A RISK BASED USER TOOL TO BUILD USER CENTERED LABELS FOR MEDICAL DEVICES

By

Eric Joseph Estrada

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Packaging – Master of Science

2021

ABSTRACT

A RISK BASED USER TOOL TO BUILD USER CENTERED LABELS FOR MEDICAL DEVICES

By

Eric Joseph Estrada

Herein we develop a user-driven, risk-based tool to inform the design of a standardized label for use with medical devices. Researchers identified 11 labeling inputs found on commercial labels and organized the inputs into a “Device Facts” box at 3 risk levels: high, medium, and low. mock labels and commercial labels were objectively compared by healthcare practitioners using a forced choice methodology where accuracy and response time served as dependent variables. Results suggested that pairwise comparisons between labels (mock vs commercial) within a given risk category (e.g. high) yielded statistically significant differences at a confidence level of 95% for time to correct response. For both medium ($p=0.0016$) and high risk information ($p<0.0001$), the mock labels yielded a quicker correct response than their commercial counterparts. Only for low risk information were the commercial labels faster ($p<0.0001$). The gains in speed made in high/moderate risk information were not attributable trade-off. Mock labels were at least as accurate as their commercial counterparts; low and high risk yielded no sign of significant difference when mock and commercial were compared and participants were significantly more accurate with questions requiring medium risk information for the mock labels.

ACKNOWLEDGEMENTS

I would like to take the time to thank my advisor, Dr. Laura Bix, who has been extremely patient and supportive of me throughout my whole time at the School of Packaging. I have come to her with several crazy ideas (some not enough) and she has helped me stay grounded and on track the whole time. I would also like to thank Dr. Susan Selke and Dr. Mark Becker who took the time to help me review my work and give me great feedback. To all of the healthcare providers, companies, and conferences who have donated material, time, and exposure, I am extremely grateful. These groups include the Association of Surgical Technologists, Cook Medical, Abbott, Smith & Nephew, Eagle Labs, HealthPack, and HcPIE. This work may have been to help them in their fields but would have been impossible to complete without their help. And finally I would also like to give my appreciation to my family, friends, and fellow peers within the HUB team. The list would go on forever, but from just a couple of words of advice to spending hours helping me review/support my research, I wanted everyone to know no words of gratitude are enough.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

In recent years, medical device labeling has been placed under the spotlight, garnering attention from various stakeholders. In comparison to highly regulated pharmaceutical labeling, there are very few requirements when it comes to medical device labels. Under Title 21 Code of Federal Regulations (CFR) Part 801 (21 CFR 801), the governing US regulations for medical device labeling, manufacturers are required to include certain key pieces of information but provided with virtually no guidance on a specific format or layout, leaving that to the discretion of the manufacturers (Gaffney, 2015). The US Food and Drug Administration (FDA) is aware of this gap and has shown increased interest in standardization of medical device labels; this is evidenced by a 2014 announcement of their study investigating 12 potential device labeling standards. They more broadly acknowledged the need for enhanced oversight in a 2013 Federal Register announcement which indicated a "growing need for medical device labeling to be delivered in a clear, concise and readily accessible format so that patients, caregivers and healthcare providers may access and utilize device labeling as efficiently and effectively as possible," (Gaffney, 2015).

Relatively new European Medical Device Regulations (MDR) suggest that this phenomenon is not unique to the US; specifically, the MDR requirements are in the middle of their five-year pre-implementation period, which began in May 2017. These MDR requirements are an overhaul of what is required on a medical device label, which has big implications for medical device manufacturers distributing their products in Europe (*Complying with Labelling as EU MDR Gets Close*, 2018). More specifically, these labeling requirements include the presence

of a Unique Device Identifier (UDI), warnings & precautions being directly printed on a medical device label (previously included in the Instructions for Use), and the inclusion of a symbol identifying the product as a medical device to name a few. Looking beyond Europe, things become even more complicated when increasing the scope globally. From dealing with the State Food and Drug Administration (S FDA) in China to the Brazilian customer protection code in Brazil, each regulatory body has their own specific requirements for each country, forcing medical device manufacturers to spend resources on localization efforts for their products (Songara, 2010).

Our work provides empirical evidence related to the performance of medical device labeling, something that will be needed to inform both regulatory and standardization decisions. Specifically, we present the development of a risk-based tool which prioritizes label information that users deem to be the most critical for the safe use of medical devices. This label standard derived from the use of the tool is termed the "Medical Device Facts Box," a design which resembles the Drug Facts Label (DFL) dictated by Title 21 CFR Subpart C §201.66 for over the counter (OTC) medication. Prioritization of the information in the Medical Device Facts Box was based on a user-assessed risk (a combination of the probability of harm and severity harm) survey which asked users to evaluate the risk associated with missing or misunderstanding eleven labeling inputs universally found on medical device labels in the US. The user-assessed risk associated with misinterpreting or missing each of the eleven medical device labeling inputs was obtained via feedback provided using a survey administered to healthcare providers. The eleven pieces of information were designed and formatted based on information gathered during an extensive review of the literature review related to the performance of medical device labels in an attempt to create label designs which emphasized information deemed crucial by healthcare

providers. The mock labels created were subsequently evaluated using a Forced Choice Task Decision method to objectively compare the performance of these labels to their commercial label counterparts. In doing so, we fill the gap in empirical evidence regarding effective labeling design for medical devices across multiple devices and types.

Chapter 2

Background

To understand the paradigm shift currently happening that involves medical device labeling, it is important to acknowledge the history of how contemporary labeling practice came to be. Since the early 20th Century, food and drug labeling has been regulated by the US Food and Drug Administration (FDA). In the beginning after multiple attempts, the US Congress tried to regulate the Food Industry with little success. It was not until the release of Upton Sinclair's *The Jungle*, that the public feverishly supported regulatory oversight of commercial producers. Sinclair's muckraking political fiction spoke of putrid conditions within the meat packing plants of Chicago with the intent of advancing social reform for the working class. Instead of Sinclair's intended message, readers fixated on the horrific details of the unsanitary conditions of the packing plants. Official government investigations discovering that Sinclair's portrayal were actually true, coupled with the outrage held by the public, led to the passage of the Pure Food and Drugs Act of 1906 (Francis, 2020). This law prohibited food and drug companies from selling products that were deemed "misbranded" or "adulterated" (Francis, 2020). Later historical events expanded these concepts to other products, including medical devices.

Although the Pure Food and Drugs Act was the first step in providing systemic regulatory oversight to those producing food and drugs, products in the marketplace illuminated shortcomings of the law. From foods still deceptively packaged and/or labeled, to drugs with false therapeutic claims, and even products that caused harm to the consumer. One such example was the "Lashlure", an eyelash dye which caused eye-related injuries to women and in a single case resulted in blindness.

However, even with many products with either false claims or damaging side effects, a new bill to replace the Pure Food and Drugs Act was not enacted until the aftermath of the elixir sulfanilamide disaster of 1937. A Tennessee drug company marketed sulfanilamide in an elixir form that held the promise of easier dosing for pediatric patients; the new dosage form, unfortunately resulted in the death of over 100 people, with a majority of the victims being children (FDA, 2018). Just as *The Jungle* drew public outrage which pressured the passage of the Pure Food and Drugs Act, the sulfanilamide disaster of 1937 also sparked public outcry which led to the release of the Federal Food Drug and Cosmetic Act of 1938 (FFDCA), and resulted in increased regulatory authority specific to drug products (FDA, 2018). This law filled many gaps that the Pure Food and Drugs Act had related to drugs and foods. The intention of the new law was that manufacturers had to provide evidence of safety and efficacy prior to marketing of new drugs and proof of fraud was no longer a necessity to challenge false marketing claims for drugs. Foods now had new food standards including tolerances for additives and residues such as pesticides. And now for the first time, cosmetics and therapeutic devices were now to be regulated as well (Janssen, 1981).

Medical devices remained a very loosely regulated industry into the 70s, until premarket problems associated with informal (and inappropriate) testing of an intrauterine device (IUD), the Dalkon Shield, catalyzed the passage of the Medical Device Amendments Act of 1976 (Johnson, 2016). The 1976 amendments to the FFDCA established a risk-based classification system for medical devices (Class I- low risk; Class II- moderate risk; Class III- high risk), where premarket burdens related to safety and efficacy and level of regulatory oversight drive classification. (Johnson, 2016).

Despite enhanced oversight of devices in the relatively recent past, a quick glance at the requirements for labeling of medical devices found in 21 CFR Part 801 in comparison to the labeling requirements found in 21 CFR Part 201 show disparity between the two; with drugs having more regulation (21 CFR Part 801) (21 CFR Part 201). The labeling for food and drugs is prescribed in great detail (dictating content and formatting, down to the details of recommended fonts, minimum types sizes and precise placement of information) (Llamas, 2020). For labeling purposes, different requirements are split between Over the Counter (OTC) and Prescription Drugs with both containing different formatting and content rules and objectives as found in 21 CFR Part 201 Subpart C and 21 CFR 201 Subpart B respectively (21 CFR Part 201 Subpart B) (21 CFR Part 201 Subpart C) . OTCs which do not require the oversight of a physician have labeling requirements which are standardized and follow the same format and content requirements set forth by the FDA including the use of a "Drug Facts Box" as per 21 CFR Part 201.66 (21 CFR Part 201.66) . As prescription drugs are only available through prescriptions, labeling may have fewer requirements; for example, adequate directions for use being exempt per 21 CFR Part 201.100 provided the requirements are met (21 CFR Part 201.100). With regards to labeling, the FDA does not currently regulate the warning and instruction labels typically seen on prescription drugs which may vary depending on the pharmacy (Llamas, 2020).

General controls are required for all medical devices sold within the US, regardless of their risk classification. The general controls prescribe the minimal requirements for medical device labeling. Requirements are located in the following sections of the Code of Federal Regulations (CFR):

- General Device Labeling - 21 CFR Part 801
- Use of Symbols - 21 CFR Part 801.15
- In Vitro Diagnostic Products - 21 CFR Part 809
- Investigational Device Exemptions - 21 CFR Part 812
- Unique Device Identification - 21CFR Part 830
- Good Manufacturing Practices - 21 CFR Part 820
- General Electronic Products - 21 CFR Part 1010

General device labeling requirements dictate that the name and place of manufacture, intended use of the device, and adequate directions for use be "clearly labeled" (801). Other required information is specific to the packaged device itself, such as, latex/natural rubber warnings. Some sections of the CFR (801.15) specify the use of symbols, or are specific to unique products or circumstances (809, 812, 1010). None of the information requirements provide mandates regarding formatting or design of the information to be presented. As such, manufacturers present what is required in a myriad of places, frequently separating components that users deem critical, and this lack of standardization has resulted in a proliferation of varied presentations and formats which has been found to be confusing for healthcare providers (Cai, 2012).

Section 830 of the requirements related to general controls dictates requirements related to Unique Device Identifiers (UDI's). UDIs represent a relatively new piece of information required for the labeling of medical devices and are intended to assist with supply chain transparency. UDI standards primarily focus on the type of information that is encoded within the number which represents the UDI (specifically, a device identifier, followed by a product

identifier) and the presentation format of the information (automatic identification data capture (AIDC) and human readable) (*Access GUDID*, 2020).

Furthermore, the FDA, with the collaboration of the National Library of Medicine (NLM), created the Global Unique Device Identification Database (GUDID) to act as a repository for device identification information that has been submitted to the FDA for devices that contain UDI. Per GUDID and the FDA, a device UDI is compromised of the Device Identifier (DI) – "A unique numeric or alphanumeric code specific to a device version or model" and Production Identifier(s) (PI) – "Numeric or alphanumeric codes that identify production information for a device…". The information required in the UDI per FDA's 2013 UDI mandate is as follows:

1. Lot or batch within which a device was manufactured
2. Serial number of a specified device
3. Expiration Date
4. Manufactured Date
5. Distinct identification code required by 21 CFR 1271.290 (c) for a human cell tissue, or cellular and tissue-based products (HCT/P) regulated as a device

(*Access GUDID*, 2020) An example of a fictitious label with common labeling information identified from GUDID is seen in Figure 1.

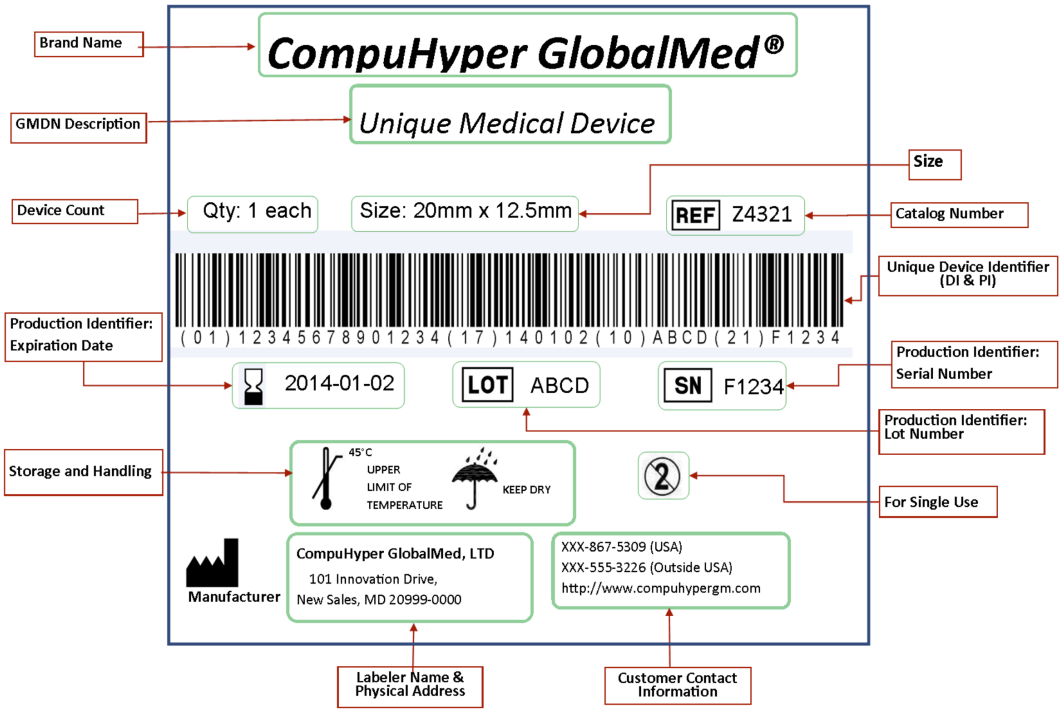


Figure 1 – GUDID Label Example (*GUD ID Diagram*, 2020)

US requirements do not dictate the placement, size and other design elements for medical device labeling information are not dictated (FDA, 2019). As a result, the labels of medical devices are incredibly varied with regard to the presentation and formatting of information that is important to their safe and effective use.

This is further compounded by an increasing complexity of medical devices and various environments for use of these products (dental and medical offices, home environments, prehospital environments, emergency departments, perioperative environments, acute care wards, ambulatory surgery centers, hospice, nursing homes, battle fields, veterinary clinics, etc.). These factors suggest that the thoughtful design of devices and their packaging to be an important goal for improving healthcare delivery (Ward, 2004). That said, any labeling standard must consider the diversity of products already available in the device market. Examples range

from simple tongue depressors and scalpels to MRI scanners and patient monitors which play a huge role in the health and safety of a patient being diagnosed or treated. As time and technology move forward, devices increasingly are employed to diagnose and monitor patients, and, as such, ensuring their safe and efficacious use is more critical than ever before (Ward & Clarkson, 2004)

With medical devices becoming more advanced with newer technologies, medical errors are becoming more unavoidable as each layer of complexity creates more opportunities for such medical errors. Medical errors are errors that occur in a medical setting; where correct practice is not being conducted and may result in patient harm (Ward & Clarkson, 2004). In the US it is estimated that medical errors can result in up to 100,000 deaths per year. A flow of the different subgroups of errors that fall under Medical Errors can be seen in Figure 2.

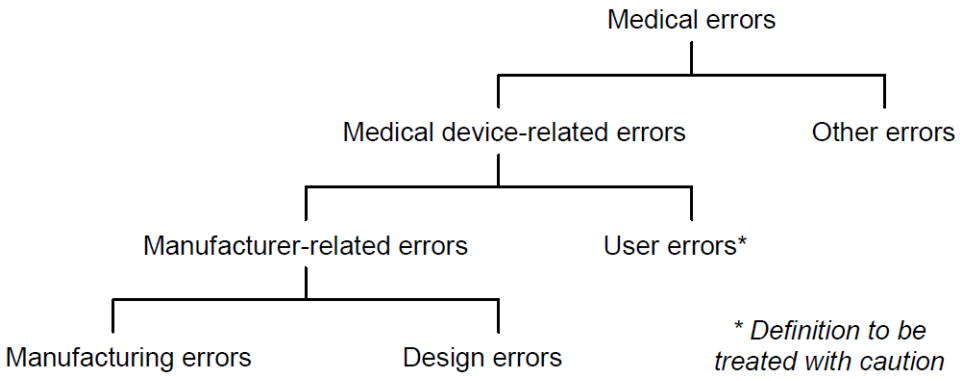


Figure 2 – Medical Error Subgroup Flowchart

A breakdown of Medical Errors Ranging from device error to manufacturer error, one of the most prevalent errors in accident research, even in areas outside of the medical industry, is user error (Ward & Clarkson, 2004). User errors occur when the device is not at fault; specifically, when the error is caused by a human. Nowadays, humans are more likely to be the biggest threat to complex and hazardous systems such as healthcare systems when compared to technical

related failures (Reason, 1995). Research into system design suggests that, while it is possible to mitigate risk caused by humans, it is nearly impossible to eliminate it (Reason, 1995). Included in the various types of user error is the misunderstanding of labeling and instructions by healthcare providers which must be taken into account when mitigating patient risk.

Chapter 3

Literature Review

The growing interest in revising the requirements for the labels of medical devices has catalyzed a small, but growing, body of research investigating both the performance of existing labels and proposed designs.

3.1 Cai's Research

Research conducted by Cai (Cai, 2012) characterized the performance of medical packaging through the lens of healthcare providers within the perioperative environment. The primary goal of the work was to identify needed areas of innovation and research. Seven focus groups were conducted with a total of 21 practitioners to evaluate medical types of packaging and performance of features, as well as how the operating room context affects packaging utility. As part of the focus groups, participants also conducted a series of activities intended to prioritize opportunities and frustrations. These activities included: *"(1) The rank ordering of different packaging features regarding importance, (2) The rating of varied aspects of packaging (quick identification, ease of opening and aseptic presentation), (3) group development of a list of frequent problems associated with healthcare packaging as well as self-surmised estimates of the frequency of occurrence of each.* Labeling emerged as one of the top problems associated with packaging according to perioperative personnel; specifically, the top three problems affiliated with packaging by perioperative personnel were identified as: aseptic presentation (41.4%), opening difficulty (31.0%), and labeling (19.0%). With regards to labeling, two broad themes were found to be consistent:

(1) Healthcare providers indicated that they preferred not to have to read label information, relying on simple heuristics like color and a visual confirmation of the product (using transparent packaging), preferring packaging that enabled users to quickly identify the contents without the need for reading.

(2) Providers indicated that information critical to the safe and effective use of devices must be clear and easily identifiable.

When asked about pieces of information healthcare providers deemed to be critical for the safe and effective use of medical devices, four pieces of information emerged. Namely: product identity, expiration date, sterility status, and whether the contents contained latex or not. Participants consistently reported that non-critical information on packaging made it harder to identify the critical pieces of information (Cai, 2012).

3.2 Seo's Research

Seo (2014) conducted a benchmarking study intended to objectively evaluate how different design approaches (grouping critical information, boxing critical information, using a simple system for color coding and using symbols) affected accuracy of the selection of a medical device and the time to correctly identify it. Prior to evaluating the efficacy of the aforementioned design factors, a benchmarking study was conducted to assess the assertion of Cai's participants that the information they deemed critical (product name, sterility status, latex status and expiration dating) was frequently scattered throughout device labels. Labels from six commercial products (all indwelling urinary catheters) provided by two companies were assessed in the benchmarking study. All six products evaluated had four columns of text on their lidstock. These were not only used to evaluate the placement of critical information on the labels during the benchmarking study, but were used also as model for the creation of labels that served as a comparative point of performance to objectively evaluate label performance. The benchmarking study reported the frequency with which the four pieces of critical information appeared in a

single column (n= 0 ; 0%); across two (n= 8; 40%) three (n=12; 60%) or across all four columns of information (n=0; 0%). Results support Cai's (Cai, 2012) findings from focus groups which suggested information critical to the safe and effective use of medical devices tended to be scattered throughout labels and that noncritical information tended to interfere with the ability to identify it.

After confirming reported labeling issues via the benchmarking study, Seo (Seo, 2014) redesigned the labels for the purpose of evaluating how varied design strategies previously discussed affected attention to critical information and selection of products using two methods: change detection, and a forced choice task. Each of the four design elements was presented in two levels (present and absent—specifically: boxed and unboxed; grouped and ungrouped; symbol present and absent; and with and without color coding), and all combinations were crossed for a total of 12 label designs tested (2 x 2 x 2 x 2).

The forced choice testing included 54 trials, each trial consisted of two labels which were designed in the same way (e.g. color present, boxing absent, critical information grouped and symbol absent) one piece of information differed between the two labels. For example, for one trial one label would indicate the presence of latex while the other did not. Participants were given instructions to choose a label based on given criteria (e.g. choose the product containing latex) as quickly as possible. The dependent variables correctness (selected correctly- binary variable) and time to correct selection (continuous variable) were recorded and analyzed. Results suggested that optimal performance occurred in the label treatments that included the use of color, symbol and grouped critical information but did not include boxing the same. These treatments (color coded, symbols used and information grouped) were more accurate (97.3% correct response) and quicker (3.53 seconds) than either of the commercial labels tested (92.0%

correct response for commercial label A 89.8% for commercial B) with accuracies of 92.0% and 89.8% as well as times of 8.92 and 8.26 seconds respectively for commercial label A and commercial label B. Results support the notion that altering design factors, specifically, color coding, symbol usage, and grouping critical information can significantly impact the performance of medical device labeling relative to the current commercial approach to design.

3.3 RTI International

Work conducted by research group RTI International, also supports standardization of medical device labels (Stifano et al., 2013). Focus groups were held to collate professional opinions and feedback on medical device labels. The focus groups worked with all aspects of labeling including primary packaging, inserts, and manuals. Some of the key recommendations for improvement included the use of:

- larger fonts
- color
- more white space in between information
- clear and concise Information

This information was utilized to create a format for pacemaker labeling which was also reviewed by healthcare practitioners which received positive feedback (Stifano et al., 2013).

3.4 Medical device Labeling Inputs

To create a standard for medical device labeling, a review of the requirements related to the same was needed. General controls are required for all medical devices sold in US commerce. Title 21 CFR Part 801 contains the requirements related to labeling. At a minimum, the following must be clearly labeled:

1. Name and Place of Business

2. Intended Use of the Device
3. Adequate Directions for Device

Labeling is also required of medical devices that contain latex or are delivered in a sterile state. Similarly, UDI information in both human readable and machine readable formats is now a requirement for all devices, regardless of risk classification (I, II or III). Our review of the CFR and a series of commercially available devices from different risk classification categories yielded the following eleven medical device labeling inputs as either required or common to medical devices sold in the US.

Table 1 – 11 Medical Device Labeling Inputs

11 Medical Device Labeling Inputs
1. Name/Identity of Medical Device
2. Name of Manufacturer, Packer, or Distributor
3. Place of Business of Manufacturer, Packer, or Distributor
4. Adequate Directions for Use
5. Unique Device Identifier
6. Device Containing Latex/Natural Rubber Warnings
7. Sterility Status
8. Storage and Handling Instructions
9. Expiration Date
10. Net Quantity/Weight/Size/Dimensions
11. Unit, Lot, Batch, or Control Number

Although there have been some recommendations for the reform of medical device labeling (Cai, 2012) (Seo, 2014) (Seo et al., 2017) (Stifano et al., 2013) and suggestions of the need for revision from FDA themselves (*Agency Information Collection Activities; Proposed Collection; Comment Request; Survey of Health Care Practitioners for Device Labeling Format and Content*, 2014), ideally, policy recommendations will be informed by empirical evidence related to design performance.

Chapter 4

Study Objectives

Given the lack of standardization for medical devices labeling and our literature review which suggests that healthcare professionals have difficulty finding information reported as critical to the safe and effective use of medical devices quickly (Cai, 2012) (Seo, 2014), *we hypothesized that we could create a more efficient label standard for medical devices.*

Our specific objectives were:

Objective 1 – To determine the information contained on medical device labeling that is required (CFR Review) and typical (review of commercial device labels) on varied medical devices sold in US commerce.

This objective, completed in the Literature Review of this thesis, identified 11 medical device labeling inputs identified in Table 1.

Objective 2 - Assess the importance of the information contained on medical device labeling (survey of healthcare providers) from a user-centered, risk-based perspective.

Objective 3 - Utilize survey findings (obtained under objective 2), survey results, and available knowledge from the literature (CFR, Previous Research regarding optimized labeling) to create a user-centered, risk-based medical device label standard which prioritizes and emphasizes the information identified as associated with the highest risks (survey results- objective 2).

Objective 4 - Conduct an objective assessment of the novel, user-centered labels (created in support of objective 3) compared to existing commercial labels for medical devices using a forced-choice test.

Chapter 5

Medical Device Labeling Input Risk Assessment Survey

In support of objective 2, a survey was conducted in order to "grade" the 11 medical device labeling inputs identified in support of Objective 1 (see Table 1) in order to categorize each input based on the risk associated with missing or misinterpreting that piece of information. Survey results were used to characterize the risk associated with a specific labeling input risk (high, medium, and low). This was done by fitting the data into an Item Response Theory Model and then grouping using K-Means Clustering.

5.1 Methodology

5.1.1 Participants

A purposeful, selective sampling technique was used to recruit participants approved under IRB# x17-1448e. Eligible participants had to:

- Be of 18 years of age or higher
- Be a surgeon, surgical technologist, registered nurse, or related healthcare practitioner.

A flyer was also posted on the AST national website; additionally, an email flyer was distributed to Registered Nurses, Surgeons, and Surgical Technicians around the Mid-Michigan area through local connections of the research team.

Participants began the process with an electronic consent form and had to consent to proceed to the data collection portion of the survey. Following the consent process, information evaluation began with a questionnaire which collected demographic information; within this section, respondents also answered questions intended to characterize their work history and

environment. Appendix A provides the IRB approved advertisement, consent form, and pre-survey questionnaire used for the survey administered in support of Objective 2.

5.1.2 Materials and Survey Structure

The survey, conducted online, was designed and delivered using the cloud-based Qualtrics' survey software (SAP; Provo, UT). Qualtrics was selected as a secure and accessible method for conducting the work. Within the recruitment advertisements, both email and traditional, an online link to the password protected survey was provided for access.

The survey tasked the participants with evaluating each of the 11 medical device labeling inputs identified in support of Objective 1 (see Table 1 and Figure 3), based on the user's assessment of the risk that missing or misinterpreting each input would cause.

