reliability of many thermometers, nurses were not knowledgeable about the thermometers they are using in their environment. This IR synthesized the body of research into one document and provides EBP recommendations which can be used by nurses and organizational leaders.

Project Purpose

Temperature assessment is integral to the care of all hospitalized adult patients. Imprecise temperature measurements may lead to unrecognized infection, increased morbidity and mortality, and increased health care costs (Dellinger et al., 2012; Hall, Williams, DeFrances, & Golosinskiy, 2011). At the practicum site, when staff nurses were asked to clarify why they chose a specific thermometer or route, they were unable to

specify any related evidence to support their practice. These practices demonstrated a knowledge-to-practice gap related to the use of clinical thermometers used on adult hospitalized patients.

The purpose of this project was to conduct an IR of the body of research related to the accuracy and reliability of clinical thermometers. The IR will provide EBP information to help narrow the knowledge-to-practice gap identified in the clinical environment. Also, a synthesis of the evidence made available to organization leaders may facilitate decision-making regarding which thermometers are best for early sepsis identification.

Methodology

The methodology selected for this project was an IR review of the literature. The IR is the broadest type of research review and can incorporate both experimental and nonexperimental designs (Whittemore & Knafl, 2005). The inclusion of different types of research can lead to a more robust understanding of the clinical question (Whittemore & Knafl, 2005).

The framework for this review was the methodology proposed by Whittemore and Knafl (2005). The five stages of the IR are (a) problem identification, (b) literature search, (c) data evaluation, (d) data analysis, and (e) presentation of the findings (Whittemore & Knafl, 2005). The literature was also categorized using the Hierarchy of Evidence for Interventional Studies described by Fineout-Overholt et al. (2010). Finally, the quality of the research was evaluated using the quality indicators and quality grade described by Hooper and Andrews (2006).

Sources of evidence. The four databases that were queried for the literature were CINHAL & MEDLINE Simultaneous Search, ProQuest Nursing & Allied Health Source, the Cochrane Database of Systematic Reviews, and PubMed. The inclusion dates for this review were articles from January 1999 to December 2015. The inclusion criteria were as follows: Peer reviewed journals, English language, human studies, critical care or intensive care unit, perioperative, ED, inpatient, and adult (19+). The exclusion criteria were as follows: Ambulatory settings, outpatient settings, prototype experimental studies, intraoperative earphone type TM thermometer, exercise related studies, healthy volunteers, and pediatrics (less than 19). The following keywords were used in each database: *temperature assessment, temperature assessment AND methods, body temperature determination, body temperature determination AND methods, thermometry,* and *thermometry AND methods AND comparison*.

The inclusion of multiple clinical areas in this review was specific to the population identified in the clinical practice question as a "hospitalized adult patient." While the ED may be considered an outpatient treatment area, it is also a significant source (high volume, high risk) for screening patients for infection or sepsis (Rajee & Sultana, 2006; Singler et al., 2013; Varney et al., 2002). Lastly, the review did not include children, as their physiologic differences limit the generalizability of the findings to adult populations.

The search results yielded 2,643 papers, and the abstracts were reviewed for their relevance to the clinical question. The initial abstract review resulted in the selection of 85 papers. A secondary review resulted in the exclusion of 38 papers. Articles were

excluded if the sample included children, outpatients, or healthy volunteers. Further exclusions included pilot studies, experimental devices, and studies using the earphone-type TM thermometers (not comparable to the other TM devices).

Findings

There were 47 articles which met the inclusion criteria for this review. Based on the Hierarchy of Evidence described by Fineout-Overholt et al. (2010), 46 of the 47 studies were categorized as level IV studies. There was one SR, with level I evidence, which met the inclusion criteria for this review. The quality of the research was evaluated using the quality indicators described by Hooper and Andrews (2006). Of the 46 level IV studies, two were determined to have a quality grade of A, while there were 20 with a grade of B, and 18 with a grade of C. The lowest quality grade, D, was assessed for five studies, as they were found to have less than two quality indicators. Most of the populations were from the ICU (21), followed by medical/surgical wards (9), perioperative patients (6), the ED (4), and oncology (3). Several articles had combined populations from more than one area such as the ICU and medical/surgical wards (2), ICU and the ED (1), and the ICU and perioperative patients (1).

Level I evidence: Systematic review. Jefferies, Weatherall, Young, and Beasley (2011) conducted a SR to evaluate the accuracy of peripheral thermometers in the detection of fever (> 37.5°C) in critically ill patients. While only three studies met the inclusion criteria of their SR, data were evaluated on seven TM thermometers (different brands and models), both O (digital) and R (digital). Five of the TM thermometers and the O thermometer were within ±0.2ºC of the PAC. The mean difference of the R to PAC

was outside the acceptable criterion. The authors concluded that the TM and O devices provided accurate temperature readings on febrile patients; the R device was not recommended.

Level IV evidence: Cohort studies.

Tympanic membrane (TM) thermometer. There were 27 studies which included an evaluation of the TM thermometer. A number of reference temperatures sites were used for comparison including the PAC, ES, BL, O (digital or mercury), and R (digital or mercury). The findings from these studies were inconsistent. Seventeen of the studies had findings which did not support the use of the TM thermometer, while 10 others found the device to be accurate enough for clinical use.

Clinically unacceptable. The TM thermometer was found to have unacceptably wide bias or variability in 17 studies and was not recommended for clinical use (Dunleavy, 2010; Farnell et al., 2005; Fountain et al., 2008; Frommelt et al., 2008; Gilbert et al., 2002; Giuliano et al., 1999; Giuliano et al., 2000; Irvin, 1999; Jensen et al., 2000; Khan, Vohra, Paul, Rosin, & Patel, 2006; Langham et al., 2009; Lawson et al., 2007; Moran et al., 2007; Nonose et al., 2012; Prentice & Moreland, 1999; Rubia-Rubia et al., 2011; Varney et al., 2002). A priori clinically acceptable differences were described in 10 studies and ranged from ± 0.2ºC to ± 0.5ºC or ± 0.2ºF to ± 0.5ºF. Farnell et al. (2005), using an expert panel, defined clinical acceptability by determining if the inaccuracy would result in a delay in care or unnecessary interventions. Clinically acceptable differences were not specified in seven studies (Dunleavy, 2010; Fountain et al., 2008; Frommelt et al., 2008; Irvin, 1999; Khan et al., 2006; Moran et al., 2007;

Nonose et al., 2012). When the quality of these studies was evaluated, seven were found to have an acceptable quality grade of A or B (Gilbert et al., 2002; Giuliano et al., 1999; Giuliano et al., 2000; Jensen et al., 2000; Lawson et al., 2007; Moran et al., 2007; Rubia-Rubia et al., 2011).

Clinically acceptable. In contrast, authors of 10 other studies concluded that the TM thermometer was clinically acceptable for use (Amatoeng et al., 1999; Hasper et al., 2011; Haugan et al., 2012; Leon et al., 2005; Rajee & Sultana, 2006; Rubia-Rubia et al., 2011; Singler et al., 2013; Smitz et al., 2000; Smitz et al., 2008; Spitzer et al., 2008). Of note, only four studies had predefined clinically acceptable differences, range ± 0.25ºC to 0.5ºC and ± 0.5ºF (Amatoeng et al., 1999; Haugan et al., 2012; Rajee & Sultana, 2006; Rubia-Rubia et al., 2011). The quality of these studies was also variable, with half receiving an acceptable quality grade of A or B (Amatoeng et al., 1999; Haugan et al., 2012; Leon et al., 2005; Rubio-Rubio et al., 2011; Singler et al., 2013).

