OPTIMIZING OTC LABELS FOR OLDER ADULTS: EMPIRICAL EVALUATION OF LABELS DESIGNED TO PROVIDE OLDER USERS THE INFORMATION THEY NEED TO MINIMIZE ADVERSE DRUG EVENTS

By

Alyssa Lee Harben

A DISSERTATION

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

Packaging- Doctor of Philosophy

2021

ABSTRACT

OPTIMIZING OTC LABELS FOR OLDER ADULTS: EMPIRICAL EVALUATION OF LABELS DESIGNED TO PROVIDE OLDER USERS THE INFORMATION THEY NEED TO MINIMIZE ADVERSE DRUG EVENTS

By

Alyssa Lee Harben

Despite the many benefits of Over-the-Counter drugs for older consumers, there are risks that accompany their use, with as many as 15% of older OTC medication users being at risk of a serious Adverse Drug Reaction. As such, there is a responsibility to develop packaging that provides the essential function of facilitating cost-effective patient care by communicating critical information at the point of purchase. Optimally designed labels garner attention to critical information regardless of whether the consumer is engaged in bottom-up processing (a habitual purchase) or top-down processing (deliberative search for specific information).

We objectively assessed four label designs to investigate the effect of highlighting critical information (defined as warnings associated with drug/drug or drug diagnosis interactions and the active ingredient within a product) and placement of the same on the front of the package (FOP label treatment). Highlighting and FOP were crossed for a total of four designs (highlight (HL) with FOP, HL without FOP, No HL (nonHL) with FOP, nonHL without FOP(current, standard practice)). These treatments were utilized to evaluate how design attributes attract attention to critical information and promote decision-making in older adults (65+) when accessing that information *was* and *was not* the participant's goal. Three studies were conducted in support of these goals. First, a change detection task, investigating the efficacy of each design strategy's ability to garner attention to critical information; dependent variables were both binary (correctly located yes/no) and continuous (time to correct identification). The final 2 studies investigated design performance from a top-down processing frame using an absolute judgement task and a dichotomous decision, forced-choice task.

Dependent variables for each of the final two experiments were accuracy and response time (reported in units of log10ms).

Overall, the results support the novel combination proposed (HL/FOP) as a strategy for communicating critical information. Change detection results support the use of HL, particularly for active ingredient information appearing on the Principal Display Panel, as indicated by a significant interaction between HL and change location for both accuracy and reaction time. In the absolute judgment task, accuracy in drug warning trials increased in the presence of HL (nonHL ME=0.738, SE=0.019 vs HL ME=0.777, SE=0.018; p=0.04), and the presence of an FOP helped garner attention to active ingredient information, evidenced by both FOP treatments (FOP/HL ME=0.910, SE=0.019, vs FOP/nonHL ME=0.908, SE=0.019) being significantly more accurate than the no FOP, nonHL treatment (ME=0.878, SE=0.023; p=0.01). There was also evidence for the efficacy of HL with significantly faster FOP/HL responses (ME=3.902, SE=0.026) than no FOP/nonHL responses (ME=3.944, SE=0.026; p=0.003). Forced choice results also suggest HL increases accuracy and decreases reaction time, evidenced by a significant main effect of HL on accuracy for drug warning trials (nonHL ME=0.952, SE=0.010 vs HL ME=0.974, SD=0.007; p=0.013), and compared to no FOP/nonHL, significantly faster reaction times induced for no FOP/HL treatment in active ingredient trials (no FOP/HL ME=3.670, SE=0.025 vs no FOP/nonHL ME=3.718, SE=0.025) and for both types of HL treatments for drug warning information trials (FOP/HL ME=4.276, SE=0.022; no FOP/HL ME=4.291, SE=0.023 vs no FOP/nonHL ME=4.392 SE=0.023). Results of a secondary analysis investigating familiarity with brand names and active ingredients indicate that participants were significantly more familiar with the brand names (M=7.5, SD=2.52) than the active ingredients (M=3.4, SD=2.54; p<0.001) for all nine of ten products reviewed. When individual brand-active ingredient pairs were investigated, only Advil-Ibuprofen had similar levels of familiarity.

ACKNOWLEDGEMENTS

To my parents for always encouraging me to read and learn and grow. Before I can remember, you read to me and once I could read on my own, you opened up my window to the world even wider by signing me up for a library card. Thank you for fostering my curiosity as a child and for your financial support throughout my time in higher education. I would not be graduating with a PhD if you had not been there supporting my love for learning before my first day of kindergarten.

To Laura Bix, these past 5 years learning from you and working with you have been life changing. Your guidance developed me into the teacher, researcher and scholar I am today. To Mark Becker, I hope your wizardry with experimental design and counterbalancing has rubbed off on me a little. Thank you. Debby Kashy, thank you for your efforts in planning and programming the statistical analysis for these projects. Rafael Auras and Susan Selke, thank you for being supportive committee members and for sharing your packaging knowledge with me. Thank you also to the recruitment partners at MSU Extension, Tri-Counties Area Agency on Aging, and RSVP. You made this work possible.

To Katie Anderson and my Packaging HUB colleagues: Paula Perez, Chancy Zhang, Eric Estrada, Tony Trier, Eric Brunk, Lanqing Liu, Andrew Nathan, Cory Wilson and Shiva Esfahanian; you made this dissertation possible by encouraging me, pilot testing experiments, assisting with data collection, and lots of laughter while learning together. To my writing group, Anicca Cox, Kathy Kim, Chris Fowler, and Adam Lyman, you made me a better writer.

Thank you, Luke for always being here. Your partnership makes everything better.

Finally, I'd like to acknowledge and thank my grandparents: thank you for inspiring my interest in making the world a little bit more accessible for older adults, for participating in my research, and for making Michigan instantly feel like home.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

ADE Adverse Drug Event

ADR Adverse Drug Reaction

AI Active Ingredient information trial

CHPA Consumer Products Healthcare Association

DD Drug-Drug or Drug-Diagnosis Warning information trial

DFL Drug Facts Label

FDA US Food and Drug Administration

FOP Front of Package Label

GSA Gerontological Society of America

HPIM Human Package Interaction Model

ISO International Standards Organization

ME Mean Estimate

NSAID Nonsteroidal Anti-inflammatory Drug

OTC Over-the-Counter

PDP Principal display Panel

REALM-R Rapid Estimate of Adult Literacy in Medicine, Revised

Rx Prescription

SE Standard Error

US United States of America

Chapter 1 Introduction

Packaging serves many functions, including: containment, convenience, communication, and protection (Yam, 2009). While the physical structure of many packages provides much of the functionality, the label of products of all types provides the consumer with important information relating to the identity, contents, information about the product and important directions for use. The importance of the label information can vary from trivial to informative to requisite for safe product use. Over the counter (OTC) are products where the information on the label helps to provide both a protective and a communicative function. Specifically, because OTC labels provide information required to mitigate potential harm from self-medication, the communication of the OTC label functions as consumer protection. Because OTC medication is not risk free, there is an obligation (ethically and legally) to develop packaging and labeling strategies that facilitate varied functions from making cost-effective choices to ensuring safe patient care by communicating information to the consumer at the point of purchase. This dissertation is focused on objectively evaluating a novel strategy for OTC labeling in an attempt to develop OTC labels that are optimized for older adult consumers that is also feasible for implementation in the United States of America (US).

Determining an OTC medication's appropriateness for a unique individual's self-care regimen is a complex decision-making process. Consumers must take on the role of health care provider and make decisions including: the identification of symptoms to be treated, the consideration set of possible treatment options; these must be considered in light of personal information such as budget and how options fit into their potentially complex mix of comorbid health conditions as well as things like dietary considerations and other medications (Bown, Kisuule, Ogasawara, Siregar, & Williams, 2000; Rolita & Freedman, 2008). Each of these tasks requires an adequate level of both health literacy and numeracy, yet research suggests that not all older adult consumers of OTC medication have

sufficient levels of health literacy (Federman, Sano, Wolf, Siu, & Halm, 2009). In light of these complexities, treatment via OTCs requires labeling that readily enables safe decision makings.

As healthcare costs continue to rise and lifespans increase (Dieleman et al., 2017; Jacobsen, Kent, Lee, & Mather, 2011), 75% of older adults state they are opting to use lower cost OTC medication to treat common maladies (Mintel, 2018). While there are many benefits accompanying the use of OTC drugs for older consumers, there are also serious health risks, with as many as 15% of older OTC medication users being at risk of a serious drug-drug interaction (Qato et al., 2008; Qato, Wilder, Schumm, Gillet, & Alexander, 2016). Holden et al.'s 2018 publication on older consumers OTC decision making proposes a model with two styles of decision making; habit-based versus deliberative (Holden et al., 2018). Habit-based OTC decision making primarily relies on the processing of product information tangentially, rather than a goal of accessing and processing detailed product information. Habit-based decision making is likely to occur when making routine purchases, including some routine OTC medication purchases, such as restocking a medicine cabinet (Holden et al., 2018). Deliberative decision making however is more involved and includes an explicit information processing goal. This type of processing is likely to occur when making OTC medication decisions to treat new ailments or if the patient is concerned about a new medication due to comorbid health concerns (Holden et al., 2018). An ideal label would function well for both types of decision making; to be optimized, label designs must perform under both scenarios of decision-making by older adults selecting OTC medication.

This set of studies address the following research questions:

1. What is the effect of highlighting on attracting attention to critical OTC information both when the highlighted information is and is not germane to the explicit goal of the patient?

2. What is the effect of moving critical information to the Principal Display Panel (PDP) of OTC packaging on attracting attention both when it is and is not the explicit goal of the patient?
3. What is the combined effect of both moving critical information to the PDP of OTC packaging and highlighting critical information on attracting attention both when it is and is not the explicit goal of the patient?

Resolving these broad questions is the foundation of the research objectives for this dissertation.

Project Overview

Work presented herein, undertakes two aims, encompassing a total of three studies. The first study in this dissertation addresses

Aim 1: identify label formatting techniques that attract attention to critical information when accessing that information *is not* the person's explicit goal. It is comprised of a change detection study (method introduced in later chapters) with the goal of determining the visual saliency of the proposed label formats when attention is likely engaged in bottom-up fashion. That is, when the viewer is engaged in bottom-up processing of the labels, not tasked with a specific goal. The second and third studies, collectively, address

Aim 2: identify label formatting techniques that attract attention to critical information when accessing that information *is* the person's explicit goal. Both studies are repeated measures studies with the goal of evaluating how label format enhances consumer knowledge of a given product, and how label format facilitates cross-product comparisons. Because each asks the viewer to engage varied and explicit pieces of information to make decisions about the product, Aim 2 objectively examines how the design of information impacts top-down processing mechanisms.

This dissertation is organized in the following way. Chapter 2 is a comprehensive literature review which evaluates the existing work and to identify the gaps in knowledge related to the following areas: the risks and benefits of OTC medication use, OTC decision making by older adults, risk perception of OTC medication, and packaging's role in facilitating safe and effective use of OTC products. Chapter 3 supports Aim 1; older adults (n=60) participated in a change detection methodology to investigate how highlighting critical information and presenting critical information on the front of OTC packaging impacts bottom-up processing of OTC labels for older adults. Chapter 4 supports Aim 2; specifically, it presents an experiment investigating how the labeling strategies support top-down processing among older adults. Older adults (n=75) performed a dichotomous, absolute judgment task where they responded to yes-no questions about OTC medications with varied label formats with accuracy of question response and time to correct response serving as dependent variables. Chapter 5 also supports Aim 2 and builds on the work started in Chapter 4 to examine how the OTC label format can help or hinder cross-product comparison. This chapter presents the results of a dichotomous forced choice task in which participants (n=49) are presented with a single question about two OTC labels, and must select the product that best answers the question posed in the experiment. As with the first experiment in Aim 2, the accuracy of correct response and time to correct response each serve as dependent variables in the analysis to objectively evaluate the performance of varied label designs.

The work was supported by an NIH R01 grant supported under the Call PD 27979 entitled: "*Optimizing OTC labels for older adults: Empirical Evaluation of Labels designed to provide older users the information they need to minimize adverse drug events.*"

Chapter 2 Literature Review

OTCs are a popular treatment alternative for many consumers because they provide a cost-effective and readily available way to treat illness and provide relief from symptoms. In the US, OTCs are available in many different retail environments, including grocery stores and online pharmacies, largely because fewer legal restrictions act as barrier to manufacturing, distribution, sale and purchase than alternative therapies (e.g. Rx). Their ubiquity is just one of many reasons why OTCs are a popular first choice for minor ailments (Consumer Healthcare Products Association, 2010).

Data collected by the Consumer Healthcare Product Association (CHPA), a trade association comprised of OTC manufacturers, suggests that 93% of adults in the United States prefer to treat their own minor ailments with an OTC before seeking medical advice, 81% report actually using OTCs as their first response to minor ailments, and 86% of adults believe OTC medication lowers the cost of healthcare for people like them (Consumer Healthcare Products Association, 2010). Overall, adults believe that OTC medication is just as safe and effective as prescription medication when taken according to directions, and healthcare professionals tended to agree (Consumer Healthcare Products Association, 2010). That said, safe and effective use of OTC products is dependent on thoughtful engagement on the part of the consumer. Specifically, self-medicating patients should read, understand, *and* follow relevant directions and warnings on the medication label, yet our review of the literature relating to consumer behavior with OTC medication use calls this axiom to question.

Surveys of consumers of OTC medication are conclusive— manufacturers cannot count on every consumer reading the label in its present form (McNeil Consumer Healthcare, 2015; Wazaify, Shields, Hughes, & McElnay, 2005). Reasons consumers have difficulty reading the comprehensive drug information provided on the package are multifactorial. They include: readability (Trivedi, Trivedi, & Hannan, 2014; Wogalter & Vigilante, 2003), font size (Hellier, Edworthy, Derbyshire, & Costello, 2006; Murty & Sansgiry, 2007; W. H. Shrank et al., 2007), inadequate color contrast

(Rousseau, Lamson, & Rogers, 1998), low levels of risk perception (Bongard et al., 2002; Cryer, Barnett, Wagner, & Wilcox, 2016; Wilcox, Cryer, & Triadafilopoulos, 2005), differing levels of involvement in the purchase of OTC medication (Reisenwitz & Wimbish, 1997), and low levels of health literacy (Mullen, Curtis, et al., 2018; M. S. Wolf, Gazmararian, & Baker, 2005; Yin et al., 2009).

Despite the reasons for incomplete engagement and use of OTC product labels, the health risks associated with taking OTC medication with incomplete knowledge remain (Hellier et al., 2006; Schmiedl et al., 2014; Trivedi et al., 2014). This tendency to not always read and act on the information provided with an OTC drug can lead to health ramifications for at-risk populations; such as, patients with comorbid conditions, patients who engage in polypharmacy, and aging patients (Lavan & Gallagher, 2016).

User Processing of OTC Warning Labels

One framework for understanding specifically how humans process information used to characterize interactions between people and packaged products is the Human-Package Interaction Model (HPIM). Adapted from information processing (Dejoy, 1991) and human computer interaction theory(Card, Moran, & Newell, 1983; Shackel, 2009) and proposed by de la Fuente (de la Fuente, 2013), the HPIM purports that within a given context (e.g. grocery store, closet in the middle of the night, driving a car, etc.) interaction occurs between package information and a person with a task or goal. This person will use their perceptual system to take in information, use their cognitive system to process the input, and, finally, engage the motor system to action (Card et al., 1983). While the person is going through the 5 stages of information processing: exposure, perception, encodation, comprehension, and action, the package is simultaneously providing both static and dynamic information to be processed. Examples of static information, information that would not change over the life of the package, including text printed on the label. Dynamic information, in contrast, changes over time. Dynamic information would include tactile feedback when the closure is turned the wrong

direction in an attempt to open or changing product attributes, such as visible fill level through the transparent package wall.

The HPIM, by its nature, suggests that processing occurs sequentially. That is, late-stage processing stages (comprehension and action) are dependent on successful completion of the early stages (exposure, perception and encodation). Despite the fact that early stages of processing are requisite for later stages, research related to the processing of information from OTC labels specifically, to date, has tended to focus on later stages of information processing (Brass & Weintraub, 2003; King et al., 2011; Murty & Sansgiry, 2007; Sansgiry, Cady, & Shubhada, 2001; Tong, Raynor, & Aslani, 2014, 2018; Trivedi et al., 2014) rather than the prerequisite early stages (Bix, Bello, Auras, Ranger, & Lapinski, 2009; Gawasane, Bix, de la Fuente, Sundar, & Smith, 2012; Raghavan, Paliwal, & Slattum, 2017). Additionally, at all stages of processing, a majority of drug labeling research has focused on the labeling of prescription products (Bailey, Navaratnam, Black, Russell, & Wolf, 2015; Bojka, Gaddy, Lew, Quinn, & Israelski, 2005; Davis et al., 2009; Davis, Wolf, Bass, Middlebrooks, et al., 2006; Lee, Ladoni, Richardson, Sundar, & Bix, 2019; Morrell, Park, & Poon, 1989; Mullen, Duhig, et al., 2018; W. Shrank, Avorn, Rolon, & Shekelle, 2007; W. H. Shrank & Avorn, 2007; Sundar, Becker, Bello, & Bix, 2012; van Beusekom, Kerkhoven, Bos, Guchelaar, & van den Broek, 2018; Webb et al., 2008; M. S. Wolf et al., 2011, 2016; M. S. Wolf, Davis, et al., 2007).

Table 2.1 summarizes the literature reviewing research focused on OTC medication labeling. It is framed by the information-processing model adapted to packaging by de la Fuente (de la Fuente, 2013) from human-computer interaction theory (Card et al., 1983) and warning processing theory (Dejoy, 1991).

Table 2.1 Information Processing of OTC Medication Labels and Older Adults	
Stage of Information Processing	*Key Findings*
Early Stages (Attention) ↕ Late Stages (Comprehension & Action)	• Age related warnings are not included on most OTC packages (Raghavan et al., 2017). • Warnings that are legally required to be conspicuous are not the most noticeable feature on OTC labels (Bix et al., 2009). • Older adults often do not access the information presented in the DFL (Liu, 2016).
	• Readability of labels is the primarily focus of work investigating the middle stages of information processing of OTC medication labels (Trivedi et al., 2014; Wogalter & Vigilante, 2003)
	• Warning wording, appropriate icons, and formatting are crucial for improving understanding risks associated with OTCs (King et al., 2011). • Warnings on OTC Ibuprofen were rated more difficult to read and understand than the Harvard Law Review (Trivedi et al., 2014). • Standardized labels outperform on consumer preference-based tests of usability but do not always outperform on comprehension metrics (Murty & Sansgiry, 2007; M. P. Ryan & Costello-White, 2017; Tong et al., 2014, 2018).

In 2013, an industry group of OTC manufacturers, the Consumer Healthcare Products Association (CHPA) and the Gerontological Society of America (GSA) together identified the dearth of information focused on OTC labeling use and decision making by older adults. This led to the formation of a panel of experts in OTCs and decision making. The panel of experts convened by CHPA and the GSA urged more research that could be utilized to develop OTC labeling optimized for older consumer use (Albert et al., 2014). Additionally, the panel suggested the need for more research specific to OTCs regarding the roles of health literacy, caregiving, and technology, as well as the role of clinicians (Albert et al., 2014) on information processing for older consumers considering

these products. A more in-depth explanation of why older adults' use of labeling is the focus of this work is presented in table 2.3.

Usability of OTC Warning Labels

A secondary framework one could use to approach the functionality of the design of OTC Warning Labels is the framework of usability. Usability, as defined by the International Standards Organization (ISO), is the, "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use*"* (ISO 9241-11, 2018). See table 2.2 below for the definitions of the components of Usability. Usability, in combination with the HPIM, affords a framework which helps investigators to dissect and evaluate specific functions of a system, as well as the process that the user must navigate to enhancements that can be made to improve ease of use.

Table 2.2 Definitions fundamental to Usability

Terminology	Definition
Effectiveness	"…the accuracy and completeness with which users achieve specified goals…Effectiveness represents the extent to which actual outcomes match intended outcomes…Lack of effectiveness can result in outcomes that could cause harm from use*"* (ISO 9241-11, 2018)
Efficiency	"…is the resources used in relation to the results achieved. These resources include: time, human effort, money and materials" (ISO 9241-11, 2018)
Satisfaction	*"*Extent to which the user's physical, cognitive and emotional responses that result from the use of a system, product or service meet the user's needs and expectations." (ISO 9241-11, 2018)

Herein we utilize the usability construct while attempting to optimize and evaluate the system (i.e. OTC medication labels). The specified user is an older adult, lay consumer of OTC medications. In the context of OTC labels, effectiveness is related to whether or not the label's intended message is received and interpreted correctly by the viewer of the label so that they can act upon the received information. Effectiveness, in this dissertation, is defined as the label capturing enough of the consumer's attention that they perceive the warning information (early stages of information processing) *and* correctly interpret the warning information (late stages of information processing), as measured by accuracy in each of the included studies. For this work, efficiency is measured via response time, as the amount of time it takes one to respond to the task at hand is a proxy for the ease or difficulty of the task.

Existing, published usability evaluations focused on OTC warning labels investigate user satisfaction (M. P. Ryan & Costello-White, 2017; Tong, Raynor, & Aslani, 2015), and thus this dissertation fills a gap in the literature by providing further insight into the efficiency and effectiveness of a standardized OTC label and addition design features intended to enhance communication.

Problems Associated with OTC Label Processing

Inherent in human information-processing is the assertion that humans have limited processing capabilities, and thus, "consumers tend to minimize their information processing effort and are consequently sensitive to any factor, including information presentation format, that affects the ease of processing," (Simonson, 1999). In order to navigate the world with limited processing capabilities, humans and primates developed visual processing systems that rely on visual saliency to prioritize specific visual information out of a scene (Treue, 2003; Veale, Hafed, & Yoshida, 2017). The visual saliency, or the amount an object stands out compared to the surrounding objects, is one component that predicts the amount of visual attention that will be allocated to it (Itti & Koch, 2000). Visual attention can either be unconsciously attracted via the overtness of the object or allocated via

conscious effort. Unconscious attention is referred to as "bottom-up processing", while attention that is consciously allocated is referred to as "top-down processing" (Itti & Koch, 2001; Kinchla & Wolfe, 1979). For a warning label to be effectively and efficiently utilized, it needs to facilitate both types of processing. The warning needs to stand out in the visual field in order to attract attention as well as provide information in a manner that is easily processed for both early-stage processing (associated with perception and encodation (see Table 2.1), later stages (comprehension and action) is also a noted problem for medication labels.

The importance of both early-stage attention garnering, and later-stage understanding is exemplified by studies focused on readability of OTC labels. Accordingly, if the label sufficiently garners attention, that is not adequate on its own: the information also needs to be able to be encoded and comprehended to be useful in consumer decision making. Despite the important role of OTC labeling as a critical source of usable information for lay consumers, OTC labels have been suggested as requiring a relatively high level of reading ability (Trivedi et al., 2014) (late stages), and feature text in too small of a font size (Murty & Sansgiry, 2007) for most older adults to access necessary information (early stages). One study (Trivedi et al., 2014) of the reading ease and grade level required to comprehend nonprescription medication labels (n=40) reports the average reading level required to understand all of the labels to be 16+/- 5 years, or the equivalent of a Bachelor's Degree. When assessed across products, the most difficult subset of labels to comprehend were Nonsteroidal Anti-inflammatory drugs (NSAIDs), also the most culpable in Adverse Drug Reactions (ADRs). Specifically, NSAIDs required a grading level of 22+/- 3 years, or the equivalent of a graduate degree (Trivedi et al., 2014). The reported reading levels significantly exceed the average education level of adults over the age of 65 in the US; as of 2015, only 26.7% (+/- 0.8%) of this population had completed a bachelor's degree or more (C. L. Ryan & Bauman, 2016).

An additional complicating factor is the methods used to evaluate the labels. Most of the studies that have investigated the usability and readability of OTC labels are qualitative assessments of OTC label formatting alternatives, and are thus measuring participants' *perceived* usability and *perceived* readability thru questions about conjectured use (Roe, Levy, Brenda, & Derby, 1999; Tong et al., 2014, 2015, 2018). While the qualitative methodologies are useful for understanding consumer preferences and estimating the likelihood of consumer acceptance of a given label format, these methodologies do not provide direct measures of the needed objective assessment of different labeling formats on noticeability, encodation, or comprehension of OTC labels. Within the limited set of studies quantitatively investigating use of a label (rather than consumer preferences), the available literature reveals a concentrated focus on late stage processing, primarily comprehension (Brass & Weintraub, 2003; Murty & Sansgiry, 2007; M. P. Ryan & Costello-White, 2017; Sansgiry et al., 2001) rather than of noticeability, or ease of encoding, both of which are prerequisite to the late stage processes (cognition).

Adverse Drug Events and Adverse Drug Reactions

A potential risk that consumers face in taking all medications (both OTC and Rx) involves suffering the consequences of an adverse drug event (ADE), defined as an injury resulting from medical intervention related to a drug (Kohn, Corrigan, & Donaldson, 2000). Within the broad category of ADEs, Adverse Drug Reactions (ADRs) are a specific subset of interest to researchers who study ways to reduce number of occurrences of ADEs. An ADR is defined as, "an appreciable harmful or unpleasant reaction resulting from an intervention related to the use of a medicinal product, which predicts hazard from further administration and warrants prevention or specific treatment, or alteration of the dosage regimen or withdrawal of the product," (Edwards & Aronson, 2000).

Traditionally, an ADR is distinct from an ADE due to the requirement of a causal relationship between the drug and the adverse occurrence rather than simply a temporal relationship between the two (International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1994). A more modern definition of an ADR further distinguishes causality through classification of whether or not the event is considered dose related or augmented (Type A), non-dose related or bizarre (Type B), dose-related and time related (Type C), time-related (Type D), withdrawal (Type E), or an unexpected failure of therapy (Type F). Type A ADRs are often predictable, and thus preventable, while type B are unpredictable and thus more difficult to prevent (Edwards & Aronson, 2000). Herein, we are studying an intervention (labeling), with the potential to enable people to make informed choices regarding the selection of OTC drugs. Specifically, this would potentially reduce Type A ADRs.

One important aspect of the definition of an ADR is that it encompasses all medicinal products, including OTCs, herbal medication and dietary supplements as well as prescription medication (Rx). Because OTCs lack the oversight of a learned intermediary (e.g. a pharmacist or prescribing physician), the label of an OTC takes on a unique role in ensuring the patient is equipped to safely and effectively use the product by providing critical information during decision making. This lack of guaranteed oversight from learned intermediaries potentially places all of the onus for identifying drug-drug or drug-diagnosis interactions with the potential to result in a Type A ADR on the consumer themselves.

ADRs are more prevalent in the population of older adults than other sectors of the population. Studying the self-medicating[1] behaviors of older adults is important not just because of the increased risk for ADRs, but also because adults over 65 are the fastest growing demographic in

[1] Self-medication is a broad term generally used to describe any use of medication that is not prescribed by a licensed professional, such as a physician or a dentist, though some researchers use more precise definitions(Jerez-Roig et al., 2014).

the United States (Jacobsen et al., 2011). As people age, health generally declines and the number of health problems requiring daily medication increases (Qato et al., 2008). Despite comprising approximately 15% of the total population, older adults consume about 35% of all prescription drugs and 30% of all OTC drugs (The Gerontological Society of America, 2013). Over 90% of adults over the age of 65 report taking at least one medication daily, and about 50% report taking 5 or more medications per day (Qato et al., 2016). This suggests that approximately half of US citizens over the age of 65 are engaged in polypharmacy, defined as, "5 or more medications used daily" (Masnoon, Shakib, Kalisch-Ellett, & Caughey, 2017), described as the tendency to take multiple medications to treat comorbid health conditions. Polypharmacy is associated with an increased risk of ADR (e.g., a reaction that results from drug-drug interaction), and because older adults are more likely to take numerous medications regularly it follows that they are at increased risk (Guthrie, Makubate, Hernandez-Santiago, & Dreischulte, 2015).

One study, (Franceschi et al., 2008) conducted in Italy documenting the prevalence and avoidability of ADRs that lead to hospitalizations of older adults found an ADR rate of 5.8% of hospital admissions, of those hospitalized, the most culpable class of medicines was indicated as NSAIDS (23.5% of ADRs, often available in OTC forms at low doses). NSAIDs were followed by oral anticoagulants (20.6%), and low dose aspirin (13.7%), another drug frequently available via OTC purchases in much of the world (Franceschi et al., 2008). A German study (Schmiedl et al., 2014) reports a hospitalization rate due to ADR of 3.2%, with most (96.1%) of the ADR admissions caused by prescription drug use, and the remaining 3.9% caused at least partially attributable to self-medication with OTCs (Schmiedl et al., 2014). In an additional investigation into the relationship between self-medication and ADRs in France, (Asseray et al., 2013) 9.8% of the *hospitalized* participants (2% of the total number of participants reporting self-medication behaviors) in the study were diagnosed with an ADR related to self-medication (Asseray et al., 2013). Overall, 63.7%

of the total number of participants in the study reported self-medication; with 50.1% of those who reported self- medicating indicating the use of OTCs. 59.9% of study participants reported taking a prescription medicine in the two weeks leading up to hospitalization (Asseray et al., 2013). When defining what is considered an ADR, it is important to include all forms of medication because an ADR can be more difficult to prevent for those self-medicating with OTC products.

Risk Factors for ADRs in older adults

In addition to the increased risk imposed by engaging in polypharmacy, older adults have increased susceptibility for ADRs for a myriad of reasons including changes in pharmacokinetics and pharmacodynamics[2] as the body ages and being at risk for lower health literacy (Davies & O'Mahony, 2015; Kobayashi, Wardle, Wolf, & Von Wagner, 2016; Veehof, Jong, & Haaijer-Ruskamp, 2000). The coupling of physiological changes compounded by increasingly complicated drug regimens have been reported by multiple researchers as probable reasons for the increased susceptibility of older adults to serious ADRs requiring hospitalization, as compared to younger patients (Atkin, Veitch, Veitch, & Ogle, 1999; Davies & O'Mahony, 2015; Lavan & Gallagher, 2016; Mannesse, Derkx, de Ridder, Man In 'T Veld, & Van Der Cammen, 2000; Nair et al., 2016; Routledge, O'Mahony, & Woodhouse, 2004). Table 1.2 summarizes some of the key findings about risk factors for ADRs that apply specifically to older adults (defined as over the age of 65), particularly emphasizing findings that investigate OTC medications.

[2] Pharmacokinetics is a term that describes how the body metabolizes or processes a drug ("Pharmacokinetics - an overview," n.d.), while pharmacodynamics is a term used to describe the effects the drug has on the body ("Pharmacodynamics - an overview," n.d.).

Table 2.3 Risk Factors for Adverse Drug Reactions	
Risk Factor	Key Findings
Age	• Age is strongly associated with an increased risk of suffering an ADR, but the literature is divided on whether numerical age is itself a causal factor (Atkin et al., 1999; Bourgeois et al., 2010; Nair et al., 2016; O'Connor et al., 2012; Oscanoa et al., 2017). • Changing pharmacodynamics and lean body mass percentage might be the causal factors associated with age (Lavan & Gallagher, 2016).
Polypharmacy	• Polypharmacy is implicated as a high risk factor for ADRs (Atkin et al., 1999; Bourgeois et al., 2010; Lavan & Gallagher, 2016; Marcum et al., 2012; O'Connor et al., 2012; Oscanoa et al., 2017). • Polypharmacy is increasing in the United States (Guthrie et al., 2015).
OTC Medication Use	• OTC medication use contains inherent risks for patients with comorbid conditions or daily medication regimens (Hess, Linnebur, Rhyne, & Valdez, 2016; Qato et al., 2008; The National Council on Patient Information and Education, 2003; Wold et al., 2005). • Many older adults use OTC medication without full knowledge of the risks that accompany the benefits (Amoako, Richardson-Campbell, & Kennedy-Malone, 2003; Wilcox et al., 2005; Wold et al., 2005). • OTCs are implicated in some ADRs that require hospitalization (Asseray et al., 2013; Franceschi et al., 2008; Schmiedl et al., 2014).
Health Literacy	• Patients with lower levels of health literacy are at risk for misinterpreting OTC drug warnings (M. S. Wolf, King, et al., 2012). • Older adults tend to have lower levels of health literacy than younger adults (Federman et al., 2009; Kobayashi et al., 2016; M. S. Wolf, Curtis, et al., 2012).
Gender	• Older women are more susceptible to ADRs than older men, likely due to the greater loss of lean body mass (Cadigan, Magaziner, & Fedder, 1989) • Older men are more likely to take anticoagulants, a drug class responsible for many ADRs and drug-drug interactions (Qato et al., 2008).
Risk Perception	• Lay consumers are less likely to perceive OTCs as risky than medical professionals (Bongard et al., 2002). • Older adults tend to have a lower perception of the risk of OTC medication than younger adults, and that risk perception appears to be informed by their history of medication use (McNeil Consumer Healthcare, 2015; Wawruch et al., 2013; Wilcox et al., 2005).

Preventing ADRs through Packaging and Labeling Regulations

One of the most commonly utilized strategies for standardized communication of product attributes and warnings is labeling and packaging. In 1966, the Fair Packaging and Labeling Act (FPLA) was passed by the US Congress to enable consumers to make informed value comparisons in the marketplace (Wall, 2002). For food, drugs, and cosmetic products, the FPLA authorized the US Food and Drug Administration (FDA)[3] to establish standardized rules for the communication of net contents of a product, the name and address of the manufacturer, distributor, or packer, and a statement of identity (21 CFR § 201.61-201.62). In addition to requiring these claims, the FPLA also defined the Principal Display Panel (PDP), as "part of a label that is most likely to be displayed, presented, shown, or examined under normal and customary conditions of display for retail sale." (15 USC Ch. 39 § 1459(f)) and required the statement of identity and the net contents to appear on the PDP (21 CFR § 201). In the 1960's-1970's the US FDA (Food and Drug Administration, 2006), acting under authority granted by the US Congress in the Federal Food Drug and Cosmetic Act, started initiatives to enhance safety and effectiveness by improving the labeling of drug products "as to render it likely to be read and understood by the ordinary individual under customary conditions of purchase and use" 21 U.S.C. 352(c). These goals were first accomplished by requiring package inserts stating risks and benefits be included with prescription drugs in 1970, and then through the creation of the Over-the-Counter Drug Review in 1972 to address drugs available without prescriptions (Food and Drug Administration, 2006).

The next major change in labeling requirements for OTC drugs came in 1999 with the introduction of the DFL, a standardized label dictating both standardized content and formatting, requirements for the labeling of OTCs, with the ultimate goal of providing consumers with "easy-to-

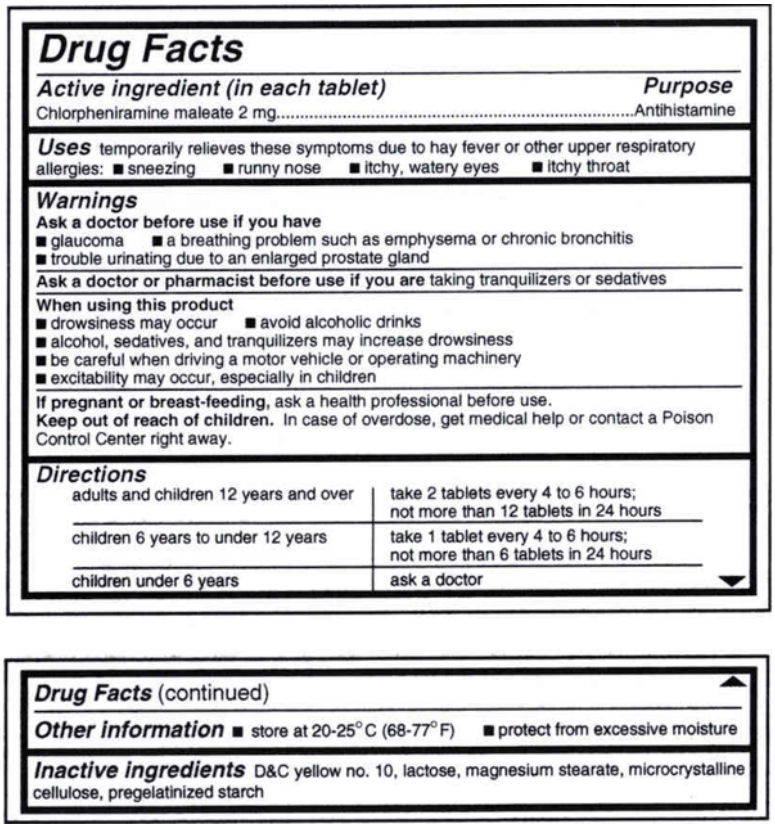
[3] The mission of the FPLA is also carried out by the Fair Trade Commission for consumer products that are not food, drugs, or cosmetics (15 U.S.C. §§ 1451-1461).

find" product information presented in a consistent manner on an information panel, directly to the right of the PDP (Food and Drug Administration, 2006). More recently, organ specific warnings and active ingredient prioritization were added to the DFL of certain OTC drugs which pose a risk for liver damage or stomach bleeding. Despite this effort, little is known about how effective these labels are at attracting consumer attention to effectively communicate information, although some research suggests that consumers tend to use other packaging attributes (such as trade dress or brand name or simple heuristics like color) to make decisions related to the purchase of OTC medication (Aker, Beck, Travis, & Harris, 2014; Gawasane et al., 2012; Liu, 2016).

The standardized DFL is dictated by 21 CFR § 201.66, which requires the use of specific headings to organize product information in the following order: Title (Drug Facts or Drug Facts Continued); Active Ingredient(s); Purpose(s); Use(s); Warning(s); Direction(s); Other information; Inactive Ingredients; and an optional heading of Questions? (or Questions or comments). An example of a DFL is included below in figure 2.1. The prescribed DFL ordering of information creates a hierarchy of information. If one assumes that consumers read, engage with, and understand the entire label, the order or placement of the information should not influence communication of the entire message, unfortunately, this assumption does not appear to be universally true as consumers ignored entire panels of product information in past experiments (Liu, 2016).

There are noted difficulties associated with the labels in their current form which impact differing stages of information processing that can be particularly problematic in older adults. Noted problems for older consumers interacting with OTC labels include: small font size and information density (Murty & Sansgiry, 2007), the relative conspicuousness of the information presented on the label, (i.e. the brand name appearing prominently, while critical safety information appears less prominently)(Bix et al., 2009; Liu, 2016), and an overarching perception that OTCs are innocuous (Hellier et al., 2006) potentially leading to consumers completely ignoring safety information.

While this dissertation focuses primarily on the legal environment in the US, the effort to optimize labeling to motivate safe and appropriate OTC use is not only a priority in the US, but a global public health effort (Mintel, 2018; Popescu, 2014). The globality of efforts is demonstrated by the recent activities of the Australian Therapeutic Goods Administration which passed regulation requiring changes to the labeling of OTC medication, which created the "Medicine Box," an equivalent to the DFL (Austrailian Government Department of Health, 2011; Austrailian Government Department of Health and Aging, 2011; Department of Health Therapeutic Goods Administration, 2016). In Canada, the regulatory authority with jurisdiction over OTC medications, Health Canada, issued updated guidance for Good Label and Package Practices for Non-prescription Drugs and Natural Health Products in 2017 (Health Canada, 2017). Examples of Canadian and Australian OTC labels are included below in figures 2.2 and 2.3.



Drug Facts

Active ingredient (in each tablet) ***Purpose***

Chlorpheniramine maleate 2 mg ..... Antihistamine

Uses temporarily relieves these symptoms due to hay fever or other upper respiratory allergies: ■ sneezing ■ runny nose ■ itchy, watery eyes ■ itchy throat

Warnings

Ask a doctor before use if you have

■ glaucoma ■ a breathing problem such as emphysema or chronic bronchitis
■ trouble urinating due to an enlarged prostate gland

Ask a doctor or pharmacist before use if you are taking tranquilizers or sedatives

When using this product

■ drowsiness may occur ■ avoid alcoholic drinks
■ alcohol, sedatives, and tranquilizers may increase drowsiness
■ be careful when driving a motor vehicle or operating machinery
■ excitability may occur, especially in children

If pregnant or breast-feeding, ask a health professional before use.
Keep out of reach of children. In case of overdose, get medical help or contact a Poison Control Center right away.

Directions

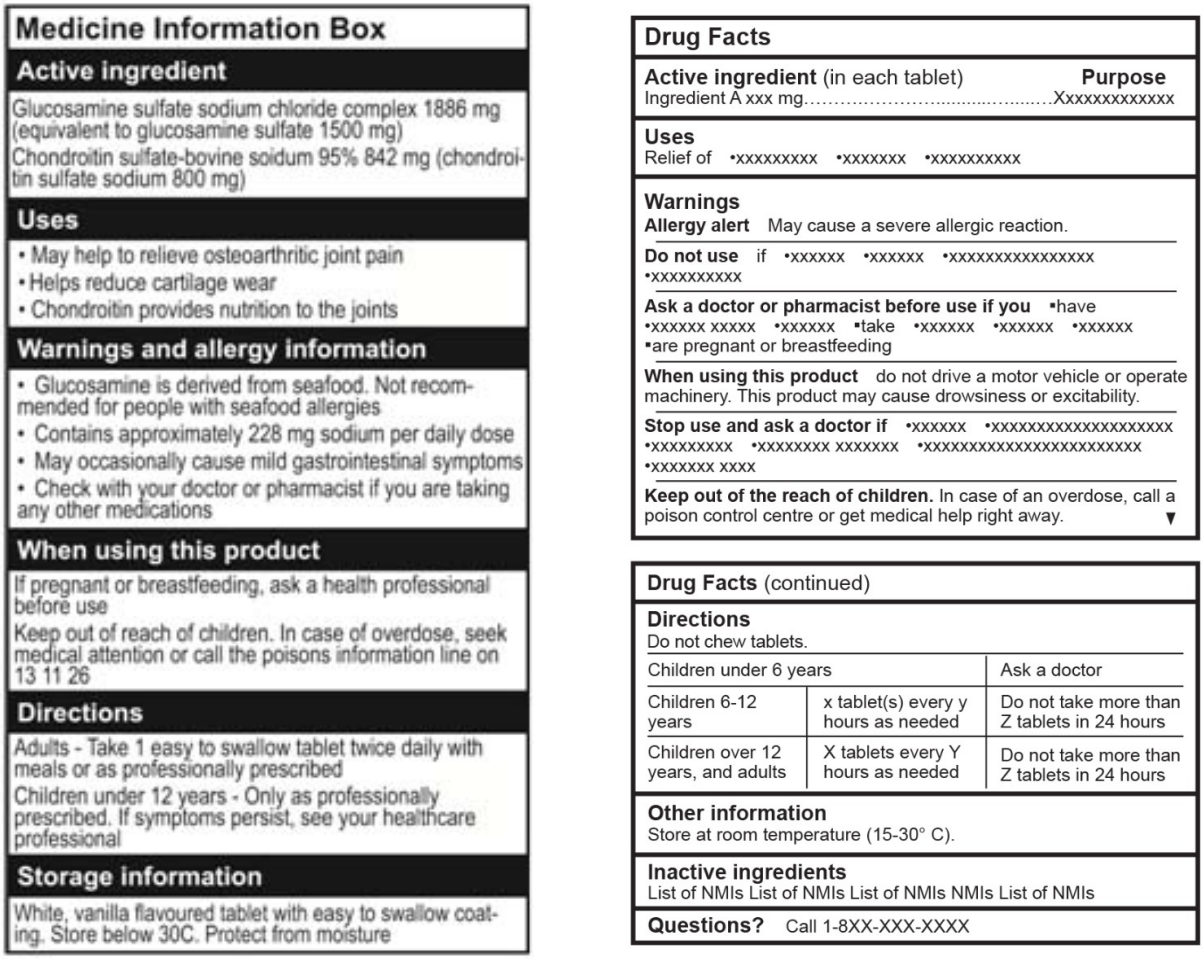
adults and children 12 years and over	take 2 tablets every 4 to 6 hours; not more than 12 tablets in 24 hours
children 6 years to under 12 years	take 1 tablet every 4 to 6 hours; not more than 6 tablets in 24 hours
children under 6 years	ask a doctor

Drug Facts (continued)

Other information ■ store at 20-25°C (68-77°F) ■ protect from excessive moisture

Inactive ingredients D&C yellow no. 10, lactose, magnesium stearate, microcrystalline cellulose, pregelatinized starch

Figure 2.1 An example Drug Facts Label (U.S. Food and Drug Administration, 2017)



Medicine Information Box

Active ingredient

Glucosamine sulfate sodium chloride complex 1886 mg (equivalent to glucosamine sulfate 1500 mg)

Chondroitin sulfate-bovine soidum 95% 842 mg (chondroitin sulfate sodium 800 mg)

Uses

- May help to relieve osteoarthritic joint pain
- Helps reduce cartilage wear
- Chondroitin provides nutrition to the joints

Warnings and allergy information

- Glucosamine is derived from seafood. Not recommended for people with seafood allergies
- Contains approximately 228 mg sodium per daily dose
- May occasionally cause mild gastrointestinal symptoms
- Check with your doctor or pharmacist if you are taking any other medications

When using this product

If pregnant or breastfeeding, ask a health professional before use

Keep out of reach of children. In case of overdose, seek medical attention or call the poisons information line on 13 11 26

Directions

Adults - Take 1 easy to swallow tablet twice daily with meals or as professionally prescribed

Children under 12 years - Only as professionally prescribed. If symptoms persist, see your healthcare professional

Storage information

White, vanilla flavoured tablet with easy to swallow coating. Store below 30C. Protect from moisture

Drug Facts

Active ingredient (in each tablet) **Purpose**
Ingredient A xxx mg……Xxxxxxxxxxxx

Uses
Relief of •xxxxxxxxx •xxxxxxx •xxxxxxxxxx

Warnings

Allergy alert May cause a severe allergic reaction.

Do not use if •xxxxxx •xxxxxx •xxxxxxxxxxxxxxxx •xxxxxxxxxx

Ask a doctor or pharmacist before use if you •have •xxxxxx xxxxx •xxxxxx •take •xxxxxx •xxxxxx •xxxxxx •are pregnant or breastfeeding

When using this product do not drive a motor vehicle or operate machinery. This product may cause drowsiness or excitability.

Stop use and ask a doctor if •xxxxxx •xxxxxxxxxxxxxxxxxxxx •xxxxxxxxx •xxxxxxxx xxxxxxx •xxxxxxxxxxxxxxxxxxxxxxxx •xxxxxxx xxxx

Keep out of the reach of children. In case of an overdose, call a poison control centre or get medical help right away. ▼

Drug Facts (continued)

Directions
Do not chew tablets.

Children under 6 years		Ask a doctor
Children 6-12 years	x tablet(s) every y hours as needed	Do not take more than Z tablets in 24 hours
Children over 12 years, and adults	X tablets every Y hours as needed	Do not take more than Z tablets in 24 hours

Other information
Store at room temperature (15-30° C).

Inactive ingredients
List of NMIs List of NMIs List of NMIs NMIs List of NMIs

Questions? Call 1-8XX-XXX-XXXX

Figure 2.2 The first example is the medicine box label from Australia. © The Commonwealth of Australia (Therapeutic Goods Administration, 2012). The second is the non-prescription drug label required in Canada (Health Canada, 2018)

Although labeling is one of the most common packaging strategies for informing consumer behavior, research suggests that it is not always practically effective at conveying hazard information (Ayanoglu, Duarte, Noriega, Teixeira, & Rebelo, 2012). Understanding mechanisms of visual processing to develop more efficient strategies for labeling that comply with current legal requirements for labeling has been identified as a crucial need (Murty & Sansgiry, 2007; M. P. Ryan & Costello-White, 2017). To further explore how at-risk consumers are utilizing packaging and labeling of medication, some of the different packaging attributes selected as variables of interest in studies investigating the safe and effective use of OTC medication are included in Table 1.3.

Table 2.4 Over the Counter Packaging and labeling challenges to safe and effective use of medication	
Packaging Attribute	Key Findings
Font Size	• The minimum font size legally allowed is too small to be considered legible for older adults (Murty & Sansgiry, 2007). • Increasing the font size of warnings and other critical information could increase noticeability (Hellier et al., 2006).
OTC Drug label use during selection	• A standardized OTC label format is rated higher on usability for consumers, but does not have an associated increase in comprehension or retention of the label (Murty & Sansgiry, 2007; M. P. Ryan & Costello-White, 2017; Tong et al., 2014, 2015, 2018). • Consumers primarily use the trade dress, price, and brand name to inform OTC medication decisions (Aker (Johnson & Drungle, 2000)et al., 2014; Harben et al., 2018; Liu, 2016). • Consumers do not utilize the entirety of the safety information available on the package of OTC products when making decisions about whether or not a product is appropriate to take (Liu, 2016).
Compliance and OTC packaging	• Compliance Packaging tends to be blister packages (Weiss, 2009). • Despite the increased likelihood that an able-bodied patient will comply with the intended drug regimen, blister packages are ranked lower in terms of ease of use for older adults (de la Fuente, Gustafson, Twomey, & Bix, 2015).

Risk Perception of OTC Products and Health Literacy

While there is evidence that format, color, and use of attention attracting signal words influence the perception of risk associated with varied products (Hellier et al., 2006), research which analyzes OTC labels and risk perception is limited. The work that is available suggests that the context in which OTC products are purchased (i.e. a grocery store rather than from behind a pharmacists' counter) influences risk perception of the products (Stevenson, Leontowitsch, & Duggan, 2008), and that healthcare professionals perceive the risks associated with OTC medication use (i.e. how much caution should be taken when consuming OTC drugs) differently than non-health professionals (Bongard et al., 2002). Specifically, the research team found that non-health professionals did not consider the OTC drugs most commonly implicated in ADRs to be risky while healthcare providers

did. Additionally, lay people are also less likely to accurately indicate the drug categories most likely to result in ADRs (NSAIDs or anticoagulants) as compared to health care professionals (Bongard et al., 2002).

Despite being more susceptible to suffering an ADR, it has been suggested that older adults are less likely than their younger counterparts to perceive OTC medication as risky; in one survey only 54% of adults over the age of 70 reported reading the labels for medications that they had used previously as important as compared to 82% of millennial adults surveyed (McNeil Consumer Healthcare, 2015). In a separate study, a large majority of older adults (75%) reported viewing OTC drugs as either "safe" or "mostly safe" (Wawruch et al., 2013). Research suggesting that familiarity with the repeated purchase and use of medical products reduces risk perception of older consumers making routine choices (Johnson & Drungle, 2000; Wogalter, Brelsford, Desaulniers, & Laughery, 1991).

All of this suggests that consumers, particularly older adults with a history of OTC use, perceive OTCs as benign. Due to this perception of safety, rather than a perception of risk, it has been proposed that effective OTC labels need to also provide "refutation text," which refutes consumer misconceptions about the absolute safety of the product they are considering (M. P. Ryan, Costa, & Cruz, 2017). This perception of safety is consistent with qualitative work that we conducted with groups of older consumers (Harben et al., 2018).

One potential reason commonly cited for the differential in risk perception between lay consumers of all ages and healthcare professionals is differing levels of health literacy. Health literacy is defined as "an individual's capacity to obtain, process, and understand basic health information and services sufficiently to make appropriate health decisions." The reported tendency for low health literacy among older adults (Federman et al., 2009; Kobayashi et al., 2016; M. S. Wolf, Gazmararian, & Baker, 2007) is especially concerning as health literacy is regarded as an influential factor in all types

of healthcare decision-making, including the amount of healthcare expenditures (Hardie, Kyanko, Busch, Losasso, & Levin, 2011); specifically, patients with lower levels tend to spend more for care comparable to those with higher literacy levels. Because OTC labels are frequently the sole source of information used by patients (Cheatham & Wogalter, 2002), health literacy is an important consideration for policy, and should be considered carefully by policy makers who create dictates such as the DFL's presentation. A limited number of studies specifically focus on OTC labeling in those at-risk for health literacy (King et al., 2011; Mullen, Curtis, et al., 2018; Yin et al., 2012). A majority of those studying the relationship of health literacy and label interpretation focus on prescription drugs and find misinterpretation of the information to be a problem (Bailey, Agarwal, Sleath, Gumusoglu, & Wolf, 2011; Davis, Wolf, Bass, Thompson, et al., 2006; M. Wolf, 2017; M. S. Wolf et al., 2016; M. S. Wolf, Davis, et al., 2007).

Lessons from Nutritional Labeling

A growing body of literature suggests that poor health literacy is not the only contributor to the inefficient transfer of information presented in the DFL. In eye tracking work (Liu, 2016) examining the effects of changing the prominence of information presented on the PDP's of OTC label (active ingredient, symptom relief, or brand name), researchers investigated older adults' ability to make safe OTC medication choices. Nearly 64% of the participants did not access the DFL to make a decision about whether or not the medication was appropriate for them to consume. Because many participants did not manipulate the virtual package to view the DFL when making decisions about appropriateness, communication of OTC safety information was never exposed by the user, interrupting the earliest stage of information processing (Liu, 2016). Researchers encouraged the development of design strategies intended to catalyze early-stage processing (attention; exposure and perception) of safety information.

Even for the consumers that do engage in information search (turn the package to view the DFL) there are still barriers to processing the information on the DFL (Carpenter & Yoon, 2012; M. P. Ryan & Costello-White, 2017; Trivedi et al., 2014; Wilson & Wolf, 2009). These barriers are primarily due to the format of the information relevant to stopping an ADR in the DFL, as it is presented in a visually dense manner that is difficult for older adults to process (Carpenter & Yoon, 2012; Wilson & Wolf, 2009) and has been noted to be too small to read for many people (Murty & Sansgiry, 2007).

We postulated that by placing information on the panel that tends to be exposed and is always attended (Lui, 2016), the PDP, in a visually salient format (highlighted text), early-stage processing (Table 2.1) would be enhanced for at-risk populations (older adults). Highlighting of key words or phrases was indicated to increase participants' performance in a task evaluating participants' evaluation of aspirin labeling claims in a previous study (M. P. Ryan et al., 2017). Because prior work suggests that consumers frequently and consistently utilize the PDP of OTC medication to make decisions regarding purchase and use, rather than flipping the carton of OTC drugs to the side to utilize the comprehensive DFL, a more purposeful PDP design offers a rich area of inquiry.

A large and growing body of work related to food labeling supports the idea of moving critical information to the PDP, commonly referred to as a "Front of Pack" (FOP) label for these products. The goal of the FOP approach is to facilitate consumer attention to nutrition information and aid cross-product comparisons, ultimately resulting in more healthful selection. An FOP presents truncated information from the comprehensive nutrition information on the package's front, or PDP. Generally, presented nutrients are closely related to with disease states (e.g. fat and saturated fat- heart disease; sugar- diabetes; salt- hypertension). (See (Hawley et al., 2013) for a review of the literature related to the efficacy of the strategy for food).

While there are many styles of nutritional FOPs, including: health logos, traffic lights, summary indicators, and warning labels (Kanter, Vanderlee, & Vandevijvere, 2018), each style has the goal of catalyzing consumer attention and understanding of nutritional information that tends to be related to disease states (e.g. saturated fat, fat, sodium, sugar). Multiple researchers have found that the simplified format of FOPs, especially when combined with a traffic-light color coding system [4] and prominent positioning garners attention readily (Bialkova et al., 2014; Bialkova, Grunert, & van Trijp, 2013; Bialkova & van Trijp, 2010; Koenigstorfer, Wąsowicz-Kiryło, Styśko-Kunkowska, & Groeppel-Klein, 2014), facilitates cross-product comparisons (Hersey, Wohlgenant, Arsenault, Kosa, & Muth, 2013; Jones & Richardson, 2007; Kelly et al., 2009) and increases healthful purchasing decisions (Levy, Riis, Sonnenberg, Barraclough, & Thorndike, 2012; Thorndike, Riis, Sonnenberg, & Levy, 2014). In short, there is evidence that these labels improve all stages of information processing when placed on the PDPs of packaged foods. In fact, while Ryan et al. did not use the language of an FOP to describe the recommendation in the study referenced earlier investigating highlighting and aspirin labels, they did suggest a directive to "always read the label" or "see new warnings information" as a potential strategy to induce meaningful engagement with the DFL (M. P. Ryan et al., 2017).

Conclusion

Based on the extensive, growing body of research available from the field of food labeling, two strategies will be used to increase visual saliency (early-stage processing, see table 2.1) and enhance cognitive processing (late stages of processing, see table 2.1) of information critical for the safe and effective use of OTC products. Research proposed herein will objectively evaluate the use of an FOP incorporating information which (if heeded) is likely to result in prevention of an ADR; additionally, based on the insights of others which suggest the use of colored highlighting also enhances

[4] Traffic Light Color Coding in this instance refers to using the colors green, yellow, and red to signal if something is healthy, less healthy, or unhealthy (Thorndike et al., 2014).

information processing, we test this postulate for critical information on the labels of OTCs. We expect that increased visual salience of the information relevant to ADR prevention will, in turn, increase the likelihood participants make more efficient and safe self-medication decisions. Testing methods are detailed in Chapters 3, 4 and 5.

Chapter 3 A Change Detection Study

Overview

Common information processing models posit that processing occurs in a serialized sequence of steps (see table 1.1 in Chapter 1). Under this construct, early stages (exposure and attention) are pre-requisite to later stages (encodation, comprehension and action). Early-stage processing is sometimes completed by involved consumers utilizing purposeful search for needed information, top-down processing of a label. However, data suggests many consumers of OTC medication do not actively seek or engage much of the comprehensive, regulated information that is required to be present on OTC packages (Harben et al., 2018; Liu, 2016; McNeil Consumer Healthcare, 2015). Most specifically, in Liu's 2016 work, approximately 64% of participants did not turn beyond the PDP to examine the more comprehensive information in the DFL(Liu, 2016). As a result, early stages, pre-requisite to further processing, are not fulfilled leaving one to wonder if this information influences decision making. While this information could be held in the consumer's memory and incorporated into the decision making process, the evidence discussed previously suggests otherwise (Liu, 2016). Our review of the literature suggests a gap in knowledge specific to how OTC labeling techniques perform in the early stages of information processing; (see table 2.1) specifically, how different labeling strategies work to garner attention to the critical information. In other words, how effective are different labeling approaches at inspiring consumer attention to critical information that is needed for the safe and effective use of OTC products?

With the gaps in the literature in mind, this study proposes a novel OTC label format inspired by the success of nutritional Front of Pack (FOP) labels at garnering attention (Becker, Bello, Sundar, Peltier, & Bix, 2015; Bix et al., 2015). The methodology presented herein was piloted with older adults in 2018 (Esfahanian, 2020) to: inform this study design, provide pilot data and as a proof of concept. Highlighting at two levels (present vs absent) was crossed with label type (FOP present vs FOP absent)

for a total of four treatments of interest (FOP with highlight; FOP without highlight; no FOP with highlight and no FOP without highlight). See figure 3.1 for examples of the four label treatments being evaluated in this study. Full sized versions of the labels are provided in Appendix A as a reference.

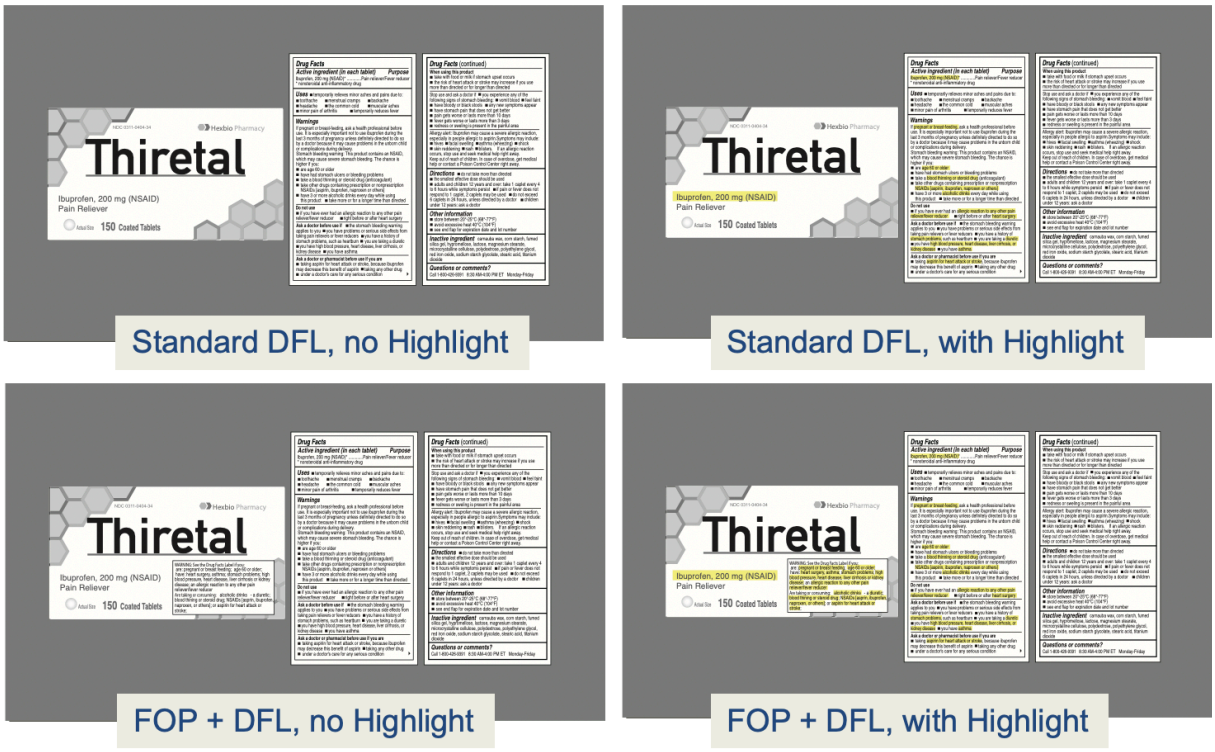


Figure 3.1 The four label treatment styles that are being evaluated in this dissertation

The overarching objective of this study is the development of design strategies for OTC labels that are effective at all stages of information processing for older adults (a population more likely to have an ADR than other sectors of the population). Proximal to this goal, herein, we objectively investigate two formatting techniques ability to attract attention to information critical for the safe and effective use of OTCs (highlighting important safety information and introducing an FOP label with important safety information) when accessing it is *not* the participant's explicit goal (bottom-up attention). The scientific hypotheses (Gotelli & Ellison, 2004) are:

Hypothesis 1: Highlighting will increase the accuracy of participants noticing changes compared to non-highlighted labels**.**

Hypothesis 2: Changes occurring in the front of pack label will increase the accuracy of participants' noticing of changes compared to comparable changes in the drug facts label.

Hypothesis 3: Highlighting will decrease the amount of time required for participants to notice changes compared to non-highlighted labels.

Hypothesis 4: Changes occurring in the front of pack label will decrease the amount of time required for participants to notice changes compared to comparable changes in the drug facts label.

Methods and Materials

The materials and methods section of this chapter first discusses the design of the experiment, secondly describes the materials and methods used to generate the experimental stimuli, thirdly discusses the recruitment and data collection procedures including the screening criteria, and finally describes the statistical analysis strategy and methodology used to analyze the data. Methods were approved by the MSU Psychology and Social Science Internal Review Board in Summer 2018 as STUDY00000832.

Experimental Design

A change detection experiment, or flicker task experiment, was conducted to examine the amount of time it takes a participant to notice different aspects of mock-branded OTC labels across the four treatments previously described. A flicker task experiment is structured so that the participant is seated in front of a computer screen that alternates between two images, with a blank screen briefly appearing in between each image (Rensink, O'Regan, & Clark, 1997). This methodology has been utilized with older adults to examine the changes in visual attention while driving (Costello, Madden, Mitroff, & Whiting, 2010; Hoffman, Atchley, McDowd, & Dubinsky, 2005; McCarley et al., 2004; Pringle, Irwin, Kramer, & Atchley, 2001; Veiel, Storandt, & Abrams, 2006), but our literature review

suggests this experiment to be among the first applying the flicker task methodology to assess the noticeability of different pieces of information on an OTC label utilizing this vulnerable population.

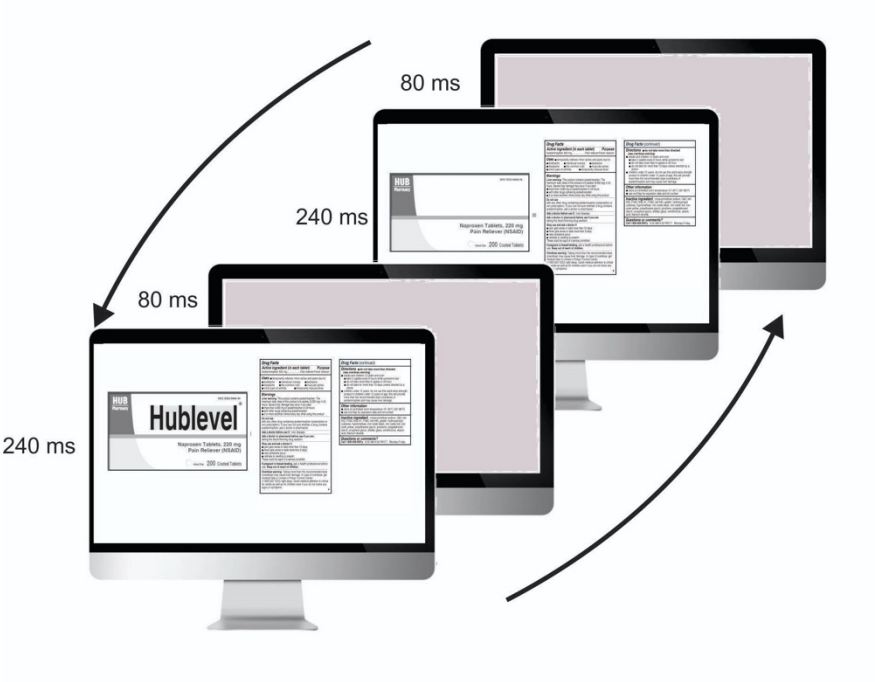


Figure 3.2 Change Detection Method: Trial depicts the cycle of images displayed and timing of the images with the standard DFL treatment with no highlighting of critical information. Reprinted from the original grant submission Call PD 27979 entitled: "*Optimizing OTC labels for older adults: Empirical Evaluation of Labels designed to provide older users the information they need to minimize adverse drug events*"

Timing was based on the work of Rensink et al. (1997) which dictates the stimulus of interest appear for 240 ms followed by a brief grey screen for 80 ms, then the stimulus image (slightly altered from the stimulus of interest) for 240 ms and followed by another grey screen. This sequence is shown in a loop such that there is one difference between the two images, which results in one aspect of the image changing as the screen appears to "flicker" in the location of the change. The participant is tasked with identifying the location of that difference and is instructed to pause the program by depressing the space bar to signal that the changes has been detected as quickly as possible. After the testing is paused, they are instructed to utilize a mouse to click in the area of change to verify that they have, indeed, accurately located it. Participants were provided 4 practice trials to get acquainted with

how to operate the experimental program, with additional help offered by a research assistant for participants who were uncomfortable operating a computer mouse.

Because the amount of time it takes to detect a change is a validated proxy for attention, (Rensink et al., 1997) change detection studies can be applied to labeling as a means of quantifying attention given different design strategies (Bix, Kosugi, Bello, Sundar, & Becker, 2010), independently of participant affect (Bendall & Thompson, 2015) or goal. Because the participants were tasked with identifying changes to the labels rather than specific content information, this study addresses early stages of information processing that are frequently neglected in the research literature relative to studies focusing on late-stage processing (i.e., comprehension of information) participants evaluating the overall stimulus for a change to the image, are invoking a bottom-up attentional process (i.e., one that is independent of the user's goals). (See chapter 1 for a more in-depth explanation of early and late stages of information processing).

Since there is no information search or processing goal associated with finding the change, this experiment separates the effect of label format from the content of the label and objectively evaluates how FOP formatting and highlighting affect the allocation of attention to OTC medication labels based on different design strategies. This methodology has been previously applied by in studies of nutritional labeling (Becker et al., 2015, 2016; Bix et al., 2015) prescription medication labeling (DeHenau, Becker, Bello, Liu, & Bix, 2016), and medical device labeling (Seo, 2014) with time to notice the changing element serving as a proxy for the locus of attentional deployment (Bix et al., 2010).

To compare the degree of noticeability of information on OTC labeling, two factors (highlighting and front of packaging warning) at two levels (highlighting present and absent and FOP design present and absent) were crossed, for a total of four treatments (a standard label, a standard label enhanced with highlighting, an enhanced label with an FOP, and an enhanced label both with an

FOP and highlighting). Documentation of the process which defined "critical information" for the safe and effective use of OTCs (i.e. highlighted and/or moved to the FOP) are included in Appendix B. For examples of the stimuli used in this experiment, please see Figure 3.1 and Appendix A.

Two types of critical information were tested in trials defined as critical: trials which contained changes to information conveying the active ingredient (AI), and changes involving drug-diagnosis or drug-drug interaction warning (DD) information. Because standard label treatments contain the active ingredient in the PDP, AI information always appeared on both the PDP and the DFL. DD information, on the other hand, only appeared on the PDP in the FOP treatment; because in standard practice, warning information does not appear anyplace other than the DFL. Critical trials were defined as trials with changes to AI or DD information that was selected for highlighting or inclusion in the front of package. The same information was considered critical in all four treatments, whether or not highlighting or the FOP were present or absent. Refer to Figure 3.3 for a diagram of the experimental structure related to critical trials. Non-critical trials served to distract participants from the objective of the study and were defined as any change that was not a critical change. Refer to Figure 3.4 for a diagram of the non-critical trial structure. For highlighted label treatments, the change was the highlighting (which is rectangular in shape behind the text) of a single warning or active ingredient appearing or disappearing. For treatments without highlighting, the changes were text of a single warning or active ingredient itself appearing and disappearing.

In order to counterbalance the effects of highlighting and the FOP on the information that was changing, and the location of the changes, the change detection task was divided into 4 experimental blocks, each comprised of 32 trials. While 128 trials allowed for a completed, counterbalanced experiment, each participant completed 64 trials, or two experimental blocks; as such, it took two participants (each who viewed two groups of trials- described below) to complete the entire block of all critical trials across all brands of mock products. As such, two versions of the

experiment were developed in order to completely counterbalance the independent variables of (highlighting (present and absent) and FOP condition (present and absent) across the four brands of drug we created and reduce fatigue effects in participants. Each of these programs alternated between participants (participants with odd participant numbers completed version A, and participants with even participant numbers completed version B) with randomized block order and trial order within each block of trials for each person. Each block of trials included 14 critical trials and 18 filler trials. There were four mock-brands, each of which featured a different active ingredient. Counterbalancing was conducted so that every participant saw all four label treatments (see figure 3.3) two times over the course of the experiment with two different mock brands, with the other version of the experiment containing the other two mock brands. Counterbalancing was conducted in this manner as it was assumed that the mock branding would not have a significantly different effect on accuracy or reaction time.

Only the critical trials were used to compare the effects on the labeling strategies on accuracy and time to correctly identify changes; however, average accuracy on noncritical trials was included in the analysis as a covariate to control for individual differences in accuracy across participants.