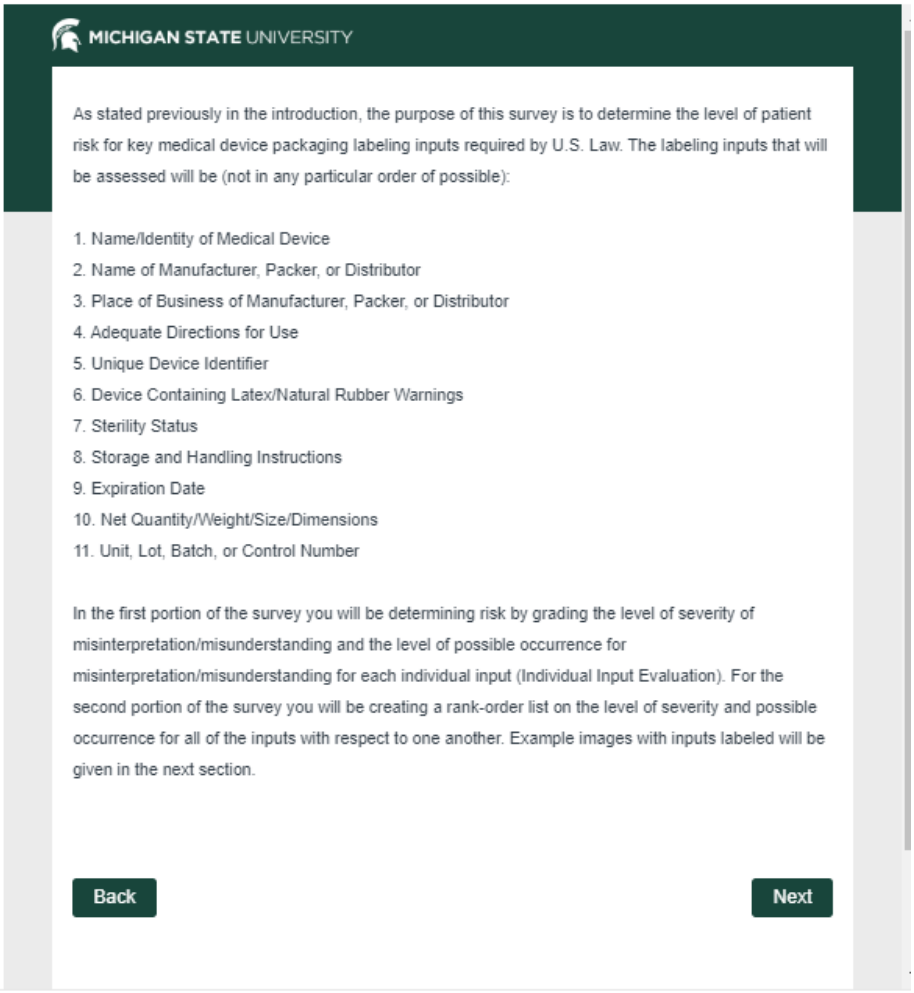


Figure 3 – 11 Medical device labeling Inputs as Presented in Survey

Risk assessment was informed by the definitions set forth in ISO 14971 (International Organization for Standardization, 2007). In accordance with this standard, there are three levels of grading for each of the two components of risk (severity of risk x likelihood of the occurrence). To cultivate a meaningful and consistent understanding and assessment associated with the terminology, survey participants were provided with these definitions from the ISO standard prior to assessing the 11 medical device labeling inputs (see Table 1 and Figure 3). Applicable definitions from the standard are presented in Table 2.

Table 2 – Risk Associated Definitions According to ISO 14971

Severity Level	Definition
Significant	Death or loss of function or structure
Moderate	Reversible of minor injury
Negligible	Will not cause injury or will injure slightly
Occurrence Level	**Definition**
High	Likely to happen often or frequently
Medium	Can happen but not frequently
Low	Unlikely to happen, rare, remote

This information was presented to participants as shown in Figure 4. Participants were informed that this section could be revisited at any point of the survey in order to review the given definitions.

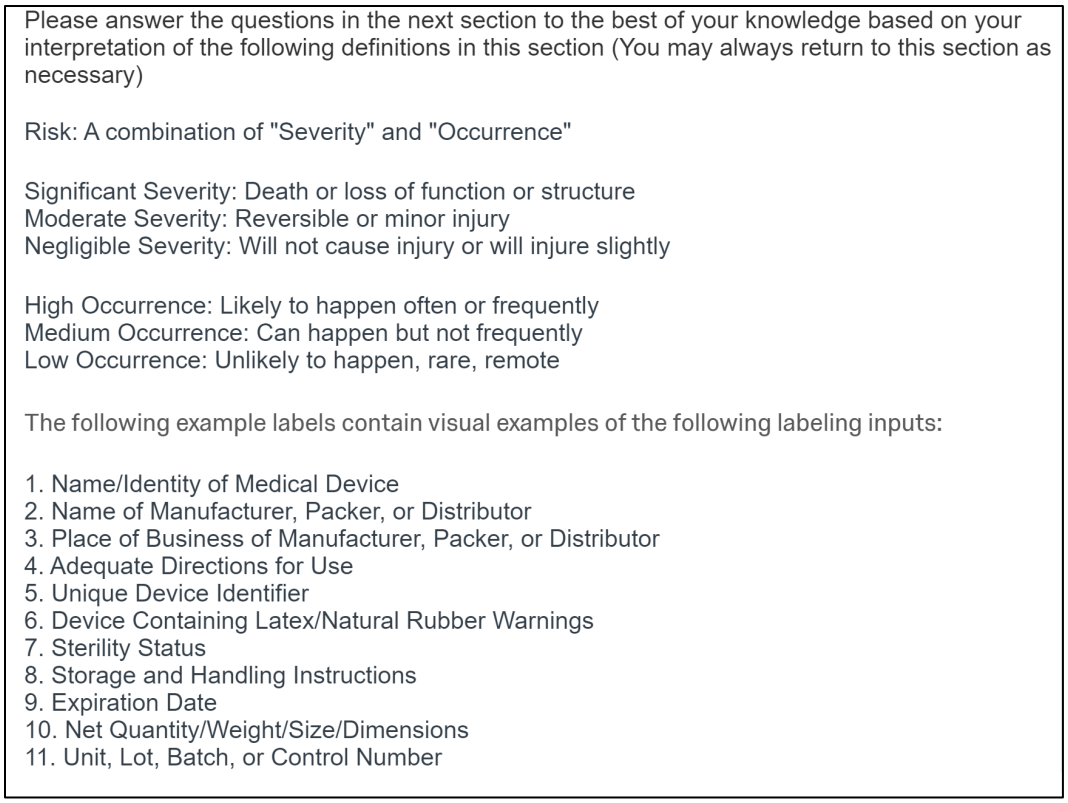


Figure 4 – ISO 14971 Survey Definitions and the 11 identified inputs common (or required) for the labeling of medical devices

Three examples of commercial medical device labels which identified the labeling inputs of interest were presented to participants to enable them to develop a better sense of the task (see Figures 5-7). These three examples were drawn from a pool of donated commercial labels. The collection of these commercial Labels is detailed in Chapter 6.

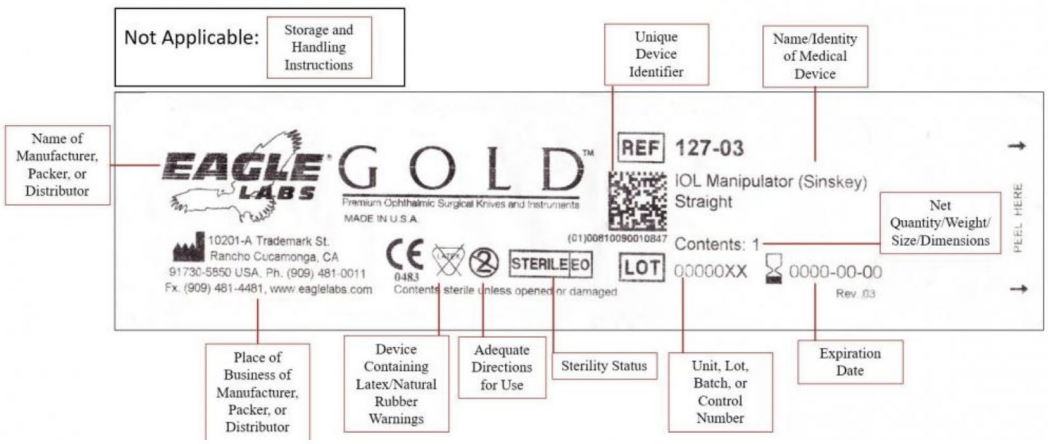


Figure 5 – Class I Medical device (regulatory category associated with the lowest levels of risk)- Commercial Label Example

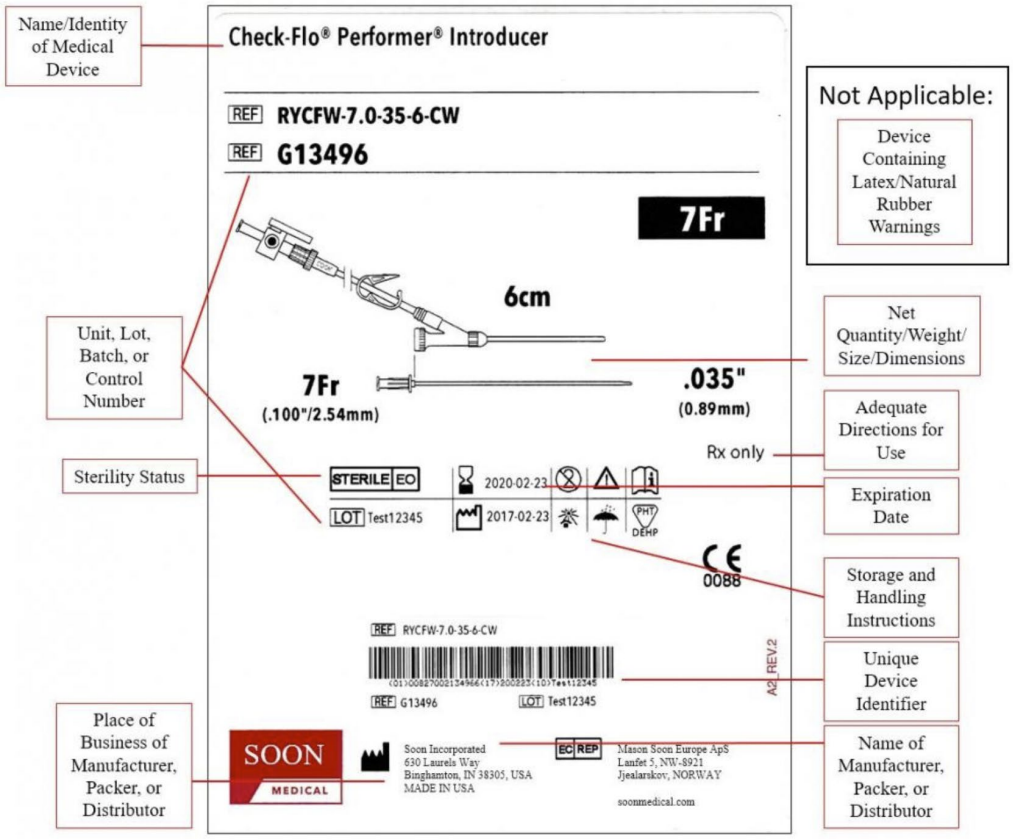


Figure 6 – Class II Medical device (regulatory category associated with elevated regulatory oversight relative to Class I because of increased levels of risk) Commercial Label Example

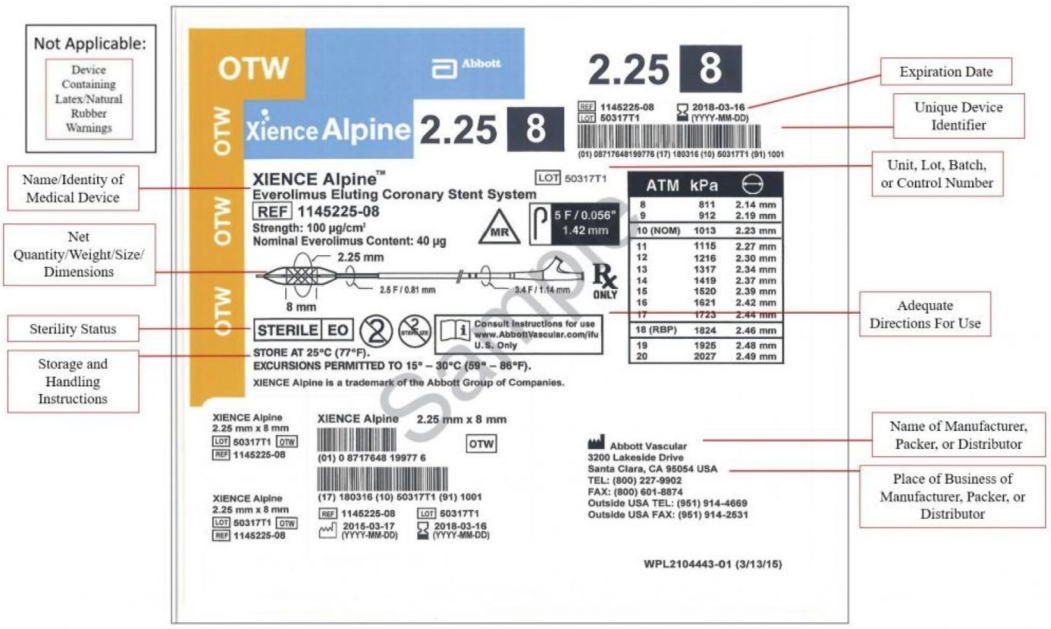


Figure 7 – Class III Medical device (the highest levels of regulatory oversight because of the levels of risk associated with failure to perform) Commercial Label Example

After participants were provided with the standardized definition for both components of risk (severity and occurrence), and had the opportunity to view the three commercial labels with the inputs of interest labeled (Figure 5-7), they were instructed to rate the risk associated with missing or misinterpreting each input on a medical device label (see Table 2 and Figure 4 for definition). With this approach, each of the 11 medical device labeling inputs had an associated severity and occurrence score. Figure 8 depicts a screenshot of how this was presented to participants during the survey.

What is the Severity and Occurrence Level for each labeling input? Select the levels for severity and occurrence in the answer choices below.
(You may return to previous sections to review definitions and examples.)

	Severity Level	Occurrence Level
Name/Identity of Medical Device	Significant Severity	Medium Occurrence
Name of Manufacturer, Packer, or Distributor	Significant Severity	Low Occurrence
Place of Business of Manufacturer, Packer, or Distributor	Moderate Severity	Low Occurrence
Adequate Directions for Use	Moderate Severity	Medium Occurrence
Unique Device Identifier	Significant Severity	Low Occurrence
Device Containing Latex/ Natural Rubber Warnings	Negligible Severity	Medium Occurrence
Sterility Status	Moderate Severity	High Occurrence
Storage and Handling Instructions	Moderate Severity	Low Occurrence
Expiration Date	Moderate Severity	Low Occurrence
Net Quantity/Weight/Size/Dimensions	Significant Severity	Medium Occurrence
Unit, Lot, Batch, or Control Number	Negligible Severity	Medium Occurrence

Figure 8 – Individual Input Severity and Occurrence Rating

The labeling inputs were also evaluated utilizing a rank ordering for both elements of risk (severity and occurrence) using the same definitions from Table 2. While collected, "rank ordering" data was not utilized for the resulting Risk Categorization of the labeling inputs. This portion of the of the study on how it was presented can be found in Appendix B.

Upon completion of the survey, participants were compensated with a $10 Amazon gift card for their time. Compensation was sent to an email disassociated from private information provided by the participant.

5.1.3 Statistical Model

Ultimately, the individual labeling input severity and occurrence ratings were used to group the 11 medical labeling inputs into categories of risk. This data was analyzed by fitting the collected ratings of severity and occurrence (each into their own item response theory model) specifically a Rasch Rating Scale Model, in R with the RSM function. Each model gave numeric scores for all 11 labeling inputs, for both severity and occurrence. Each labeling input's severity and occurrence score was paired and plotted on a graph (severity on the x axis and occurrence on the y). The plotted data was then grouped using the K-Means clustering function in R. The plotted coordinates of severity and occurrence were separated into three groups to align with the three levels of risk as defined in ISO 14971 (International Organization for Standardization, 2007) (see Table 2).

5.2 Results

5.2.1 Characterization of Participants

One hundred and thirty-six healthcare providers were recruited in support of the survey with *all* completing the survey in its entirety (100%). Results of participant characterizations are shown in Figure 9-Figure 11. One-hundred and one respondents were women and 35 men. A majority of the participants had at least 10 years of experience overall as a healthcare provider. The most frequently reported occupation was surgical technologist (n = 104; 76.5%) (See Figure 10).

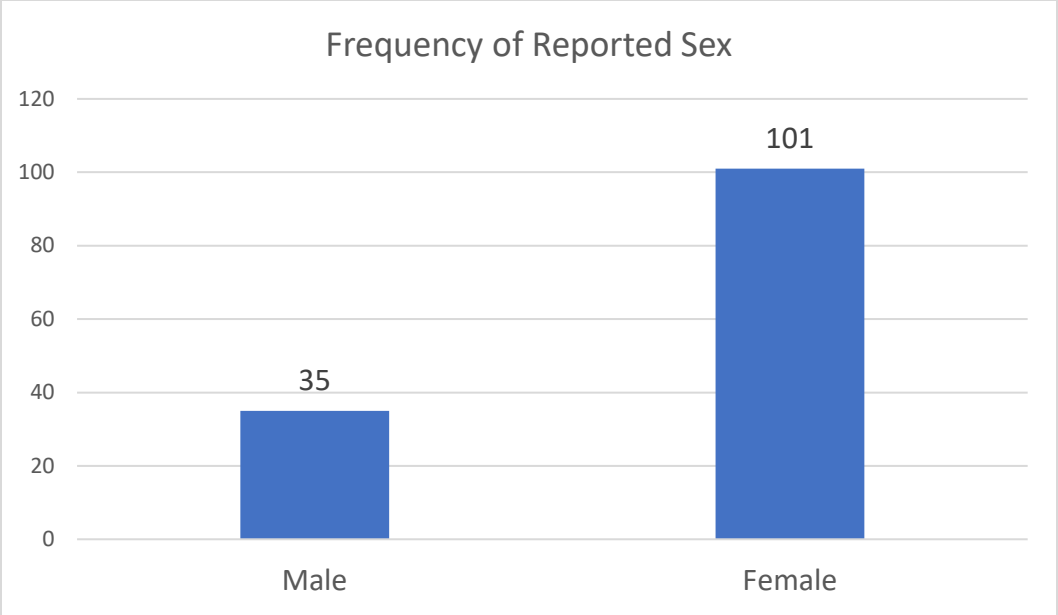


Figure 9 – Survey Participants- Frequency of Reported Sex

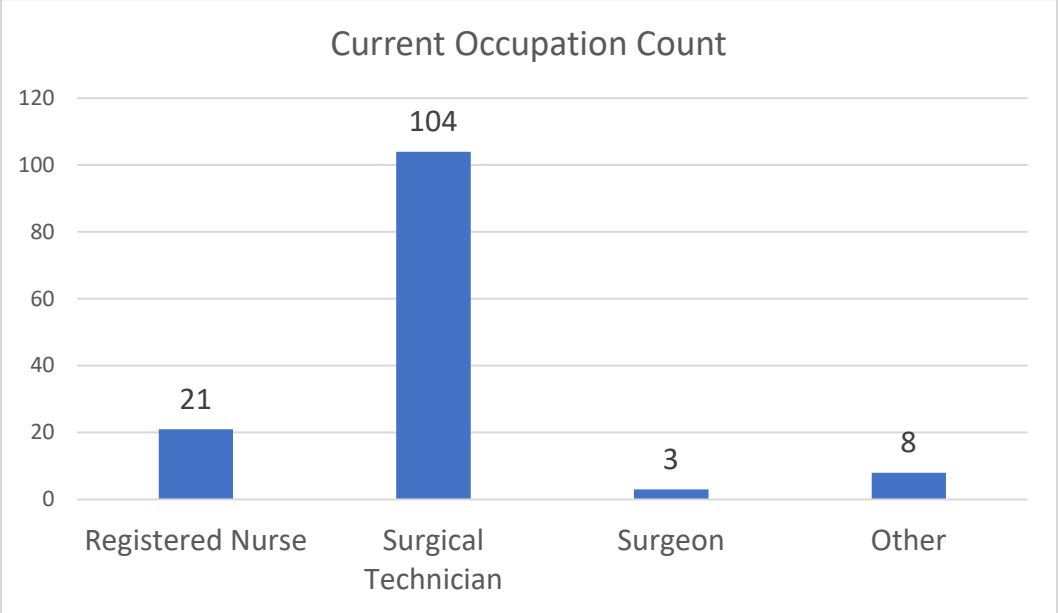


Figure 10 – Survey Participants- Frequency of reported occupation

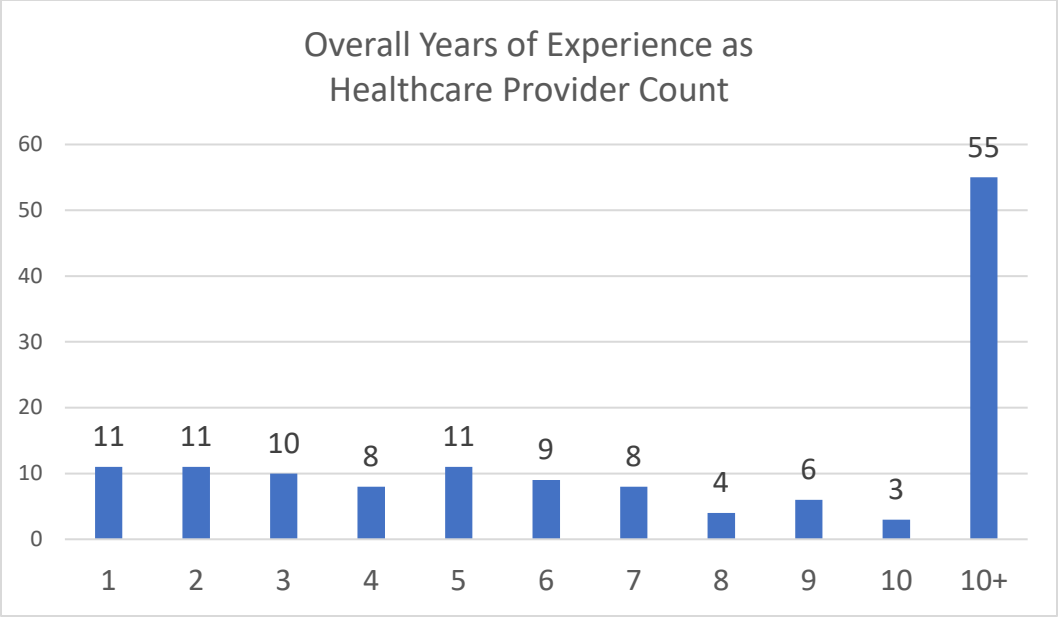


Figure 11 – Survey Participants- Overall Years of Experience as a Healthcare Provider

5.2.2 Survey Results and Analysis

The responses from the survey were analyzed using the Rasch Scale Rating Model Item Response Theory using R. For both models for severity and occurrence outputs, a higher score would indicate a higher severity/occurrence level and vice versa for a lower score. The score output from both the severity and occurrence Rasch Scale Rating Model are shown in Table 3 and Table 4 respectively.

Table 3 – Severity Item Response Theory R Output

Labeling Input	Score
Name/Identity of Medical Device	0.75236
Name of Manufacturer, Packer, or Distributor	-0.97945
Place of Business of Manufacturer, Packer, or Distributor	-1.34299
Adequate Directions for Use	1.80642
Unique Device Identifier	0.11935
Device Containing Latex/Natural Rubber Warnings	2.31681
Sterility Status	1.83048
Storage and Handling Instructions	1.06919
Expiration Date	1.49108
Net Quantity/Weight/Size/Dimensions	-0.41660
Unit/Lot/Batch/Control Number	-0.16058

Table 4 – Occurrence Item Response Theory R Output

Labeling Input	Score
Name/Identity of Medical Device	0.46023
Name of Manufacturer, Packer, or Distributor	-0.29439
Place of Business of Manufacturer, Packer, or Distributor	-0.85143
Adequate Directions for Use	1.77562
Unique Device Identifier	0.63324
Device Containing Latex/Natural Rubber Warnings	1.98626
Sterility Status	1.53016
Storage and Handling Instructions	1.30864
Expiration Date	1.92230
Net Quantity/Weight/Size/Dimensions	-0.15975
Unit/Lot/Batch/Control Number	0.06778

The scores from both models were paired for each labeling input and plotted with severity on x-axis and occurrence on y, as shown in Figure 12.

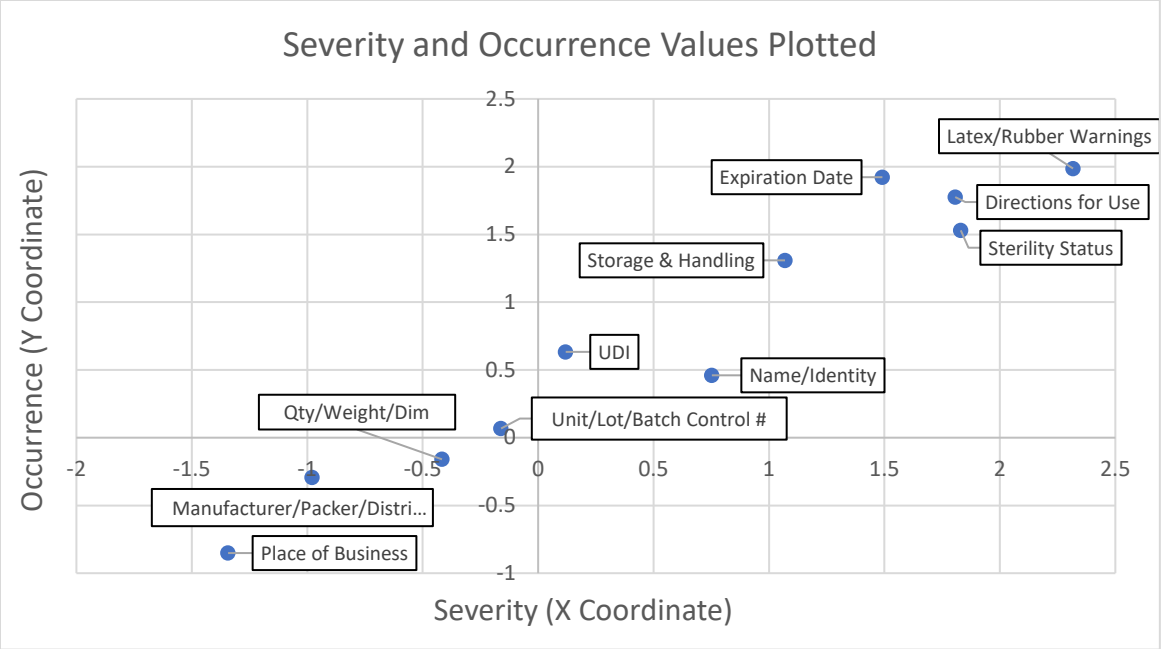


Figure 12 – Severity and Occurrence -IRT Values Plotted for each of the 11 labeling inputs

Using K-Means clustering in R, labeling inputs were grouped into 3 categories aligning with the three levels of risk classified in ISO 14971 (International Organization for Standardization, 2007); namely: high, medium, and low risk levels (see Figure 13). This grouping into 3 categories was pre-determined prior to running the K-Means Clustering algorithm.

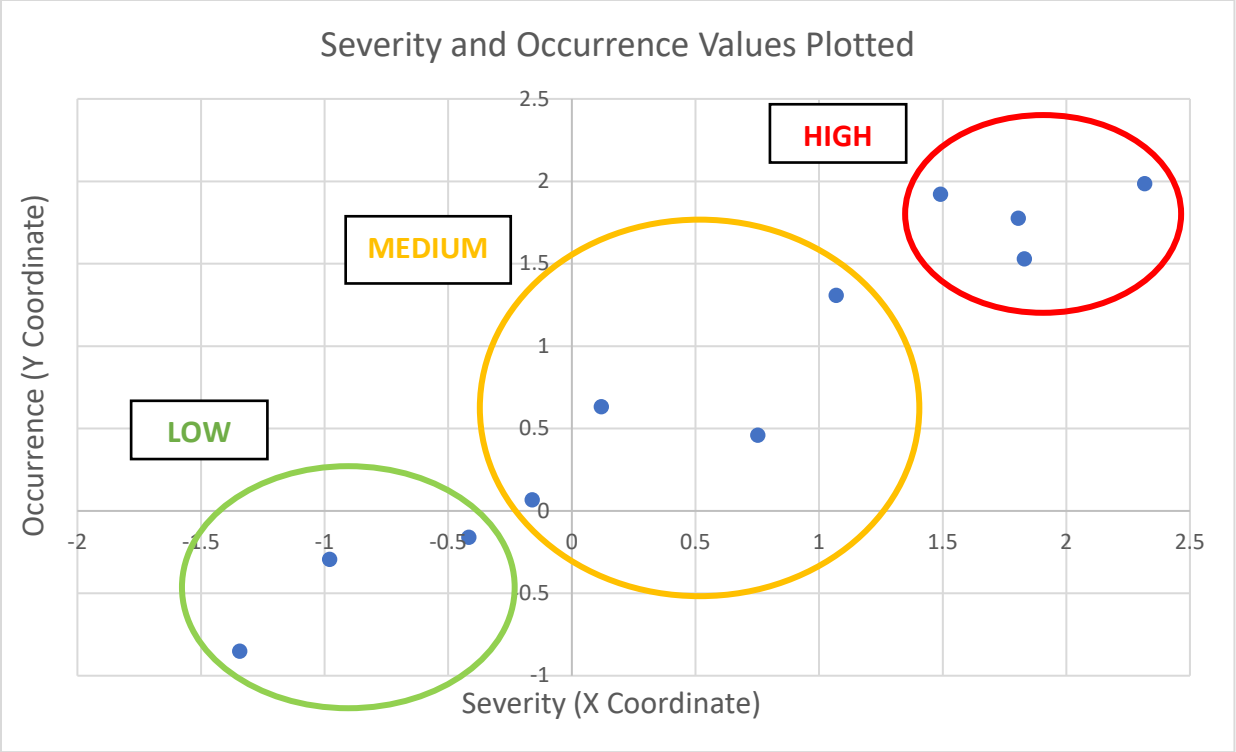


Figure 13 – K-Means Clustering Groups by Risk Level based on Severity and Occurrence ratings collected from Study Participants

Table 5 depicts the information by resultant groups of risk category in a table format.