Diagnostic accuracy. There were eight studies which included an evaluation of the diagnostic accuracy of the TM thermometer to detect fever or infection (Leon et al., 2005; Prentice & Moreland, 1999; Rajee and Sultana, 2006; Rubia-Rubia et al., 2011; Singler et al., 2013; Smitz et al., 2000; Smitz et al., 2009; Varney et al., 2002). Fever was defined in seven studies and ranged from 37.5ºC to 39ºC; one group used an adjudicated diagnosis of infection (Singler et al., 2013). The conclusions regarding the diagnostic accuracy of the TM thermometer were also mixed.

Unacceptable diagnostic accuracy. When sensitivity to predict fever was evaluated (Leon et al., 2005; Rajee & Sultana, 2006; Singler et al., 2013), the values

ranged from 51% to 73%. Out of those patients with the adjudicated diagnosis of infection (*n* = 105), 22.8% had TM fevers >38.ºC whereas R fever was present in 43% (Singler et al., 2013). The last group (Varney et al., 2002) found that of 19 occurrences of R fever, that were afebrile by TM or O, 68% (*n* = 13) required admission.

Acceptable diagnostic accuracy. Rubia-Rubia et al. (2011), Smitz et al. (2000), and Smitz et al. (2009) concluded that the TM could provide acceptable diagnostic accuracy for fever. The sensitivity for fever ranged from 86% to 98%. Of note, Smitz et al. concluded that although the TM thermometer can predict R fever, the predictive accuracy of the TM thermometer was dependent upon operator technique and the quality of the equipment.

Disposable chemical (CH) thermometers. Eight papers evaluated the accuracy and reliability of the CH dot thermometer (such as TempaDot ™ or NexTemp™) for use in other clinical areas. The findings for use of the CH thermometer were mixed. Additionally, half of the authors recommended limited or cautionary use, so their findings are described separately.

Clinically unacceptable. Counts et al. (2014) and Fallis et al. (2006) compared the CH to the O thermometer and found the CH was not acceptable for clinical use. A priori clinically acceptable differences were described in both studies (± 0.3ºC). Counts et al. reported 21% (*n* = 10) differences > ± 0.5ºC and 13% (*n* = 6) differences ≥ ± 1.0ºC. Fallis et al. found that 91% of CH readings either overestimated (81%) or underestimated (10%) the O temperature. Additionally, Fallis et al. reported 65 instances where the CH

thermometers failed to demonstrate any color change. In evaluating the quality of these studies, both were found to have a quality grade of B.

Clinically acceptable. Rubia-Rubia et al. (2011) used the PAC as the reference and Rajee and Sultana (2006) used the R as the reference. Clinical acceptability was defined in both studies (± 0.2ºC). Both found a nonsignificant bias for both agreement and repeatability between the CH thermometer and the reference temperature. The quality evaluation of these two studies were mixed; the study by Rubia-Rubia et al. received a B, while the study by Rajee and Sultana received a C.

Cautionary or limited use. Farnell et al. (2005), Fountain et al. (2008), and Frommelt et al. (2008) found the CH to have a narrow bias while Potter et al. (2003) determined acceptability with a strong correlation. However, there were sufficient readings which were ≥ 1.0º (F or C) to cause concern for use in the clinical environment. Fountain et al. noted that 30% (n = 18) differed from the O temperature by 1.0 ºF and 3% (n = 2) differed by 2.0 ºF. Potter et al. noted that 25% of the CH temperatures measured either overestimated (11.8%) or underestimated (10.8%) body temperature by 0.4ºC or more. CH thermometer may be useful for screening or isolation, but abnormal findings should be validated with another thermometer. All four studies in this group received a quality grade of C.

Temporal artery thermometer (TAT). There were 19 studies which included a review of the TAT thermometer. Reference sites for comparison included the PAC, BL, O, and TM. The conclusions regarding this device are also varied.

Clinically unacceptable. Authors from 12 studies concluded that the TAT device was not accurate enough for clinical use (Bodkin et al., 2104; Counts et al., 2014; Fetzer et al. 2008; Fountain et al., 2008; Frommelt et al., 2008; Kimberger et al., 2007; Marable et al., 2009; Rubia-Rubia et al., 2011; Singler et al., 2013; Stelfox et al., 2010; Winslow et al., 2012; Wolsfon et al., 2013). Clinically acceptable differences were established in 10 studies, with a range of ± 0.2ºC to 1.0ºC and ± 0.5ºF to 0.6ºF. Fountain et al. (2008) and Frommelt et al. (2008) did not define clinically acceptable differences. Significantly, the TAT was found to have a lack of agreement especially in the hypo- and hyperthermic ranges (Kimberger et al., 2007; Stelfox et al., 2010; Winslow et al., 2012).

No-Touch (NT) infrared forehead thermometer. Only two studies which evaluated the NT met the inclusion criteria for this review (Duncan et al., 2008; Woodrow et al., 2006). Duncan et al. (2008) found the NT to be reliable (between NT readings, $r = 0.94$), but the agreement was poor. Woodrow et al. (2006) found the NT to have an acceptable agreement, however they reported that the *t* test comparisons (TAT to TM oral and TAT to TM core) were statistically and significantly greater ($t = 7.038$; $p < 0.001$ and $t = -6.736$; $p < 0.001$). The conclusion by both groups of authors was that this thermometer was not accurate enough for use in acutely ill adults. Given the paucity of data on accuracy and reliability, this thermometer is not recommended for use.

Axillary (AX) thermometry. There were 13 studies which included an evaluation of the accuracy of AX temperatures. A variety of reference temperatures were used including PAC, BL, O, and R. As with the other clinical thermometers, the findings were

inconsistent. However, most authors concluded the AX route should not be used for routine clinical assessments.

Clinically unacceptable. The findings from 10 studies demonstrated that AX thermometry is clinically inaccurate for use in acutely ill adults (Barringer et al., 2011; Jensen et al., 2000; Khan et al., 2006; Langham et al., 2009; Lawson et al., 2007; LeFrant et al., 2003; Marable et al., 2009; Mason et al., 2015; Moran et al, 2007; Myny et al., 2005). Clinically acceptable differences ranged from 0.3ºC to 0.5ºC and 0.2ºF to 0.5ºF. Lefrant et al. (2003), Moran et al. (2007), Khan et al. (2006), and Barringer et al. (2011) did not report clinically acceptable differences in their studies. The quality of the research was variable. Five studies were found to have an acceptable quality grade of A or B (Jensen et al., 2000; Lawson et al., 2007; Marable et al., 2009; Moran et al., 2007; Myny et al., 2005).

Clinically acceptable. The conclusions reached in three studies was that the AX was accurate for use in acutely ill adults. Reference temperatures included the PAC and BL (Nonose et al., 2012) and the PAC alone (Rubia-Rubia et al., 2011) while Smith (2003) compared AX (gallium-in-glass) to AX (mercury). Only one group identified a clinically acceptable difference of 0.2ºC (Rubia-Rubia et al., 2011). Nonose et al. (2012) and Smith determined acceptability with correlational statistics ($r = 0.64$ and $r = 0.886$), respectively. Rubia-Rubia et al. (2011) created a value score (which included accuracy, ease of use, cost, speed, and durability) and determined the AX to have the strongest results. Significantly, the authors noted the required dwell time for this device was 12 minutes, a significant limitation given the time constraints for nurses. The quality of

these three studies was also variable, Rubia-Rubia et al. and Smith had a quality grade of B, while the study by Nonose et al. had a quality grade of C.

Recommendations

The review of the literature did not reveal a clinical thermometer with better accuracy and reliability than any other. There were inconsistent findings for all the thermometers, except the NT device. Given the lack of evidence supporting any one thermometer, the recommendations will address each device specifically.

