ABSTRACT

Based on a Class 2 limited risk hospital based clinical trial, and a subsequent project retrospective, Briggs and Ramer propose an expanded clinical trial protocol. Such an expanded protocol would be especially helpful for evaluating De Novo devices: new inventions which require new processes for full hospital integration. Since these new processes would often require training of nurses and supporting professionals, Briggs and Ramer suggest that nurse investigators could be very effective in carrying out such expanded protocol studies. We briefly describe the FDA approval process, the role of the nurse in evidence based medical device evaluation, the Ramer, et al. clinical trial, the proposed expanded clinical trial protocol, and candidate categories of devices that might employ the limited risk Class 2 Medical Device clinical trial protocol. The investigators look forward to carrying out such a medical device clinical trial in the near future.

Key Words: Medical devices, FDA, De Novo, Clinical trials, Protocol, Nurse investigators

1. INTRODUCTION

The regulation of both chemical substances and medical devices in the United States lies with the Food and Drug Administration (FDA), where the primary concern is safety. Procedures for the FDA approval of chemical substances and medical devices have historically been very different. Chemical substances have most frequently been approved based upon large scale clinical trials; whereas medical devices have been given approval through a variety of approaches depending on assessment of risk. In clinical practice, responsibility for medical prescriptions typically lies with the primary care provider, supported by the pharmacist for chemical substances, as well as nurses and biotechnical specialists for medical devices used in a hospital setting.

Medical devices are subdivided by the FDA, U.S. Food and Drug Administration[1] into Three classes. Class 1 devices are self-certifiable devices like reading glasses, walkers, wheelchairs, and walking canes; materials like cotton balls or Q tips. Class 2 devices include specialized hearing aids, diagnostic ultrasound equipment, user friendly fundus cameras, and near infrared vein viewers. Class 3 devices include implantable prosthetic joints, MRI conditional pacemakers, cardiac monitoring devices, cardiac occluders, stents and artificial heart valves. In the European Union (EU), Class 2 devices are further classified into 2a and 2b based upon potential risks associated with the device.[2]

This leads to three levels of approval processes:

(1) Low risk Class 1 devices are usually exempt from formal testing, grandfathered in with an individual device exemption (IDE) if equivalent in function to existing approved products.

(2) Class 2 devices are seldom given an IDE exemption. Instead, FDA approval involves a process that includes both development and validation stages. These stages
include a feasibility study followed by a pivot study, leading to specialized validation protocols for each family of devices.

(3) Class 3 devices not only require feasibility studies and extensive bench testing for safety, but completion of a premarket evaluation (PMA) that requires one or more clinical trials.

It might be possible to have much simpler approval procedures for certain class 2 devices. Rather than going through specialized processes, unique to each type of device, a standardized process could be developed—especially for class 2a (low risk) devices. Such a process would include standardized clinical trial protocols somewhat similar, but far more focused than those used for chemical substances. They would provide the prospective performance evidence for instrument usage effectiveness as well as an evidence basis for implementation guidelines. These protocols would be of special interest to nurses and hospital administrators, since they could provide information about costs and benefits as well as assurances of safety. Extending a standardized approach to class 2b and class 3 devices is also possible, but is more problematic. Class 2a devices involving new technology, like the VeinViewer Ramer and Briggs[3] studied, involve almost no risk of harm, so that the manufacturer is almost certain to be approved. The clinical trial almost becomes a marketing trial as well: a win-win. Class 2b and Class 3 devices involve risk of harm, so failing to find a clear difference may lead to a costly FDA decline. The stakes are much higher.

2. MEDICAL DEVICE RESEARCH AND EVIDENCE BASED PRACTICE IN NURSING

Nurses have been using devices to assist in patient care since the advent of the thermometer and blood pressure cuffs. Examples of the many devices are not limited to restraints, bandages, electronic thermometers, intravenous pumps (IV), or IV assistive devices. Even though these devices are generally assumed to be safe as well as accurate, they may never have been evaluated with careful experimental comparisons. Lack of experimental evidence from carefully controlled studies often leaves the nurses giving and receiving ad hoc incomplete training. There is little thought given to the complex consequences of a device development process and evaluation. Even the FDA may be misled about the safety of medical devices, since their focus is harm from the device itself rather than harm from a process that relies on the utility or the precision of the device. Is the process of a device evaluation currently being carried out?