Table 5 – Medical Device Labeling Input Risk Groups

High Risk
Adequate Directions for Use
Device Containing Latex/Natural Rubber Warnings
Sterility Status
Expiration Date
Medium Risk
Name/Identity of Medical Device
Unique Device Identifier
Storage and Handling Instructions
Unit/Lot/Batch/Control Number
Low Risk
Name of Manufacturer, Packer, or Distributor
Place of Business of Manufacturer, Packer, or Distributor
Net Quantity/Weight/Size/Dimensions

Using the groupings identified from the K-Means clustering of the 11 medical device labeling inputs (see Table 1), mock labels were designed to emphasize high risk information based on previous research (Cai, 2012) (Seo, 2014) (Seo et al., 2017) (Stifano et al., 2013), with the goal of improving label performance (objective 3).

Chapter 6

Mock Label Creation

Information gained from the results of the medical device labeling input risk assessment survey (Chapter 5- Objective 2) and the Literature Review were leveraged to create risk-based, user-focused mock labels (Objective 3). Mock designs included all of the 11 labeling elements previously identified and represented a redesign of existing commercial labels following a simple, yet purposeful, set of design rules.

6.1 Methodology

6.1.1 Labels for Testing

Labels of commercially available devices were collected through a call that was sent out using the research team's connections available via the LinkedIN® network; interested parties provided labels as donations (see Acknowledgements). A complete list of the commercial labels donated is shown in Appendix C. 18 donated commercial labels comprised of all three risk classes (I, II and II) were converted to mock labels which emphasized higher risk information. The FDA classification (based on risk) and the product categories affiliated with the commercial devices used in the creation of the mock labels are depicted in Figure 14 and Figure 15. There was a good mixture of Class II and Class III devices, 9 and 8 respectively. In contrast, Class I, the regulatory category comprising the lowest risk devices, had only one label donated. Cardiovascular devices comprised the largest category represented (44% of the donated labels employed for use in the study).

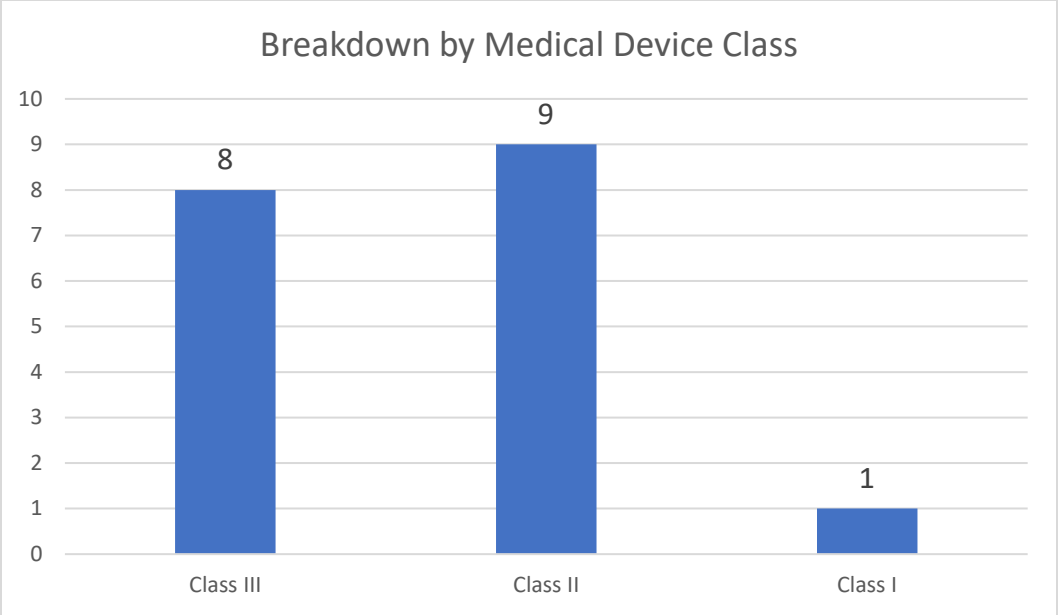


Figure 14 – Donated Labels by Regulatory Classification

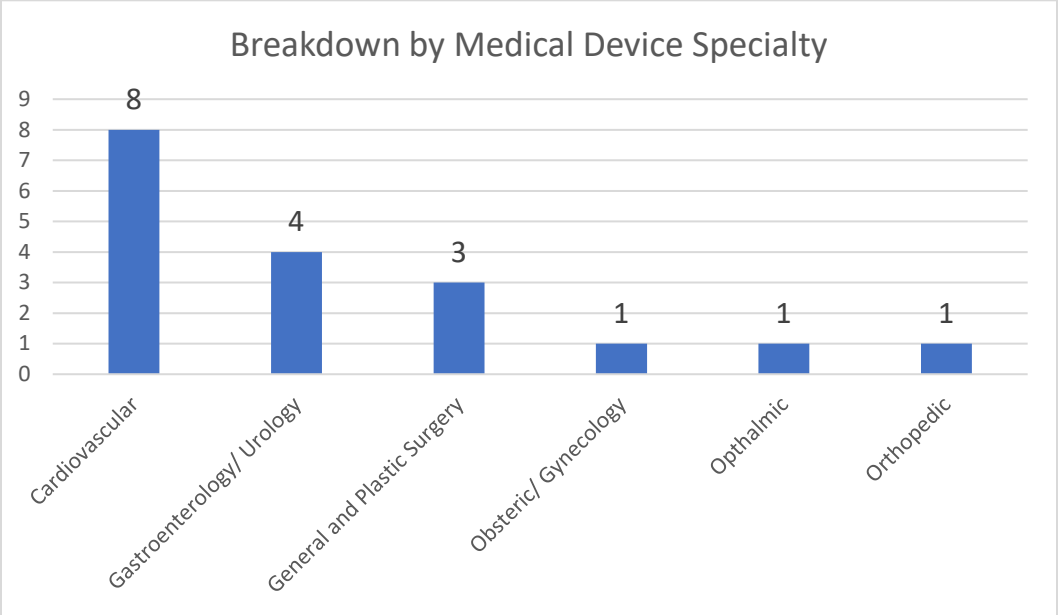


Figure 15 – Donated Labels by Device Specialty

6.1.2 Content & Formatting of Mock Labels

To create and modify the mock labels which comprised the test stimulus, Adobe Illustrator (Version 24; San Jose, CA), was used. Redesigned device labels, or "mock labels"

utilized a “Medical Device Facts” box similar to the “Drug Facts Label” (DFL). Drug Facts Labels are required for the vast majority of over the counter medications (OTCs) sold in the US (21 CFR part 201.66). Our novel “Device Facts” box contained all 11 medical device labeling inputs (Table 1) identified as required or typically included on medical devices labels through our assessment of the regulations and label review (objective 1). These 11 inputs were characterized using the analysis from the survey responses into “high, medium or low” risk groups (objective 2- See Table 5) and formatted accordingly to prioritize appropriate information (objective 3). Within the 11 Inputs, the name/identity of the medical device was consistently placed at the top of the mock labels.

Design rules (Table 6) were developed to provide a consistent frame for conversion of the commercial label into its “mock label” counterpart to be tested.

Table 6 – Mock Label Design Rules

Design Rules for Conversion from Commercial to Mock Labels
Content
All original content from commercial label will be present in the mock counterpart
Symbols utilized in the commercial label will be used in the mock counterpart and will be accompanied by identifying English text
A box titled "Device Facts" will contain the identified medical device labeling inputs, excluding Device Name, which will be placed above the box
Formatting
Medical device labeling inputs will be grouped by risk category and displayed in the following order, either Top to Bottom or Left to Right dependent on the original label format (vertical or horizontal): 1. High 2. Medium 3. Low
Risk Category Font Sizing Ratio: High: 200% Medium: 150% Low: 100% (Base) All fonts sizes on mock labels can vary, but above ratios must be maintained. Font size of low risk category is used as the base level size. Example: High: 14 pt. font Medium: 10.5 pt. font Low: 7 pt. font
Sans Serif or Equivalent Style (Helvetica, etc.) font used. Font is standard used for Drug Facts Label found on OTCs.
Latex/Natural Rubber Warnings and Sterility Status color coded as follows: Green: Sterile or Latex Free Red: Non-Sterile or Containing Latex/Rubber
Medical device labeling input title bold faced when applicable
Same size/footprint of label space as the commercial label

Work by Seo (2014) (2017) informed the design factors which were intended to enhance attention to critical information; specifically, components indicated by healthcare providers to result in high risk if misinterpreted or missed utilized design factors which Seo suggests garner attention (symbols, color, and grouping).

An example of a mock label, Device 6, following the Design Rules (Table 6) and a commercial label for the same device are shown in Figure 16Figure 17, respectively.

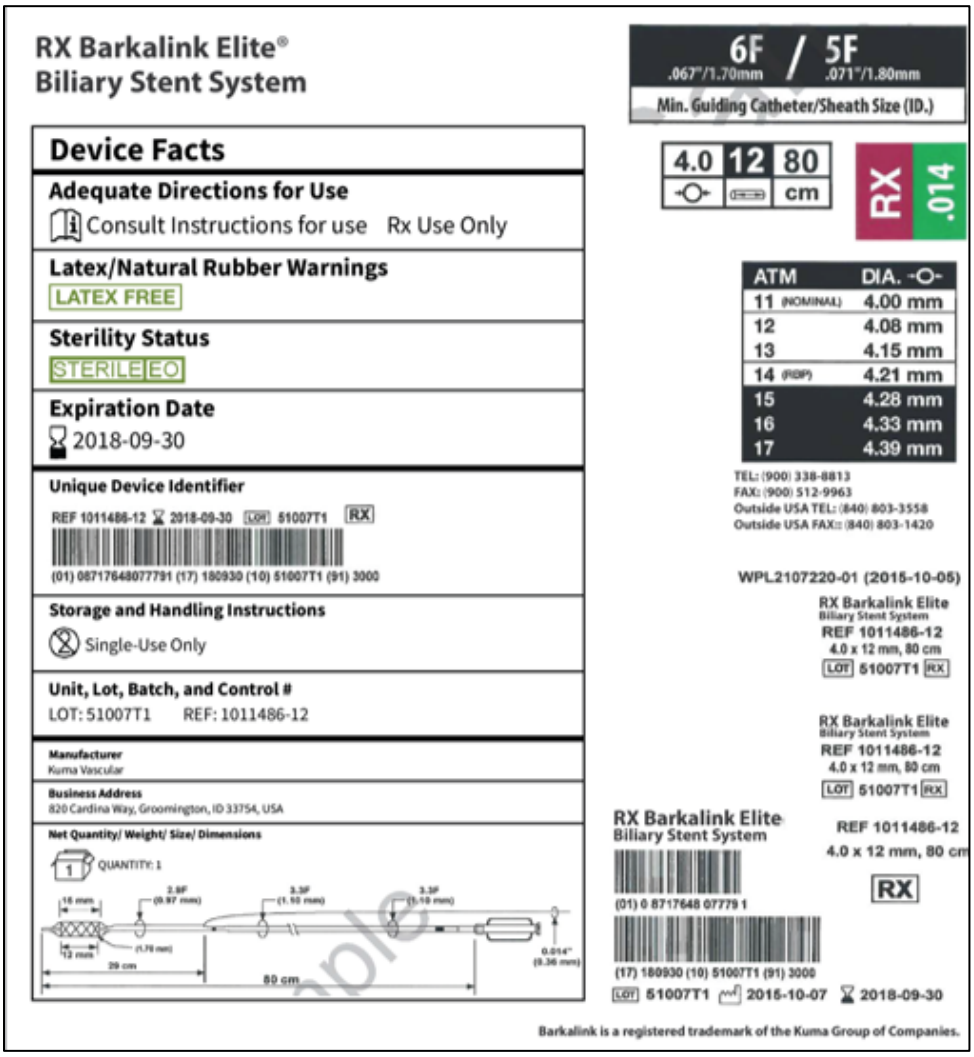


Figure 16 – Mock Label Example, Device 6

Figure 17 – Commercial Label Example

The mock labels applying the design rules presented Table 6 identified are presented in Figure 18-Figure 23.

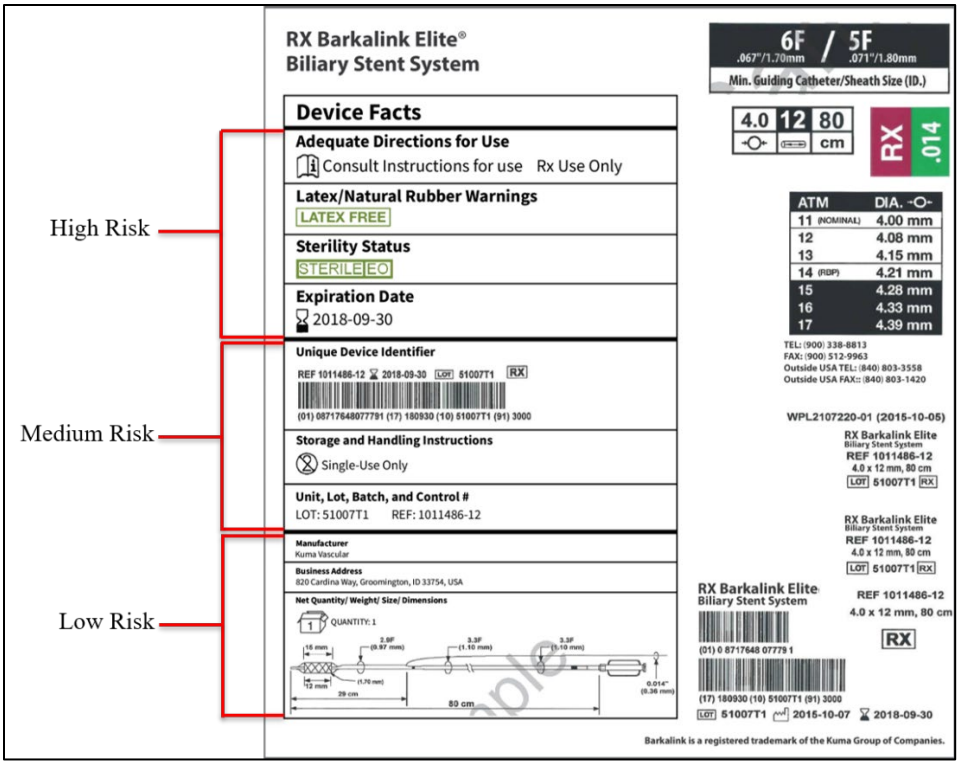


Figure 18 – Mock Label Example with Risk Categories Labeled

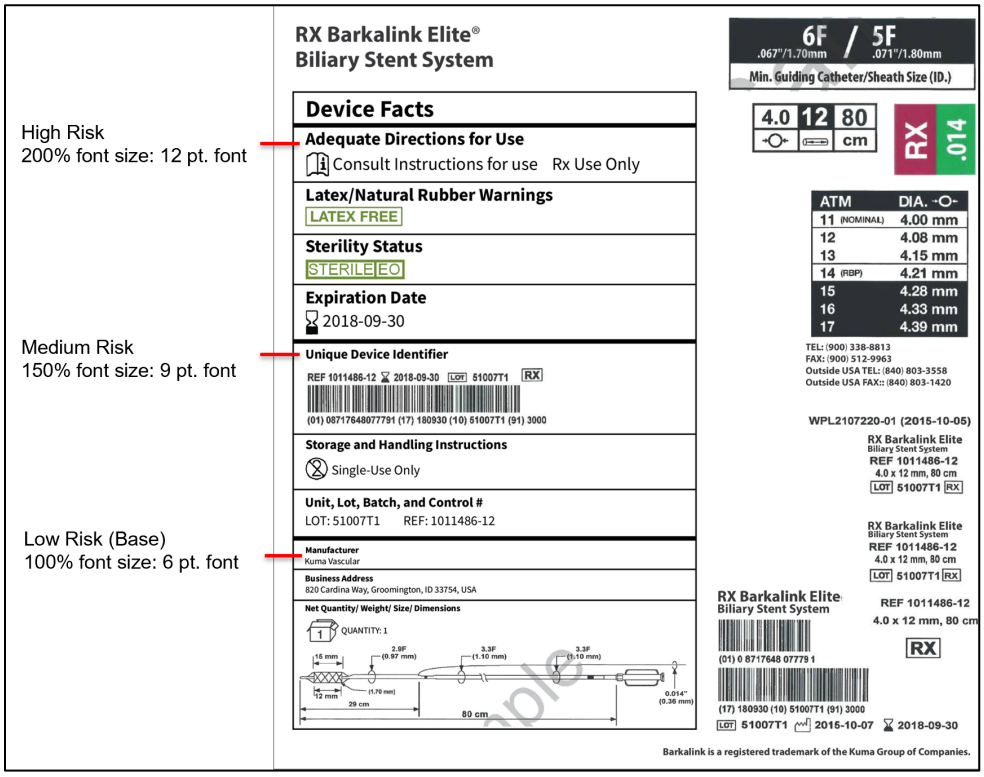


Figure 19 – Mock Label Example with Risk Category Font Size Ratios

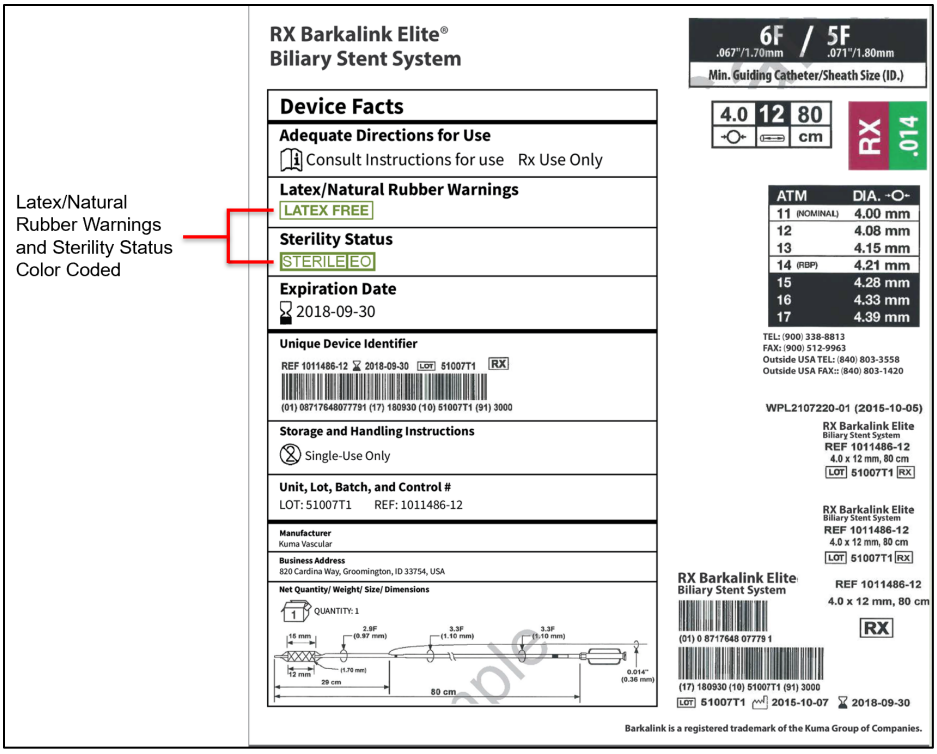


Figure 20 – Mock Label Example with Color Coded Latex/Natural Rubber Warnings and Sterility Status Labeled

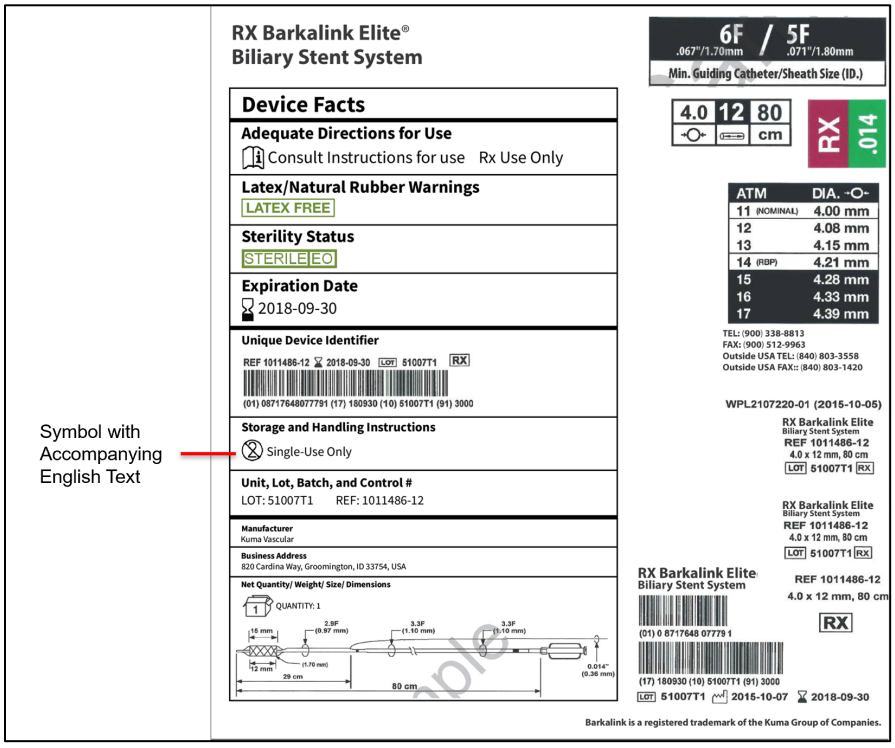


Figure 21 – Mock Label Example with Symbols and Accompanying English Text

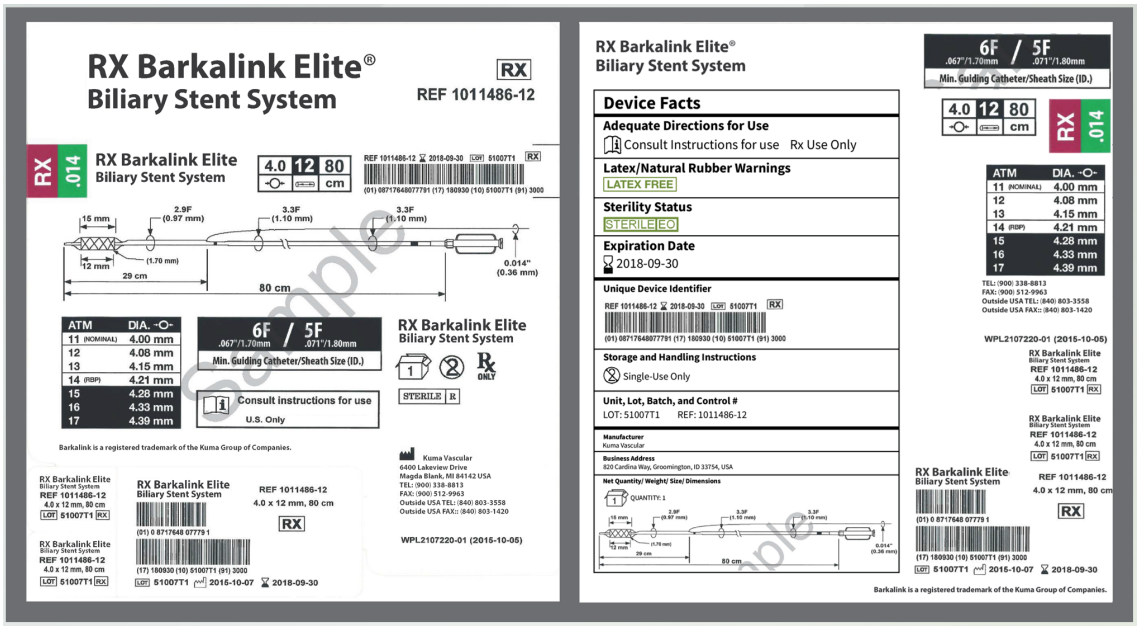


Figure 22 – Mock Label Example Showing Same Footprint as Commercial Label

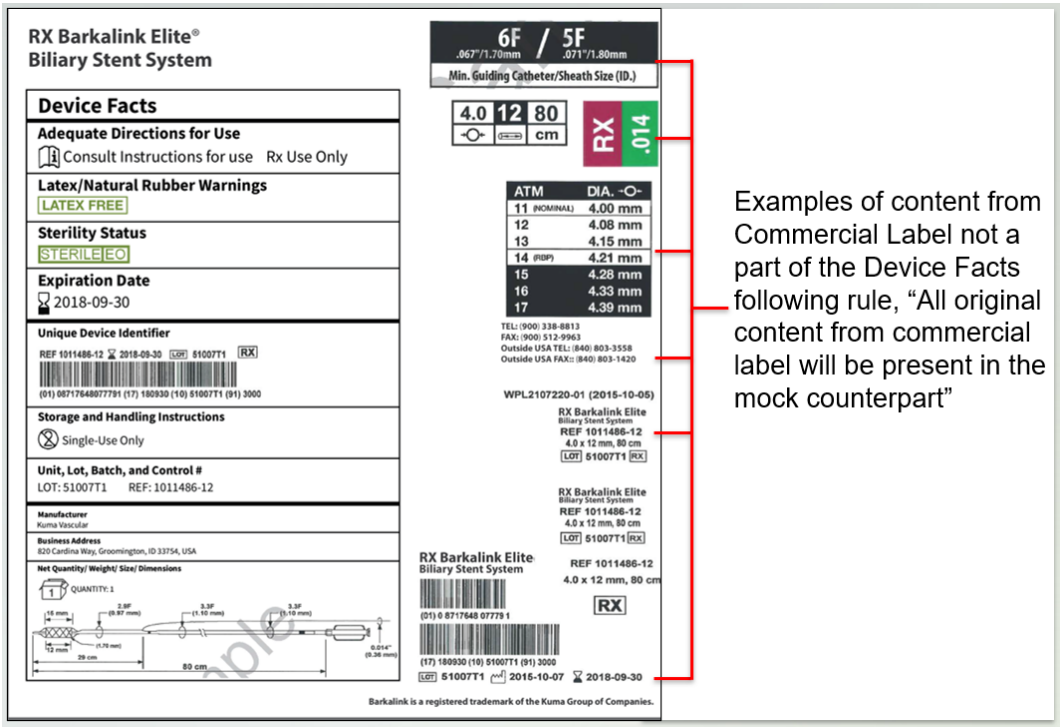


Figure 23 – Mock Label Example with Original Content, Non-Device Facts Box Information, Shown on Label

6.2 Results

A total of 18 mock labels were created using the guidelines outlined in Section 6.1 and summarized in Table 6. The complete gallery of mock labels are depicted in Figure 24–Figure 41. While 20 commercial labels were donated in total, only Devices 2-19 were used and, as such, the numbering sequence is reflected in the Figure captions shown below.