Tympanic (TM) thermometer. One systematic review (Jefferies et al., 2011) found that five of seven models of TM thermometer were accurate. The accuracy associated with different models was also described by other researchers (Giuliano et al., 2000; Haugan et al., 2012; Smitz et al., 2009). Another factor associated with the variability and accuracy of the TM thermometer is technique (Amoeteng et al., 1999; Gilbert et al., 2002; Giuliano et al., 1999; Giuliano et al., 2000; Smitz et al., 2009). Where the TM thermometer is in use and to maximize clinical accuracy, ensure adequate training and emphasize the importance of using the correct technique. Additionally, organizational leaders should consider annual skills validation for this device. Organization leaders considering the purchase of the TM thermometer should consider evidence which specifies a superior model (described above).

Disposable chemical (CH) thermometers. The CH thermometer was found to be accurate, but highly variable. Use of the CH thermometer should be limited for specific clinical situations (such as isolation). When the patient's condition requires the use of this

thermometer, any abnormal temperatures should be validated by another, more reliable, thermometer.

Temporal artery thermometers (TAT). Accuracy of the TAT thermometer was limited to patients who are normothermic. Given the wide variety of clinical areas where hyperthermia or hypothermia are of concern, this is a significant device limitation. The use of this thermometer for acutely ill adults is not recommended.

NT forehead thermometers. Only two studies met the inclusion criteria for this review. The paucity of data on the NT thermometer also creates limitations for use. More research is needed on the NT thermometer before it can be recommended for use.

Axillary thermometry. AX thermometry is often used as an alternate site for patients who are combative or confused, or when the oral site cannot be used. However, the research does not support the use of the AX route for routine patient temperature assessments. When the oral route is contraindicated, an alternative noninvasive device should be used (Sund-Levander & Grodzinsky, 2013).

Implications for Nursing Practice

It is clear that the technological advances in clinical thermometers will continue. These advances may outpace the ability of researchers to validate devices for accuracy and reliability. However, is it also clear that faster more noninvasive thermometers do not necessarily equate to better clinical results. As new devices are developed, time and care should be used to ensure they are used appropriately.

An anecdotal experience I had highlights the problem with the rapid employment of a clinical thermometer without adequate training. During a medical

appointment, a nursing assistant obtained vital signs including a temperature. The device he used was a noncontact infrared thermometer and which he aimed it at the carotid artery. When asked about this new thermometer and technique (aiming at the carotid artery), the nursing assistant relayed that the thermometers were new, sent out for use only recently. No training or user manuals were available, as the device was "self-explanatory." Given my interest in the topic, a picture of the device was obtained with the goal of adding to my knowledge and to the EBP project.

After several internet searches, I verified that the device was a NT temporal artery thermometer. The company representative provided the user manual, which specified that the correct technique was to aim the thermometer at the forehead. The representative also specified that temperature accuracy could not be guaranteed if the device was not used to its specifications.

Analysis of Self

Practitioner

The pursuit of this degree has been one of the most challenging and personally satisfying endeavors of this chapter of my career. When I began my DNP journey, I considered myself to be an expert nurse and clinician. However, after beginning the coursework for my DNP, I realized that my level of professional development was actually quite narrow.

According to Zaccagnini and White (2011), the DNP-prepared nurse is effective in "(1) translating research into practice, (2) quality improvement and patient-centered care, (3) evaluation of practice, (4) research methods and technology, (5) participation in

collaborative research, and (6) disseminating findings" (p. 68). While I would not consider myself an expert in each of these areas, I have experienced tremendous personal and professional growth throughout the program. In particular, and with this project, I developed in my ability to evaluate a body of research and synthesize the information into a scholarly product that can be used by other nurses.

Scholar

I find this quote from Boyer (1992) to be one of my favorites "we need to relate theory and research to the realities of life" (p. 90). I think it applies well to the current clinical environment. The nurses I trained and worked with wanted information or evidence on current practices. It is difficult for nurses to value research that is not applicable to their practice.

Boyer's (1992) redefinition of scholarship to include not only discovery but also integration, application and teaching was also a powerful message to me. Before beginning by DNP journey, my perspective of scholarship was that it was about research (discovery). As an ICU nurse and an educator, I can see myself in each of the roles described by Boyer. I feel even more capable now, as I complete my DNP project and finish my degree.

Project Completion

Zaccagnini and White (2011) described the importance of analyzing and understanding a clinical issue within the boundaries of a system. Leaders must understand "the structure within the system" as well as "patterns of behavior" in order to identify the best way to affect any change (Zaccagnini & White, 2011, p. 43). The

description of organizational and systems analysis is consistent with the second essential characteristics of the DNP-prepared nurse (AACN, 2006). Throughout this DNP journey and through the development of my project, one of my greatest challenges was my limited exposure to the nuances of the civilian healthcare system. As a career military nurse, I was able to operate and affect change within the military healthcare system. However, knowledge of the military system did not translate well to a civilian, for-profit organization.

As a project manager, one of the most valuable lessons I learned through my practicum experiences was the critical importance of understanding the system. Although I had many years of clinical experience, I was at a disadvantage in effecting changes in my early projects because I did not understand the system. Similarly, as my DNP project progressed, there were system issues (at the practicum site) which required a number of changes to my project.

Most importantly, I have learned that each step in the project was a new learning opportunity. While I found some of the changes frustrating, there was new insight to be gained in looking at the project in a new light. Ultimately, with the guidance of my committee chair and a great deal of hard work, I think I have developed an excellent product. I believe my DNP project will add to the body of knowledge regarding clinical thermometers and can serve as a resource for clinical nurses and organizational leaders.

Summary

Temperature accuracy is relevant in all areas of nursing practice. In the hospitalized adult patient, temperature changes can signal early indicators of infection or

of sepsis (Dellinger et al., 2012). Inaccuracy in temperature assessment, either because of poor operator technique or device limitations, can result in missed opportunities to identify and treat infection early. These missed opportunities can result in increased morbidity and mortality, hospital length of stay, and increased health care costs (Dellinger et al., 2012).

Often, clinical thermometers are chosen for use because they give rapid results and the noninvasive nature of the device. However, these factors do not necessarily equate to improved accuracy. This IR review provided a synthesis of the current evidence on the accuracy of clinical thermometers. The synthesis of the literature is a resource for clinical nurses and helps to bridge the knowledge-to-practice gap observed in the clinical environment. In addition, the IR can serve as a resource for organizational leaders who may be considering the purchase of new clinical thermometers.