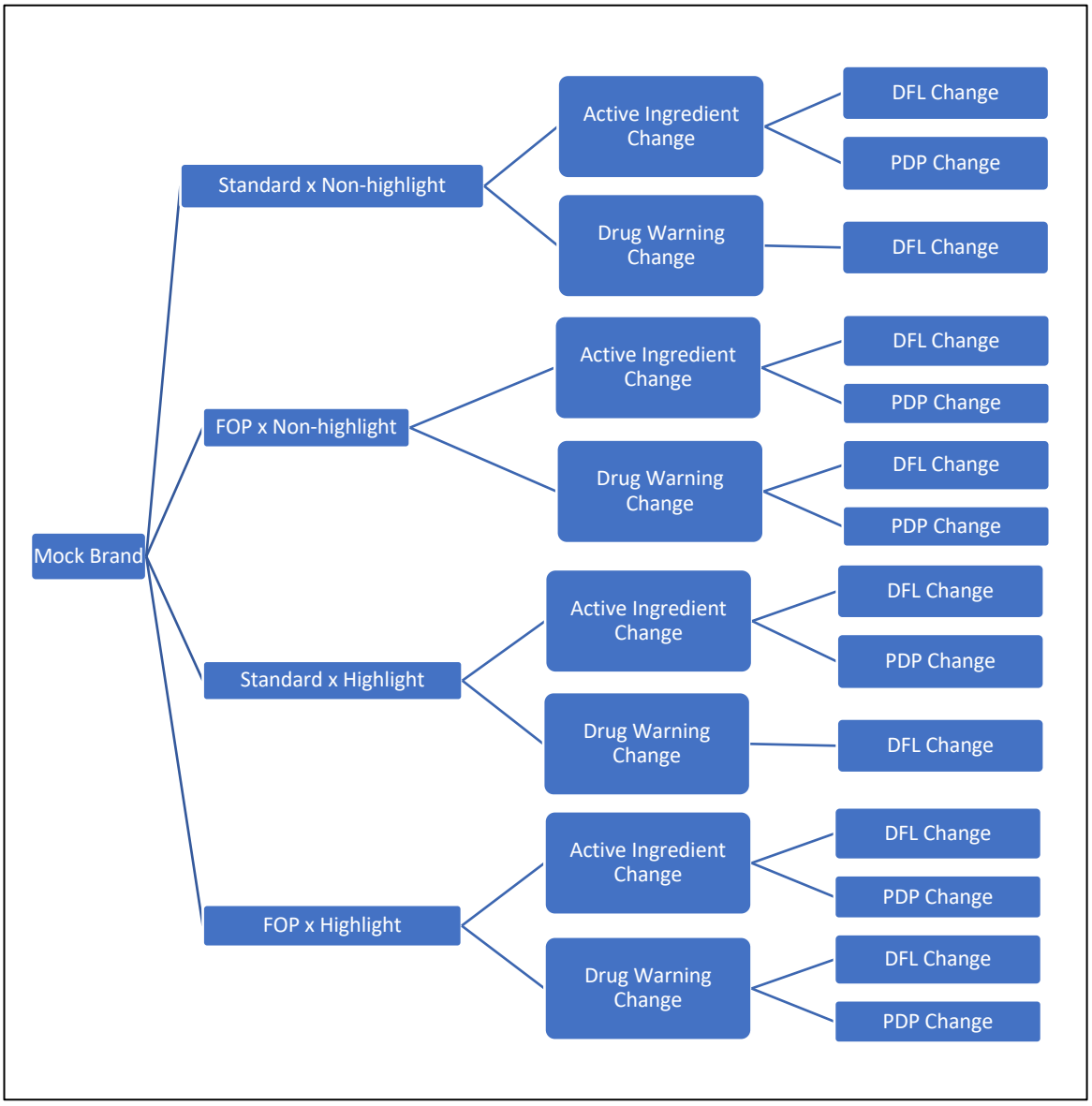


Figure 3.3 Diagram of Change Detection Critical Trial Structure

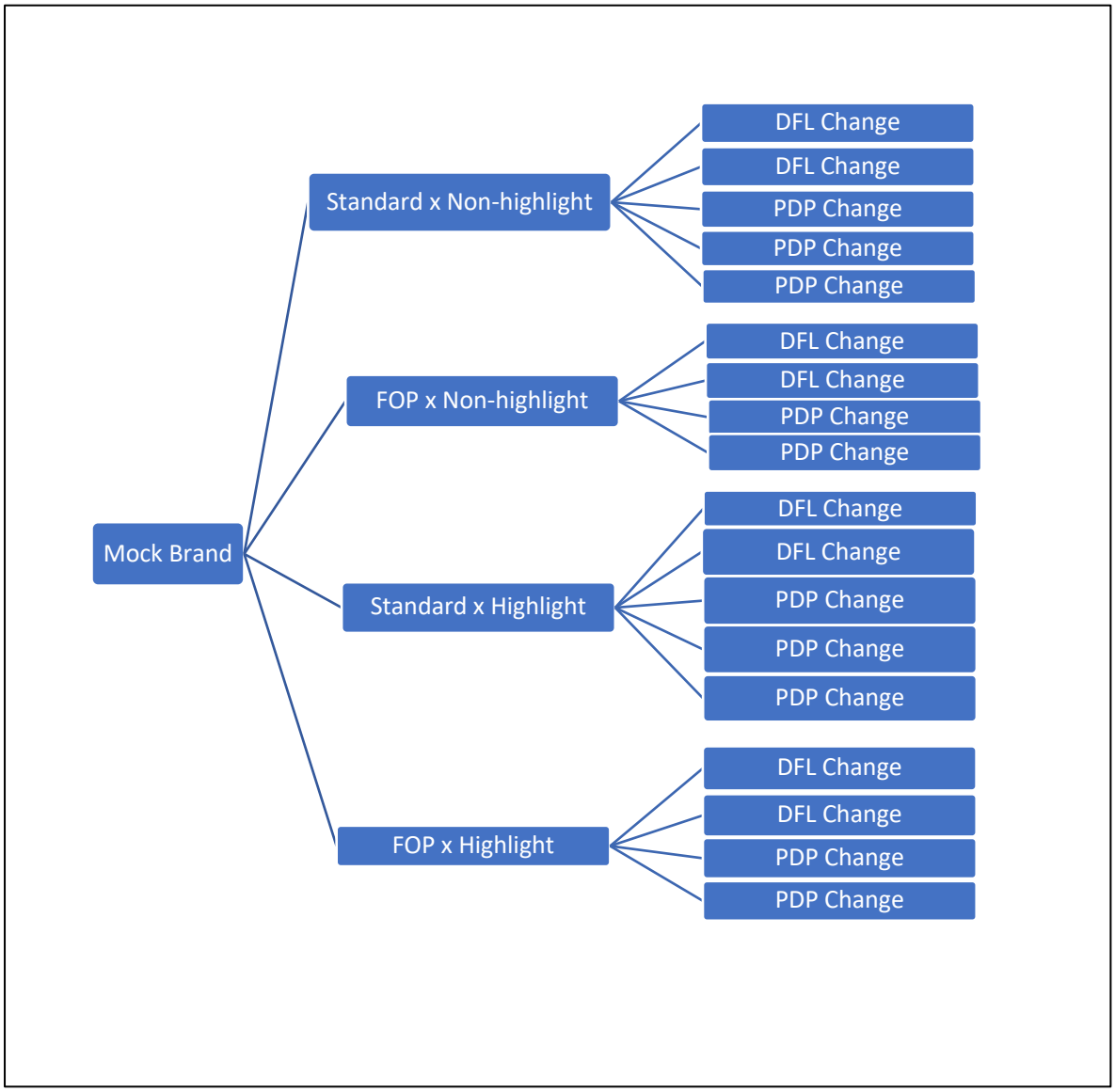


Figure 3.4 Diagram of Change Detection Non-Critical Trial Structure

Materials

In this experiment, we created four, single active-ingredient, mock brands, each of which contained one of four active ingredients for OTCs sold in the US. They were: acetaminophen, ibuprofen, omeprazole, and phenylephrine. Experimental stimuli were designed and developed in Adobe Illustrator (Adobe Systems version 7, Incorporated San Jose, CA). Included in Appendix A are the critical trials stimuli used in this experiment, along with a base image file for each treatment. Appendix B elaborates on the procedure used to determine what information was chosen for highlighting or placement in the FOP.

The experiment was programmed and run using E-Prime version 3 (Psychology Software Tools, Sharpsburg, PA).The program was run on two styles of laptops: the Dell Latitude 5490 BTX, with an 8th Gen Intel Core i5-8350U (Quad Core, 6M Cache, 1.7GHz, 15W, vPro), running Windows 10 Professional at 2400MHz with 8 GB of RAM, and the Dell Latitude 5480, XCTO, and also with an 8th Gen Intel Core, 2X8GB of RAM, running Windows 10 Professional. Both models displayed the experiment at a resolution of 1920x1080 with 14" screens.

Recruitment and Data Collection Procedures

An effect size of d=0.46 for differences between highlighted and non-highlighted labels was used to estimate the power for this study. Power calculations were based on previous change detection work which utilized a community sample of an at-risk populations (Becker et al., 2016). The power calculations suggested a sample of 48 (recruiting 60 participants before attrition) would allow provided confidence > 0.85 at the stated effect size. Proposed recruitment targets resulted in a target population of 60 participants needed in order to detect anticipated differences and allow for 20% attrition.

Before beginning the 64 research trials, participants were provided 4 practice trials intended to acquaint them with the needed keystrokes as well as provided an opportunity to pose targeted questions regarding the experiment to the research team. For each trial, there was an equally likely

chance for the change to occur on the PDP or the DFL (i.e. each of the change locations occurred in the same number of trials- 32 for each participant for each location). Participants were given 18 seconds to identify the change by pressing the space bar to stop the reaction timer. Those that did not find the change prior to time out were coded as "timeout". For changes detected within the 18 second limit, participants were asked to click on the location of the change. Change locations were defined rectangularly with X,Y pixel coordinates sampled from the image files. There was a range of 75 pixels in each direction around the change in which a click would record as a "hit", every other click location would be counted as a "miss". If the change was detected prior to time out and recorded as a "hit," the reaction time was recorded and included as a variable for analysis.

Prior to participation in the full study, participants received a test of memory and concentration that was also to screen participants unable to provide informed consent (Short Blessed Test, (Katzman et al., 1983)). After an informed written consent, and passing the cognitive screening, participants completed a survey that included: demographics (age, gender, race and ethnicity, native language, annual income, and educational attainment), a health literacy screening (Rapid Estimate of Adult Literacy in Medicine, Revised (REALM-R)(Bass, Wilson, & Griffith, 2003))(recorded as their score on the REALM-R, with participants scoring 9 or greater being dismissed due to an inability to provide informed consent), near point visual acuity (Sloan Pocket Size Near Vision Card with Continuous Text by Precision Vision in Woodstock IL)(recorded as the NPVA of the smallest line the participant was able to read when holding the card approximately 18 inches away from their face), and ability to see color (recorded as a binary yes-no variable, a no was recorded if participants were unable to distinguish the number of >2 plates) (Pseudo-Isochromatic Plates by Richmond Products, Southeast Albuquerque NM). See the data sheet in appendix E.

To characterize OTC usage, we also collected self-reported familiarity with a series of active ingredients commonly found in OTCs in the US and perceived appropriateness of common over the

counter active ingredients for the participant's use. See the data sheet in appendix E for the complete list of OTC medications, and exact wording of the familiarity and appropriateness measures. Methods were approved by the MSU Psychology and Social Science Internal Review Board as a portion of the procedures approved as part of IRB STUDY00000832. Recruitment materials, consent forms, and the data sheet are included in as supplemental files.

The population that was sampled for this study, and each of the subsequent studies reported herein, were older adult consumers (age 65+) had used OTC medications within the previous 12 months, were legally sighted, purchased and managed their own medications, willing and able to travel to the testing locations, and capable of rendering informed consent (as indicated by the Short Blessed screening). Study participants were screened from eligibility if they had a history of epilepsy or seizures because the change detection methodology results in a flashing stimulus or if their Short Blessed score was greater than 8. The participants were recruited from multiple locations in the state of Michigan, including: the greater Lansing area, Wayne County (Detroit), Kent County(Grand Rapids), and Genesee County (Flint). Recruitment was supported by MSU Extension and Wayne County's Area Agency on Aging programs targeted at seniors.

Statistical Analysis

Data was analyzed using two models, a multi-level Restricted Maximum Likelihood model that assessed the continuous variable, time to correctly identify a change, and a Logistic Multilevel Model that assessed the binary variable, correctly identified prior to time out (y/n). For both types of analysis, two models were analyzed to validate the approach. First, the model was run including only primary variables of interest (main effects of highlighting, label format (FOP-yes/no), and information that changed (AI or DD, the location of the change (PDP or DFL), and interactions between those variables) as predictors. Second, the models were run again using the primary study variables included in the first run, as well as the nuisance variables of participants' age and average reaction time for

noncritical trials, this second model with age and non-critical reaction time is what is reported in this dissertation as age is fundamental to the research objectives. Average reaction time of noncritical trials was included to account for individual differences in each participant's ability to find changes as well as the inherent variability in reaction times for different people.

While the treatment effects were tested for inclusion as random effects in the model, the models treating the treatments as random effects would not run, suggesting that any variation in treatment effects between participants was noise. This insufficient random variation necessitated the inclusion of the treatment effects as fixed effects; subject was the only variable included as a random effect. For the Multilevel Restricted Maximum Likelihood model used for reaction time data, trials were treated as repeated observations and compound symmetry was used as the model of the residuals. This imposes the required equal variance assumption. For the Logistic Multilevel Model used to analyze accuracy data, trials were treated as repeated observations, and compound symmetry was used as the model of the residuals.

Reaction time data was log transformed and only hits, i.e. trials that were correctly identified prior to time out, were included in the reported analysis. Because of differences inherent in labels (the standard labels did not include DD information in both the PDP and DFL locations), analysis of the AI and DD information were run separately. The confound imposed by honoring the standard/realistic label which does not have warning information present on the PDP, results in a different number of critical trials based on label type and information type (DD or AI)(see figure 3.1). Consider, for instance, the active ingredient information. Each of the four label treatments (HL/FOP; HL/No FOP; No HL/FOP; No HL/No FOP) occurs an equal number of times across locations of change (DFL versus the PDP) because the information appears in both locations. However, because the drug warning information does not appear on the PDP in current commercial conditions (the standard treatments) there are inequal numbers of cells to analyze between the two types of

information (AI and DD) to which changes took place. This difference in the number of cells between AI and DD trials was addressed by completing two separate analyses: one involving the changes to both AI and DD trials that allowed for inclusion of a location of change interaction term, and one that separated AI and DD trials to assess the main effects associated with the label types but losing an interaction term on the location of the change for the DD trials only. As the AI information always appeared on both the PDP and the DFL, the AI only analysis included the interaction term with location and the treatments. Reaction Time and Accuracy analysis are presented for both the comprehensive analysis including both AI and DD trials and the separate AI and DD analyses in this results sections.

Results

In the spring and summer of 2019, 60 participants were recruited for this study in 3 locations across the lower peninsula of Michigan; after screening for inability to provide informed consent[5] and removing incomplete data, 57 participants were included in the final analysis. Participant recruitment and data collection in Wayne County provided 6 participants (7 recruited, 1 dropped due to inability to provide informed consent), 27 participants were included from the Kent location (28 recruited, 1 withdrew due to technical issues with the computer program), and 24 from Ingham county recruitment efforts (25 recruited, 1 withdrew due to technical issues with the computer program) participants. The sample had a modal REALM-R score of 8 (SD=1.19, range 1-8), with three participants with a score less than 6, indicating a risk for low health literacy. The mean age of participants was 71.4 years old (SD=6.93), and the sample was 30% male (n=16) and 70% female (n=41). The sample was 76.7% (n=46) white, and 13.3% (n=8) African American. One participant (0.02%) reported being Hispanic, while 53 participants did not report being Hispanic. Three

[5] A Short Blessed Test score of 9 or more was the threshold for being unable to provide informed consent. The Short Blessed Test is a short assessment of memory and concentration and includes a measure of being orientated to time. It is used as a screening measure for cognitive decline or dementia.

participants did not report their race or ethnicity. The sample included for analysis had a mean Short Blessed Test score of 1.09 (SD= 1.79), a range of 0-8, and a median of 0. See table 3.1 for presentation of the descriptive statistics.

Table 3.1 Sample Description	
Characteristic	**N (%) or Mean (SD)**
Gender	
Men	16 (30.0%)
Women	41 (70.0%)
Race	
White	46 (76.7%)
Black or African American	8 (13.3%)
Did not report	3 (5.0%)
Ethnicity	
Hispanic	1 (0.02%)
Non-hispanic	53 (93.0%)
Did not report	3 (5.0%)
Age	71.4 (6.93, Range 65-100)
Short Blessed Test Score (0= no impairment, >8 impairment on par with dementia)	Mode= 0 (Range 0-8)
REALM R (Scores <6 are at risk for poor health literacy)	Mode= 8 (Range 1-8)
Near Point Visual Acuity	Mode= 20/32 (Range 20/20-20/50)
Ability to see Color	
Yes	54 (94.7%)
No	3 (5.3%)
TOTAL included in analysis	57 (100.0%)

While all of the described variables characterizing participants were tested for inclusion in the final model as covariates using correlation of the variable with the response variables of overall accuracy and reaction time, only age was significantly correlated with either of the response variables (accuracy and reaction time), and, thus, kept in the final model. All other potential covariates had r

values less than 0.17, and p values greater than 0.18. Mean differences of these variables were examined using independent group t-tests to investigate differences in accuracy or reaction time between sex (male versus female), education (dichotomized into two groups: some college or more versus High School or less), and race (white versus non-white) and the outcome variables. No evidence of a significant effect was detected in the outcome variables of accuracy or on the reaction time when the aforementioned factors were assessed.

Age (as a continuous variable) and accuracy were correlated at $r = -0.513$, $p < 0.001$, with age indicted to have a significant effect ($p=0.039$) on accuracy for all trial types, whereby older participants were less accurate. Thus, to be consistent and include the same independent variables for both dependent variables, age was included in each analysis of reaction time and accuracy.

Additionally, to control for individual variation in performance in a change detection task between participants, a co-variate of performance in non-critical trials was included in each analysis. For the accuracy analysis for trials with a dependent variable of accuracy in **critical** trials, accuracy in **non-critical** trials was included as a co-variate. For the reaction time analysis with a dependent variable of reaction time in critical trials, reaction time in non-critical trials was included as a covariate.

There are three sets of analyses with results from the Change Detection Task presented herein. The first investigates the effects of highlighting, FOP labeling, and location of change for the trials with changes involving only the Active Ingredient (AI). The second investigates the effects of highlighting, FOP labeling, and location of change for the changes in information occurring in Drug-Drug Warning or Drug Diagnosis Warning (DD) changes only, with an unequal number of cells in the analysis due to the lack of a DD change on the PDP in the treatments without an FOP. The final set of analyses included **both** AI and DD change types, and looked at the overall effect of highlighting, FOP labeling, and location of change across both types of information.

Active Ingredient Results

The AI is present on both locations (the PDP and the DFL) in all label treatments. This balance enabled straightforward analysis of the main effects of highlighting, the addition of a FOP, the location of the change, and all possible interactions of those three main effects. Accuracy results are presented first, as the accuracy is determined by recording the percentage of responses which correctly identified the change prior to timing out at 18 seconds. The continuous variable, reaction time, is a subset of accuracy, as it is a record of how quickly a participant was able to correctly locate a change prior to timing out, and, thus, only includes trials with successfully detected changes, not misses or timeouts.

Accuracy in AI trials is the response variable in this analysis, with predictor variables of: FOP (present or absent), highlighting (present or absent), the interaction between highlighting and FOP, location (PDP or DFL), the interaction between highlighting and location, the interaction between FOP and location, the three-way interaction between highlighting, FOP, and location, age (continuous), and accuracy in non-critical trials. Table 3.2 provides the results from this analysis.

Table 3.2 Fixed Effects for Active Ingredient Trials Only with Accuracy as the Dependent Variable

Source	Df 1	Df 2	F	Sig.
Corrected Model*	9	894	9.639	0.000
FOP effect	1	894	0.344	0.557
Highlight effect	**1**	**894**	**8.568**	**0.004**
FOP effect x Highlight effect	1	894	1.324	0.250
Location effect	**1**	**894**	**47.045**	**0.000**
FOP effect x Location effect	1	894	0.422	0.516
HL effect x Location effect	**1**	**894**	**12.320**	**0.000**
FOP effect x Highlight effect x Location effect	1	894	0.050	0.823
Age	1	894	3.619	0.057
Accuracy Noncritical Trials	**1**	**894**	**19.805**	**0.000**

* "Corrected Model" results are included in within-subject designs. "The F-test for the corrected model is a test of whether the model as a whole accounts for any variance in the dependent variable." (IBM Support, 2020) It is not an independent variable in the model.

Figure 3.5 presents the results related to the AI-only analysis using accuracy as the dependent variable. There were main effects of Highlight ($p=0.004$), Location of Change ($p<0.001$), Accuracy in non-critical trials ($p<0.001$), and a significant interaction effect of highlighting by location was apparent ($p<0.001$). The presence of highlighting enhanced the ability of participants to accurately detect changes when changes to the AI information took place in the DFL location across all treatments (DFL changes: HL ME=0.607, SE=0.042, nonhighlighted ME=0.377, SE=0.041). In looking at AI changes that took place in the PDP, the benefits of highlighting are less clear (ME=0.719, 0.038 for highlighted, ME=0.736, SW=0.037 for unhighlighted). The efficacy of highlighting the information appears to be influence by the location, with highlighting yielding more accurate ability to detect the change prior to timing out when said change occurs in the DFL but this benefit does not hold for changes to the same information within the PDP. One possible explanation for this interaction term is that the AI information which appears on the PDP is larger than it is in its appearance within the DFL (see Figure. 3.1). It could be conjectured, that, as a result, it is already performing well with little opportunity to improve accuracy within the PDP location as compared to the DFL. Table 3.3 includes the full results of the model for this analysis and figure 3.5 for a graphical representation of AI trial accuracy results.

Accuracy in non-critical trials was also significant in this model. The coefficient of the predictor variable accuracy in non-critical trials was 0.124. The positive sign on this coefficient indicates that participants with higher accuracy in non-critical trials performed significantly better in the critical AI trials. Again, this variable was included to account for individual variation in change detection task skill, and it is unsurprising that participants who performed more accurately in one type of trial also preformed more accurately in another.

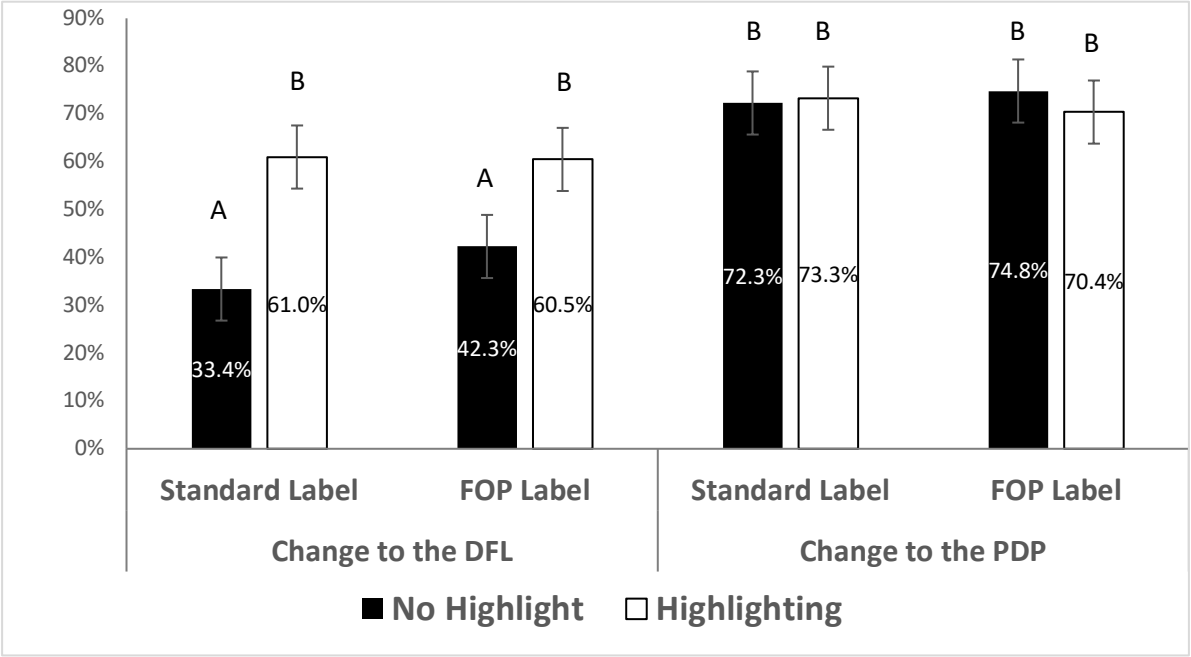


Figure 3.5 Accuracy for AI trials only. Treatments with different letters above them are significantly different from each other at the alpha= 0.05 level

In the next analysis, reaction time in correctly identified AI trials is the response variable in this analysis, with predictor variables of: FOP (present or absent), highlighting (present or absent), the interaction between highlighting and FOP, location (PDP or DFL), the interaction between highlighting and location, the interaction between FOP and location, the three-way interaction between highlighting, FOP, and location, age (continuous), and reaction time in non-critical trials.

Results for the reaction time analysis examining changes to the AI information revealed that highlighting (p= 0.002), presence of an FOP (p=0.002), and location of change (p=0.000) had a significant effect on the time to detect a change at α=0.05. A significant 2-way interaction between the location of the change (PDP or in the DFL), and highlighting was identified as well (p=0.036) (See Table 3.3). As the interaction term included two of the main effects, this result will be discussed within the context of the significant interaction term, and the main effect of FOP. Additionally, reaction time in non-critical trials was significant (p=0.023). Trends in the data were as expected, specifically, as

average reaction time in noncritical trials increases, so does the reaction time in critical trials. As this variable was included to account for individual variation in change detection task skill it is unsurprising that participants who responded quicker in one type of trial also responded quicker in another.

Table 3.3 Type III Tests of Fixed Effects for Active Ingredient Trials Only with Log10 Reaction Time as the Dependent Variable				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	50.824	40388.750	0.000
FOP effect	**1**	**488.667**	**10.031**	**0.002**
Highlight effect	**1**	**506.425**	**9.820**	**0.002**
FOP effect * Highlight effect	1	486.727	.002	0.962
Location effect	**1**	**510.553**	**135.799**	**0.000**
FOP effect * Location effect	1	489.134	.006	0.940
Highlight effect * Location effect	**1**	**496.478**	**4.422**	**0.036**
FOP effect * Highlight effect * Location effect	1	488.980	.985	0.322
Age	1	78.483	1.235	0.270
Reaction Time Noncritical Trials	**1**	**56.780**	**5.435**	**0.023**

The significant main effect of FOP and significant 2-way interaction between highlighting and location suggest two things; first that the FOP is effective at garnering attention, as participants were slower at finding AI changes when an FOP was present (ME= 3.766 log10ms, SE=0.021 versus FOP present vs ME= 3.702 log10ms, SE=0.021 FOP absent).

Secondly, when we examine the impacts of highlighting, unlike the previous analysis that investigated the effect of highlighting on accuracy which indicated more accuracy benefit to highlighting in the DFL than in the PDP (see Table 3.2 and Figure 3.5), a clear benefit of highlighting appears across all treatments and locations of change for the reaction time data; people are faster at detecting the changes in the highlighted conditions. That said, it is important to note that the location mediates the beneficial reaction time effect of highlighting, with highlighting being more impactful in attracting attention and reducing reaction time on the more prominent PDP (HL PDP ME=3.560 log10ms, SE= 0.024, unhighlighted PDP ME=3.667, SE=0.024) than the less prominent DFL (HL

DFL ME=3.844 log10ms, SE=0.025, unhighlighted DFL ME=3.865log10ms, SE=0.025)[6]. This is not necessarily solely a location effect because, as noted previously, the AI information is larger (and subsequently, the highlighting is larger) than that which appears in the DFL. See figure 3.6 for a graphical representation of AI trial reaction time results and table 3.2 for the full results of the model.

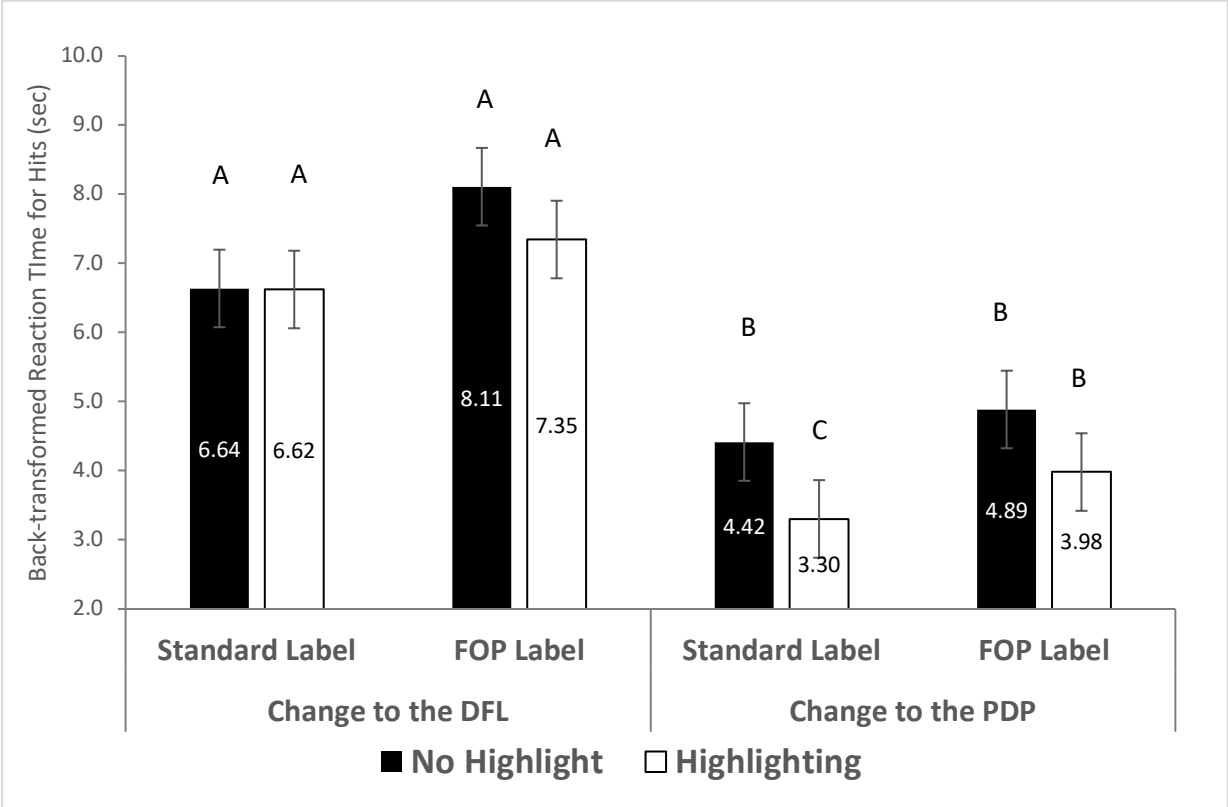


Figure 3.6 Back-transformed reaction time for detecting changes in the AI trials only, error bars represent 95% confidence intervals. Treatments with a different letters above them are significantly different from each other at the alpha= 0.05 level

Drug-Drug and Drug-Diagnosis Interaction Warning Results

While the AI trials were perfectly counterbalanced, with an equal number of changes occurring in the DFL versus the PDP (due to the presence of AI information in both locations regardless of

[6] Prominence in this instance is referring to both font size and density of the information. The AI on the PDP is a larger font size, and has less surrounding information competing for attention than the smaller font sized AI on the information dense DFL.

treatment type), the nature of the standard label designs, which were drafted from those commercially in use at the time of testing, made achieving balance impossible for the DD information. Limiting the standard treatment to the form typical in commercial practice, resulted in only 6 treatment-change location combinations when changes occurred in DD information. As a result, pairwise comparison and contrasts were used in the following analysis.

In addition to the unbalanced number of treatment-change location combinations for DD trials, there was an error in the E-Prime Programming that wasn't discovered until after data collection was complete. This impacted version two of the two programs code for trials comprised of a single active ingredient in the highlighted, FOP label condition. Participants who completed the problematic version of the program saw three trials involving the change to the DD in the highlighted condition appearing on the PDP (within the FOP), and only 1 trial with the change to this information in the DFL.

Participants who completed version one (of the two versions of the experiment that comprised a complete block) had a balance in the location of the critical trials; specifically, two trials where the highlighted DD change was located within the FOP (on the PDP) and two trials with the highlighted DD change located in the DFL. In other words, for each participant that utilized program one, the four pieces of data for the DD information in treatments that were highlighted and contained an FOP (two in the PDP – on the FOP and two in the DFL) were recorded (refer back to figure 3.3 for an illustration of the different trial types), but for the erroneous version, the four pieces of data that were recorded included one for a change in the DFL, and 3 for a change in the FOP on the PDP.

Although this error occurred for only a single active ingredient in one of the two programs run, it did result in a reduced number of observations related to changes to the DD information in highlighted treatments where an FOP was present, and as such, a slightly smaller denominator was used to calculate the accuracy percentage in the relevant treatment data (Highlighted, FOP present).

Figure 3.7 visually presents the results of the accuracy analysis for the DD information; note that the large effect of highlighting for the FOP label with a change in the DFL condition is the trial type in which the programming error occurred and thus isn't a reliable result and will not be discussed further.

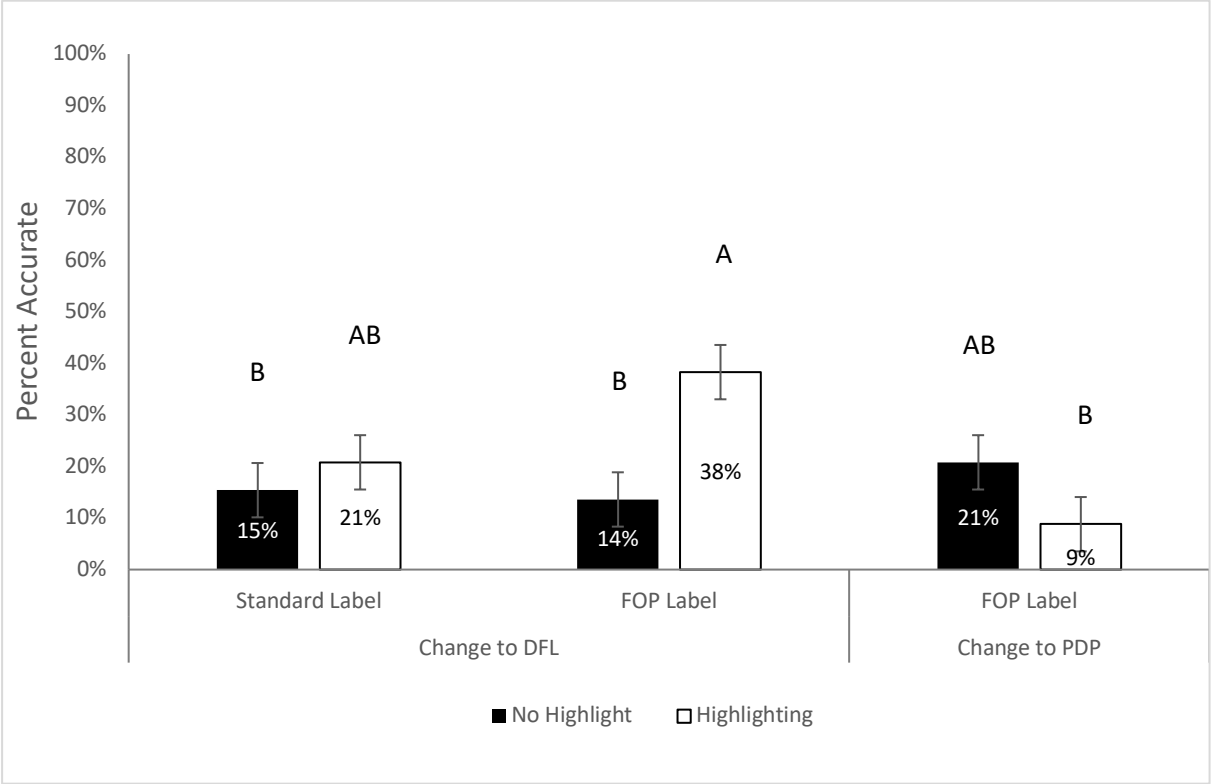


Figure 3.7 Accuracy for DD Trials Only. Treatments with different letters above them are statistically significantly different at the $\alpha = 0.05$ level. Note that there are only 6 possible combinations of treatment and change location for DD trials, as DD warnings only appeared in the DFL without the treatment of an FOP

Table 3.4 Mean Estimates for Drug Warning Trials with Accuracy as the Dependent Variable				
	Mean		95% Confidence Interval	
Label Treatment	Estimates	Std. Error	Lower Bound	Upper Bound
FOP highlight DFL	0.383 A*	0.056	0.280	0.498
FOP highlight PDP	0.088 B	0.025	0.050	0.150
FOP non-highlight DFL	0.136 B	0.033	0.033	0.214
FOP non-highlight PDP	0.208 AB	0.039	0.141	0.295
standard highlight DFL	0.208 AB	0.040	0.139	0.298
standard non-highlight DFL	0.154 B	0.035	0.097	0.236
* Means followed by different letters are significantly different from each other at the $\alpha = 0.05$ level.				

As reaction time data is a subset of accuracy data, the programming error was less impactful on the reaction time results, as even without a programming error, there are different numbers of correct responses to include for each label treatment's average response time, which is a continuous variable. While changes to unhighlighted DD information located in the PDP (FOP ME=3.908 log10ms, SE=0.036) were not indicated to result in reaction times that were significantly different from those collected for highlighted DD that occurred in the PDP (FOP ME=3.833 log10ms, SE= 0.050), and reaction times were significantly faster for the highlighted DD change in the FOP than for all remaining treatments (See Table 3.6 for all Mean Estimates/Standard Errors and pairwise comparisons). Results indicate that changes involving the DD information in the FOP (unhighlighted ME=3.908 log10ms, SE=0.036, highlighted ME=3.833 log10ms, SE= 0.050) were detected faster than the changes to a non-highlighted, standard package with a change in the DFL (ME=4.062 log10ms, SE= 0.044). The changes in the FOP were detected faster than the DFL with no significant difference between highlighted or non-highlighted conditions, meaning the FOP is beneficial, whether or not it is highlighted. When changes occurred to DD information within the DFL for all other treatments (highlighted x standard, highlighted x FOP, nonhighlighted x FOP), reaction times did not differ significantly from the standard, non-highlighted package. Figure 3.8 shows the reaction time analysis and Table 3.4 presents the estimated means and confidence intervals of the 6 different label treatment and location interactions analyzed for the DD analysis.

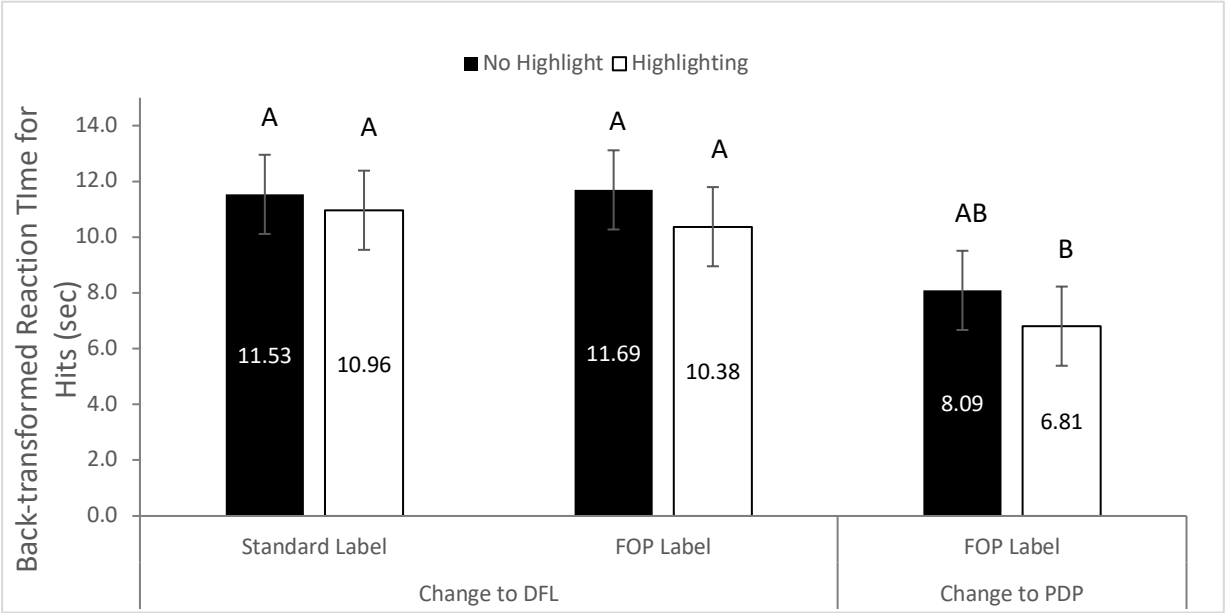


Figure 3.8 Back-transformed reaction time for DD Trials Only, error bars represent 95% confidence intervals. Treatments with different letters above them are statistically significantly different at the α =0.05 level

Table 3.5 Mean Estimates for Drug Warning Trials Only with Log10 Reaction Time as the Dependent Variable					
				95% Confidence Interval	
Label Treatment	Mean Estimates	Std. Error	df	Lower Bound	Upper Bound
FOP highlight DFL	4.016 A	0.032	125.611	3.952	4.08
FOP highlight PDP	3.833B	0.05	126.901	3.734	3.932
FOP non-highlight DFL	4.068 A	0.046	122.811	3.978	4.158
FOP non-highlight PDP	3.908 AB	0.036	118.887	3.836	3.979
standard highlight DFL	4.040A	0.037	122.762	3.967	4.113
standard non-highlight DFL	4.062A	0.044	125.646	3.975	4.149
* Means followed by different letters are significantly different from each other at the α =0.05 level.					

Comprehensive Analysis Results

The analysis which included examined changes that occurred to both pieces of critical information (AI and DD), termed the comprehensive analysis, was done for both dependent variables, accuracy and time to correctly detect a change. In this analysis, accuracy is the response variable with

predictor variables of: information type (AI or DD), FOP (present or absent), highlighting (present or absent), the interaction between highlighting and FOP, the interaction between highlighting and information type, the interaction between FOP and information type, the three way interaction between highlighting, FOP, and information type, age (continuous), and accuracy in non-critical trials. Location is not included as it in unbalanced in the DD portion of the critical trials.

Results suggest that accuracy to successful detection changes is significantly influenced by the type of critical information that is changing ($p<0.001$). In other words, respondent's ability to successfully detect changes prior to timing out was significantly influenced by whether the change was occurring in AI or DD information, with AI trials being more accurately detected than DD trials (ME= 0.609, SE= 0.024 vs ME= 0.171, SE=0.020 respectively). Figure 3.9 graphically presents the comprehensive accuracy results, and table 3.7 contains the full results of the statistical model.

Table 3.6 Fixed Effects for All Critical Trials with Accuracy on Critical Trials as the Dependent Variable

Source	Df 1	Df 2	F	Sig.
Corrected Model	9	1,579	30.081	0.000
Information Type (AI or DD)	**1**	**1,579**	**237.239**	**0.000**
FOP Effect	1	1,579	0.121	0.728
Highlight effect	**1**	**1,579**	**7.750**	**0.005**
FOP Effect * Highlight effect	1	1,579	1.115	0.291
FOP Effect * Information Type	1	1,579	0.038	0.845
Highlight Effect * Information Type	1	1,579	0.523	0.470
FOP Effect * Highlight Effect * Information Type	1	1,579	0.011	0.915
Age	**1**	**1,579**	**4.273**	**0.039**
Accuracy Noncritical Trials	**1**	**1,579**	**24.093**	**0.000**

This analysis indicates that the presence of highlighting increases accuracy across information type (unhighlighted ME= 0.322, SE=0.027, HL ME= 0.402, SE= 0.029). Along with increased accuracy attributable to highlighting, the type of critical information also had an effect on accuracy of detection, with AI changes (ME= 0.609, SE= 0.024) being more accurately detected than DD changes (ME= 0.171, SE=0.020), see figure 3.9 for a graphical representation of these results and table 3.6 for

the full results of the statistical model. Additionally, age and accuracy in non-critical trials were significant. The coefficient for age is -0.039, meaning as age increases, accuracy decreases. The coefficient for accuracy in noncritical trials is 0.101, meaning as accuracy in noncritical trials increases, accuracy in critical trials increases as well. As accuracy in noncritical trials is included to account for individual variation in skill in a change detection task, it is unsurprising that participants who are more accurate in non-critical trials are more accurate in critical trials.

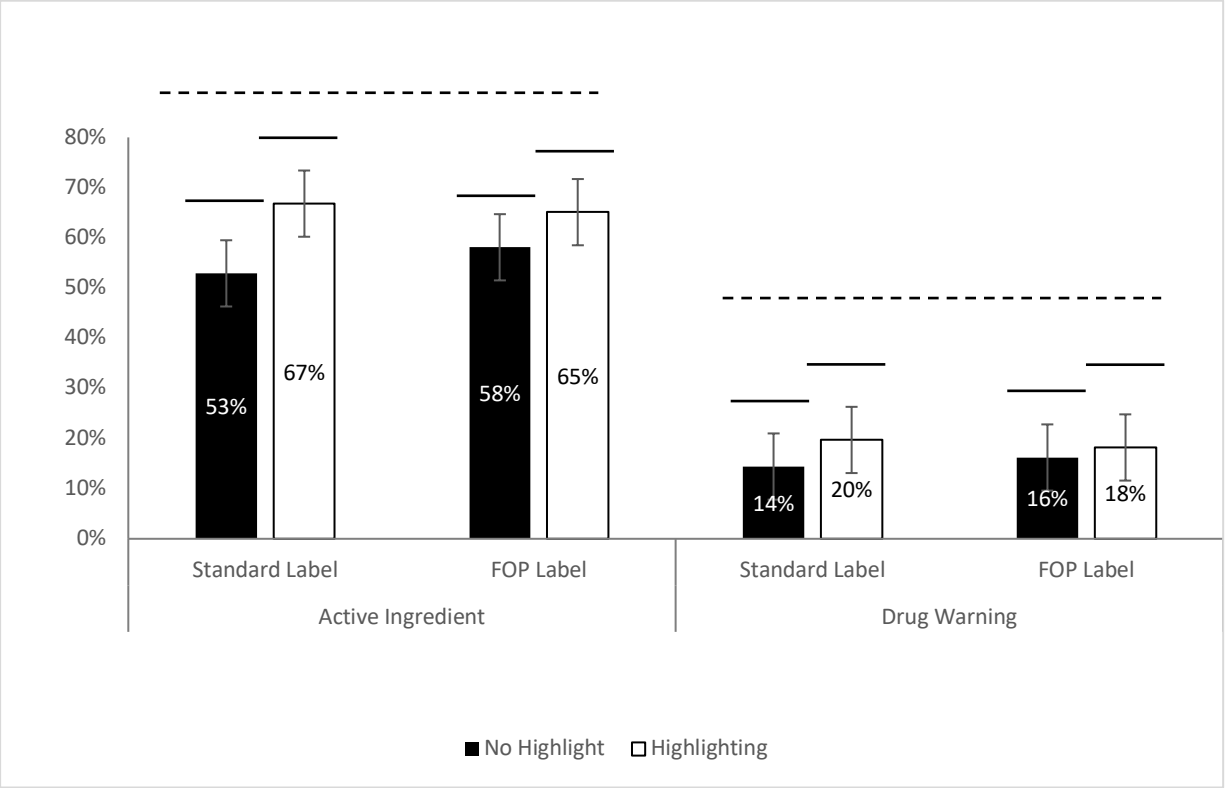


Figure 3.9 Change Content Accuracy, all trials, error bars represent 95% confidence intervals. Solid line represents the main effect of highlighting, and the dashed line represents the main effect of information type

For the comprehensive reaction time analysis, reaction time in correctly identified critical trials is the response variable, with predictor variables of: information type (AI or DD), FOP (present or absent), highlighting (present or absent), the interaction between highlighting and FOP, the interaction between highlighting and information type, the interaction between FOP and information type, the

three way interaction between highlighting, FOP, and information type, age (continuous), and reaction time in non-critical trials. Location is not included as it in unbalanced in the DD portion of the critical trials.

The comprehensive analysis of treatment influence of reaction time reinforces that what information is changing (AI info versus DD info) has more of an effect on reaction time than the novel format of the information (highlighted vs. non-highlight or FOP vs. standard). This is not surprising due to the size confounds that exist between the AI information and DD information, with the AI changes on the PDP being more prominent than any DD change. Specifically changes to the AI were detected significantly faster (ME= 3.712 log10ms, SE= 0.015) than those to the DD across locations (ME=4.009log10ms, SE= 0.026) (table 3.7 and figure 3.10). Additionally, there is a significant interaction between FOP and Information Type, with the presence of an FOP decreasing reaction time for DD changes (FOP absent ME=4.057log10ms, SE 0.040, FOP present ME=3.962 log10ms, SE 0.029) but increasing reaction time for AI changes (FOP absent ME=3.677log10ms, SE 0.019, FOP present ME=3.746 log10ms, SE 0.019). Again, this is not surprising as the AI information does not appear in the FOP, and thus the FOP competes with the AI information for attention.

Table 3.7 Type III Tests of Fixed Effects for All Critical Trials with Log10 Reaction Time as the Dependent Variable

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	95.077	51938.619	0.000
Information Type (AI_DD)	**1**	**658.304**	**129.135**	**0.000**
FOP Effect	1	636.292	0.261	0.609
Highlight (HL) effect	1	660.678	0.89	0.346
FOP Effect * HL effect	1	638.06	0.001	0.976
FOP Effect * AI_DD	**1**	**639.121**	**9.862**	**0.002**
HL Effect * AI_DD	1	647.548	0.884	0.347
FOP Effect * HL Effect * AI_DD	1	638.417	0.001	0.976
Age	1	96.406	0.264	0.608
Reaction Time Noncritical Trials	**1**	**61.52**	**6.804**	**0.011**

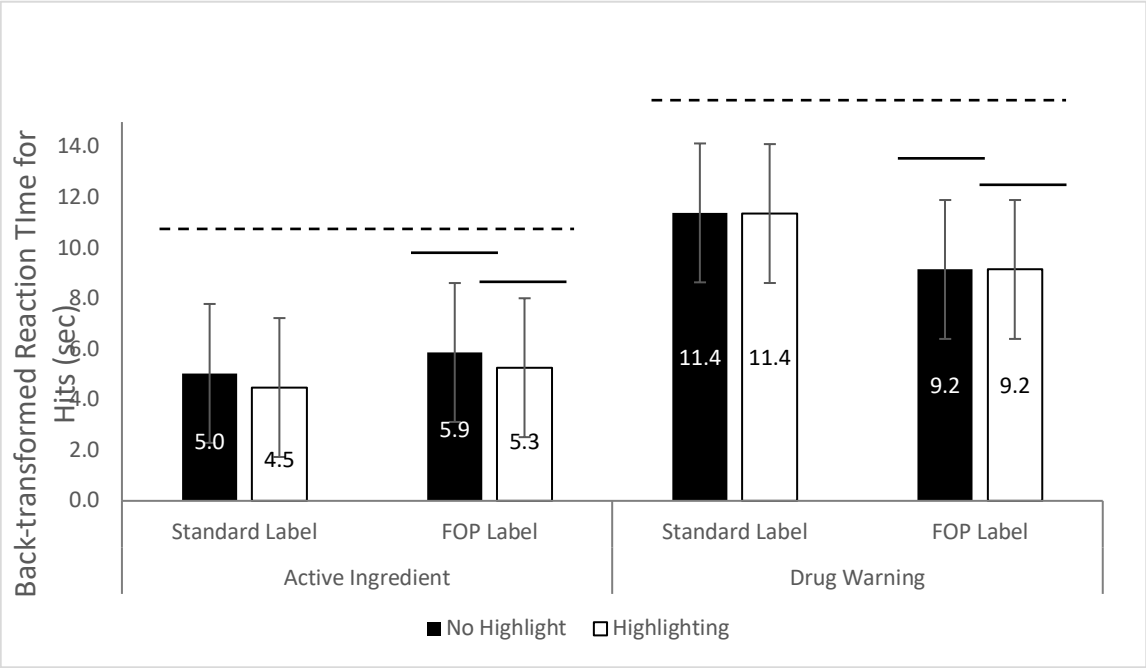


Figure 3.10 Change Content Reaction Time back-transformed from log10ms into seconds, all trials, error bars represent 95% confidence intervals back-transformed from log10ms into seconds. Solid line represents the interaction between FOP and information type, and the dashed line represents the main effect of information type

Additionally, reaction time in non-critical trials was significant (p=0.023). Unsurprisingly, the trends in the data show that as average reaction time in noncritical trials increases, so does the reaction time in critical trials. Again, reaction time in non-critical trials was included to account for individual differences in participants' ability in a change detection task.

Conclusion

Finally, to close the results section related to the change detection methodology employed, below is a presentation of the results in terms of the four scientific hypotheses presented earlier this this chapter:

Hypothesis 1: *Highlighting information determined to be critical to the safe and effective use of a product will increase the accuracy of participants noticing changes compared to non-highlighted labels.* This hypothesis was supported by the comprehensive accuracy analysis that included both AI and DD change types, and the analysis that investigated changes to the critical information, AI.

Hypothesis 2: *Changes occurring in the front of pack label will increase the accuracy of participants' noticing of changes to critical information compared to changes to the same information in the drug facts label.* Partially due to the error in the program that impacted the balancing of the highlighted, FOP trials with DD changes, this hypothesis is unsupported.

Hypothesis 3: *Highlighting information deemed to be critical to the safe and effective use of an OTC will decrease the amount of time required for participants to notice changes to the same information in non-highlighted labels.* This hypothesis was supported by the results of the AI information only analysis which found evidence of highlighting decreasing reaction time, with a caveat that the benefits of highlighting for AI information are mediated by the location of the highlighting (PDP versus DFL) and the presence or absence of an FOP. The beneficial effect of highlighting was more pronounced in the DFL than the PDP location, and reaction time was increased in the presence of an FOP.

Hypothesis 4: *Changes occurring in the front of pack label will decrease the amount of time required for participants to notice changes compared to comparable changes in the drug facts label.* This hypothesis was supported reaction time analysis for changes to the DD information where the evidence suggests that FOP presence results in decreasing reaction time compared to changes that occurred in the DFL. The mean reaction time to detect a change in the highlighted FOP was significantly different than the mean reaction times to detect changes in the DFL, whether or not the changes were highlighted.

Discussion and Implications

Overall, results suggest that our design changes (highlighting and using an FOP) to the standard OTC label garner attention and improve the likelihood that older adults notice information crucial to making informed healthcare decisions. The results of this assessment of the bottom-up processing of OTC medication labels by older adults will be discussed further in Chapter 6 in tandem with the results of the two investigations into the effects of these label formats on enhancing the efficiency for top-down processing tasks.

Most apparent is the conclusion that highlighting is helpful in attracting older adults' attention to information, especially when the information being highlighted is the active ingredient included in the DFL. This conclusion is based on the results from the AI Accuracy Analysis and supports the findings by King et al in their investigation of highlighting as a possible improvement to acetaminophen labels(King et al., 2011). Highlighting was also indicated to enhance accuracy for both AI and DD trials, while the effect of highlighting is less clear for DD trials than for AI trials. As older adults are likely undertaking habit-based decision making when selecting OTCs (Holden et al., 2018), these results support the notion that labels designed to garner attention could interrupt decisions being made on "auto-pilot" and facilitate more deliberative processes.

The results indicate that a label optimized for older adults making OTC purchasing and use decisions might be a label in which there is an FOP including the warnings on the front, but with highlighting only on the DFL to minimize the chance that the highlighting has diminishing returns on attracting attention. Relative salience of the highlighting is important to consider when designing an optimized OTC label, as research in the area of highlighting text for studying suggests the more highlighting is present, the less any one piece of highlighted information stands out (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). Because of the diminishing returns of highlighting, careful consideration of what information is highlighted or moved to the FOP should be undertaken to maximize the attention garnering benefits and facilitate use of the information. Further work to investigate the effects of highlighting and prioritization of warnings on the FOP when participants are tasked with using the critical information to respond to a task will be conducted in later chapters to better understand if highlighting and the FOP are beneficial in realistic usage scenarios.

The combined results from AI trials only and DD trials indicate the FOP label is efficient at attracting attention to warnings because the warnings in the FOP are more prominent than the DFL, this assertion is based on DD reaction time results. Additionally, the FOP did not have a statistically

significant inhibiting effect decreasing the noticeability of AI information, meaning its presence on PDP the did not significantly distract from other important health information. However, further investigation into whether the warnings in the FOP should highlighted in addition to appearing on the PDP or if only the warnings in the DFL should be highlighted is needed, in part due to the program error. These four label treatments will be further analyzed to assess their effects on use of product information in the next two chapters.

Chapter 4 An Absolute Judgment Task

Overview

The objective of this study was to investigate how label formatting techniques (FOP and highlighting) attract attention to critical information (DD and AI) when accessing that information is the person's explicit goal (top-down processing). To address this objective, this binary, absolute judgment test was conducted which consisted of a series of trials containing one of four labeling formats (2 treatments at 2 levels)(unhighlighted, no FOP, highlighted no FOP, unhighlighted with FOP, highlighted with FOP) with a question that could be answered in binary fashion related to the about the medication that could be answered using the label. The question served as the explicit goal motivating the participants to search for information about the medication.

Herein, we continued to objectively investigate two label treatments at two levels (four treatments in total) however this study differs from the Change Detection (presented in Chapter 3) as the task driven nature of this experiment engages the top-down attentional processing (Kinchla & Wolfe, 1979) as opposed to the bottom up processing mechanisms engage during the change detection testing (discussed in Chapter 3). The same label treatments (unhighlighted, no FOP, highlighted no FOP, unhighlighted with FOP, highlighted with FOP) assessed in the Change Detection study are assessed in this work, see figure 3.1 for an example of the treatments. The scientific hypotheses (Gotelli & Ellison, 2004) are:

> ***Hypothesis 1:*** Highlighting information determined to be critical to the safe and effective use of OTCs will increase the accuracy of participants and older responses to yes/no questions about the about the drug compared to non-highlighted labels.
>
> ***Hypothesis 2:*** The front of pack label will increase the accuracy of participants responding to yes/no questions about the drug compared to labels without the front of pack label.

Hypothesis 3: Highlighting will decrease the amount of time required for participants to accurately respond to yes/no questions about the drug compared to non-highlighted labels.

Hypothesis 4: The front of pack label will decrease the amount of time required for participants to accurately respond to yes/no questions about the drug compared to labels without the front of pack label.

In addition to the primary objectives detailed in the grant application, we also postulated the prior use and familiarity with a specific active ingredient had the potential to influence response accuracy and time to correct responses. This secondary line of inquiry examines the degree to which older adults are more or less familiar with 10 different OTC brand names as opposed to the 10 corresponding OTC active ingredients, and what effect that familiarity has on performance in the labeling task. This interest in the effect of familiarity was born out of researcher's experience with collecting the information about participants background and baseline familiarity with active ingredients during the survey portion of the study. Additionally, examining participants' familiarity with branding versus active ingredients will provided evidence about whether the mock-branding was sufficient in masking participants' prior knowledge about OTC medications from the labeling task.

Methods and Materials

The materials and methods section of this chapter first discusses the materials and methods used to generate the experimental stimuli, and secondly, describes the design of the experiment. Next, it discusses the recruitment and data collection procedures including the screening criteria , and finally, describes the statistical analysis strategy and methodology used to analyze the data. Methods were approved by the MSU Psychology and Social Science Internal Review Board in Summer 2018 as STUDY00000832.

Materials

Experimental stimuli were designed and developed in Adobe Illustrator (Adobe Systems version 7, Incorporated San Jose, CA). The experiment was programmed and run using E-Prime version 3 (Psychology Software Tools, Sharpsburg, PA). The program was run on two styles of laptops: the Dell Latitude 5490 BTX, with an 8th Gen Intel Core i5-8350U (Quad Core, 6M Cache, 1.7GHz, 15W, vPro), running Windows 10 Professional at 2400MHz with 8 GB of RAM, and the Dell Latitude 5480, XCTO, also with an 8th Gen Intel Core, 2X8GB of RAM, running Windows 10 Professional. Both models displayed a resolution of 1920x1080 with 14" screens.

Experimental Design

We conducted an absolute judgement task featuring yes/no questions which required information from the OTC medication labels to objectively investigate how label treatments of highlighting (present absent) and FOP (present absent) impacted participant accuracy and reaction time related to question response (a task driven objective). Questions included in the study are included in Appendix E. As with the Change Detection study presented in the previous chapter, two types of critical safety information trials were examined: those that involved the active ingredient (AI) information, and those that required the drug-diagnosis or drug-drug interaction warning (DD) information. Additionally, distraction trials served to investigate the potential distracting effect of the presence of an FOP or highlighting when participants were searching for other (i.e. non AI or DD) information in the DFL. The trials investigating the distraction effect featured questions about the use of the medication, a nonhighlighted piece of information in the first panel of the DFL.

The absolute judgement task was designed so that each participant completed a total of 144 trials (See Figure 4.1 for an example trial); 128 of which were critical trials, defined as those which involved the participant's ability to answer questions that required information critical to the safe and effective us, namely, AI and DD. See Table 4.1 for a list of mock products created with their active

ingredients. The remaining 16 trials were considered distraction effect trials. Within each block of 18 trials, 16 trials were critical trials examining the effectiveness, and 2 trials examined if the presence of an FOP served to "distract" participants from the more comprehensive information in the DFL. There were no breaks in participation between blocks; the blocks function was to spread out the occurrences of active ingredient and treatments to limit learning or run order effects. The order of blocks was randomized, and the order of trials within blocks was also randomized. Due to error when programming the experiment, the first 40 participants did not complete distraction effect trials, and thus the final data set only analyzes the distraction effect trials for the final 35 participants.

Trials consisted of a single label paired with a yes/no question specifically crafted for that product (see Appendix E for a full list of questions used). Responses to the question were recorded by participants pressing the "z" key to respond no and the "m" key to respond yes. When the participant pressed the key to respond, the experiment moved to the next trial. The same two treatments at two levels from the previous chapter were utilized again in this study. Refer to Figure 3.1 and Appendix C for illustration of these treatments. For the enhanced label treatments, the information needed to correctly respond to the question was in the FOP, if it was a DD question, and highlighted in both the DFL and the FOP (if present), for those trials that included highlighting Appendix B details the process used to determine what information was deemed critical.

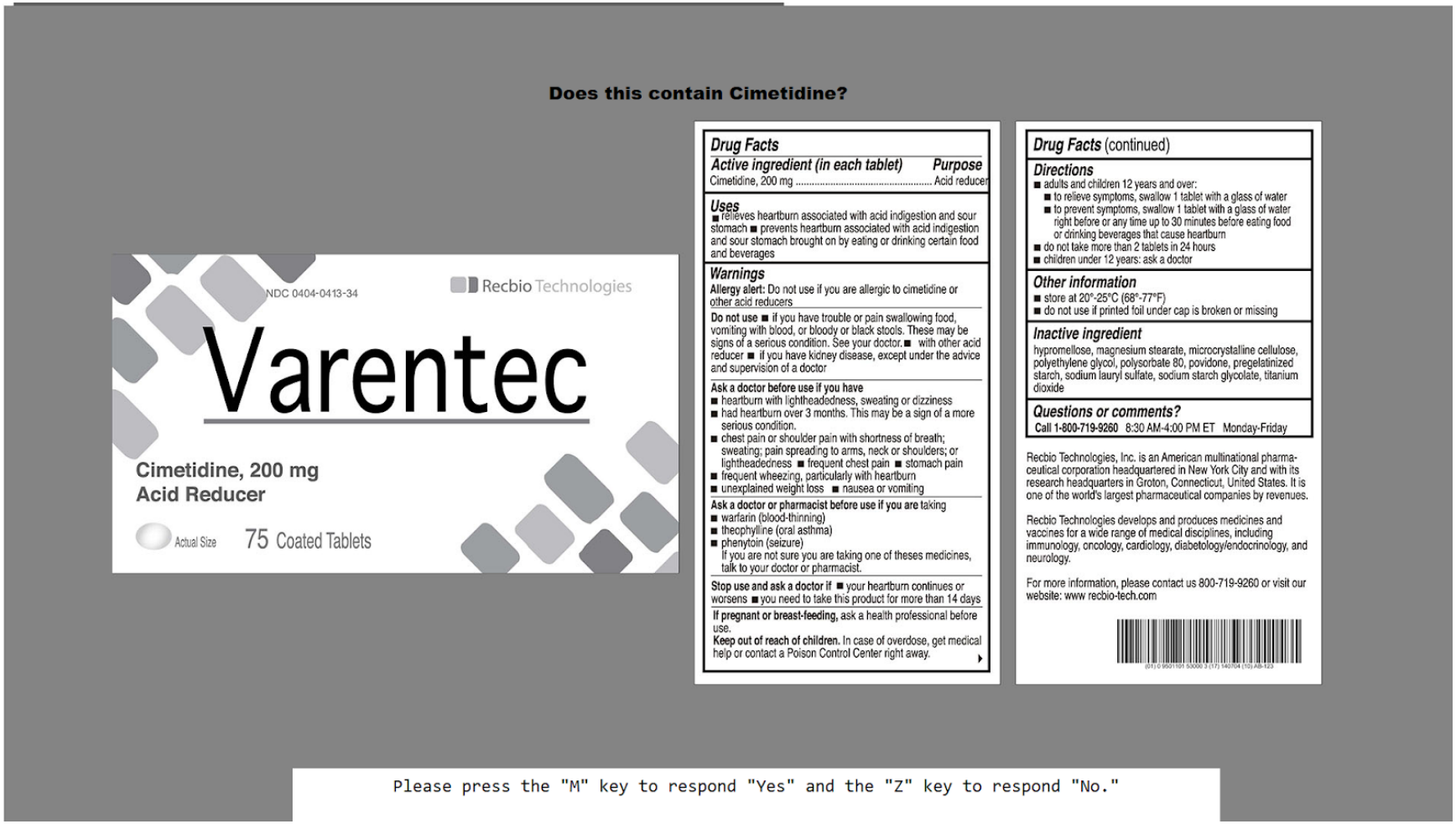


Figure 4.1 Example task participants will complete during this study. The stimuli presented is the treatment of a standard label with an active ingredient question

The distraction trials were conducted in a similar manner. Instead of questions regarding AI or DD information, the question required information found within the DFL related to the subheading "uses." This information was unhighlighted and appeared only in the DFL. The distraction effect trials examined if the presence of the FOP and/or highlighting acts as an inadvertent hindrance, or a distraction, to finding the desired information. Thus, all of the distraction effect trials featured the FOP (either with or without highlighting), with the information necessary to accurately answer the question in an un-highlighted portion of the DFL. See Figure 4.2 for an illustration of the trials that appeared in each block.

To carefully inform the design of experiments, active ingredients were paired to better manage the total number of questions participants would respond to so that the correct response was a "yes" for one, and a "no" for its pair. Two yes questions and two no questions were developed for each active ingredient, which the inverse used for its partner drug. For example, one pair of active ingredients was Acetaminophen and Ranitidine. The questions about the Acetaminophen label with the correct answer of "yes" (Such as, "Does this medicine contain Acetaminophen?"), would also be displayed for the Ranitidine label, where the correct answer was "no," and vice versa. Ibuprofen was paired with Dextromethorphan, Naproxen was paired with Omeprazole, and Cimetidine was paired with Phenylephrine. Questions used in the study are listed in Appendix E. 32 mock brands[7] were developed for 8 active ingredients (table 4.1).

[7] Each mock product was composed of a single, active ingredient commonly sold in US commerce at the time of the study

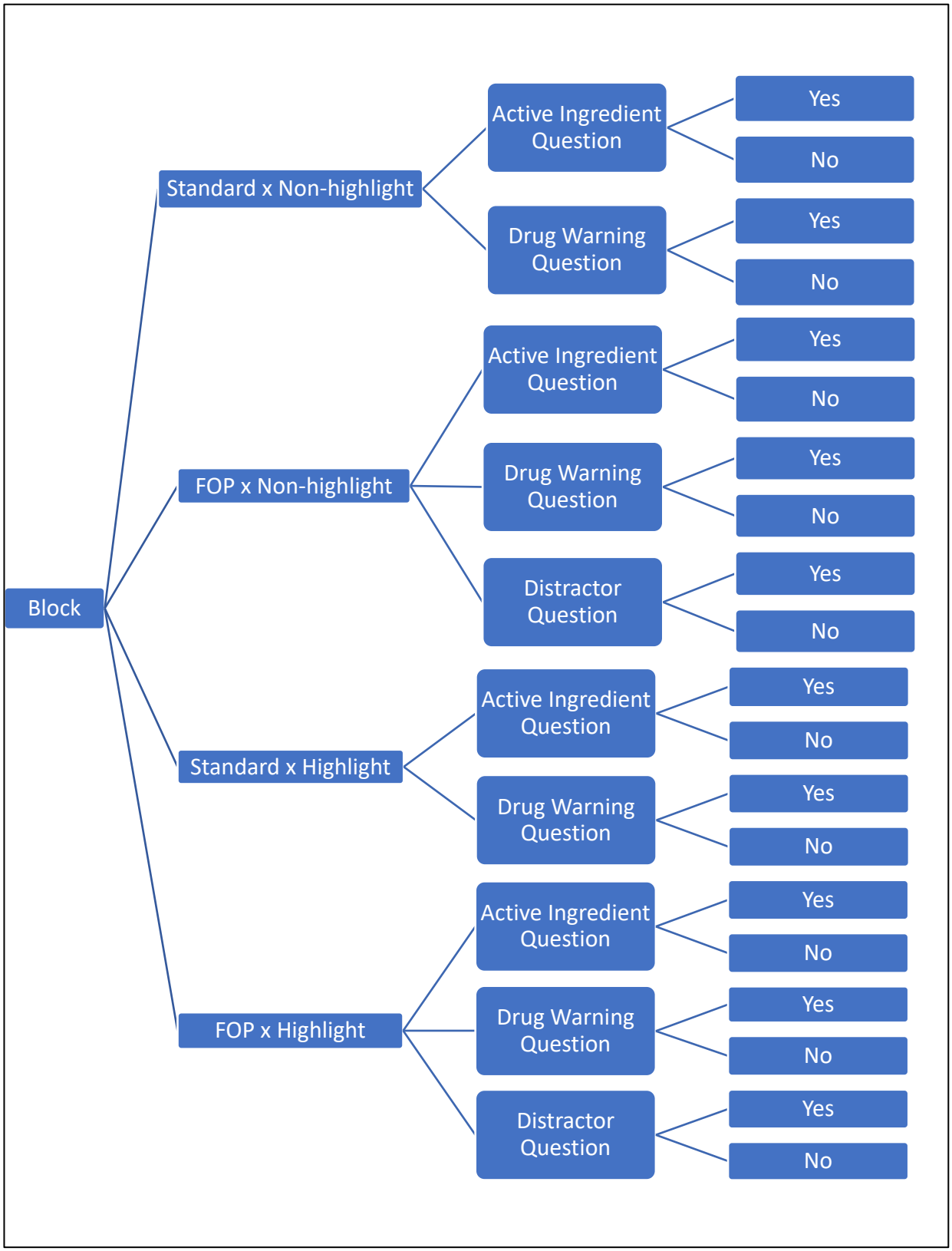


Figure 4.2 Diagram of Trial Structure for the Yes/No Absolute Judgement Task

Table 4.1 List of Active Ingredients and indications used in study		
Active Ingredient	Indication	Mockbrand Names
Acetaminophen	Pain Reliever/ Fever Reducer	Alendor, Zabinor, Lufnor, Gallicor
Cimetidine	Antiacid	Toftec, Varentec, Saridac, Xerbec
Dextromethorphan	Cough Suppressant	Cadoxtin, Clemdan, Circussin, Garswen
Phenylephrine	Decongestant	Corrigan, Rutaven, Enrallen, Hubarrin
Ibuprofen	Pain Reliever/ Fever Reducer	Thiretal, Rheidol, Bodrell, Hexidvil
Naproxen	Pain Reliever/ Fever Reducer	Naddel, Alladail, Dibidal, Harbenal
Omeprazole	Antiacid	Dantic, Shastic, Monach, Clindach
Ranitidine	Antiacid	Baxoden, Albanac, Lazarec, Recantac

To control for potential confounds with mock brand names or active ingredient, the experiment was divided into 8 blocks of 18 trials. Each block featured no repeat brand names and equal numbers of active ingredient or drug warning questions, and yes or no correct responses; the allocation of what label treatment was paired with each mock brand, question type, and correct responses were counterbalanced across 4 participants by developing 4 versions of the experiment's E-prime program. Thus, while version 1 might have a Highlight+FOP label of Mockbrand 1 used in a trial featuring an active ingredient question with the correct response of yes, for the same treatment level, Mockbrand 1 in version 2 would have an active ingredient question with the correct response of no, version 3 would have a drug warning question with the response of yes, and version 4 would have a drug warning question with the response of no. This counterbalancing ensured that any confounds between mock brand and label treatment were removed from the experiment.