There is a very limited literature broadly evaluating device research strategies in nursing. McConnell[4] lists the principles of assessing devices as need, safety, effectiveness, economic appraisal and social impact. Giuliano[5] takes the approach of matching the correct patient need to technology or device. She then goes on by generally describing how to evaluate the technology or foundations for device research; however, Giuliano does not discuss the experimental design “how to” or required statistical data analysis of device research. She does raise the issue of bias if the device evaluation is funded by the manufacturer without a standardized protocol in place.[5] McCarthy and Shaban[6] anticipate the authors’ interests by including discussions of device research phases as well as classes of devices.

In the nineties, McConnell[4] did several studies exploring the methods used to implement devices in Australia. The focus of this research was methods used in educating nurses. These studies included evaluations of whether or not nurses felt prepared to use the devices.[7–9] For example, they found that the orientation to an IV pump was more comprehensive than a Foley catheter.[9] McConnell[9] discusses how to evaluate a device, but not how to use prospective quantitative research designs to evaluate a device. These writings do not discuss the outcomes of the research supposedly done during the device development process.

More recently, research has focused on the clinical implications of devices. Rosenthal et al.[10] conducted an international retrospective study on nosocomial infection rate in intensive care units (ICU) in the Americas and developing nations using existing data bases. The most frequent infections occur in ventilators, central line catheters and in-dwelling urinary catheters. In their study, Rosenthal et al.[10] details the rate of infection per 1,000 days in these three devices and compares the rate between different countries. The relevance of both the McConnell[9] and the Rosenthal, et al.[10] work is limited. None of these studies contemplate the modern American FDA processes[11] that often involve elaborate protocols appropriately approved by an Institutional Review Board (IRB).

McCarthy and Shaban[6] outlined an approach for nurses to lead trials in an international setting. Ramer et al.[11] accepted the McCarthy & Shaban[6] challenge by evaluating the VeinViewer,[12] an innovative device that helps a nurse locate a vein. Implementation of VeinViewer technology would involve the development of both training and implementation protocols. A nurse evaluator lead team would also need to determine whether or not the VeinViewer would be considered an asset by clinical patients, and cost effective by administrators. Subsequent evaluation of the Ramer et al.[11] study has led Ramer and Briggs[3] to contemplate the development of a more general Clinical Trials protocol for
similar Class 2 medical devices.

3. A DEVICE STUDY EXEMPLAR: THE VEINVIEWER STUDY

3.1 Introduction
Ramer and Briggs[3] have recently been working with an example of class 2a medical devices with very limited risk. The specific device they tested was a vein viewer. The vein viewer helps the nurse find a vein for more reliable vascular access. The vein viewer and similar devices would lend themselves to a standardized clinical trial protocol. Such a protocol would allow a vendor to demonstrate value both through efficiency and patient comfort. Through a single-site clinical trial, Ramer, et al.[11] have demonstrated the feasibility of a multisite clinical protocol capable of affording multiple independent and dependent variables.

3.2 The problem
The study by Ramer, et al.[11] aims to successfully integrate near infrared-based visualization technologies into a pediatric outpatient hematology/oncology clinic in an underserved population. For such patients, indwelling catheters are not deemed safe for outpatient chemotherapy which results in patients requiring IV access at each clinic visit whenever blood is needed for laboratory studies or chemotherapy is required. Even though NIR-based vascular access assistive devices are used by many, there is still a need to review the existing research methods used with this type of technology and evaluate the benefits to both the institution and the patient.

3.3 Clinical implementation
Obtaining IV access has long been a specialized nursing skill that requires clinical knowledge plus psychomotor coordination. Unsuccessful IV access is frustrating for nurses and is widely known to cause anxiety for the patient. A 2008 nursing survey by the Eztel-Hardman group found that nurses considered anxious pediatric patients to be a difficult patient population in which to successfully obtain IV access.[13] Additionally, unsuccessful IV placements can adversely affect patient satisfaction scores. IV assistive devices can assist nurses in successful IV starts as well as save time. For the patient to have only one attempt to start an IV can be less frightening and increase patient satisfaction.[11]

3.4 Type of study
This was a simple device group vs. comparison group study evaluating the VeinViewer against standard methods for IV access without an assistive device. The dependent variable was the effect of the VeinViewer on venous access procedural time as compared with standard methodology. Quantitative data was gathered on patient and nurse satisfaction.[11]