References

Abolnik, I. Z., Kithas, P. A., McDonnald, J. J., Soller, J. B., Izrailevsky, Y. A., & Granger, D. L. (1999). Comparison of oral and tympanic temperatures in a Veterans Administration outpatient clinic. *American Journal of the Medical Sciences*, *317*(5), 301. doi:10.1016/S0002-9629(15)40532-4

Agency for Healthcare Research and Quality. (2013). *How patient and family engagement benefits your hospital.* Retrieved from http://www.ahrq.gov/professionals/systems/hospital/engagingfamilies/index.html

Ahmadnia, H., Mojahedi, M. J., Dalooee, M. K., & Ghanbarizadeh, S. R. (2010). A comparison between urinary bladder temperature and rectal, axillary and oral temperatures following kidney transplantation. *Saudi Journal of Kidney Diseases and Transplantation*, *21*(6), 1135. Retrieved from http://www.sjkdt.org/

American Academy of Pediatrics. (2007). *Fever and your child*. Retrieved from http://patiented.aap.org/content2.aspx?aid=5107

American Association of Colleges of Nursing. (2006). *The essentials of doctoral education for advanced nursing practice.* Retrieved from http://www.aacn.nche.edu/publications/position/DNPEssentials.pdf

Amoateng-Adjepong, Y., Del Mundo, J., & Manthous, C. A. (1999). Accuracy of an infrared tympanic thermometer. *Chest Journal*, *115*(4), 1002–1005. doi:10.1378/chest.115.4.1002

Arslan, G. G., Eser, I., & Khorshid, L. (2011). Analysis of the effect of lying on the ear on body temperature measurement using a tympanic thermometer. *Journal of*

Pakistan Medical Association, 61, 1065–1068. Retrieved from http://www.jpma.org.pk/

Bahr, S. J., Senica, A., Gringras, L., & Ryan, P. (2010). Clinical Nurse Specialist-led evaluation of temporal artery thermometers in acute care. *Clinical Nurse Specialist, 24*(5), 238–244. doi:10.1097/NUR.0b013e3181ecda4b

Barringer, L. B., Evans, C. W., Ingram, L. L., Tisdale, P. P., Watson, S. P., & Janken, J. K. (2011). Agreement between temporal artery, oral and axillary temperature measurements in the perioperative period. *Journal of PeriAnesthesia Nursing, 26*(3), 143–150. doi:10.1016/j.jopan.2011.03.010

Birriel, B. (2013). *Rapid identification of sepsis - The value of screening tools*. Retrieved from http://www.sccm.org/Communications/Critical-Connections/Archives/Pages/Rapid-Identification-of-Sepsis---The-Value-of-Screening-Tools.aspx

Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *327*(8476), 307–310. Retrieved from http://www.medicine.mcgill.ca/epidemiology/hanley/tmp/Surveys/ba.pdf

Bland, J. M., & Altman, D. G. (2010). Statistical methods for assessing agreement between two methods of clinical measurement. *International Journal of Nursing Studies*, *47*(8), 931–936. doi:10.1016/j.ijnurstu.2009.10.001

Bock, M., Hohlfeld, U., Von Engeln, K., Meier, P. A., Motsch, J., & Tasman, A. J. (2005). The accuracy of a new infrared ear thermometer in patients undergoing

cardiac surgery. *Canadian Journal of Anesthesia*, *52*(10), 1083–1087. doi:10.1007/BF03021609

Bodkin, R. P., Acquisto, N. M., Zwart, J. M., & Toussaint, S. P. (2014). Differences in noninvasive thermometer measurements in the adult emergency department. *American Journal of Emergency Medicine*, *32*(9), 987–989. doi:10.1016/j.ajem.2014.05.036

Brown, C. G. (2014). The Iowa model of evidence–based practice to promote quality care: An illustrated example in oncology nursing. *Clinical Journal of Oncology Nursing*, *18*(2), 157–159. doi:10.1188/14.CJON.157-159

Brown-Mahoney, C. & Odom, J. (1999). Maintaining intraoperative normothermia: A meta-analysis of outcomes with costs. *American Association of Nurse Anesthetists Journal, 67*(2), 155–164. Retrieved from https://www.aana.com/newsandjournal/Documents/maintainingintraoperative_0499_p155–164.pdf

Browne, G. J., Currow, K., & Rainbow, J. (2001). Practical approach to the febrile child in the emergency department. *Emergency Medicine, 13*, 426–435. Retrieved from http://intranet.emergency.med.ufl.edu/med_students/reading_assignment/40_practical%20approach%20to%20the%20febrile%20child%20in%20the%20emergency%20dept.pdf

Calonder, E. M., Sendelbach, S., Hodges, J. S., Gustafson, C., Machemer, C., Johnson, D., & Reiland, L. (2010). Temperature measurement in patients undergoing colorectal surgery and gynecology surgery: A comparison of esophageal core,

temporal artery, and oral methods. *Journal of PeriAnesthesia Nursing*, *25*(2), 71–78. doi:10.1016/j.jopan.2010.01.006

Camboni, D., Philipp, A., Schebesch, K. M., & Schmid, C. (2008). Accuracy of core temperature measurement in deep hypothermic circulatory arrest. *Interactive Cardiovascular and Thoracic Surgery, 7*, 922–924. doi:10.1510/icvts.2008.181974

Centers for Disease Control and Prevention. (2010). *National Hospital Ambulatory Medical Care Survey: 2010 emergency department summary tables.* Retrieved from http://www.cdc.gov/nchs/data/ahcd/nhamcs_emergency/2010_ed_web_tables.pdf

Centers for Medicare and Medicaid. (n.d.). *About CMS*. Retrieved from https://www.cms.gov/about-cms/about-cms.html

Centers for Medicare and Medicaid. (2012). *Frequently asked questions hospital value-based purchasing program.* Retrieved from http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/Downloads/FY-2013-Program-Frequently-Asked-Questions-about-Hospital-VBP-3-9-12.pdf

Chan, E. Y., Lee, Y. K., Poh, T. H., Ng, I. L., & Prabhakaran, L. (2011). Translating evidence into nursing practice: Oral hygiene for care dependent adults. *International Journal of Evidence-Based Healthcare*, *9*(2), 172–183. doi:10.1111/j.1744-1609.2011.00214.x

Chase, M., Klasco, R. S., Joyce, N. R., Donnino, M. W., Wolfe, R. E., & Shapiro, N. I. (2012). Predictors of bacteremia in emergency department patients with suspected infection. *American Journal of Emergency Medicine, 30*, 1691–1697. doi:10.1016/j.ajem.2012.01.018

Counts, D., Acosta, M., Holbrook, H., Foos, E., Hays-Ponder, K., Macairan, O., Thomas, L.,… Twiss, E. J. (2014). Evaluation of temporal artery and disposable digital oral thermometers in acutely ill patients. *Medsurg Nursing*, *23*(4), 239. Retrieved from http://www.medsurgnursing.net/cgi-bin/WebObjects/MSNJournal.woa

Creagh-Brown, B., Armstrong, D., & Jackson, S. (2005). The use of the TempaDot thermometer in routine clinical practice. *Ageing, 3*, 297–299. doi:10.1093/ageing/afi069

Cronin, K., & Wallis, M. (2000). Temperature taking in the ICU: Which route is best? *Australian Critical Care*, *13*(2), 59–64. doi:10.1016/S1036-7314(00)70623-0

Davie, A., & Amoore, J. (2010). Best practice in the measurement of body temperature. *Nursing Standard, 24*(42), 42–49. doi:10.7748/ns2010.06.24.42.42.c7850

Dellinger, R. P., Levy, M. M., Rhodes, A., Annane, D., Gerlach, H., Opal, S. M., Sevransky, J. E.,... Moreno, R. (2012). Surviving sepsis campaign: International guidelines for management of severe sepsis and septic shock, 2012. *Intensive Care Medicine*, *39*(2), 165–228. doi:10.1097/CCM.0b013e31827e83af

Doody, C. M., & Doody, O. (2011). Introducing evidence into nursing practice: Using the IOWA model. *British Journal of Nursing, 20*(11), 661–664. Retrieved from http://ulir.ul.ie/bitstream/handle/10344/1801/Doody.pdf?sequ

Dowding, D., Freeman, S., Nimmo, S., Smith, D., & Wisniewski, M. (2002). An investigation into the accuracy of different types of thermometers. *Professional Nurse (London, England)*, *18*(3), 166–168. Retrieved from http://www.nursingtimes.net/

Duncan, A. L., Bell, A. J., Chu, K., & Greenslade, J. H. (2008). Can a non-contact infrared thermometer be used interchangeably with other thermometers in an adult emergency department? *Australasian Emergency Nursing Journal*, *11*(3), 130–134. doi:10.1016/j.aenj.2008.05.003