Within the 16 critical trials in a block, each of the four label treatments (HL/no FOP; HL/FOP; NoHL/noFOP; NoHL/FOP) appeared four times (Figure 4.2). Within the four appearances comprised of the same treatment (e.g. HL/no FOP), each type of question (active ingredient based (y/n) or drug warning based (y/n)) appeared such that the correct numbers of yeses and nos were counterbalanced (i.e. 1 active ingredient correct response yes, one active ingredient correct response no, 1 drug warning correct response yes, 1 drug warning correct response no) . Labels for 8 active ingredients (see Table 4.1 for a list of the active ingredients employed in this study) were

developed with 4 brands per active ingredient. Each of the eight active ingredients appeared two times over the course of the 16 critical trials with each appearance of an active ingredient corresponding to a different mock brand.

Within each block of 18 trials, there were no repeating mock brands in an effort to limit participant learning specific to a given product, thereby forcing participants to use the labels to answer the questions. Additional analysis of participants' familiarity with active ingredients versus name brands was conducted to determine whether or not the mock branding was sufficient in masking prior knowledge about OTCs (See supplemental files).

Recruitment and Data Collection Procedures

The power estimates for the absolute judgement task were also based on previous work (Bix, Seo, Ladoni, Brunk, & Becker, 2016) with an effect size of d=0.84. But, as the previous work was conducted with younger, surgical technicians, (subject matter experts in medical device selection and opening), the power estimate was conducted with 50% of the measured effect size, or d=0.42 to be conservative in our recruitment efforts. 60 participants (recruiting 75 participants before attrition) allowed for a calculated detection at d=0.42 with power > 0.85.

Participants were recruited in accordance with the recruitment methods presented in Chapter 3, and were located in the Greater Lansing or Greater Flint areas. Qualifications to participate in this study were: manage your own medication, be age 65 or older, be legally sighted, and have consumed at least 1 OTC medication in the past year. 75 participants over the age of 65 were recruited to participate in this study in accordance with the estimated sample size developed using power calculations and accounting for up to 20% attrition.

In addition to the absolute judgement task, study participants also completed a survey that included: demographics (age, gender, race and ethnicity, native language, annual income, and educational attainment), a health literacy screening ((REALM-R) (Bass et al., 2003)), a test of memory

and concentration that was also used to screen participants unable to provide informed consent (Short Blessed Test (Katzman et al., 1983)), near point visual acuity (Sloan Pocket Size Near Vision Card with Continuous Text by Precision Vision in Woodstock IL) and ability to see color (Pseudo-Isochromatic Plates by Richmond Products, Southeast Albuquerque NM. Chapter 3 includes more precision description of how these measures were conducted. Participants were also asked to self-report familiarity with and perceived appropriateness of common OTC active ingredients. See the recruitment materials, consent forms, and the data sheet included as supplemental files for the exact language of the survey.

Statistical Analysis

Both the accuracy of the answers and the time it took participants to correctly answer were recorded as dependent variables of interest. Primary Analysis of trials requiring AI information and those which required DD information were analyzed separately to account for the prominence of realistically presented AI information relative to the DD. For both AI and DD trials, a Linear Mixed Model was used for the Reaction Time analyses, and a Binary Logistic Mixed Model was used for accuracy analysis, as this study is a within-subjects design. Two types of analysis were conducted for trials involving each type of information (the AI analysis and the DD analysis): the first compared the effects of highlighting and FOP in a 2x2 factorial design, and the second compared highlighted or FOP containing labels against the standard (i.e. the current commercial practice of an non-highlighted, no FOP label). Reaction time was truncated at 120,000 msec and then log (base 10) transformed prior to analysis to meet normality assumptions. Covariates in the final model included: sex, education level (as a binary variable of at least some college, or high school or less), race (as a binary variable of white, non-Hispanic or an individual from another racial or ethnic group), age, Near Point Visual Acuity, and Health Literacy (in the form of participants' Realm R score).

The final analysis incorporated participants' familiarity level related to common active ingredients for OTCs sold in the US, and popular brand name impacted accuracy in the absolute judgment task. Three sets of analyses were conducted related to these data. First, a familiarity score for active ingredient was calculated for each participant by summing the total number of "yes" responses in the survey questions asking about familiarity with active ingredients. This process was repeated to calculate a familiarity score for brand names with the "yes" responses to survey questions about familiarity with brands. Per the guidance of statistical consulting, these two scores were compared using Wilcoxon Signed-Rank Test, a non-parametric method of assessing significant differences between two variables.

Secondly, familiarity with each active ingredient and brand name were compared individually to examine if there were different levels of familiarity with different drugs. This analysis utilized McNemar's Test to test for significant differences between familiarity with brand names of OTCs and Active Ingredients of OTCs. Each pair (OTC active ingredient and corresponding OTC brand name commonly affiliated with that name) were tested independently in pairwise fashion to examine if there were differences within the individual brand-active ingredient pairs. Finally, familiarity with the active ingredient was included in the statistical model as a fixed effect to examine for possible effects of familiarity on either dependent variable (reaction time or accuracy) for both AI and DD trial types. The familiarity variable for this analysis was constructed by assigning each trial the binary familiarity rating the participant gave to the active ingredient in the trial.

Results

Over 6 months, 75 adults in the mid-Michigan area were recruited to participate in this study. Of the 75 recruited adults, 2 were screened out using the eligibility parameters (see IRB approved consent and advertisement Appendix G) due to age ineligibility (under age 65); 4 were released after consent was collected due to scores on the Short Blessed Test (9 or greater) that indicated an inability

to provide informed consent, and 1 withdrew from the experiment early due to technical difficulties. As such, data analysis included results collected from a total of 68 participants.

Testing was performed at 3 locations[8]: the first in Lansing MI, the second in East Lansing MI, and the third in Flint MI. Descriptive information about the participants is included below in table 4.2. The final sample included in analysis had 18 men and 50 women with a mean age of 71.95 years (SD=5.76). The racial and ethnic background of the sample was 57.4% White, 38.2% Black, 1.5% Asian, and 1.5% Native American, with 7.4% of participants reporting being Hispanic or Latino. Overall, participants preformed fairly accurately in this experiment, with an overall average accuracy rate of 80.1% across all completed trials.

[8] The testing in Lansing MI was hosted by the RSVP Foster Grandparents program at their office. Testing in East Lansing was hosted by the HUB Lab at the School of Packaging. Testing in Flint was hosted by MSU Extension of Genesee County.

Table 4.2 Description of the Sample	
Characteristic	N (%), Mean (SD, min, max)
TOTAL included in analysis	68 (100.0%)
Gender	
Men	18 (26.5%)
Women	50 (73.5%)
Race	
White	39 (57.4%)
Asian	1 (1.5%)
Black	26 (38.2%)
Native American or Alaskan Native	1 (1.5%)
Other	1 (1.5%)
Ethnicity	
Hispanic or Latino	5 (7.4%)
Not Hispanic or Latino	63 (92.6%)
Age	71.95 (SD 5.76, min 65, max 88)
REALM R (Scores <6 are at risk for poor health literacy)	8 (min 1, max 8)*
Visual Acuity	Mode= 20/32 (range 20/16 to 20/63)
Ability to see color	
Yes	66 (97.1%)
No	2 (2.9%)
Education Level	
Middle School	1 (1.5%)
High School	30 (44.1%)
Associate Degree	16 (23.5%)
Bachelor's degree	11 (16.2%)
Master's degree	7 (10.3%)
Doctoral Degree	3 (4.4%)
Income	$20,000 (SD $55,831.74, min$1,300, max $250,000)**
Native Language	
English	65 (95.6%)
Spanish	2 (2.9%)
Did not disclose	1 (1.5%)
* Central tendencies followed by a single asterisk are modes, as the type of information being recorded did not lend itself to a mean or median ** Central tendencies followed by two asterisks are medians, as the range was large and a mean would have been too heavily influenced by outliers	

Primary Analysis: Effect of label designs

Active Ingredient Results

The first analysis of the efficacy of the labeling intervention centered on the accuracy of participants responses to questions that relied on AI information. Accuracy in AI question trials was the response variable, and the predictor variables were: highlighting (present or absent), Label Type (FOP or Standard), the interaction between Highlighting and Label Type, race/ethnicity (Non-Hispanic white versus non-white or Hispanic), sex, education (dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

In this accuracy analysis, there was a significant main effect of label type, age and health literacy. Whereby, people with greater levels of health literacy were more accurate (coefficient of 0.222) and older participants less accurate than younger participants (coefficient of -0.79). Investigation of the main effect of label type suggested that the presence of an FOP improved accuracy regarding AI questions in the presence of an FOP was M = .919, SE = .019 compared to M = .880, SE = .022 for trial accuracy generated by the standard labels ($p<0.001$). Results of this analysis are included below in tables 4.3 and 4.4.

Table 4.3 Main effects of the AI Accuracy Analysis with a 2x2 Binary Logistic Mixed Model

Effect	F	df numerator	df denominator	p
Highlighting	.15	1	4178	0.694
Label Type (FOP/standard)	**13.84**	**1**	**4178**	**0.000**
Highlight X Label Type	.02	1	4178	0.887
Race/Ethnicity	1.95	1	4178	0.163
Sex	.71	1	4178	0.399
Education	.06	1	4178	0.805
Age	**6.59**	**1**	**4178**	**0.010**
Health Literacy (Realm R)	**5.19**	**1**	**4178**	**0.023**
Visual acuity	.41	1	4178	0.521

Table 4.4 2x2 AI Accuracy Binary Logistic Model		
Label Treatment	Estimated Marginal Mean	SE
Highlight+ FOP	.910	.021
Highlight+standard	.883	.020
Non-highlight+FOP	.908	.021
Non-highlight+standard	.878	.020

The second analysis of AI question trials investigated the effect of our label treatments on reaction time. Reaction Time in AI question trials was the response variable, and the predictor variables were: highlighting (present or absent), Label Type (FOP or Standard), the interaction between Highlighting and Label Type, race/ethnicity (Non-Hispanic white versus non-white or Hispanic), sex, education (dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

Results of the reaction time for the dependent variable, time to correct response for AI trials, revealed main effects of Label Type (p= 0.004), Health Literacy (Realm-R) (p=0.026), and Near Point Visual Acuity (p= 0.036). For the significant main effect Label type, the reaction time in units of log10ms for FOP treatments was M = 3.909, se = .025 and for standard M = 3.938, se = .025. Unsurprisingly, trends in the data show that as health literacy increased, reaction time decreased, and that as visual acuity worsened, reaction time increased. Results of this analysis are included in tables 4.5 and 4.6.

Table 4.5 Main effects of 2x2 AI Reaction Time Linear Mixed Model				
Effect	F	Df numerator	Df denominator	p
Highlighting (present/absent)	1.48	1	3568	0.224
Label Type (FOP/standard)	**8.14**	**1**	**3569**	**0.004**
Highlight x Label Type	.01	1	3568	0.913
Race/Ethnicity	3.80	1	58	0.056
Sex	.57	1	57	0.452
Education	.04	1	58	0.849
Age	.03	1	58	0.860
Health Literacy (Realm R)	**5.23**	**1**	**60**	**0.026**
Visual acuity	**4.59**	**1**	**58**	**0.036**

Table 4.6 2x2 AI Reaction Time Linear Mixed Model		
Label Treatment	Estimated Marginal Mean (log10 msec)	SE
Highlight + FOP	3.902	.026
Highlight + standard	3.933	.026
Non-highlight + FOP	3.916	.026
Non-highlight + standard	3.944	.026

As the goal of these studies is to investigate a strategy to improve the communication of critical OTC safety information additional analysis was conducted to assess whether or not the three label treatments differed significantly from the standard practice label that represents current labeling practice. These analyses included the same response variables with a predictor variable that included all four label treatments. The demographic predictor variables remained the same.

For the dependent variable, accuracy of response to questions requiring AI information, Post-hoc Bonferroni corrected pairwise tests compared the unhighlighted standard to the other three treatments. This analysis indicated that both of the FOP present treatments (highlighted FOP ME=0.910, SE= 0.019 and unhighlighted FOP ME=0.908, SE=0.019) resulted in significantly higher accuracy results compared to non-highlighted standard (ME= 0.878, SE= 0.023 (figure 4.3)). There was no evidence of a significant difference between the unhighlighted standard and the highlighted standard. These results are presented in table 4.7. Age and Health Literacy resulted in significant effects in this model as well, with age having a coefficient of -0.79 and health literacy having a coefficient of 0.222. Thus, as age increased accuracy decreased, and as health literacy increased, accuracy increased.

Table 4.7 AI Accuracy Results of 4 cell model results to compare each cell against standard practice				
Effect	F	Df numerator	Df denominator	p
Four Label Treatments	**4.68**	**3**	**4178**	**0.003**
Race/Ethnicity	1.95	1	4178	0.163
Sex	.71	1	4178	0.399
Education	.06	1	4178	0.805
Age	**6.59**	**1**	**4178**	**0.010**
Health Literacy (Realm R)	**5.19**	**1**	**4178**	**0.023**
Visual acuity	.41	1	4178	0.521

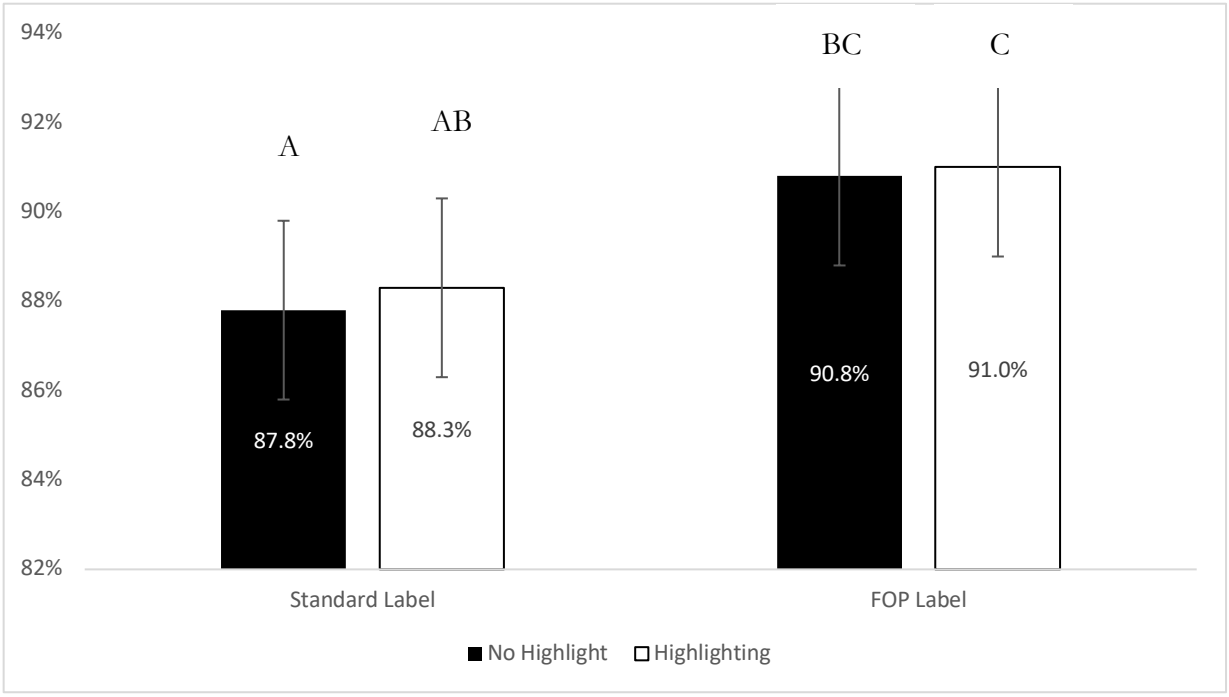


Figure 4.3 Estimated Accuracy for AI Trials. Variables with different letters above them are significantly different from the Non-highlight Standard label at the alpha=0.05 level using post-hoc Bonferroni corrections

This process of comparing the label treatments to standard practice was repeated in reaction time to assess whether or not the labeling treatments represented an improvement in reaction time when compared to the current labeling practice. With reaction time to correct response to AI questions as the dependent variable, the FOP with highlighting resulted in significantly faster responses (ME=3.902 log10ms, SE=0.026) than the standard label (no FOP/no highlight (ME= 3.944 log10ms, SE=0.026 p=0.024)(back-transformed data is presented in figure 4.4). There was no evidence of a significant difference in response time when the standard label without highlighting was compared to the other two treatments (standard highlighted ME= 3.933 log10ms, SE= 0.026 (p=1) and FOP no highlight ME=3.916, SE=0.026 (p=0.314)). Unsurprisingly, again, trends in the data suggest that as health literacy increased, reaction time decreased, and that as visual acuity worsened, reaction time increased. These results are presented below in table 4.8.

Table 4.8 AI Reaction Time Results of 4 cell model results to compare each cell against standard practice				
Effect	F	Df numerator	Df denominator	p
Four Label Treatments	**3.21**	**3**	**3568**	**0.022**
Race/Ethnicity	3.80	1	58	0.056
Sex	.57	1	57	0.452
Education	.04	1	58	0.849
Age	.03	1	58	0.860
Health Literacy (Realm R)	**5.23**	**1**	**60**	**0.026**
Visual acuity	**4.59**	**1**	**58**	**0.036**

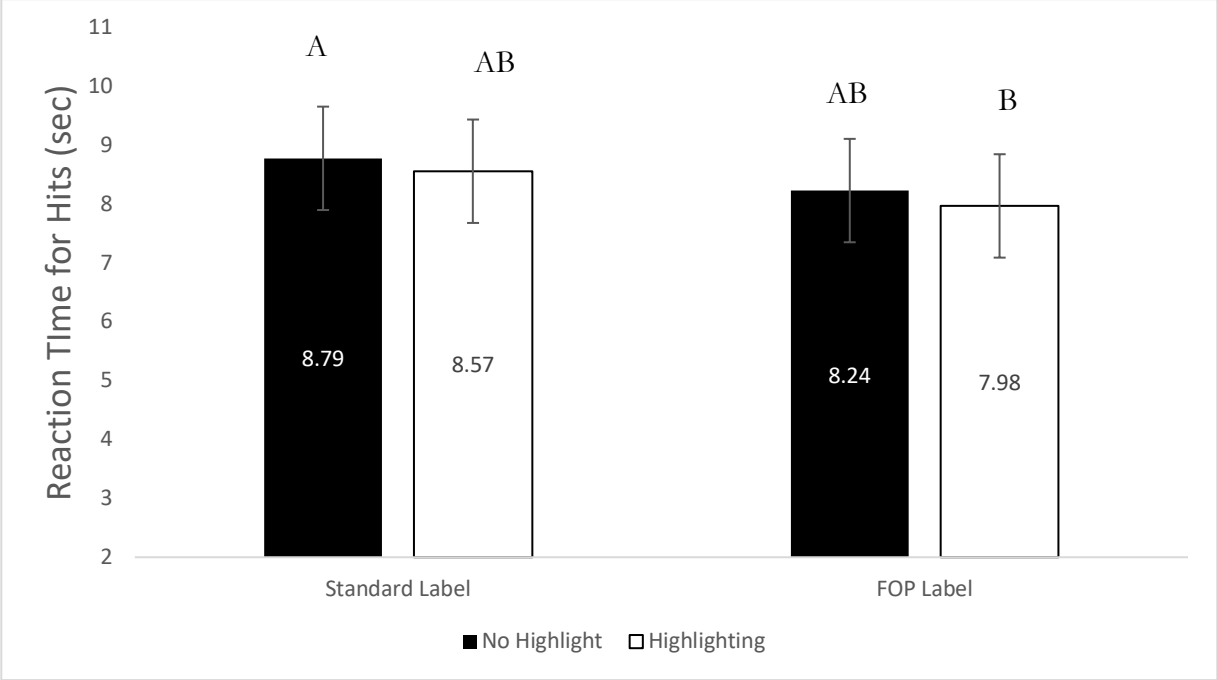


Figure 4.4 Estimated Marginal Means of Reaction time for Active Ingredient Trials with Standard, non-highlight as the comparison. Variables with different letters above them are significantly different at the alpha=0.05 level using post-hoc Bonferroni corrections. Error bars represent 95% confidence intervals

Drug-Drug and Drug-Diagnosis Interaction Warning Results

Analysis was repeated to examine the trials featuring DD questions. In the first of these analyses, accuracy in DD question trials was the response variable, and the predictor variables were: highlighting (present or absent), Label Type (FOP or Standard), the interaction between Highlighting and Label Type, race/ethnicity (Non-Hispanic white versus non-white or Hispanic), sex, education

(dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

In the accuracy analysis for responses to questions that required DD information, there was a significant label type main effect (p=0.001) and a significant highlighting main effect (p= 0.003), as well as main effects of age (p<0.001) and health literacy (p<0.001). A coefficient of 0.228 suggested. that people with greater levels of health literacy were more accurate than those at risk for poor health literacy. Consistent with the literature, the coefficient of -0.061 suggested older participants were less accurate than their younger counterparts. For the significant Highlighting main effect (p=0.003), the accuracy mean for highlighted trials was ME= 0.777, SE = .018 and ME = .738, SE = .019 for non-highlighted trial questions that relied on DD information to accurately answer. For the significant label type main effect (0.001), the accuracy mean for FOP treatments was ME= 0.779, SE= .018, and for standard it was ME=0.735, SE = 0.020. Results of this analysis is included in tables 4.9 and 4.10.

Table 4.9 Main effects of 2x2 DD Accuracy with a 2x2 Binary Logistic Mixed Model

Effect	F	Df numerator	Df denominator	p
Highlighting	**8.80**	**1**	**4182**	**0.003**
Label Type (FOP/standard)	**11.34**	**1**	**4182**	**0.001**
Highlight X Label Type	.00	1	4182	0.997
Race/Ethnicity	1.00	1	4182	0.318
Sex	2.00	1	4182	0.158
Education	.22	1	4182	0.643
Age	**14.83**	**1**	**4182**	**0.000**
Health Literacy (Realm R)	**17.91**	**1**	**4182**	**0.000**
Visual acuity	1.89	1	4182	0.169

Table 4.10 Estimated Marginal Means 2x2 DD Accuracy Binary Logistic Model

Label Treatment	Estimated Marginal Mean	SE
Highlight + FOP	.797	.019
Highlight + standard	.755	.021
Non-highlight + FOP	.760	.021
Non-highlight + standard	.714	.022

In the second of these analyses, reaction time for correct response in DD question trials was the response variable, and the predictor variables were: highlighting (present or absent), Label Type (FOP or Standard), the interaction between Highlighting and Label Type, race/ethnicity (Non-Hispanic white versus non-white or Hispanic), sex, education (dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

Results of the reaction time for correct response analysis for trials which required DD information revealed a significant main effect of Highlighting. Reaction time estimates for trials with DD information highlighted was ME = 4.323, SE= .030 and for non-highlighted ME= 4.367, SE = .030 ($p<0.001$). Results of this analysis are included below in tables 4.11 and 4.12.

Table 4.11 Main effects of 2x2 DD Reaction Time Linear Mixed Model

Effect	F	Df numerator	Df denominator	p
Highlighting	**15.77**	**1**	**2997**	**0.000**
Label Type (FOP/standard)	.77	1	2997	0.381
Highlight X Label Type	1.61	1	2997	0.204
Race/Ethnicity	2.39	1	59	0.127
Sex	.08	1	58	0.779
Education	.69	1	59	0.411
Age	2.96	1	60	0.090
Health Literacy (Realm R)	.04	1	62	0.849
Visual acuity	.06	1	60	0.814

Table 4.12 2x2 DD Reaction Time Linear Mixed Model

Label Treatment	Estimated Marginal Mean (log10 msec)	SE
Highlight + FOP	4.335	.026
Highlight + standard	4.311	.026
Non-highlight + FOP	4.365	.026
Non-highlight + standard	4.370	.026

Again, additional analysis was conducted to assess whether or not the three label treatments differed significantly from the standard practice label that represents current labeling practice as the goal of these studies is to improve the communication of critical OTC safety information compared

to the current practice. These analyses included the same response variables with a predictor variable that included all four label treatments. The demographic predictor variables remained the same.

For DD accuracy, post-hoc means tests using a Bonferroni correction indicated that the non-highlighted, standard labels (No Highlight, No FOP- standard label; ME = 0.714 , SE=0.022) differed significantly (p=0.034) from Highlighted FOP (ME = 0.797, SE=0.019) (Figure 4.5). The differences between the current commercial label treatment (standard label) and the other two treatments (standard with highlight (ME= 0.755, SE=0.021) and FOP without highlighting (ME=0.760 SE=0.021) did not indicate a significant difference in accuracy using the Bonferroni correction. These results are presented below in tables 4.13.

Table 4.13 DD Accuracy Results of 4 cell model results to compare each cell against standard practice

Effect	F	Df numerator	Df denominator	p
Label treatment	**6.66**	**3**	**4182**	**0.000**
Race/Ethnicity	1.00	1	4182	0.318
Sex	2.00	1	4182	0.158
Education	.22	1	4182	0.643
Age	**14.83**	**1**	**4182**	**0.000**
Health Literacy (Realm R)	**17.91**	**1**	**4182**	**0.000**
Visual acuity	1.89	1	4182	0.169

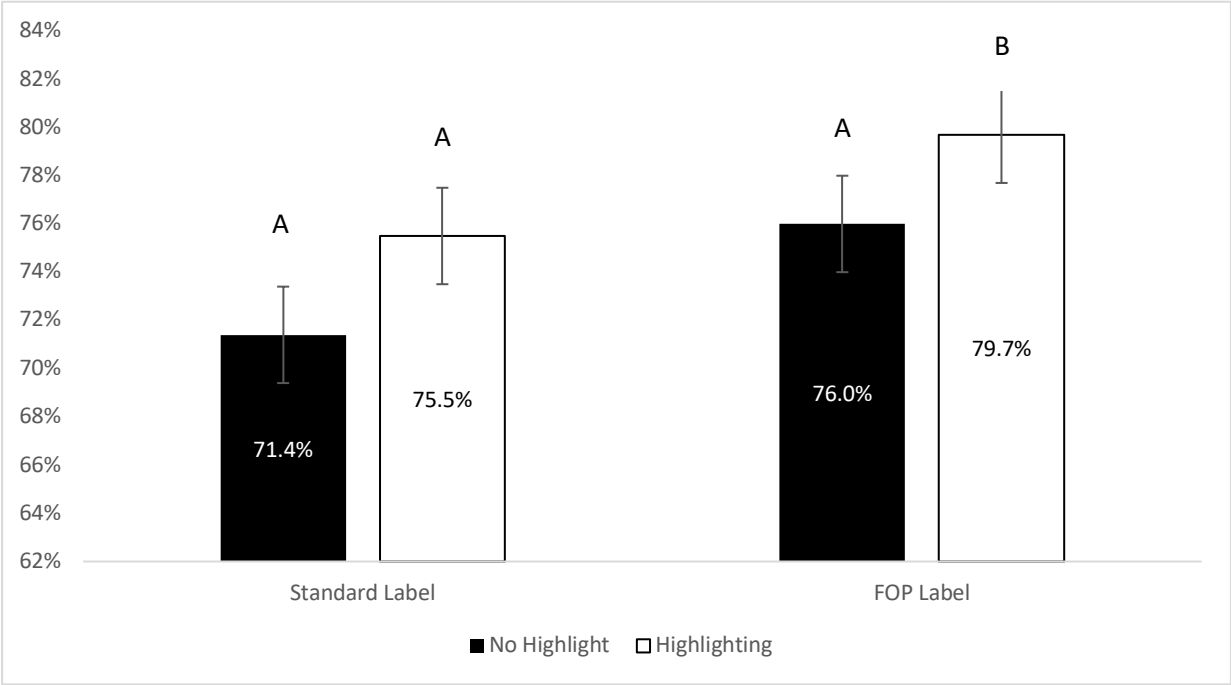


Figure 4.5 Accuracy Estimated Marginal Means for DD trials. Variables with different letters above them are significantly different from the Non-highlight Standard label at the alpha=0.05 level using post-hoc Bonferroni corrections

When we compared all of the labels against the standard, commercial practice (non-highlighted, no FOP) for RT (figure 4.6) for questions which relied on DD information using a post-hoc mean tests with Bonferroni correction, it was found that the standard practice cell (No Highlight, standard label; ME=4.370 log10ms, SE=0.031) differed significantly (p=0.002) from the Highlighted standard (ME=4.311 log10ms, SE=0.031) but not from Highlighted version with an FOP (M =4.335 log10ms, SE= 0.031) (p=0.174), or from the non-highlighted FOP (ME = 4.365 log10ms, SE=0.031)(p = 1)(see figure 4.6 for presentation of these results back transformed into units) These results are presented below in tables 4.15 and figure 4.6. Across the reactions which required accessing the DD information, the standard label with highlighting present performed significantly faster than all other treatments. However, it is important to remember that reactions to this label type (requiring DD) were significantly less accurate (ME= 0.755, SE=0.021) than those garnered by the highlighted label with an FOP present (ME = 0.797, SE=0.019) See Figure 4.5.

Table 4.14 DD Reaction Time Results of 4 cell model results to compare each cell against standard practice:				
Effect	F	Df numerator	Df denominator	p
Label Treatment	**5.98**	**3**	**2997**	**0.000**
Race/Ethnicity	2.39	1	59	0.127
Sex	.08	1	58	0.779
Education	.69	1	59	0.411
Age	2.96	1	60	0.090
Realm R	.04	1	62	0.849
Visual acuity	.06	1	60	0.814

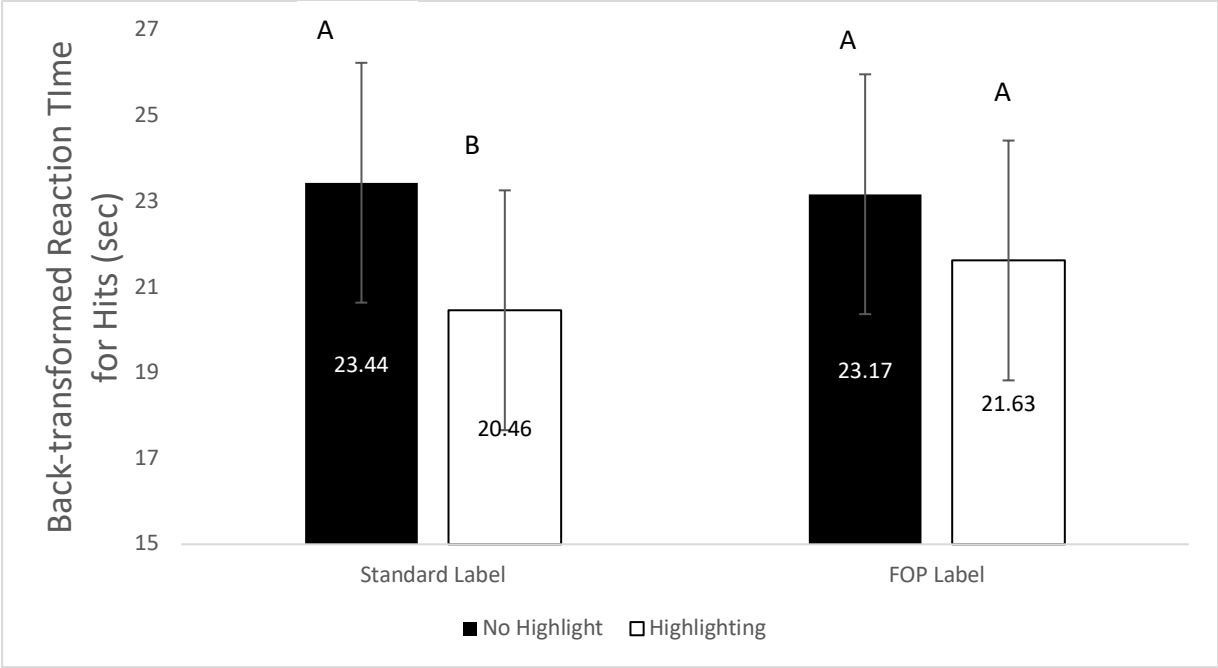


Figure 4.6 Back-transformed Reaction Time Estimated Marginal Means for DD trials. Variables with different letters above them are significantly different from the Non-highlight Standard label at the alpha=0.05 level using post-hoc Bonferroni corrections. Error bars represent a 95% confidence interval

Secondary Analysis: Familiarity with OTC active ingredients

Secondary analysis was conducted to examine: if there were significant differences between the distribution of participants' familiarity with brand names versus the distribution of their familiarity with active ingredients, if there were differences in the distributions of participants' familiarity with individual active ingredients and the corresponding most common brand name used in the United

States to sell that active ingredient, and finally if there was an effect of participants pre-existing familiarity with the active ingredients used in the study and reaction time and accuracy.

Familiarity Score

Overall, participants were significantly more familiar with brand names (mean=7.5, SD2.52) than they were with active ingredients (mean=3.4, SD=2.54 ($p<0.001$). Tables 4.25 and 4.26 present the results of a non-parametric hypothesis test for the Aggregated Active Ingredient Familiarity and Aggregated Brand Familiarity variables are presented. There was a significant difference between the medians of these variables. Figure 4.7 illustrates the distribution of familiarity scores for both active ingredients and brand names.

Table 4.15 Theory Test Summary for Difference between aggregated familiarity score for Active Ingredients versus aggregated familiarity score for Brands			
Null Hypothesis	Test	Sig.	Decision
The median of differences between AI Familiarity and Brand Familiarity equals 0.	Related-Samples Wilcoxon Signed Rank Test	0	Reject the null hypothesis.
Asymptotic significances are displayed. The significance level is .050.			

Table 4.16 Related-Samples Wilcoxon Signed Rank Test Summary	
Total N	68
Test Statistic	2007.5
Standard Error	145.688
Standardized Test Statistic	6.861
Asymptotic Sig. (2-sided test)	0

For all but one of the ten active ingredient-brand pairs we investigated, a significant difference was indicated when self-reported familiarity was analyzed. For all nine pairs that returned significant differences, people were more familiar with the brand name than the active ingredient that it typically contains. Only the active ingredient, Ibuprofen, had the same distribution of familiarity with the active ingredient as with the brand name. The other nine active ingredient-brand pairs had significantly different distributions related to familiarity scores, with more participants being familiar with the

brands than with the active ingredients. Visual presentation of the distribution of familiarity with OTC Active Ingredients versus Brand Names for all of the brands are presented below, and visual presentation of the pairwise comparisons are included in Appendix D.

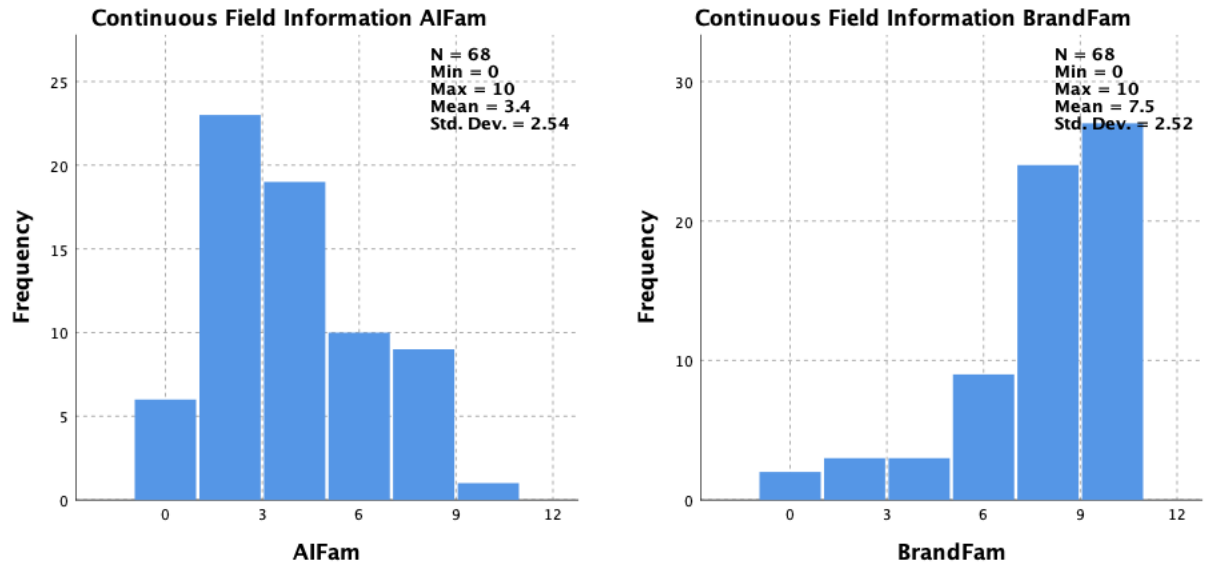


Figure 4.7 Distributions of participants overall Active Ingredient familiarity (AIFam) versus overall brand familiarity (BrandFam)

Differences in Familiarity Between Individual Active ingredients and Corresponding Brand names

After examining if there were differences between participants' familiarity with active ingredients and brand names across all products that we tested (Appendix G), differences in familiarity associated with specific pairs (i.e. a brand name paired with the active ingredient that it contained) was investigated. The results of this analysis are presented below in table 4.27.

Table 4.17 Familiarity with Active Ingredient typically affiliated with product versus Familiarity with Brand Hypothesis Test Summary				
Typical Active Ingredient-Brand	Null Hypothesis	Test	Sig.	Decision
Acetaminophen and Tylenol	The distributions of different values across Familiar Acetaminophen and Familiar Tylenol are equally likely.	Related-Samples McNemar Change Test	0.000a	Reject the null hypothesis.
Phenylephrine and Sudafed	The distributions of different values across Familiar Phenylephrine and Familiar Sudafed are equally likely.	Related-Samples McNemar Change Test	0	Reject the null hypothesis.
Cimetidine and Tagamet	The distributions of different values across Familiar Cimetidine and Familiar Tagamet are equally likely.	Related-Samples McNemar Change Test	0.001a	Reject the null hypothesis.
Diphenhydramine and Benadryl	The distributions of different values across Familiar Diphenhydramine and Familiar Benadryl are equally likely.	Related-Samples McNemar Change Test	0	Reject the null hypothesis.
Ranitidine and Zantac	The distributions of different values across Familiar Zantac and Familiar Ranitidine are equally likely.	Related-Samples McNemar Change Test	0	Reject the null hypothesis.
Omeprazole and Prilosec	The distributions of different values across Familiar Prilosec and Familiar Omeprazole are equally likely.	Related-Samples McNemar Change Test	0.012a	Reject the null hypothesis.
Dextromethorphan and Robitussin	The distributions of different values across Familiar Dextromethorphan and Familiar Robitussin are equally likely.	Related-Samples McNemar Change Test	0	Reject the null hypothesis.
Naproxen and Aleve	The distributions of different values across Familiar Naproxen and Familiar Aleve are equally likely.	Related-Samples McNemar Change Test	0	Reject the null hypothesis.
Ibuprofen and Advil	The distributions of different values across Familiar Ibuprofen and Familiar Advil are equally likely.	Related-Samples McNemar Change Test	0.549a	Retain the null hypothesis.
Guaifenesin and Mucinex	The distributions of different values across Familiar Guaifenesin and Familiar Mucinex are equally likely.	Related-Samples McNemar Change Test	0	Reject the null hypothesis.
Asymptotic significances are displayed. The significance level is .050. If the significance level is followed by an "a" an Exact significance is displayed for this test.				

Effects of familiarity on performance in the labeling task

The final set of analyses that included familiarity examined whether or not participant familiarity with the active ingredient present in each trial had an effect on either of the dependent variables studied (either reaction time to correct response or response accuracy). In this section, only the familiarity results will be discussed, as the main effects of the label formats were addressed earlier. This set of analyses was conducted by adding the fixed effect variable of familiarity and all appropriate interactions into the Linear Mixed Model for reaction time analysis and Binary Logistic Mixed Models for the response accuracy analysis. As previously mentioned, familiarity analysis was conducted for two reasons: first, to examine whether or not participants had differing levels of familiarity with OTC active ingredients compared with OTC brand names, and second, to see if familiarity impacted performance on the absolute judgment task (Study 3A). This second analysis provides insight into the efficacy of the mock branding strategy to mask prior knowledge.

The reaction time analysis that included familiarity and investigated how accurately (and how fast) participants were able to respond to questions that required information about the AI are presented below in table 4.28. Trials that required accessing the AI information resulted significant interaction between Label Type and Familiarity ($p=0.002$) when the dependent variable was reaction time to a correct response. This interaction is graphed in back transformed units in figure 4.7. Specifically, when participants were unfamiliar with the active ingredient, speed to accurate response was the same across label types, however, when participants indicated that they were familiar with the active ingredient, their reaction time was significantly slower in answering questions that required the AI when the standard label was present than for trials that included the FOP label.

Table 4.18 Main effects for AI reaction time with a 2x2 AI Linear Mixed Model with familiarity included				
Highlight	1.332	1	3547.823	0.249
Label Type (FOP/standard)	**12.373**	**1**	**3548.548**	**0**
Familiarity Effect	0.012	1	3604.444	0.914
Highlight X Label Type	0.028	1	3547.643	0.866
Highlight X Familiarity Effect	0.004	1	3548.254	0.949
Label Type X Familiarity Effect	**9.23**	**1**	**3548.937**	**0.002**
Highlight X Label Type X Familiarity Effect	0.189	1	3548.076	0.664
Race/Ethnicity	3.818	1	58.079	0.056
Sex	0.549	1	57.004	0.462
Education	0.037	1	58.101	0.847
Age	0.038	1	58.194	0.847
Health Literacy (Realm R)	**5.198**	**1**	**60.196**	**0.026**
Visual acuity	**4.591**	**1**	**58.321**	**0.036**
Effect	F	Df numerator	Df denominator	p

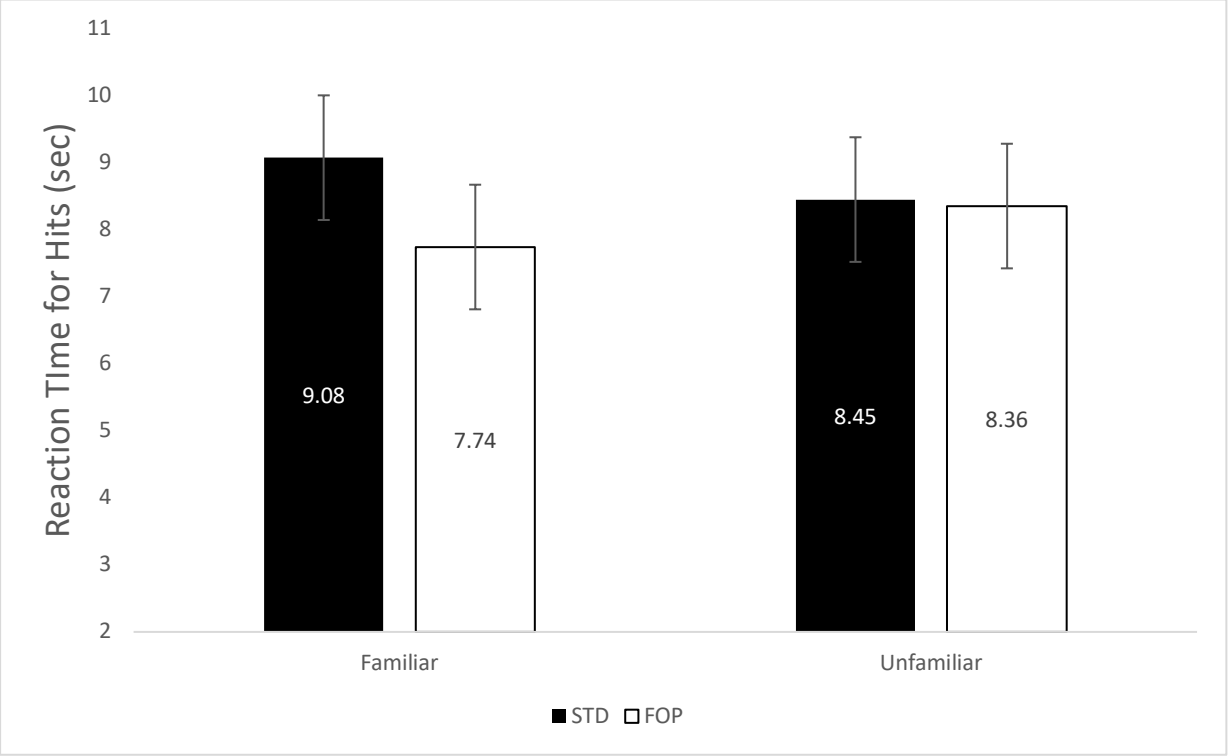


Figure 4.8 Reaction Time: AI trial type, Label Type X Familiarity. Error bars represent a 95% confidence interval

The accuracy analysis for these trials (those which required AI information also identified a significant interaction with label type and familiarity (p=0.025) (see Table 4.29). Like with the Reaction Time analysis, the novel FOP treatments were benefits were related to whether or not the participant was familiar with the medication (See figure 4.8). As with the reaction time, the novel treatment was more effective for active ingredients were familiar to participants.

Table 4.19 Main effects for AI Accuracy with a 2x2 DD Binary Logistic Mixed Model with familiarity included

Effect	F	Df numerator	Df denominator	p
Highlight	0.208	1	4158	0.648
Label Type (FOP/standard)	**18.606**	**1**	**4158**	**0**
Familiarity Effect	0.411	1	4158	0.521
Highlight X Label Type	0.071	1	4158	0.79
Highlight X Familiarity Effect	0.002	1	4158	0.964
Label Type X Familiarity Effect	**5.04**	**1**	**4158**	**0.025**
Highlight X Label Type X Familiarity Effect	1.052	1	4158	0.305
Race/Ethnicity	1.926	1	4158	0.165
Sex	0.597	1	4158	0.44
Education	0.061	1	4158	0.805

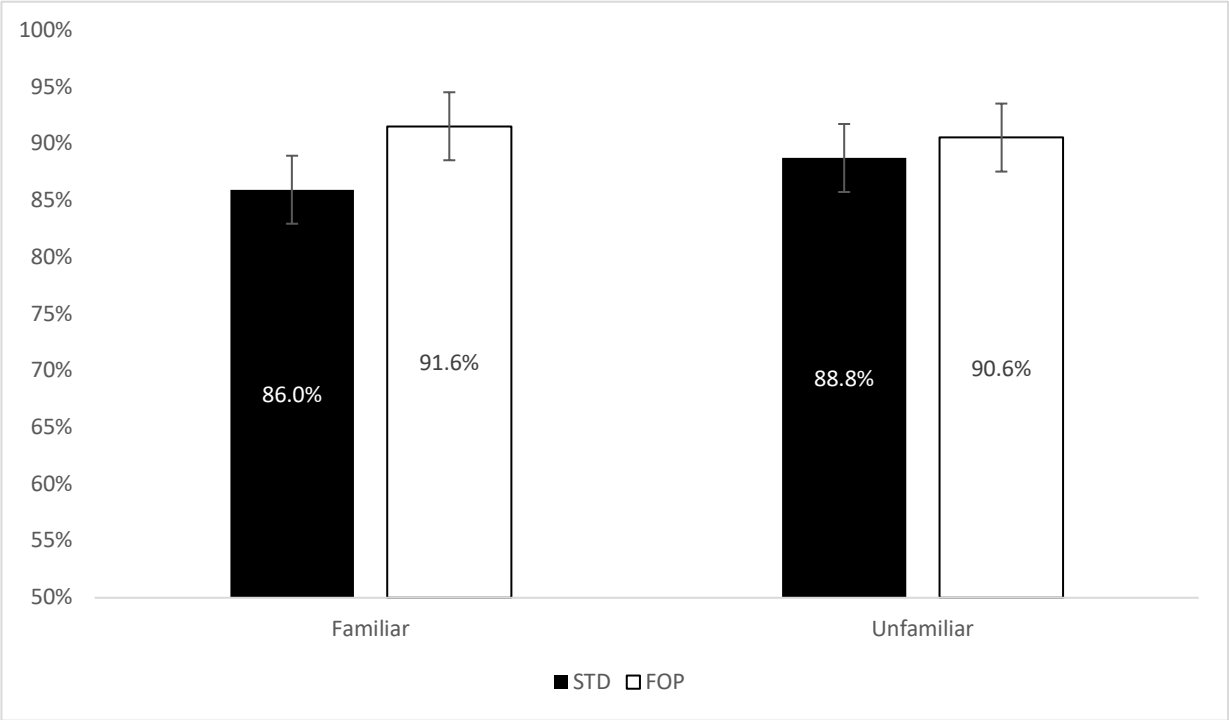


Figure 4.9 Accuracy: AI, Label Type x Familiarity

This same type of analysis was conducted for information that required accessing the DD. For the dependent variable of reaction time for correct response, in trials that required the DD a significant interaction was identified between Highlighting and Familiarity (p=0.027) (See Table 4.28). This interaction is graphed in figure 4.9. Findings suggest a beneficial effect of highlighting specific to whether or not the participant was familiar with the medication (See figure 4.9). As with the previous discussion, the impact of a novel approach is mediated by prior familiarity with the active ingredient. However, in the previous case (questions related to AI information- See Figure 4.9), the novel label (FOP presence) was more beneficial to those who reported prior familiarity with an active ingredient compared to active ingredients that they reported as unfamiliar. When the trials that required DD information were examined for an effect on reaction time, the significant interaction was related to highlighting and familiarity. The trend of the interaction was consistent. That is the novel treatment (in this case highlighting) had a larger positive effect for those active ingredients that were familiar to participants than for active ingredients that they did not know (See Figure 4.10).

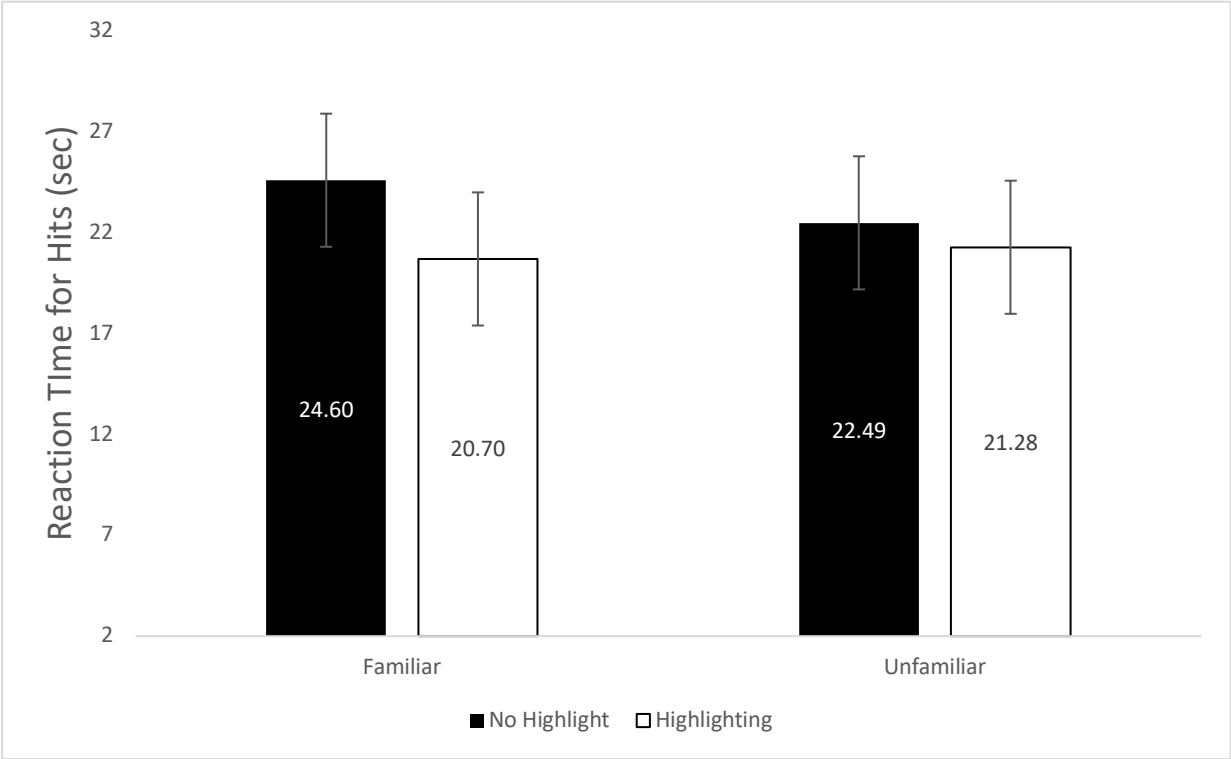


Figure 4.10 Reaction Time: DD Highlight x Familiarity

Table 4.20 Main effects for DD reaction time with a 2x2 DD Linear Mixed Model with familiarity included				
Effect	F	Df numerator	Df denominator	p
Highlight	**18.452**	**1**	**2976.744**	**0**
Label Type (FOP/standard)	0.836	1	2976.892	0.361
Familiarity Effect	1.039	1	3026.598	0.308
Highlight X Label Type	1.741	1	2976.758	0.187
Highlight X Familiarity Effect	**4.9**	**1**	**2976.769**	**0.027**
Label Type X Familiarity Effect	0.251	1	2976.934	0.616
Highlight X Label Type X Familiarity Effect	0.003	1	2976.567	0.955
Race/Ethnicity	2.463	1	58.79	0.122
Sex	0.099	1	58.34	0.754
Education	0.676	1	58.678	0.414
Age	2.975	1	59.986	0.09
Health Literacy (Realm R)	0.037	1	61.67	0.848
Visual acuity	0.052	1	59.552	0.82

Summary of Results

Overall, the results of this study indicate that enhancing the standard OTC label with highlighting and an FOP would improve the likelihood that older adults correctly interpret the critical health information presented on OTC medication labels (response accuracy). In terms of the scientific hypotheses (Gotelli & Ellison, 2004) described at the beginning of the chapter:

Hypothesis 1: *Highlighting will increase the accuracy of participants responding to yes/no questions about the labels compared to non-highlighted labels*. This hypothesis was supported for DD information type, and through an interaction between highlighting and label type for the AI information type.

Hypothesis 2: *The front of pack label will increase the accuracy of participants responding to yes/no questions about the labels compared to labels without the front of pack label.* This hypothesis was supported by both the AI and DD results, as there were significant label type main effects for accuracy in both AI and DD information types.

Hypothesis 3: *Highlighting will decrease the amount of time required for participants to accurately respond to yes/no questions about the labels compared to non-highlighted labels*. This hypothesis was partially supported by the beneficial results of highlighting on time to respond for DD information questions. However, as a benefit of highlighting alone for AI information type was not detected, this hypothesis was only partially supported.

Hypothesis 4: *The front of pack label will decrease the amount of time required for participants to accurately respond to yes/no questions about the labels compared to labels without the front of pack label.* This hypothesis was partially supported by the beneficial results of highlighting on time to respond for AI information questions. However, as a benefit of highlighting alone for DD information type was not detected, this hypothesis was only partially supported.

Discussion and Implications

The results of this assessment of the top-down processing of OTC medication labels by older adults will be discussed further in Chapter 6 in tandem alongside the results of the change detection (bottom-up) and cross product comparison (top-down) tasks. Chapter 6 will discuss the results of all three experiments, as each focus on a specific attribute that is necessary to fully determine an optimized OTC label design. In this particular study, the results of the assessment of the addition of highlighting and a FOP are promising.

While it is surprising that in AI trials the FOP significantly improved performance in terms of both speed and accuracy as opposed to highlighting alone (as the AI information does not appear in the FOP), it is a promising sign that the FOP was not hindering participants ability to effectively process the labels. The highlighted FOP was significantly better than standard practice in facilitating answering AI questions in terms of both speed and accuracy of responses. Unsurprising is the conclusion that highlighting is helpful to older adults' attempting to answer DD questions, as the highlighting was designed to draw attention to the requisite DD information. As the FOP is beneficial in AI trials and highlighting is beneficial in DD trials, a label optimized to communicate both AI and DD information would feature both highlighting and an FOP, according to this study which assesses products in isolation, without comparison to another product.

Additionally, it appears familiarity with an OTC induces a bias to miss information presented on a standard label. As participants who were familiar with an active ingredient did worse relative to participants who were unfamiliar in standard label formats (standard for AI trial types and non-highlighted in DD trial types), undertaking label optimization strategies seems to be even more necessary considering mock brands were used in this experiment, and participants were overwhelmingly familiar with OTC brands. This study provides some evidence that the FOP facilitates the use of the label as a refutation text (M. P. Ryan et al., 2017), by encouraging more careful

consideration of the product information on the label rather than reliance on prior knowledge, whether or not that information is appearing in the FOP.

Moreover, there are major implications for pharmacists and healthcare practitioners in the results of the analysis of familiarity with brand names versus the active ingredients the brands contain. While these results confirm a common assumption, little work has been published concerning the ramification of the differences in patient familiarity with active ingredients versus brand names (Aker et al., 2014; Hanoch, Gummerum, & Brass, 2007; Kauppinen-Räisänen, Owusu, & Bamfo, 2012), and how advise from a healthcare provider or pharmacist can adapt to account for differing levels of familiarity. Developing improved labels that emphasize critical safety information, such as the active ingredient, could help reduce barriers to compliance with recommended OTC treatments when name brand alternatives aren't available.

Chapter 5 A Dichotomous Cross-Product Comparison Forced Choice Task

Overview

The objective of this study served to evaluate how the presence of an FOP label and highlighting critical information affected the attention of older adults when that information is explicit to the viewer's goal of comparing the information presented on two OTC labels. As consumers of OTC medications select a drug for consumption from a large option set, assessing how label design impacts cross-product comparisons is crucial for the development of an optimized label. The results presented within this chapter are preliminary as the outbreak of the novel coronavirus, COVID-19, resulted in a temporary suspension of human subject data collection starting in the spring of 2020, extending through spring 2021.

While this study compliments the study presented in Chapter 4, the previously presented study utilized an absolute judgement task to assess how OTC labels presented in two treatments, each at two levels (highlight- present and absent; FOP- present and absent), impacted the ability of older adults to utilize critical information. This chapter focuses on a forced-choice task to further assess the same treatments' performance in a different, decision-making task. In this study, participants were shown a question which required accessing critical information to select the most appropriate product. Participants were presented two mock-products drug labels (both at the same treatment level; i.e. both highlighted with an FOP present), and asked to select the medication appropriate for the specific question (Figure 5.1). Because the question required accessing critical information, it engaged and assessed the designs ability to engage, the top-down processing mechanisms of the user while also assessing the ability to facilitate cross-product comparisons by older consumers.

Herein, we continue to objectively investigate top-down processing (Kinchla & Wolfe, 1979) of the 2x2 crossing of the two label treatments (FOP present/absent; Highlight present/absent), however this study differs from the absolute judgement task presented in Chapter 4 as it is a cross-

product comparison, forced choice procedure and simulates an information processing goal that might be conducted while shopping. In this study, the focus of our investigation is the extent of the label treatments' facilitation of searching for, interpreting, and comparing information critical for the safe and effective use of OTCs. The same label treatments assessed in the previous two chapters are being assessed in this work, see figure 3.1 for an example of the treatments. The scientific hypotheses (Gotelli & Ellison, 2004) are:

Hypothesis 1: Highlighting will increase the accuracy of participants choosing between the labels compared to non-highlighted labels.

Hypothesis 2: The FOP label will increase the accuracy of choosing between the labels compared to labels without the front of pack label.

Hypothesis 3: Highlighting will decrease the amount of time required for participants to choose between the labels compared to non-highlighted labels.

Hypothesis 4: The FOP label will decrease the amount of time required for participants to choose between the labels compared to labels without the front of pack label.

Methods and Materials

The methods and materials section of this chapter first discusses the design of the experiment, secondly describes the materials and methods used to generate the experimental stimuli, thirdly discusses the recruitment and data collection procedures including the screening criteria, and finally describes the statistical analysis strategy and methodology used to analyze the data. Methods were approved by the MSU Psychology and Social Science Internal Review Board in Summer 2018 as STUDY00000832.

Experimental Design

In order to determine how the presence of an FOP label and highlighting critical information affect the attention of older adults when comparing two OTC products, this study evaluated the four

label designs discussed previously, but with the labels presented in pairs. Participants were instructed to choose the appropriate product given a question which appeared on the screen simultaneously on the screen with the stimuli. The trials in this study were developed by pairing trials from the previous forced choice task of yes/no questions and rewriting the question to be a forced-choice comparison of the two products (same level of label of design in flattened images of PDP and DFL), see Figure 5.1 for an example of a trial and Appendix C for the list of questions. As with the previous experiments, critical trials contained questions that required AI information, specifically, relating to the name of ingredient or amount of ingredient, or DD information, relating to either drug-diagnosis or drug-drug warnings. As with the previous study, non-critical trials featured questions that asked about the uses of the medications. This study leverages the same pairs of drugs that were utilized previously (table 5.1); recall that there were 128 critical trials and 16 trials investigating the distraction effect of the treatments for each participant to complete in the yes/no forced choice task. Due to the paired nature of trials in this study, there were 64 critical and 8 distraction effect trials for each participant, see table 5.1 for the pairings. The preliminary results of 49 participants[9] who completed the 72 trials previously described are presented.

The same eight active ingredients and 32 brand names (table 4.1) used in the yes/no task were used in this cross-product comparison task, see table 5.1 below for a visual representation of how the active ingredients were paired. Active ingredients were paired for question development in the same manner as the yes/no forced choice task. Acetaminophen always appeared alongside Ranitidine, Ibuprofen alongside Dextromethorphan, Omeprazole alongside Naproxen, and Cimetidine alongside

[9] The power estimates for this study are the same as those in Chapter 4, and were based on previous work (Bix et al., 2016) with an effect size of d=0.84. To account for the younger age and skill of the previous sample, the power estimate was conducted with 50% of the measured effect size, or d=0.42. 60 participants (recruiting 75 participants before attrition) should allow detection at d=0.42 with power > 0.85. However, due to the IRB's decision to stop collecting human subject data in March of 2020, only 49 participants had participated at the time of this write up of the results.

Phenylephrine. As mentioned, pairs of active ingredients were selected so that the active ingredients were not from the same product category (e.g. pain relievers), so that they were less likely to contain similar contraindications. This pairing structure afforded unambiguous questioning about the different active ingredients, example questions are included below for each of the active ingredients included in the study. Questions used in the study are listed in Appendix E.

Table 5.1 Examples of the pairings of active ingredients and their purposes alongside example questions about each drug in the pair

Active Ingredient 1 (AI1)	**Active Ingredient 2 (AI2)**	**Example Question with AI1 as Correct Response**	**Example Question with AI2 as Correct Response**
Acetaminophen pain reliever/fever reducer	Ranitidine antiacid	Which medication should someone avoid consuming 3 or more alcoholic drinks while taking?	Which medication should someone consult a doctor before taking if they suffer from chest pain and shortness of breath?
Cimetidine antiacid	Phenylephrine nasal decongestant	Which medication should someone contact a doctor about if they have nausea or vomiting?	Which medication should someone consult their doctor about if they have a chronic cough with too much phlegm?
Dextromethorphan cough suppressant	Ibuprofen pain reliever/fever reducer	Which medication should be avoided by someone using a prescription for Parkinson's disease?	Which medication has a higher chance of stomach bleeding when taking the medication, if age 60 or older?
Naproxen pain reliever/fever reducer	Omeprazole antiacid	Which medication should someone consult their doctor about before taking if they suffer from kidney disease?	Which medication should someone avoid if they have pain swallowing food?

Counterbalancing was completed so that for active ingredient in each label treatment, there was an occurrence of every possible trial scenario (see figure 5.2 for visualization of the counterbalancing). The experiment was counterbalanced so that the side of the screen that each drug

label in a pair was displayed (left versus right) and which side the correct response was displayed appeared an equal number of times for each active ingredient pair. Additionally, the questions about the drugs appeared an equal number of times across the experiment. In order for a completely counterbalanced experimental design to be used without increasing chances of run order effects due to either learning or exhausted participants, 4 versions of the program were developed. As in the previous two studies, both accuracy of the responses and reaction time to correct response (the amount of time between onset of the stimuli presentation and participant response) were recorded and analyzed. Figure 5.2 demonstrates the degree to which this experiment was counterbalanced.

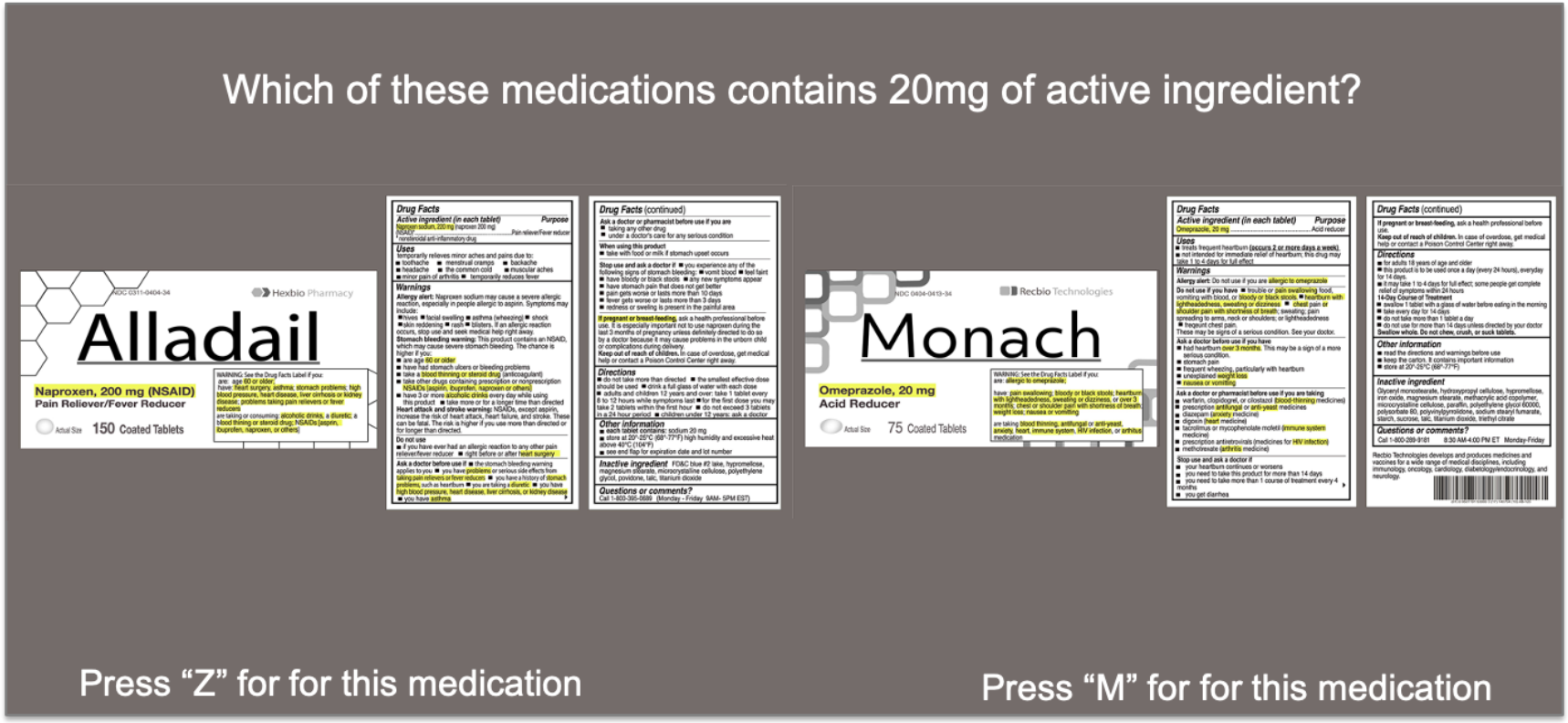


Figure 5.1 Example of a trial with the highlight and FOP treatment in the cross-product comparison task. This experiment was displayed on 34” ultra-wide screen monitors so that the OTC medication labels could appear side by side and be fully legible

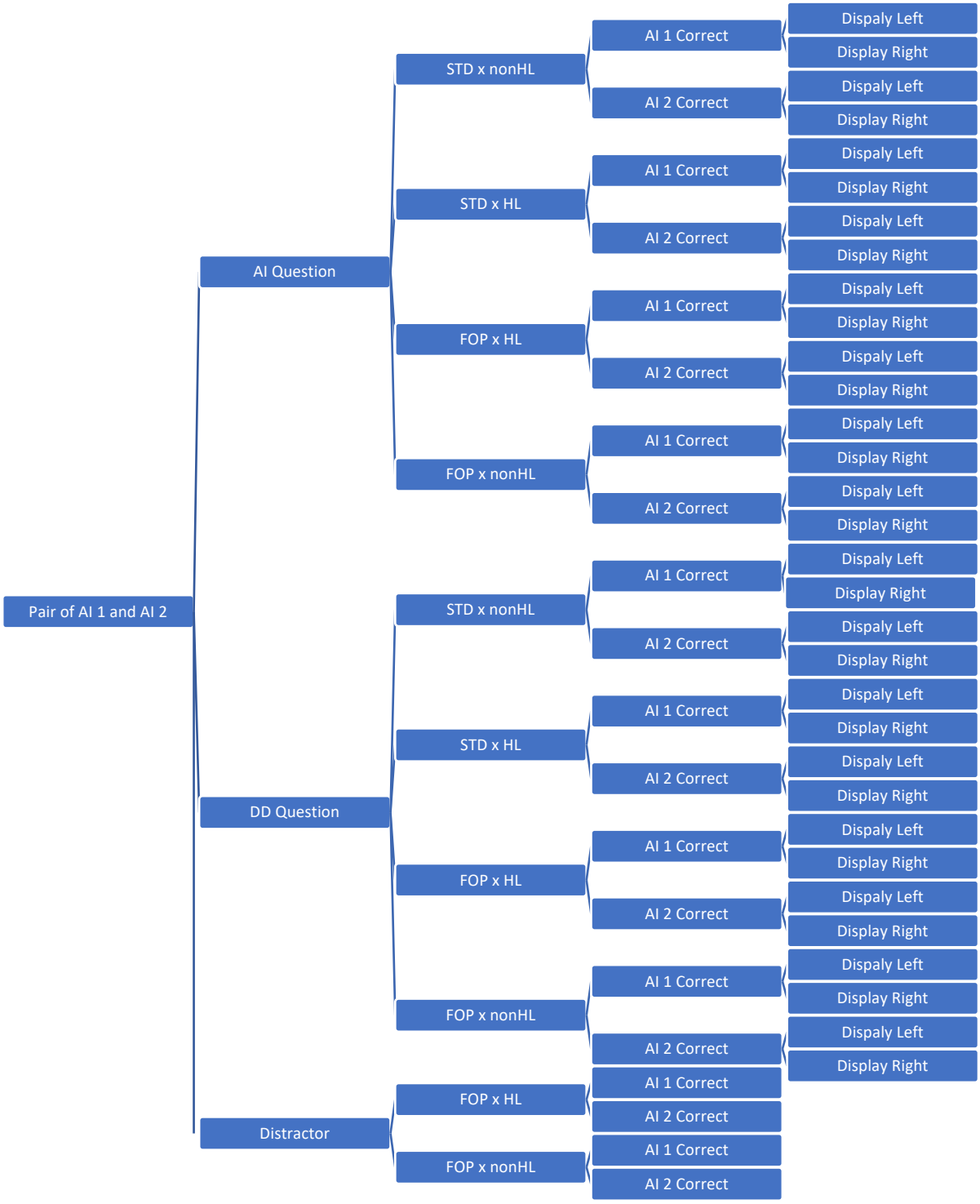


Figure 5.2 Diagram of all trials for a single ordered pair of active ingredients. As there are 4 pairs of active ingredients used in this study, this diagram would be multiplied by 4, resulting in 144 trials for this ordering of the active ingredient pairs

Materials

Experimental stimuli were designed and developed in Adobe Illustrator (Adobe Systems version 7, Incorporated San Jose, CA) and were based on commercial designs and informed by Title 21CFR subpart C §201.66 which specifies both content and formatting specific to OTC drugs. The experiment was programmed and run using E-Prime version 3 (Psychology Software Tools, Sharpsburg, PA). As with the previous studies, two styles of laptops: the Dell Latitude 5490 BTX, with an 8th Gen Intel Core i5-8350U (Quad Core, 6M Cache, 1.7GHz, 15W, vPro), running Windows 10 Professional at 2400MHz with 8 GB of RAM, and the Dell Latitude 5480, and the Dell XCTO, also with an 8th Gen Intel Core, 2X8GB of RAM, running Windows 10 Professional were used. The laptops were used in tandem with 34" ultra-wide-screen monitors with a resolution of 3440 x 1440 at 60 Hz. The use of these monitors for this study enabled the research team to present the two labels simultaneously side by side without compromising image clarity or font size within the labels; this was done by dividing the resolution of the width of the screen in two, and then using the aspect ratio of the stimuli to solve for the ideal height of the label. This resulted in the labels being saved in a resolution of 1720px by 741px, and displayed side by side in that resolution at runtime. For each trial, the question was displayed for 3 seconds before the two images appeared. Participants pressed either the "z" key to select the image on the left or the "m" key to select the image on the right. When participants pressed the key, the experiment automatically moved to the next trial.