Participants During the open period, fifty-three (53) participants were enrolled, twenty-seven (27) of which were randomly assigned to the VeinViewer group and twenty-six (26) of which were randomly assigned to the standard methods group. The average age of participants was 13.1 (range 1-21) years of age and was not significantly different between randomization groups ($p = .789$). The distribution of sex, height, weight, and BMI was likewise comparable between randomization groups ($p > .05$, all). The majority of the study population were of Hispanic descent ($n = 48$), but there were two (2) subjects who defined as African American and three (3) subjects of Asian descent.[11]

Aside from the primary cancer diagnosis, additional comorbidities for this population included depression, dermatological findings, diabetes, other endocrine issues and or hematological issues. The majority of subjects enrolled were diagnosed with acute lymphoblastic leukemia; however, other types of leukemia, lymphomas and sarcomas were noted. Subjects also presented with a range of hematological diagnoses which ranged from anemia to neutropenia.[11]

It is known that recent venous access procedures can hinder subsequent procedures. Therefore, information about recent IV placement or phlebotomy procedures was tracked. Most subjects or their families reported that their last venous access attempt was greater than 1 week previous to the study appointment. However, approximately 9 subjects had access procedures performed within the past week with 2 of those happening in the 12-24 hour period prior to the study time. Arm dominance was also noted with most subjects exhibiting right-handed preference.[11]

Of the fifty-three (53) study participants, forty-four (44) were in clinic for phlebotomy procedures while and nine (9) required IV therapy. Across all subjects, fifty-one (51) of the study access sites were considered optimal and two sites (2) were considered secondary, less preferred areas of access as determined by the study nurse: thirty-eight (38) in the hand, fourteen (14) in the antecubital fossa and one (1) in the foot.[11]

3.5 Performance results
A clearly significant difference in procedural time between groups was demonstrated with nurses using the VeinViewer requiring less time ($p < .05$).[11]

3.6 Satisfaction results
Subjects rated themselves equal in terms of attempts, difficult access status, and nurse time spent and perceptions of venous
access pain ($p \geq .05$). However, subjects rated nurses using the VeinViewer as having significantly more skill than nurses who did not use the VeinViewer ($p < .05$). Additionally, patients gave significantly better scores for ‘overall experience’ to the VeinViewer group ($p < .05$).[11]

Nurses using the VeinViewer generally saw the device in a positive light with a majority of users agreeing that the device provided greater patient ease, ease of use, and more venous options. Nurses either agreed or strongly agreed most of the time that they would want to have the device available to them for their next vascular access attempt. Nurses either disagreed or strongly disagreed most of the time that subjects were intimidated by the device or that the device made them change their point of catheter/needle insertion.[11]

### 3.7 Implications

Ramer and Briggs[3] have shown interest in carrying out such a model clinical protocol as a multisite Clinical Trial with an appropriate medical device.

### 4. Medical Instrument Protocol Candidates

The entire spectrum of Medical devices and approval procedure can be overviewed and studied in great detail on the U.S. Food and Drug Administration[11] website: [http://www.fda.gov/MedicalDevices/default.htm](http://www.fda.gov/MedicalDevices/default.htm). Our special interest is in medical devices, especially new innovative medical instruments that are of marginal risk and used in a medical setting by nurses with patients, like the VeinViewer. Here is a very brief excerpted summary of the FDA processes:

- **Devices that are submitted to the FDA for approval by manufacturers go through a suggested step by step process:**
  1. **Classification** as Class 1, Class 2, or Class 3 device.
  2. **Is the product is the same of functionally equivalent** to an existing approved device? If so, the manufacturer may simply apply for and receive what is called a 510(k) exemption. Nearly all Class 1 devices, and many Class 2 devices, will be exempt from a 510(k); most would require no more than a 510(k) application.
  3. **Some Class 2 devices and most Class 3 devices require a Pre Market Approval process (PMA)** that assesses instrument safety through bench testing, and patient safety and effectiveness through clinical trials.
  4. **New devices that have no functional equivalence in the medical world, like the VeinViewer are not easily classified based upon steps 1-3 as described above.** They are tested by means of a De Novo process that encompasses much of the rigor of a step 3 PMA even if the instrument poses no more than a Class 2a risk.