Dunleavy, K. (2010). Which core body temperature method is most accurate? *Nursing 2010, 40*(12), 18–19. doi:10.1097/01.NURSE.0000390678.95642.7f

Dzarr, A. A., Kamal, M., & Baba, A. A. (2009). A comparison of infrared tympanic thermometry, oral and axilla with rectal thermometry in neutropenic patients. *European Journal of Oncology Nursing, 13*, 250–254. doi:10.1016/j.ejon.2009.03.006

Environmental Protection Agency. (2015). Mercury thermometers. Retrieved from http://www.epa.gov/mercury/mercury-thermometers

Fallis, W. M. (2005). The effect of urine flow rate on urinary bladder temperature in critically ill adults. *Heart & Lung: The Journal of Acute and Critical Care*, *34*(3), 209–216. doi:10.1016/j.hrtlng.2004.10.001

Fallis, W. M., Hamelin, K., Wang, X., & Symonds, J. (2006). A multimethod approach to evaluate chemical dot thermometers for oral temperature measurement. *Journal of Nursing Measurement*, *14*(3), 151–162. Retrieved from http://www.springerpub.com/journal-of-nursing-measurement.html

Farnell, S., Maxwell, L., Tan, S., Rhodes, A., & Philips, B. (2005). Temperature measurement: Comparison of non-invasive methods used in adult critical care. *Journal of Clinical Nursing*, *14*(5), 632–639. doi:10.1111/j.1365-2702.2004.00916.x

Fetzer, S. J., & Lawrence, A. (2008). Tympanic membrane versus temporal artery temperatures of adult perianesthesia patients. *Journal of PeriAnesthesia Nursing*, *23*(4), 230-236. doi:10.1016/j.jopan.2008.01.010

Fineout-Overholt, E., Melnyk, B. M., Stillwell, S. B., & Williamson, K. M. (2010). Evidence-based practice, step by step: Critical appraisal of the evidence Part III. *American Journal of Nursing, 110*(11), 43–51. doi:10.1097/01.NAJ.0000390523.99066.b5

Flynn-Makic, M. B., VonRueden, K. T., Rauen, C. A., & Chadwick, J. (2011). Evidence-based practice habits: Putting more sacred cows out to pasture. *Critical Care Nurse, 31*(2). 38–61. doi:10.4037/ccn2011908

Fountain, C., Goins, L., Hartman, M., Phelps., Scoles, D., & Hays, V. (2008). Evaluating the accuracy of four temperature instruments on an adult inpatient oncology unit. *Clinical Journal of Oncology Nursing, 12*(6), 983–987. doi:10.1188/08.CJON.983-987

Frommelt, T., Ott, C., & Hays, V. (2008). Accuracy of different devices to measure temperature. *Medsurg Nursing, 17*(3), 171–182. Retrieved from http://www.technologyinstructor.us/nursing1/journals/Journal%20Article%20-%20Frommelt,%20Accuracy%20of%20diff%20devices%20to%20measure%20temperature.pdf

Gallimore, D. (2004). Reviewing the effectiveness of tympanic thermometers. *Nursing Times, 100*(32), 32–34. Retrieved from www.lexisnexis.com/hottopics/lnacademic

Gasim, G. I., Musa, I. R., Abdien, M. T., & Adam, I. (2013). Accuracy of tympanic temperature measurement using an infrared tympanic membrane thermometer. *BMC Research Notes*, *6*(1), 194. doi:10.1186/1756-0500-6-194

Giantin, V., Toffanello, E. D., Enzi, G., Perissinotto, E., Vangelista, S., Simonato, M., Ceccato, C.,... & Sergi, G. (2008). Reliability of body temperature measurements in hospitalised older patients. *Journal of Clinical Nursing*, *17*(11), 1518–1525. doi:10.1111/j.1365-2702.2007.02140.x

Gilbert, M., Barton, A. J., & Counsell, C. M. (2002). Comparison of oral and tympanic temperatures in adult surgical patients. *Applied Nursing Research*, *15*(1), 42–47. doi:10.1053/apnr.2002.29528

Giuliano, K. K., Giuliano, A. J., Scott, S. S., MacLachlan, E., Pysznik, E., Elliot, S., & Woytowicz, D. (2000). Temperature measurement in critically ill adults: A comparison of tympanic and oral methods. *American Journal of Critical Care*, *9*(4), 254–261. Retrieved from http://ajcc.aacnjournals.org/

Giuliano, K. K., Scott, S. S., Elliot, S., & Giuliano, A. J. (1999). Temperature measurement in critically ill orally intubated adults: A comparison of pulmonary artery core, tympanic, and oral methods. *Critical Care Medicine, 27(10)*, 2188–2193. doi:10.1097/00003246-199910000-00020

Göbölös, L., Philipp, A., Ugocsai, P., Foltan, M., Thrum, A., Miskolczi, S., Pousios, D.,... & Ohri, S. K. (2013). Reliability of different body temperature measurement sites during aortic surgery. *Perfusion*, *29*(1), 75–81. doi:10.1177/0267659113497228

Goldstein, B., Giroir, B., & Randolph, A. (2005). International pediatric sepsis consensus conference: Definitions for sepsis and organ dysfunction in pediatrics. *Pediatric Critical Care Medicine, 6*, 2–8. doi:10.1097/01.PCC.0000149131.72248.E6

Grove, S. K., Burns, N., & Gray, J. R. (2013). *The practice of nursing research: Appraisal, synthesis, and generation of evidence* (7th ed.). St. Louis, MO: Saunders Elsevier

Hall, M. F. (2008). *Looking to improve financial results? Start by listening to patients*. Retrieved from http://pressganey.com/Documents/Press%20Room/1008_HFM_Hall.pdf

Hall, M. J., Williams, S. N., DeFrances, C. J., & Golosinskiy, A. (2011). Inpatient care for septicemia or sepsis: A challenge for patients and hospitals. *NCHS Data Brief* 2011: 1–8. Retrieved from http://www.cdc.gov/nchs/

Hamilton, P. A., Kasbekar, R. S., & Monro, R. (2013). Clinical performance of infrared consumer-grade thermometers. *Journal of Nursing Measurement*, *21*(2), 166–177. doi:10.1891/1061-3749.21.2.166

Harioka, T., Matsukawa, T., Ozaki, M., Nomura, K., Sone, T., Kakuyama, M., & Toda, H. (2000). “Deep-forehead” temperature correlates well with blood temperature. *Canadian Journal of Anesthesia*, *47*(10), 980–983. doi:10.1007/BF03024869

Hasper, D., Nee, J., Schefold, J. C., Krueger, A., & Storm, C. (2011). Tympanic temperature during therapeutic hypothermia. *Emergency Medicine Journal*, *28*(6), 483–485. doi:10.1136/emj.2009.090464

Haugan, B., Langerud, A. K., Kalvøy, H., Frøslie, K. F., Riise, E., & Kapstad, H. (2012). Can we trust the new generation of infrared tympanic thermometers in clinical practice? *Journal of Clinical Nursing*, *22*(5–6), 698–709. doi:10.1111/j.1365-2702.2012.04077.x

Hausfater, P., Zhao, Y., Defrenne, S., Bonnet, P., & Riou, B. (2008). Cutaneous infrared thermometry for detecting febrile patients. *Emerging Infectious Diseases*, *14*(8), 1255. doi:10.3201/eid1408.080059