Recruitment and Data Collection Procedures

After reviewing and signing the consent form (see supplemental files) and before starting the forced choice, cross product comparison task, study participants also completed a paper based survey (included as a supplemental file) that included: demographics (age, gender, race and ethnicity, native language, annual income, and educational attainment), a health literacy screening ((REALM-R) (Bass et al., 2003)), a test of memory and concentration that was also used to screen participants unable to

provide informed consent (Short Blessed Test) (Katzman et al., 1983), near point visual acuity (Sloan Pocket Size Near Vision Card with Continuous Text by Precision Vision in Woodstock IL) and ability to see color (Pseudo-Isochromatic Plates by Richmond Products, Southeast Albuquerque NM), and familiarity with, and perceived appropriateness of, common over the counter active ingredients. As detailed in previous chapters, if participants scored above an 8 on the Short Blessed Test, they were dismissed from the study and provided with the participant incentive.

After completing the survey (see supplemental files), participants sat at a computer station for the cross-product comparison task. When participants first sat at the computer station, a research assistant asked if they needed the monitor adjusted, and helped adjust the monitor up or down depending on the participant's need for comfort and preference for viewing. Once the participant was comfortable in their computer station, the research assistant opened the correct version of the E-prime program and input the participant number to start the program. Next, participants were greeted with the following message "Welcome to this Over-the-Counter Medication Labeling Experiment. Thank you for your willingness to participate. Please press any key to continue." After pressing a key, participants then saw these instructions before starting the task,

> "In this experiment, each trial will consist of two pictures of Over-the-Counter medicine labels and a question about the medicine. Please select the medicine that best answers the question as quickly as possible. Press the Z key to select the medicine label on the left, and the M key to select the medicine label on the right. Please ask the researcher any questions you might have before you press the space bar to continue."

Research assistants then either answered participant questions about the task, or instructed the participant to press the space bar to start if they had no questions. On the bottom of the screen for each trial was a reminder to press Z for the medicine on the left, and M for the medicine on the right.

Participants for this study were recruited in accordance with the recruitment methods presented in Chapter 3, and were located in the Greater Flint, Greater Lansing or Greater Grand

Rapids areas[10]. Qualifications to participate in this study were: manage your own medication, be age 65 or older, be legally sighted, and have consumed at least 1 OTC medication in the past year. Forty-nine out of 75 participants over the age of 65 had been recruited at the time of this writing. A goal of 60 participants (after attrition) was established in accordance with the estimated sample size developed using power calculations described in Chapter 4. However, data collection was suspended due to the outbreak of the novel coronavirus, COVID-19 and the shutdown of in-person human subject data collection at Michigan State University in the spring and summer of 2020.

Statistical Analysis

Both the accuracy of the answers and the time it takes for the participant to answer correctly were recorded and used as dependent variables in analysis. Like in the previous two studies, Analysis of AI trials and DD trials were conducted separately to account for the inevitable confound that occurs across the AI and DD trials due to the relative prominence of AI information as compared to the DD information (see the labels in Appendix C). For trials that required AI information and for trials that required DD information, a Linear Mixed Model was used for the reaction time analyses (time to correct response), and a Binary Logistic Mixed Model was used for analysis of the accuracy information. For both dependent variables (accuracy and reaction time to correct response) for both types of critical information (AI and DD), two types of analysis were conducted: the first compared utilized a 2x2 factorial design to investigate how factors of interest (highlighting and FOP) influenced performance, and the second compared how treatments with enhanced labels (Highlight x Standard, non-highlight x FOP, or Highlight x FOP) performed relative to current practice (Non-highlight, no FOP treatments). Reaction time was truncated at 120,000 milliseconds (120 seconds) and then log (base 10) transformed prior to conducting analyses in order to meet normality assumptions.

[10] Participants in the Flint area were recruited with the help of MSU Extension of Genesee County and participants in the Grand Rapids area were recruited with the help of MSU Extension of Kent County. Participant in the Lansing area were recruited through local senior centers.

Covariates in the final model included: sex, education level (as a binary variable of at least some college, or high school or less), age, Near Point Visual Acuity, and Health Literacy (in the form of participants' Realm R score). While racial and ethnic background data were collected as part of the survey, limited sample variability resulted in it being omitted from the analysis.

A final analysis examined the distraction effect trials. In this analysis, the distraction effect trials were added to the Linear Mixed Model (for reaction time) or Binary Logistic Mixed Model (for accuracy). Two analysis were conducted. The first to compare distractors to the standard label with DD questions and the second to compare the distractors to the FOP labels with DD questions.

Results

Preliminary results obtained from 49 participants recruited from the greater-Lansing, greater-Flint, and greater-Grand Rapids areas during the winter of 2020 are presented. Due to orders to cease data collection with all human subjects due to the COVID 19 pandemic (occurring in the early spring of 2020), presented results only comprise the first 49 participants collected. Overall, participants performed very accurately, with an overall average accuracy of 96% on all trials. No participants were dismissed due to Short Blessed Test Score, and no participants withdrew early. Descriptive information about the participants is included below in table 5.2. The final sample included in analysis has 13 men and 36 women with an average age of 73.67 (SD 5.10) years. The racial and ethnic background of the sample was 93.8% White (n=46), 2.0% Black (n=1), and 4.1% Asian (n=2), with 0% (n=0) of participants reporting being Hispanic or Latino.

Table 5.2 Description of the Sample	
Characteristic	N (%), Mean (SD, min, max)
TOTAL included in analysis	49 (100.0%)
Age	73.67 (SD 5.10, min 65, max 88)
Gender	
Men	13 (26.5%)
Women	36 (73.5%)
Race	
White	46 (93.8%)
Asian	2 (4.1%)
Black	1 (2.0%)
Ethnicity	
Hispanic or Latino	0 (0%)
Not Hispanic or Latino	49 (100%)
Visual Acuity	Min 20/16, max 20/50
REALM R (Scores <6 are at risk for poor health literacy)	8 (min 2, max 8)*
Highest Education Level	
High School	13 (26.5%)
Associate Degree	11 (22.4%)
Bachelor's degree	14 (28.6%)
Master's degree	7 (14.3%)
Doctoral Degree	4 (8.2%)
Estimated Household Income	$51,124.5 (SD $29,093.42, min $1,229, max $150,000)
Native Language**	
English	46 (93.8%)
Slovak	1 (2.0%)
Chinese	2 (4.1%)
* Central tendencies followed by a single asterisk are modes, as the type of information being recorded did not lend itself to a mean or median ** Percentages don't add up to 100% because one participant did not report their native language.	

Active Ingredient Results

The first analysis of the cross-product comparison task centered on the accuracy of participants responses to questions that relied on AI information. Accuracy in AI question trials was the response variable, and the predictor variables were: highlighting (present or absent), Label Type (FOP or Standard), the interaction between Highlighting and Label Type), sex, education (dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

As the main effects in Table 5.3 illustrate, no evidence of any significant effects was detected when the accuracy of responses that required the AI information were analyzed. This was likely due to a ceiling effect; participants performed very accurately in the AI experimental tasks, with all four label treatments having accuracy rates between 99-100% (see Figure 5.3).

Table 5.3 Main Effects AI question type 2x2 model results (Accuracy)				
Effect	F	Df Numerator	Df Denominator	p
Highlighting	1.621	1	1559	0.203
Label Type (FOP/standard)	0.706	1	1559	0.401
Highlight X Label Type	0.706	1	1559	0.401
Sex	0.442	1	1559	0.506
Education	0.027	1	1559	0.869
Age	2.254	1	1559	0.133
Health Literacy (RealmR)	1.133	1	1559	0.287
Visual acuity	0.912	1	1559	0.340

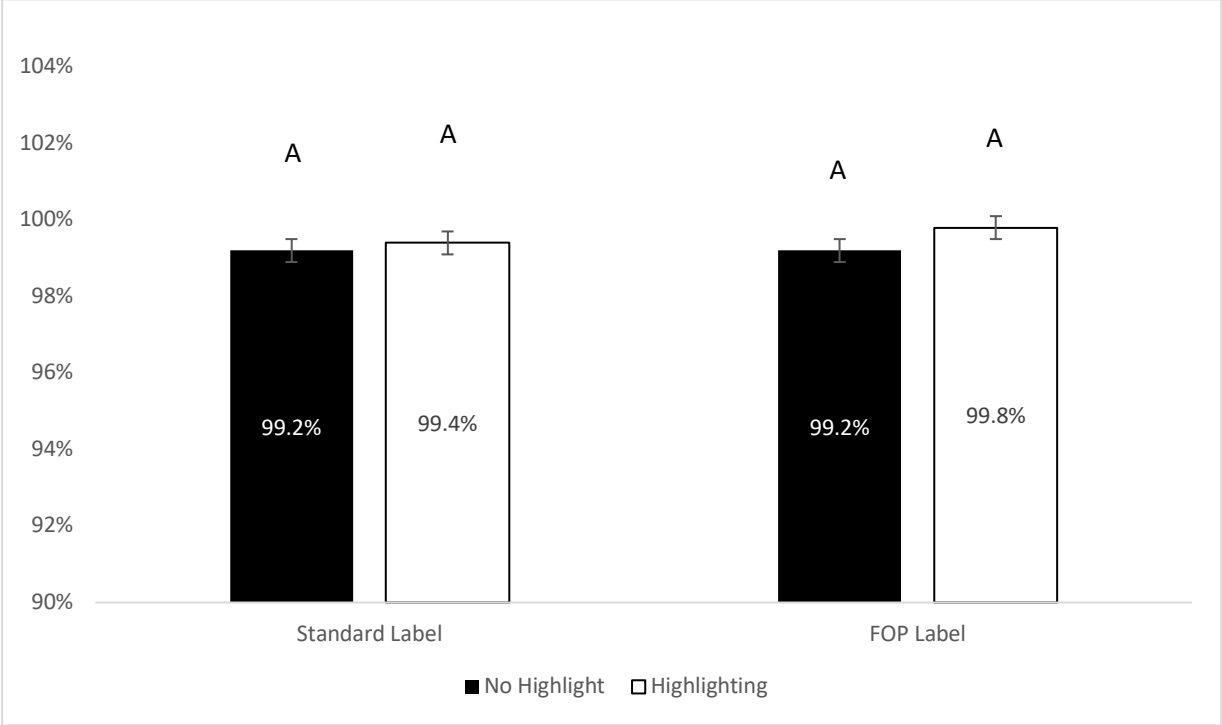


Figure 5.3 Estimated Marginal Means of accuracy for AI trials. No significant differences between the label treatments were detected in accuracy. Means with different letters signify statistically significant differences in a post-hoc means tests at a 95% confidence

The means of the accuracy rates for each label treatment are presented in figure 5.3. Like in the previous chapter, additional analysis was conducted to compare the standard label to the three labels with treatments to assess novel labels performance compared to current practice. Table 5.4 presents the results of this comparison of each label treatment to standard practice.

Table 5.4 AI question 4 cell model results to compare each cell against standard practice: Accuracy				
Effect	F	Df Numerator	Df Denominator	p
Four Label Treatments	0.614	3	1559	0.606
Sex	0.442	1	1559	0.506
Education	0.027	1	1559	0.869
Age	2.254	1	1559	0.133
Health Literacy (RealmR)	1.133	1	1559	0.287
Visual acuity	0.912	1	1559	0.340

The next analysis examined reaction time to correct product selection in trials with questions that relied on AI information. Reaction time to correctly answered AI question trials was the response variable, and the predictor variables were: highlighting (present or absent), Label Type (FOP or Standard), the interaction between Highlighting and Label Type, sex, education (dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

The analysis that focused on reaction times for correct responses related to questions that required accessing AI information indicated highlighting (p=0.001), sex (p=0.046), and Health Literacy (p= 0.024) all had a significant effect (see table 5.5). Specifically, mean reaction time was faster in the Highlight trials than in Non-highlight trials (Highlight ME = 3.686 log10 msec, se = .023; Non-highlight ME = 3.727 log10 msec, se = .023, p= 0.001); women performed quicker than men in this task (coefficient of -0.468, p=0.046), and the significant RealmR (included in the model as a continuous variable) effect indicated that people with higher levels of health literacy performed the cross product comparison significantly more quickly than people with lower levels of health literacy (coefficient of 0.0016, p=0.024).

For AI information, which appears in both the PDP and DFL but is separate from the FOP, the main effect of the label type was marginally significant (p=0.057). Specifically, the trend suggests that people tend to be faster on standard trials relative to FOP trials, (FOP ME = 3.719 log10 msec, SE = .023 and standard ME = 3.694 log10 msec, SE = .023 (p=0.057))(See table 5.5). While it is possible that this is because of the efficacy of the FOP at garnering attention, it reinforces the need to complete data collection to increase the sample size to the planned minimum of 60 participants, as more power could likely clarify this.

Table 5.5 Main Effects AI question type reaction time 2x2 model results (effects indicated to be statistically significant at α =0.5 are presented in bold

Effect	F	Df Numerator	Df Denominator	p
Highlighting	**10.26**	**1**	**1503**	**0.001**
Label Type (FOP/Standard)	3.62	1	1503	0.057
Highlight X Label Type	0.30	1	1503	0.583
Sex	**4.24**	**1**	**43**	**0.046**
Education	0.06	1	43	0.813
Age	3.42	1	43	0.071
Health Literacy (RealmR)	**5.49**	**1**	**43**	**0.024**
Visual acuity	0.49	1	43	0.489

Again, additional analysis was conducted to compare the standard label to our novel, three labels treatments to assess if there was improvement in reaction time for novel labels when compared to current practice. When comparing the label treatments to standard, commercial practice (No Highlight, standard label; ME = 3.718log10ms, SE=0.025) in a means tests using a Bonferroni correction, the reaction time for correct responses to questions which required AI information on Highlighted, standard label treatments differed significantly (ME= 3.670log10ms, SE=0.025, p=0.048). However, there was no evidence of a difference between the standard, commercial label when reaction time responses were compared to either the FOP without highlights (ME=3.736 log10ms, SE=0.025, p= 1) or the FOP with highlights (ME=3.701log10ms, SE=0.025, P=1) for the trials that required the AI information. In addition to the comparison against the standard label,

comparing the reaction times of correct responses requiring AI information yielded evidence of significance between the standard Highlight, ME = 3.670, SE=0.025 and FOP non-Highlight, ME = 3.736, SE=0.025 (p=0.002).

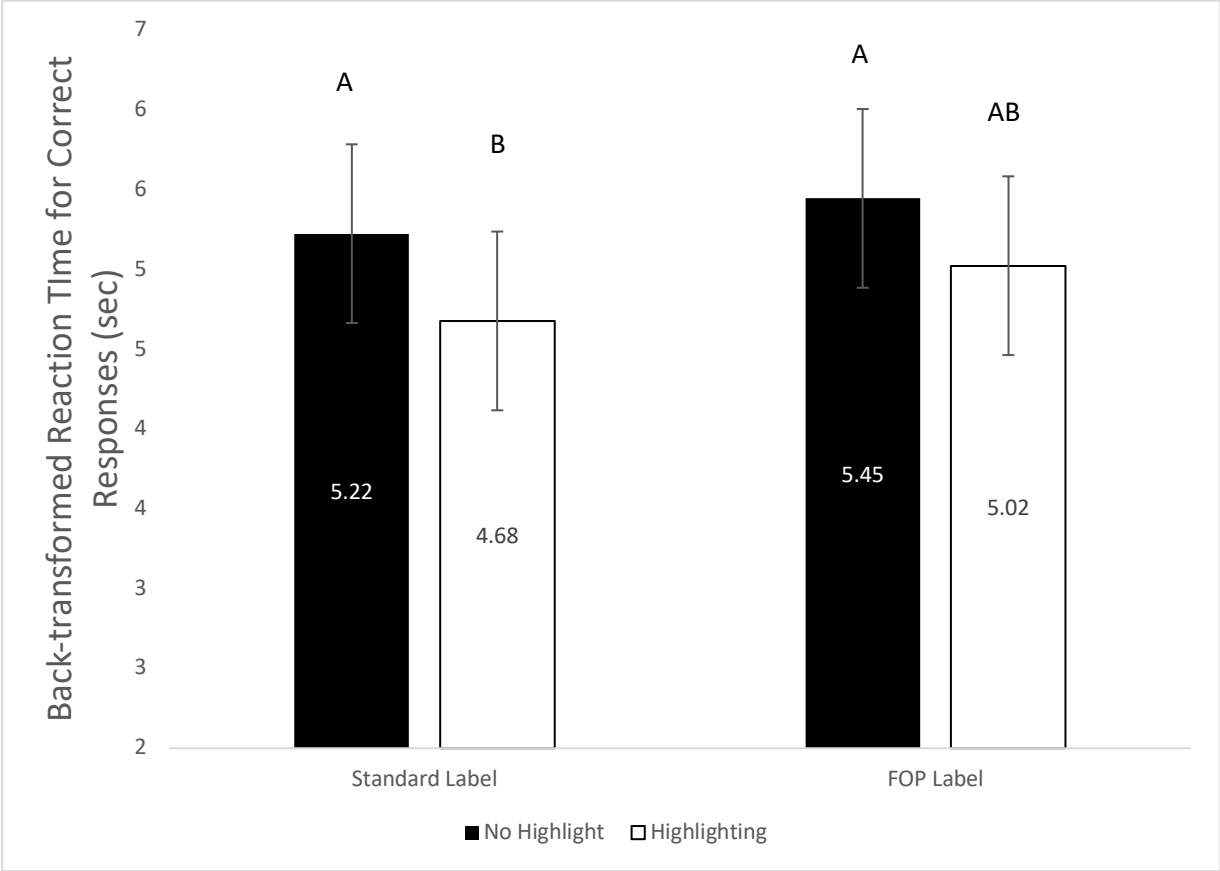


Figure 5.4 Estimated Mean Reaction time for correct responses. Results for AI trials only, comparing each treatment to standard practice (standard, no highlight). Means with different letters signify statistically significant differences in a post-hoc means tests at a 95% confidence

Trends in the data show that as health literacy increased, reaction time decreased. However, only 4 participants had a RealmR score less than 8; although in the expected direction, the significant effect Health Literacy (RealmR) should be interpreted with caution due to the limited sample available for inference. Additionally, trends in the data revealed more spread in the reaction time of male participants than female participants. However, the proportion of males to females in the sample is unbalanced, with approximately twice as many women as men. Purposeful sampling of both men and

at populations at risk for health literacy should be prioritized when data collection is finished to examine whether these preliminary results remain. These results are presented in table 5.4 and the back transformed estimated means are presented in figure 5.4.

Although the experiment yielded no evidence of an effect of labeling treatment on accuracy of response which required AI information (again, likely due to a ceiling effect), reaction time data do suggest highlighting to be a promising strategy for enhancing attention to the AI information, although readers are (again) cautioned that data collection is incomplete.

Drug-Drug and Drug-Diagnosis Interaction Warning Results

The subsequent analysis focused on accurate selection of product during a dichotomous choice for questions that relied on DD information. Accuracy in DD question trials was the response variable, and the predictor variables were: highlighting (present or absent), Label Type (FOP or Standard), the interaction between Highlighting and Label Type, sex, education (dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

When examining the selection accuracy of participants across trials that required accessing the critical DD information, a significant effect of highlighting ($p=0.03$), education ($p=0.026$) and of age ($p=0.008$) were noted. Accuracy was improved with highlighting, (Highlight was ME = 0.974, se = .007 and for Non-highlight ME = 0.952, se = .010; $p=0.013$), younger participants were more accurate than the older counterparts (coefficient of -0.110, $P=0.008$), and more educated participants were more accurate (coefficient of 0.448, $p=0.26$). These results are presented in table 5.6 and figure 5.5.

Table 5.6 Main Effects DD question type 2x2 model results: Accuracy				
Effect	F	Df Numerator	Df Denominator	p
Highlighting	**6.235**	**1**	**1559**	**0.013**
Label Type (FOP/standard)	0.423	1	1559	0.515
Highlight X Label Type	1.427	1	1559	0.232
Sex	3.210	1	1559	0.073
Education	**4.988**	**1**	**1559**	**0.026**
Age	**7.101**	**1**	**1559**	**0.008**
RealmR	3.057	1	1559	0.081
Visual acuity	0.068	1	1559	0.794

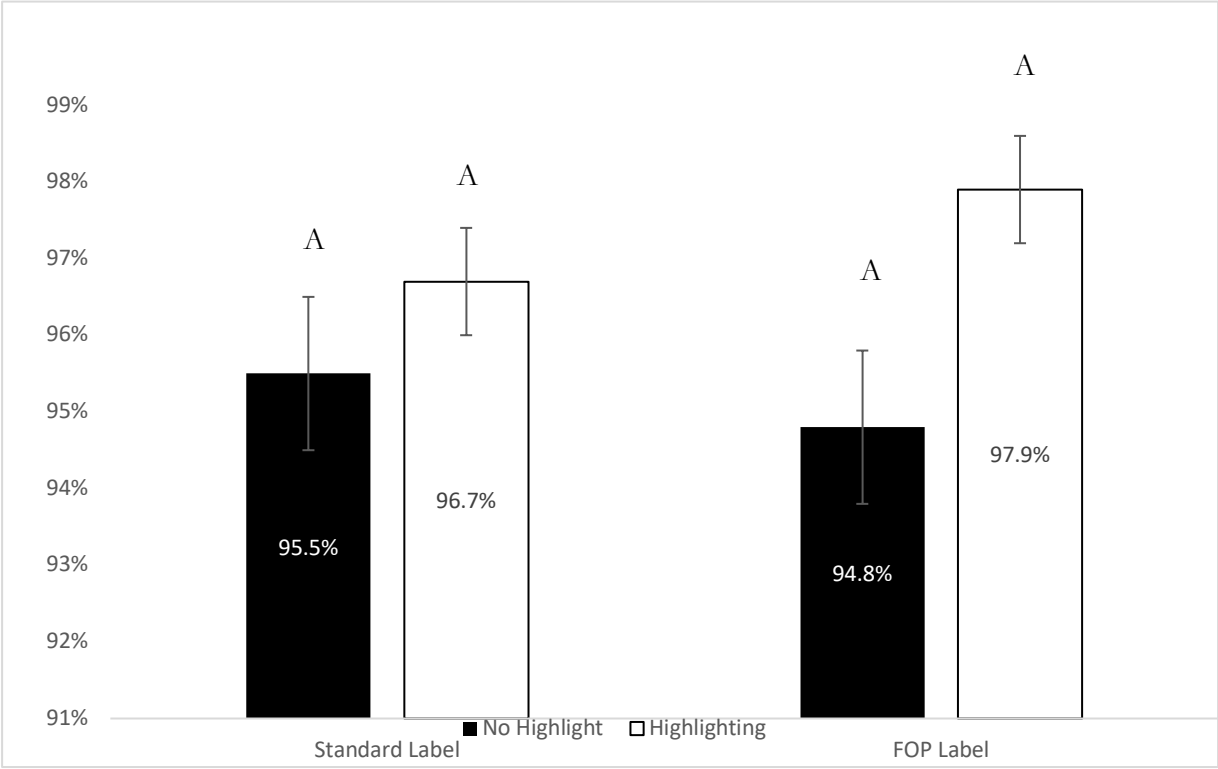


Figure 5.5 Estimated accuracy means for DD trials. No significant differences between the label treatments were detected in accuracy

Again, additional analysis was conducted to compare the standard label to the three novel treatments to assess if there was improvement in accuracy for trials that relied on DD information when novel labels were compared to current practice.

No evidence of a significant effect was detected at α=0.05 when the commercial standard was compared with novel treatments (standard x Non-highlight vs standard x Highlight; FOP x Non-

highlight; and FOP x Highlight). However, a marginal difference was noted (p=0.07) when the responses that required DD information in the trials with FOPs present in each highlighting format (highlighted and not) were compared (ME= 94.8%, SE=0.12 for non-highlighted) as well as FOP x Highlight (ME=97.9%, SE=0.007) . Like the AI Accuracy findings, this again reinforces the need to complete data collection to increase the sample size to the planned minimum of 60 participants, as more power will clarify whether highlighting critical information has a significant effect on participants selection accuracy when comparing products' warning information at α=0.05. The results of this test are presented below in table 5.7.

Table 5.7 DD question 4 cell model results to compare each cell against standard practice: Accuracy				
Effect	F	Df Numerator	Df Denominator	p
Four Label Treatments	2.358	3	1559	0.070
Sex	3.210	1	1559	0.073
Education	**4.988**	**1**	**1559**	**0.026**
Age	**7.101**	**1**	**1559**	**0.008**
Health Literacy (RealmR)	3.057	1	1559	0.081
Visual acuity	0.068	1	1559	0.794

The next analysis of forced choice task centered on the reaction time of participants correct responses to questions that relied on DD information. Reaction time to correctly answered DD question trials was the response variable, and the predictor variables were: highlighting (present or absent), Label Type (FOP or Standard), the interaction between Highlighting and Label Type, sex, education (dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

As with the analysis related to questions that were dependent on viewing the active ingredient (AI), the analysis of reaction time related to accurate selections requiring DD information also revealed a significant highlighting effect (p<0.000). Mean reaction time for Highlighted treatments was ME = 4.283 log10 msec, SE = .020 compared with for Non-highlighted treatments ME= 4.383 log10 msec, SE = .030. The results of this analysis are included below in table 5.8. None of the covariates presented

evidence of a significant effect in the performance in the DD trials when the dependent variable was reaction time to correct response.

Table 5.8 Main Effects DD question type 2x2 model results: Reaction Time				
Effect	F	Df Numerator	Df Denominator	p
Highlighting	**40.332**	**1**	**1410**	**0.000**
Label Type (FOP/standard)	1.005	1	1410	0.316
Highlight X Label Type	0.004	1	1410	0.947
Sex	0.456	1	43	0.503
Education	0.909	1	43	0.346
Age	1.432	1	43	0.238
Health Literacy (RealmR)	2.788	1	45	0.102
Visual acuity	0.324	1	43	0.572

As with the AI trials, post-hoc means tests using a Bonferroni correction were conducted to compare the average reaction times (to select the correct product) in trials requiring DD information for the standard, commercial label to the other treatments of interest to the research study. Analysis indicated that the standard commercial trials (No Highlight, standard label; M = 4.392, SE= 0.023) differed significantly from Highlighted standard (ME = 4.291, SE= 0.023 $p<0.000$) and from highlighted FOP (ME= 4.276, SE= 0.022 $p<0.000$) but not from the Non-highlight FOP (ME= 4.375, SE= 0.023). Figure 5.6 presents the back-transformed estimated means and Table 5.9 presents the result of this multiple comparison.

Table 5.9 DD question 4 cell model results to compare each cell against standard practice: Reaction Time				
Effect	F	Df Numerator	Df Denominator	p
Four Label Treatments	**13.827**	**3**	**1410**	**0.000**
Sex	0.456	1	43	0.503
Education	0.909	1	43	0.346
Age	1.432	1	43	0.238
Health Literacy (RealmR)	2.788	1	45	0.102
Visual acuity	0.324	1	43	0.572

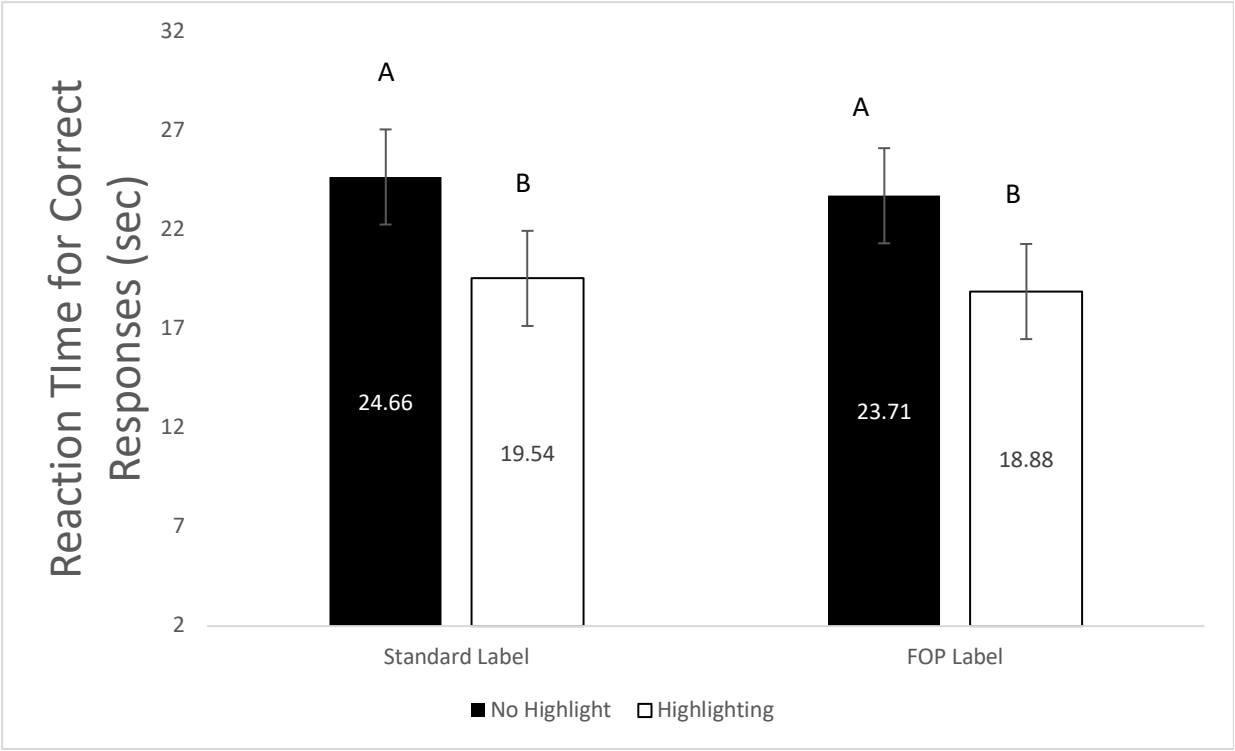


Figure 5.6 Reaction time results for DD trials only, comparing each treatment to standard practice (standard, no highlight) Means with different letters signify statistically significant differences in a post-hoc means tests using a Bonferroni correction at an alpha=0.05

Design Features as a Potential Distraction

We also investigated the notion that novel treatments (highlight and FOP) had the potential to be so effective at drawing viewer attention that they could actually divert it from other information present on the label (noncritical information). As such, in addition to analyzing the critical trials, distraction effect trials[11] were also analyzed to examine potential distracting effects of the FOP or highlighting when searching for nonprioritized information (i.e. information that was non-critical and therefore, not included in the FOP or highlighted).

[11] As detailed in the experimental design subsection of the methods section and figure 5.2, the experiment included both critical trials and distraction trials. In critical trials with the label optimization treatments, the information necessary to answer the question was either highlighted, included in the FOP or both highlighted AND in the FOP. In distraction trials, the FOP was present, however the information needed to answer the question was present only as nonhighlighted text in the DFL. The label treatments that appeared in distractor trials were either highlighted or nonhighlighted FOP labels.

Two comparisons were conducted for this purpose (figures 5.6 and 5.7). In the first, results from distraction effect trials[11], which required non-critical information (specifically, uses) were compared to the results obtained from standard trials without an FOP which required DD information. This included both standard labels and highlight standard labels (no FOP present) in order to compare searching for non-prioritized information in the DFL in the presence of an FOP (uses- the distractors) compared to searching for information in the DFL without the presence of an FOP (drug warning information). This comparison yields insight into if the FOP labels hinder information search relative to standard practice. Illustration of this comparison is in figure 5.6.

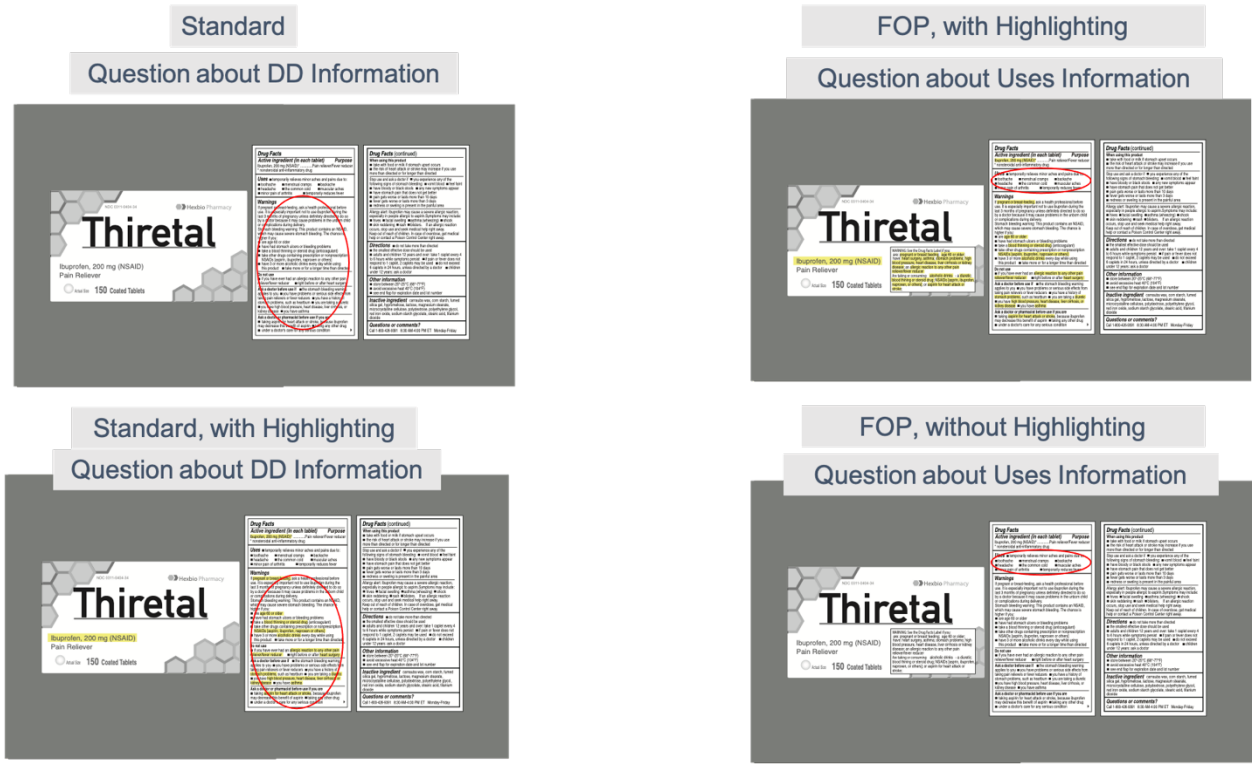


Figure 5.7 Standard label, DD information compared to Distractor trial. The red circles indicate the location of the information which would have been used to respond to the question

The second comparison was with DD information type, FOP labels, to examine the difference between speed and accuracy when FOP label optimization strategies are present and the information is both relevant (DD information) and irrelevant (Uses information) to the question being asked.

Comparison of the dependent variables (accuracy and time to correct selection) by information type (relevant verses irrelevant) yield insights on the potential for accentuating critical information to "distract" from other elements of the label (in this case uses). Theoretically, if the critical information outperforms the non-critical, which is located on the same panel and roughly the same size, it would suggest that the accentuating factors distract from those that don't carry the same emphasis. Illustration of this comparison is in figure 5.7.

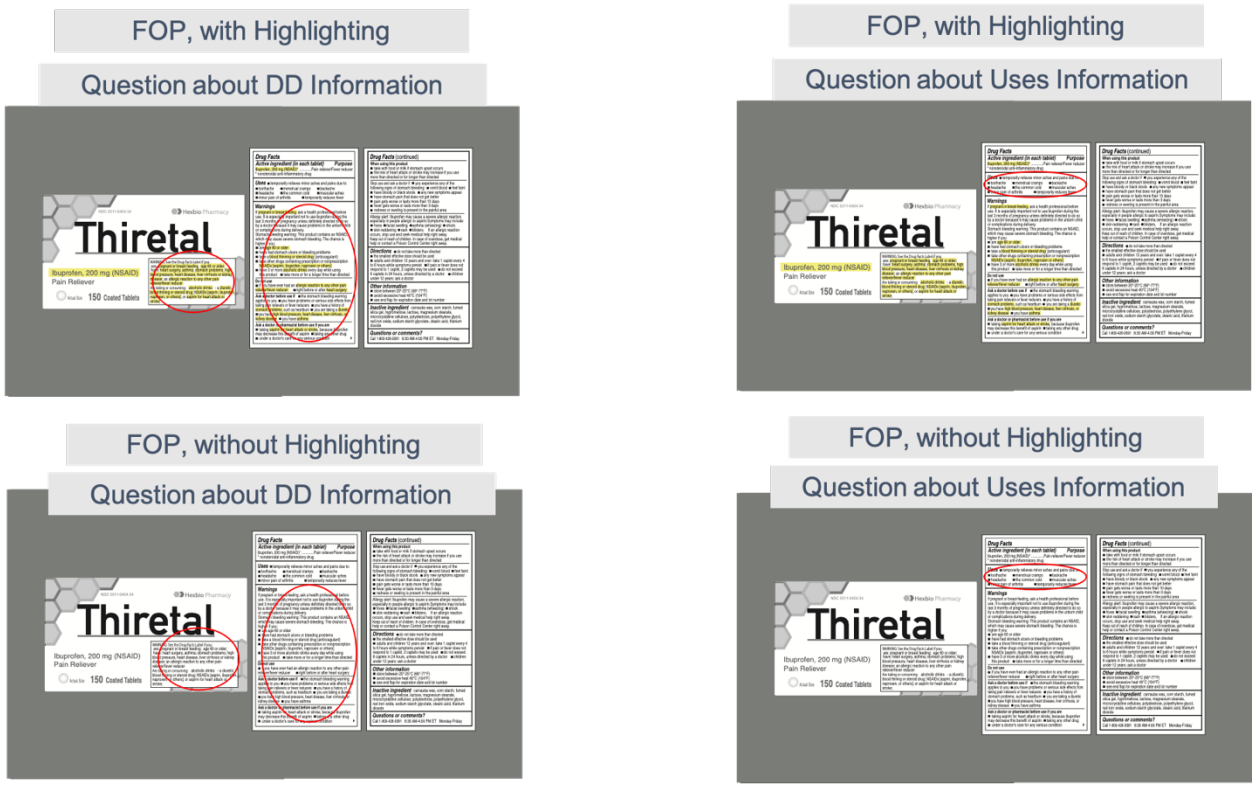


Figure 5.8 FOP label, DD information compared to Distractor trial. The red circles indicate the location of the information which would have been used to respond to the question

Comparison against Standard Practice

The first analysis of the distraction trials compared the accuracy of participants responses in both standard label DD question trials and both distractor trials. Accuracy was the response variable, and the predictor variables were: highlighting (present or absent), distraction effect (distractor trial verses DD Standard trial), the interaction between Highlighting and distraction effect, sex, education

(dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

In this accuracy analysis of the distraction effect trials, there was a main effect of distraction and visual acuity, but no interactions with highlighting. The main effect for distraction effect indicates that people were significantly more accurate in DD-standard trials than in trials that relied on the uses information (distraction effect) (DD-standard ME =0.959, SE =0.011 and Distraction ME=0.881, SE=0.023). People with worse visual acuity performed less accurately (coefficient=-0.094). These results are included below in tables 5.10 and 5.11.

Table 5.10 Accuracy Main Effects 2x2 model results: from analysis comparing DD standard question type with distractor trials

Effect	F	Df Numerator	Df Denominator	p
Highlighting	0.101	1	1167	0.751
Distraction effect	**22.828**	**1**	**1167**	**0.000**
Highlight X Distraction effect	1.240	1	1167	0.266
Sex	1.446	1	1167	0.229
Education	0.673	1	1167	0.412
Age	1.400	1	1167	0.237
Health Literacy (RealmR)	0.101	1	1167	0.751
Visual acuity	**9.906**	**1**	**1167**	**0.002**

Table 5.11 Accuracy Estimates

Treatment	ME	SE
Highlight distraction effect	0.871	.029
Highlight DD standard	0.965	.012
Non-highlight distraction effect	0.890	.027
Non-highlight DD standard	0.952	.014

The second analysis of the distraction trials was an assessment of reaction time to correctly select the product for DD standard question trials and questions that required uses information (distraction effect trials). Reaction Time to correct product selection was the response variable, and the predictor variables were: highlighting (present or absent), distraction effect (distractor trial verses

DD Standard trial), the interaction between Highlighting and distraction effect, sex, education (dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

In the reaction time analysis of trials requiring uses information, there was a significant distraction effect main effect; specifically, these trials were significantly faster than the DD information type standard x Non-highlight trials (Distraction ME = 4.070 log10 msec. SE = .023 and DD-standard ME = 4.342 log10 msec, SE=.020, $p<0.000$). A significant interaction ($p<0.000$) was identified between the distraction effect variable and the highlight variable (see Table 5.9 and 5.10). Exploration of the interaction suggests that the difference between distraction effect and DD-standard is larger in the non-highlighted trials than in the Highlight trials. These results indicate that the presence of the FOP or Highlight do not significantly penalize information not included in FOP or highlighted more than the standard practice of prioritizing no information. These results are presented below in tables 5.12 and 5.13

Table 5.12 Main Effects DD question type 2x2 model results: Reaction Time

Effect	F	Df Numerator	Df Denominator	p
Highlighting	0.089	1	1013	0.765
Distraction effect	**201.352**	**1**	**1018**	**0.000**
Highlight X Distraction effect	**25.057**	**1**	**1013**	**0.000**
Sex	0.373	1	43	0.545
Education	0.217	1	43	0.643
Age	1.353	1	43	0.251
Health Literacy (RealmR)	2.725	1	44	0.106
Visual acuity	0.091	1	43	0.765

Table 5.13 Reaction Time Estimated Marginal Means for Correct Responses

Treatment	Mean Estimates (log10msec)	SE
Highlight distraction effect	4.115	.028
Highlighted standard label DD question	4.292	.023
Non-highlight distraction effect	4.025	.028
Non-highlighted standard label DD question	4.393	.023

Comparison against FOP Labels with Relevant, Question related Information

The next analysis of the distraction trials was an assessment of accuracy of participants responses in both DD FOP question trials and distractor trials. Accuracy was the response variable, and the predictor variables were: highlighting (present or absent), distraction effect (distractor trial (uses information) verses FOP trial(DD information)), the interaction between Highlighting and distraction effect, sex, education (dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

In this second accuracy analysis of the distractor trials, there was a main effect of distractor ($F=56.530$, $p<0.001$) and visual acuity ($F=14.534$, $p<0.001$), and an interaction with distractor and highlighting ($F=5.644$, $p=0.018$) (see table 5.14). This significant interaction is explored in table 5.15 below. This interaction indicates that people were significantly more accurate trials with prioritization of relevant information (DD information questions with FOP) than in distractor trials without prioritization of relevant information (uses information questions with FOP). Additionally, highlighting enhanced performance for FOP trials with DD information, while hindering performance for distractor trials (DD-FOP, Highlighted mean = .988, SE= .005, DD-FOP, non-Highlighted mean = .968, SE= .008 and Distractor non-highlighted Mean = .888, SE= .025, Distractor highlighted Mean = .869, SE= .027). People with worse visual acuity performed worse on these trials.

Table 5.14 Accuracy Main Effects comparing FOP Label DD question type and distractors 2x2 model results

Effect	F	Df Numerator	Df Denominator	p
Highlighting	2.624	1	1951	0.105
Distraction effect vs DD-standard	**56.530**	**1**	1951	**0.000**
Highlight X Distraction effect vs DD-standard	**5.644**	**1**	1951	**0.018**
Sex	0.280	1	1951	0.597
Education	1.856	1	1951	0.173
Age	1.025	1	1951	0.312
Health Literacy (RealmR)	1.282	1	1951	0.258
Visual acuity	**14.534**	**1**	**1951**	**0.000**

Table 5.15 Accuracy Estimated Marginal Means for distraction analysis when distraction trials are compared to trials with FOP labels and DD information		
Treatment	Mean Estimates	SE
Highlight distraction effect	.869	.027
Highlighted DD FOP question	.988	.005
Non-highlight distraction effect	.888	.025
Non-highlighted DD FOP question	.968	.008

The final analysis of the distraction trials was an assessment of reaction time of participants responses. Reaction Time was the response variable, and the predictor variables were: highlighting (present or absent), distraction effect (distractor trial verses DD FOP trial), the interaction between Highlighting and distraction effect, sex, education (dichotomized into some college or more versus high school or less), age, health literacy (Realm-R score value), and visual acuity.

Table 5.16 Main Effects comparing FOP Label DD question type and distractors 2x2 model results: Reaction Time				
Effect	Df Numerator	Df Denominator	F	p
Highlighting	1	1013	0.483	0.487
Distraction effect	**1**	**1018**	**6.297**	**0.012**
Highlight X Distraction effect	**1**	**1013**	**9.567**	**0.002**
Sex	1	43	0.800	0.376
Education	1	43	0.008	0.928
Age	1	43	1.707	0.198
Health Literacy (RealmR)	1	44	3.149	0.083
Visual acuity	1	43	0.218	0.643

In this reaction time analysis of distractor trials, there was a significant distractor main effect ($F=6.297$, $p=0.012$). A significant interaction ($F=9.567$, $p=0.002$) (see table 5.16) was identified between the distractor variable and the highlight variable. Exploration of the interaction (table 5.17) suggests that the difference between distractor and DD-FOP is bigger in the non-highlighted trials than in the highlight trials. These results suggest that highlighting hinders performance on the distraction trials while it enhances performance on the critical warning trials. However, in distraction

trials, when the highlighting was not present, participants were quicker to respond than on critical trials with an FOP but without highlighting.

Table 5.17 Time to Correct Response Estimated Marginal Means for distraction analysis when distraction trials are compared to trials with FOP labels and DD information		
Treatment	M log10 msec	se
Highlight distractor	4.117	.035
Highlight DD FOP	3.984	.022
Non-highlight distractor	4.028	.035
Non-highlight DD FOP	4.041	.023

In conclusion, there is a potential for the addition of an FOP and highlighting of critical information to distract from other information in the DFL. However, the distraction effect of the FOP is minimal when compared to the benefits of improving the standard practice in OTC labeling so that more attention is given to critical information. While participants were less 120accurate in distraction trials compared to DD questions with standard labels, participants responded more quickly to distraction questions than standard label DD questions without highlighting. The results of this distraction analysis emphasize the importance of the process of distinguishing which information presented in the DFL is to be highlighted or prioritized in the FOP.

Discussion and Implications

To reiterate, because of the preliminary nature of this analysis due to COVID-19 human subject data collection restrictions, readers are cautioned that reported results are likely underpowered and subject to change. Future analysis upon completing data collection will be forthcoming. This discussion begins by revisiting the scientific hypothesis (Gotelli & Ellison, 2004) of this experiment in light of the results:

Hypothesis 1: *Highlighting will increase the accuracy of participants choosing between the labels compared to non-highlighted labels*. This hypothesis was partially supported by the main effect of highlighting for DD accuracy results, however this finding did not generalize to the active ingredient information. Thus,

finishing data collection with the full sample size is necessary to establish whether or not the lack of a significant effect of highlighting on active ingredient cross product comparison is an artifact of a lack of power.

Hypothesis 2: *The front of pack label will increase the accuracy of choosing between the labels compared to labels without the front of pack label.* There was no evidence that the front of pack label increased accuracy in the cross-product comparison task. Overall, participants were extremely accurate in this task, leaving little margin for improvements in accuracy.

Hypothesis 3: *Highlighting will decrease the amount of time required for participants to choose between the labels compared to non-highlighted labels.* Both AI and DD results demonstrated a beneficial effect of highlighting on reaction time for correct responses. Highlighting significantly improved the processing of AI and DD information compared to trials with non-highlighted labels.

Hypothesis 4: *The front of pack label will decrease the amount of time required for participants to choose between the labels compared to labels without the front of pack label.* There was marginal evidence that the front of pack label increased accuracy for the AI information type in the cross-product comparison task. Finishing data collection with the full sample size is necessary to establish whether or not the marginal front of package benefit on active ingredient cross product comparison is an artifact of a lack of power.

Current results suggest that enhancing the standard OTC label with highlighting potentially improves older adults' ability to conduct cross-product comparisons of critical health information presented on OTC medication labels. Highlighting proved to be beneficial in facilitating cross-product comparison for both types of critical information (AI and DD), suggesting that highlighting of critical information is a productive design path to pursue. While there was not enough power to detect any significant effects related to the presence of FOP label designs in this preliminary analysis, data trends suggest that the FOP, in tandem with highlighting, might be useful particularly for trials that required

the critical DD information. Additionally, as gender was a significant covariate in DD reaction time analysis, data collection should be completed with a goal of recruiting more men to balance the sample and determine whether or not the effect of gender remains with a balanced sample.

The results of this assessment of the top-down processing of OTC medication labels by older adults will be discussed further in Chapter 6 in tandem with the results of the change detection (bottom-up) and absolute judgment (yes/no) task (top-down). Chapter 6 will discuss the results of all three experiments, as each focus on a specific attribute that is necessary to fully determine an optimized OTC label design. The results of this study, in combination with the results of the two previously reported studies, yield evidence which can be used to make recommendations about OTC label designs that are likely to enhance information processing for older consumers who are known to be at risk from the ill effects of adverse drug reactions (ADRs). Specially, we have investigated how varied designs impact how: consumers notice critical health information on an OTC, facilitate consumer search for critical health information, and enhance consumers' ability to conduct cross-product comparison of OTC products to select the most appropriate option.

Chapter 6 Discussion

Implications

OTC medication packaging serves a vital role in facilitating safe and effective use of drugs by older adults. The primary implication of these studies is that there is room to continue improve the labeling of OTCs to better communicate OTC risks to older consumers. These findings suggest that the strategy of requiring highlighting mandated by the FDA in CFR 21§201.326 to improve communication of the risks of acetaminophen and NSAIDs could be expanded to better emphasize active ingredient information in the DFL in general. Research by Goyal et al. evaluating the highlighting of the active ingredient acetaminophen, and the addition of product specific organ warnings (for acetaminophen, specifically a liver warning) improved risk perception of consumers(Goyal, Rajan, Essien, & Sansgiry, 2012). If this strategy is widely implemented, the benefit of highlighting increasing risk perception, in addition to increasing the usability of the DFL, would help older consumers better understand the risks associated with OTC medication use. Additionally, these results suggest that the FOP strategy found to be useful when communicating nutritional information is worth exploring in other product categories where easily understood information could increase consumer's ability to compare health and safety information more efficiently.

Packaging and labeling are powerful tools in communicating with consumers both in the retail environment at time of purchase, and at time of use. In the era of the COVID-19 pandemic, the retailing of OTC drug products is changing alongside retailing of food and other consumer goods with accelerated shifts to e-commerce (Bhargava et al., 2020). Consumer behavior changes in response to the changing retail environment can influence what face of a package consumers interact with before purchase, as consumers only have access to the product information retailers include on the ecommerce platform, not necessarily all faces of a package. Especially as OTCs are one of two product categories in which consumers intend to continue to purchase online (Bhargava et al., 2020), this rapid

switch to online retailing of OTCs should be included in examinations of the DFL's effectiveness. This change in what label information is easy to access, both in brick and mortar stores due to sanitation concerns and on ecommerce platforms, only reinforces the importance of the label drawing attention to critical warnings. The implication of these findings that simple updates to the PDP and DFL can improve the likelihood that older adults notice critical information and comprehend that critical safety information should compel the FDA to routinely examine whether or not the labeling status quo is doing enough to ensure consumer safety.

Finally, the results presented in chapter 4 indicate that there are significant differences between older consumers familiarity with brand names and active ingredients of OTC medications. Healthcare professionals working with this at-risk population should not assume that communicating the risks of a specific OTC to a patient by active ingredient will be as effective as communicating those risks incorporating both the active ingredient and branding information

Review of Research Questions, Objectives, and Results

The three experiments presented in this dissertation were guided by the following research questions: 1. What is the effect of highlighted OTC label formatting on attracting attention to critical information both when it is and is not the explicit goal of the patient? 2. What is the effect of moving critical information to the PDP of OTC packaging on attracting attention both when it is and is not the explicit goal of the patient? 3. What is the combined effect of both moving critical information to the PDP of OTC packaging and highlighting critical information on attracting attention both when it is and is not the explicit goal of the patient? As whether or not accessing the information is part of the participants goal is related to whether top-down or bottom-up attentional processes are being utilized, the research questions were all addressed by the summation of the three experiments presented in this dissertation. The overall goal of these studies was to provide benchmarking for these

novel labeling strategies and determine a single optimized labeling format to be evaluated in a more ecologically valid manner.

In review, the results from the change detection study indicate highlighting was an effective strategy for attracting attention to information in the DFL in both AI and DD type trials, and the FOP was found to be effective at attracting attention in DD trials. The absolute judgement and dichotomous forced choice tasks also indicated that highlighting was helpful for facilitating participants' use of drug warning information for both single product and cross product comparison tasks, and that the FOP helped facilitate use of active ingredient information in the single product absolute judgement task. The mechanism by which the FOP facilitated use of the active ingredient information is unclear, as the active ingredient did not appear in the FOP, but accuracy in AI trials did increase with the presence of the FOP. The results of these experiments are presented in full in their corresponding chapters.

Discussion of Results in Context of Theory

When we revisit the results of the three studies presented in this dissertation with the context of usability and the Human Package Interaction Model presented in depth in Chapter 2, the results of the studies can be better interpreted. The studies presented in this dissertation were aimed at the perception stage of information processing as well as the comprehension stage of information processing (see chapter 2, table 2.1). The change detection study investigated the allocation of attention of critical label information in different labeling formats (Bix et al., 2010). The allocation of attention to critical information on an OTC label is directly linked to the likelihood that a consumer of that product will access that information without searching for it specifically. Thus, promoting a label format that improves the relative prominence of the information, such as the FOP or HL strategies that were found to be more noticeable (see Chapter 3) would increase early stages of information processing (i.e. attention).

The absolute judgement and dichotomous forced choice tasks investigated the efficacy of the same four labeling strategies in the later stages of information processing, comprehension (understanding the message of the label) and action (selecting the correct response to the prompt)(see chapter 2, table 2.1). This pair of studies found highlighting to be useful for facilitating later stage processing actions, such as making cross-product comparison which required critical information (AI and DD) or in an absolute judgment task which required them to answer a product specific question related to the critical AI or DD information while viewing a single product. Additionally, these results can be interpreted as the presence of label optimizing strategies improving the speed at which these actions occurred (see Chapters 4 and 5).

In addition to the lens of the Human Package Interaction Model, a novel contribution to the field of OTC labeling assessment was the application of usability framework to OTC medication labels. The effectiveness and efficiency components of usability were also examined in this dissertation (See table 2.2). Response time results, the variable representing efficiency, from each of the studies support the label format of HL FOP. While the accuracy results, or the variable representing effectiveness, are less differentiated between the label formats in the top-down studies (Chapters 4 and 5) as participants were overwhelmingly accurate in their interpretations of the labeling information, the accuracy results from the bottom-up study (Chapter 3) also support HL and FOP strategies as methods for improving the usability of OTC medication labels.

Justification of The Selected Optimized Label Format- The Highlighted DFL and FOP

The primary objective of this dissertation was to identify an optimized label format that attracts attention to critical information whether or not accessing that information is a participant's goal. The optimized label format has been identified to be the Highlight x FOP label that combines both the highlighting strategy and the FOP strategy for improved communication. This determination was made after analyzing the results of the 3 studies included in this study in tandem. This labeling

format was selected based on the following evidence from the change detection study; HL was found to be effective in attracting attention to changes in the DFL for both AI and DD information and the FOP was found to be effective at attracting attention for DD information. Evidence for the Highlight x FOP format from study presented in Chapter 4 is that the FOP was found to be effective for AI information and highlighting was found to be effective for DD information, especially when the HL was paired with an FOP. Finale evidence from the preliminary results from study presented in Chapter 5 suggest HL facilitates comparison between two products' AI and DD information.

Table 6.1 Summary of evidence for Highlighted FOP Label	
Highlighting Evidence	• Experiment 1: Highlighting is effective in attracting attention to changes in the DFL for both AI and DD information Experiment 3: HL facilitates comparison between two products' AI and DD information
FOP Evidence	• Experiment 1: the FOP is effective at attracting attention for DD information • Experiment 2: the FOP is effective for facilitating use of AI information
Highlighting and FOP Evidence	• Experiment 2: the combination of highlighting and an FOP is effective for DD information.

The findings of the study presented in Chapter 4 in support of the Highlight x FOP strategy are especially promising as the results suggest that the presence of the FOP on the PDP does not detract from the critical AI information, despite providing additional information to process in limited space. It appears that the presence of the FOP improved consumers' ability to search for the AI information, though the mechanism of how the presence of the FOP influenced information search

is unknown and should be further explored with an experiment utilizing an eye-tracking methodology to examine scan paths.

Limitations

While the results of the studies presented in this dissertation are promising, there are also inherent limitations to the work. The first limitation is the mock branding used in all three of the presented studies. Because no real brands were used and we controlled for potential color effects by using grey scale images, generalizability to the broader OTC market is limited. As branding is known to influence how consumers perceive medications (Fraeyman, 2015; Halme, Linden, & Kääriä, 2009), these methods should be replicated with real brands to examine whether or not the results hold. Additionally, the mock brands that were utilized were grayscale, which could have increased the relative visual salience of the highlighting, biasing the results (Milosavljevic, Navalpakkam, Koch, & Rangel, 2012). Repeating these studies with branded stimuli would provide insight into whether or not the benefits of highlighting remain when the highlighting is not the only non-grayscale component of the stimuli.

The second and third limitations are due to researcher error. First, a small programming error in the change detection study limited the ability to answer some research questions to the fullest extent possible due to unbalanced occurrence of the labeling treatments. This error was described in detail in chapter 3, starting on page 47. Secondly, distraction effect trials were not included in the first dichotomous Yes/No forced choice task until half of the recruited participants had completed the study, as detailed in chapter 4 starting on page 66. Finally, the sampled participants included in this study featured limited diversity due to the difficulty in recruiting people of color in general, and men of any background (see tables 3.1, 4.2 and 5.2). As gender has been found to be linked to risk awareness of OTC drug consumers (Calamusa et al., 2012), this limitation in the sample could be leading to a

population that is more likely to perceive OTCs as risky, potentially influencing their behavior in the direction of caution.

Conclusions

In conclusion, the work presented in this dissertation supports further investigation of the labeling format of HL and the FOP for use in OTC medication labeling. These investigations of labeling format optimization strategies and the effect of those formats on participant performance on tasks involving both top-down and bottom-up processing of OTC labels have provided necessary evidence to support more the need for more ecologically valid research on consumer's use of optimized OTC labels, with the goal of reducing the prevalence of ADRs associated with OTC medication.

Suggestions for Further Study

One question opened by the results of these studies is: why does the FOP facilitate attention to the active ingredient information that is not presented within the FOP? To address that question, future studies should investigate the effects of FOP style and placement on the facilitation of attention to the active ingredient and warning information using an *eye tracking methodology to investigate scan paths and order of information access*. Future research should include measures *of participants' inherent risk perception of OTC medication*, as their perceptions of the likelihood of an adverse reaction is likely to influence their OTC usage behavior (Hoy & Levenshus, 2018) both inside and outside of the laboratory setting. Developing an understanding of participants' risk perception of OTCs and whether or not that influences their behavior with OTC labels in the laboratory setting could provide more evidence for best labeling practice. To further understand the extent to which the FOP labeling strategy improves consumer understanding of risks associated with OTC medication and facilitates safer decision making, researchers should do the following: *test with real brands*, include an evaluation of participants' risk perception of OTC medication, test with consumers making evaluations for themselves, *test within*

a product category to make it more realistic and then see if this optimization is relevant globally beyond the United States' DFL. Applying the same objective measures utilized in this study of mock-branded products to real brands would allow researchers to compare what the beneficial effect of the Highlight x FOP label is when consumers have access to the familiar information they profess to use when making decisions (Harben et al., 2018).

Furthermore, there should be an investigation into how well this optimized labeling format performs when participants are tasked with evaluating a medication's safety for themselves, including their own health history and medical concerns in the decision-making process. The studies in this dissertation provided crucial benchmarking for the optimized labeling strategy, but without an assessment of whether or not the labeling format improves participant's application of the labeling information, it is unknown the true extent to which this format could improve public health.

Additionally, the cross-product comparisons utilized in the study presented in Chapter 5 were cross-category product comparisons. To better simulate the decision making of consumers in a retail environment, a cross-product comparison of products within a category should be considered. Requiring consumers to choose between two analgesics, two antihistamines, or two antacids would better replicate the types of decisions consumers of OTCs make regularly. As there are some active ingredients in product categories that are safer for older adults than others (Fick et al., 2019), this would also provide valuable insight into how to better communicate risks that increase with age: such as the risk of stomach bleeding, or the risk of an anticholinergic effect.

Finally, as the retail environment for medication is shifting to include more ecommerce as both a compliment and substitute for in person shopping, evaluating how the PDP of a label is displayed online and developing standards as to how OTC medication information is communicated to consumers online is an area urgently in need of regulatory action. Adding these evaluations of the label's benefits would afford external validity and allow broader recommendations to regulators. The

approach presented within this dissertation should be broadened to other standardized OTC labeling formats used globally to see if the strategy of adding HL and FOP to other formats improves understanding of the label content in a global context.