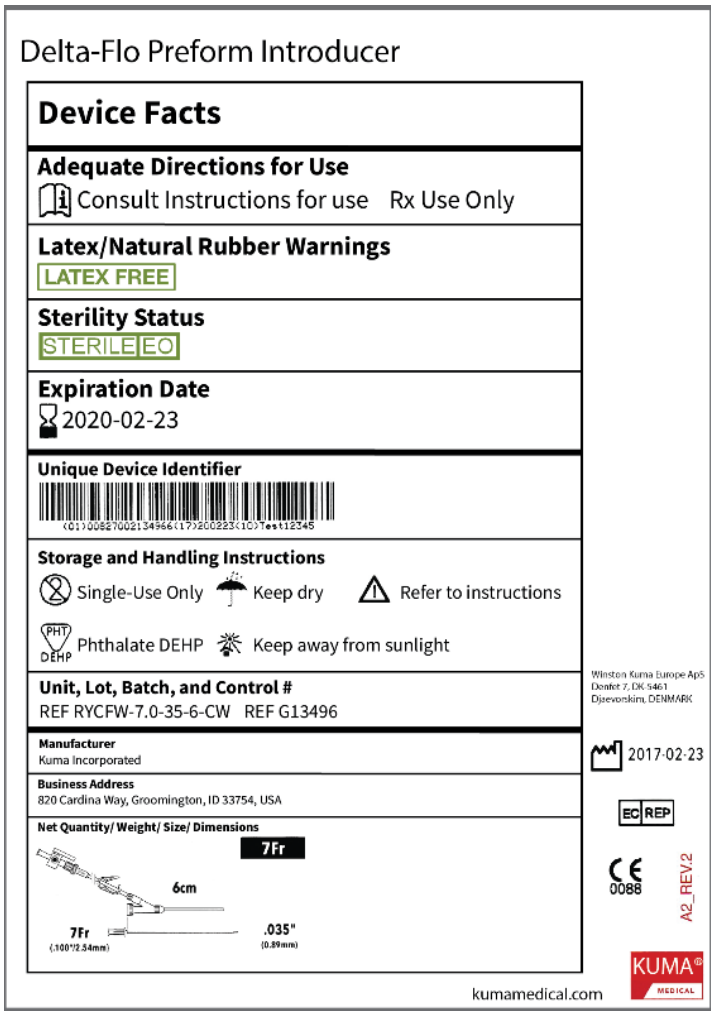


Figure 24 – Device 2 Mock Label

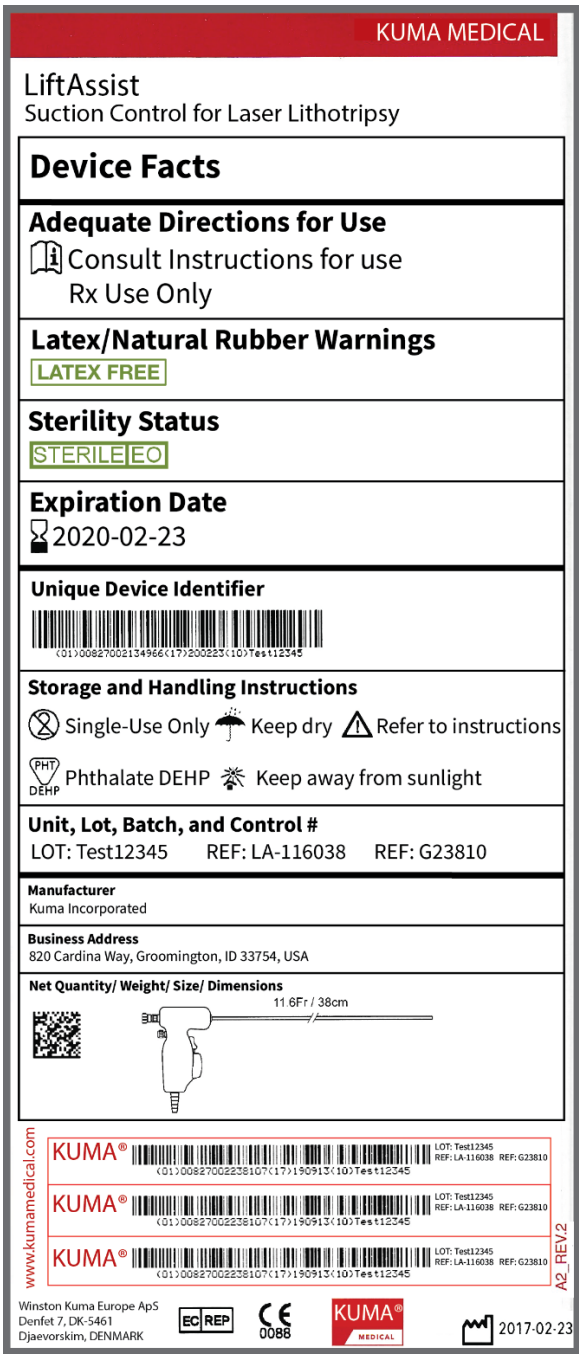


Figure 25 – Device 3 Mock Label

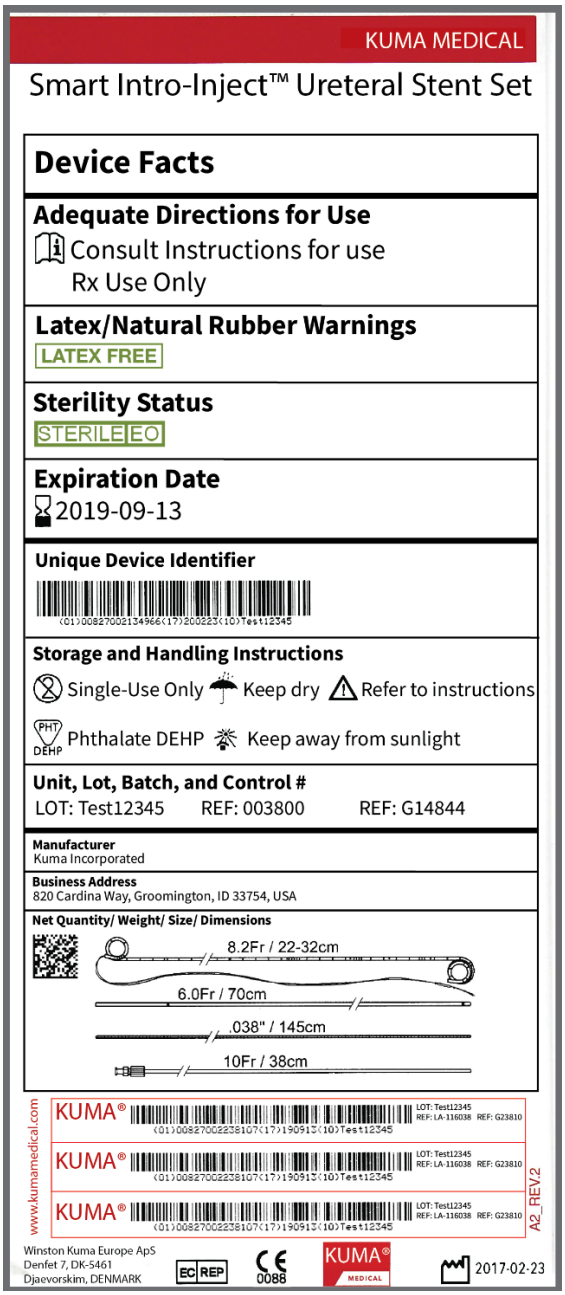


Figure 26 – Device 4 Mock Label

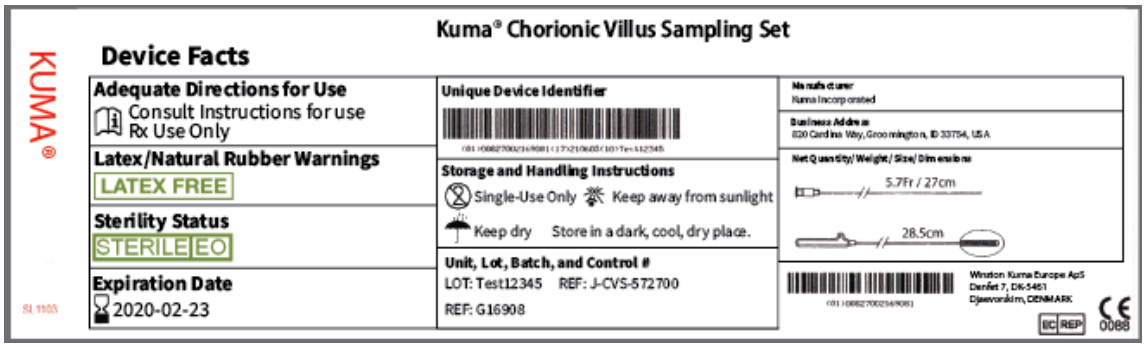


Figure 27 – Device 5 Mock Label

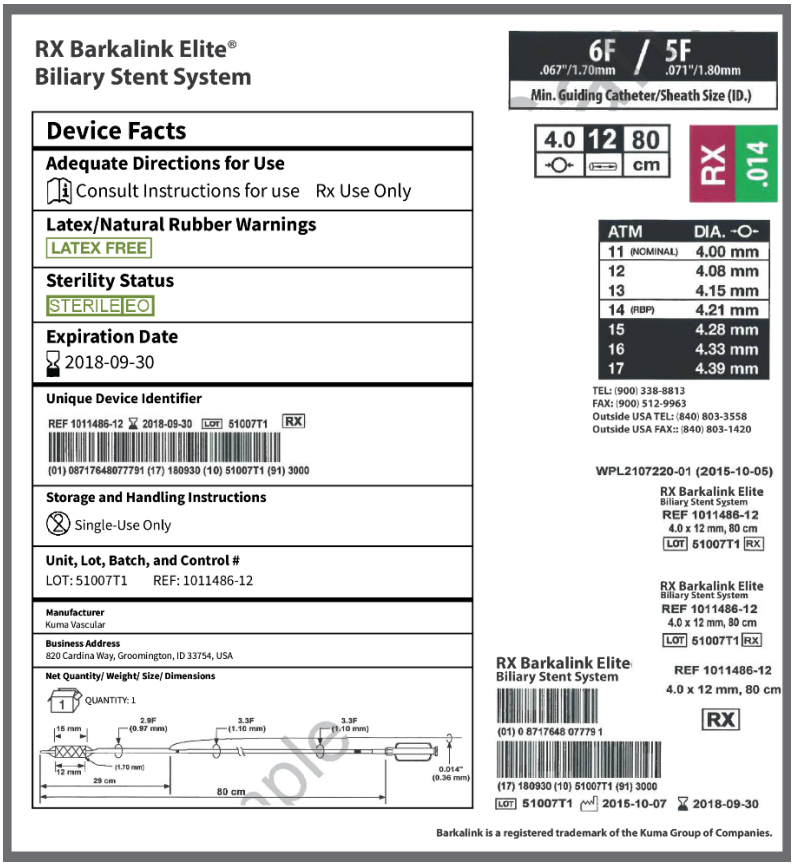


Figure 28 – Device 6 Mock Label

Figure 29 – Device 7 Mock Label

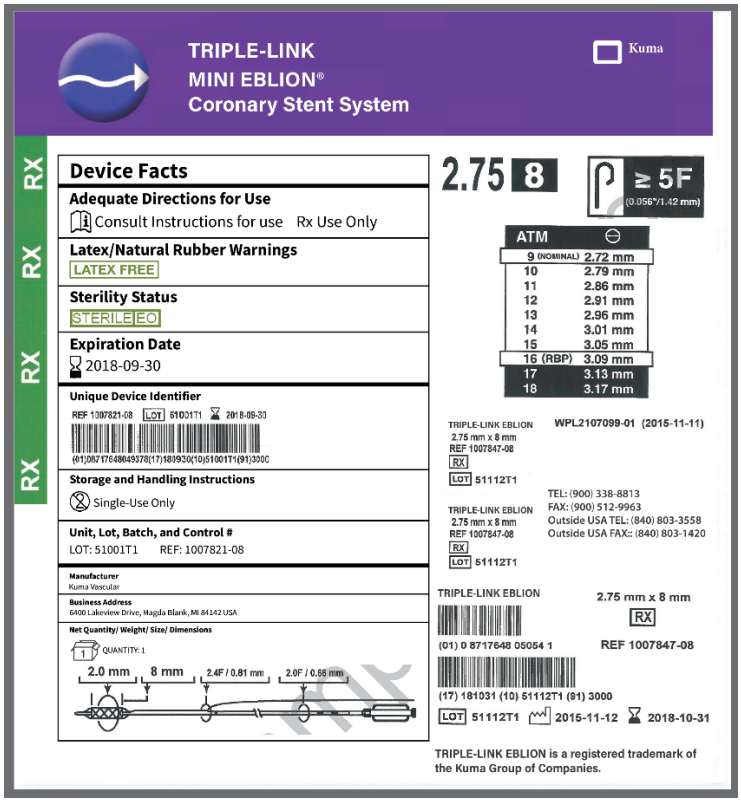


Figure 30 – Device 8 Mock Label

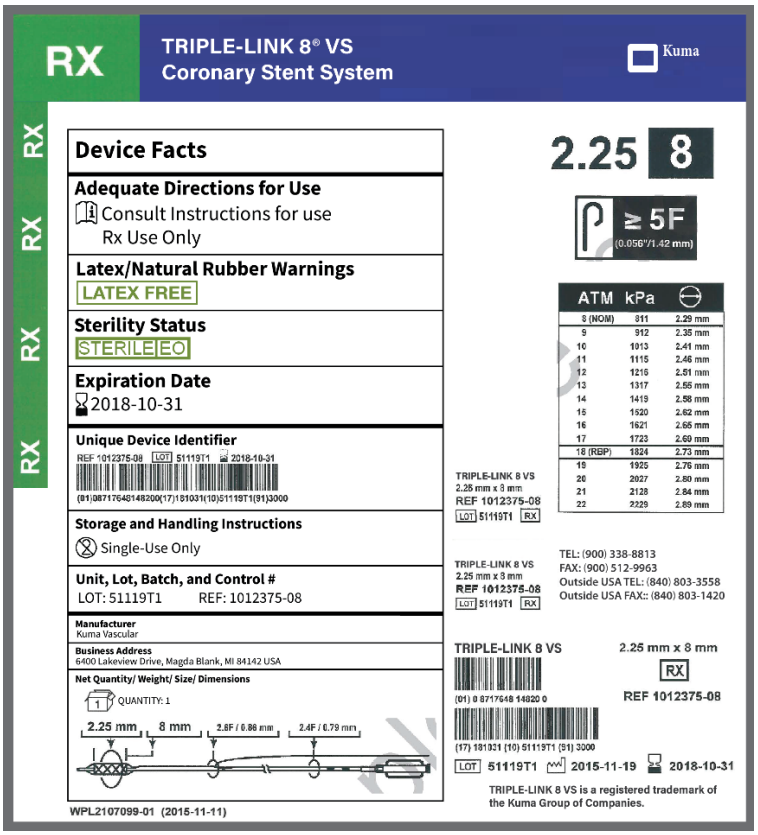


Figure 31 – Device 9 Mock Label

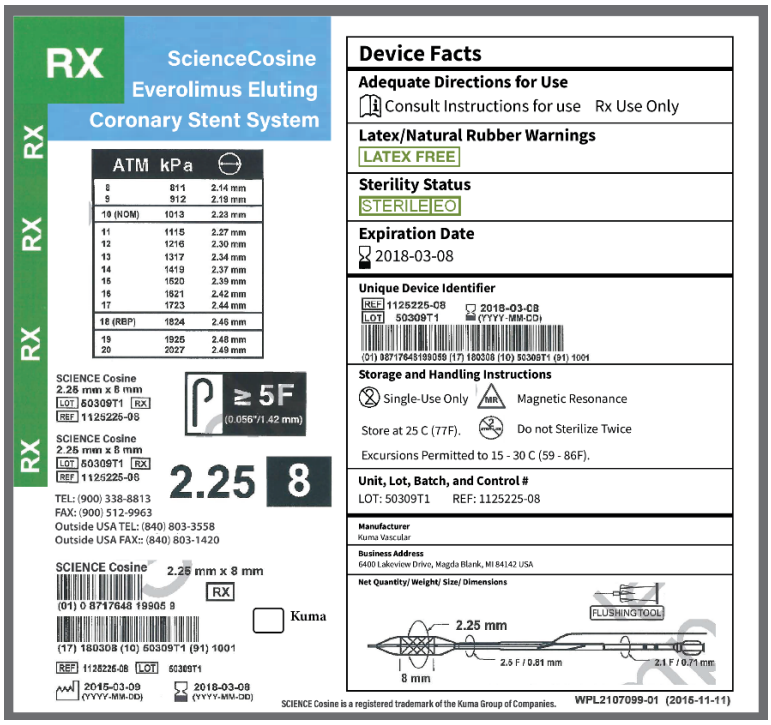


Figure 32 – Device 10 Mock Label

Figure 33 – Device 11 Mock Label

Figure 34 – Device 12 Mock Label

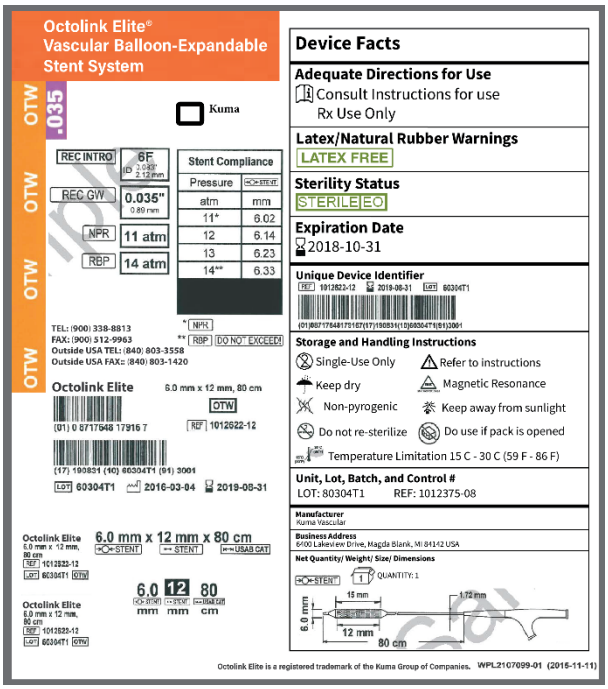


Figure 35 – Device 13 Mock Label

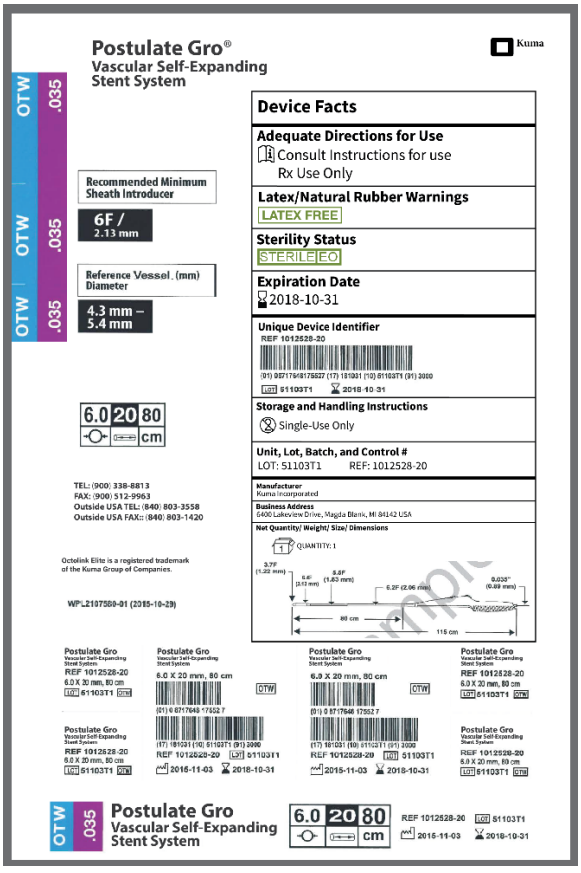


Figure 36 – Device 14 Mock Label

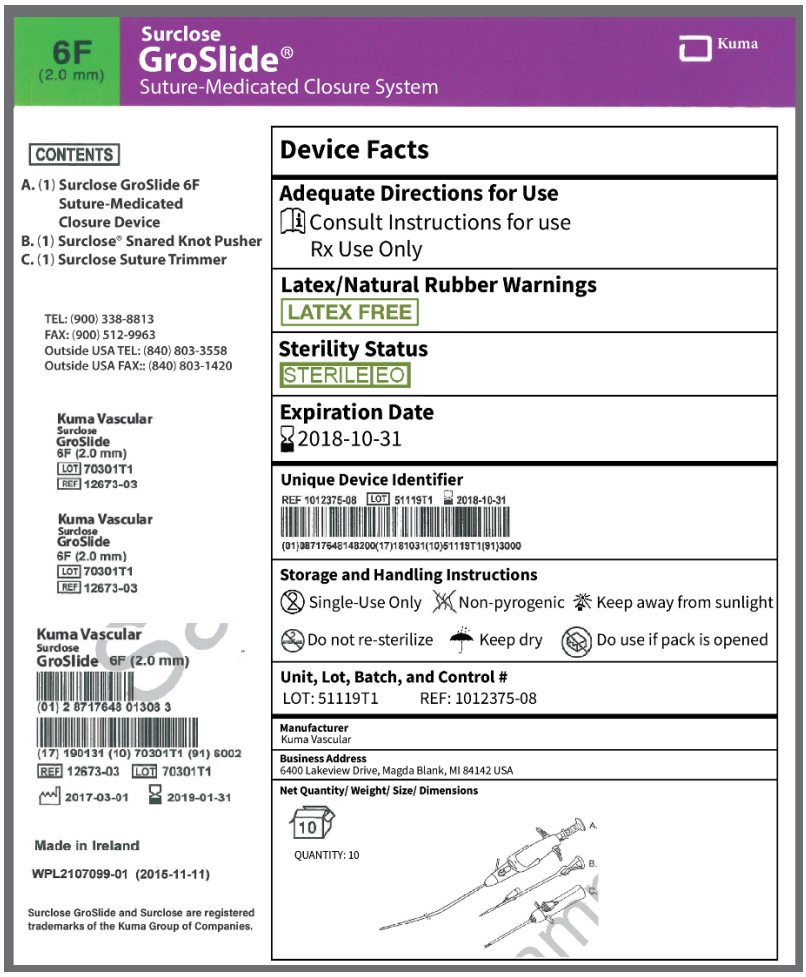


Figure 37 – Device 15 Mock Label

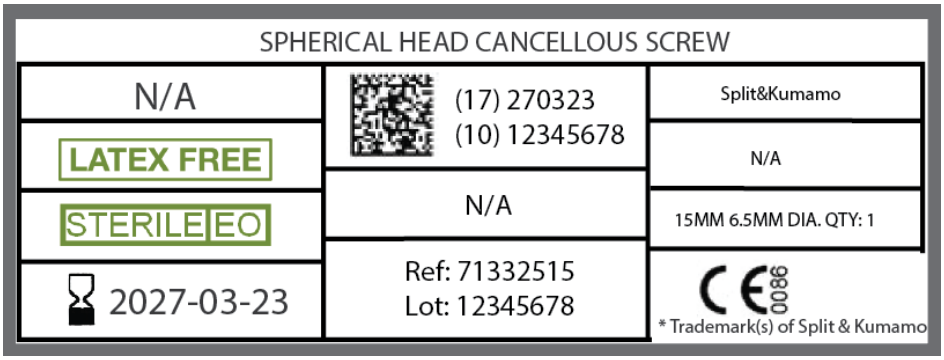


Figure 38 – Device 16 Mock Label

IOL Manipulator (Sinskey) Straight

Adequate Directions for Use N/A	**Unique Device Identifier** (01)00810090010847	**Manufacturer** Kuma Labs Premium Opthalmic Surgical Knives and Instruments
Latex/Natural Rubber Warnings LATEX FREE		**Business Address** 09190-B Landmark St. Blanco Sukamango, MA 80629-4749 USA
Sterility Status STERILE EO	**Storage and Handling Instructions** Single-Use Only	**Net Quantity/ Weight/ Size/ Dimensions** Quantity: 1
Expiration Date 2020-02-23	**Unit, Lot, Batch, and Control #** LOT: 00000XX REF: 127-03	CE 0483 Ph. (898) 370-9900 Fx. (898) 370-3370 www.kumalabs.com MADE IN U.S.A. Rev 03

PEEL HERE

Figure 39 – Device 17 Mock Label

Figure 40 – Device 18 Mock Label

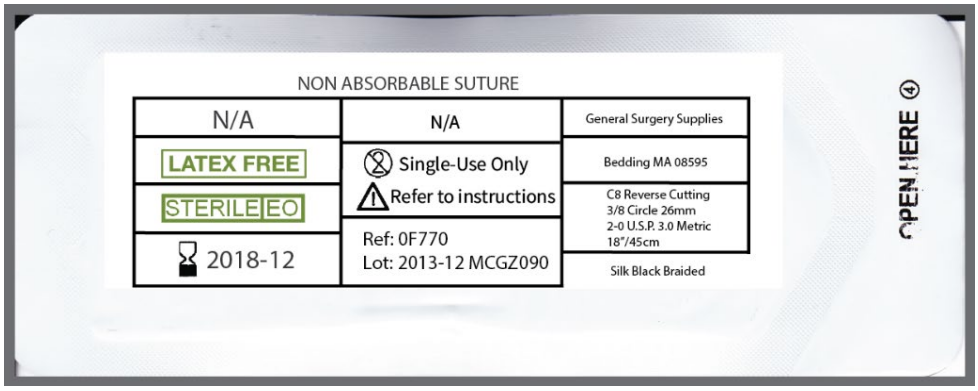


Figure 41 – Device 19 Mock Label

Chapter 7

Commercial vs. Mock Label Comparison

The mock labels that were created were objectively compared against their commercial forerunners utilizing a forced-choice task decision.

7.1 Methodology

7.1.1 Participants

The study was conducted using procedures approved under IRB # x17-1664e. Healthcare practitioners were recruited using a targeted email sent to surgeons, surgical technologists, and registered nurses within the Mid-Michigan region. An extra email blast was sent to 602 members of the Association of Surgical Technologists (AST) in the mid-Michigan area; specifically those identified by the AST database as having addresses in Ingham, Eaton, Livingston, Clinton, Jackson, and Shiawassee county. A research team also attended the Association of Surgical Technologist Michigan Assembly Fall 2018 Conference held in Mackinac City, Michigan on September 15, 2018, where conference participants were recruited through the distribution of an IRB approved flier, and interested, eligible attendees were tested on site. To participate in the study subjects had to be:

- At least 18 years of age
- A surgeon, surgical technologist, registered nurse, or related healthcare practitioner. Surgical Technologist/Registered Nurse.
- Students were permitted to participate in the event they had conducted a practical experience.

Upon arrival at an appointment (or as a walk up in the case of the conference testing), participants were given an IRB consent form to consider before proceeding. A member of the

research team explained the estimated time that the study would take, that the participant could skip any portion of the study or withdraw without consequence at any time and provided participants the opportunity to ask study-related questions. After signing the written consent document, participants began a paper-based questionnaire to collect basic demographics and job-related information. Following the consent form and questionnaire, participants completed a Forced Choice Task Decision Test at a computer workstation. $40 was given as a cash incentive to all participants. For advertisement, consent form, and pre-test questionnaire used for Forced Choice Task Decision Test, see Appendix D.

7.1.2 Materials and Forced Choice Task Decision Experimental Design

E-Prime 3.0 (Psychology Software Tools; Pittsburgh, PA) was used to create a program that ran as a two-alternative, forced-choice decision task (2 AFC) (in support of Objective 4). Adobe Illustrator (Version 24; San Jose, CA), was used to create the trials; each trial included two images of the same label at the same level of treatment (mock or commercial) on the computer screen. These two labels were identical except for 1 altered piece of information (e.g. one was sterile, the other was not; one expired one had not; one had latex, one did not, etc.). A prompt was given at the top of the screen to instruct the participant to select one of the images based on the labeling input that was altered, i.e. “Please select the device that is sterile”. Figure 42 and Figure 43 depict a trial presented in the commercial and mock label formats, respectively.