Huang, J. J., & Kurz, A. (2001). Body warmer and upper extremities position affect the accuracy of cutaneous thermometers during anesthesia. *Journal of Clinical Anesthesia*, *13*(4), 241–243. doi:10.1016/S0952-8180(01)00243-4

Haxton, D., Doering, J., Gingras, L., & Kelly, L. (2012). Implementing skin-to-skin contact at birth using the Iowa model. *Nursing For Women's Health*, *16*(3), 220–230. doi:10.1111/j.1751-486X.2012.01733.x

Healthcare Improvement Scotland. (2012). What is the clinical and cost effectiveness of temporal artery thermometers compared with infrared in-ear thermometers for

temperature measurement in routine clinical practice? *Technologies Scoping Report, 11*. Retrieved from http://www.healthcareimprovementscotland.org/our_work/technologies_and_medicines/earlier_scoping_reports/technologies_scoping_report_11.aspx

Hicks, C. & Hennessy, D. (1997). Mixed messages in nursing research: Their contribution to the persisting hiatus between evidence and practice. *Journal of Advanced Nursing 25*(3): 595–601. doi:10.1046/j.1365-2648.1997.1997025595.x

Höcker, J., Bein, B., Böhm, R., Steinfath, M., Scholz, J., & Horn, E. P. (2012). Correlation, accuracy, precision and practicability of perioperative measurement of sublingual temperature in comparison with tympanic membrane temperature in awake and anaesthetised patients. *European Journal of Anaesthesiology (EJA)*, *29*(2), 70-74. doi:10.1097/EJA.0b013e32834cd6de

Hooker, E. A., & Houston, H. (1996). Screening for fever in an adult emergency department: Oral vs tympanic thermometry. *Southern Medical Journal*, *89*(2), 230–234. doi:10.1097/00007611-199602000-00016

Hooper, V. D., & Andrews, J. O. (2006). Accuracy of noninvasive core temperature measurement in acutely ill adults: The state of the science. *Biological Research for Nursing, 8*(24). doi: 10.1177/1099800406289151

Hutton, S., Probst, E., Kenyon, C., Morse, D., Friedman, B., Arnold, K., & Helsley, L. (2009). Accuracy of different temperature devices in the postpartum population. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, *38*(1), 42–49. doi:10.1111/j.1552-6909.2008.00302.x

Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academies Press

Irvin, S. M. (1999). Comparison of the oral thermometer versus the tympanic thermometer. *Clinical Nurse Specialist, 13*(2), 85–89. doi:10.1097/00002800-199903000-00011

Jefferies, S., Weatherall, M., Young, P., & Beasley, R. (2011). A systematic review of the accuracy of peripheral thermometry in estimating core temperatures among febrile critically ill patients. *Critical Care and Resuscitation, 13*(3), 194–199. Retrieved from http://www.cicm.org.au/journal/2011/september/ccr_13_3_010911-194.pdf

Jensen, B. N., Jensen, F. S., Madsen, S. N., & Lossl, K. (2000). Accuracy of digital tympanic, oral, axillar, and rectal thermometers compared with standard rectal mercury thermometers. *European Journal of Surgery, 166*(11), 848-851. Retrieved from http://link.springer.com/journal/10353

Johnson, M., Gardner, D., Kelly, K., Maas, M., & McCloskey, J. (1991). The Iowa model: A proposed model for nursing administration. *Nursing Economics*, *9*(4), 255–262. Retrieved from http://www.nursingeconomics.net/cgi-bin/WebObjects/NECJournal.woa

Khan, T. A., Vohra, H. A., Paul, S., Rosin, M. D. & Patel, R. L. (2006). Axillary and tympanic membrane temperature measurements are unreliable early after cardiopulmonary bypass. *European Journal of Anaesthesiology, 23*, 551–554. doi:lO.i017/SO265O2I5050O1936

Khorshid, L., Eşer, İ., Zaybak, A., & Yapucu, Ü. (2005). Comparing mercury-in-glass, tympanic and disposable thermometers in measuring body temperature in healthy young people. *Journal of Clinical Nursing*, *14*(4), 496–500. doi:10.1111/j.1365-2702.2004.01076.x

Kimberger, O., Cohen, D., Illievich, U., & Lenhardt, R. (2007). Temporal artery versus bladder thermometry during perioperative and intensive care unit monitoring. *Anesthesia and Analgesia, 105*(4), 1042–1047. doi:10.1213/01.ane.0000281927.88935.e0

Kimberger, O., Thell, R., Schuh, M., Koch, J., Sessler, D. I., & Kurz, A. (2009). Accuracy and precision of a novel noninvasive core thermometer. *British Journal of Anaesthesia*, 103(2), 226–231. doi:10.1093/bja/aep134

Kistemaker, J. A., Den Hartog, E. A., & Daanen, H. A. M. (2006). Reliability of an infrared forehead skin thermometer for core temperature measurements. *Journal of Medical Engineering & Technology*, *30*(4), 252–261. doi:10.1080/03091900600711381

Kowal, C. (2010). Implementing the Critical Care Pain Observation Tool using the Iowa model. *Journal of the New York State Nurses Association*, *41*(1), 4–10. Retrieved from http://www.nysna.org/nursing-practice/journal-new-york-state-nurses-association#.VyEB6vkrKM8

Krom, Z. R., Batten, J., & Bautista, C. (2010). A unique collaborative nursing evidence-based practice initiative using the Iowa model: A clinical nurse specialist, a health

science librarian, and a staff nurse's success story. *Clinical Nurse Specialist*, *24*(2), 54–59. doi:10.1097/NUR.0b013e3181cf5537

Kumar, A., Roberts, D., Wood, K. E., Light, B., Parrillo, J. E., Sharma, S., Suppes, R.,... & Gurka, D. (2006). Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock*. *Critical Care Medicine*, *34*(6), 1589–1596. doi:10.1097/01.CCM.0000217961.75225.E9

Kurz, A., Sessler, D. I., & Lenhardt, R. (1996). Perioperative normothermia to reduce the incidence of surgical-wound infection and shorten hospitalization. *New England Journal of Medicine, 334*(19), 1209 – 1215. doi:10.1056/NEJM199605093341901

Langham, G. E., Maheshwari, A., Contrera, K., You, J., Mascha, E., & Sessler, D. I. (2009). Noninvasive temperature monitoring in postanesthesia care units. *Journal of the American Society of Anesthesiologists*, *111*(1), 90–96. doi:10.1097/ALN.0b013e3181a864ca

Lawson, L., Bridges, E. J., Ballou, I., Eraker, R., Greco, S., Shively, J., & Sochulak, V. (2007). Accuracy and precision of noninvasive temperature measurement in adult intensive care patients. *American Journal of Critical Care, 16*(5), 485–496. Retrieved from http://ajcc.aacnjournals.org/

Lefrant, J. Y., Muller, L., Coussaye, J. E., Benbabaali, M., Lebris, C., Zeitoun, N., Mari, C., Saissi, G.,…& Eledjam, J. J. (2003). Temperature measurement in intensive care patients: Comparison of urinary bladder, oesophageal, rectal, axillary and

inguinal methods versus pulmonary artery method. *Intensive Care Medicine, 29*, 414–418. doi:10.1007/s00134-002-1619-5

León, C., Rodríguez, A., Fernández, A., & Flores, L. (2005). Infrared ear thermometry in the critically ill patient: An alternative to axillary thermometry. *Journal of Critical Care*, *20*(1), 106-110. doi:10.1016/j.jcrc.2004.08.011

Liu, C. C., Chang, R. E., & Chang, W. C. (2004). Limitations of forehead infrared body temperature detection for fever screening for severe acute respiratory syndrome. *Infection Control and Hospital Epidemiology*, *25*(12), 1109–1111. doi:10.1086/502351