Two recent examples of 2016 approved De Novo devices somewhat similar to the innovative VeinViewer are the Pediatric Vision Scanner and the EarLens Contact Hearing Device. The Pediatric Vision Scanner[14] (De Novo number [DEN]1300521) can carry out a preliminary evaluation of children’s vision for strabismus to anticipate possible needs for visual training or surgery to ensure the proper development of binocular vision. The EarLens Contact Hearing Device[14] (DEN 150002) enables properly trained hearing impaired users to improve their understanding of spoken language.

5. **Approved medical devices are subject to continuous modification, requiring both post market surveillance and testing.** The surveillance is focused on negative outcomes, very unlikely for Class 2a devices. Testing may be necessary when a device has changed sufficiently. If the device were approved through a De Novo process, it might require little more than a 510(k) application the second time around. However, it might be helpful to do a completely different type of testing in the post market environment for some medical devices: comparing whole families of similar products against efficacy standards. Such testing would need to be supported financially by independent sponsors such as charitable foundations or the CDC rather than the manufacturers both for comparative purposes and to establish consistent training protocols.

### 5. Proposed Instrument Protocol for Medical Clinical Trials

This protocol is an expansion of an improved version of the Ramer, et al.[11] VeinViewer study, with modifications suggested by Ramer and Briggs[3] and expansions from single to multiple sites. Rather than creating unique protocols for each medical instrument, one could customize a standardized protocol making any Class 2 medical device appropriately testable with a scaled approach, and diverse devices easier to compare with each other. Such a standardized approach would make it easier for diverse IRBs to cooperate seamlessly in a multisite clinical trial. Each individual site from this revised protocol can be considered a module in a larger design. Viewed this way, if the corrected Ramer and Briggs[3] study had been carried out on five sites with significant findings, it could be said to have been successfully replicated five times. Such replications are very powerful evidence of validity. Five successive replications with the same outcome is a statistically significant binomial event.

The data from all of the sites could be alternately be collapsed into a database and evaluated for differences in scores...
between experienced and inexperienced nurses, different types of patients, etc. One could incorporate as many sites (or modules) as needed depending on the specific device being evaluated. Such data could be analyzed with various models of ANOVA, leading to evidence for possible difference in skill levels of nurses or effectiveness with different patient groups. For examples of alternative design possibilities and data treatment methods, see Shadish, Cook, and Campbell.\[15\]

Class 2a medical device clinical studies are prospective experimental designs that typically begin before multisite expansion with two primary independent variables, nurse/patient, and two primary dependent variables, performance/satisfaction. The study will always compare performance/satisfaction between at least two groups—an experimental group, and a comparison group. The experimental group will use the instrument, and the comparison group will carry out the existing best practice. If this study is done only at one site, such as Ramer, et al.,\[11\] neither the nurses nor the patients are likely to adequately reflect performance across the full range of user environments. However, even Ramer, et al.\[11\] goes well beyond the bench studies and qualitative assessments often performed to obtain FDA approval for medical instruments.

5.1 Modules (or sites)
Each module consists of a minimum # of nurse/patient pairs providing standardized performance/satisfaction scores at one site. Increasing the # of modules greatly increases the number of statistically valid comparisons that can be made. Even two modules allow comparisons between patient/nurse pairs involved in different treatment activities. As the number of modules increase, it becomes possible to compare performance across different experience levels of nurses, ages of patients, different clinical environments, different models/versions of instruments, and alternative nurse/user procedures. Five to ten modules would seem a convenient benchmark for expanded Class 2a Instrument clinical trial protocols, depending on number of patients required per site.

5.2 Population/sample size
The size of the population of patients to be served by a medical device can play a significant role in determining the scale of a clinical trial. The scale, in turn, can play a determining role in expanding a clinical trial beyond a single site. It would appear that the VeinViewer could serve a sizeable population, and would benefit from multisite staging.

The sample size of each group of patients depends on the estimated size of the difference measured in bench testing. With a measured time of 30 seconds in practice trials with the VeinViewer and 40 seconds without, one can estimate the percentage difference and then calculate the number of patients needed at a specific alpha level (usually \( p < .05 \)) using power calculations based on Tchebysheff’s theorem.\[16\] Our protocol for a single site, revised for power, would have required 30 completed trials for each of the two groups, or 60 completed patient trials for each site.