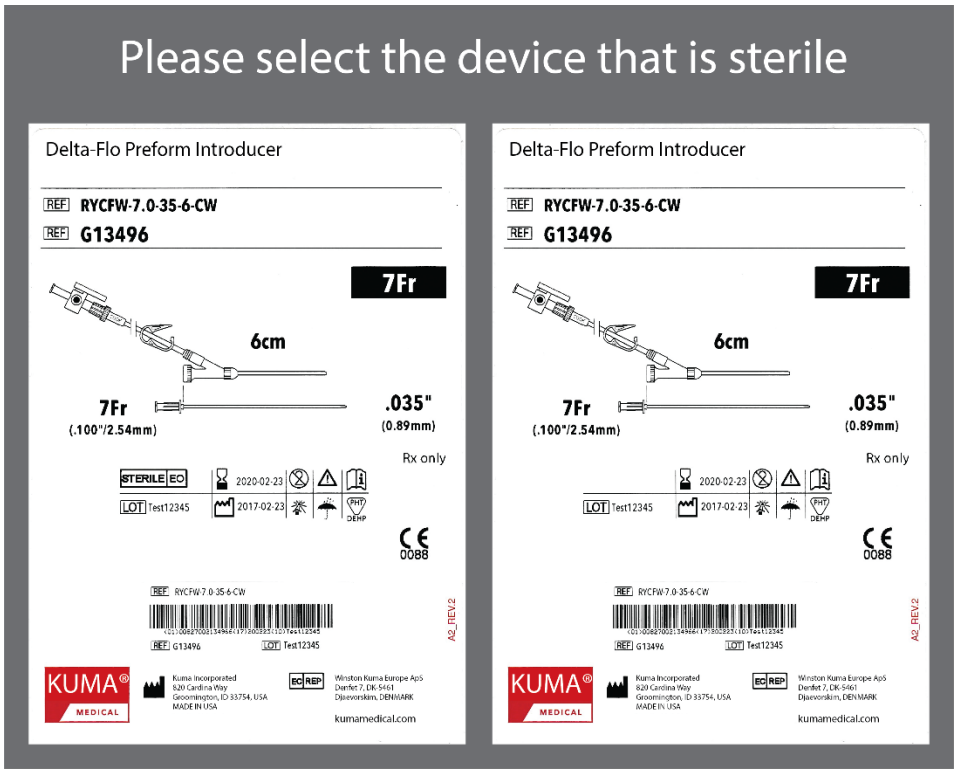


Figure 42 – Commercial Label Forced Choice Trial Example

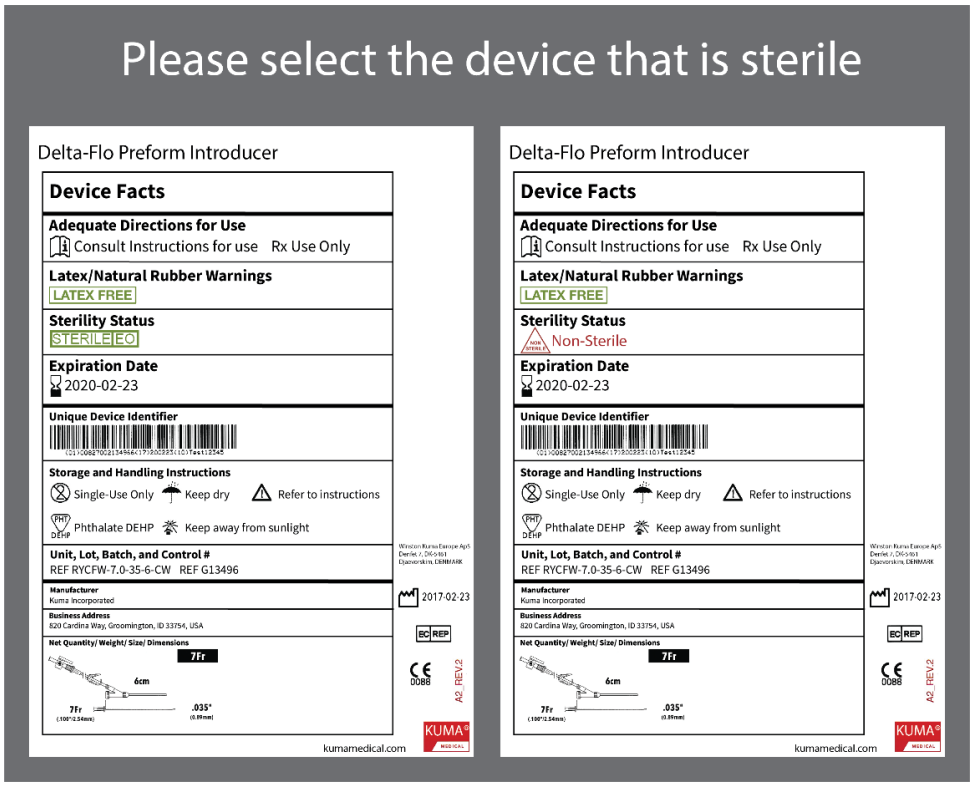


Figure 43 – Mock Label Forced Choice Trial

A total of 216 trials were created for the study which was completed by every participant. The total breakdown of trials follows:

Trials = Device Label (18 labels) × Label Version (2 versions-mock verse commercial) × Labeling Input (2 from each risk category- high, medium, and low (total of six))

Where:

> **18 Device Labels** – Corresponding to the donated device labels selected for use (see Figure 24, Figure 41 and Appendix C)
>
> **2 Label Versions** – Corresponding to commercial or mock designs
>
> **6 Labeling Inputs** – Corresponding to the specific labeling input being questioned. Two labeling inputs were used from each risk category.

For the specific combination of each trial, see Appendix E. For counterbalancing/blocking purposes, a total of four tests were created, each test being a combination of counterbalancing to mitigate run-order effect of test trials and blocking to mitigate positional effect of the correct choice for trials. A participant completed one of the following tests:

1. Even × Part A
2. Odd × Part A
3. Even × Part B
4. Odd × Part B

For the counterbalancing to mitigate run-order, denoted by either Part A or Part B. The trials were split into two sets, with trials from half of the devices tested placed into one while the remaining trials were placed into a second set as depicted in Table 7 and Table 8. Trials 1-6 and

121-126 were reserved for Device 1 while Trials 115-120 and 235-240 were reserved for Device 20. Commercial trials created from Device 1 were used in the test as part of the training sequence mentioned later.

For the study, Part A participants first completed all trials (randomized) from Table 7, took a break, and then completed all trials (randomized) from Table 8. The same was done for Part B participants in reverse, starting with Table 8 and then completing the study with Table 7.

Table 7 – Forced Choice Task Part A First Set/Part B Second Set

Device #	Commercial Trial #	Mock Trial #
Device 2	Trials 7-12	Trials 127-132
Device 3	Trials 13-18	Trials 133-138
Device 6	Trials 31-36	Trials 151-156
Device 8	Trials 43-48	Trials 163-168
Device 10	Trials 55-60	Trials 175-180
Device 12	Trials 67-72	Trials 187-192
Device 14	Trials 79-84	Trials 199-204
Device 17	Trials 97-102	Trials 217-222
Device 19	Trials 109-114	Trials 229-234

Table 8 – Forced Choice Task Part A Second Set/Part B First Set

Device #	Commercial Trial #	Mock Trial #
Device 4	Trials 19-24	Trials 139-144
Device 5	Trials 25-30	Trials145-150
Device 7	Trials 37-42	Trials 157-162
Device 9	Trials 49-54	Trials 169-174
Device 11	Trials 61-66	Trials 181-186
Device 13	Trials 73-78	Trials 193-198
Device 15	Trials 85-90	Trials 205-210
Device 16	Trials 91-96	Trials 211-216
Device 18	Trials 103-108	Trials 223-228

A high-level representation is given in Figure 44 with a more detailed representation of specific trials for each group shown in Appendix F.

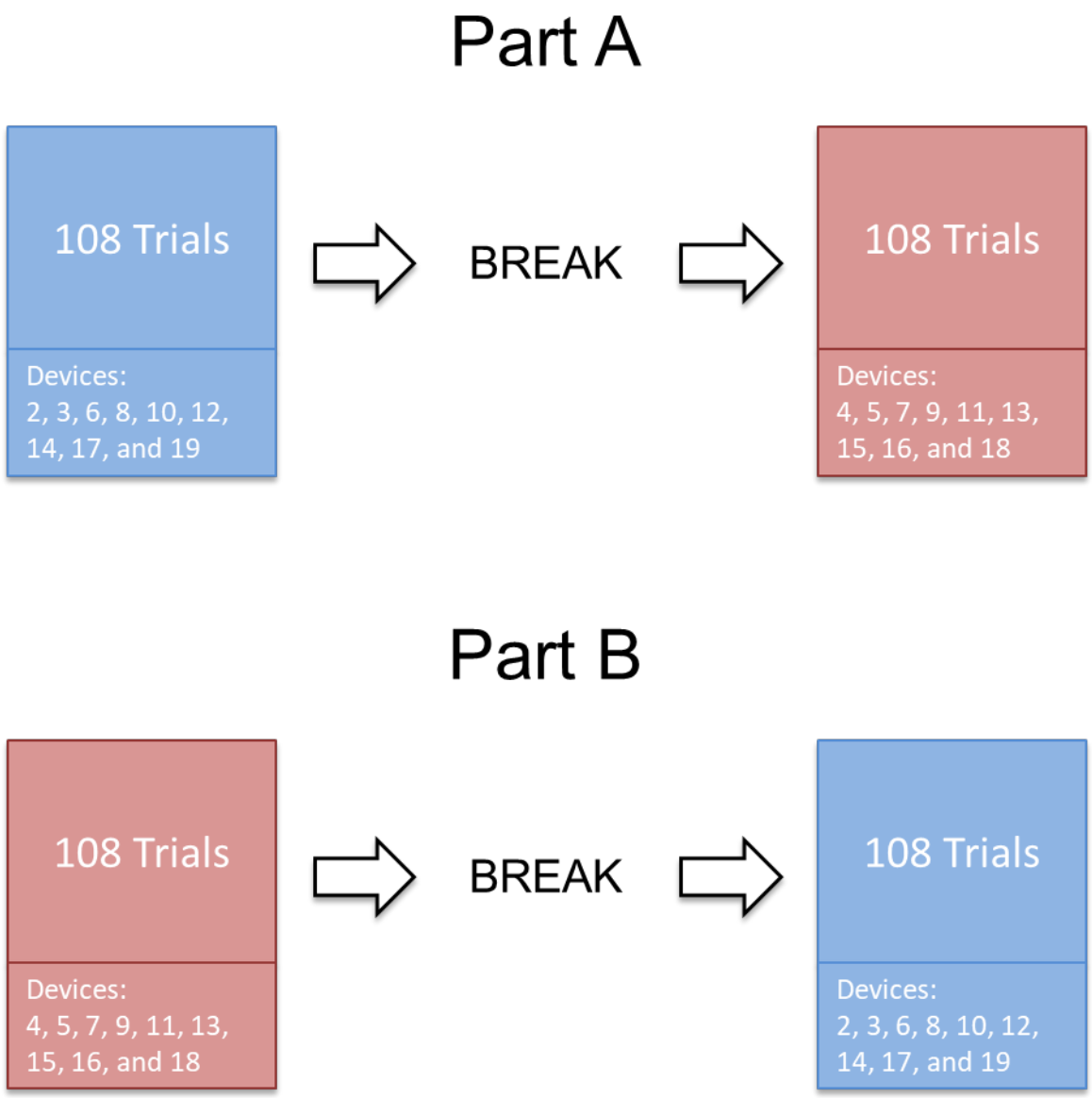


Figure 44 – Part A/Part B Counterbalance

For the blocking to mitigate positional effects related to the side of the correct choice, this meant having one group of participants having the correct choice appear on one side (Left or Right / Top or Bottom) while the other group of participants had their correct position (for the same trial comprised of a given device question combination) appear in the other location. These groups were named Even and Odd. A high-level example is given in Figure 45 (Note this

example is for visualization purposes only and does not represent the actual correct choice positioning used) with the actual representation of specific trials for each group refer to Appendix G.

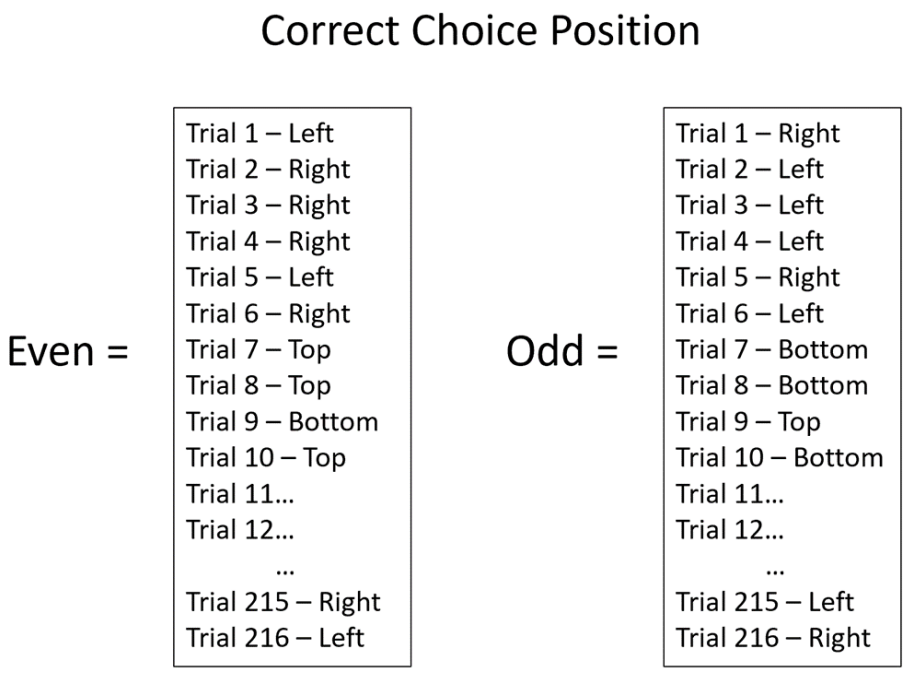


Figure 45 – Even/Odd Blocking Example

Prior to the start of the test, a training sequence was completed utilizing 3 commercial trials created from device 1, the device that was not included in the analyzed data. For this training sequence, directions were provided on the computer screen and verbally as well. Participants were instructed to select the correct image using the arrow keys (left for selecting the left image and right for selecting the right image, and top and bottom labels utilizing the up and down arrows as fast as possible; the program was created to "timeout" at 45 seconds from initial presentation of each slide if the participant had not yet made a selection. Trials that resulted in a "timeout" were treated as an incorrect image selection. Upon completion of the

training sequence, research participants were provided another opportunity to ask any questions about the testing or seek clarification from a member of the research team.

Because trials depicted a given label treatment for both of the presented stimulus (i.e. both mock or both commercial within a trial), comparing the dependent variables (accuracy of selection/time to correct selection) across the trials enabled us to objectively evaluate the performance of the mock designs relative to the existing commercial standards.

7.1.3 Statistical Model

Statistical analysis of data was completed utilizing R. The main focus was on the effect of two independent variables, label version (mock vs. commercial) and risk category (high vs. medium vs. low). The response outcomes were the proportion of correct responses (binary variable) and correct response time (continuous variable).

The response outcome, proportion of correct responses (the binary variable), was fitted to a generalized linear mixed model. As the data was binary (correct/incorrect) the results were expressed in the probability of having a correct response in terms of log odds. Log odds are defined as:

$$\log\left[\frac{p}{1-p}\right]$$

Where P is the probability of the participant giving a correct response. The log odds have a positive correlation to the probability of having a correct response; in other words, as the log odds value increases, so does the probability of a correct response. The same occurs in the inverse, where the log odds value decreases, so does the probability of a correct response. The full, generalized linear mixed model used for the analysis of proportion of correct responses is shown below:

$$\log\left[\frac{p}{1-p}\right]_i = a + a_s + a_y + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + e_i$$

Where:

$\log[p/1\text{-}p)]_i$ = probability of Correct Response in log odds

a = intercept for entire model

a_s = intercept for subject (Random Effect)

a_y = intercept for device label (Random Effect)

b_1 = slope for label version (Fixed Effect)

x_{1i} = x value for label version (Fixed Effect)

x_{1i} = x value for label version (Fixed Effect)

b_2 = slope for risk category (Fixed Effect)

x_{2i} = x value for risk category (Fixed Effect)

b_3 = slope for interaction of label version and risk category

$x_{1i}x_{2i}$ = x value for interaction of label version and risk category

e_i = random variance

Label version (one of a total of two possibilities; commercial or mock) and risk category (one of a total of three possibilities; high, medium, or low) were considered fixed effects while subject (one of a total of forty-two possibilities; participants in study) and device label (one of a total of eighteen possibilities; donated labels from companies) were considered random effects. The interaction of label version and risk category was also included in the model.

The response outcome, time to correct response (a continuous variable) was fitted to a linear mixed model. To address normality assumptions, the correct response time data was first log transformed before being fitted to the model. The full linear mixed model used for the analysis of correct response time is shown below:

$$y_i = a + a_s + a_y + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + b_4 x_{3i} + e_i$$

Where:

y_i = specific value for correct response time

a = intercept for entire model

a_s = intercept for subject (Random Effect)

a_y = intercept for device label (Random Effect)

b_1 = slope for label version (Fixed Effect)

x_{1i} = x value for label version (Fixed Effect)

b_2 = slope for risk category (Fixed Effect)

x_{2i} = x value for risk category (Fixed Effect)

b_3 = slope for interaction of label version and risk category

$x_{1i}x_{2i}$ = x value for interaction of label version and risk category

b_4 = slope for correct response position (Fixed Effect)

X_{3i} = x value for correct response position (Fixed Effect)

e_i = random variance

Label version (one of a total of two possibilities; commercial or mock), risk category (one of a total of three possibilities; high, medium, or low), and correct response position (one of a total of two possibilities; Even or Odd subgroup) were considered fixed effects while subject (one of a total of forty-two possibilities; participants in study) and device label (one of a total of eighteen possibilities; donated labels from companies) were considered random effects. The interaction of label version and risk category was also included in the model.

To analyze the role of risk categories within label version, post-hoc testing was performed using emmeans in R for pairwise comparisons utilizing odds ratio for the proportion of correct responses model and as well as the time to correct response.

7.2 Results

7.2.1 Characterization of participants

As with the survey testing, participants enrolled in the Forced Choice Task Decision testing were characterized demographically and their work history was collected. Demographic information is presented in Figure 46-Figure 48. The participants were mainly female (n=33; 79%) with a majority reporting their occupation as a "Certified Surgical Technologist" (n=37; 88%). When asked about years of experience, 10+ years of experience was reported more frequently than any other category (n=16; 38%).

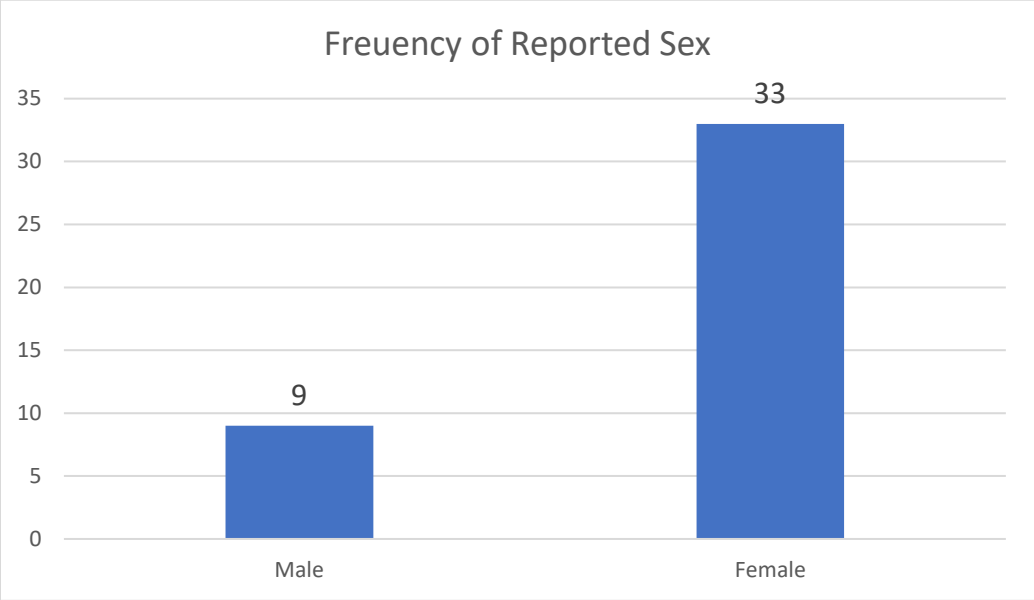


Figure 46 – Forced Choice Participant Frequency of Reported Sex

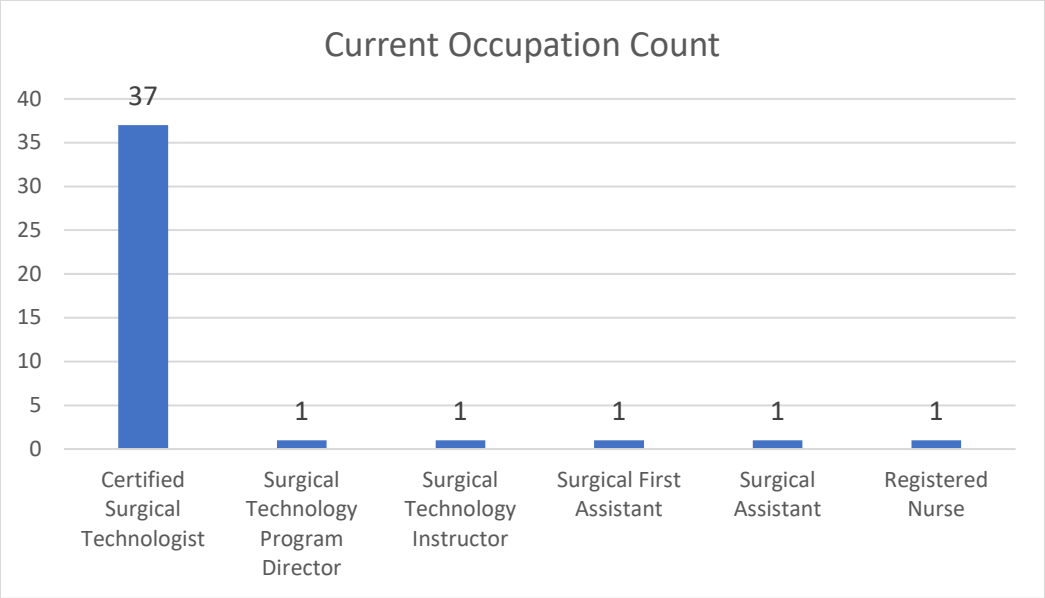


Figure 47 – Forced Choice Participation Current Occupation Count

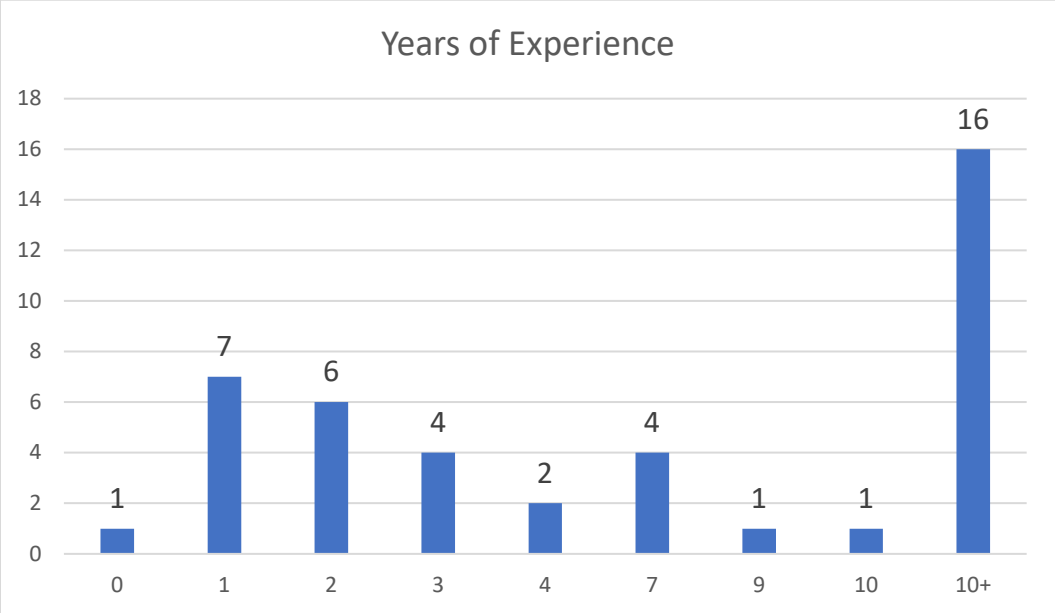


Figure 48 – Forced Choice Participant Years of Experience

Beyond basic demographics, participants also were characterized by their vision both through a Near Point Visual Acuity test, as well as a Color Differentiation Ability test. The results of these tests are presented in Figure 49 and Figure 50. A majority of the participants had some level of

vision loss; n=34 (81%) of participants were measured to have less than 20/20 vision, with only 8 out of 42 participants having 20/20 vision. In contrast, the results of the Color Differentiation Ability test identified no participants with a measured color deficiency.

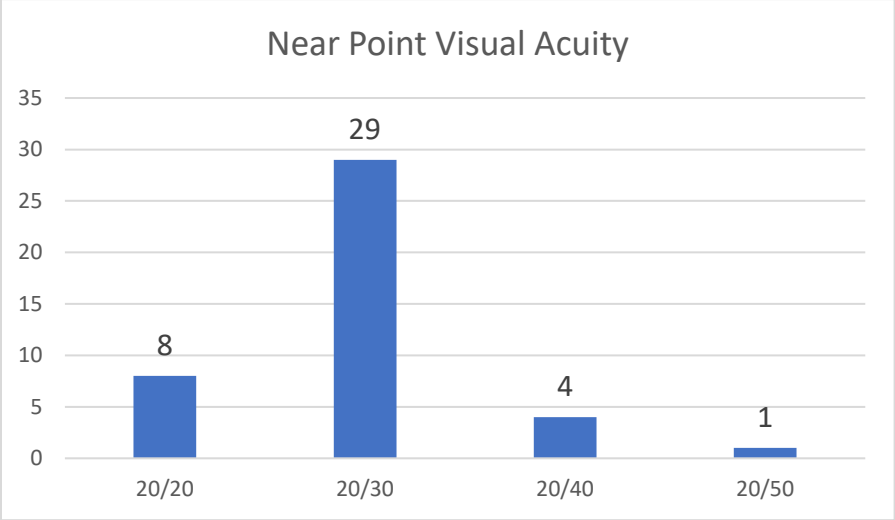


Figure 49 – Near Point Visual Acuity Test

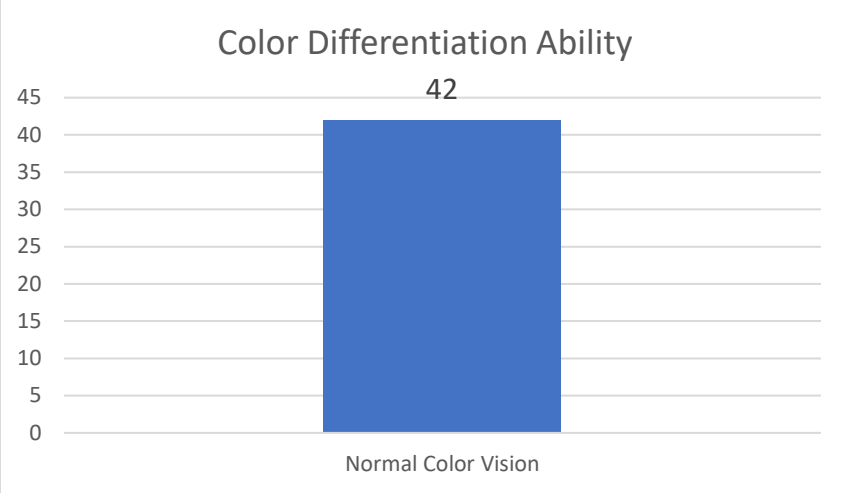


Figure 50 – Color Differentiation Ability Test

7.2.2 Forced Choice Task Decision Results and Analysis

Estimates representing the proportion of correct responses (binary variable) are shown by risk category and label version in Figure 51 and Table 9 with a confidence of 95%. Odds Ratio significance, between label versions, has been included in Figure 51.