Lu, S. H., Dai, Y. T., & Yen, C. J. (2009). The effects of measurement site and ambient temperature on body temperature values in healthy older adults: A cross-sectional comparative study. *International Journal of Nursing Studies*,*46*(11), 1415–1422. doi:10.1016/j.ijnurstu.2009.05.002

Manian, F. A., & Griesenaur, S. (1998). Lack of agreement between tympanic and oral temperature measurements in adult hospitalized patients. *American Journal of Infection Control, 26,* 428–430. doi:10.1016/S0196-6553(98)70040-9

Marable, K., Shaffer, L., Dizon, V., & Opalek, J. (2009). Temporal artery scanning falls short as a secondary, noninvasive thermometry method for trauma patients. *Journal of Trauma Nursing*, *16*(1), 41-47. doi:10.1097/01.JTN.0000348069.71593.9d

Masamune, T., Yamauchi, M., Wada, K., Iwashita, H., Okuyama, K., Ino, H., Yamakage, M.,... & Matsukawa, T. (2011). The usefulness of an earphone-type infrared

tympanic thermometer during cardiac surgery with cardiopulmonary bypass: Clinical report. *Journal of Anesthesia*, *25*(4), 576–579. doi:10.1007/s00540-011-1144-1

Mason, T. M., Reich, R. R., Carroll, M. E., Lalau, J., Smith, S., & Boyington, A. R. (2015). Equivalence of temperature measurement methods in the adult hematology/oncology population. *Clinical Journal of Oncology Nursing*, *19*(2), E36–E40. doi:10.1188/15.CJON.E36-E40

Mattie, A. S., & Webster, B. L. (2008). Centers for Medicare and Medicaid services' "never events": An analysis and recommendations to hospitals. *Health Care Manager, 27*(4)*,* 338–349. doi:10.1097/HCM.0b013e31818c8037

McConnell, E., Senseney, D., George, S. S., & Whipple, D. (2013). Reliability of temporal artery thermometers. *Medsurg Nursing*, *22*(6), 387. Retrieved from http://www.medsurgnursing.net/cgi-bin/WebObjects/MSNJournal.woa

Modell, J. G., Strong, C. A., & Hagood, L. (1999). Hope for the infrared tympanic thermometer: One model outperforms the others. *Southern Medical Journal*, *92*(7), 737–738. Retrieved from http://sma.org/smj-home/

Moran, J. L., Peters, J.V., Solomon, P. J., Grealy, B., Smith, T., Ashforth, W., Wake, M., … & Peisach, A. R. (2007). Tympanic temperature measurements: Are they reliable in the critically ill? A clinical study of measures of agreement. *Critical Care Medicine, 35(1),* 155–164. doi:10.1097/01.CCM.0000250318.31453.C8

Myny, D., De Waele, J., Defloor, T., Blot, S., & Colardyn, F. (2005). Temporal scanner thermometry: A new method of core temperature estimation in ICU patients. *Scottish Medical Journal, 50*(1), 15–18. doi:10.1177/003693300505000106

National Institute of Standards and Technology. (2011). *Promoting alternatives to mercury thermometers.* Retrieved from http://www.nist.gov/pml/div685/grp01/mercury_alternatives.cfm

Nguyen, A. V., Cohen, N. J., Lipman, H., Brown, C. M., Molinari, N. A., Jackson, W. L., Kirking, H.,... & Roberts, R. R. (2010). Comparison of 3 infrared thermal detection systems and self-report for mass fever screening. *Emerging Infectious Diseases*, *16*(11), 1710–1717. doi:10.3201/eid1611.100703

Nimah, M. M., Bshesh, K., Callahan, J. D., & Jacobs, B. R. (2006). Infrared tympanic thermometry in comparison with other temperature measurement techniques in febrile children. *Pediatric Critical Care Medicine*, *7*(1), 48–55. doi:10.1097/01.PCC.0000185476.35550.B2

Nonose, Y., Sato, Y., Kabayama, H., Arisawa, A., Onodera, M., Imanaka, H., & Nishimura, M. (2012). Accuracy of recorded body temperature of critically ill patients related to measurement site: A prospective observational study. *Anaesthesia and Intensive Care*, *40*(5), 820. Retrieved from http://www.aaic.net.au/

Norman, D. C. (2000). Fever in the elderly. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America, 31*(1), 31, 148–151. doi:10.1086/313896

Nuckton, T. J., Goldreich, D., Wendt, F. C., Nuckton, J. G., & Claman, D. M. (2001). A comparison of 2 methods of measuring rectal temperatures with digital thermometers. *American Journal of Critical Care*, *10*(3), 146–150. Retrieved from http://ajcc.aacnjournals.org/

O'Brien, D. L., Rogers, I. R., Holden, W., Jacobs, I., Mellett, S., Wall, E. J., & Davies, D. (2000). The accuracy of oral predictive and infrared emission detection tympanic thermometers in an emergency department setting. *Academic Emergency Medicine, 7*(9), 1061–1064. doi:10.1111/j.1553-2712.2000.tb02101.x/pdf

Oermann, M. H. & Hays, J. C. (2011). *Writing for publication in nursing* (2nd ed.). New York, NY: Springer Publishing Company

Onur, O. E., Guneysel, O., Akoglu, H., Aydin, Y. D., & Denizbasi, A. (2008). Oral, axillary, and tympanic measurements in older and younger patients with and without fever. *European Journal of Emergency Medicine, 15*, 334–337. doi:10.1097/MEJ.0b013e3283036cfe

Ostrowsky, B., Ober, J., Wenzel, R., & Edmond, M. (2003). The case of the cold thermometers. *American Journal of Infection Control, 31*(1), 57–59. doi:10.1067/mic.2003.10

Pearce, J. M. S. (2002). A brief history of the clinical thermometer. *QJM: An International Journal of Medicine, 95*(4): 251–252. doi: 10.1093/qjmed/95.4.251

Potter, P., Schallom, M., Davis, S., Sona, C., & McSweeney, M. (2003). Evaluation of chemical dot thermometers for measuring body temperature of orally intubated

patients. *American Journal of Critical Care*, *12*(5), 403–408. Retrieved from http://ajcc.aacnjournals.org/

Prentice, D., & Moreland, J. (1999). A comparison of infrared ear thermometry with electronic predictive thermometry in a geriatric setting. *Geriatric Nursing*, *20*(6), 314–317. doi:10.1053/gn.1999.v20.103926001

Rabbani, M. Z., Amir, M., Malik, M., Mufti, M., Bin Pervez, M., & Iftekhar, S. (2010). Tympanic temperature comparison with oral mercury thermometer readings in an OPD setting. *Journal of the College of Physicians and Surgeons Pakistan*, *20*(1), 33–36. Retrieved from http://www.jcpsp.pk/

Rajee, M., & Sultana, R. V. (2006). NexTemp thermometer can be used interchangeably with tympanic or mercury thermometers for emergency department use. *Emergency Medicine Australasia*, *18*(3), 245–251. doi: 10.1111/j.1742-6723.2006.00847.x

Rubia-Rubia, J., Arias, A., Sierra, A., & Guirre-Jaime, A. (2011). Measurement of body temperature in adult patients: Comparative study of accuracy, reliability, and validity of different devices. *International Journal of Nursing Studies, 48*, 872–880. doi:10.1016/j.ijnurstu.2010.11.003

Russell, C. L. (2005). An overview of the integrative research review. *Progress in Transplantation, 15*(1), 8–13. doi:10.7182/prtr.15.1.0n13660r26g725kj