APPENDICES

APPENDIX A:

Examples of each active ingredient label used in the Change Detection Study, in all four treatments along with the corresponding critical changes

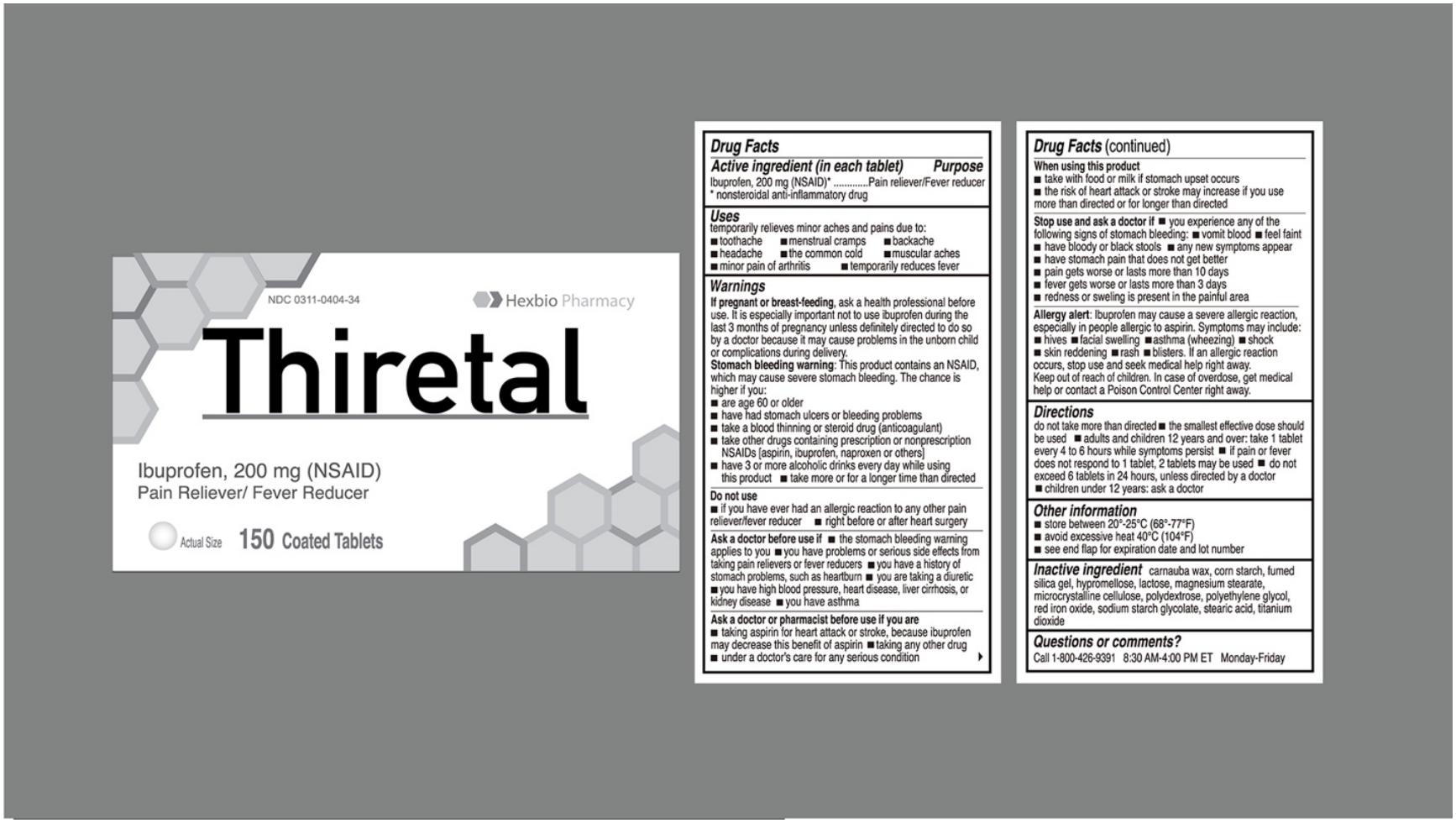


Figure A.1 Example Standard Label for Ibuprofen, with neither the FOP or Highlight Treatment

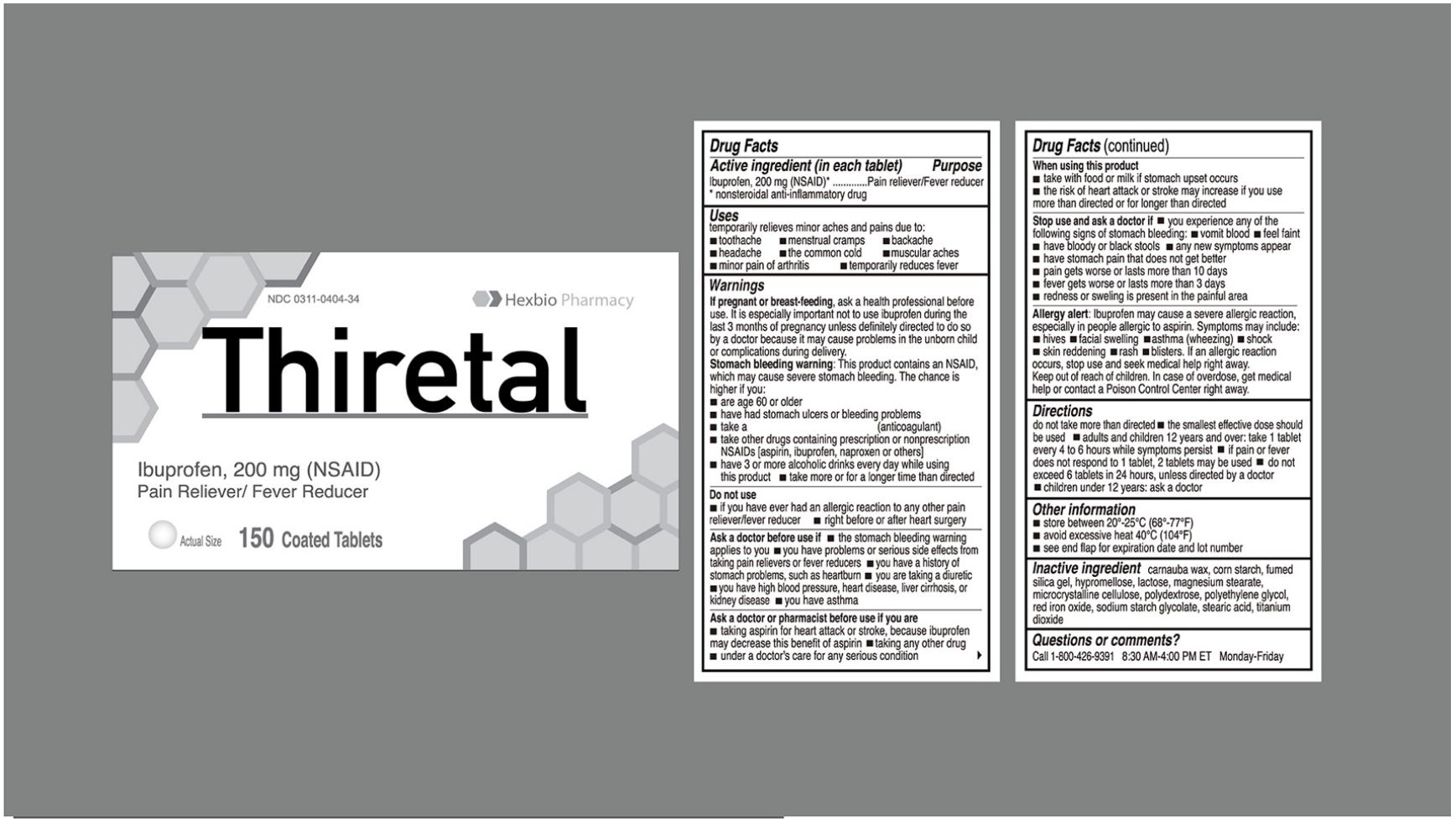


Figure A.2 DD Change in DFL

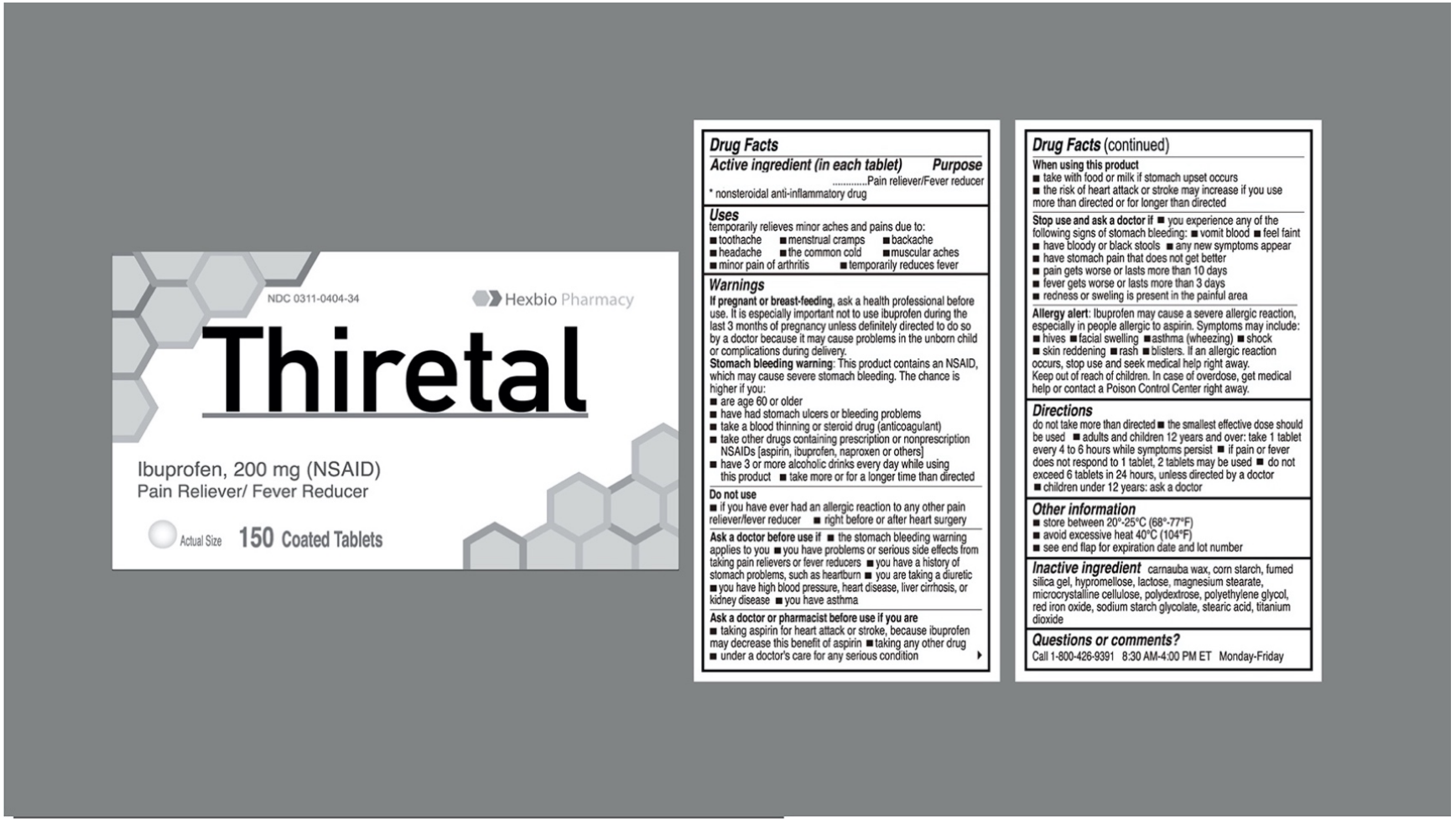


Figure A.3 AI Change in DFL

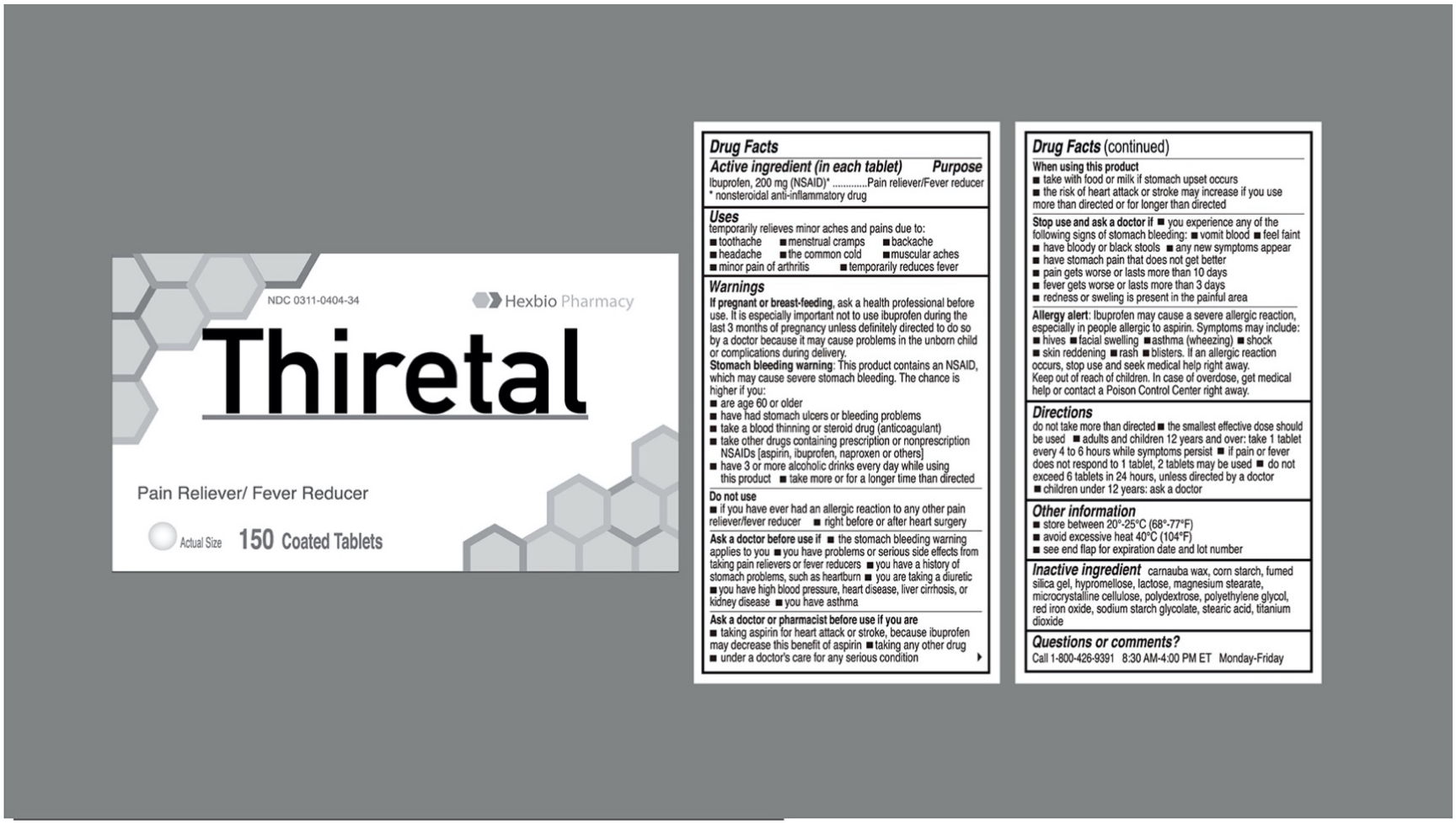


Figure A.4 AI Change on PDP

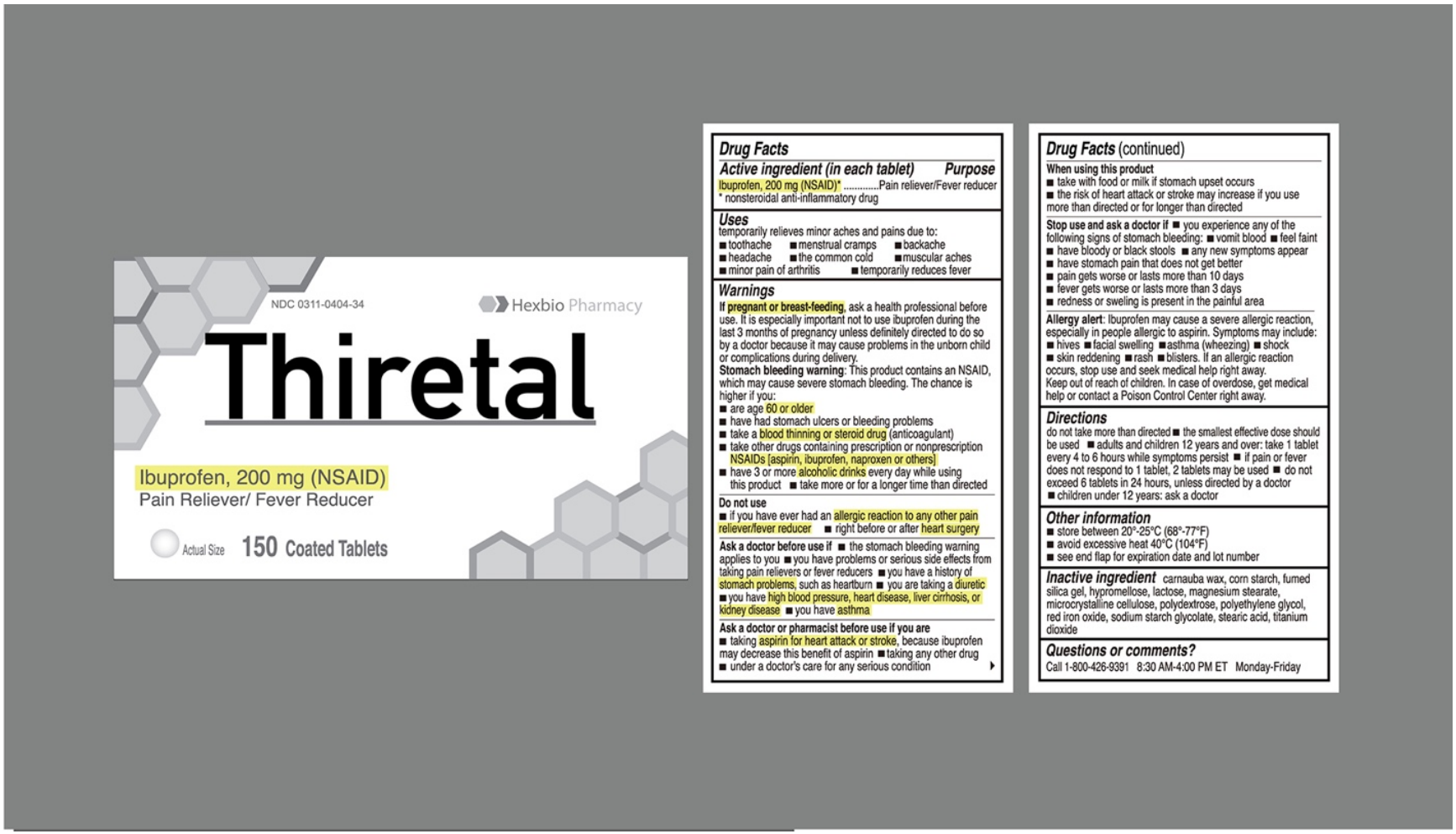


Figure A.5 Example Label for Ibuprofen, enhanced with the Highlight Treatment

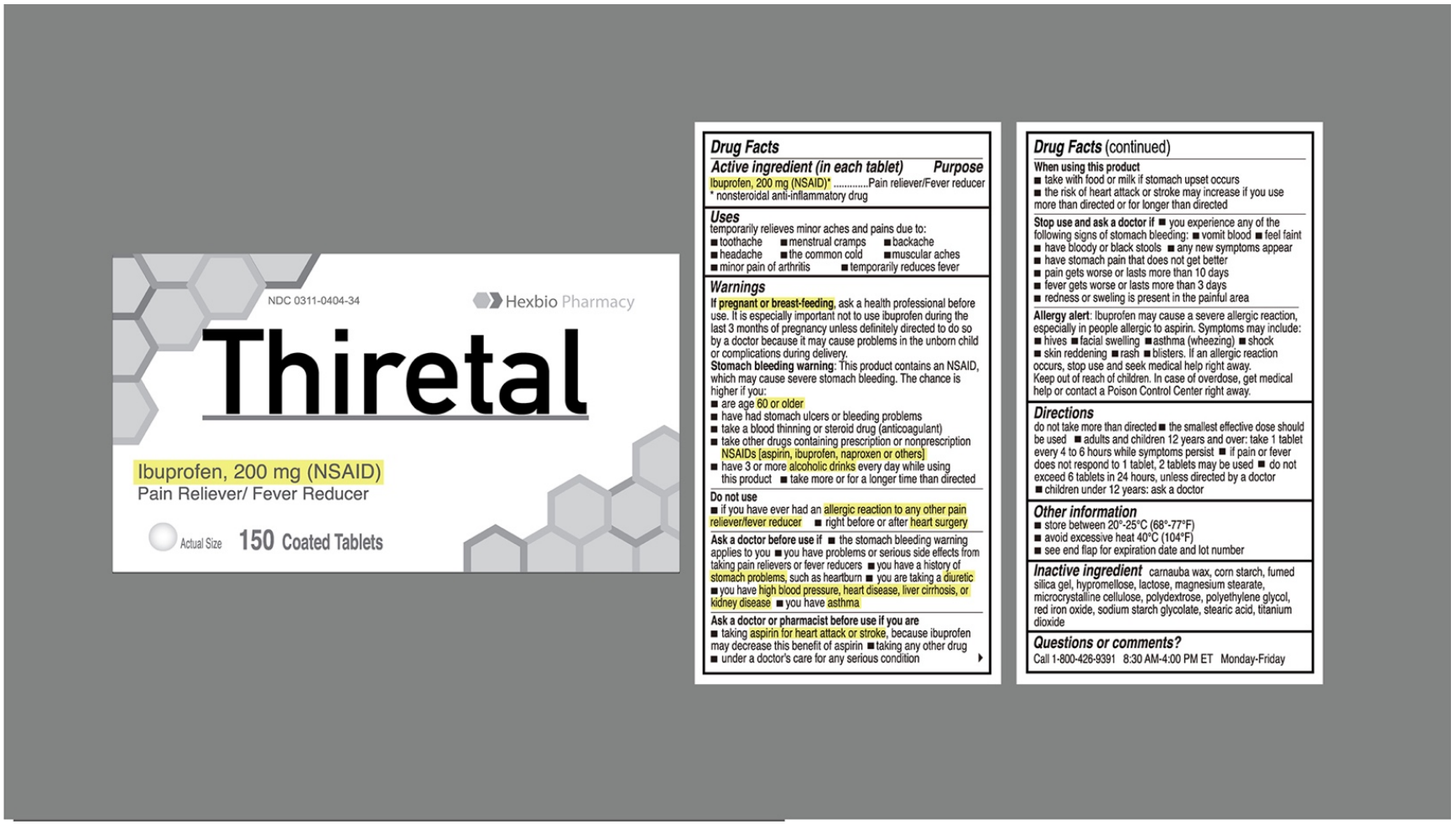


Figure A.6 DD Change in DFL

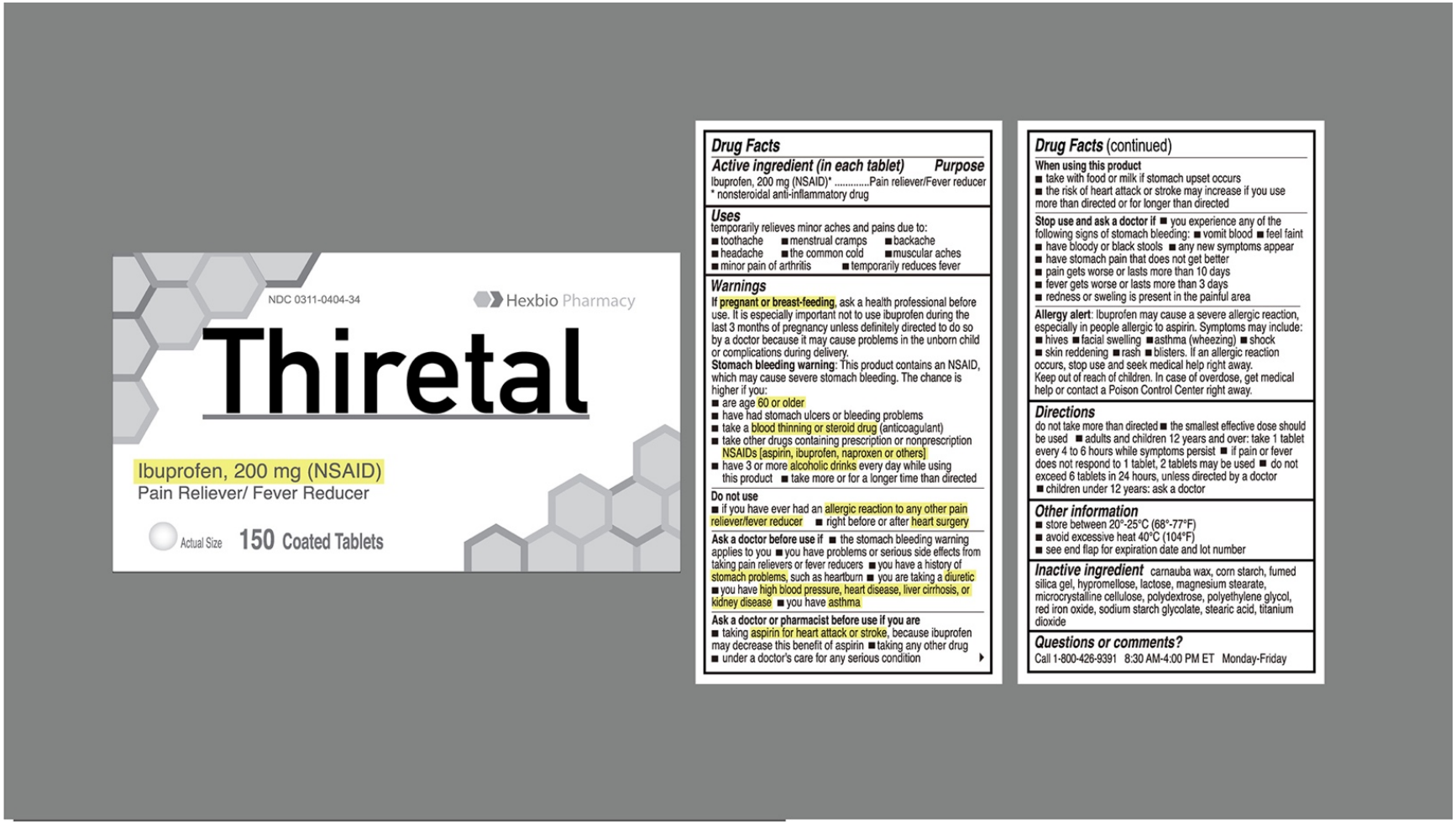


Figure A.7 AI Change in DFL

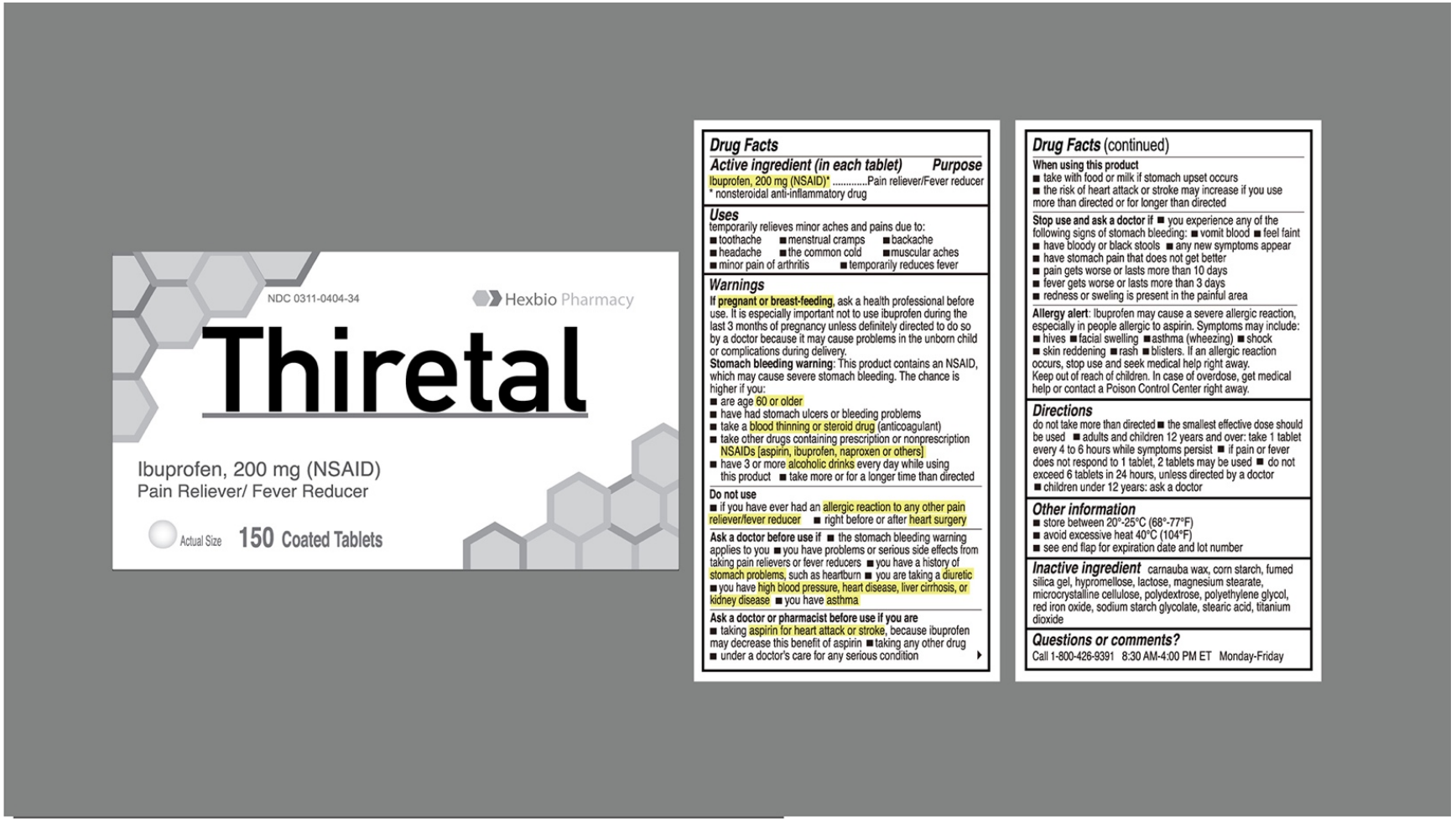


Figure A.8 AI Change on PDP

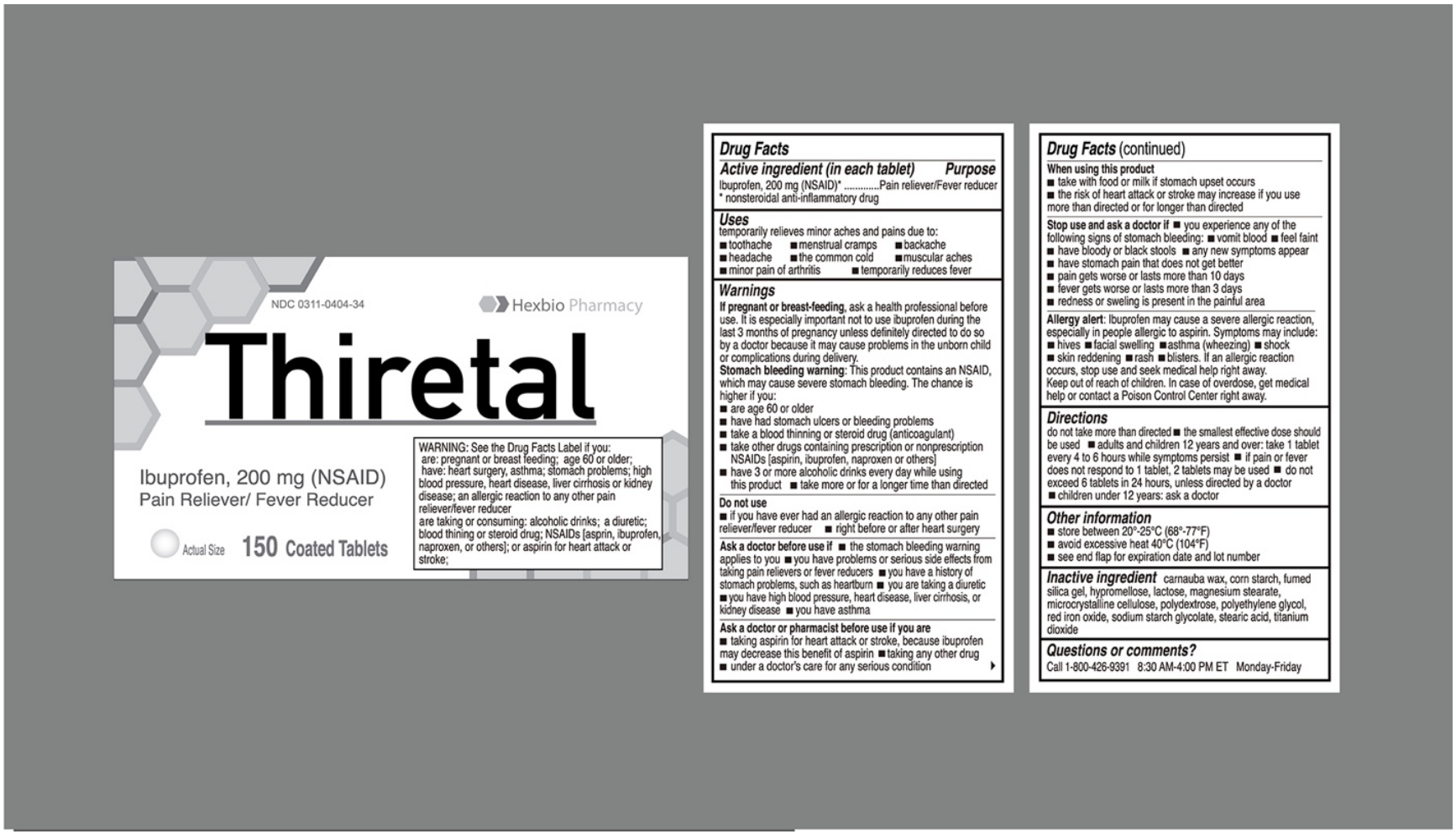


Figure A.9 Example Label for Ibuprofen, enhanced with the FOP Treatment

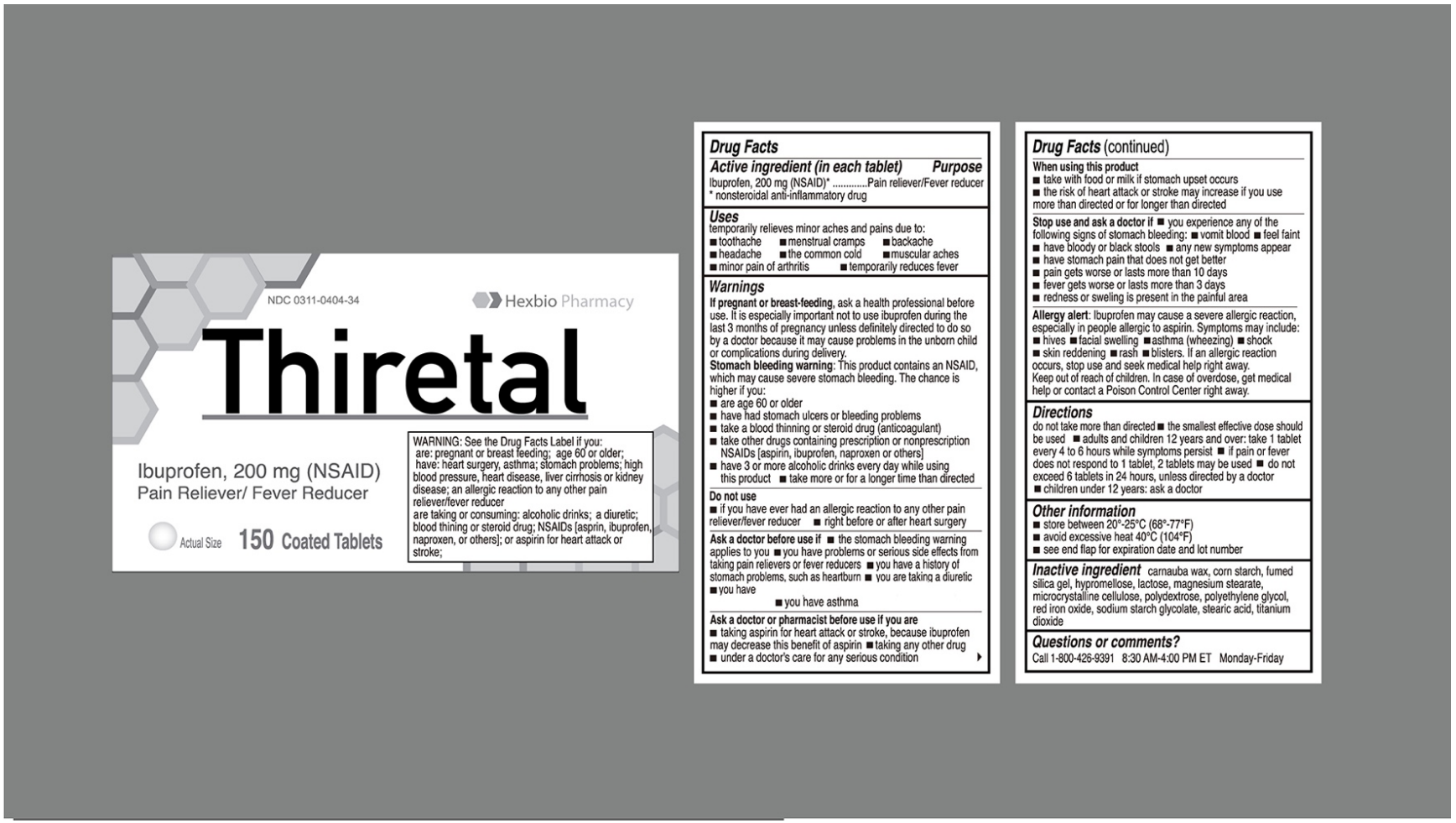


Figure A.10 DD Change in DFL

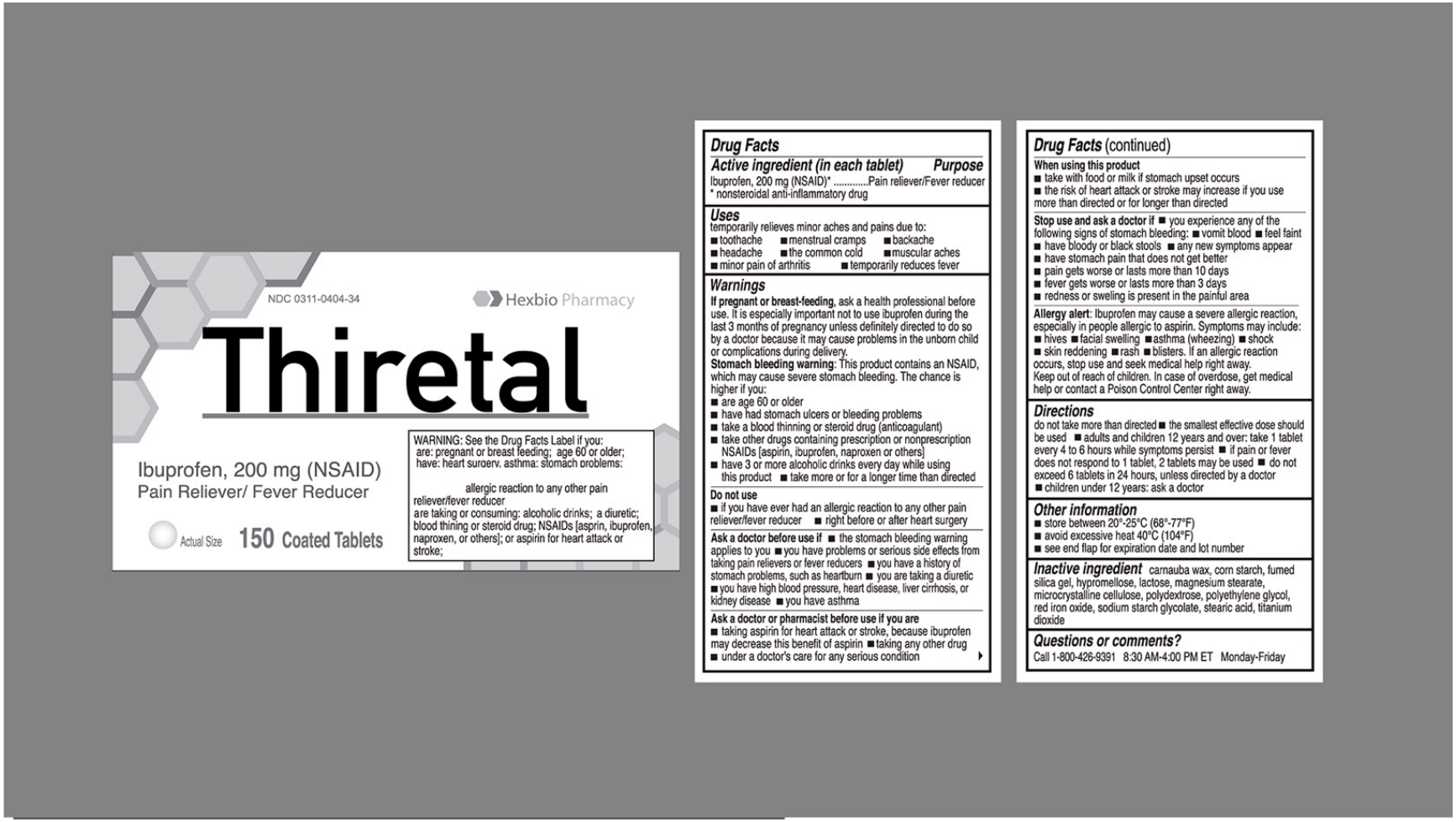


Figure A.11 DD Change on PDP

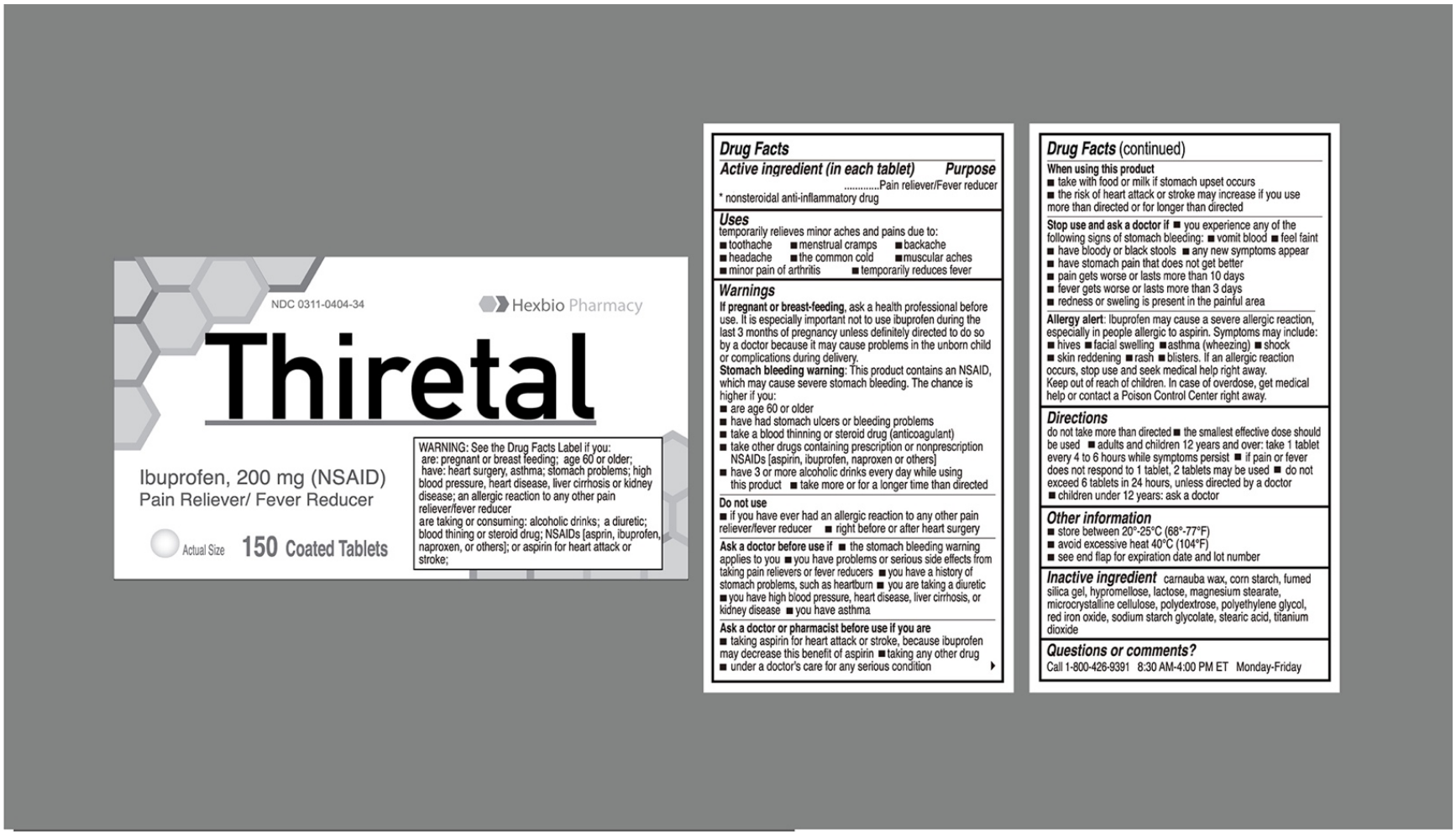


Figure A.12 AI Change on DFL

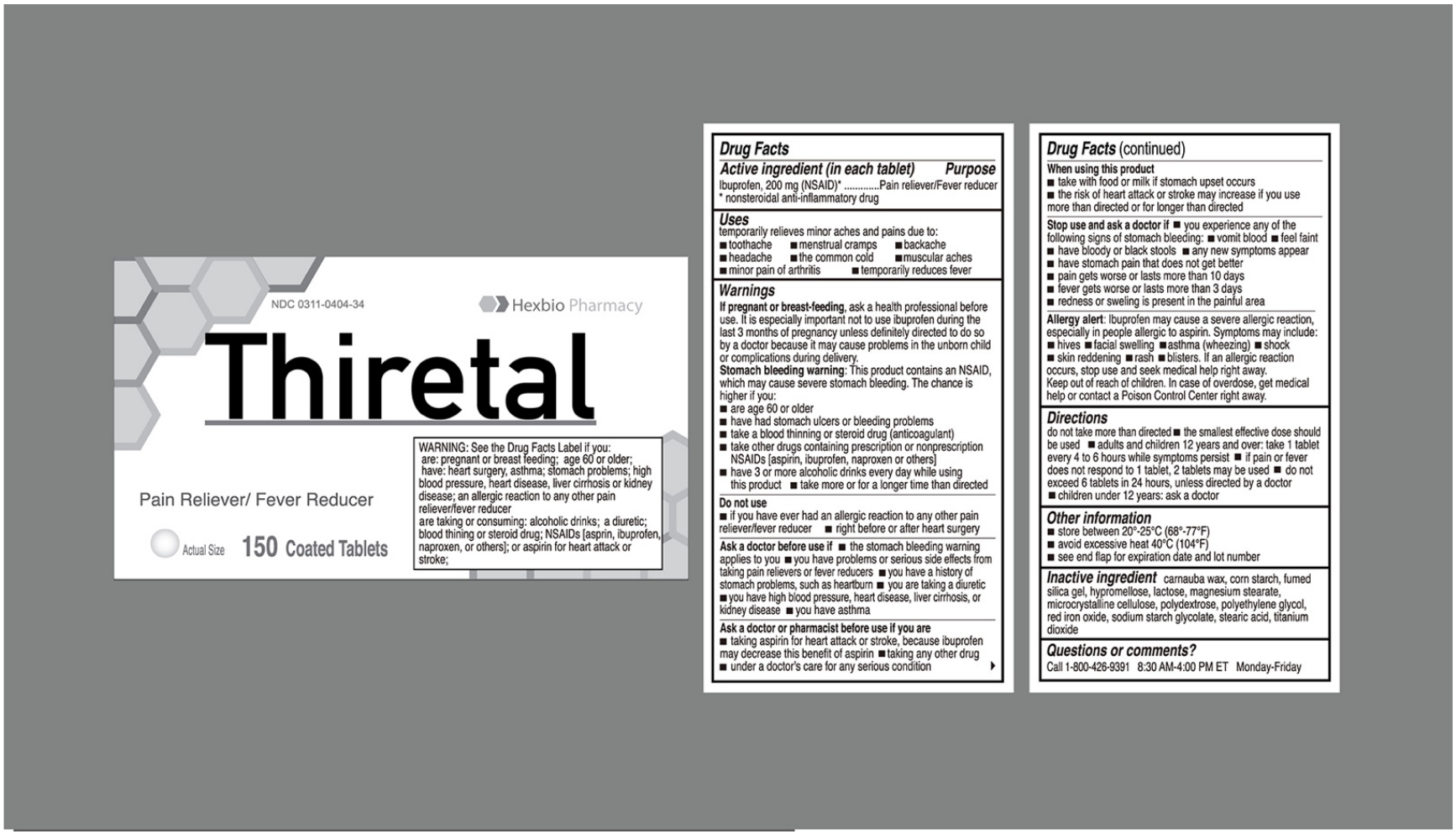


Figure A.13 AI Change on PDP

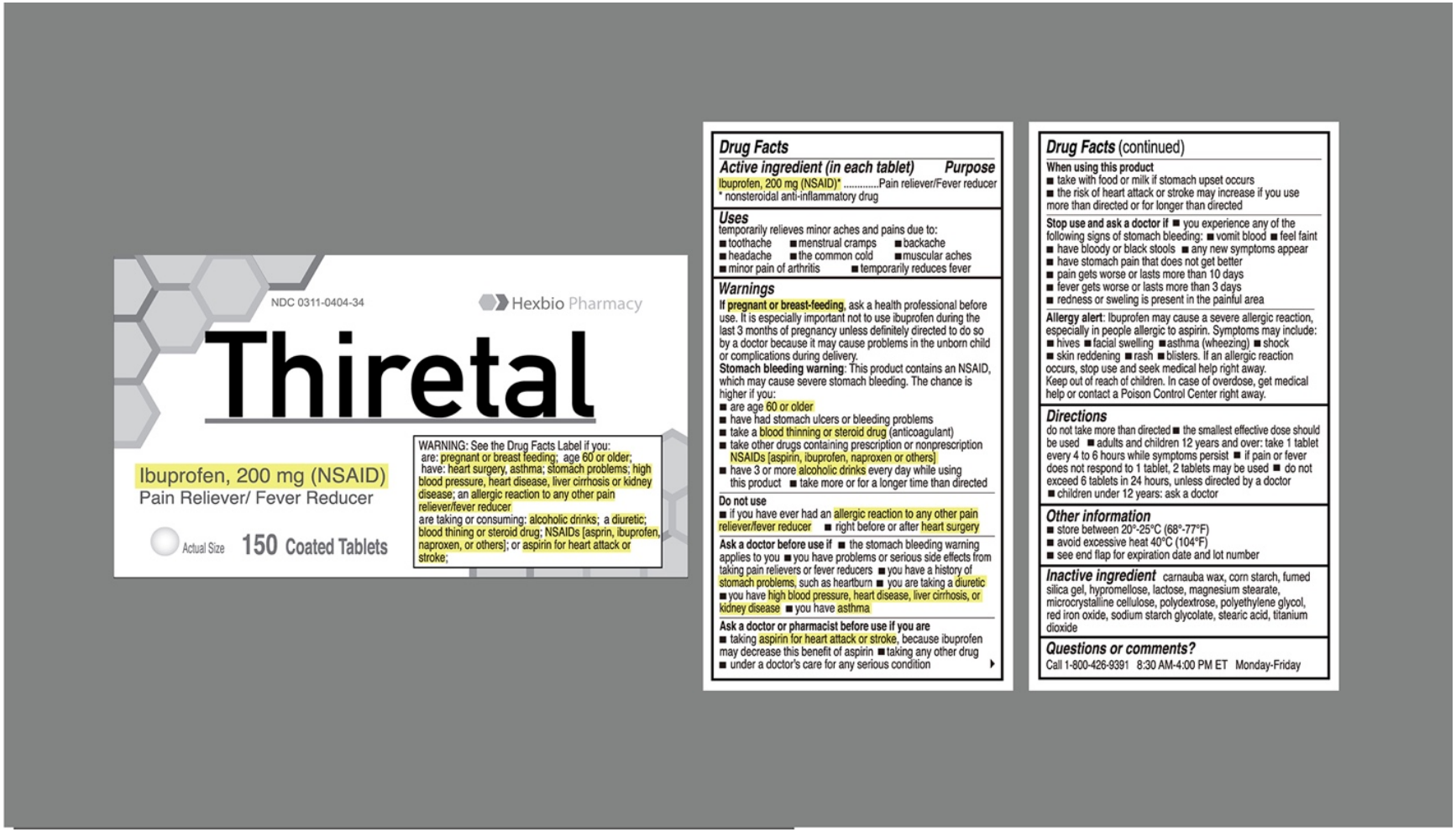


Figure A.14 Example Label for Ibuprofen, enhanced with both the FOP or Highlight Treatment

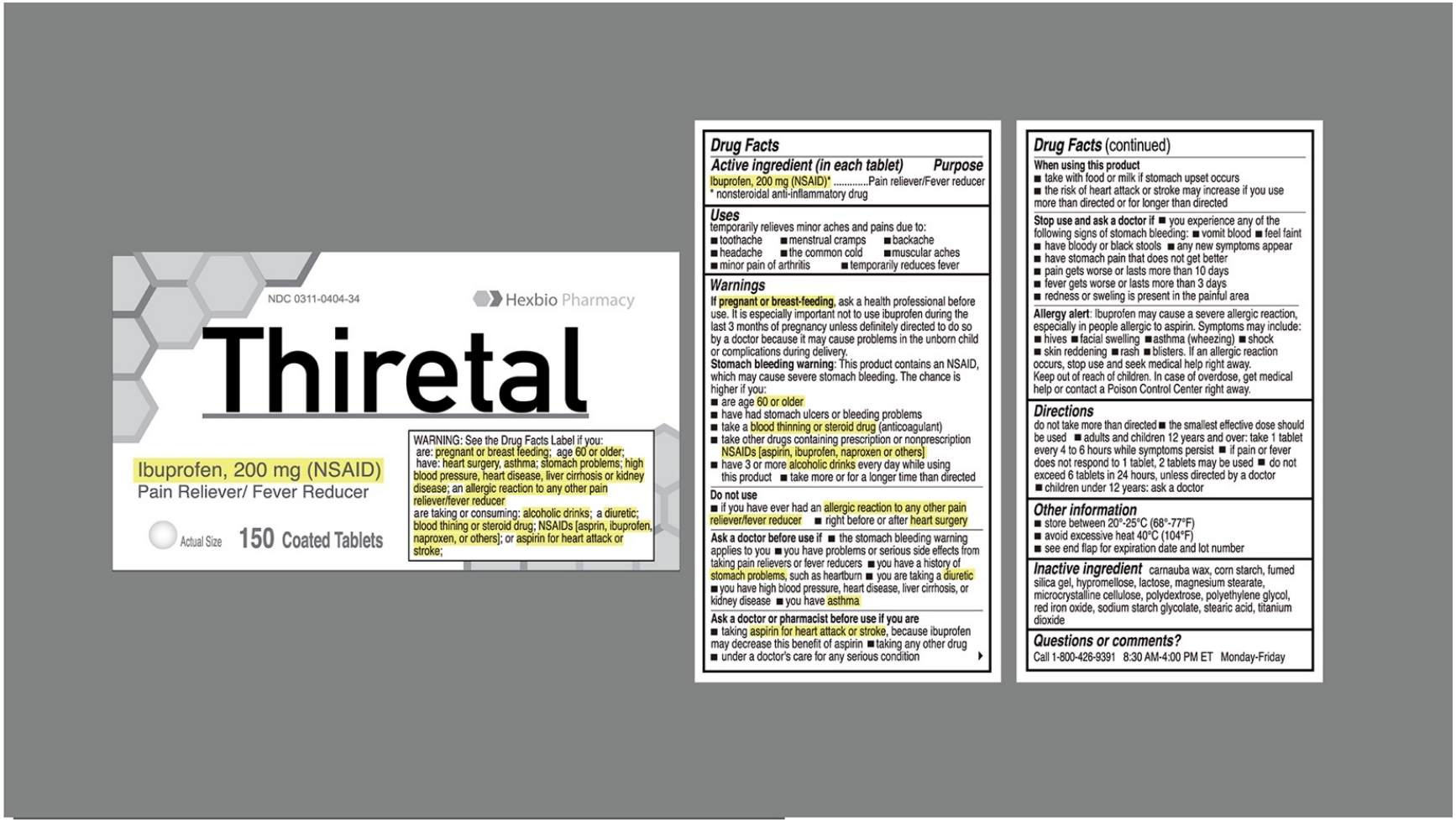


Figure A.15 DD Change in DFL

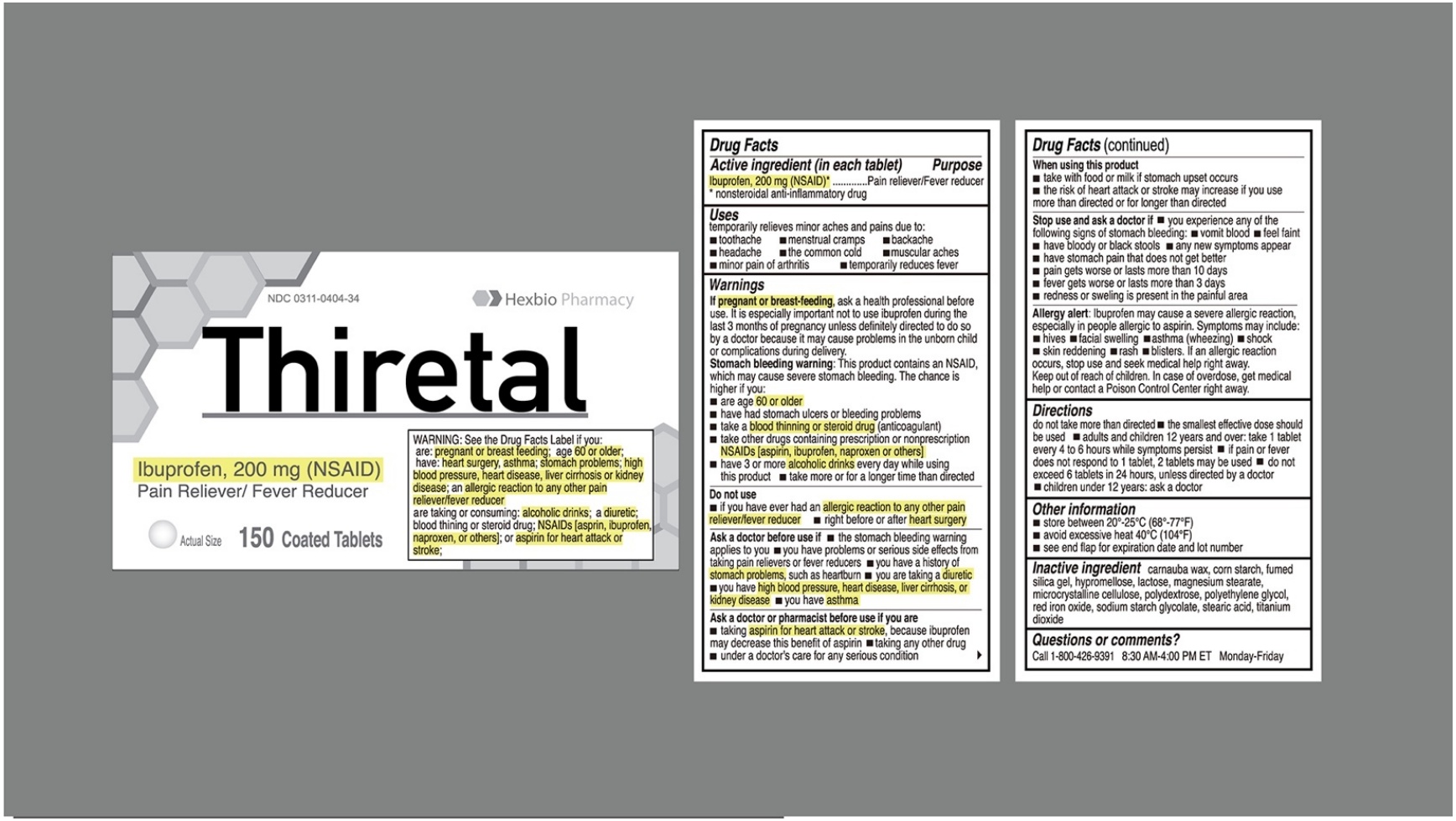


Figure A.16 DD Change on PDP

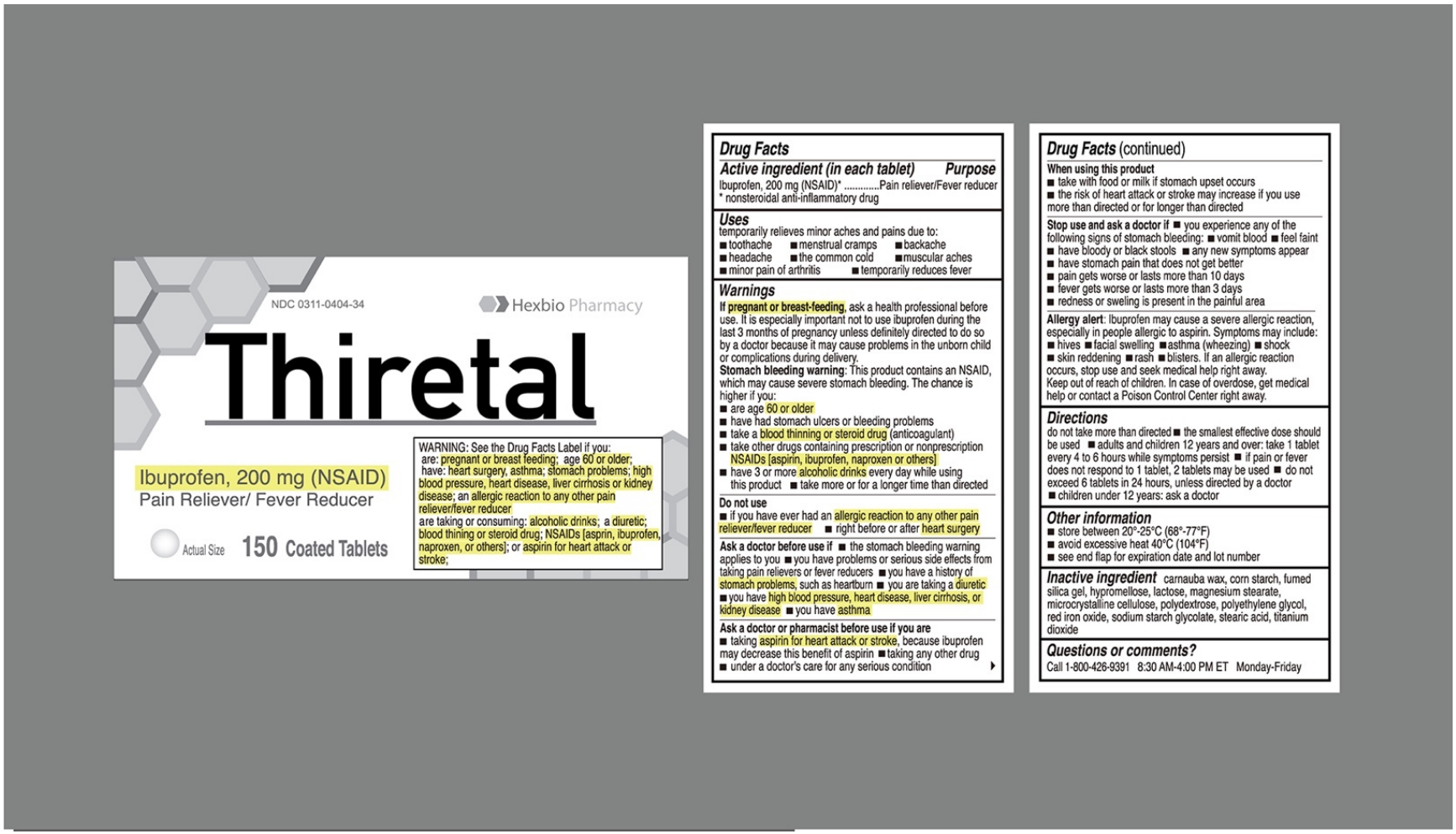


Figure A.17 AI Change on DFL

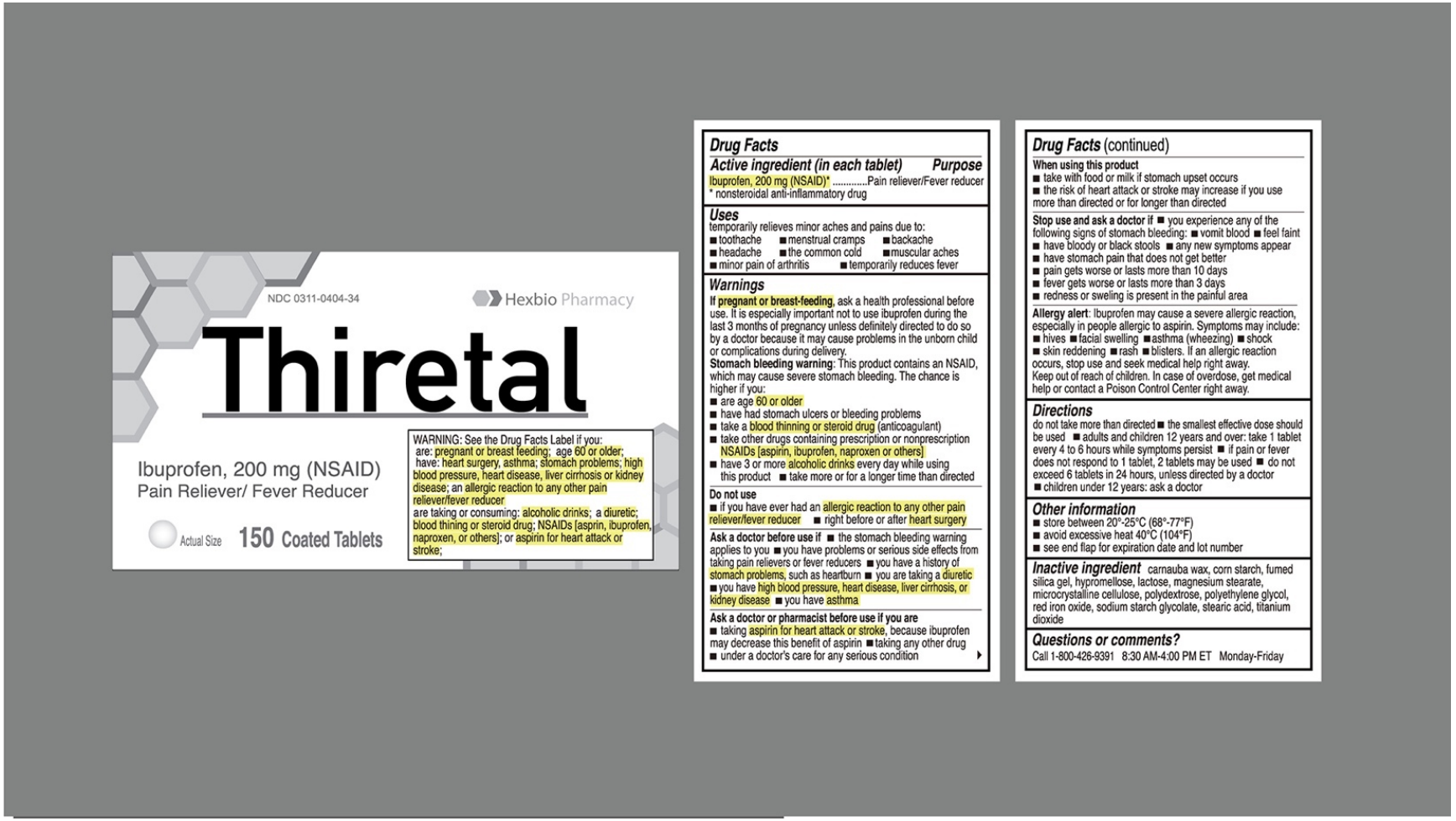


Figure A.18 AI Change on PDP

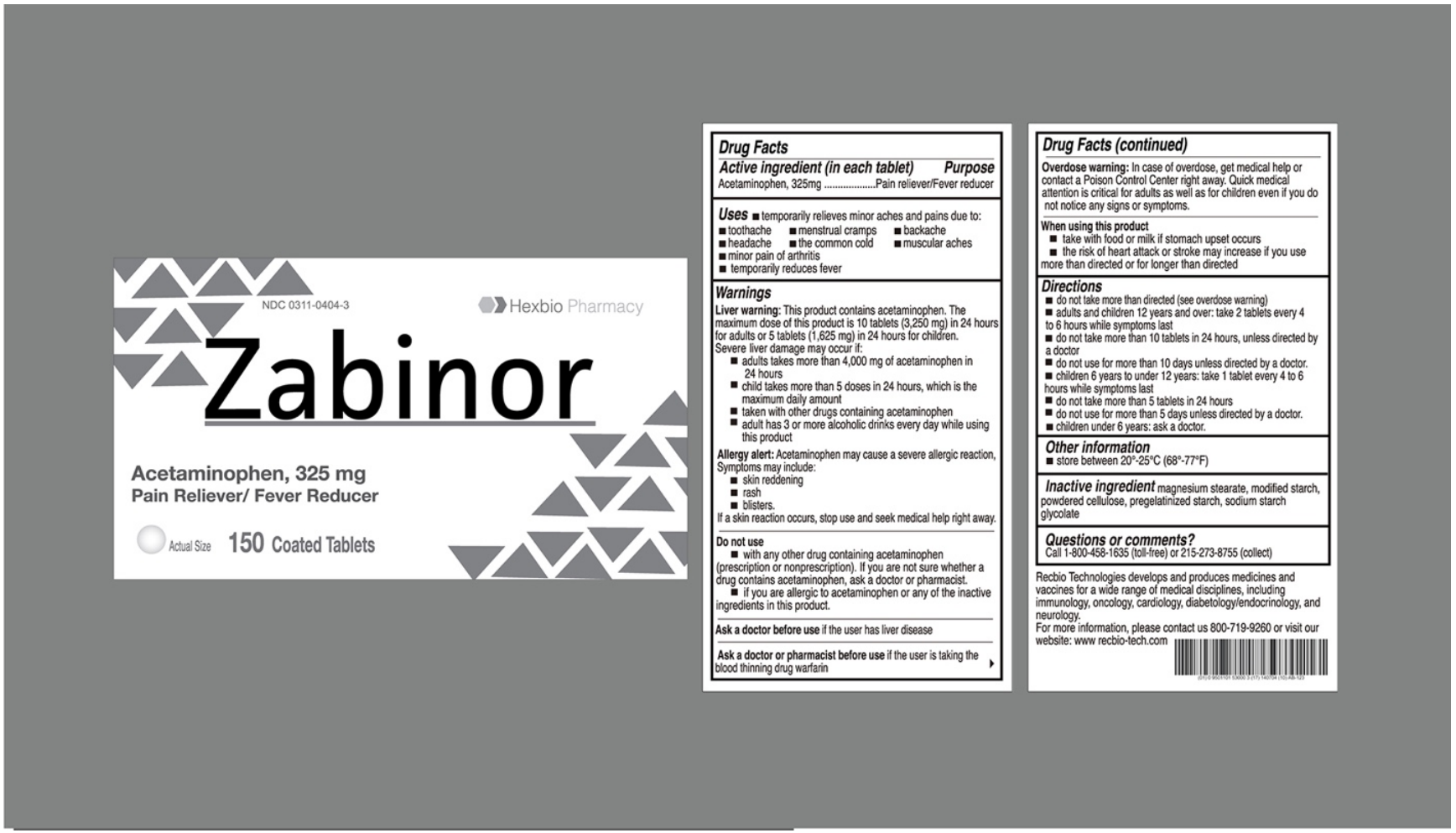


Figure A.19 Example Standard Label for Acetaminophen, with neither the FOP or Highlight Treatment