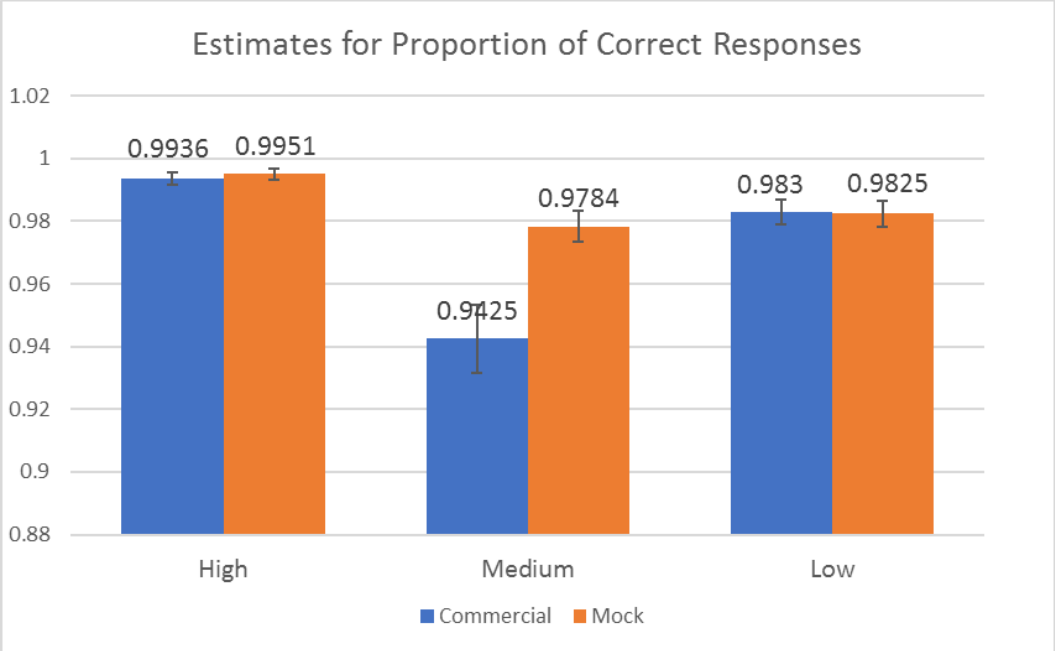


Figure 51 – Estimates for Proportion of Correct Responses

Table 9 – Estimates for Proportion of Correct Responses Tabulated

Label Version and Risk Category	Probability	Standard Error	LCL	UCL
Mock, High	0.9951	0.001737	0.9902	0.9975
Comm., High	0.9936	0.002055	0.9880	0.9966
Mock, Medium	0.9784	0.004885	0.9665	0.9862
Comm., Medium	0.9425	0.010778	0.9174	0.9603
Mock, Low	0.9825	0.004158	0.9722	0.9891
Comm., Low	0.9830	0.004066	0.9729	0.9894

To determine if significant differences existed between specific combinations of label version (commercial/mock) and risk category (high/medium/low), pairwise comparisons using odds ratio were run. For odds ratio, the following rules applied (Szumilas, 2010).

$$\text{Odds Ratio} = \frac{\text{Item A}}{\text{Item B}}$$

Results are interpreted for odds ratio reporting as follows:

Odds Ratio = 1 No difference in the proportion of correct responses odds between Item A and Item B

Odds Ratio > 1 Item A has higher odds of having a correct response compared to Item B

Odds Ratio < 1 Item A has lower odds of having a correct response compared to Item B

Comparisons within risk categories between their commercial and mock versions can be seen in Table 10 and Figure 52 and comparisons of risk categories within each of the commercial and mock versions can be seen in Table 11 and Figure 53.

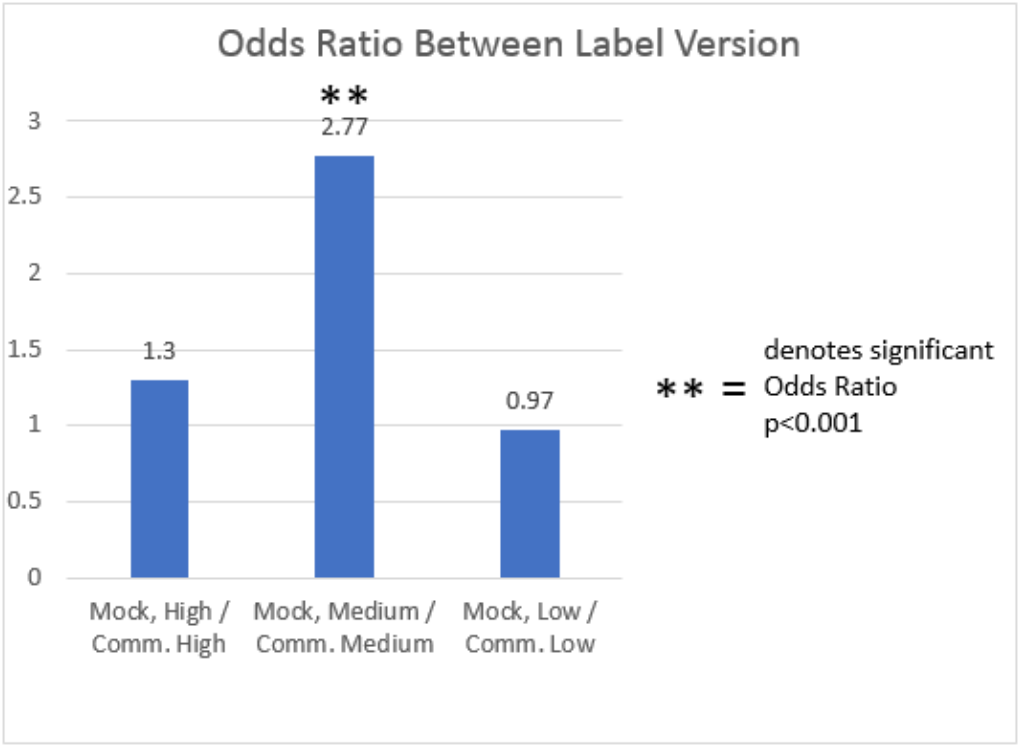


Figure 52 – Pairwise Comparison, Risk Categories Between Label Versions, Proportion of Correct Responses

Table 10 – Pairwise Comparisons, Risk Categories Between Label Versions, Proportion of Correct Responses

Risk Category Between Label Version	Odds Ratio	Standard Error	Z-Ratio	P-Value	Interpretation
Mock, High / Comm., High	1.30	0.537	0.646	0.9875	No significant difference is apparent when the odds of a correct response for a high risk labeling inputs on the mock labels are compared to a high risk labeling inputs on the commercial labels
Mock, Medium / Comm., Medium	**2.77**	**0.503**	**5.600**	**<0.0001**	**The odds of having a correct response related to a medium risk labeling inputs on a mock labels are significantly greater than having a correct response on a medium risk labeling inputs on a commercial label**
Mock, Low / Comm., Low	0.97	0.232	-0.126	1.0000	No significant difference is apparent when the odds of a correct response for a low risk labeling input on the mock labels are compared to low risk labeling inputs on the commercial labels

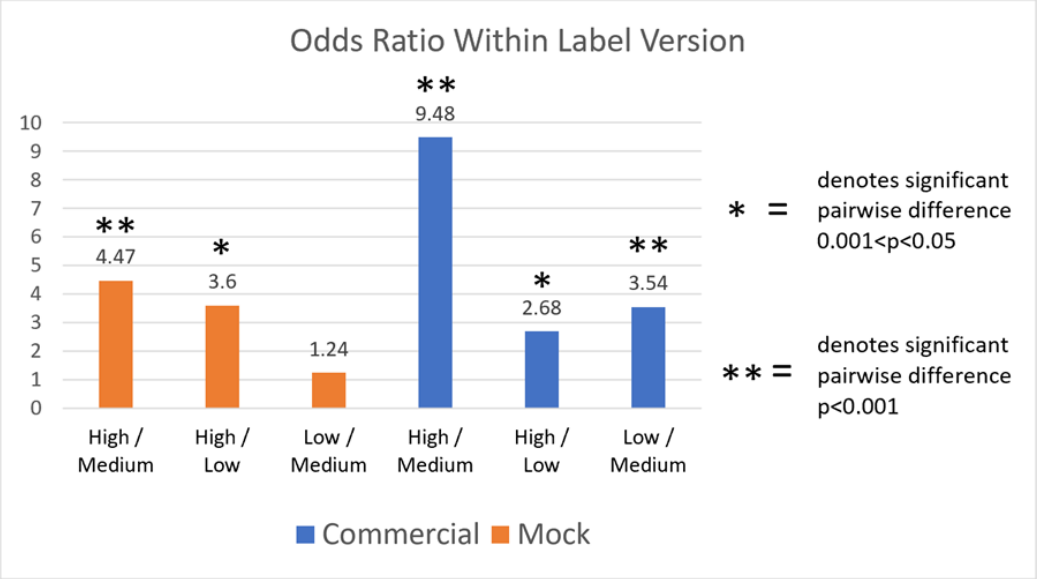


Figure 53 -Pairwise Comparisons, Risk Categories Within Label Versions Proportion of Correct Responses

Table 11 – Pairwise Comparisons, Risk Categories Within Label Versions, Proportion of Correct Responses

Risk Category Within Label Version	Odds Ratio	Standard Error	Z-Ratio	P-Value	Interpretation
Mock, High / Mock, Medium	**4.47**	**1.541**	**4.342**	**0.0002**	**The odds of having a correct response related to a high risk labeling inputs are significantly greater than having a correct response related to medium risk labeling inputs on mock label versions**
Mock, High / Mock, Low	**3.60**	**1.268**	**3.642**	**0.0037**	**The odds of having a correct response related to high risk labeling inputs are significantly greater than having a correct response on a low risk labeling input on mock label versions**
Mock, Low / Mock, Medium	1.24	0.281	0.949	0.9336	No significant difference is apparent when the odds of a correct response for a low risk labeling input is compared to a medium risk labeling input on mock label versions
Comm., High / Comm., Medium	**9.48**	**2.745**	**7.767**	**<.0001**	**The odds of having a correct response related to a high risk labeling input is significantly greater than having a correct response on a medium risk labeling input on commercial label versions**
Comm., High / Comm., Low	**2.68**	**0.860**	**3.073**	**0.0258**	**The odds of having a correct response related to a high risk labeling input is significantly greater than having a correct response on a low risk labeling input on commercial label versions**
Comm., Low / Comm., Medium	**3.54**	**0.697**	**6.407**	**<.0001**	**The odds of having a correct response related to a low risk labeling input is significantly greater than having a correct response on a medium risk labeling input on mock label versions**

For pairwise comparisons between label versions (see Table 10), the only significant difference relates to medium risk labeling inputs when a mock version compared is compared to medium risk labeling inputs on commercial versions. This is possibly attributed to symbol usage mainly being used for medium risk labeling inputs. In the Design Rules provided in Table 6, we utilized the rule that English text must redundantly present the information communicated by the symbols for mock labels; this was because research (Seo et al., 2017) has suggested internationally-recognized symbols to be poorly recognized by healthcare providers. By contrast, a great deal of symbols in the commercial version did not incorporate English text.

When pairwise comparisons were made (within label versions) to investigate how risk categories impacted the proportion of correct responses (see Table 10), significant differences were apparent between most risk categories, with the exception of the comparison of mock, low and mock, medium (P=0.9936).

To further assess the performance of the mock labels we created, we also utilized the time to correct response as a dependent variable of interest. To meet normality assumptions related to the continuous variable, time to correct response, data was log transformed. The estimates for this model are shown in Figure 54 and Table 12 with a confidence level of 95%. Results of the pairwise analysis that compares between label versions (mock verses commercial) are included in Figure 54.

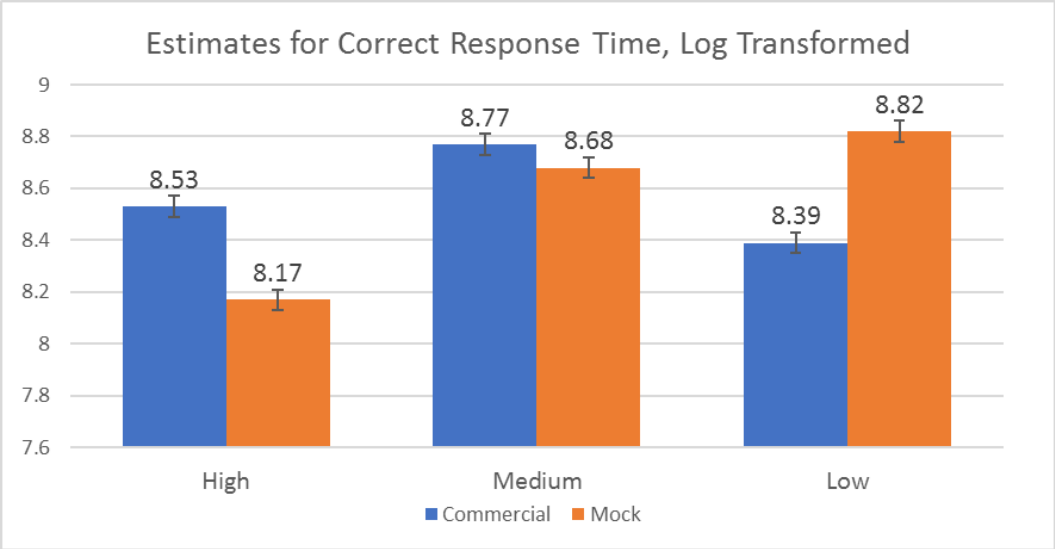


Figure 54 – Estimates for Correct Response Time

Table 12 – Estimates for Correct Response Time Tabulated

Label Version and Risk Category	Estimate (log transformed)	Standard Error	LCL	UCL
Mock, High	8.17	0.0400	8.09	8.25
Comm., High	8.53	0.0400	8.45	8.60
Mock, Medium	8.68	0.0401	8.61	8.76
Comm., Medium	8.77	0.0402	8.70	8.85
Mock, Low	8.82	0.0401	8.74	8.89
Comm., Low	8.39	0.0401	8.31	8.47

To determine if there were significant differences in time to correct response by label version and risk category, pairwise comparisons were conducted. For these comparisons, the estimates represent the difference between the two combinations. As an example:

Item A – Item B = 0 No difference in the time to correct response

Item A – Item B = positive value Item A has slower response time to correct response than Item B

Item A – Item B = negative value Item A has quicker response time to correct response than Item B

Comparisons within a risk category across commercial and mock versions can be seen in Table 13 and Figure 55 and comparisons within a label version across risk categories are presented in

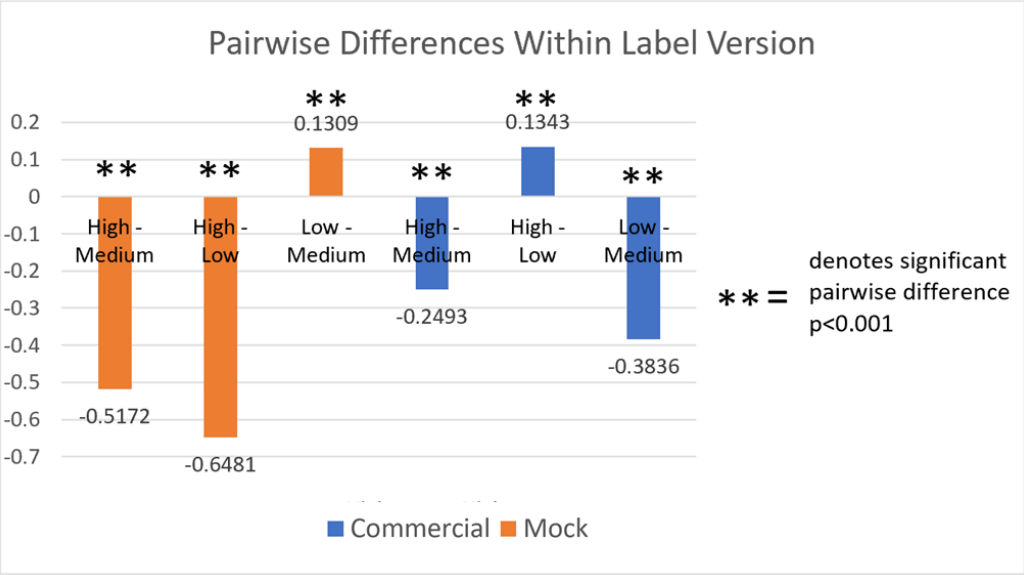


Figure 56 – Pairwise Comparisons, Risk Category Within Label Version, Correct Response Time

Table *14* and Figure 56.

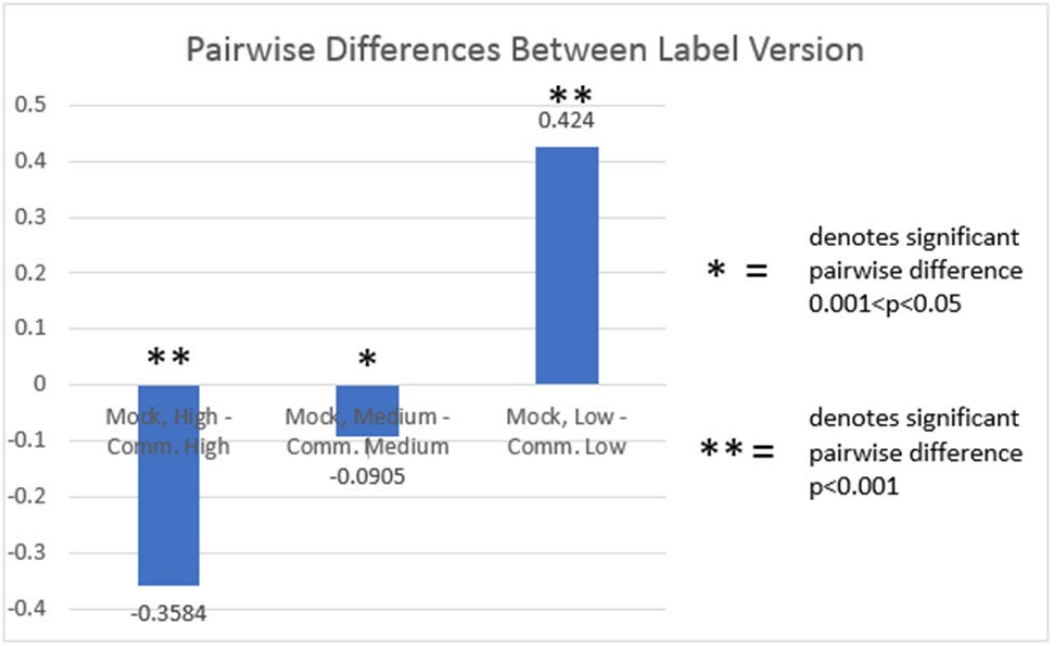


Figure 55 – Pairwise Comparisons, Risk Category Between Label Version, Correct Response Time

Table 13 – Pairwise Comparisons, Risk Category Between Label Version, Correct Response Time

Risk Category Between Label Version	Estimate	Standard Error	Z-Ratio	P-Value	Interpretation
Mock, High - Comm., High	**-0.3584**	**0.0230**	**-15.615**	**<0.0001**	**Participants were significantly faster to correctly identify high risk labeling inputs on mock labels than high risk labeling inputs on commercial labels**
Mock, Medium - Comm., Medium	**-0.0905**	**0.0235**	**-3.855**	**0.0016**	**Participants were significantly faster to correctly identify medium risk labeling inputs on mock labels than medium risk labeling inputs on a commercial labels**
Mock, Low - Comm., Low	**0.4240**	**0.0231**	**18.327**	**<0.0001**	**Participants were significantly faster to correctly identify low risk labeling inputs on**

					commercial labels than low risk labeling inputs on mock labels

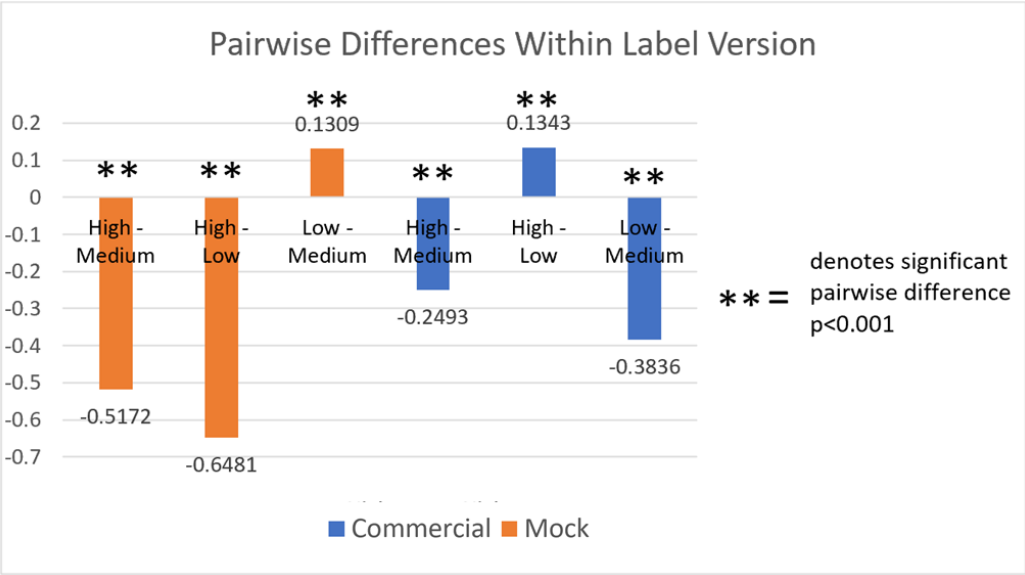


Figure 56 – Pairwise Comparisons, Risk Category Within Label Version, Correct Response Time

Table 14 – Pairwise Comparisons, Risk Category Within Label Version, Correct Response Time

Risk Category Within Label Version	Estimate	Standard Error	Z-Ratio	P-Value	Interpretation
Mock, High - Mock, Medium	**-0.5172**	**0.0231**	**-22.416**	**<0.0001**	**Participants were significantly faster to correctly identify high risk labeling inputs than medium risk labeling inputs on mock label versions**
Mock, High - Mock, Low	**-0.6481**	**0.0230**	**-28.130**	**<0.0001**	**Participants were significantly faster to correctly identify high risk labeling inputs than low risk labeling inputs on mock label versions**
Mock, Low - Mock, Medium	**0.1309**	**0.0232**	**5.650**	**<0.0001**	**Participants were significantly faster to correctly identify medium risk labeling inputs than**

					low risk labeling inputs on mock label versions
Comm., High - Comm., Medium	**-0.2493**	**0.0234**	**-10.667**	**<0.0001**	**Participants were significantly faster to correctly identify high risk labeling inputs than medium risk labeling inputs on commercial label versions**
Comm., High - Comm., Low	**0.1343**	**0.0230**	**5.827**	**<0.0001**	**Participants were significantly faster to correctly identify low risk labeling inputs than high risk labeling inputs on commercial label versions**
Comm., Low - Comm., Medium	**-0.3836**	**0.0235**	**-16.356**	**<0.0001**	**Participants were significantly faster to correctly identify low risk labeling inputs than medium risk labeling inputs on commercial label versions**

Results support the idea that we created a label that enabled users to find high risk information (P<0.0001) and medium risk (P=0.0016) more quickly than the commercial counterparts (see Table 13). Low risk, by contrast, was significantly slower to be correctly identified for mock versions compared to the commercial counterparts (<0.0001). This is likely attributable to the design Rules from Table 6 used to create the mock labels, which dictated that we emphasize medical device labeling inputs (see Table 1) in the categories identified in Objective 2 to be higher risk.

This is further supported by the findings presented in

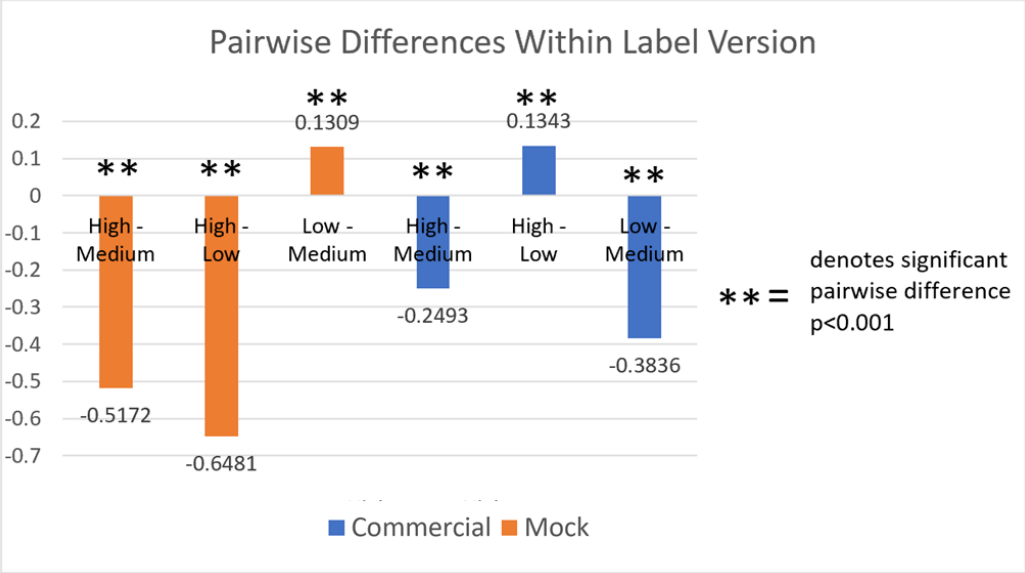


Figure 56 – Pairwise Comparisons, Risk Category Within Label Version, Correct Response Time

Table *14* which suggests differences in time to correctly identify a label when different risk categories within the same label version are compared. Data was back transformed for further discussion and are presented in Table 15.

Table 15 – Estimates for Correct Response Time Log Transformed Back

Label Version and Risk Category	Estimate (seconds)	LCL (seconds)	UCL (seconds)
Mock, High	3.53	3.26	3.83
Comm., High	5.06	4.68	5.43
Mock, Medium	5.88	5.49	6.37
Comm., Medium	6.44	6.00	6.97
Mock, Low	6.77	6.25	7.26
Comm., Low	4.40	4.06	4.77

The main takeaway that can be postulated from pairwise comparisons from both correct response and correct response time, is that mock labels were able to perform better for higher risk information (high and medium risk) when compared to their commercial counterparts. Similar accuracy between the label versions show that Healthcare Practitioners will spend the

necessary time to find key pieces of information, at the same time, our findings show that it is possible to create faster/more efficient mock labels without sacrificing said accuracy.

Chapter 8

Conclusions

This study suggests that using simple guidelines and formatting rules derived from recommendations of previous research (Cai, 2012) (Seo, 2014) (Seo et al., 2017) (Stifano et al., 2013) to emphasize information that users deem important for the safe use of medical devices, label designs which speed information processing can be created. Our objective evaluation of mock designs provides evidence that the performance of commercial medical device labels can be significantly improved and details a method that can be used to evaluate prioritization of information. And with more evolving devices and multiple stakeholders, standardization of medical device labeling will always remain as a moving target and complex undertaking. However, with the current climate of patient safety and a user centered approach, it is necessary to continue to find objective proof of what increases label performance and what also inhibits it.

Chapter 9

Future Research & Limitations

The main limitation in the research was the supplied commercial labels that were converted to a mock format. Only 18 different device labels were used in the study, which in comparison to the variety of labels currently out in the market, is a small percentage. Furthermore, regarding the label pool used, the labels were split across only 6 manufacturers, 44.45% of the 18 tested labels represented to Cardiovascular devices. Future research utilizing a more diverse range of medical device labels is recommended.

A secondary limitation was the scope at which design factor effects were analyzed. Per the study's design, only the overall performance for label version (mock vs. commercial) and label version x risk category (mock vs. commercial + high risk vs. medium risk vs. low risk) for accuracy and response time was measured. The effect of size, placement, color, and boxing was not isolated and measured in the study. Several specific design factors were measured on their own in previous studies (Seo, 2014), although not with varying degrees on the same label, for example regarding font over different risk categories: high risk highest % font size, medium risk middle % font size, low risk lowest % font size on the same space. A future study may focus on tweaking these constraints to reach optimal efficiency between these design elements on the same label.

One future study may focus on training to a specific mock label format. Before Forced Choice Task Decision testing, participants were not trained nor had the formatting/content rules of the "Device Facts" Box disclosed to them. Prior training to a mock format may help improve response accuracy and time and is an avenue which could warrant further research.

APPENDICES

APPENDIX A: Medical device labeling Input Assessment Survey Advertisement, Consent Form, and Pre-Survey Questionnaire

Figure 57 – Online Survey Advertisement

Healthcare Practitioner Volunteers Needed!

Medical Device Labeling Input
Patient Risk Assessment Survey

For participating in a **30-minute survey** you will be compensated with a **$10 gift card** from Amazon!

We are looking for registered nurses, surgical technicians, surgeons, or related medical practitioners to take part in our survey!

NO PRIOR SPECIALIZATION IN LABELING REQUIRED!