Schey, B. M., Williams, D. Y., & Bucknall, T. (2009). Skin temperature as a noninvasive marker of haemodynamic and perfusion status in adult cardiac surgical patients:

An observational study. *Intensive and Critical Care Nursing*, *25*(1), 31–37. doi:10.1016/j.iccn.2008.05.003

Schmäl, F., Loh-van den Brink, M., & Stoll, W. (2006). Effect of the status after ear surgery and ear pathology on the results of infrared tympanic thermometry. *European Archives of Oto-Rhino-Laryngology and Head & Neck*, *263*(2), 105–110. doi:10.1007/s00405-005-0966-6

Sehgal, A., Dubey, N. K., Jyothi, M. C., & Jain, S. (2002). Comparison of tympanic and rectal temperature in febrile patients. *Indian Journal of Pediatrics*, *69*(4), 305–308. doi:10.1007/BF02723215

Sener, S., Karcioglu, O., Eken, C., Yaylaci, S., & Ozsarac, M. (2012). Agreement between axillary, tympanic, and mid-forehead body temperature measurements in adult emergency department patients. *European Journal of Emergency Medicine*, *19*(4), 252–256. doi:10.1097/MEJ.0b013e32834c584

Sessler, D. I. (2008). Temperature monitoring and perioperative thermoregulation. *Anesthesiology, 109*(2), 318–338. doi:10.1097/ALN.0b013e31817f6d76.

Shin, J., Kim, J., Song, K., & Kwak, Y. (2013). Core temperature measurement in therapeutic hypothermia according to different phases: Comparison of bladder, rectal, and tympanic versus pulmonary artery methods. *Resuscitation*, *84*(6), 810–817. doi:10.1016/j.resuscitation.2012.12.023.

Siberry, G. K., Diener-West, M., Schappell, E., & Karron, R. A. (2002). Comparison of temple temperatures with rectal temperatures in children under two years of age. *Clinical Pediatrics, 41*(6), 405–414. doi:10.1177/000992280204100605

Singh, V., Sharma, A., Khandelwal, R., & Kothari, K. (2000). Variation of axillary temperature and its correlation with oral temperature. *Journal of the Association of Physicians of India*, *48*(9), 898–900. Retrieved from http://www.japi.org/

Singler, K., Bertsch, T., Heppner, H. J., Kob, R., Hammer, K., Biber, R., ... & Christ, M. (2013). Diagnostic accuracy of three different methods of temperature measurement in acutely ill geriatric patients. *Age and Ageing*, *42*(6), 740–746. doi:10.1093/ageing/aft121

Smith, L. S. (2003). Using low-tech thermometers to measure body temperatures in older adults: A pilot study. *Journal of Gerontological Nursing*, *29*(11), 26–33. doi:10.3928/0098-9134-20031101-08

Smith, L. S. (2004). Temperature measurement in critical care adults: A comparison of thermometry and measurement routes. *Biological Research for Nursing, 6*(2), 117–125. doi:10.1177/1099800404268917

Smitz, S., Giagoultsis, T., Dewé, W., & Albert, A. (2000). Comparison of rectal and infrared ear temperatures in older hospital inpatients. *Journal of the American Geriatrics Society*, *48*(1), 63–66. doi:10.1111/j.1532-5415.2000.tb03030.x

Smitz, S., Van de Winckel, A., & Smitz, M. F. (2009). Reliability of infrared ear thermometry in the prediction of rectal temperature in older inpatients. *Journal of Clinical Nursing, 18*(3), 451–456. doi:10.1111/j.1365-2702.2008.02565.x

Spitzer, O. P. (2008). Comparing tympanic temperatures in both ears to oral temperature in the critically ill adult. *Dimensions of Critical Care Nursing*, *27*(1), 24–29. doi:10.1097/01.DCC.0000304671.83874.05

Stelfox, H. T., Straus, S. E., Ghali, W. A., Conly, J., Laupland, K., & Lewin, A. (2010). Temporal artery versus bladder thermometry during adult medical-surgical intensive care monitoring. *BioMed Central Anesthesiology, 10*(13), 2–7. doi:10.1186/1471-2253-10-13

Sund-Levander, M., Forsberg, C., & Wahren, L. K. (2002). Normal oral, rectal, tympanic and axillary temperature in adult men and women: A systematic literature review. *Scandinavian Journal of Caring Sciences, 16*(2), 122–128. doi:10.1046/j.1471-6712.2002.00069.x

Sund-Levander, M., & Grodzinksy, E. (2009). Time for a change to assess and evaluate body temperature in clinical practice. *International Journal of Nursing Practice, 15*, 214–249. doi:10.1111/j.1440-172X.2009.01756.

Sund-Levander, M. & Grodzinksy, E. (2013). Assessment of body temperature measurement options. *British Journal of Nursing, 22*(16), 942–950. doi:10.12968/bjon.2013.22.16.942

Taylor-Piliae, R. E. (1999). Utilization of the Iowa Model in establishing evidence-based nursing practice. *Intensive and Critical Care Nursing*, *15*(6), 357–362. doi:10.1016/S0964-3397(99)80029-9

Terry, A. J. (2012). *Clinical research for the doctor of nursing practice.* Sudbury, MA: Jones & Bartlett Learning

Titler, M. G., Kleiber, C., Steelman, V. J., Rakel, B. A., Budreau, G., Everett, L. Q., Buckwalter, K. C.,... & Goode, C. J. (2001). The Iowa model of evidence-based practice to promote quality care. *Critical Care Nursing Clinics of North*

America, *13*(4), 497–509. Retrieved from http://lgdata.s3-website-us-east-1.amazonaws.com/docs/2600/564557/Titler_article_2001.pdf

Protection of Human Subjects, 45 C.F.R. § 46 (2009).

Varney, S. M., Manthey, D. E., Culpepper, V. E., & Creedon, J. F. (2002). A comparison of oral, tympanic, and rectal temperature measurement in the elderly. *Journal of Emergency Medicine*, *22*(2), 153–157. doi:10.1016/S0736-4679(01)00457-7

Washington, G. T., & Matney, J. L. (2008). Comparison of temperature measurement devices in post anesthesia patients. *Journal of PeriAnesthesia Nursing*, *23*(1), 36–48. doi:10.1016/j.jopan.2007.10.001

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, *52*(5), 546–553. doi:10.1111/j.1365-2648.2005.03621.x

Winslow, E. H., Cooper, S. K., Haws, D. M., Balluck, J. P., Jones, C. M., Morse, E. C., …& Kelly, P. A. (2012). Unplanned perioperative hypothermia and agreement between oral, temporal artery, and bladder temperatures in adult major surgery patients. *Journal of PeriAnesthesia Nursing, 27*(3), 165–180. doi:10.1016/j.jopan.2012.01.012

Wolfson, M., Granstrom, P., Pomarico, B., & Reimanis, C. (2013). Accuracy and precision of temporal artery thermometers in febrile patients. *MedSurg Nursing, 22*(5), 297–302. Retrieved from http://www.medsurgnursing.net/cgi-bin/WebObjects/MSNJournal.woa

Woodrow, P., May, V. Buras-Rees, S., Higgs, D., Hendrick, J., Lewis, T., Whitney, S.,… & McHenry, M. (2006). Comparing no-touch and tympanic thermometer temperature recordings. *British Journal of Nursing*, *15*(18), 1012–1016. doi:10.12968/bjon.2006.15.18.22028

Wunderlich, C. A. (1871). *On the temperature in diseases*. London, England: The New Sydenham Society.

Zaccagnini, M., & White, K. (Eds.). (2011). *The doctor of nursing practice essentials*. Sudbury, MA: Jones & Bartlett Publishers.