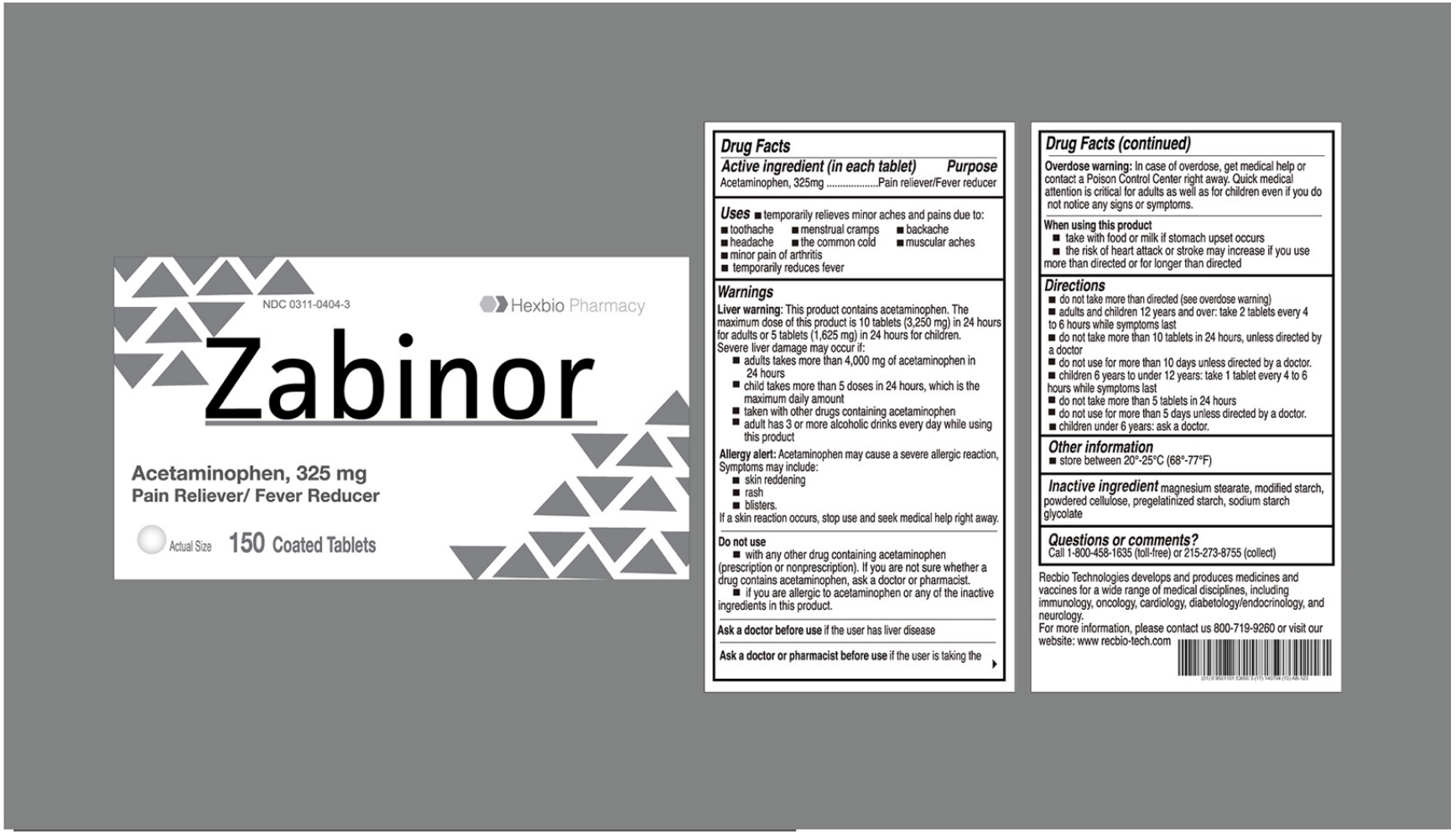


Figure A.20 DD Change in DFL

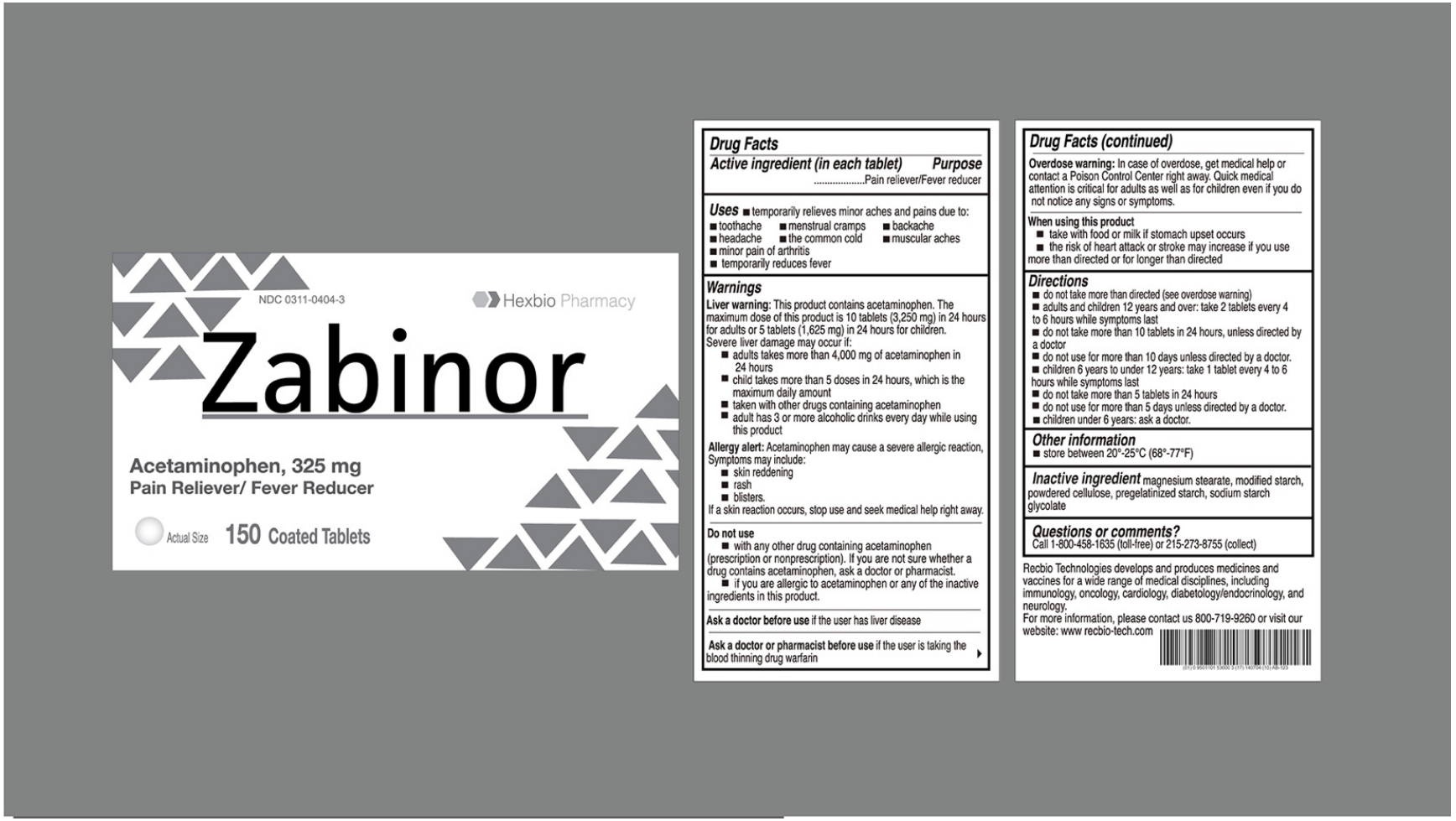


Figure A.21 AI Change in DFL

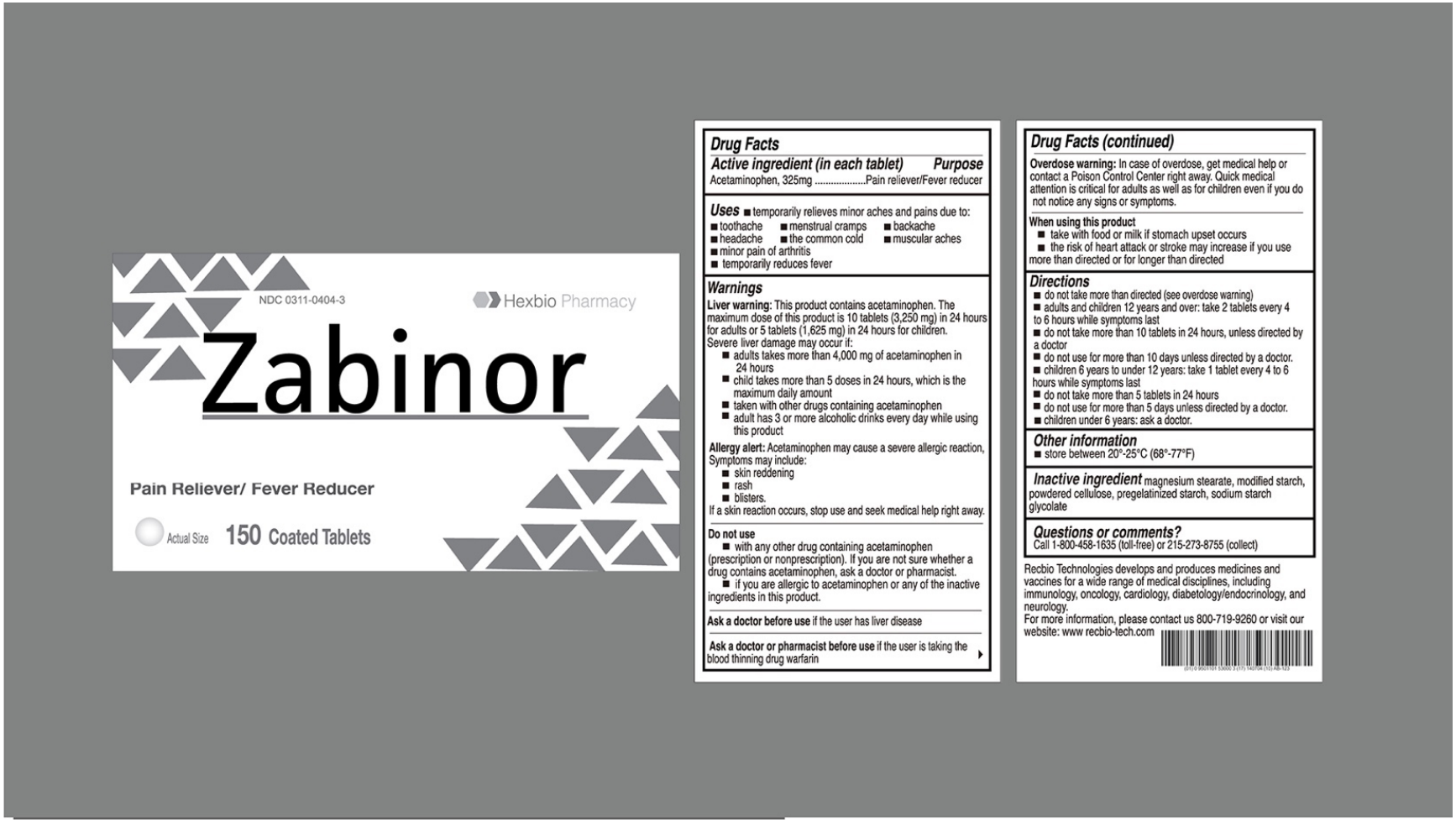


Figure A.22 AI Change on PDP

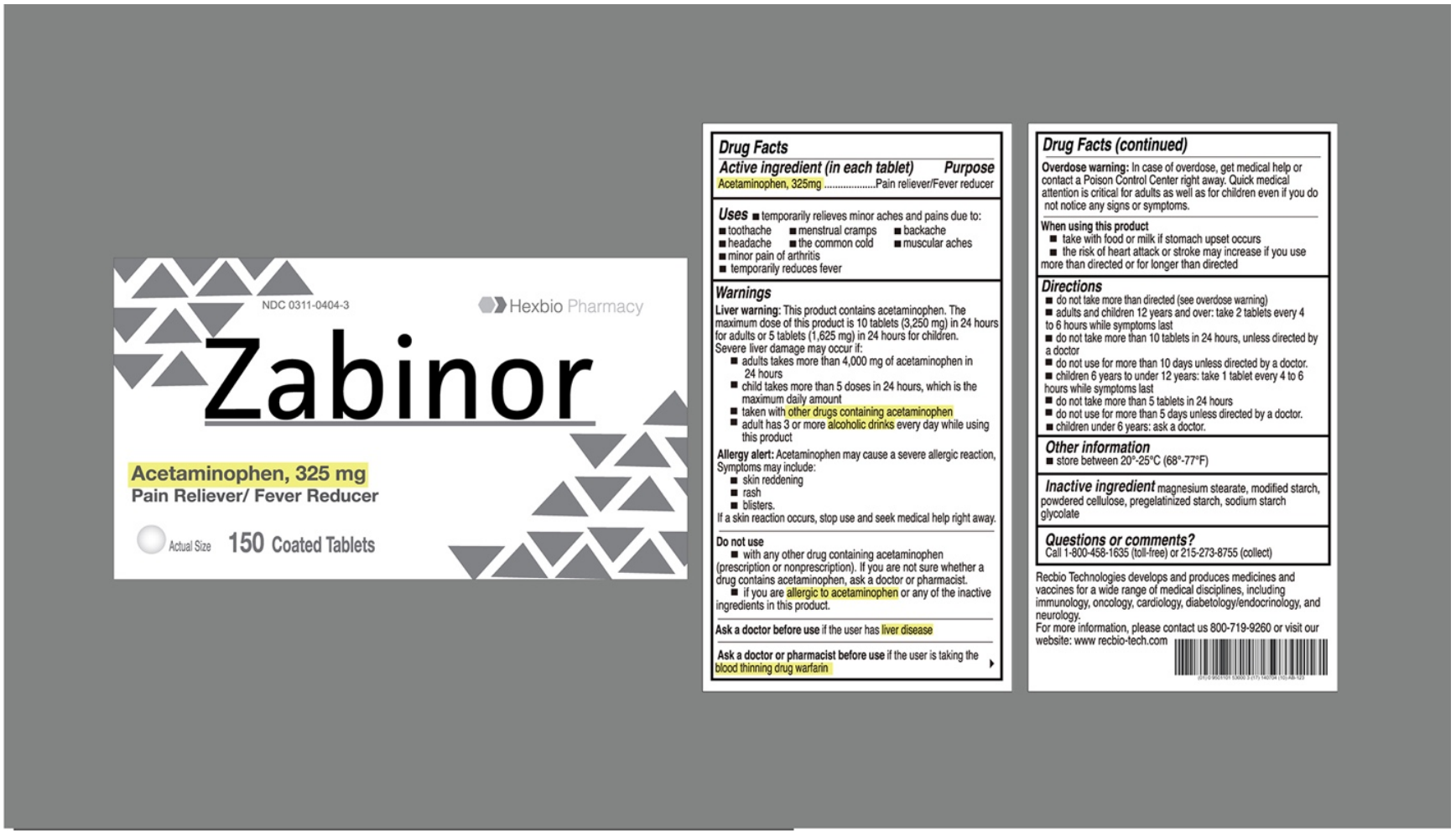


Figure A.23 Example Label for Acetaminophen, enhanced with the Highlight Treatment

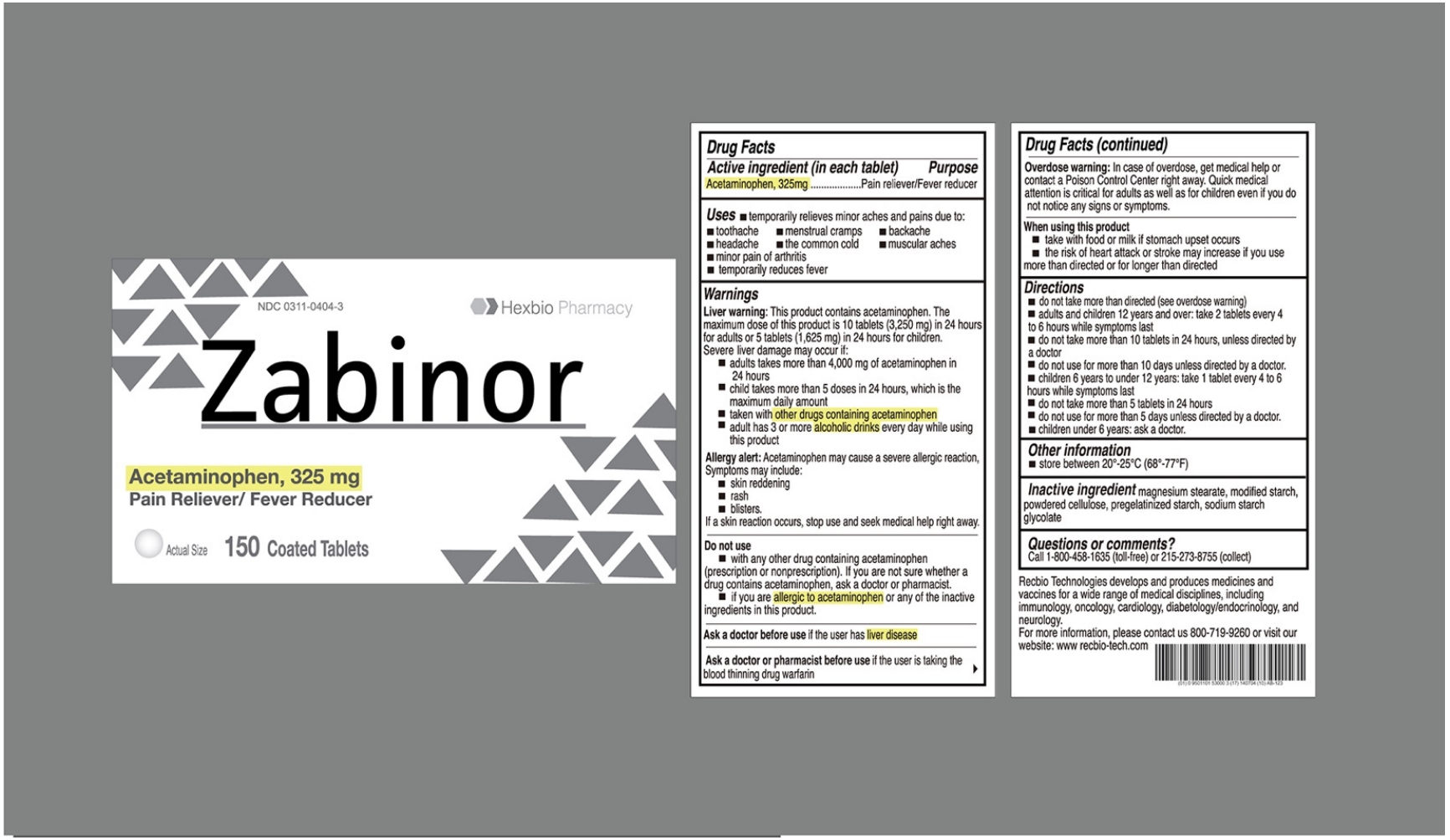


Figure A.24 DD Change in DFL

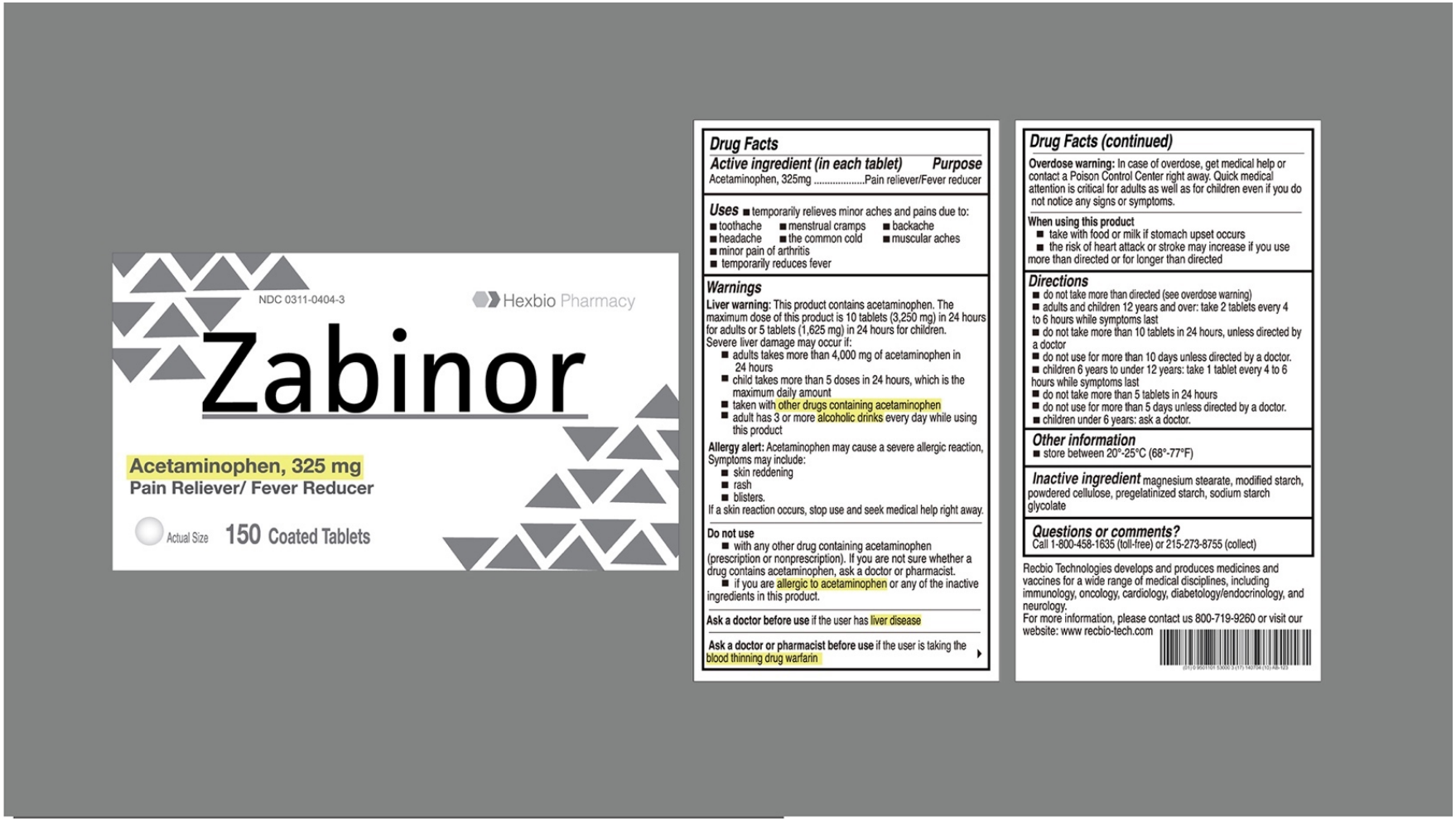


Figure A.25 AI Change in DFL

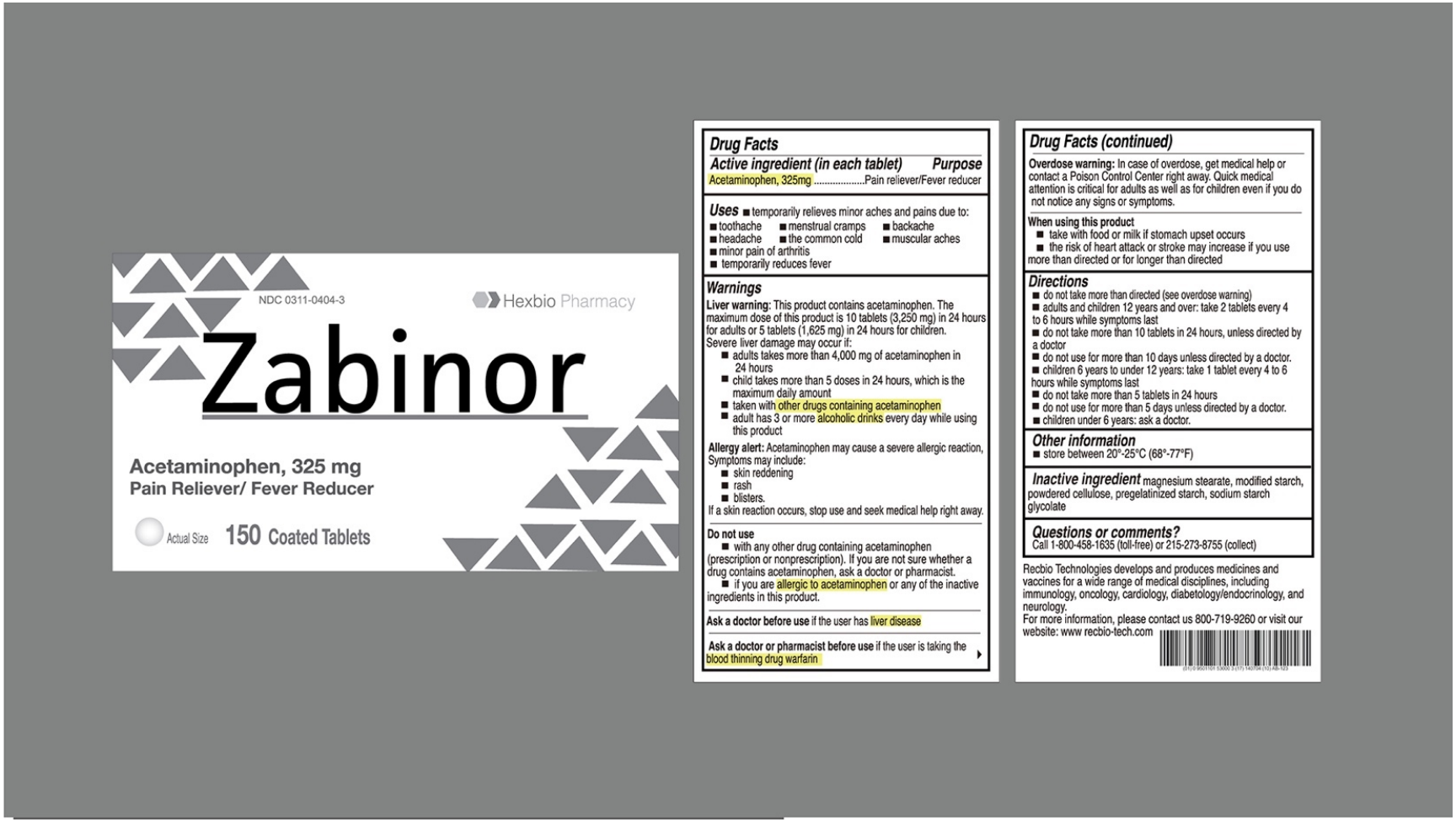


Figure A.26 AI Change on PDP

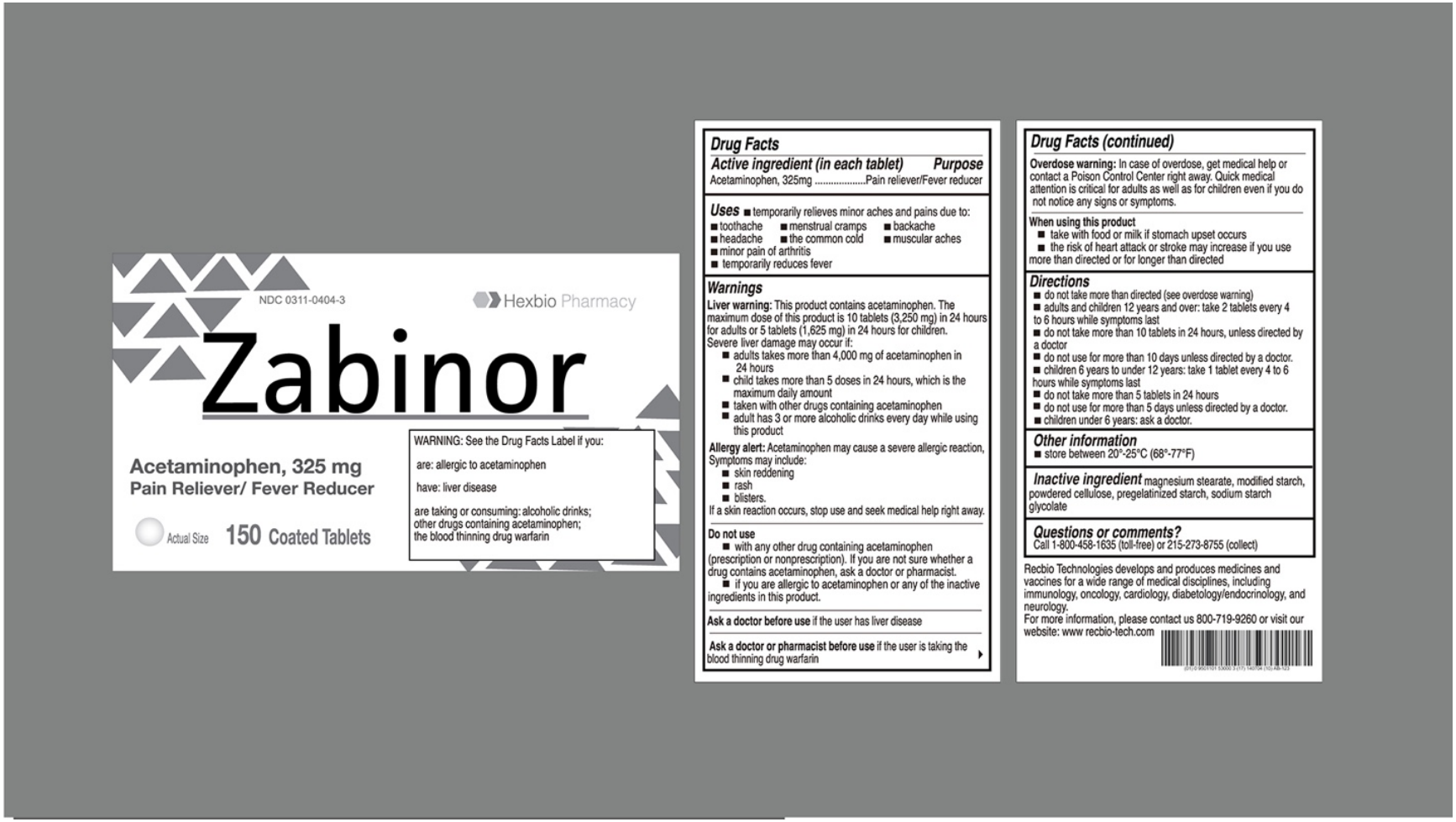


Figure A.27 Example Label for Acetaminophen, enhanced with the FOP Treatment

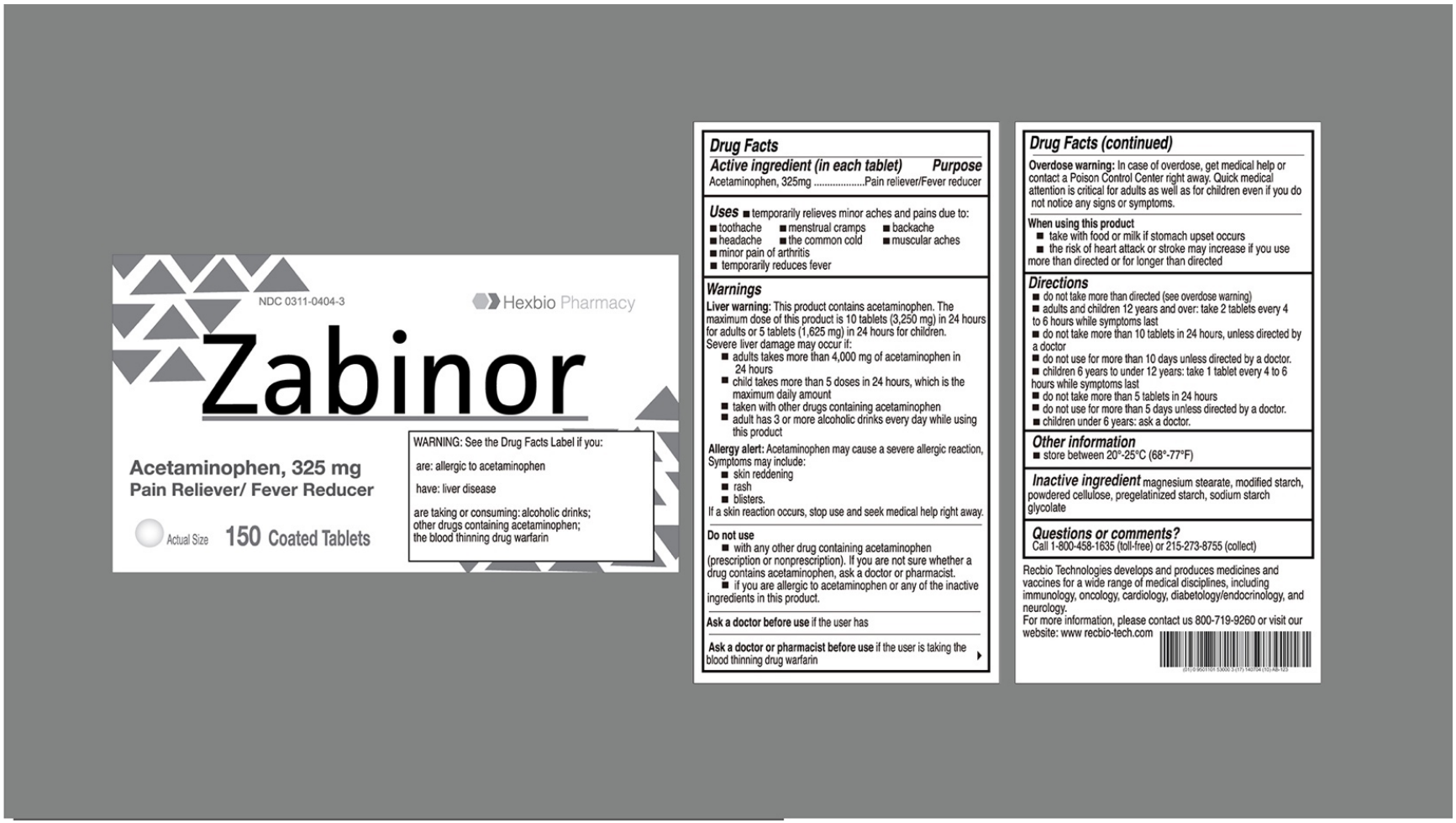


Figure A.28 DD Change in DFL

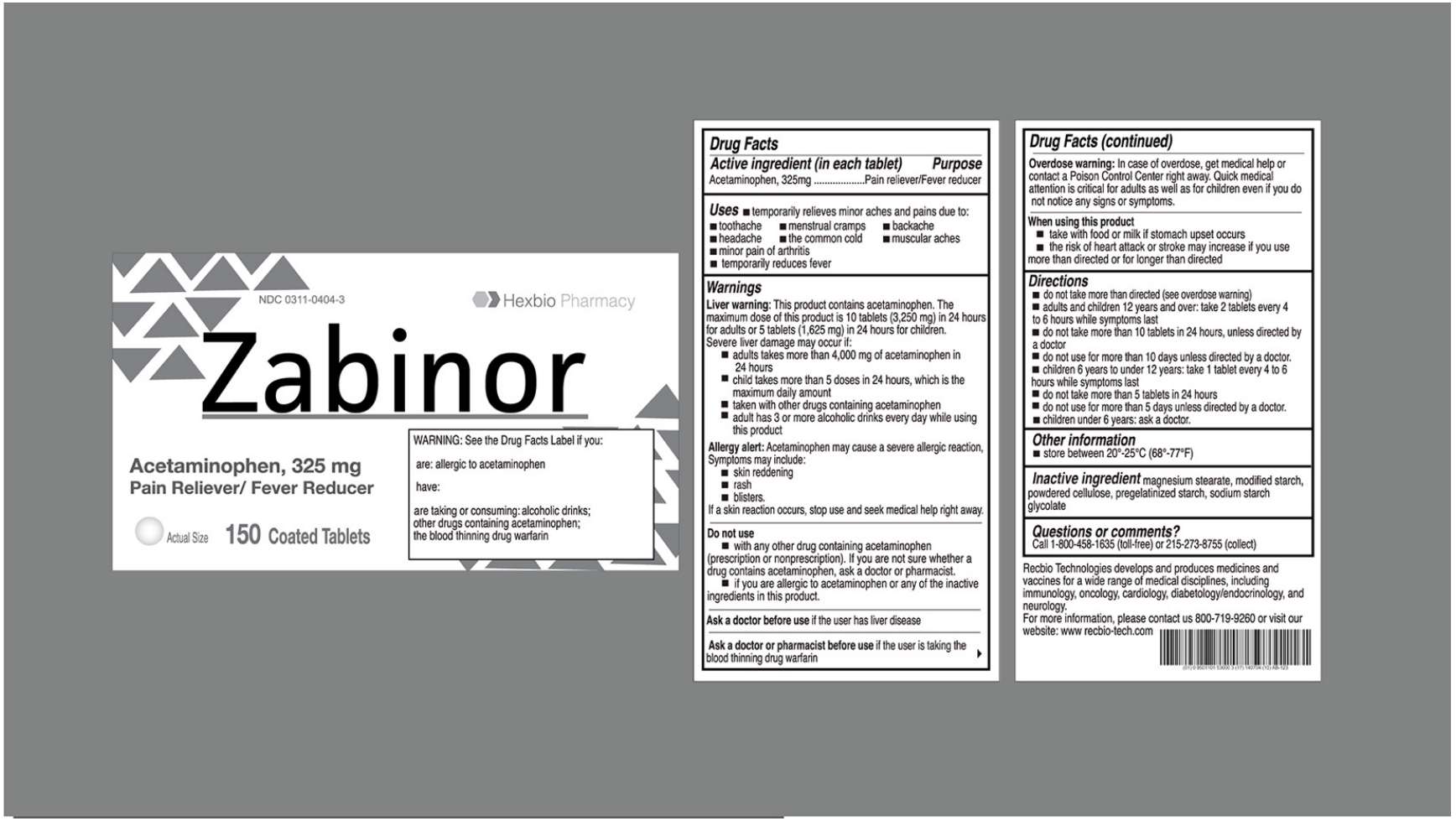


Figure A.29 DD Change on PDP

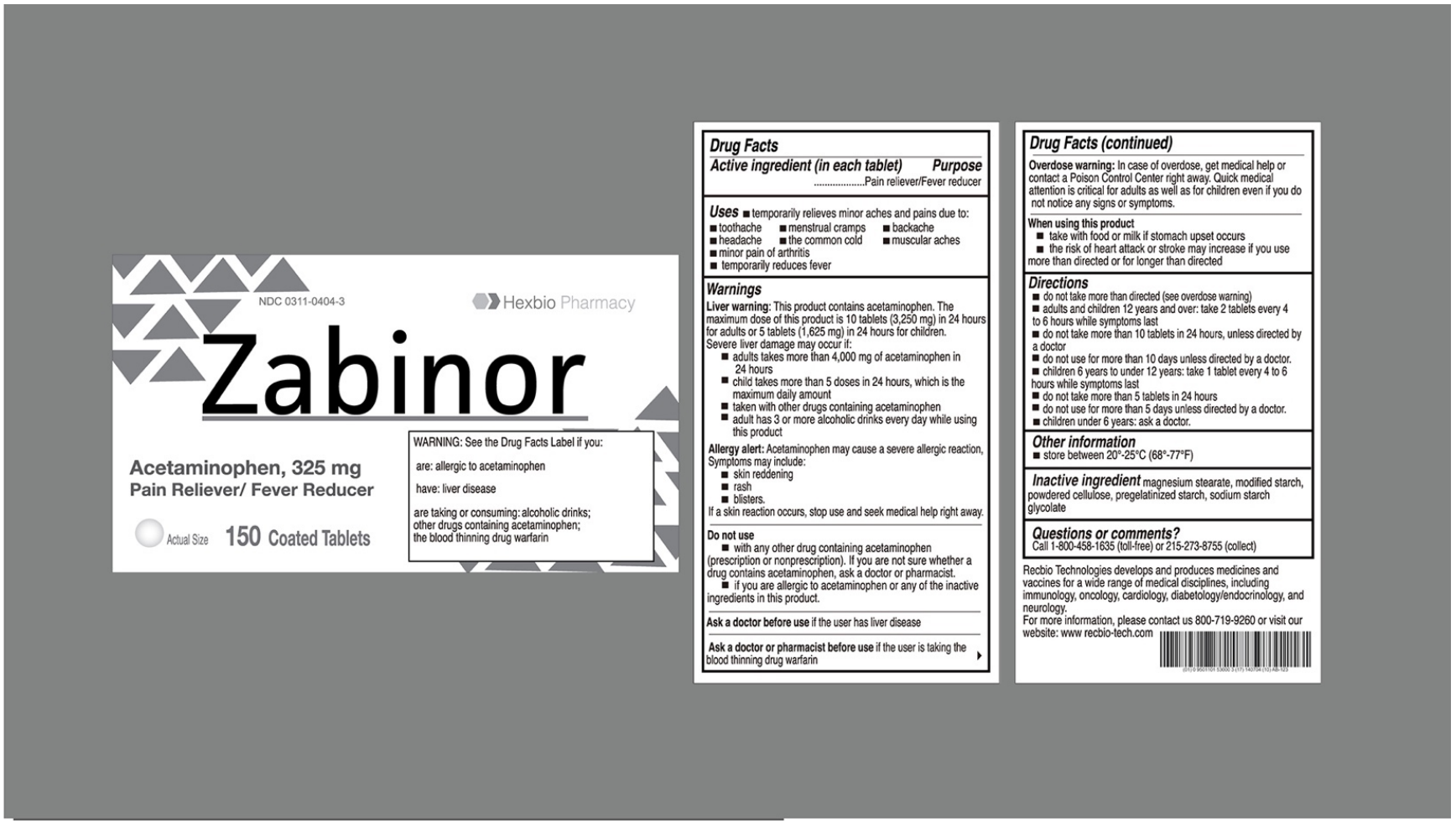


Figure A.30 AI Change in DFL

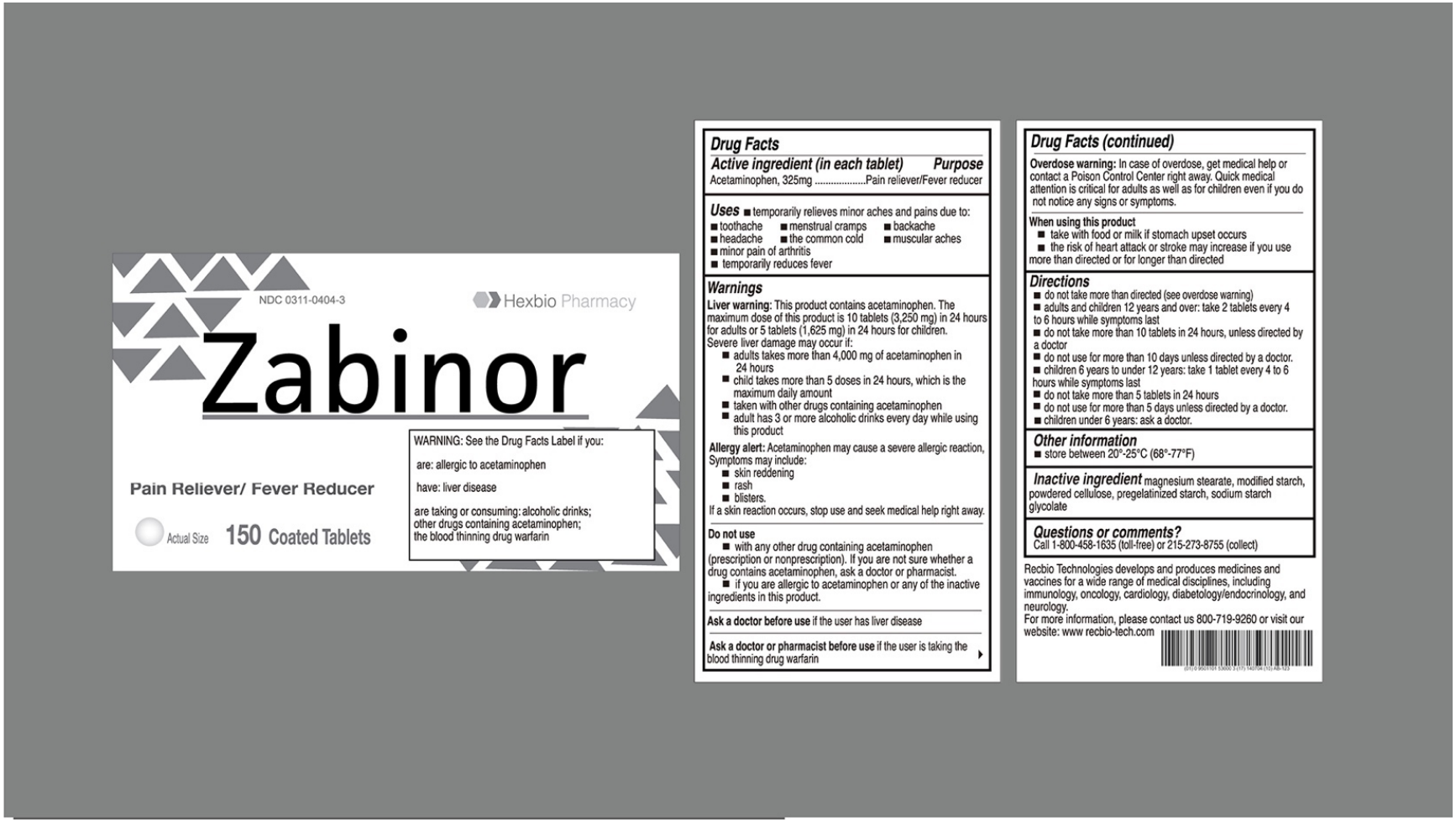


Figure A.31 AI Change on PDP

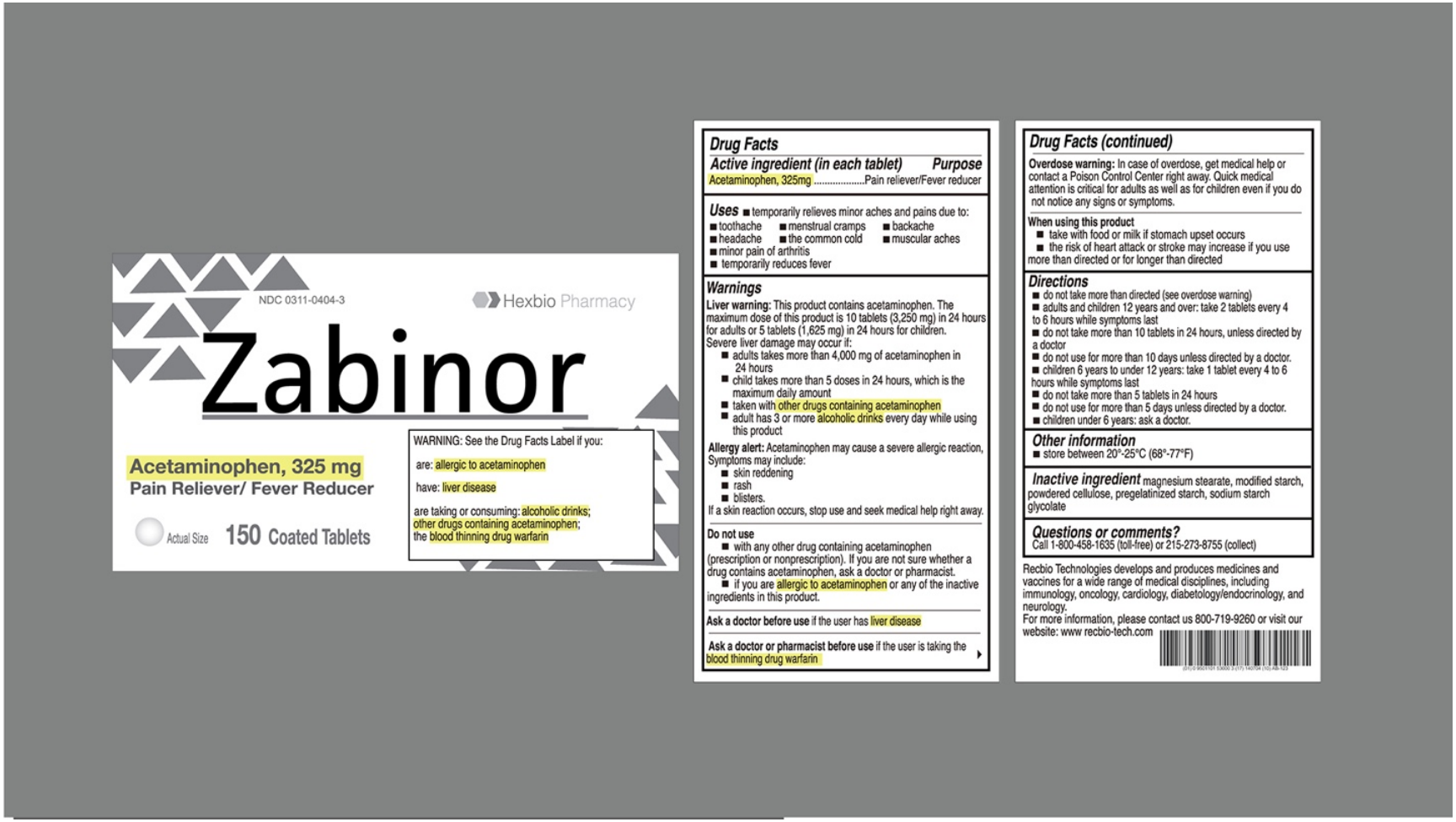


Figure A.32 Example Label for Acetaminophen, enhanced with both the FOP or Highlight Treatment

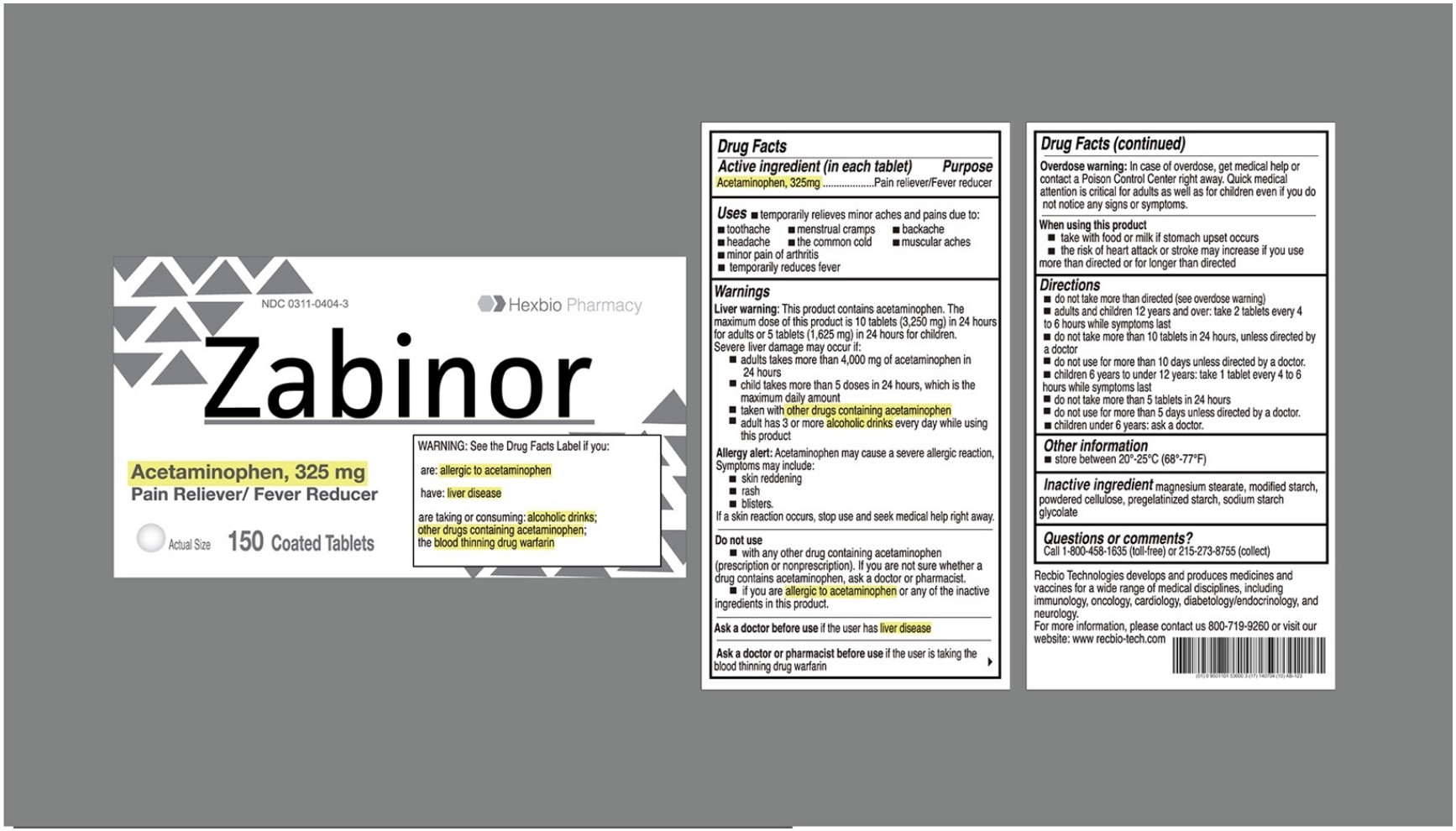


Figure A.33 DD Change in DFL

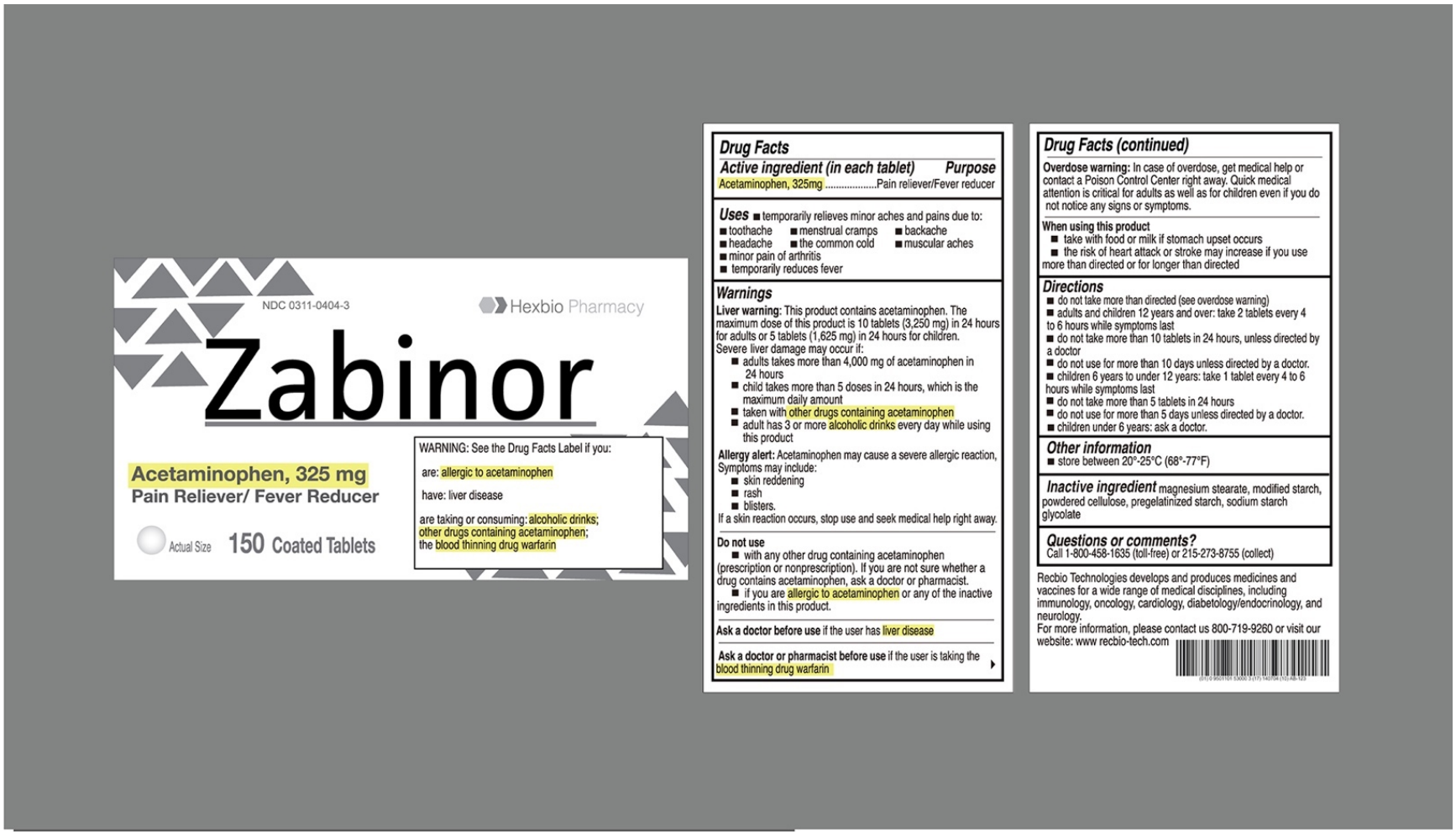


Figure A.34 DD Change on PDP

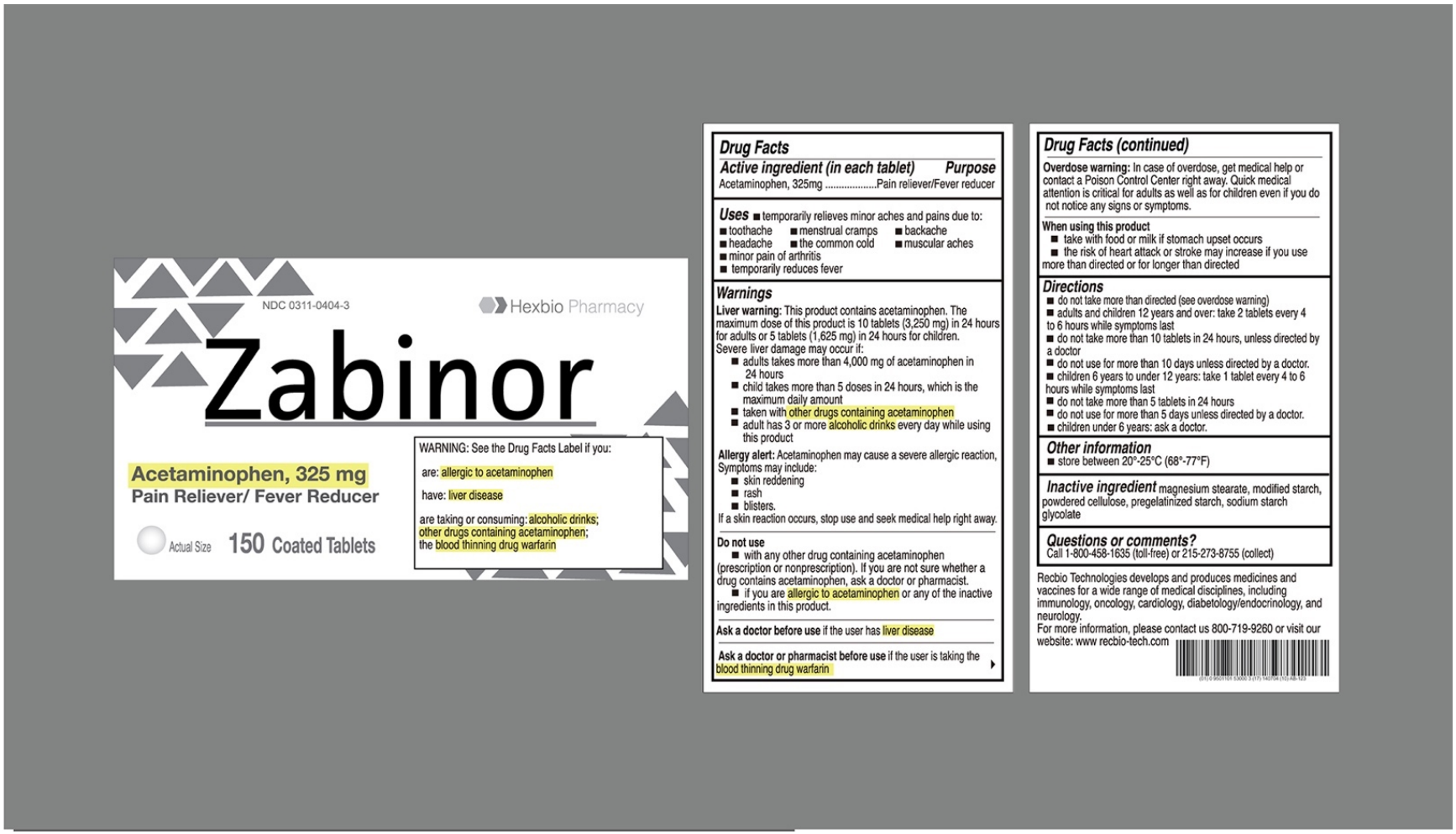


Figure A.35 AI Change in DFL

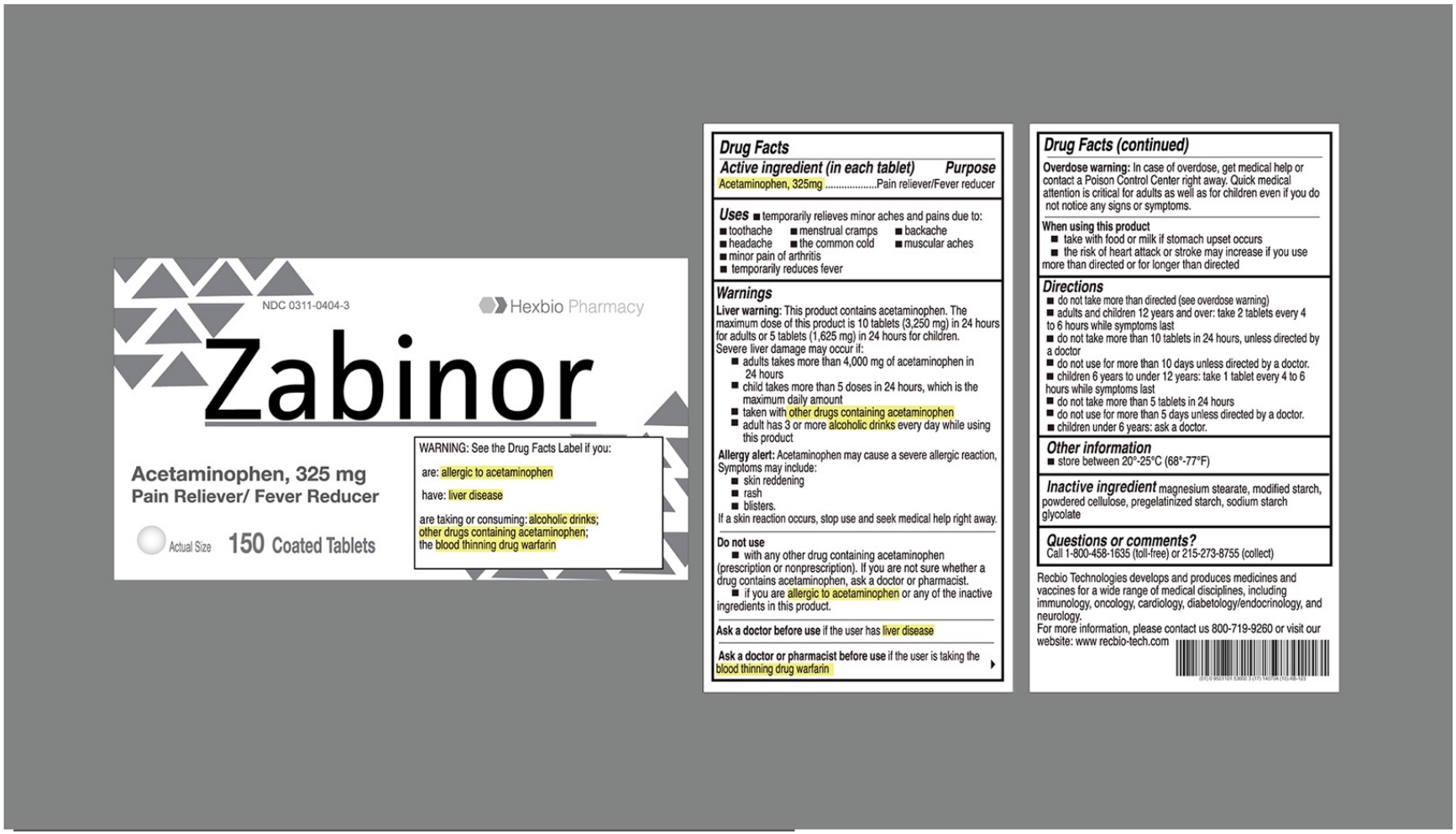


Figure A.36 AI Change on PDP

Figure A.37 Example Standard Label for Phenylephrine, with neither the FOP or Highlight Treatment

Figure A.38 DD Change in DFL

Figure A.39 AI Change in DFL

Figure A.40 AI Change on PDP

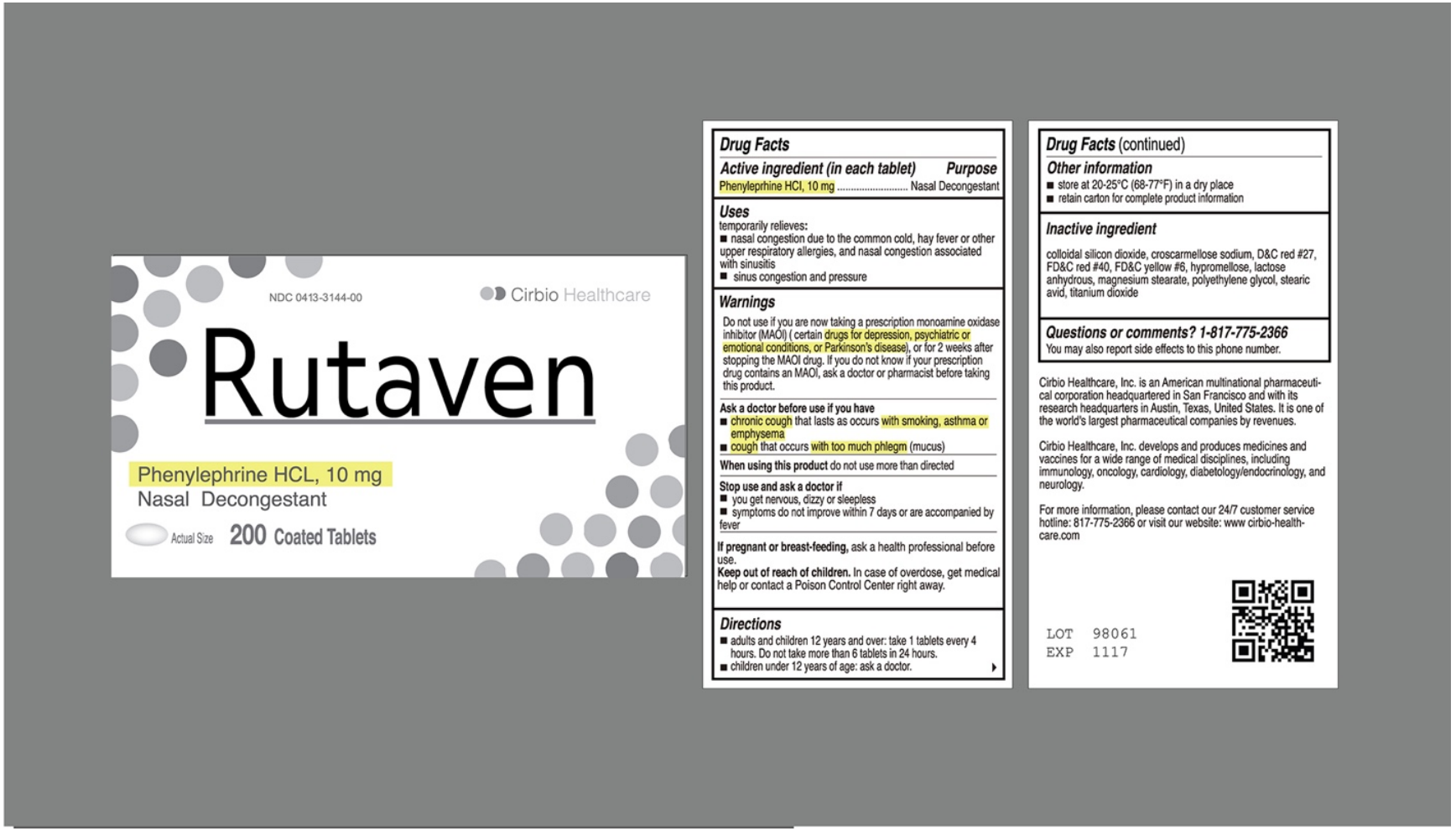


Figure A.41 Example Label for Phenylephrine, enhanced with the Highlight Treatment