If you are interested please contact either:

Dr. Laura Bix: bixlaura@msu.edu XXX-XXX-XXXX

Or

Eric Joseph Estrada: estrad43@msu.edu XXX-XXX-XXXX

Michigan State University School of Packaging
Packaging Healthcare, Universal Design, and BioMechanics Team
448 Wilson Road, East Lansing, MI 48824

Figure 58 – Survey Flyer

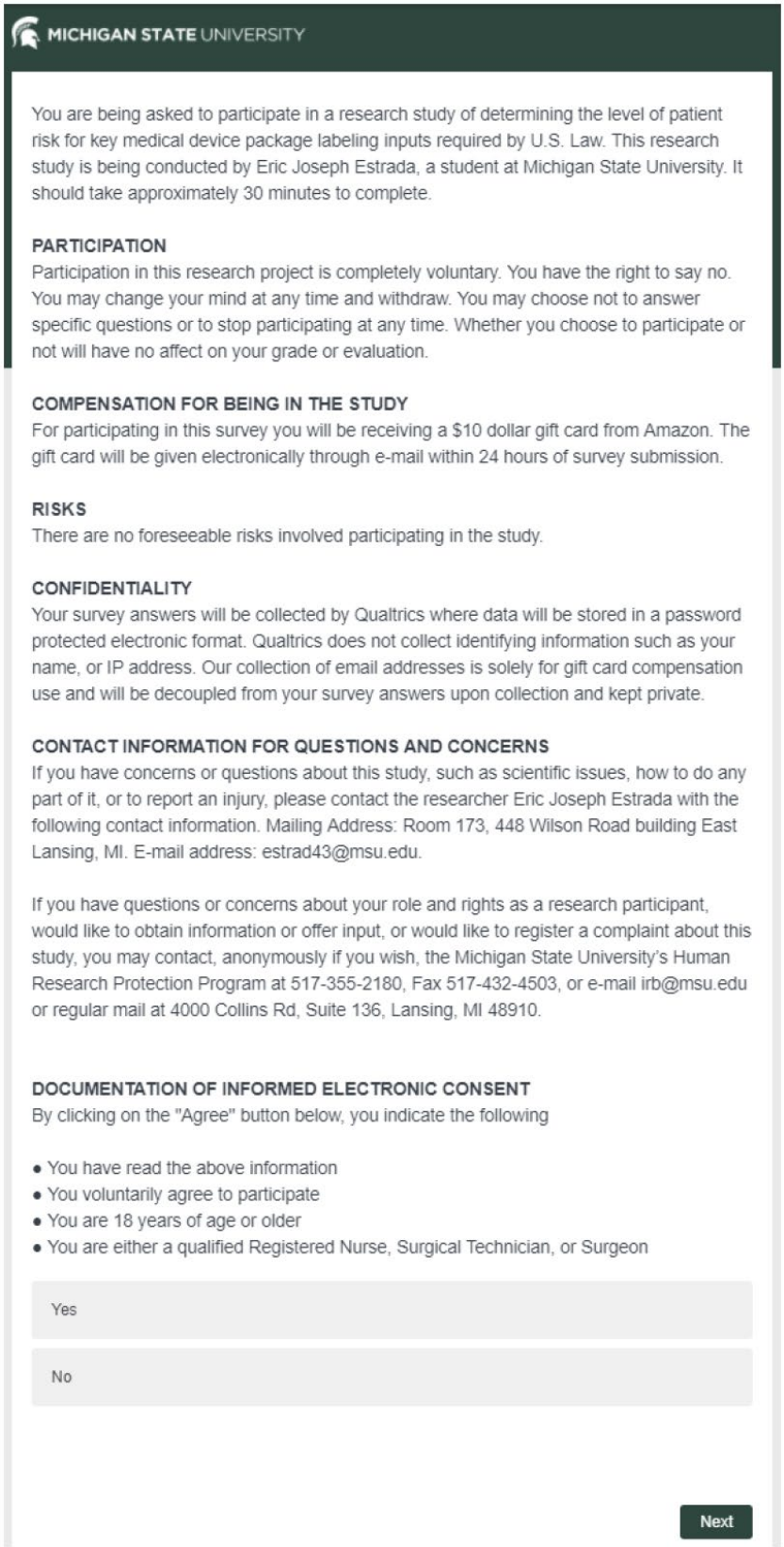


Figure 59 – Survey Consent Form

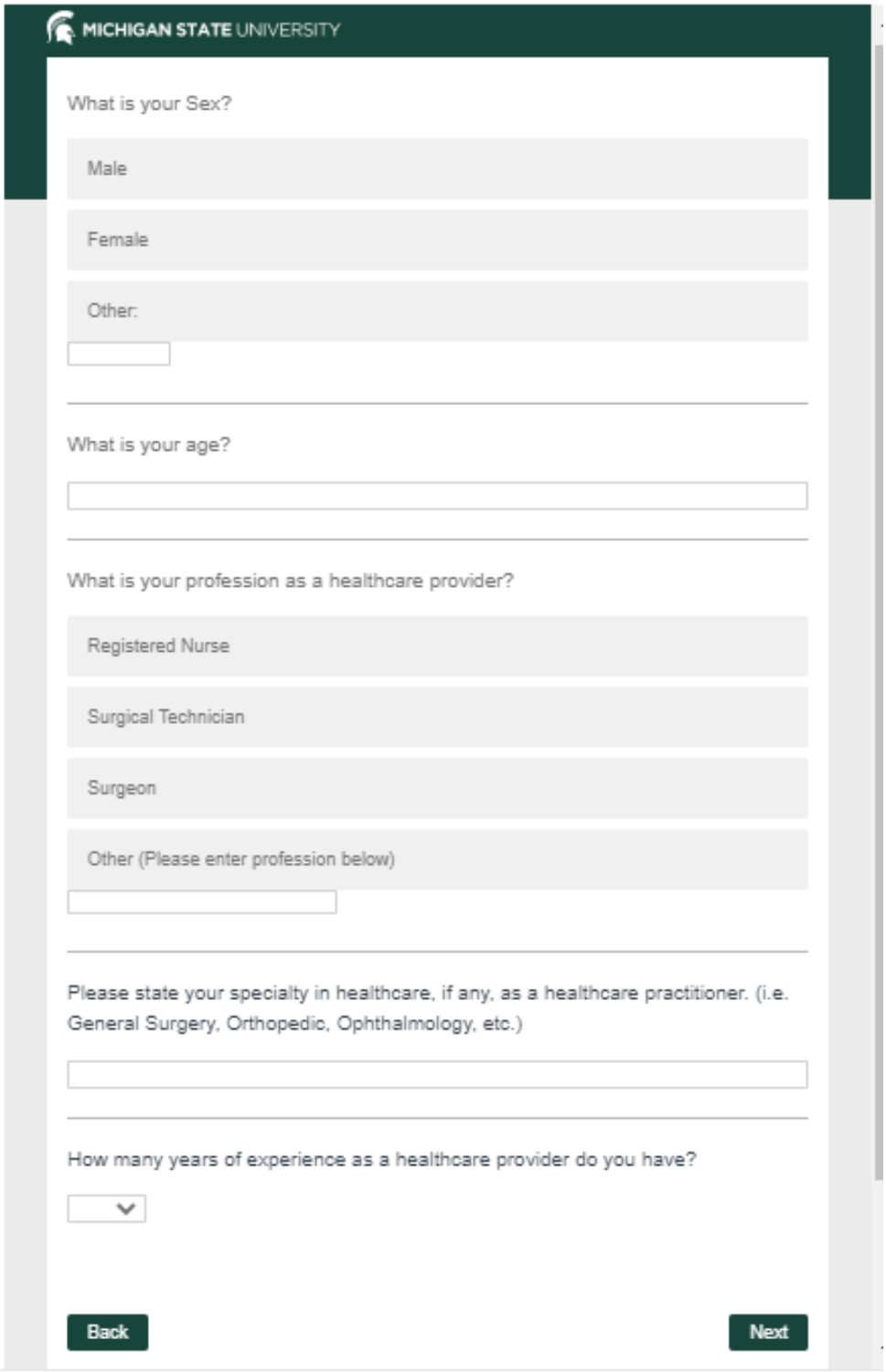


Figure 60 – Pre-Survey Questionnaire

APPENDIX B: Medical device labeling Input Assessment Survey Rank Ordering

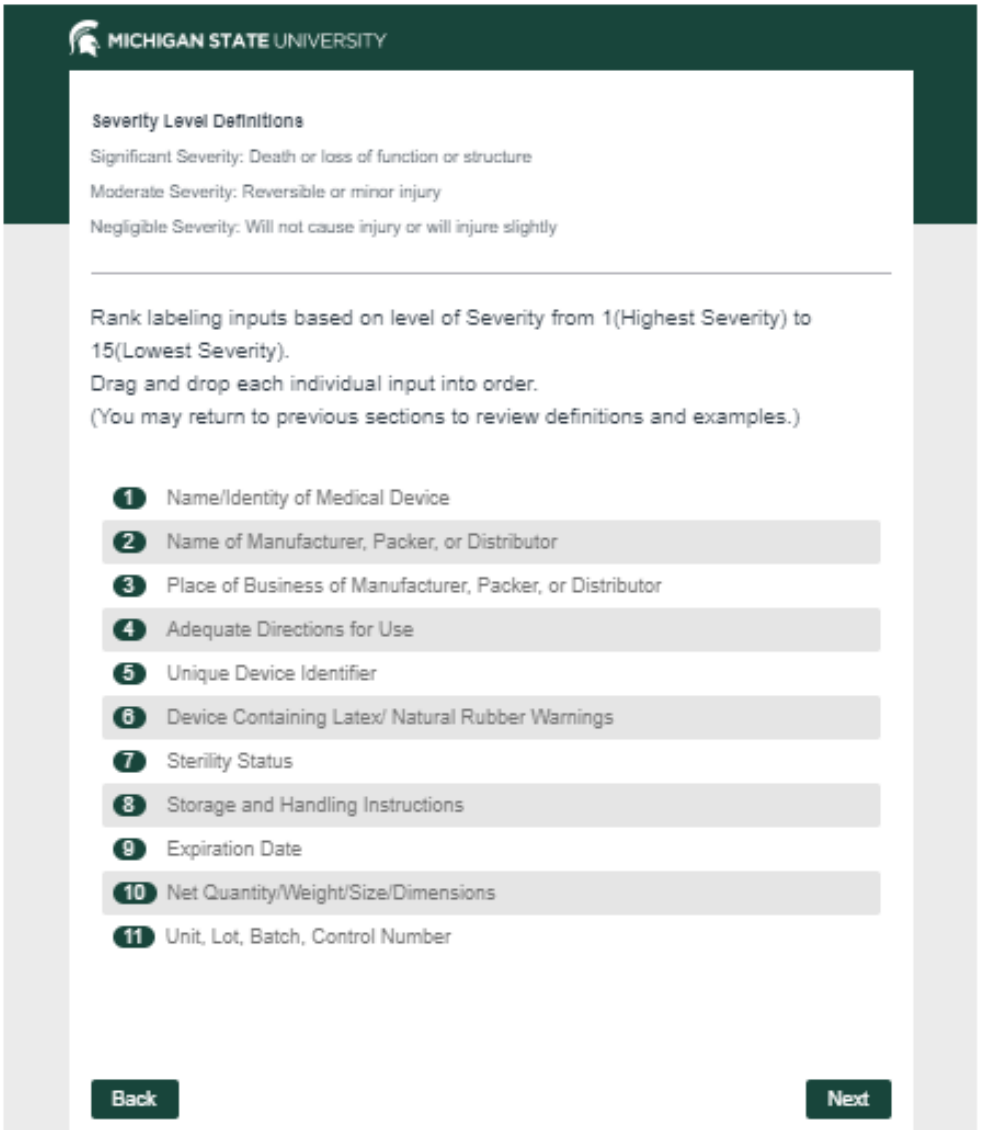


Figure 61 – Medical Device Labeling Input Rank Ordering by Severity

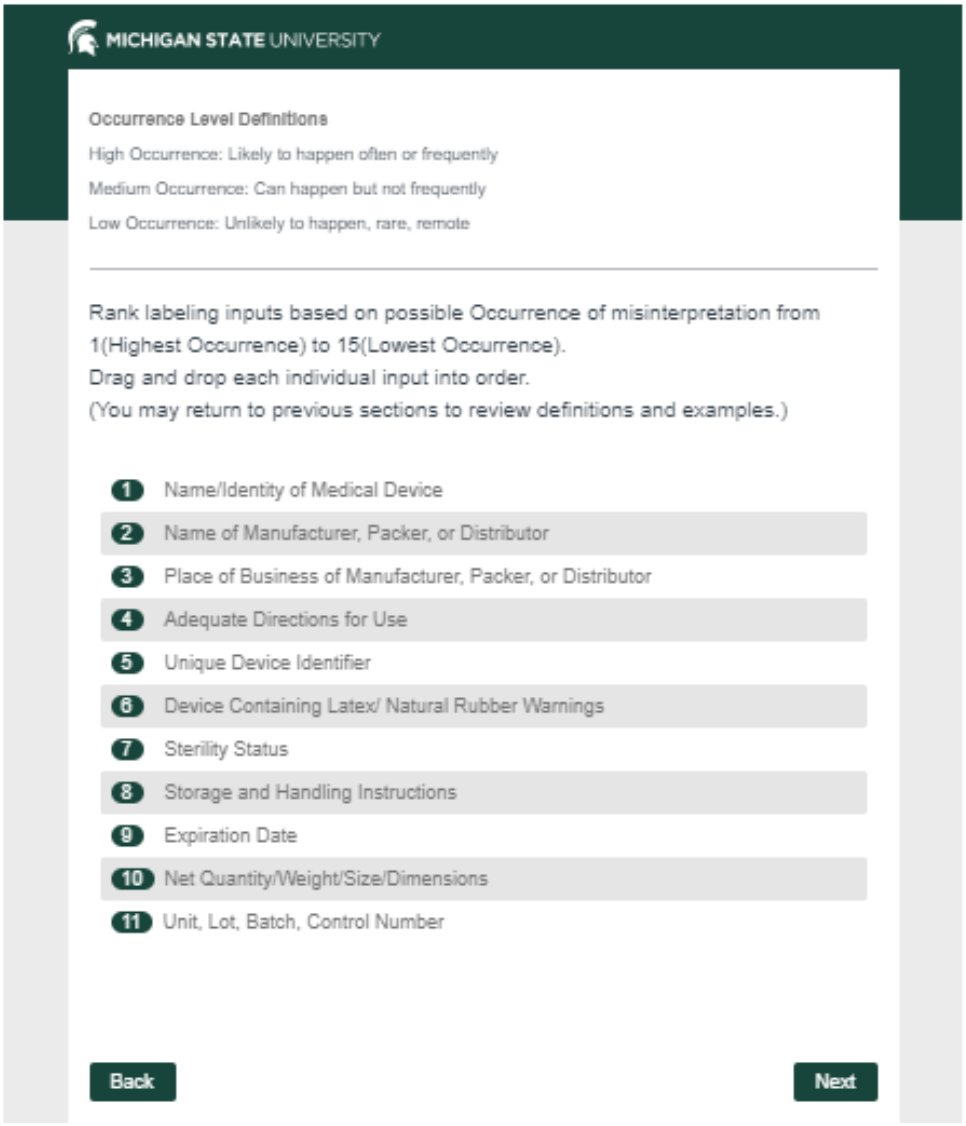


Figure 62 – Medical Device Labeling Input Rank Ordering by Occurrence

APPENDIX C: Mock Label Creation Donated Device Labels

18 out of 20 donated device samples were used for the forced task decision test. The 18 commercial device images used are shown in the figures below:

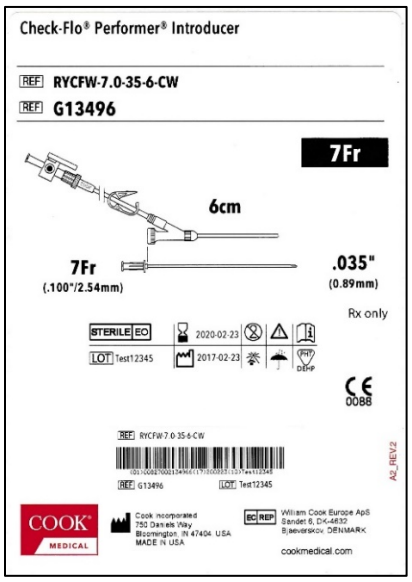


Figure 63 – Device 2 Commercial Label

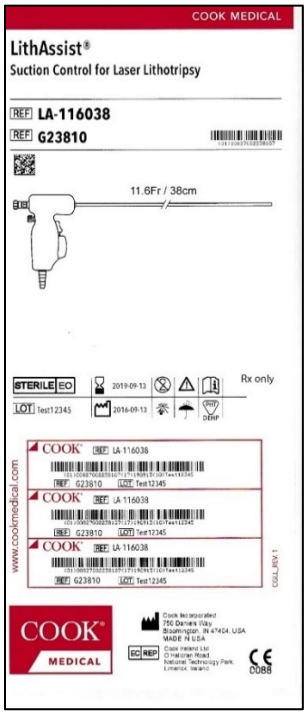


Figure 64 – Device 3 Commercial Label

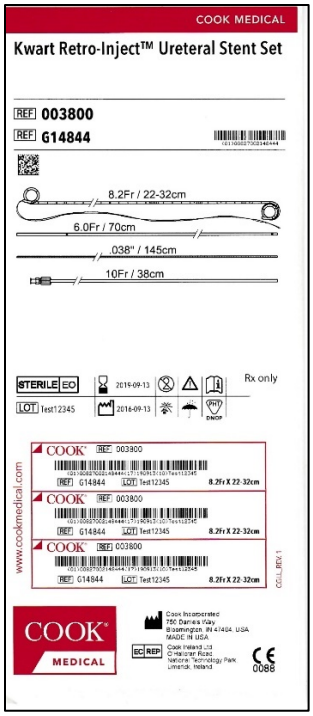


Figure 65 – Device 4 Commercial Label

Figure 66 – Device 5 Commercial Label

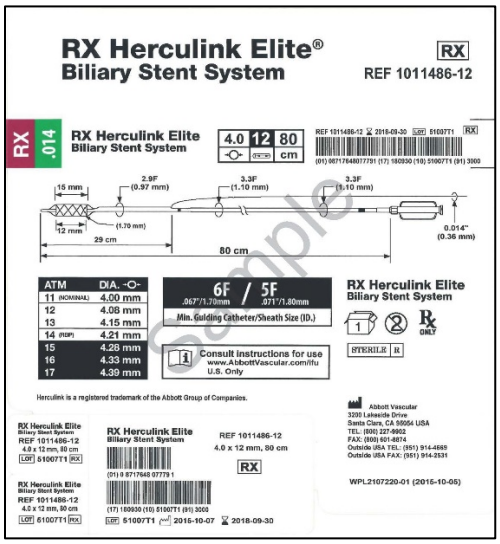


Figure 67 – Device 6 Commercial Label

Figure 68 – Device 7 Commercial Label

Figure 69 – Device 8 Commercial Label

Figure 70 – Device 9 Commercial Label

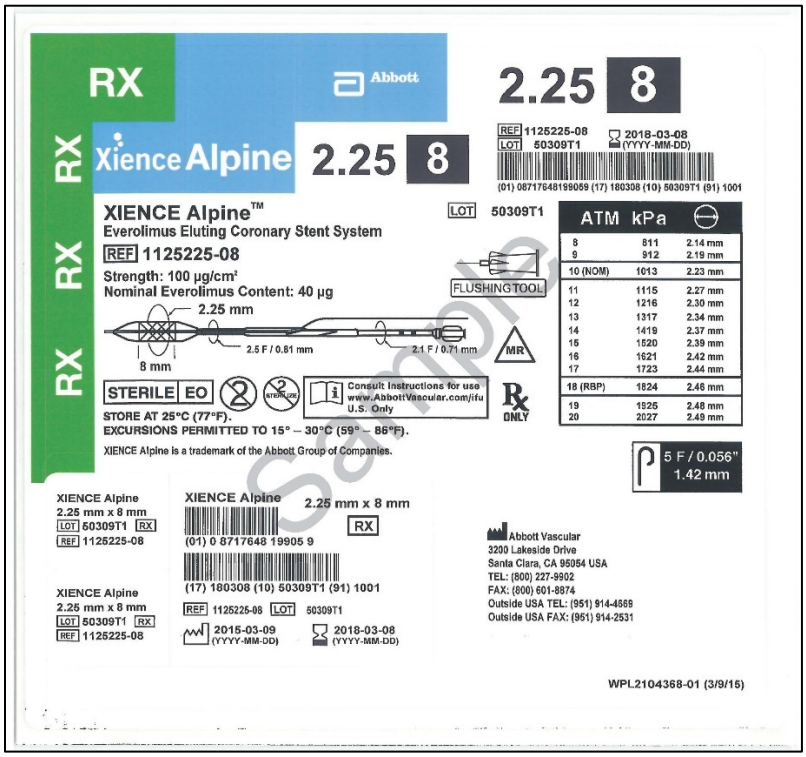


Figure 71 – Device 10 Commercial Label

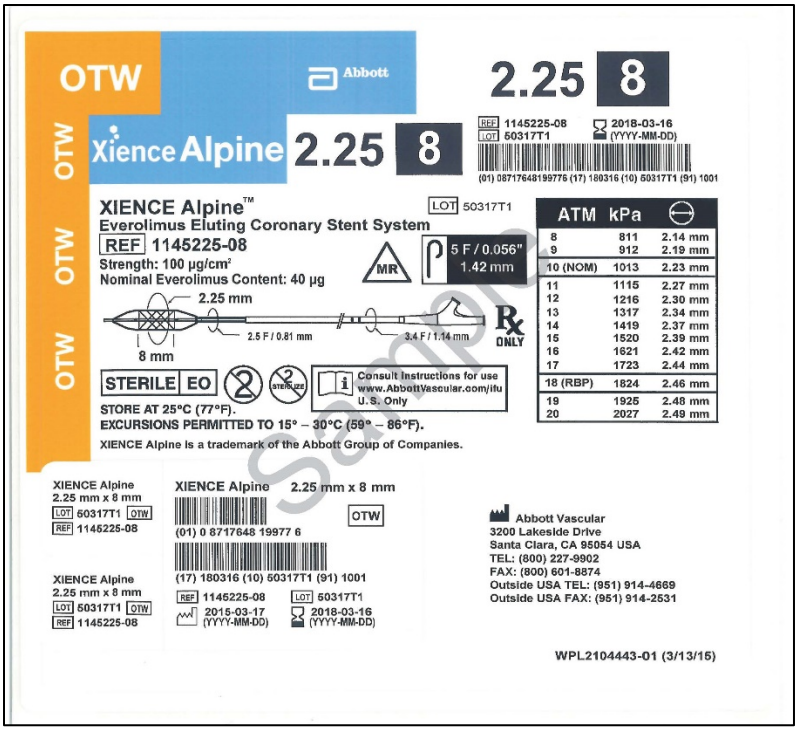


Figure 72 – Device 11 Commercial Label

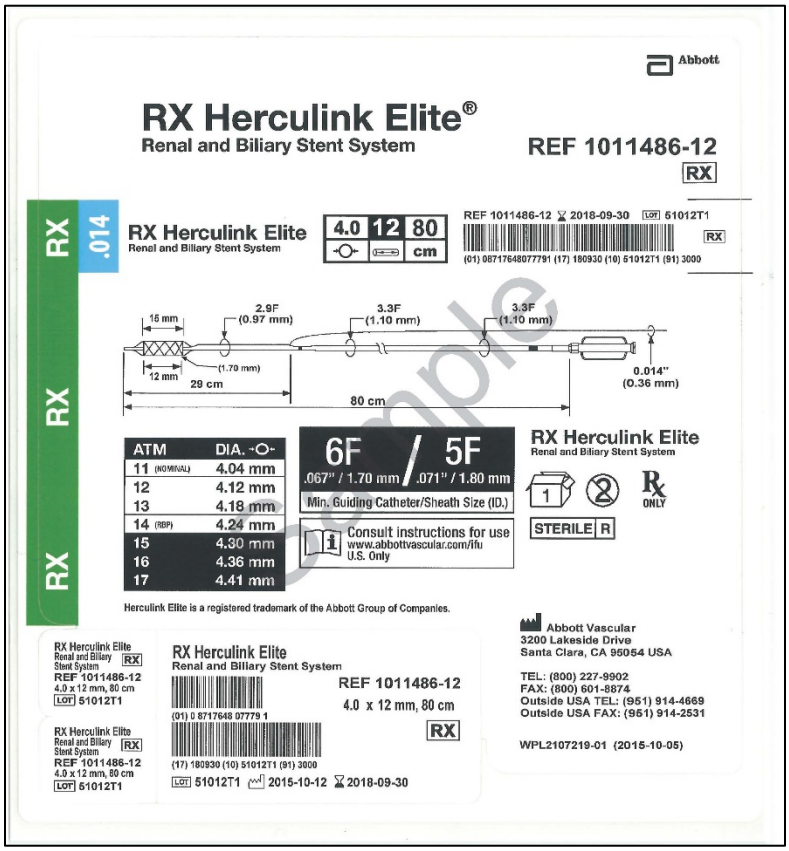


Figure 73 – Device 12 Commercial Label

Figure 74 – Device 13 Commercial Label

Figure 75 – Device 14 Commercial Label

Figure 76 – Device 15 Commercial Label

Figure 77 – Device 16 Commercial Label

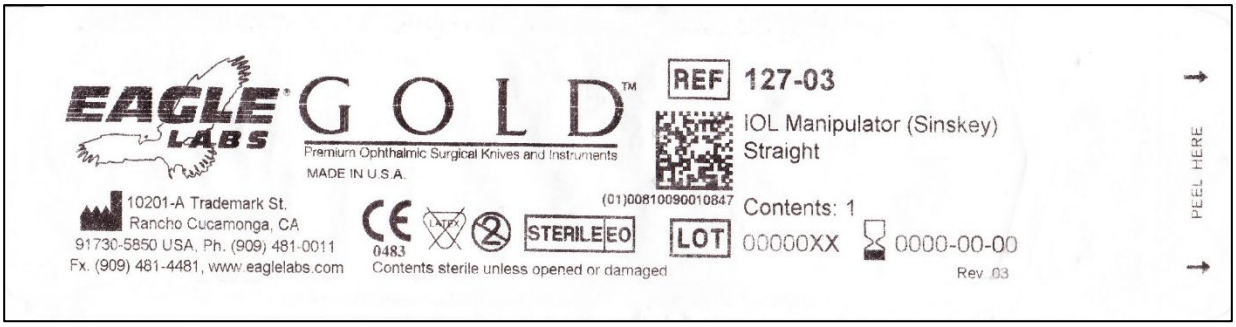


Figure 78 – Device 17 Commercial Label

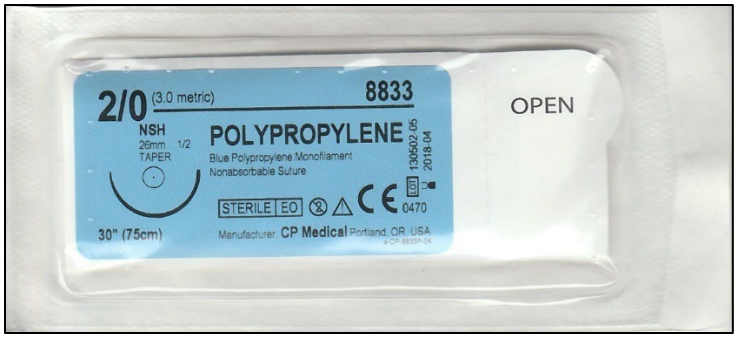


Figure 79 – Device 18 Commercial Label

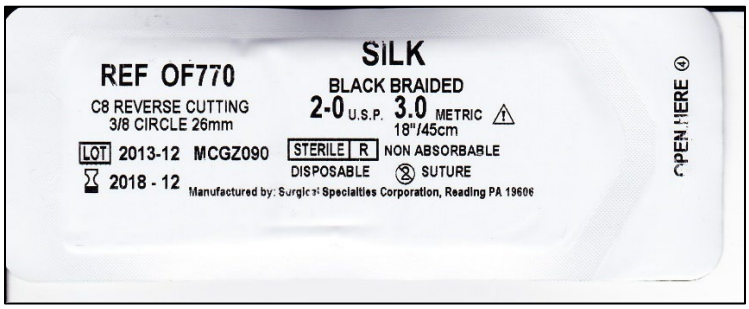


Figure 80 – Device 19 Commercial Label

APPENDIX D: Forced Choice Advertisement, Consent Form, and Pre-Test Questionnaire

Figure 81 – Forced Choice Online Advertisement

Healthcare Practitioner Volunteers Needed!

Research Study on Medical Device Labeling

For participating you will be compensated with **$40 in Cash**

We are looking for registered nurses, surgical technicians, surgeons, or related medical practitioners to take part in our survey!

To participate you must have no history of seizures!

You will be tasked to detect changes in medical device packaging labels that appear on a computer screen.

(Will take no more than 1.5 hours to complete)

NO PRIOR SPECIALIZATION IN LABELING REQUIRED!

If you are interested please contact either:

Dr. Laura Bix: bixlaura@msu.edu XXX-XXX-XXXX

Or

Eric Joseph Estrada: estrad43@msu.edu XXX-XXX-XXXX

Michigan State University School of Packaging
Packaging Healthcare, Universal Design, and BioMechanics Team
448 Wilson Road, East Lansing, MI 48824

Figure 82 – Forced Choice Flyer

Michigan State University
School of Packaging
Study Title: Medical Device Facts Box: Efficacy on Label Performance (Forced Choice Task Study)

INSTRUCTIONS AND RESEARCH CONSENT FORM

You are being asked to participate in a research project. Researchers are required to provide a consent form to inform you about the study, to convey that participation is voluntary, to explain risks and benefits of participation, and to empower you to make an informed decision. You should feel free to ask the researchers any questions or concerns you may have during the experiment.