If the Class 2a Medical Device Protocol we are suggesting were used with either The EarLens Contact Hearing Device\[13\] (DEN 150002), or The Pediatric Vision Scanner\[1\] (DEN 1300521), the potential population of patients would be much smaller and more specific than our VeinViewer study, and would require fewer test sites or modules.

5.3 Dependent variables
It is very important statistically that the dependent variables chosen be full ratio quantitative variables, preferably with a normal or near normal underlying distribution. Empirically, this is almost always impossible with performance data. If one is scoring the dependent variable time emphasizing speed, scores tends to peak close to minimal time, and the resulting distribution is positively skewed toward the fastest times; if one is scoring the same dependent variable emphasizing accuracy, scores tend to peak close to optimal accuracy, and the resulting distribution is positively skewed toward longer times, in the direction being explored. This leads to the possibility that outliers in the data (always in the wrong direction) can completely obscure real statistically significant results. In experimental studies, it is possible to minimize or even eliminate outliers through good designs that control or counterbalance for such unwanted effects. Knowing what type of distributions to expect, it is also possible to scale data to more closely approximate normality, and to minimize the consequences of outliers.

For satisfaction data, it is possible to approximate normality with interval scaling: either behaviorally anchored agree/disagree Likert responses, or scales from 1 to 10. These often are supplemented with qualitative participant driven responses to open ended questions. Such responses may be skewed similarly to accuracy or time, but outliers tend to be a much less serious problem.

Expanding our protocol to include diverse medical devices, we must consider devices different dependent variables. If our Protocol had been applied to The Pediatric Vision Scanner\[11\] (DEN 1300521), the dependent variable might have involved sensitivity/selectivity score comparisons, expressed in per cent. If it had been The EarLens Contact Hearing Device\[1\] (DEN 150002), the dependent variable would have been the score on a reading comprehension test (interval
5.4 Independent variables
The independent variables, the nurse/patient pairs can be studied in some depth in multisite clinical trials, but they always are the nuclear unit of study for any individual site—and can be sufficient independent variables for a complete study, as in Ramer et al. Performance of these pairs is always nested within different practices, that might involve very different patient groups and concerns. Patients might be young or old, reasonably healthy, or very sick. Nurses might be newly trained or experienced familiar or unfamiliar with the type of instruments being tested.

5.5 Validation
Validation of a Class 2a Medical Instrument consists of obtaining a statistically significant performance comparison and confirming satisfaction report by nurse and patient. In Ramer et al., the performance of finding a vein with the Christie Instruments VeinViewer was significantly faster than without—both patients and nurses expressed more satisfaction when the VeinViewer was used. In a multisite clinical trial, many more comparisons could be made which would enhance the usefulness and implementation of the new technology. Expanding the Class 2a Medical Device protocol to multiple sites would always strengthen the validation, but also provide additional information to help device manufacturers to market and hospitals to implement.

6. Applications of the Medical Device Protocol
(1) Product categories of special interest, either for U.S. Food and Drug Administration De Novo or post market testing:
   A. Motorized wheel chairs, which promise increased mobility for the elderly and impaired might be tested under through PMA clinical trials, as De Novo products, or against efficacy cost/benefit standards—or might even be eligible for 510 (k) approval as an improved product. However, they might be further evaluated through surveillance reports, since there are instances in which users experience harm or injury in an accident.
   B. Workplace ergonomic devices, including specialized backrests, chairs, and wrist protectors for carpal tunnel could be evaluated post market against cost/benefit efficacy standards that would include training protocols both for healthcare professionals and users.
   C. Ultrasound devices are employed for many clinical diagnostic purposes, including observations of fetal development in pregnant women. Their diagnostic efficacy could be studied post market for sensitivity/specificity.
   D. There are now many software solutions available both for training and for hospital management. Training can be provided through simulations that anticipate critical team coordination, including emergency childbirth.

(2) Special benefits to manufacturers and hospitals when the risk is low: Right now, good bench testing for performance and safety, and some evidence of clinical effectiveness is sufficient for product approval—as long as the product poses minimal risk. Patient and nurse satisfaction, of interest both to the hospital and manufacturer, comes only when a product is tried out with the kind of protocols that demand consistent performance.

CONFLICTS OF INTEREST DISCLOSURE
The authors declare that there is no conflict of interest.

REFERENCES