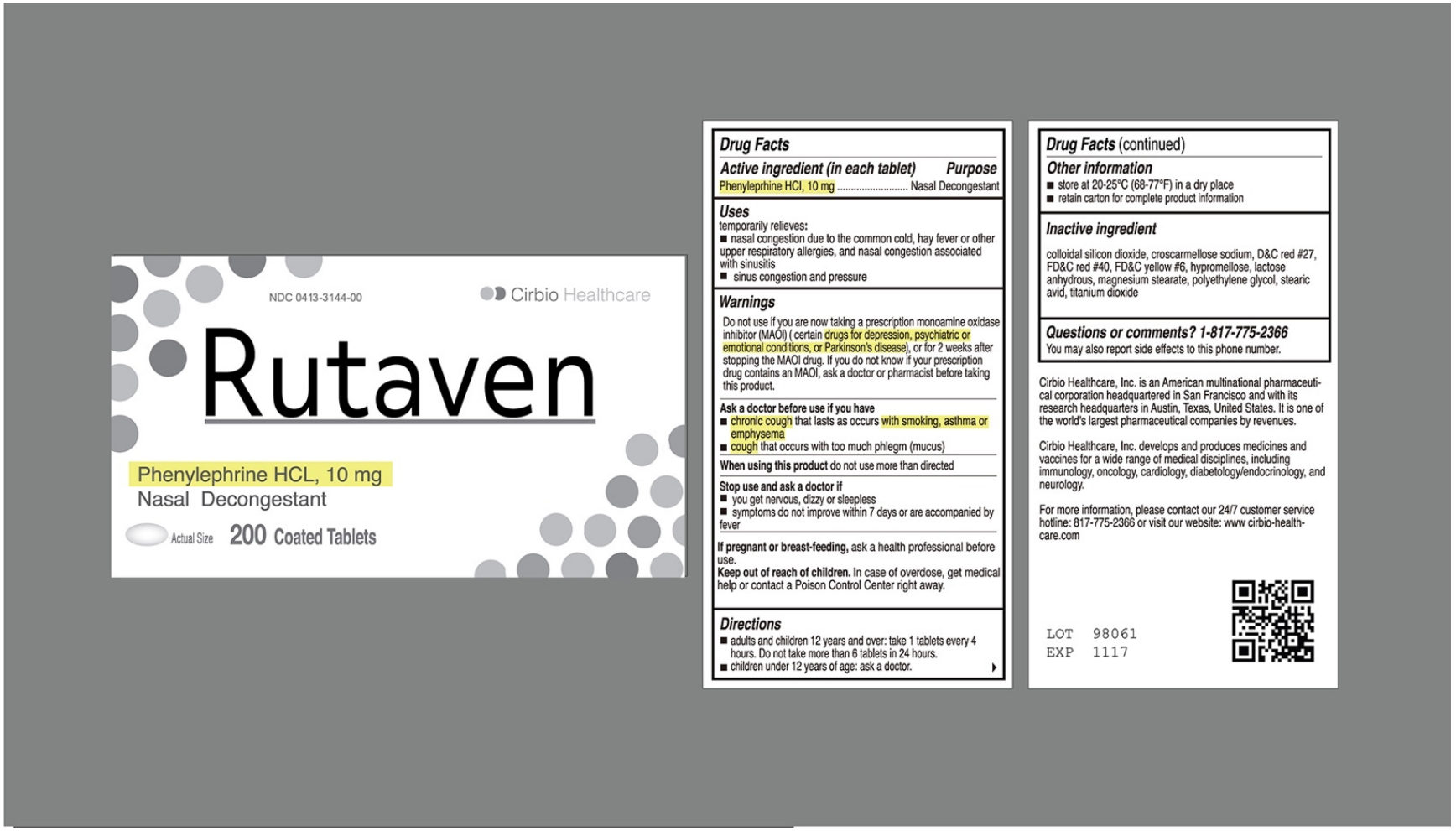


Figure A.42 DD Change in DFL

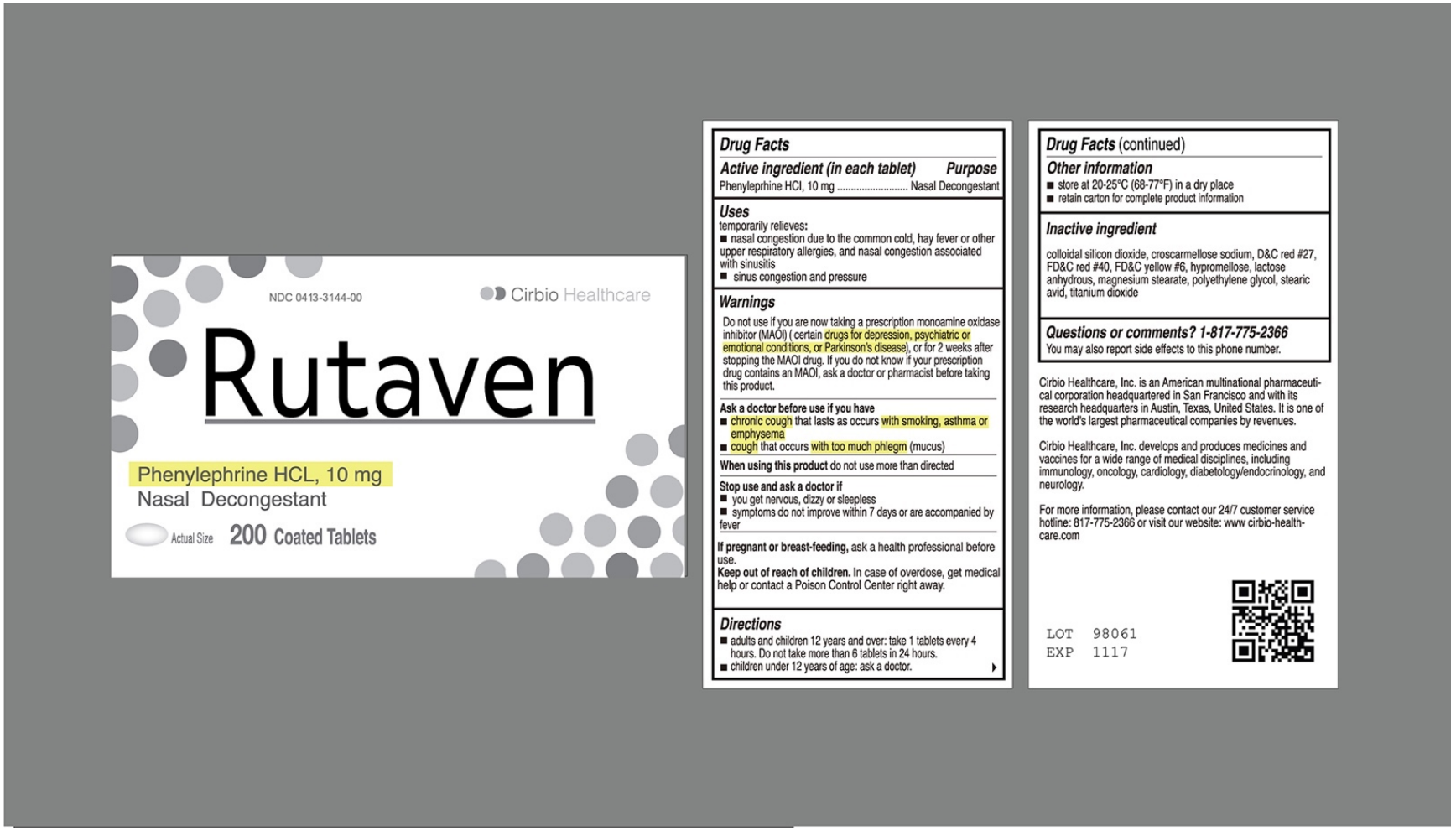


Figure A.43 AI Change in DFL

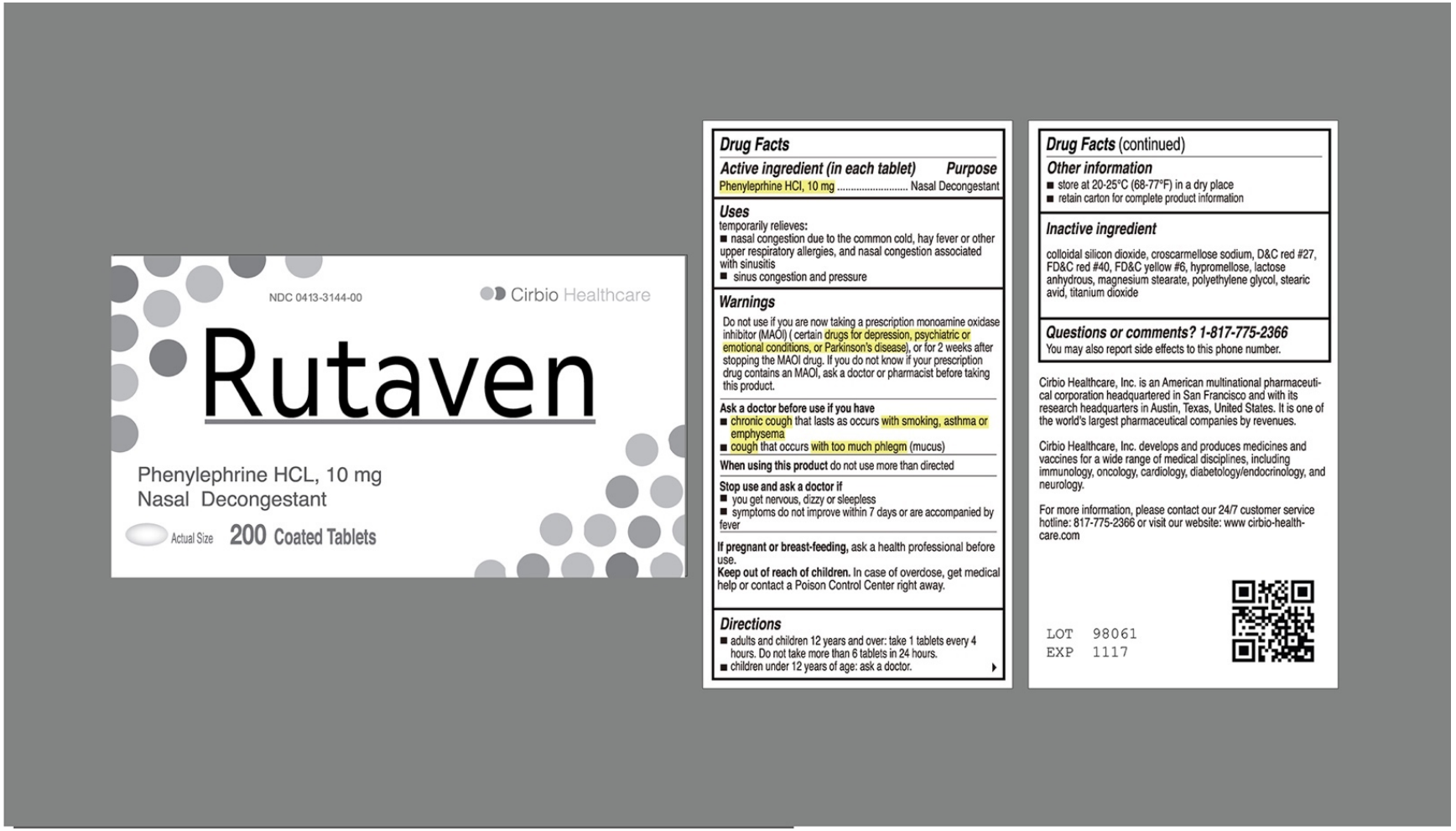


Figure A.44 AI Change on PDP

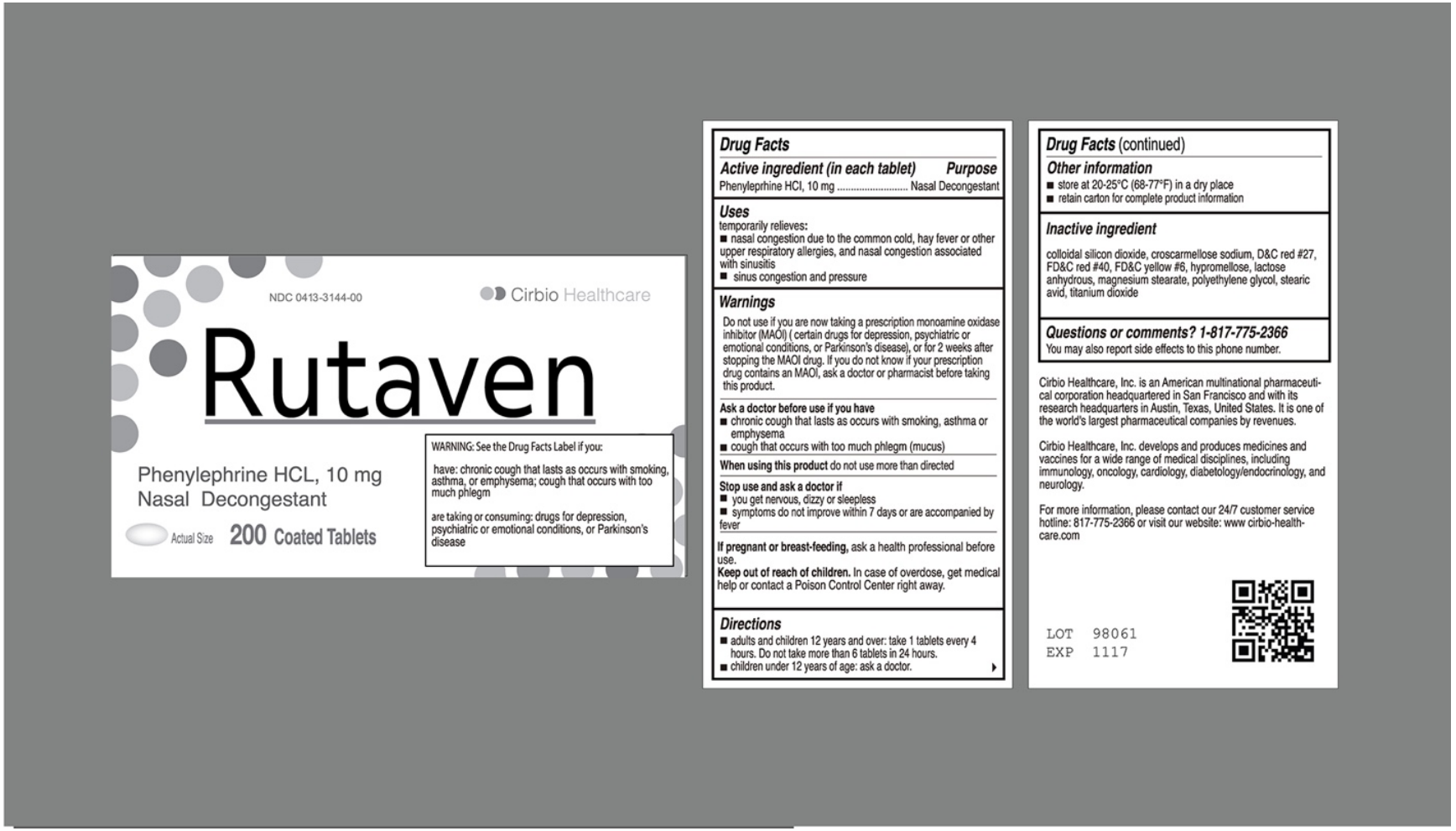


Figure A.45 Example Label for Phenylephrine, enhanced with the FOP Treatment

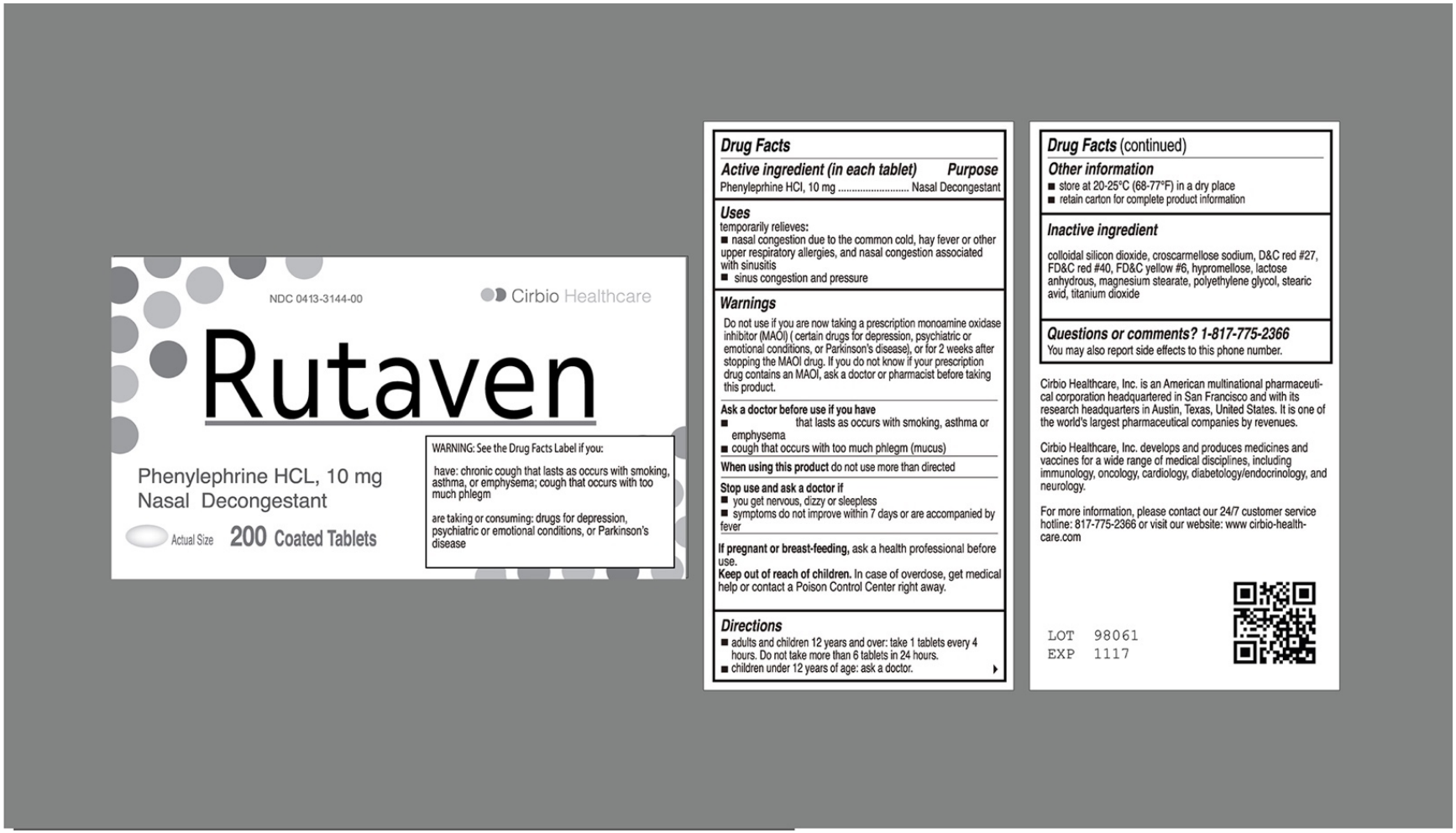


Figure A.46 DD Change in DFL

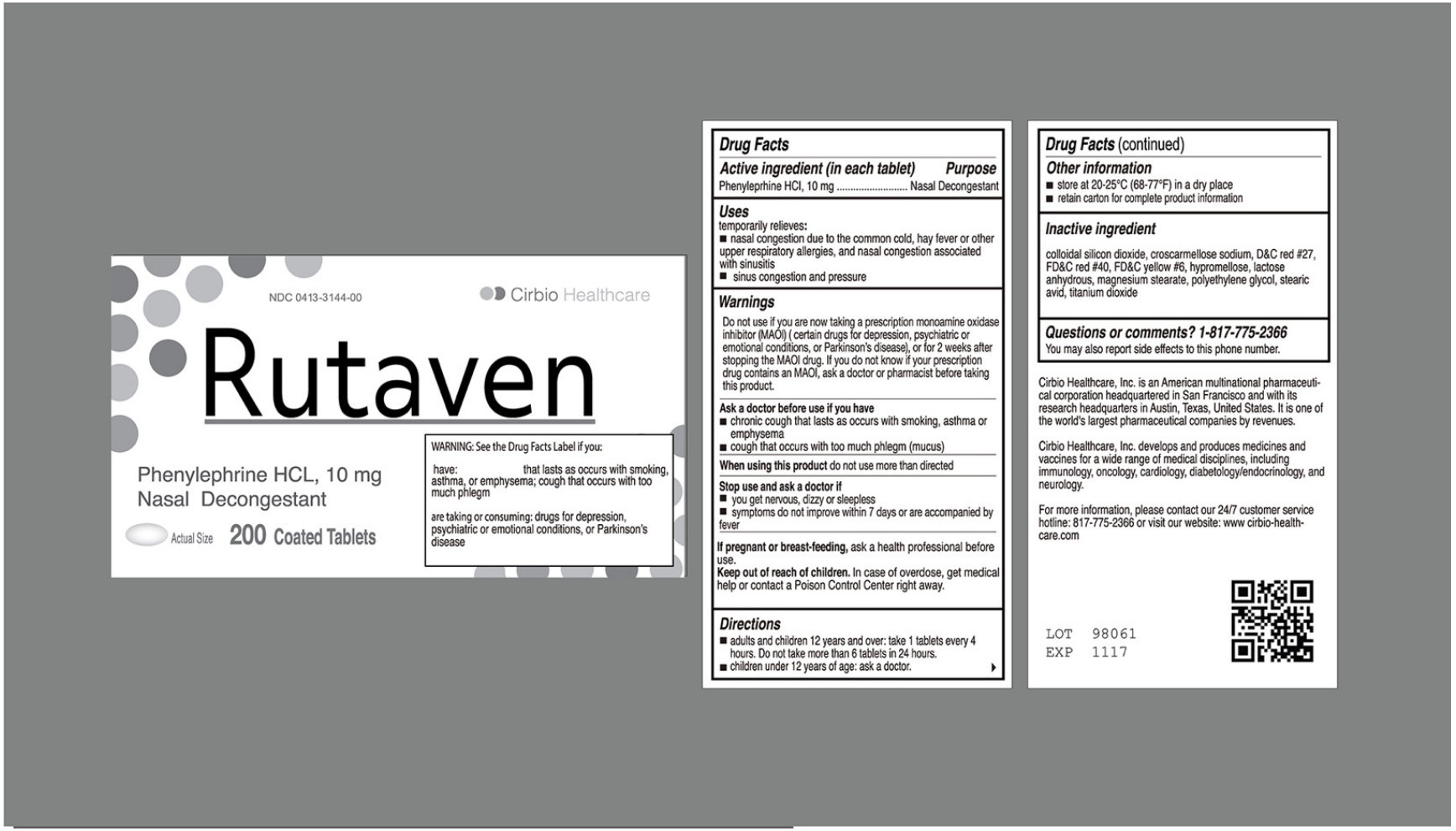


Figure A.47 DD Change on PDP

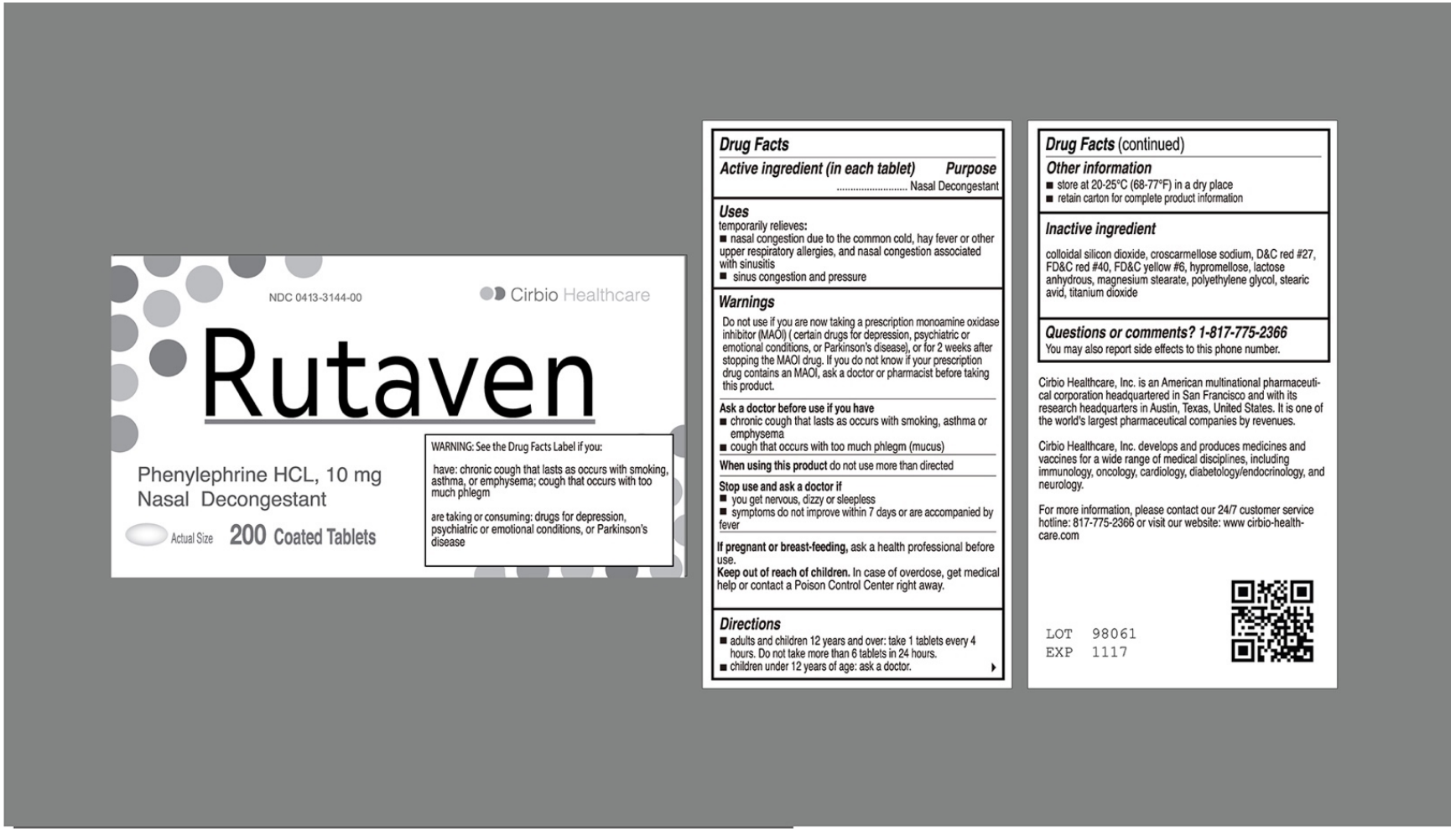


Figure A.48 AI Change in DFL

Figure A.49 AI Change on PDP

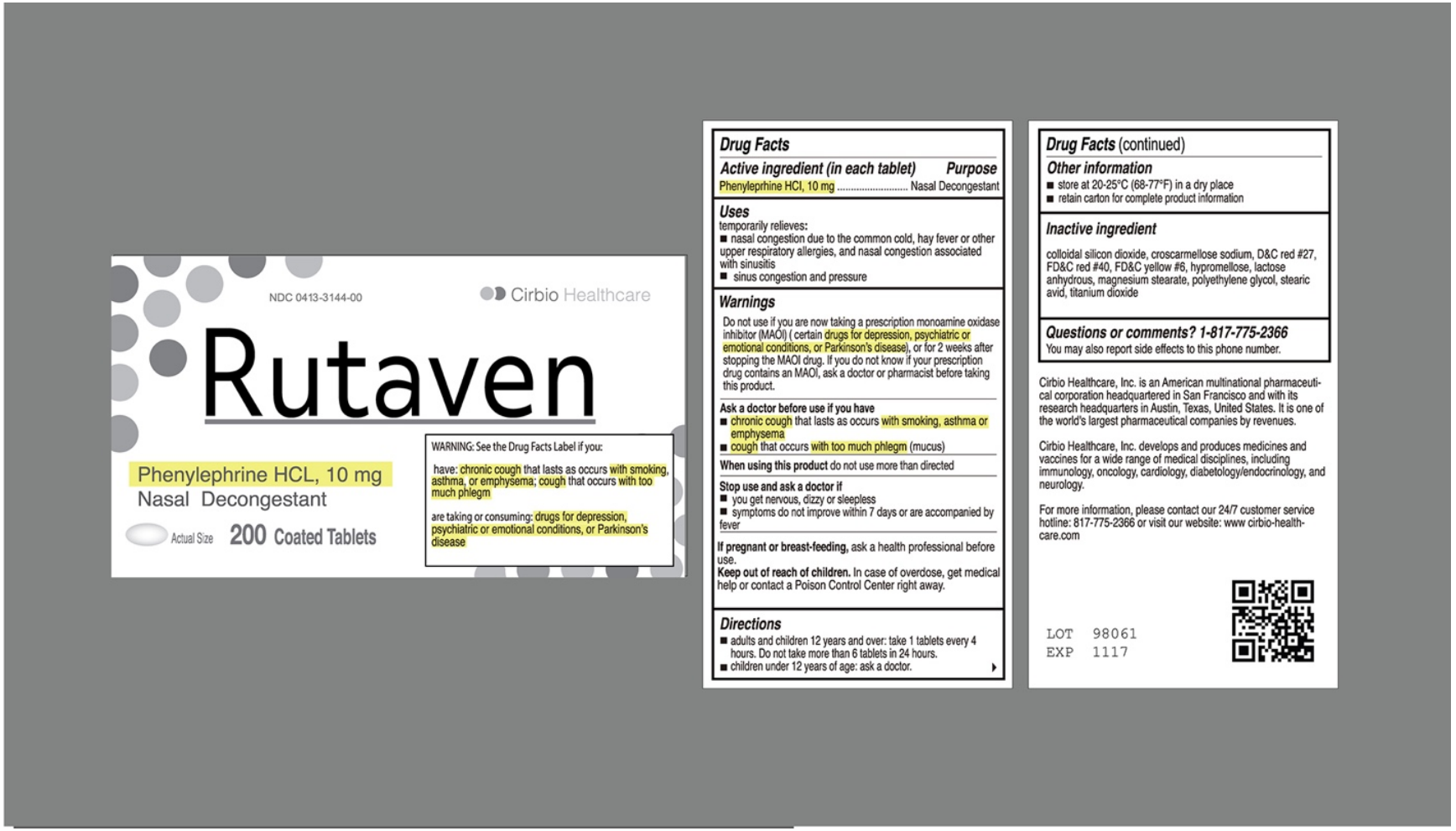


Figure A.50 Example Label for Phenylephrine, enhanced with both the FOP or Highlight Treatment

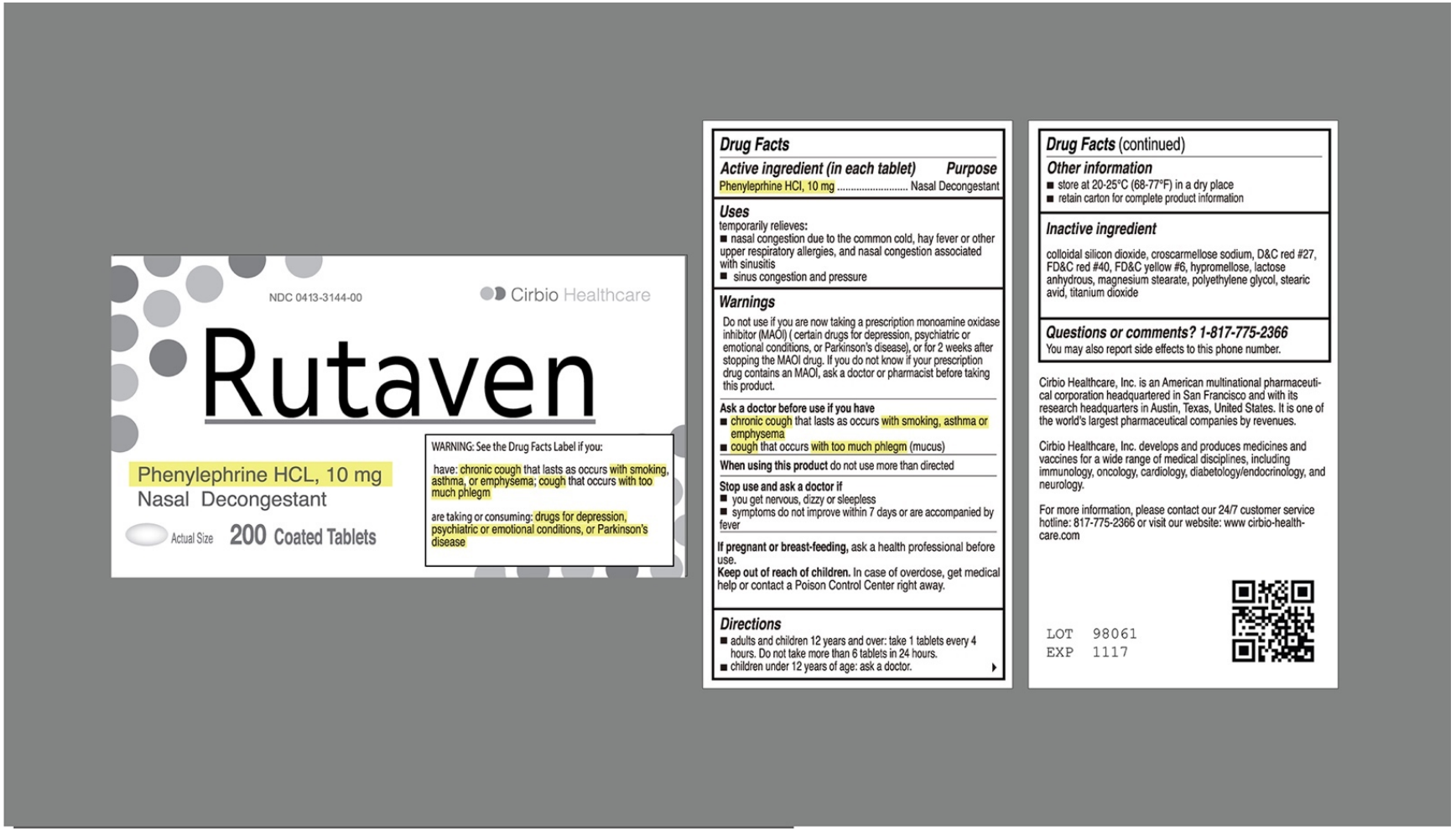


Figure A.51 DD Change in DFL

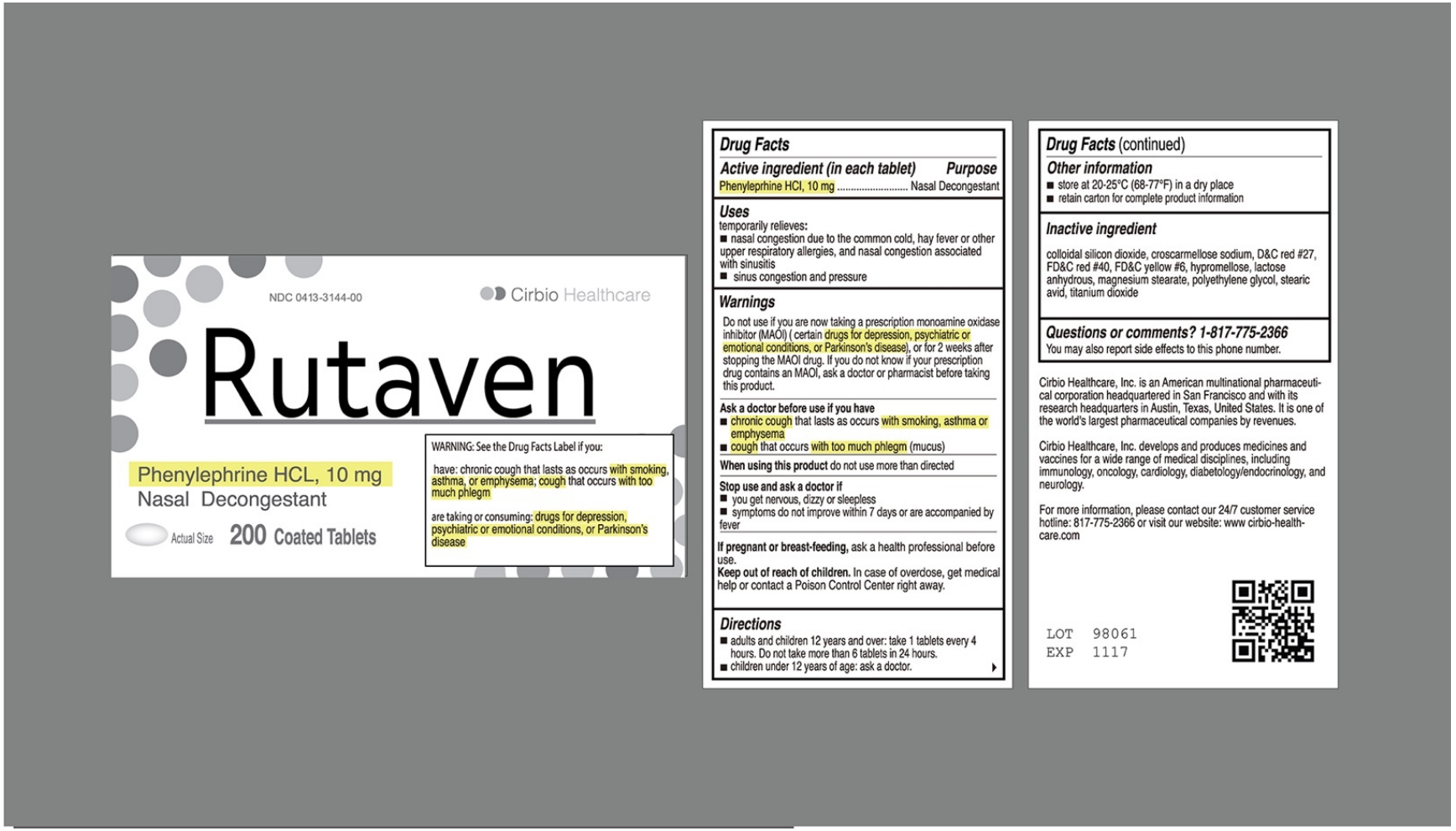


Figure A.52 DD Change on PDP

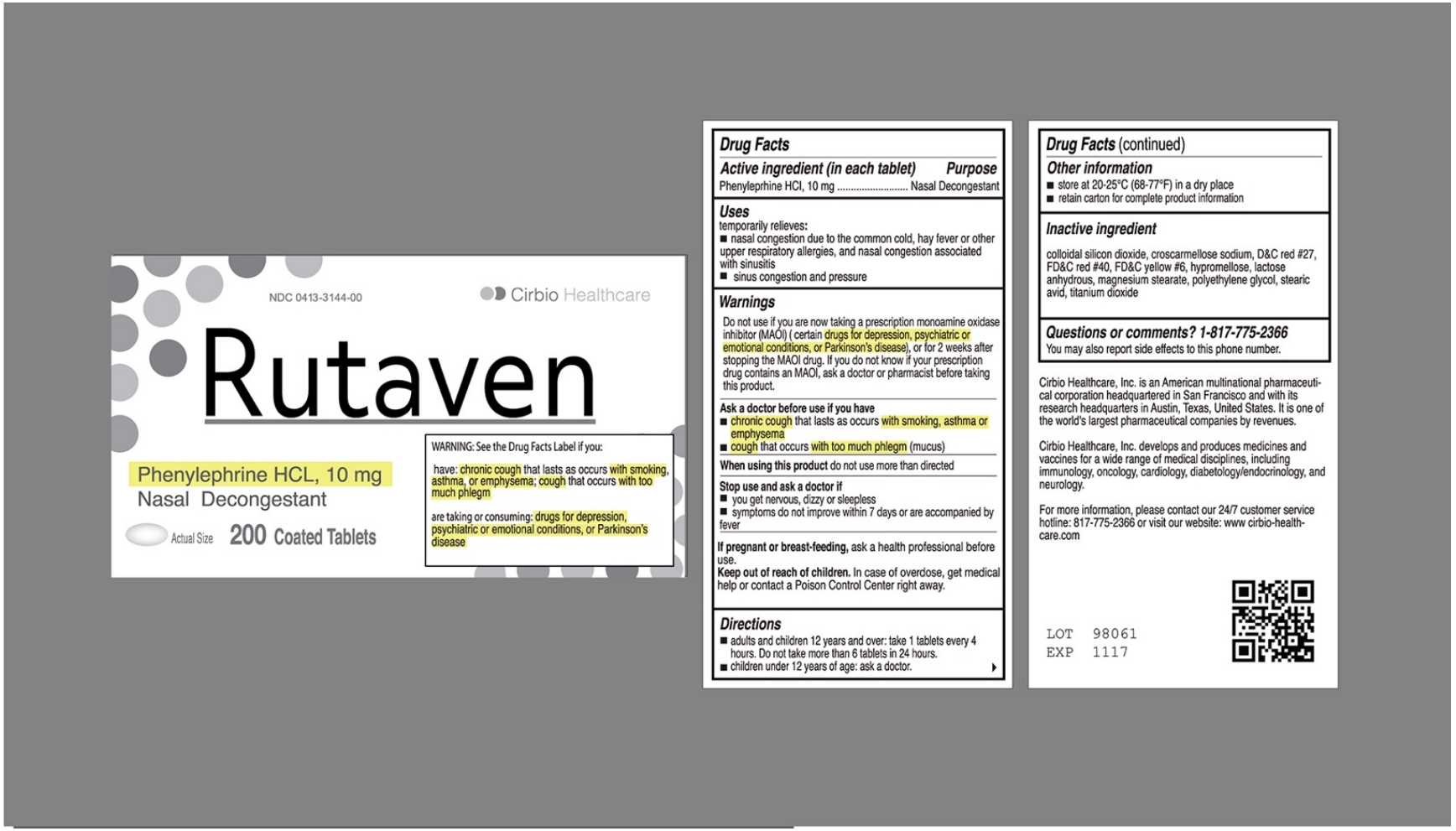


Figure A.53 AI Change in DFL

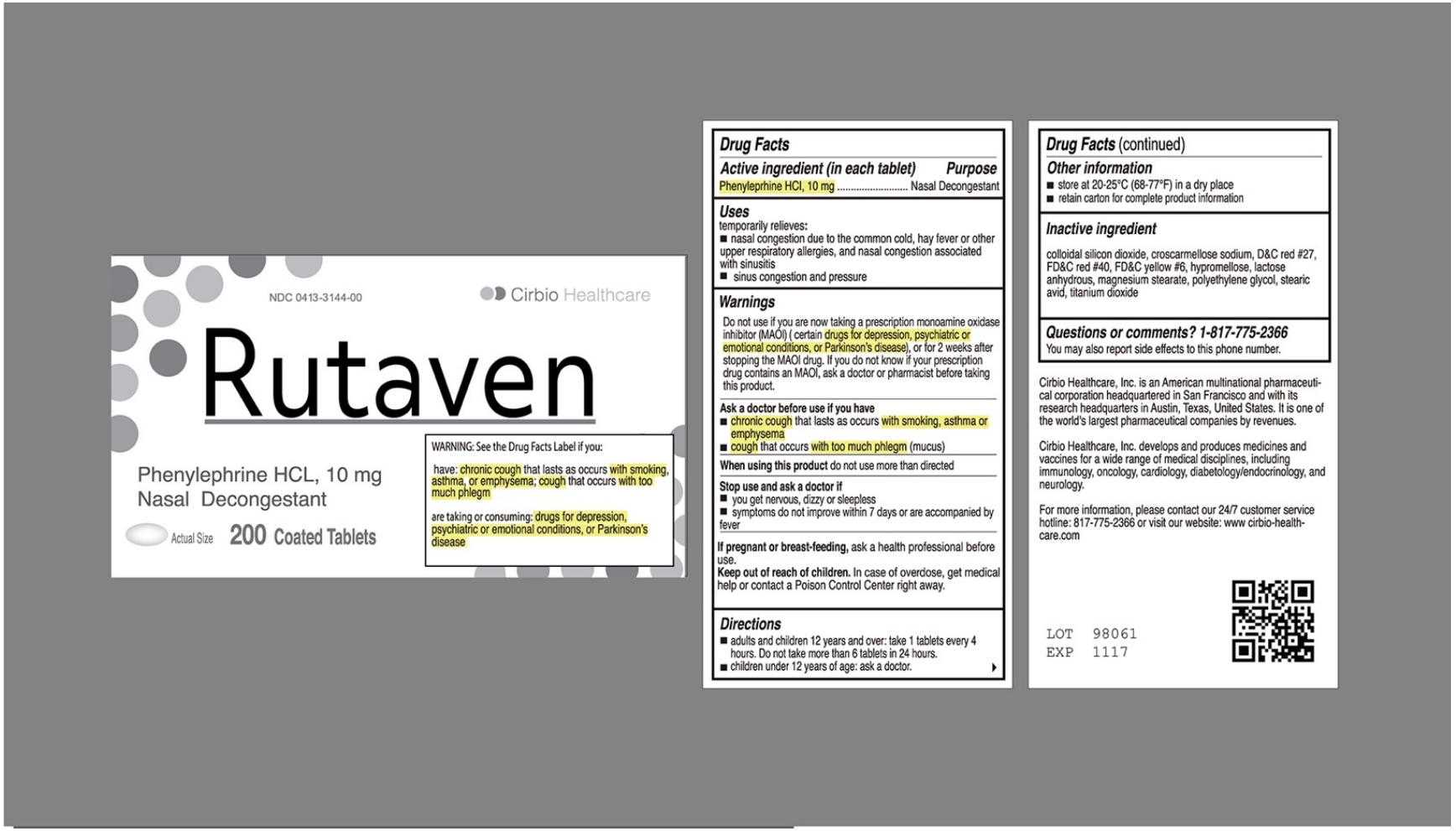


Figure A.54 AI Change on PDP

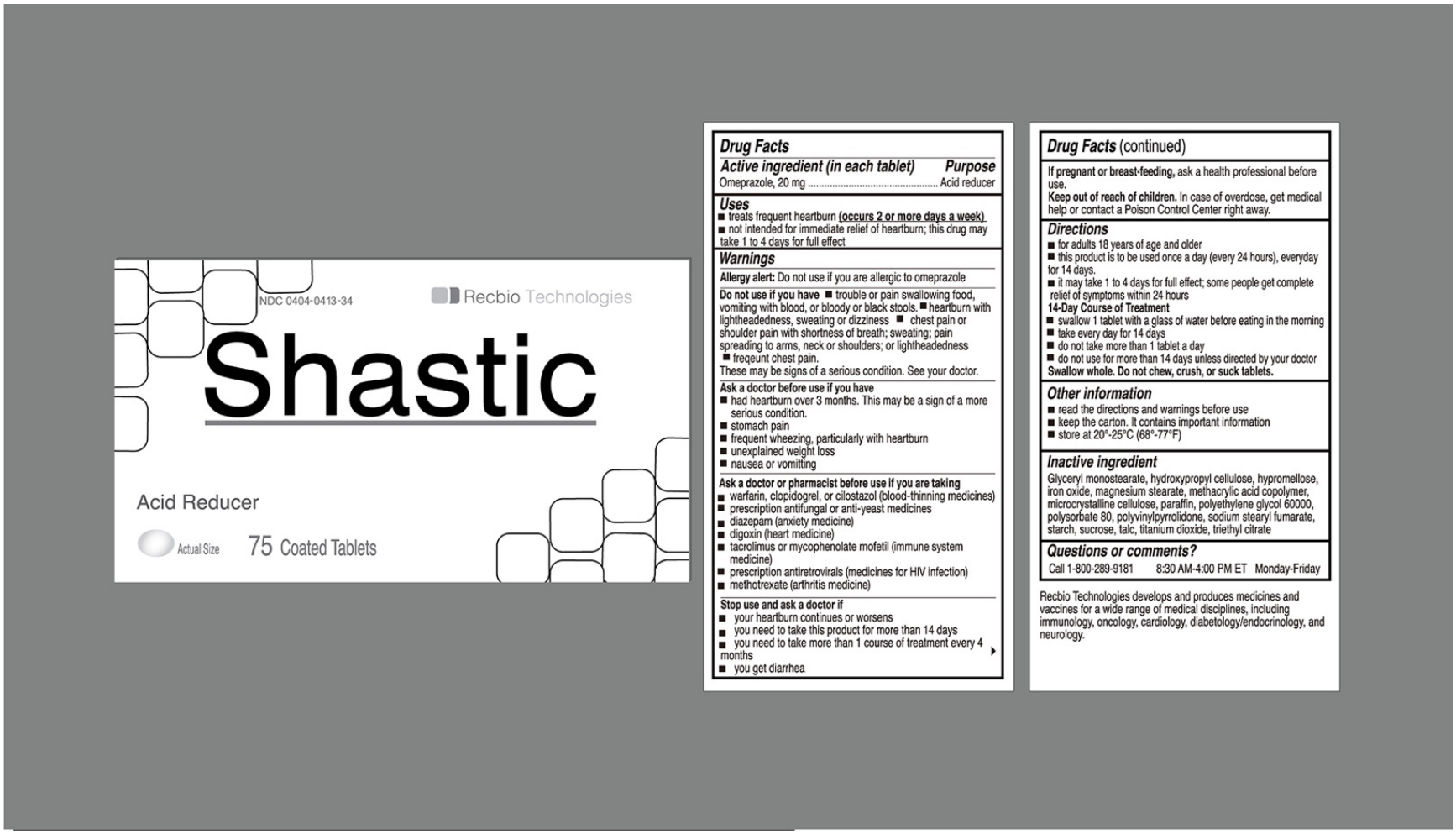


Figure A.55 AI Change on PDP

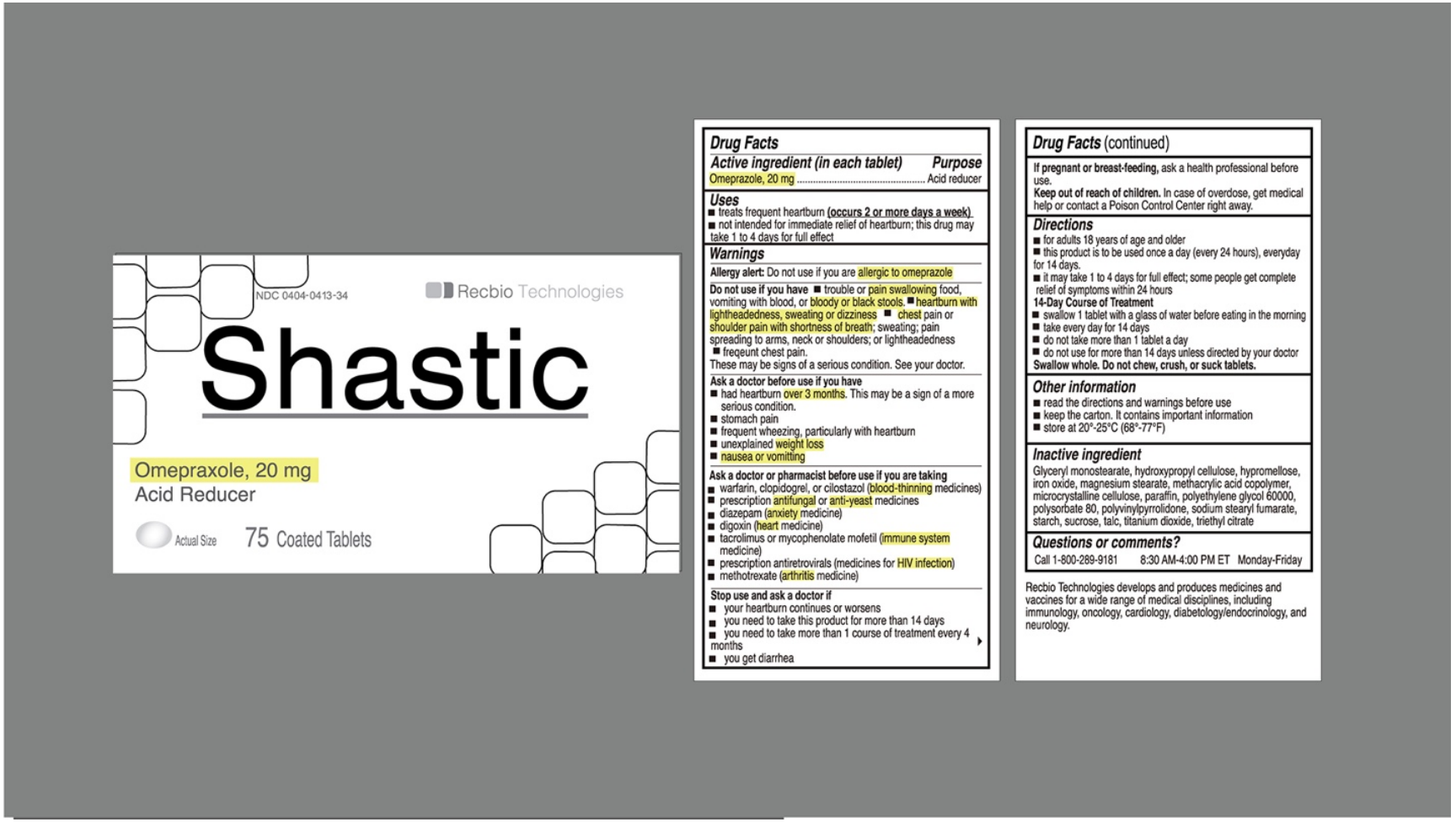


Figure A.56 Example Label for Omeprazole, enhanced with the Highlight Treatment

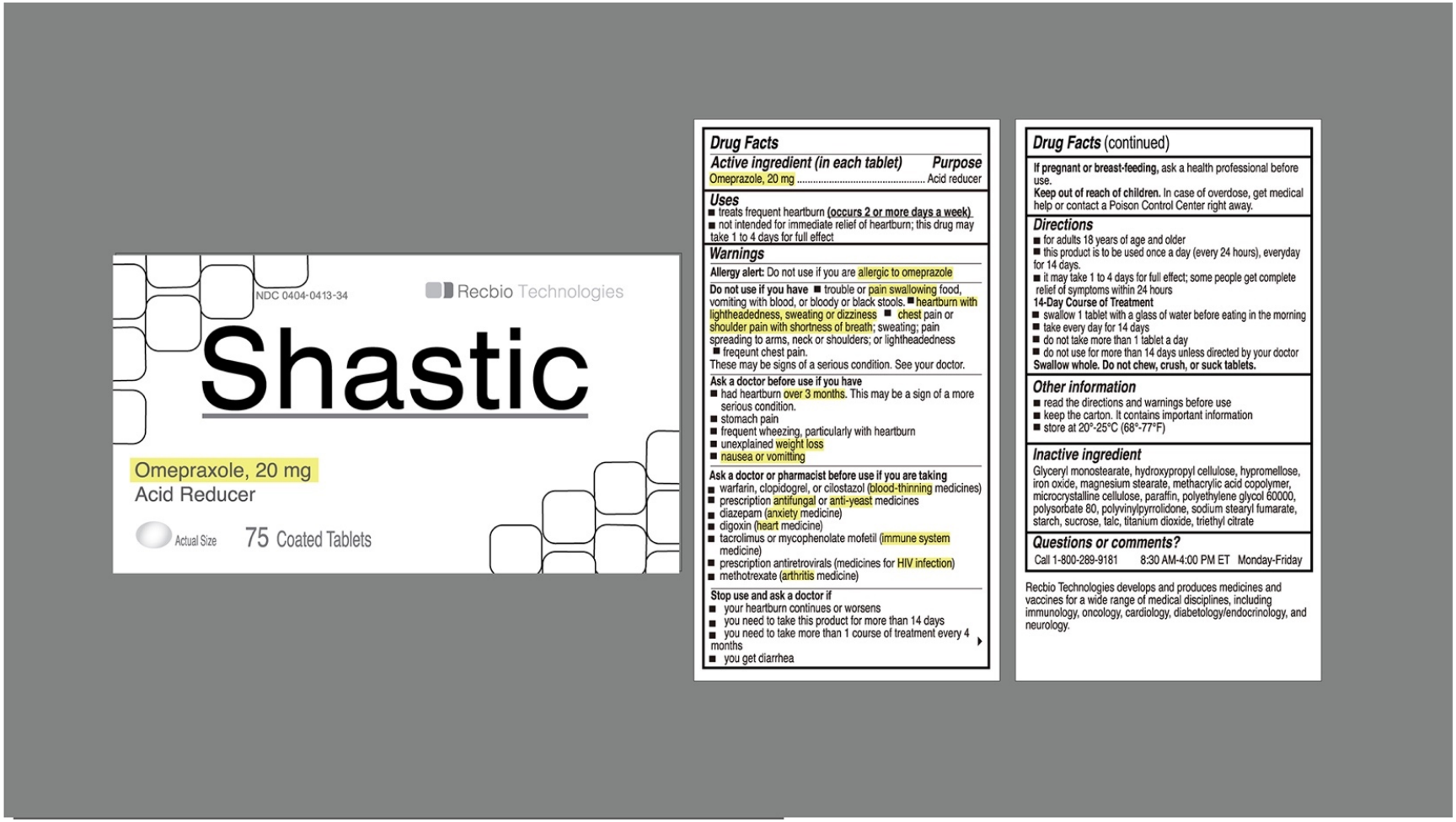


Figure A.57 DD Change in DFL

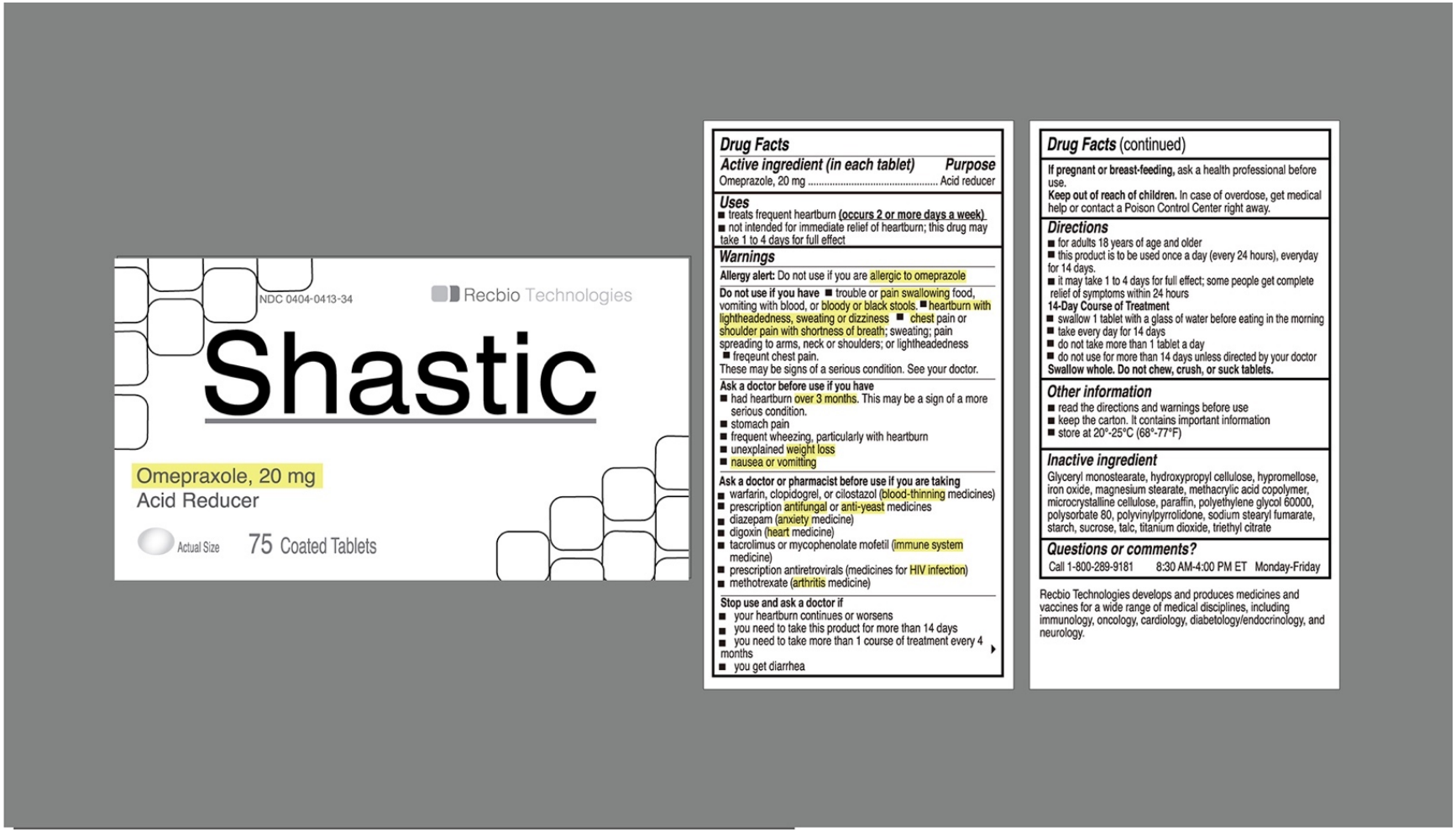


Figure A.58 AI Change in DFL

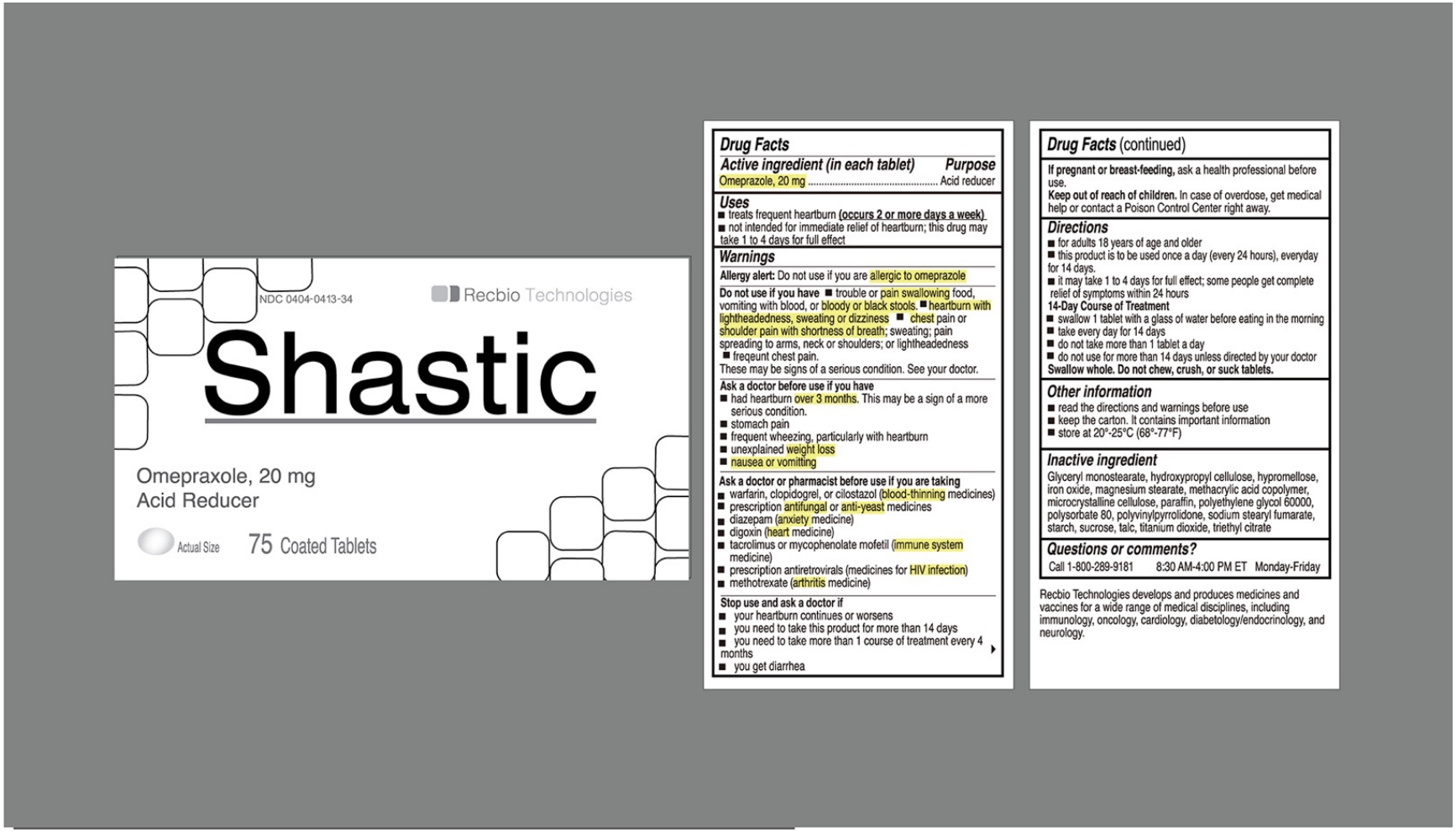


Figure A.59 AI Change on PDP

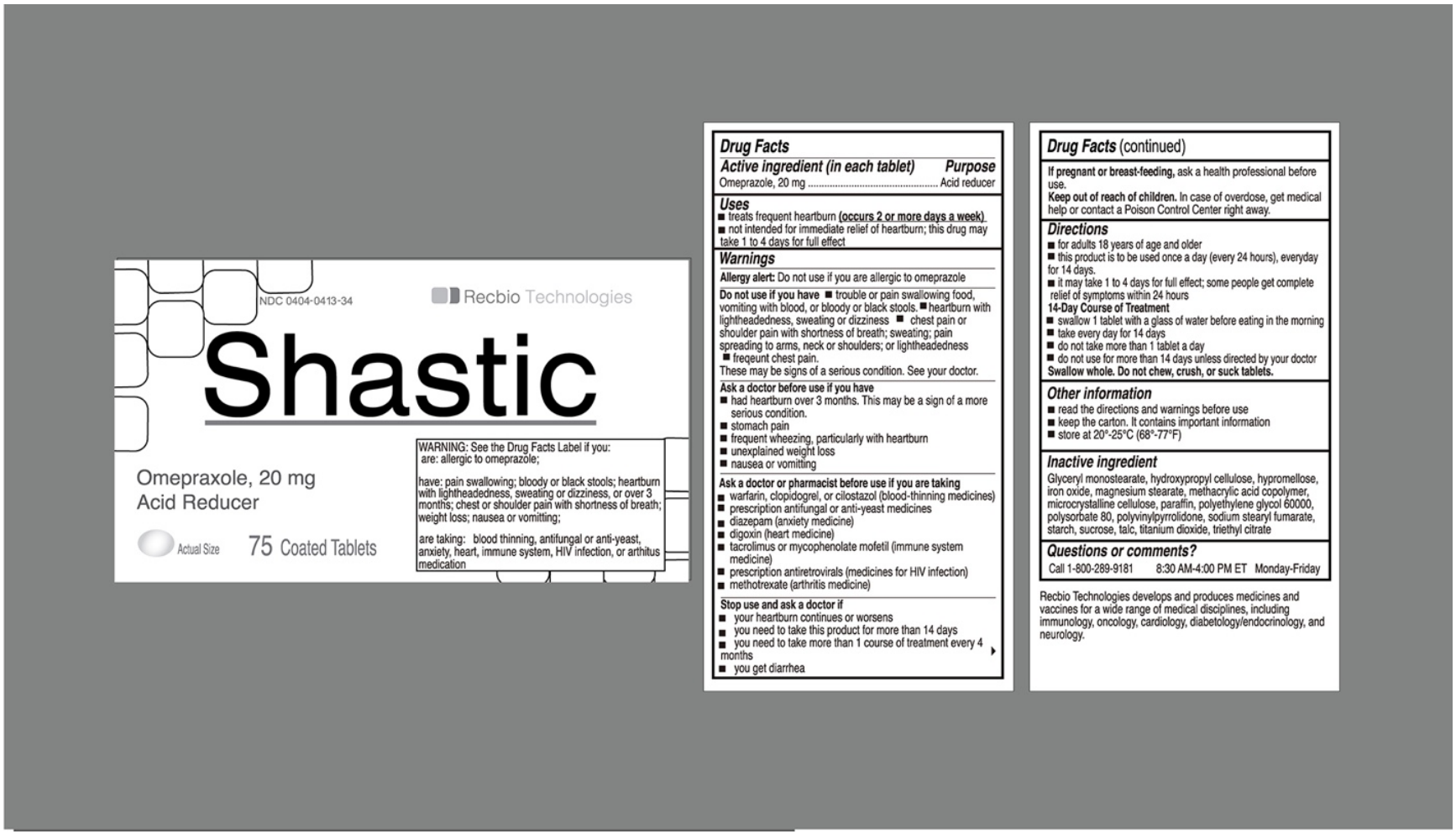


Figure A.60 Example Label for Omeprazole, enhanced with the FOP Treatment

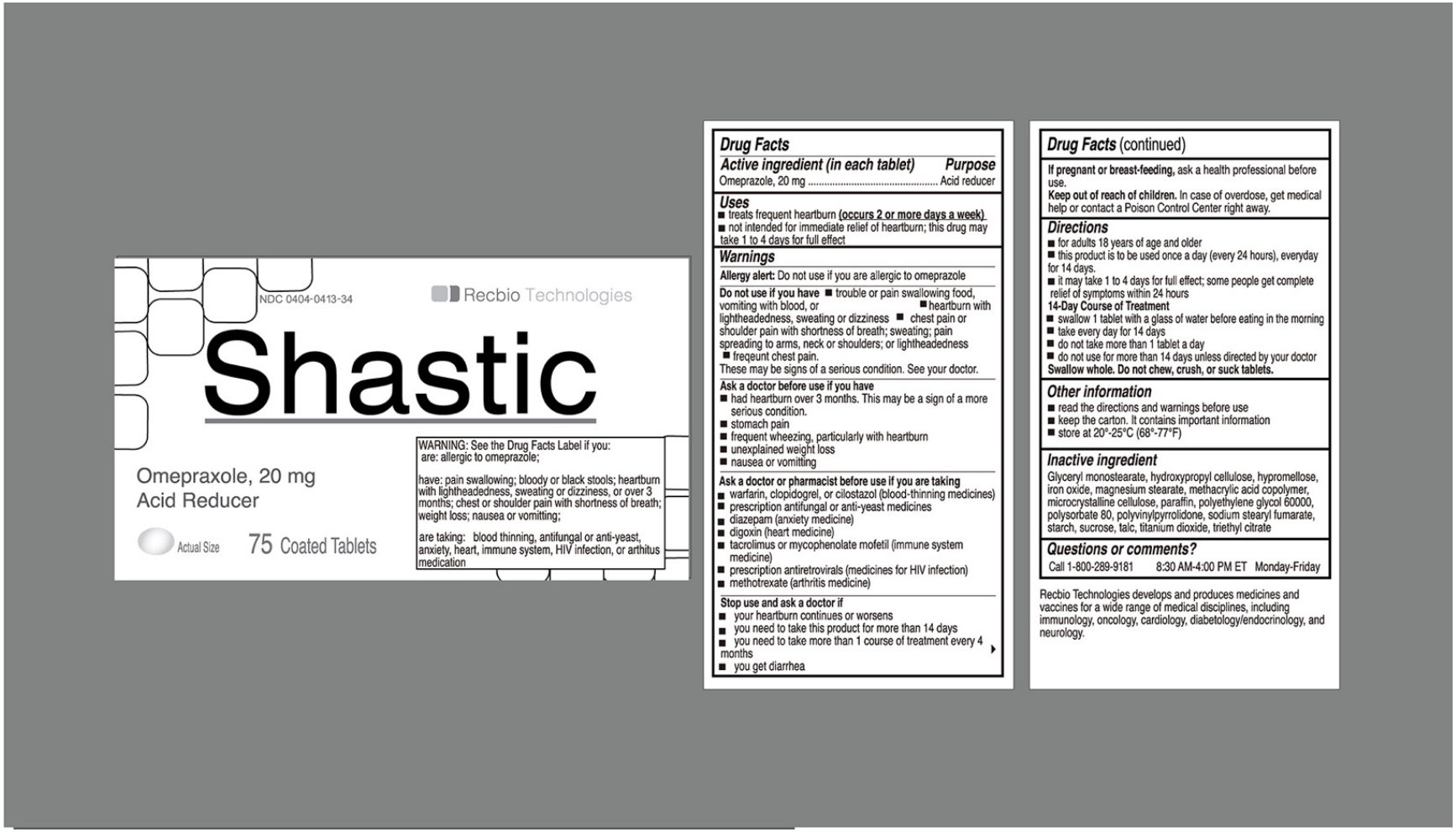


Figure A.61 DD Change in DFL

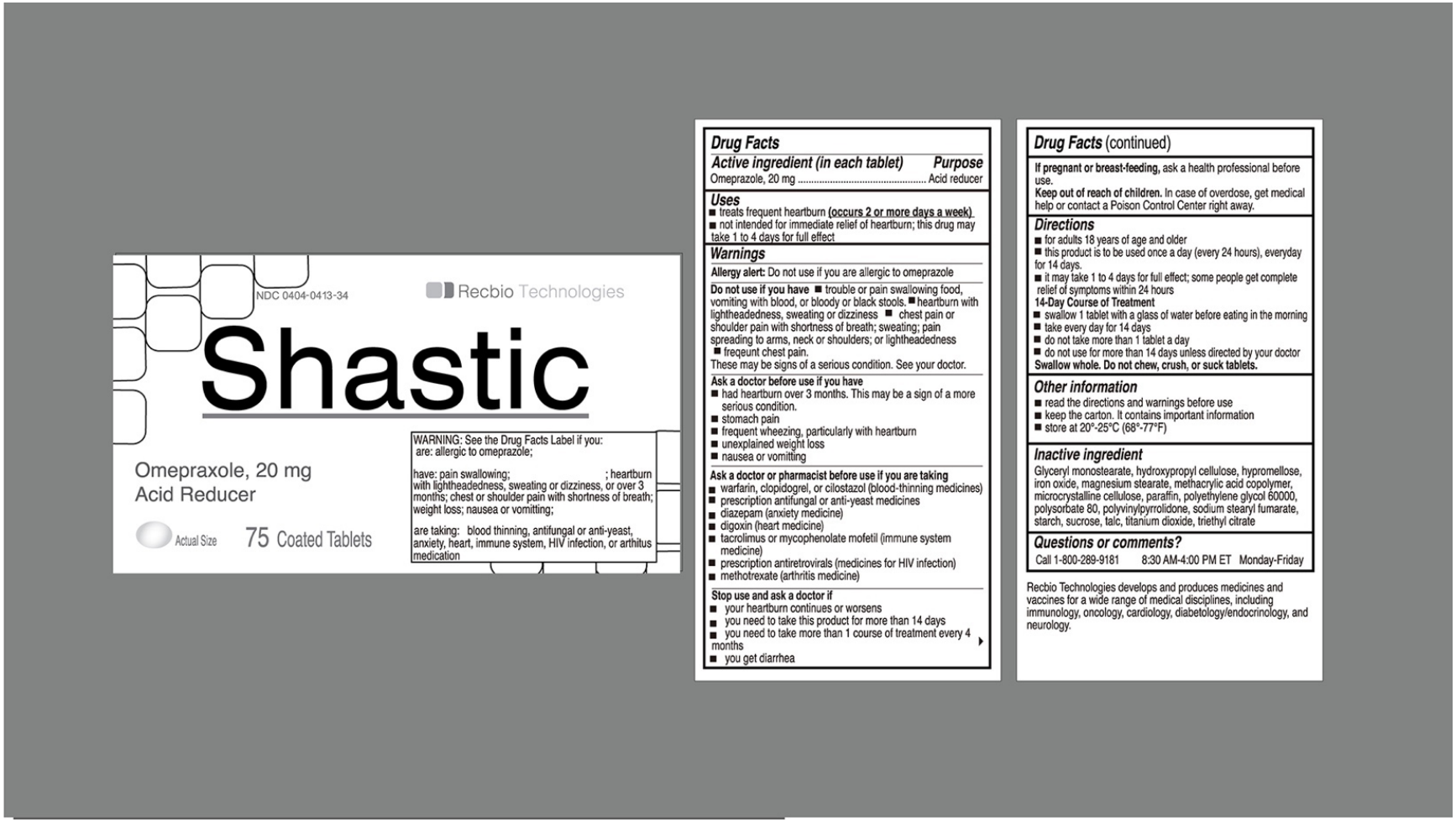


Figure A.62 DD Change on PDP

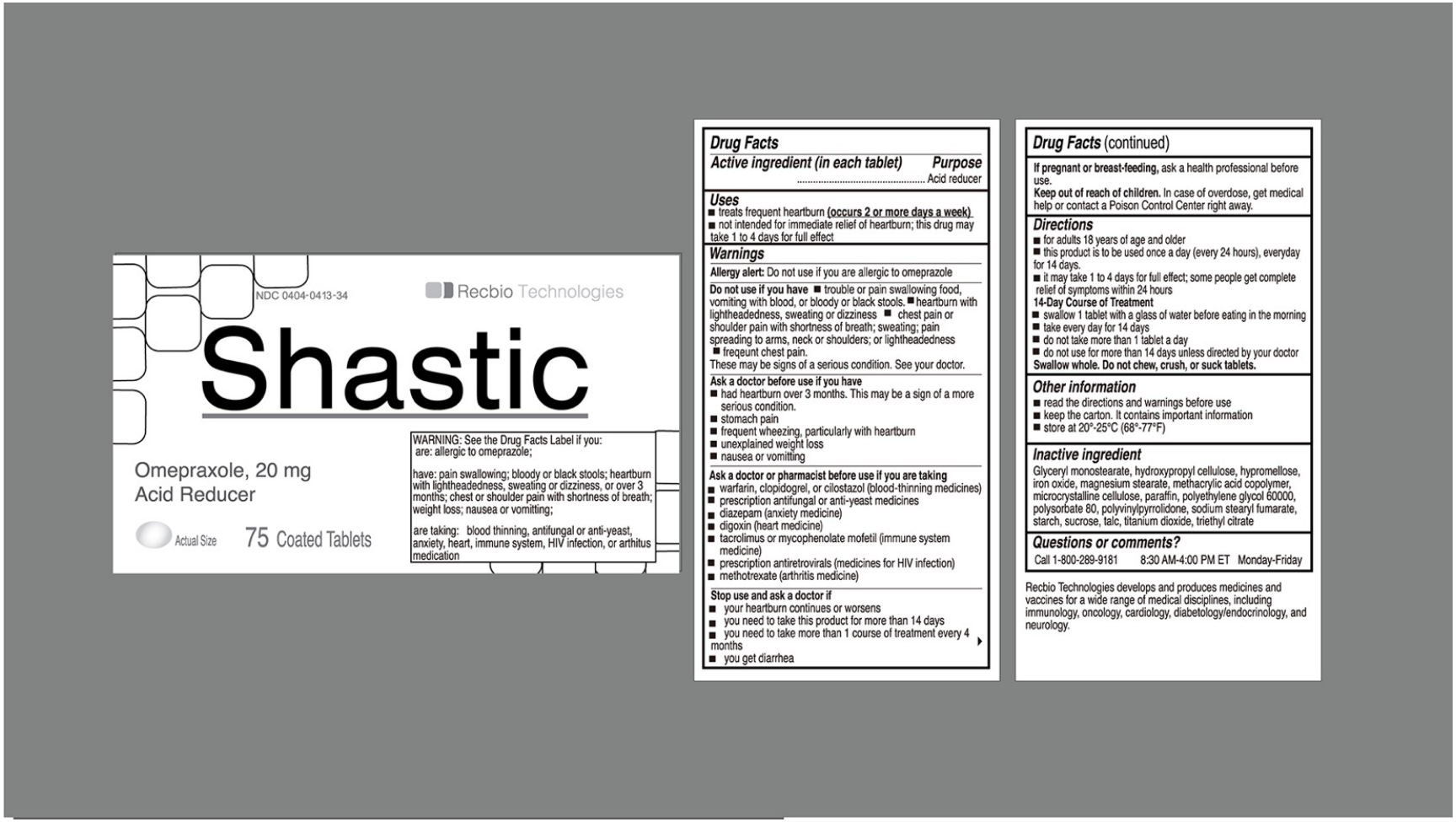


Figure A.63 AI Change in DFL

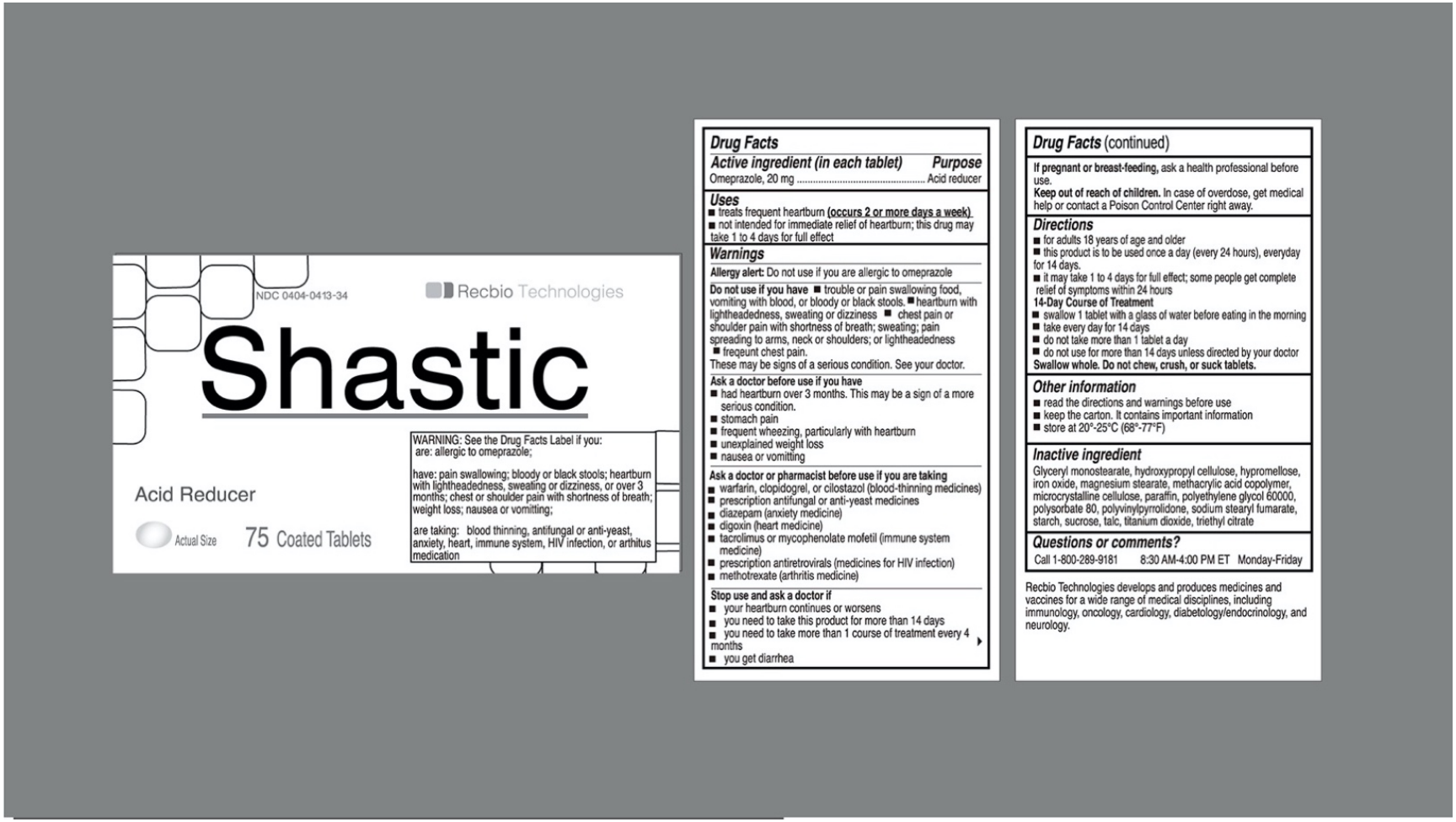


Figure A.64 AI Change on PDP

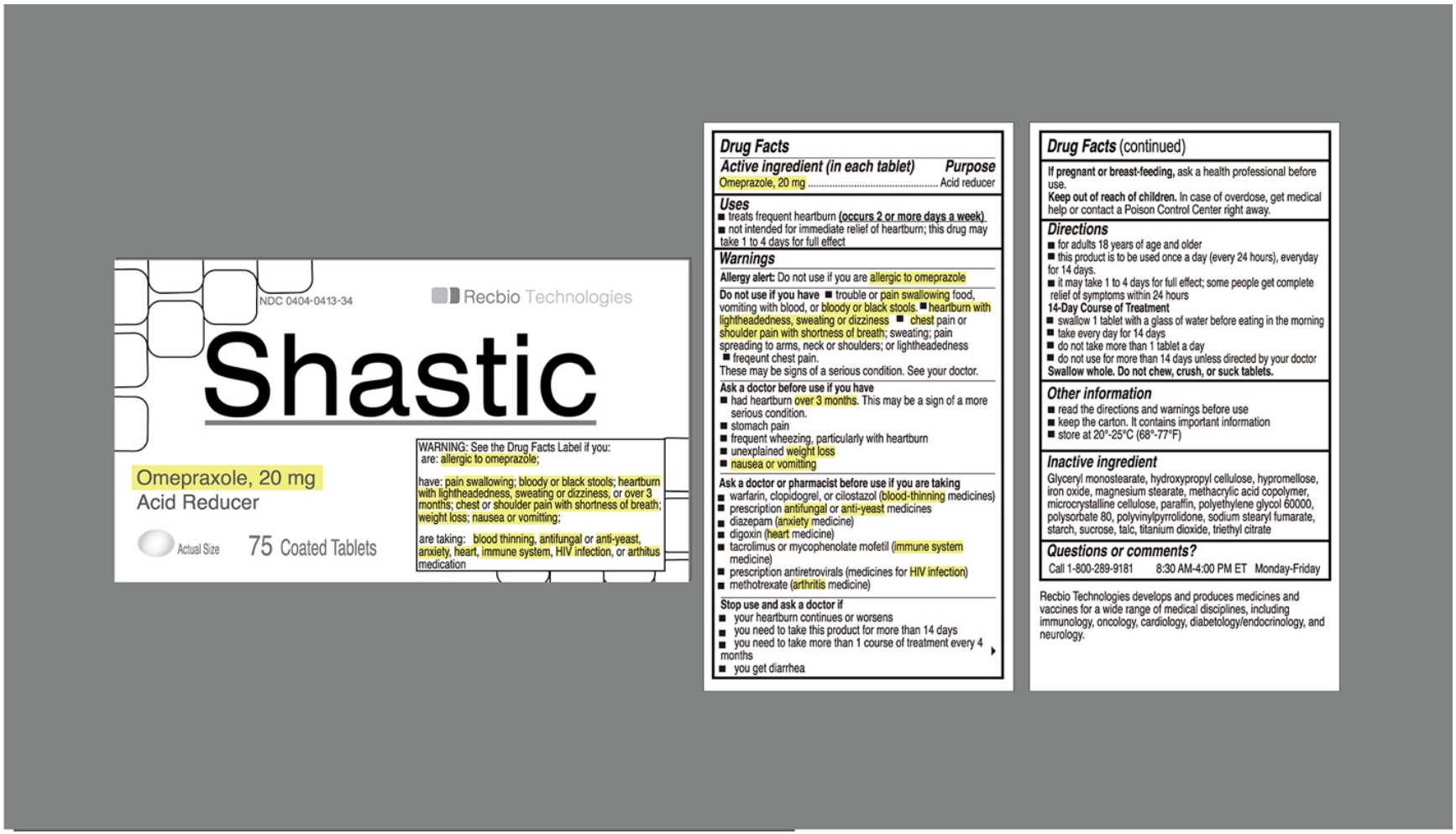


Figure A.65 Example Label for Omeprazole, enhanced with both the FOP and Highlight Treatment