1. **PURPOSE OF RESEARCH:** You are being asked to participate in an experiment to investigate how well the formats of different medical device labels work. More specifically, current medical device labels will be compared to mock medical device labels which are altered versions of the originals. Branding and addresses of original labels have been altered to protect company identity and assets.

2. **TO PARTICIPATE IN THIS STUDY YOU MUST:**
   a. Be 18 years old or older
   b. Not be legally blind
   c. Be either a registered nurse, surgical technician, surgeon, or be a medical practitioner in a related field

2. WHAT YOU WILL DO: If you agree to participate in this study, the following events will take place. We will ask you to answer some basic questions about yourself. Your visual acuity, basic health literacy, and your ability to see color will be tested. You will sit in front of a computer screen. On the computer screen, a test image with two nearly identical medical device labels will appear. The difference being one key labeling input will be the opposite of one another, i.e. medical device is sterile vs. medical device is not sterile. You will be prompted at the top of the computer screen to find and select one of the devices based on the different key input. This will be done by pressing on either the left or right arrow key. If you cannot find the difference within the time the allotted time (45 seconds per label), the software will move you to the next trial in the test. This process will repeat for a series of trials. The research should take no more than 1.5 hours of your time. In exchange for your participation in this study, you will receive $40.

3. POTENTIAL BENEFITS: There will be no direct benefit to you from these procedures. However, it is our goal to understand what factors make certain parts of a label more noticeable than others so that we can develop labels that provide important information to people in ways that they are likely to see it.

4. POTENTIAL RISKS: There are no foreseeable potential risks involved with this study. However, In the event that you are uncomfortable with any of the tasks, you may elect to skip a portion of the study, or discontinue altogether.

5. PRIVACY AND CONFIDENTIALITY: The data for this project will be tied to subject number, not name. Although the researchers, research staff, and the Institutional Review Board will have access to the data, neither the researchers nor anyone else will be able to link your data to you. Participant confidentiality will be protected to the maximum extent allowable by law. Paper records will be kept in Dr. Bix's office for as long as required by publishers or at least three years after the study closes whichever is longer; digital records will be housed on computers in our laboratories (Packaging). Data would be provided (deidentified) to publications that deemed it a necessary part of due diligence and is also accessible to the IRB.

6. YOUR RIGHTS TO PARTICIPATE, SAY NO, OR WITHDRAW: Participation in this research is completely voluntary. Refusal to participate will involve no penalty or loss of benefit. You may also refuse to answer particular questions. You may change your mind at any time, for any reason, and withdraw without penalty or loss of compensation.

Figure 83 – Forced Choice Consent Form Part 1

7. COSTS AND COMPENSATION FOR BEING IN THE STUDY: You will receive $40 in exchange for your participation in this study.

8. CONTACT INFORMATION FOR QUESTIONS AND CONCERNS: If you have any concerns or questions about this research study, such as scientific issues, how to do any part of it, or to report an injury, please contact the researcher (Laura Bix, PhD 448 Wilson, 114 Packaging East Lansing MI 48824 517 355-4556 bixlaura@msu.edu)

If you have questions or concerns about your role and rights as a research participant, would like to obtain information or offer input, or would like to register a complaint about this study, you may contact, anonymously if you wish, the Michigan State University's Human Research Protection Program at 517-355-2180, Fax 517-432-4503, or e-mail irb@msu.edu or regular mail at 408 W. Circle Drive, 207 Olds Hall, MSU, East Lansing, MI 48824.

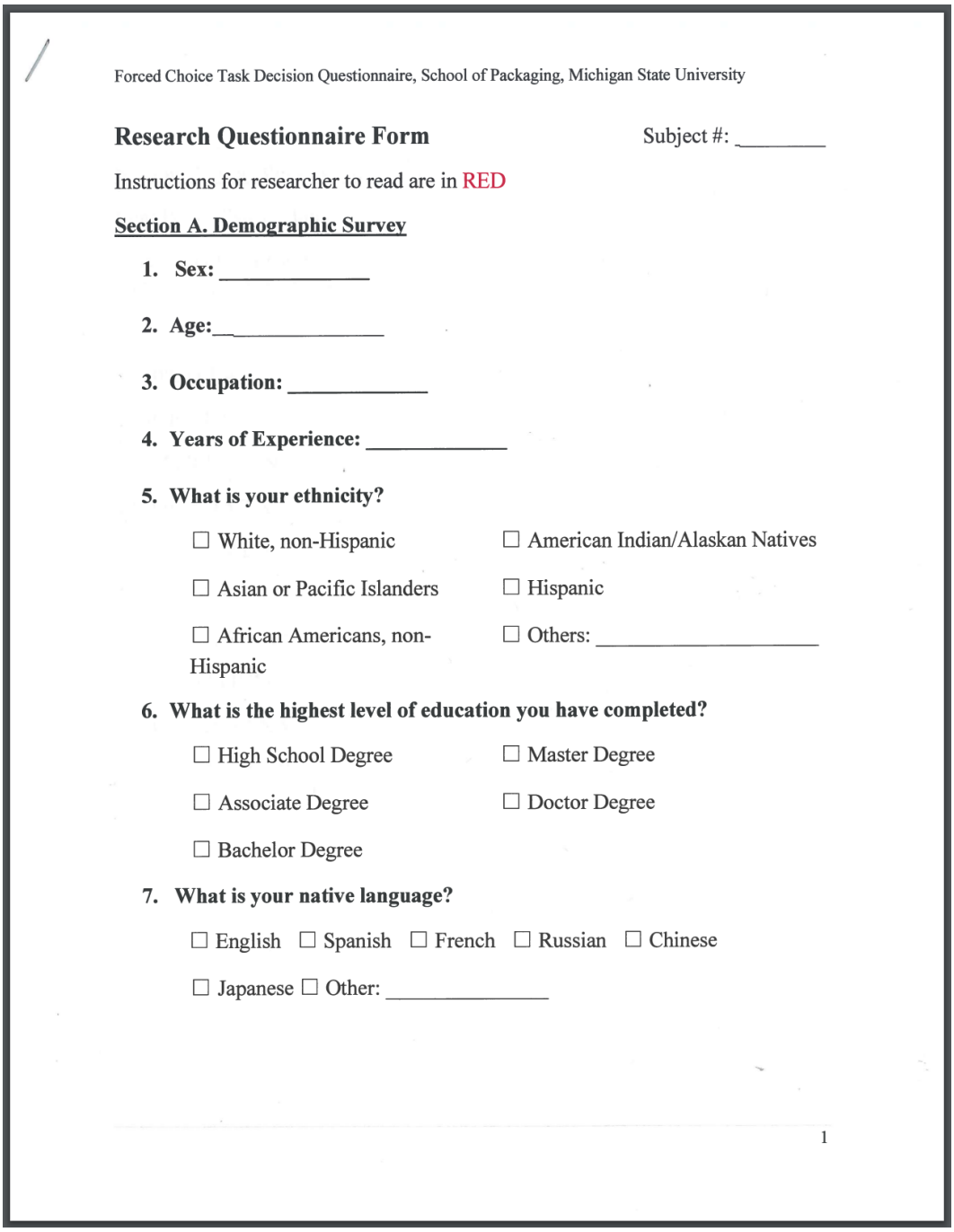
9. DOCUMENTATION OF INFORMED CONSENT: I voluntarily agree to participate in the study.

_____ _____

Signature Date

You will be provided with a copy of this consent form for your records.

Figure 84 – Forced Choice Consent Form Part 2



Forced Choice Task Decision Questionnaire, School of Packaging, Michigan State University

Research Questionnaire Form Subject #: _____

Instructions for researcher to read are in RED

Section A. Demographic Survey

1. **Sex:** _____
2. **Age:**_____
3. **Occupation:** _____
4. **Years of Experience:** _____
5. **What is your ethnicity?**

☐ White, non-Hispanic ☐ American Indian/Alaskan Natives

☐ Asian or Pacific Islanders ☐ Hispanic

☐ African Americans, non-Hispanic ☐ Others: _____

6. **What is the highest level of education you have completed?**

☐ High School Degree ☐ Master Degree

☐ Associate Degree ☐ Doctor Degree

☐ Bachelor Degree

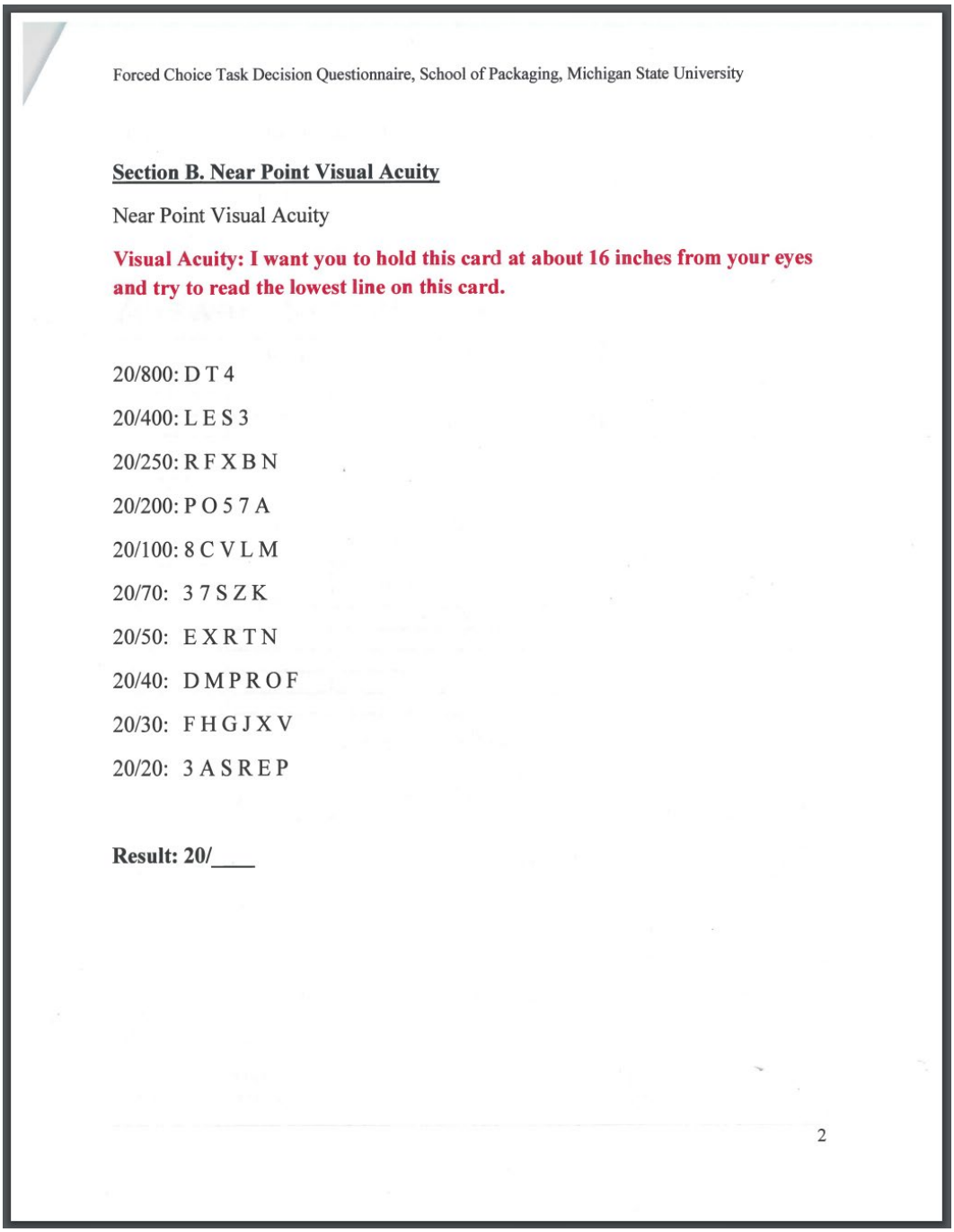
7. **What is your native language?**

☐ English ☐ Spanish ☐ French ☐ Russian ☐ Chinese

☐ Japanese ☐ Other: _____

1

Figure 85 – Forced Choice Pre-Survey Questionnaire Part 1



Forced Choice Task Decision Questionnaire, School of Packaging, Michigan State University

Section B. Near Point Visual Acuity

Near Point Visual Acuity

Visual Acuity: I want you to hold this card at about 16 inches from your eyes and try to read the lowest line on this card.

20/800: D T 4

20/400: L E S 3

20/250: R F X B N

20/200: P O 5 7 A

20/100: 8 C V L M

20/70: 3 7 S Z K

20/50: E X R T N

20/40: D M P R O F

20/30: F H G J X V

20/20: 3 A S R E P

Result: 20/____

2

Figure 86 – Forced Choice Pre-Survey Questionnaire Part 2

Section B. Color Differentiation Ability

Please hold each of these 75 cm (measure with string) from your eyes and read the number that appears to you. If no number is apparent, please say "pass".

Write the number that the subject states for each trial on this form. Put an x through incorrect trials and a checkmark across the plates that are correct.

Answers to each plate

Plate	Normal Person	Person with Red-Green Deficiencies				Person with Total Color Blindness and Weakness
1	12	12				12
2	8	3				X
3	29	70				X
4	5	2				X
5	3	5				X
6	15	17				X
7	74	21				X
8	6	X				X
9	45	X				X
10	5	X				X
11	7	X				X
12	16	X				X
13	73	X				X
14	X	5				X
15	X	45				X
		Protan		Deutan		
		Strong	Mild	Strong	Mild	
16	26	6	(2) 6	2	2 (6)	
17	42	2	(4) 2	4	4 (2)	

The mark X shows that the plate cannot be read. Blank spce denotes that the reading is indefinite. The numerals in parenthesis show that they can be read but they are comparatively unclear.

As assessment of the readings of plates 1 to 15 determines the normality or defectiveness of color vision.
If 13 or more plates are read normally, the color vision is regarded as normal.
If only 9 or less than 9 plates are read normally, the color vision is regarded as deficient.
However, in reference to plates 14 and 15, only those who read the numerals 5 and 45 and read them easier than those on plates 10 and 9 are recorded as abnormal readings.

It is rare to find a person whose recording of normal answers is 14-16 plates. An assessment of such a case requires the use of other color vision tests, including the anomaloscope.

Figure 87 – Forced Choice Pre-Survey Questionnaire Part 3

APPENDIX E: Forced Choice Trial Combination

Shown below is the combination of each trial 1-240. Each trial is a combination of device label, label version, and labeling input. Devices 2-19 were used in the Forced Choice Task Decision Test and as such only trials 7-114 (commercial) and 127-234 (mock) were used.

Table 16 – Forced Choice Trial Combination

Trial # (Commercial Label Version)	Device Label #	Labeling Input	Device Label #	Trial # (Mock Label Version)
1	Device 1	Sterility Status	Device 1	121
2		Expiration Date		122
3		Storage and Handling		123
4		Unit, Lot, Batch, and Control Number		124
5		Place of Business		125
6		Net Quantity/Weight/Size/Dimensions		126
7	Device 2	Sterility Status	Device 2	127
8		Expiration Date		128
9		Storage and Handling		129
10		Unit, Lot, Batch, and Control Number		130
11		Place of Business		131
12		Net Quantity/Weight/Size/Dimensions		132
13	Device 3	Sterility Status	Device 3	133
14		Expiration Date		134
15		Storage and Handling		135
16		Unit, Lot, Batch, and Control Number		136
17		Place of Business		137
18		Net Quantity/Weight/Size/Dimensions		138
19	Device 4	Sterility Status	Device 4	139
20		Expiration Date		140
21		Storage and Handling		141
22		Unit, Lot, Batch, and Control Number		142
23		Place of Business		143

Table 16 (cont'd)

24		Net Quantity/Weight/Size/Dimensions		144
25	Device 5	Sterility Status	Device 5	145
26		Expiration Date		146
27		Storage and Handling		147
28		Unit, Lot, Batch, and Control Number		148
29		Place of Business		149
30		Net Quantity/Weight/Size/Dimensions		150
31	Device 6	Sterility Status	Device 6	151
32		Expiration Date		152
33		Storage and Handling		153
34		Unit, Lot, Batch, and Control Number		154
35		Place of Business		155
36		Net Quantity/Weight/Size/Dimensions		156
37	Device 7	Sterility Status	Device 7	157
38		Expiration Date		158
39		Storage and Handling		159
40		Unit, Lot, Batch, and Control Number		160
41		Place of Business		161
42		Net Quantity/Weight/Size/Dimensions		162
43	Device 8	Sterility Status	Device 8	163
44		Expiration Date		164
45		Storage and Handling		165
46		Unit, Lot, Batch, and Control Number		166
47		Place of Business		167
48		Net Quantity/Weight/Size/Dimensions		168
49	Device 9	Sterility Status	Device 9	169
50		Expiration Date		170
51		Storage and Handling		171
52		Unit, Lot, Batch, and Control Number		172
53		Place of Business		173
54		Net Quantity/Weight/Size/Dimensions		174

Table 16 (cont'd)

55	Device 10	Sterility Status	Device 10	175
56		Expiration Date		176
57		Storage and Handling		177
58		Unit, Lot, Batch, and Control Number		178
59		Place of Business		179
60		Net Quantity/Weight/Size/Dimensions		180
61	Device 11	Sterility Status	Device 11	181
62		Expiration Date		182
63		Storage and Handling		183
64		Unit, Lot, Batch, and Control Number		184
65		Place of Business		185
66		Net Quantity/Weight/Size/Dimensions		186
67	Device 12	Sterility Status	Device 12	187
68		Expiration Date		188
69		Storage and Handling		189
70		Unit, Lot, Batch, and Control Number		190
71		Place of Business		191
72		Net Quantity/Weight/Size/Dimensions		192
73	Device 13	Sterility Status	Device 13	193
74		Expiration Date		194
75		Storage and Handling		195
76		Unit, Lot, Batch, and Control Number		196
77		Place of Business		197
78		Net Quantity/Weight/Size/Dimensions		198
79	Device 14	Sterility Status	Device 14	199
80		Expiration Date		200
81		Storage and Handling		201
82		Unit, Lot, Batch, and Control Number		202
83		Place of Business		203
84		Net Quantity/Weight/Size/Dimensions		204
85	Device 15	Sterility Status	Device 15	205
86		Expiration Date		206

Table 16 (cont'd)

87		Storage and Handling		207
88		Unit, Lot, Batch, and Control Number		208
89		Place of Business		209
90		Net Quantity/Weight/Size/Dimensions		210
91	Device 16	Sterility Status	Device 16	211
92		Expiration Date		212
93		Storage and Handling		213
94		Unit, Lot, Batch, and Control Number		214
95		Place of Business		215
96		Net Quantity/Weight/Size/Dimensions		216
97	Device 17	Sterility Status	Device 17	217
98		Expiration Date		218
99		Storage and Handling		219
100		Unit, Lot, Batch, and Control Number		220
101		Place of Business		221
102		Net Quantity/Weight/Size/Dimensions		222
103	Device 18	Sterility Status	Device 18	223
104		Expiration Date		224
105		Storage and Handling		225
106		Unit, Lot, Batch, and Control Number		226
107		Place of Business		227
108		Net Quantity/Weight/Size/Dimensions		228
109	Device 19	Sterility Status	Device 19	229
110		Expiration Date		230
111		Storage and Handling		231
112		Unit, Lot, Batch, and Control Number		232
113		Place of Business		233
114		Net Quantity/Weight/Size/Dimensions		234
115	Device 20	Sterility Status	Device 20	235
116		Expiration Date		236
117		Storage and Handling		237

Table 16 (cont'd)

118		Unit, Lot, Batch, and Control Number		238
119		Place of Business		239
120		Net Quantity/Weight/Size/Dimensions		240

APPENDIX F: Forced Choice Run Order Counterbalancing

The figure below represents the trial set sequence for Part A, Part B would start with Part A's Second Set followed with a break and then finishing the test with Part A's First Set.

Part A First Set						**Part A Second Set**		
7	80	165				19	86	171
8	81	166				20	87	172
9	82	167				21	88	173
10	83	168				22	89	174
11	84	175				23	90	181
12	97	176				24	91	182
13	98	177				25	92	183
14	99	178				26	93	184
15	100	179				27	94	185
16	101	180				28	95	186
17	102	187				29	96	193
18	109	188	→	BREAK	→	30	103	194
31	110	189				37	104	195
32	111	190				38	105	196
33	112	191				39	106	197
34	113	192				40	107	198
35	114	199				41	108	205
36	127	200				42	139	206
43	128	201				49	140	207
44	129	202				50	141	208
45	130	203				51	142	209
46	131	204				52	143	210
47	132	217				53	144	211
48	133	218				54	145	212
55	134	219				61	146	213
56	135	220				62	147	214
57	136	221				63	148	215
58	137	222				64	149	216
59	138	229				65	150	223
60	151	230				66	157	224
67	152	231				73	158	225
68	153	232				74	159	226
69	154	233				75	160	227
70	155	234				76	161	228
71	156					77	162	
72	163					78	169	
79	164					85	170	

Figure 88 – Forced Choice Run Order Counterbalancing

APPENDIX G: Forced Choice Correct Choice Position Blocking Odd Block

The Figure Below represents the blocking for the Odd group. The Even Group has the correct position's flipped (for example: Trial # 7's correct position would be LEFT.

Table 17 – Forced Choice Correct Choice Position Odd Block

Trial #	Correct Position (Commercial)		Trial #	Correct Position (Mock)
1	RIGHT		121	RIGHT
2	RIGHT		122	LEFT
3	LEFT		123	RIGHT
4	LEFT		124	LEFT
5	LEFT		125	RIGHT
6	RIGHT		126	RIGHT
7	RIGHT		127	RIGHT
8	LEFT		128	LEFT
9	RIGHT		129	LEFT
10	LEFT		130	LEFT
11	LEFT		131	LEFT
12	LEFT		132	RIGHT
13	LEFT		133	LEFT
14	RIGHT		134	LEFT
15	LEFT		135	RIGHT
16	LEFT		136	LEFT
17	RIGHT		137	RIGHT
18	LEFT		138	RIGHT
19	RIGHT		139	RIGHT
20	RIGHT		140	LEFT
21	RIGHT		141	RIGHT
22	LEFT		142	LEFT
23	RIGHT		143	RIGHT
24	LEFT		144	RIGHT
25	LEFT		145	RIGHT
26	LEFT		146	LEFT
27	RIGHT		147	RIGHT
28	RIGHT		148	LEFT
29	RIGHT		149	LEFT
30	LEFT		150	RIGHT
31	LEFT		151	RIGHT
32	LEFT		152	RIGHT
33	RIGHT		153	RIGHT
34	RIGHT		154	RIGHT

Table 17 (cont'd)

35	RIGHT		155	LEFT
36	LEFT		156	LEFT
37	RIGHT		157	LEFT
38	RIGHT		158	LEFT
39	RIGHT		159	LEFT
40	RIGHT		160	RIGHT
41	LEFT		161	RIGHT
42	LEFT		162	LEFT
43	LEFT		163	LEFT
44	LEFT		164	LEFT
45	LEFT		165	RIGHT
46	RIGHT		166	RIGHT
47	RIGHT		167	LEFT
48	LEFT		168	RIGHT
49	LEFT		169	RIGHT
50	RIGHT		170	LEFT
51	LEFT		171	RIGHT
52	RIGHT		172	LEFT
53	LEFT		173	LEFT
54	LEFT		174	LEFT
55	LEFT		175	RIGHT
56	LEFT		176	LEFT
57	LEFT		177	RIGHT
58	RIGHT		178	RIGHT
59	LEFT		179	LEFT
60	RIGHT		180	RIGHT
61	LEFT		181	RIGHT
62	RIGHT		182	LEFT
63	RIGHT		183	LEFT
64	LEFT		184	RIGHT
65	RIGHT		185	RIGHT
66	LEFT		186	LEFT
67	LEFT		187	RIGHT
68	RIGHT		188	LEFT
69	LEFT		189	RIGHT
70	RIGHT		190	RIGHT
71	LEFT		191	LEFT
72	RIGHT		192	LEFT
73	LEFT		193	LEFT
74	LEFT		194	RIGHT
75	RIGHT		195	LEFT
76	RIGHT		196	RIGHT
77	RIGHT		197	RIGHT

Table 17 (cont'd)

78	RIGHT		198	LEFT
79	RIGHT		199	RIGHT
80	RIGHT		200	LEFT
81	RIGHT		201	LEFT
82	LEFT		202	RIGHT
83	LEFT		203	LEFT
84	LEFT		204	RIGHT
85	LEFT		205	RIGHT
86	RIGHT		206	RIGHT
87	LEFT		207	LEFT
88	LEFT		208	LEFT
89	RIGHT		209	RIGHT
90	LEFT		210	LEFT
91	LEFT		211	LEFT
92	LEFT		212	LEFT
93	LEFT		213	LEFT
94	LEFT		214	RIGHT
95	RIGHT		215	LEFT
96	RIGHT		216	RIGHT
97	RIGHT		217	RIGHT
98	RIGHT		218	LEFT
99	LEFT		219	LEFT
100	RIGHT		220	RIGHT
101	LEFT		221	LEFT
102	LEFT		222	RIGHT
103	RIGHT		223	RIGHT
104	LEFT		224	RIGHT
105	RIGHT		225	LEFT
106	LEFT		226	RIGHT
107	LEFT		227	RIGHT
108	RIGHT		228	LEFT
109	LEFT		229	LEFT
110	LEFT		230	LEFT
111	RIGHT		231	RIGHT
112	RIGHT		232	LEFT
113	LEFT		233	RIGHT
114	RIGHT		234	LEFT
115	RIGHT		235	LEFT
116	LEFT		236	RIGHT
117	RIGHT		237	LEFT
118	RIGHT		238	RIGHT
119	LEFT		239	LEFT
120	RIGHT		240	LEFT

REFERENCES

REFERENCES

Access GUDID. (2020). U.S. National Library of Medicine.

Agency Information Collection Activities; Proposed Collection; Comment Request; Survey of Health Care Practitioners for Device Labeling Format and Content. (2014). Food and Drug Administration, HHS.

Cai, J. (2012). *Perceptions of Medical Device Packaging Used by Operating Room Personnel*. Michigan State University.

Complying with labelling as EU MDR gets close. (2018). Med-Tech Innovation. https://www.med-technews.com/news/complying-with-labelling-as-eu-mdr-gets-close/

Francis, M. (2020). *How Upton Sinclair's "The Jungle" Unintentionally Spurred Food Safety Laws*. WTTW Chicago Public Media.

Gaffney, A. (2015). How can Medical Device Labeling be Standardized? FDA Study Aims to Find Out. *Regulatory Focus*.

GUD ID Diagram. (2020). U.S. National Library of Medicine. https://accessgudid.nlm.nih.gov/assets/help/gudiddiagram-a883a224f3002763596a0c5673840fa92d376c310e506b00820314ebad56c95d.png

Janssen, W. (1981). *The Story of the Laws Behind the Labels*. FDA Consumer Magazine.

Johnson, J. (2016). FDA Regulation of medical devices. *Congressional Research Service*.

Llamas, M. (2020). *How to Read a Drug Label*. Drugwatch.

Reason, J. (1995). Understanding adverse events: human factors. *Quality in Healthcare*, *4*, 80–89.

Seo, D. C. (2014). *Assessing the Effect of Varied Design Elements on Information Processing in Medical Device Labels*. Michigan State University.

Seo, D. C., Ladoni, M., Brunk, E., Becker, M. W., & Bix, L. (2017). Do Healthcare Professionals Comprehend Standardized Symbols Present on Medical Device Packaging?: An Important Factor in the Fight Over Label Space. *Packaging Technology and Science*, *30*(3), 61–73. https://doi.org/10.1002/pts.2199

Songara, R. (2010). Need for harmonization of medical devices: a review. *Journal of Advanced Pharmaceutical Technology & Research*, *1*(2), 127–144.

Stifano, T., Fu, M., Alexander, J. P., Weger, S. A., Southwell, B. G., Burke, M. F., & Thaker, S. (2013). Healthcare provider preferences for medical device labeling. *Biomedical Instrumentation and Technology*, *47*(SPRING), 42. https://doi.org/10.2345/0899-8205-47.s1.42

Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Academie Canadienne de Psychiatrie de l'enfant et de*

l'adolescent, *19*(3), 227–229.

Ward, J. R., & Clarkson, P. J. (2004). An analysis of medical device-related errors: prevalence and possible solutions. *Journal of Medical Engineering & Technology*, *28*(1), 2–21. https://doi.org/10.1080/0309190031000123747