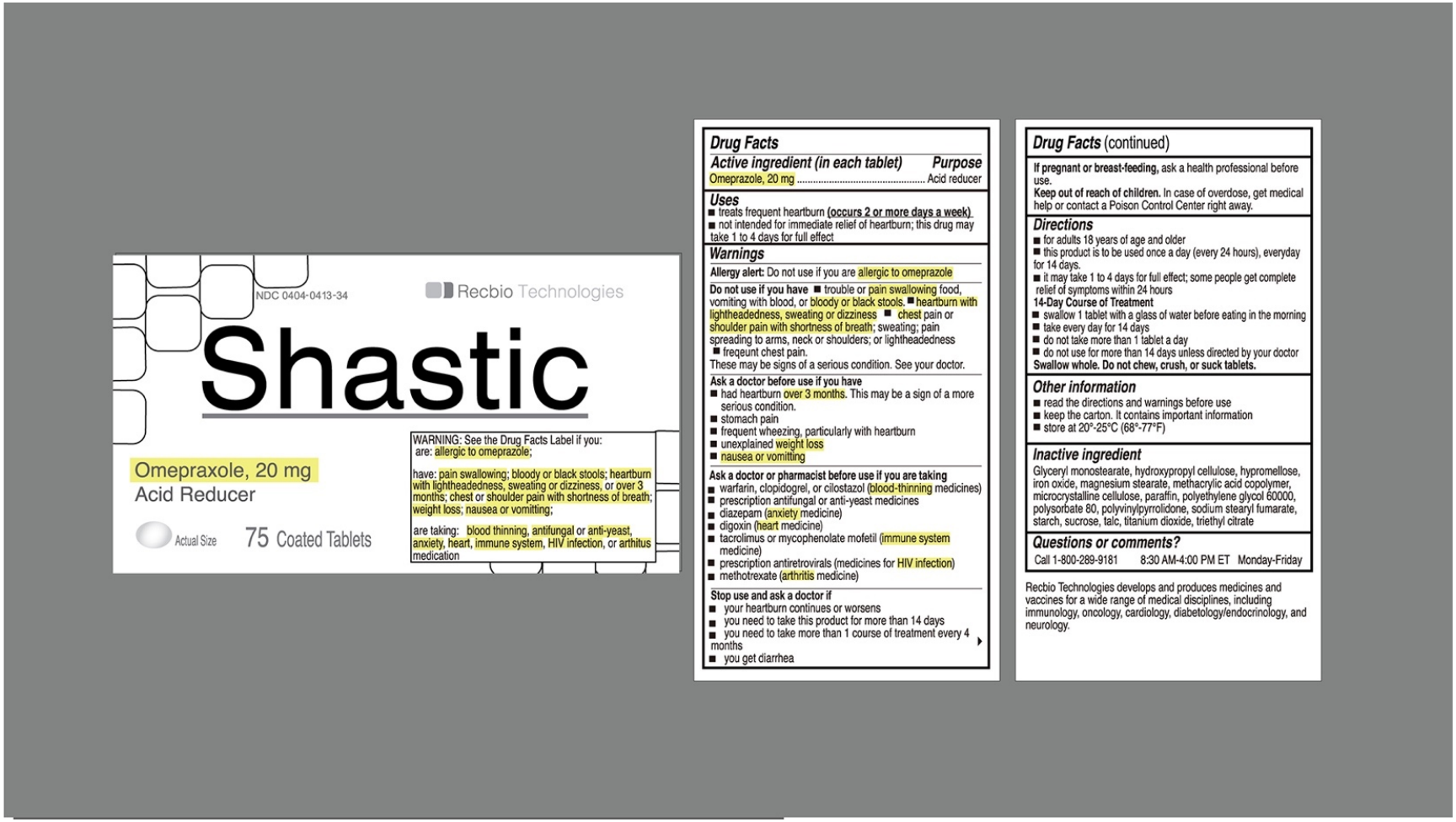


Figure A.66 DD Change in DFL

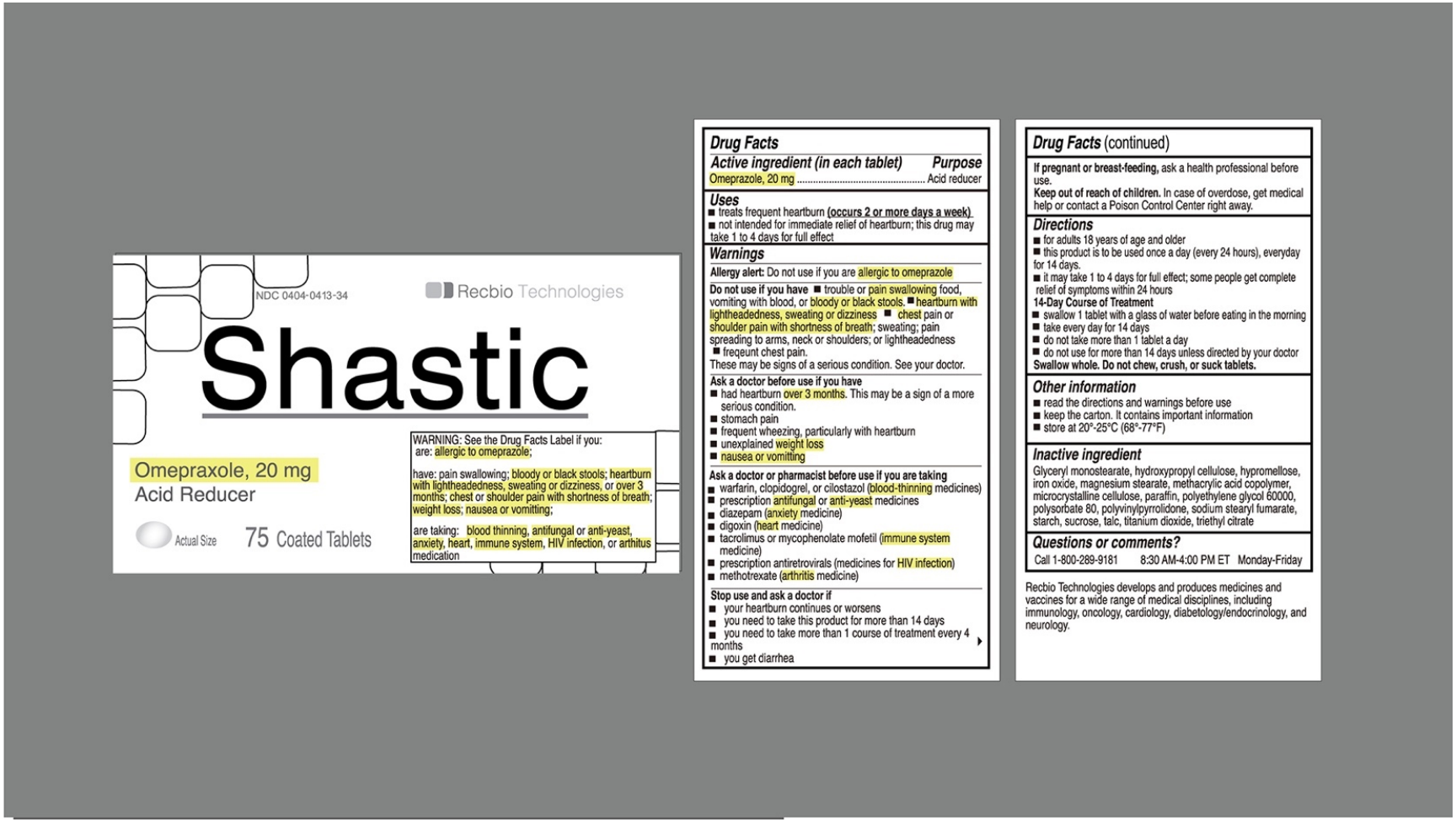


Figure A.67 DD Change on PDP

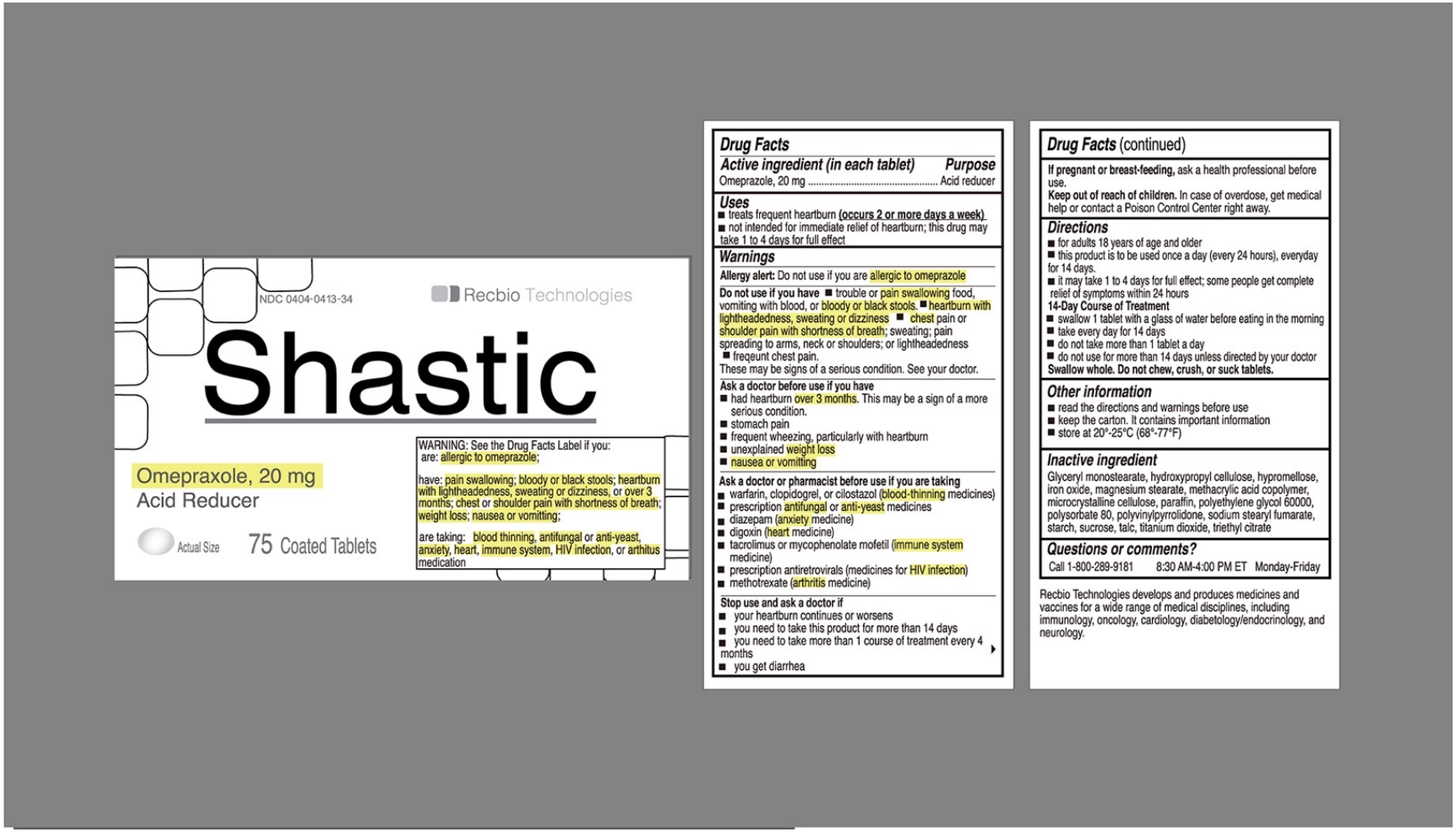


Figure A.68 AI Change in DFL

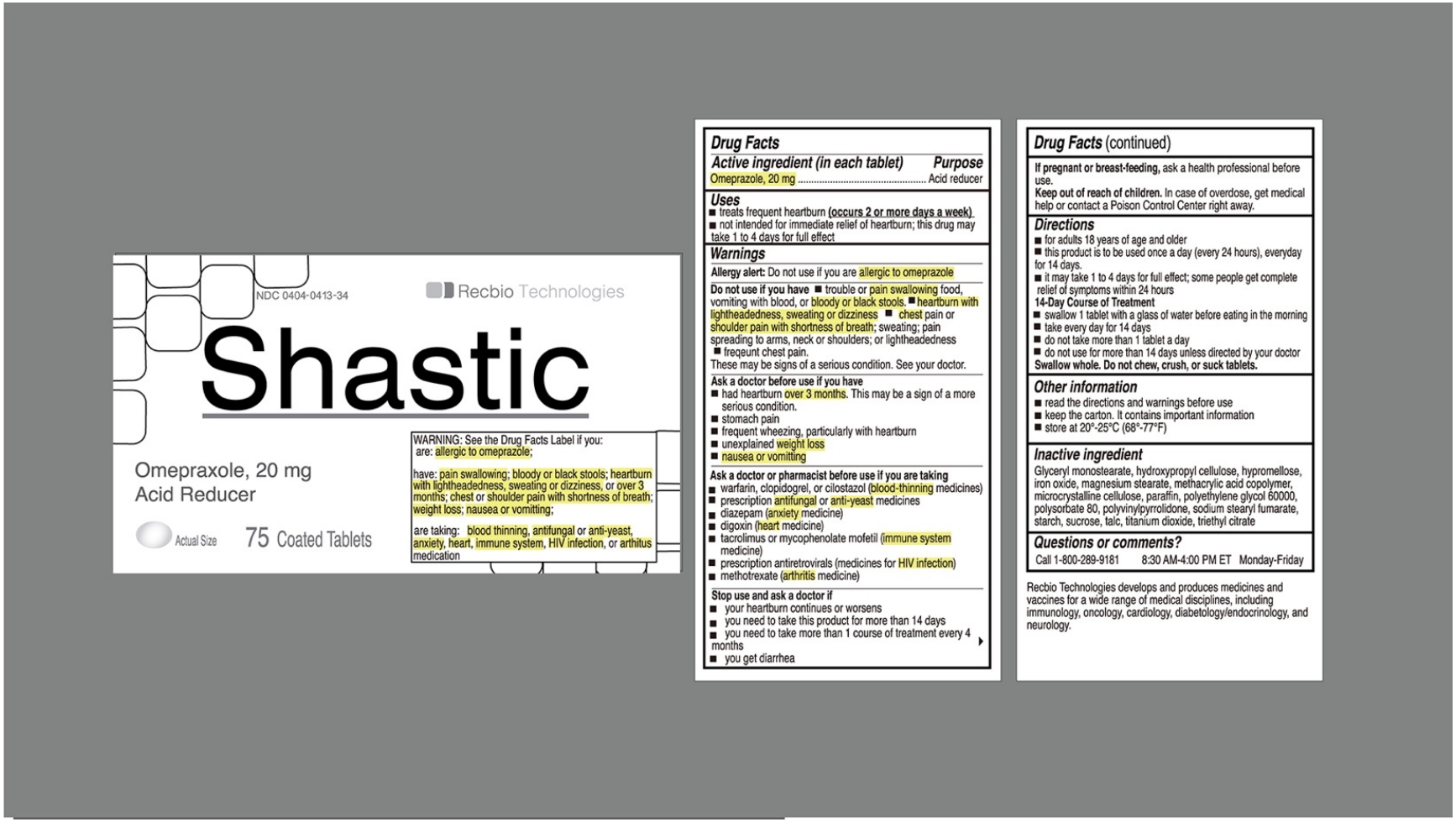


Figure A.69 AI Change on PD

APPENDIX B:

A rationale for the selection of information to highlight or include in the front of package warning

In order to standardize the selection process for what information should be highlighted or brought to the front of package when developing experimental stimuli the following method was developed. With the overarching goal of this research being a reduction of preventable ADEs attributed to OTC drugs, the information highlighted must be relevant to halting the purchase of a drug that is inappropriate for personal use due to either pre-existing or comorbid conditions or other courses of treatment. The types of information to be highlighted are: the active ingredient, warnings related to the OTC drug being contraindicated for a pre-existing diagnosis or conditions, and warnings related to drug-drug interactions between the OTC drug and other common medications. The active ingredient will be highlighted in the DFL and on the PDP.

The term diagnosis used within this context is referring to clinical diagnosis done by a medical professional examining the physical signs, symptoms, and test results and then interpreting those for the patient (Llewelyn, Ang, Lewis, & Al-Abdullah, 2007). Preexisting medical conditions, which might be considered risk factors, such as age, weight, or specific comorbid symptoms, while not considered a diagnosis in this framework, are considered conditions relevant to the prevention of an inappropriate purchase and would be highlighted. Thus, if a warning was indicated for patients with "high blood pressure" that would be considered a diagnosis warning and highlighted, a warning was for people "over age 60" that would be considered a condition warning and would be highlighted, but general statements such as "under a doctor's care for any serious condition" would not be specific enough to be highlighted.

A simplified checklist of questions used to determine whether something would be considered a drug diagnosis contraindication warning for highlighting purposes is:

- ✓ Would noticing this warning (if applicable) prevent the purchase of an OTC drug that is inappropriate for the patient to safely consume?

- ✓ Does this warning feature a specific condition, symptom or diagnosis the consumer would recognize themselves as having (if application)?
- ✓ Is this a redundant warning based on other information already selected for highlighting?

If the answer to each of these questions is yes, the requirements are met and the warning is eligible for highlighting.

Drug-drug interaction warnings specifically refer to a warning that explicitly calls out another specific drug class or drug by name. The most broad category will be highlighted if both a specific drug and the drug class are listed. Some DFLs for specific active ingredients include additional warnings about accidental overdose. In order to not confuse drug-drug interaction with over-dose warnings, only the drug interaction warnings focused on taking an additional medication concurrently will be highlighted. Thus, a warning for patients "taking Warfarin or other blood thinners", " blood thinners" would be highlighted, but a warning of "do not take more than directed" would not be highlighted.

A checklist of questions used to determine whether something would be considered a drug-drug interaction warning for highlighting purposes is:

- ✓ Would noticing this warning (if applicable) prevent the purchase of an OTC drug that is inappropriate for the patient to safely consume?
- ✓ Does this warning feature a specific medicine or class of medicines by name?

If the answer to both of these questions is yes, the requirements are met and the warning is eligible for highlighting.

All information that is highlighted in the Drug Facts Label will also be highlighted in the FOP. The highlights will be the same size on both panels. Additional context words and phrases will be included to make the FOP warnings understandable, but only the content words will be highlighted. Warning Information will be selected to appear in the FOP if the following criteria apply: if noticing

this warning (if applicable) would prevent the purchase of an OTC drug that is inappropriate for the patient to safely consume, AND if the warning feature a condition or diagnosis the consumer would recognize themselves as having OR the warning feature a specific medicine or class of medicines by name.

APPENDIX C:

Stimuli of each active ingredient label used in the absolute judgement and forced choice studies. An example of each mock-brand is included once, though in the studies each mock-brand appeared in all four treatments.

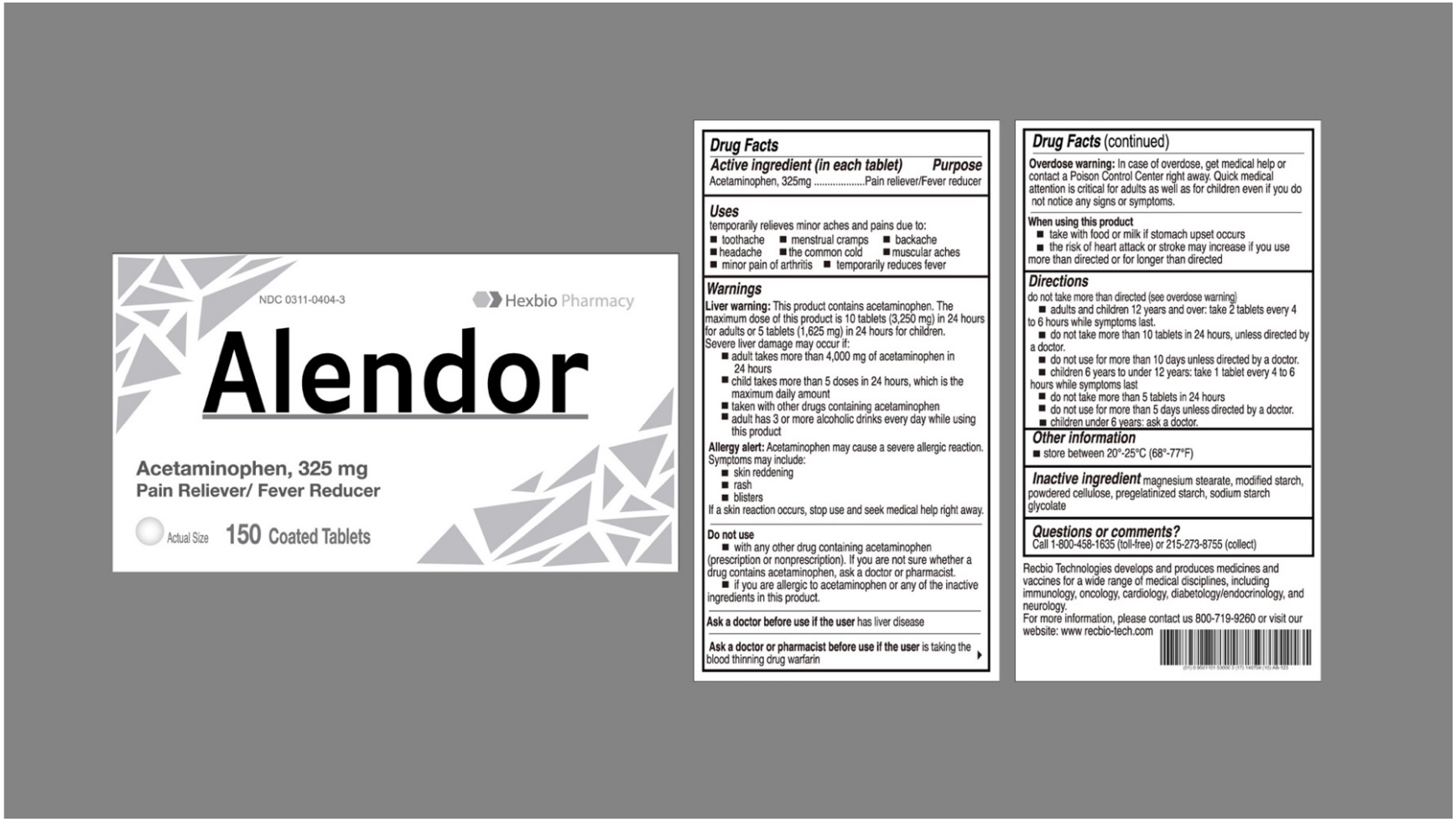


Figure C.1 Example Label for Acetaminophen, without the Highlight or FOP Warning Treatment

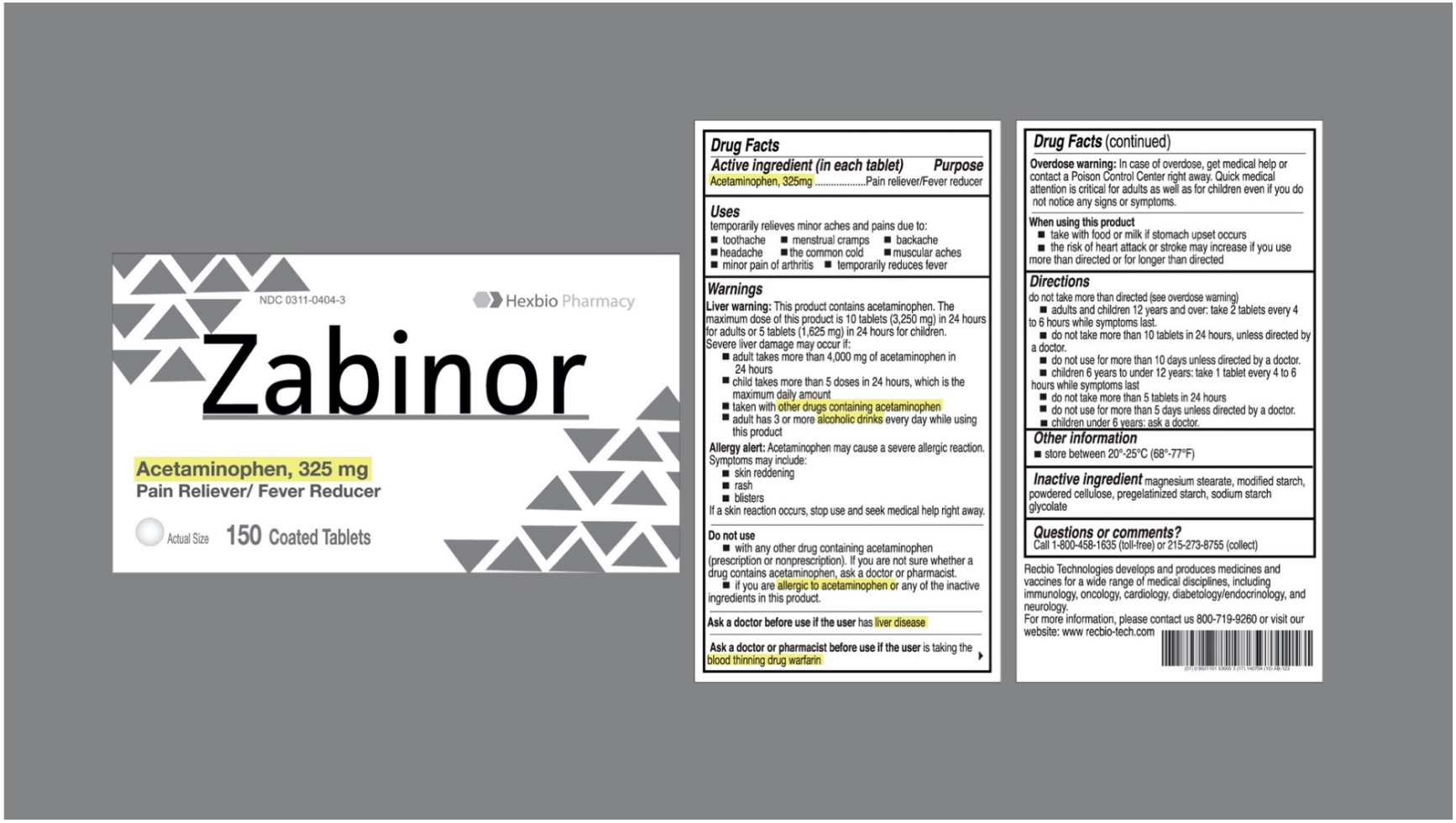


Figure C.2 Example Label for Acetaminophen, enhanced with the Highlight Treatment

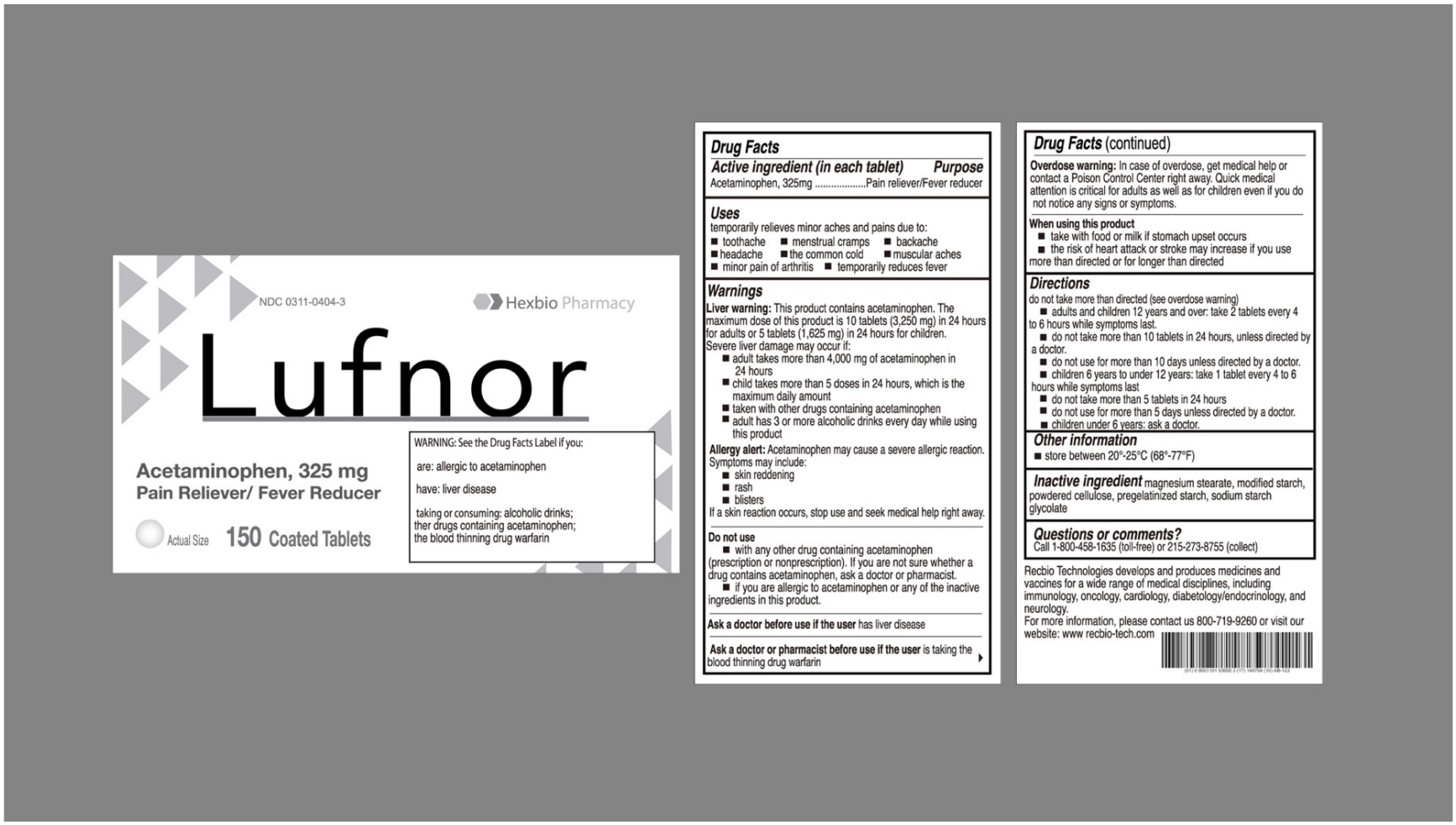


Figure C.3 Example Label for Acetaminophen, enhanced with the FOP Treatment

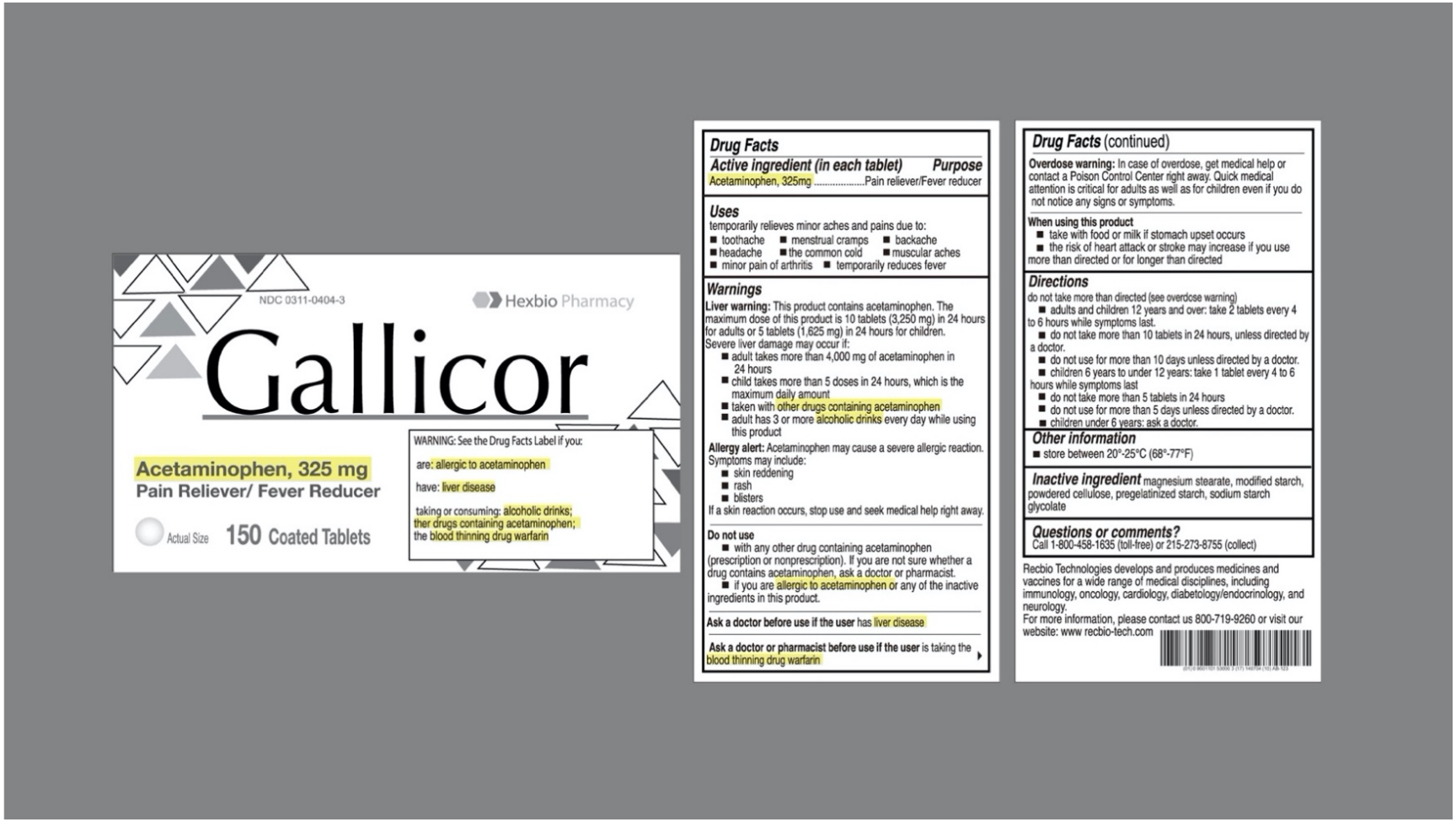


Figure C.4 Example Label for Acetaminophen, enhanced with both the FOP or Highlight Treatment

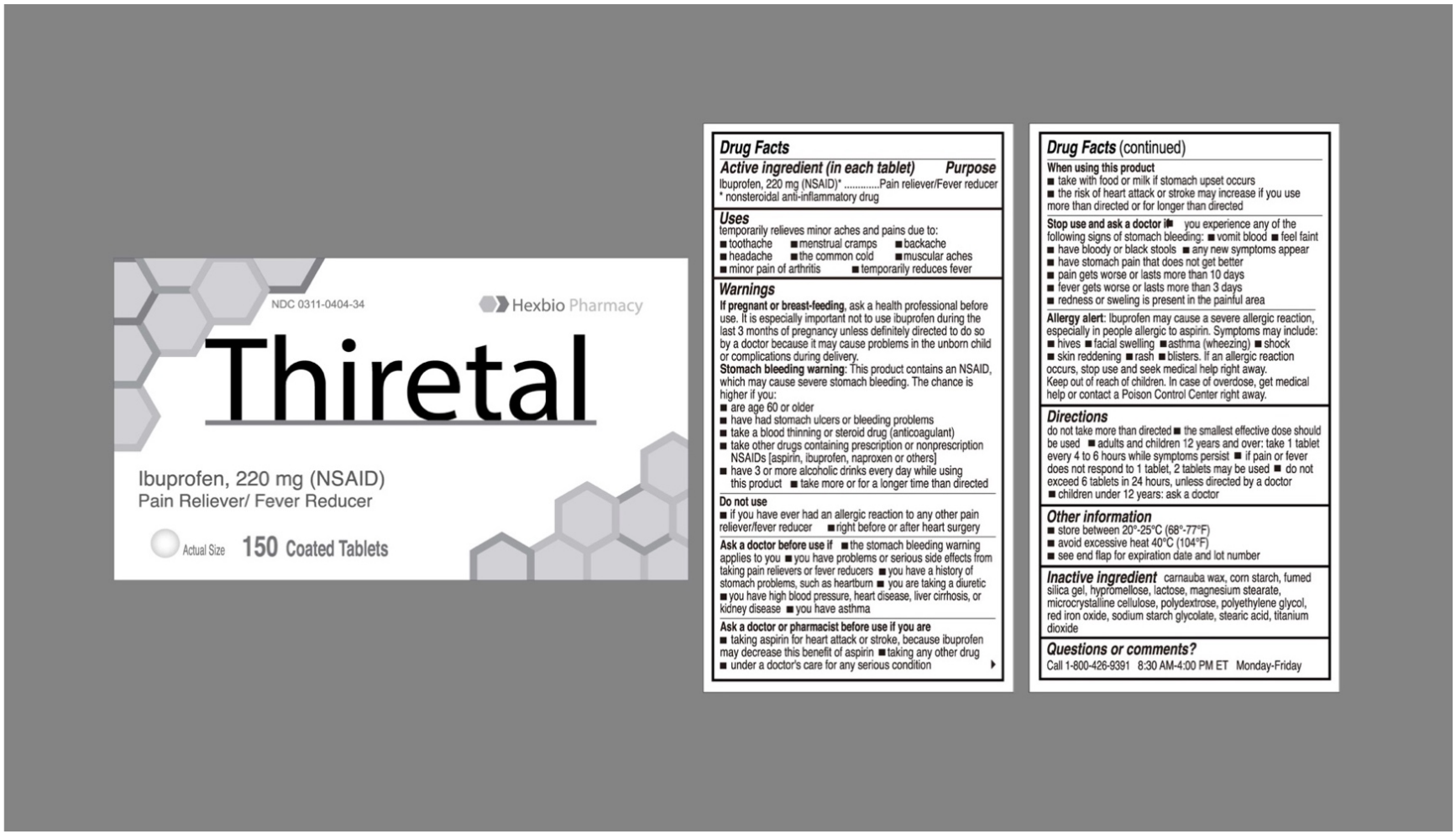


Figure C.5 Example Standard Label for Ibuprofen, with neither the FOP or Highlight Treatment

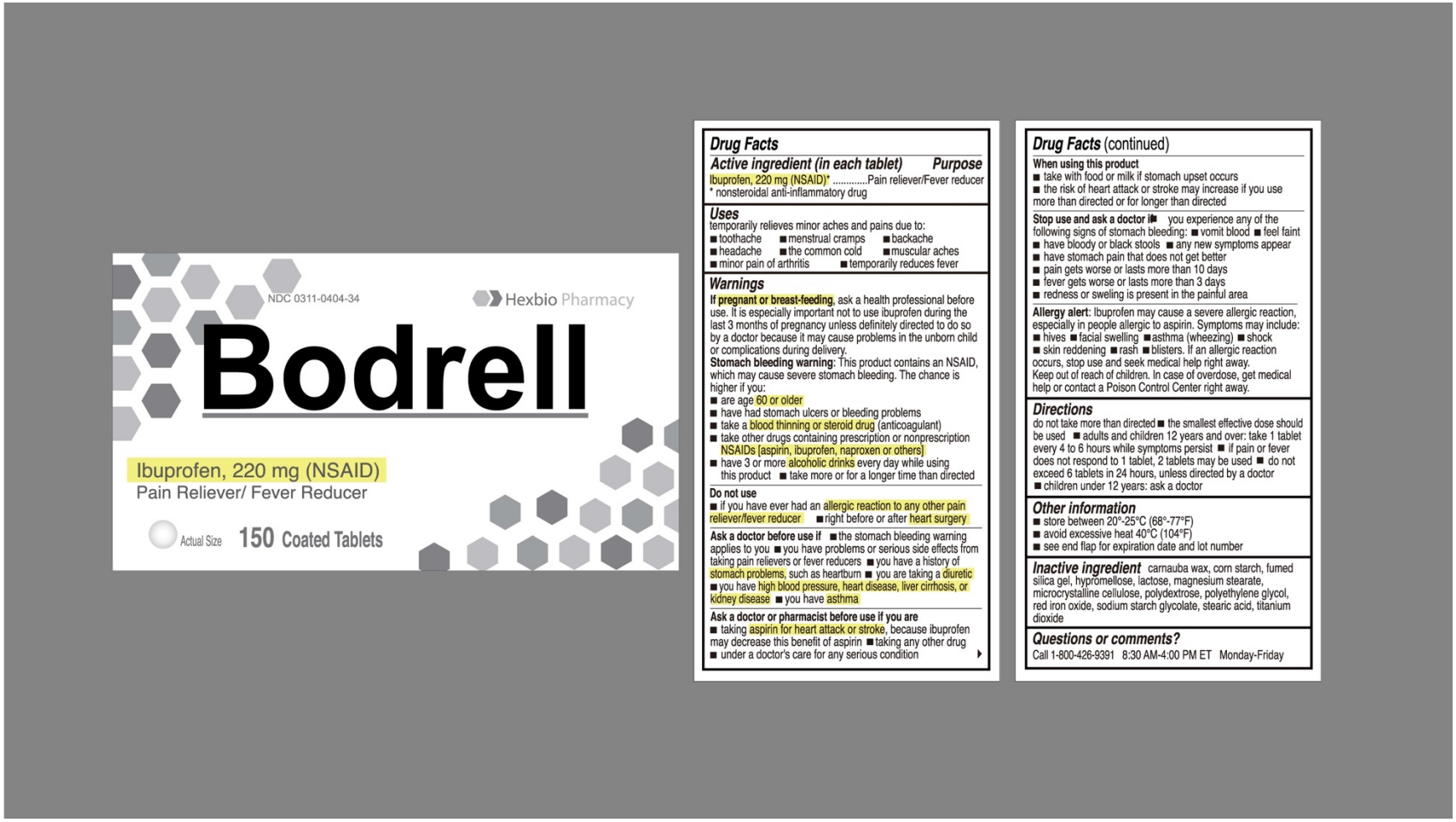


Figure C.6 Example Label for Ibuprofen, enhanced with the Highlight Treatment

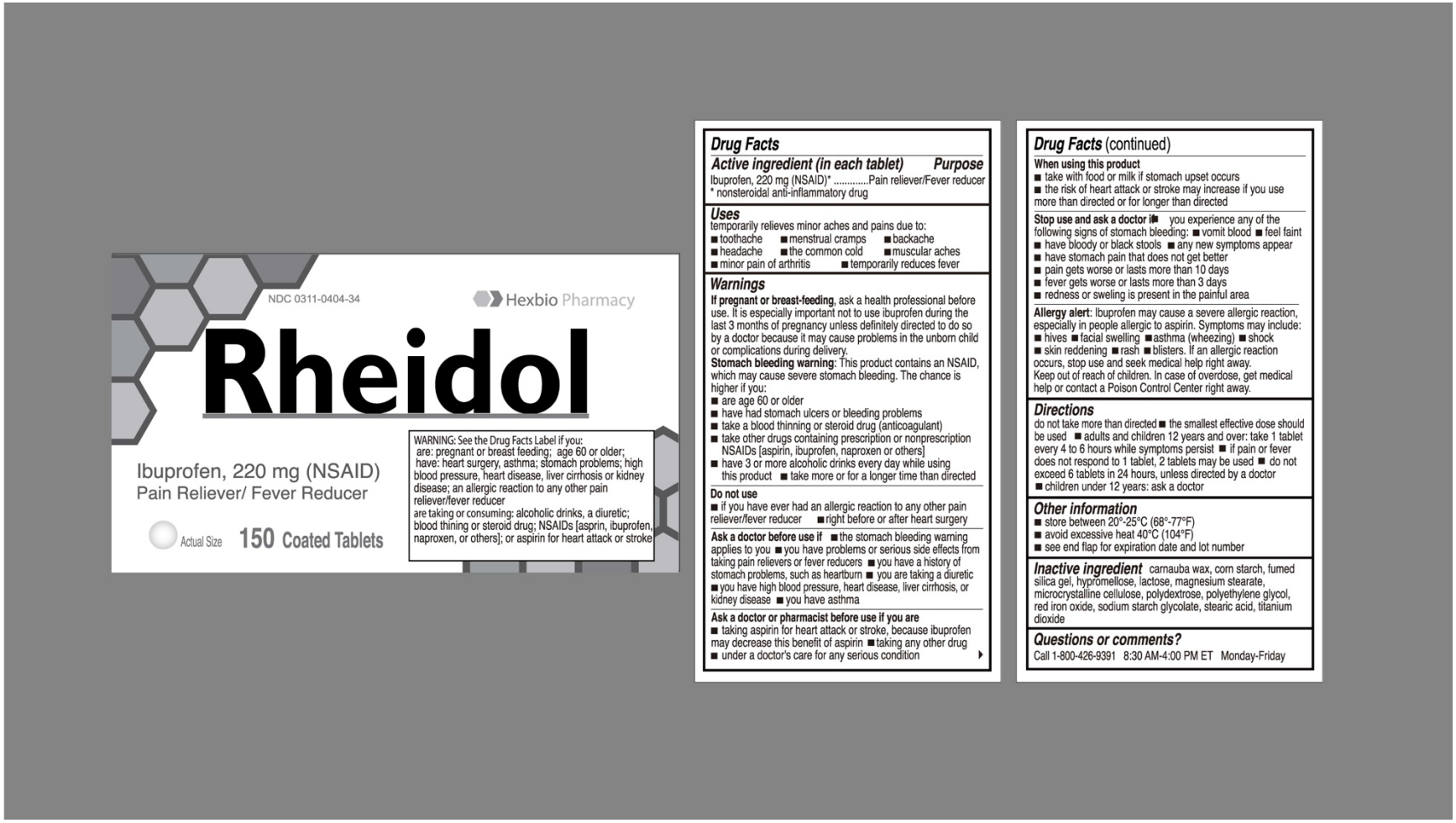


Figure C.7 Example Label for Ibuprofen, enhanced with the FOP Treatment

Figure C.8 Example Label for Ibuprofen, enhanced with both the FOP or Highlight Treatment

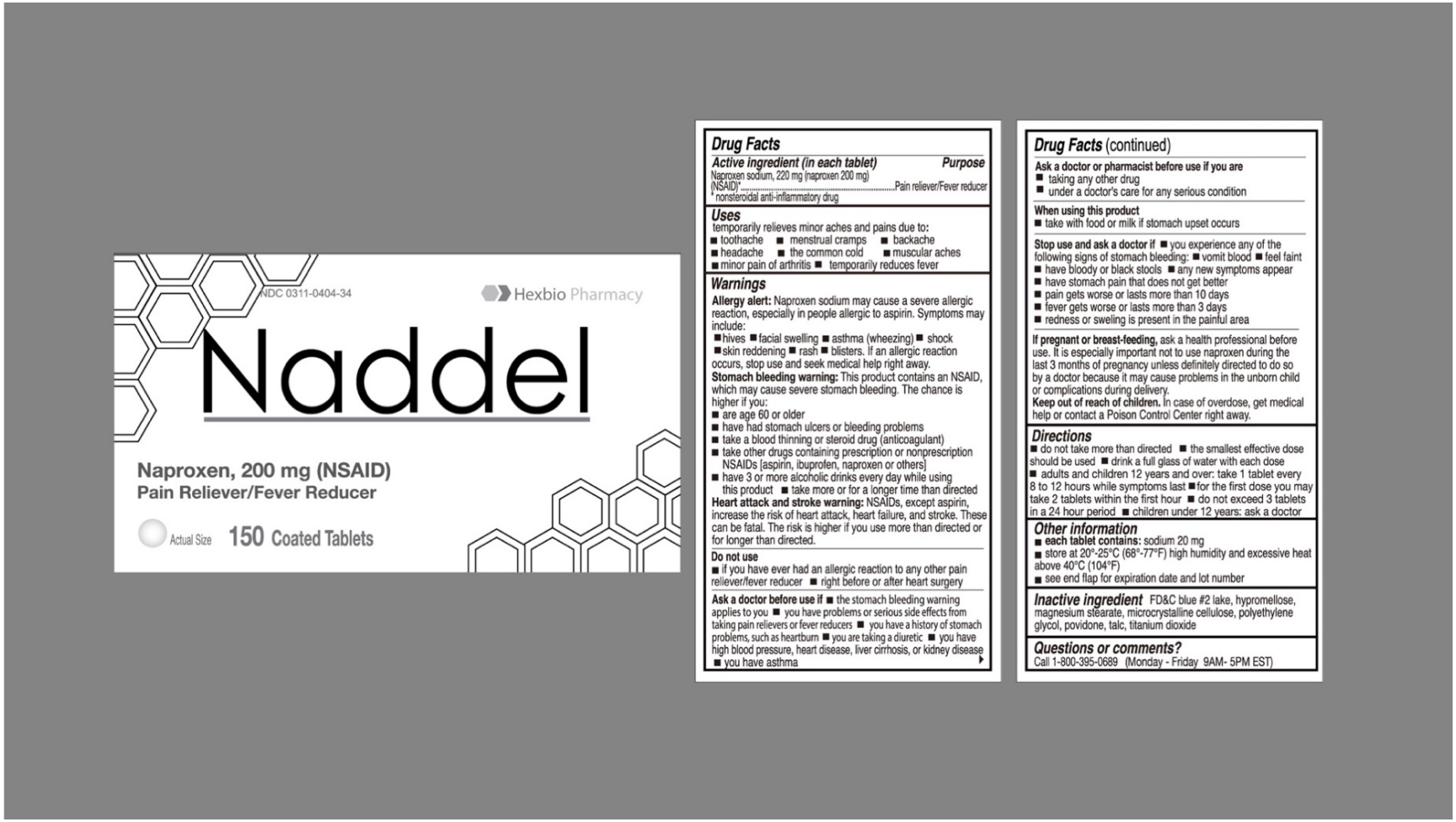


Figure C.9 Example Label for Naproxen, without the Highlight or FOP Warning Treatment

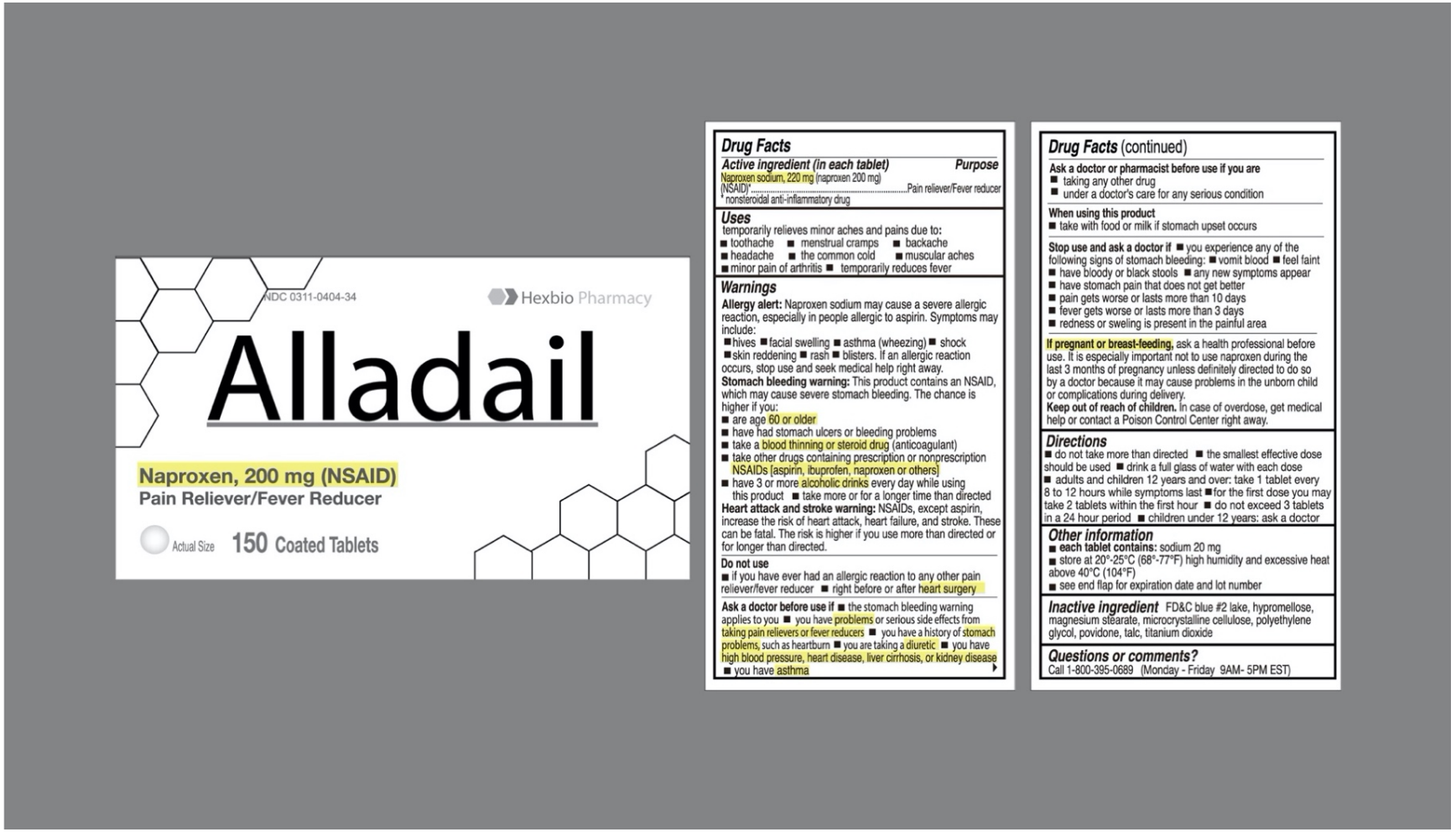


Figure C.10 Example Label for Naproxen, enhanced with the Highlight Treatment

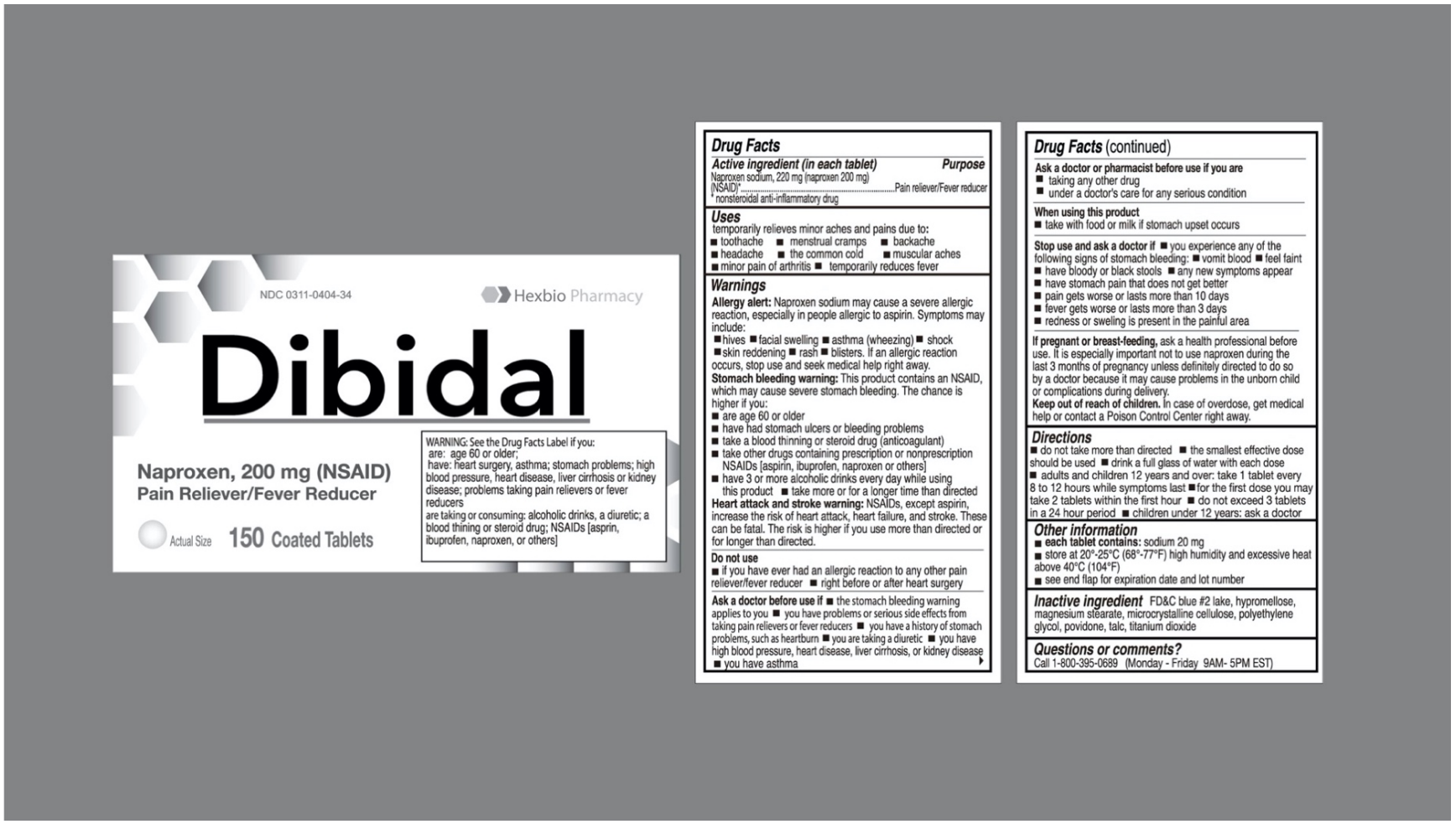


Figure C.11 Example Label for Naproxen, enhanced with the FOP Treatment

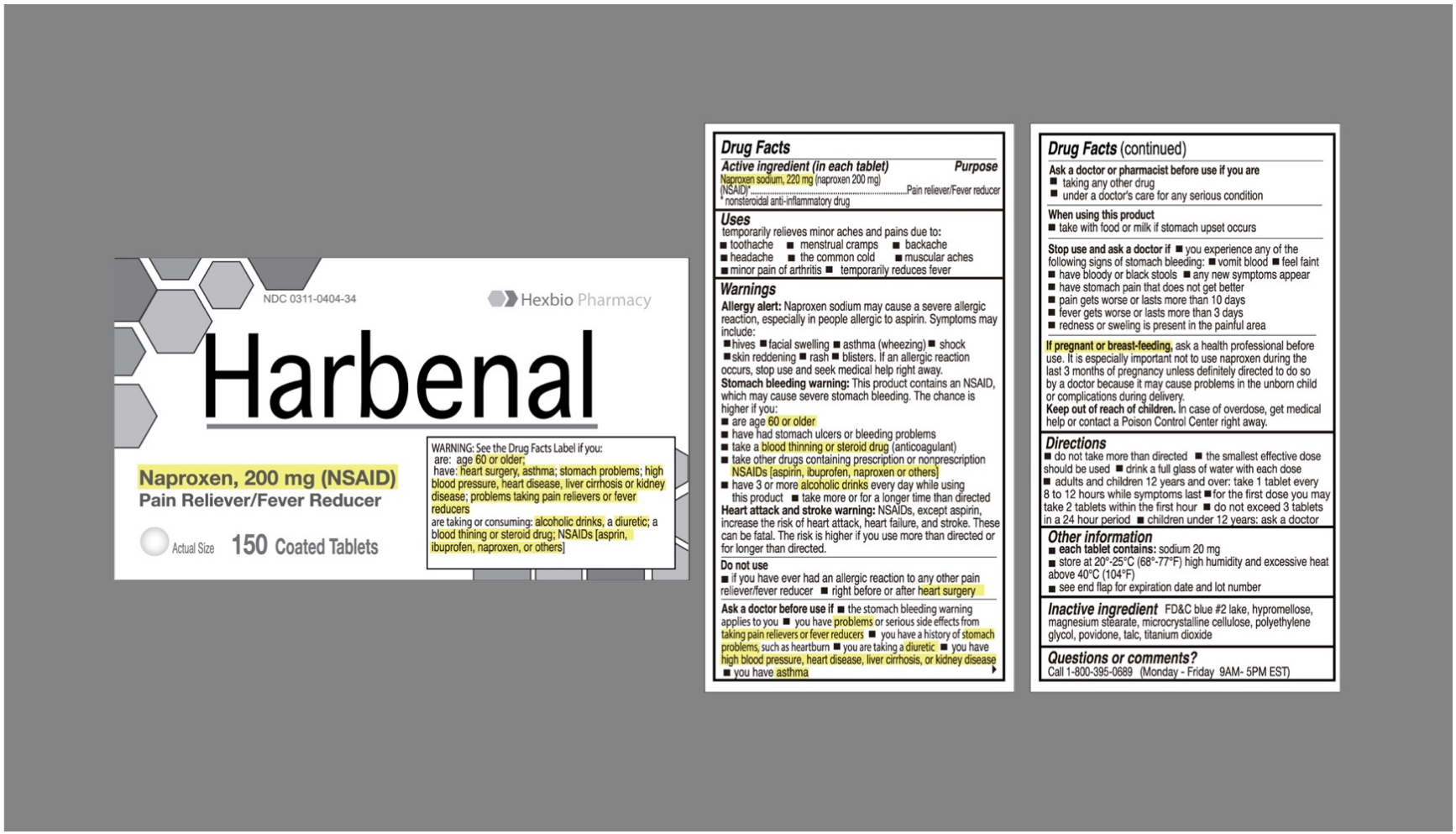


Figure C.12 Example Label for Naproxen, enhanced with both the FOP or Highlight Treatment

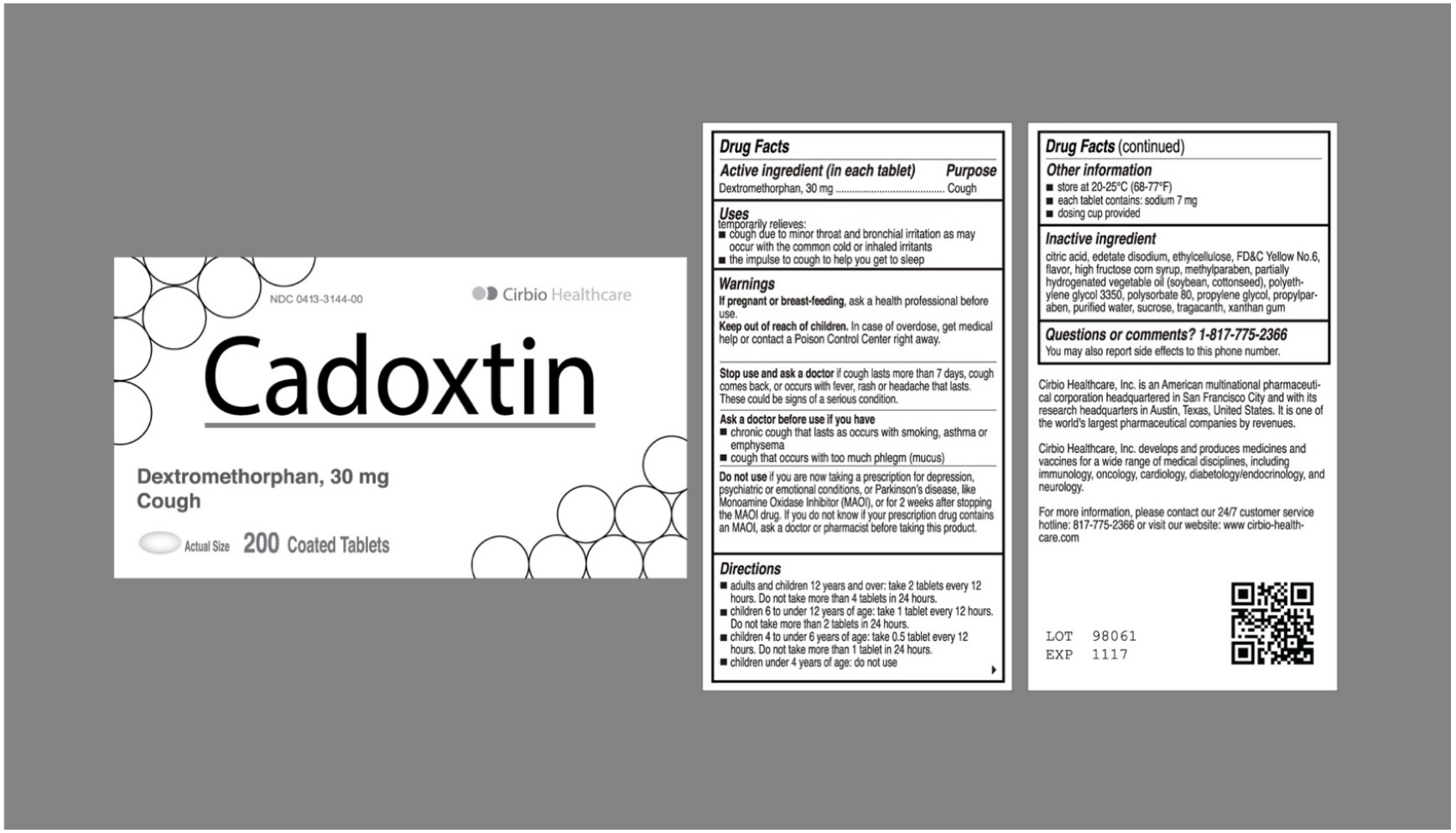


Figure C.13 Example Label for Dextromethorphan, without the Highlight or FOP Warning Treatment

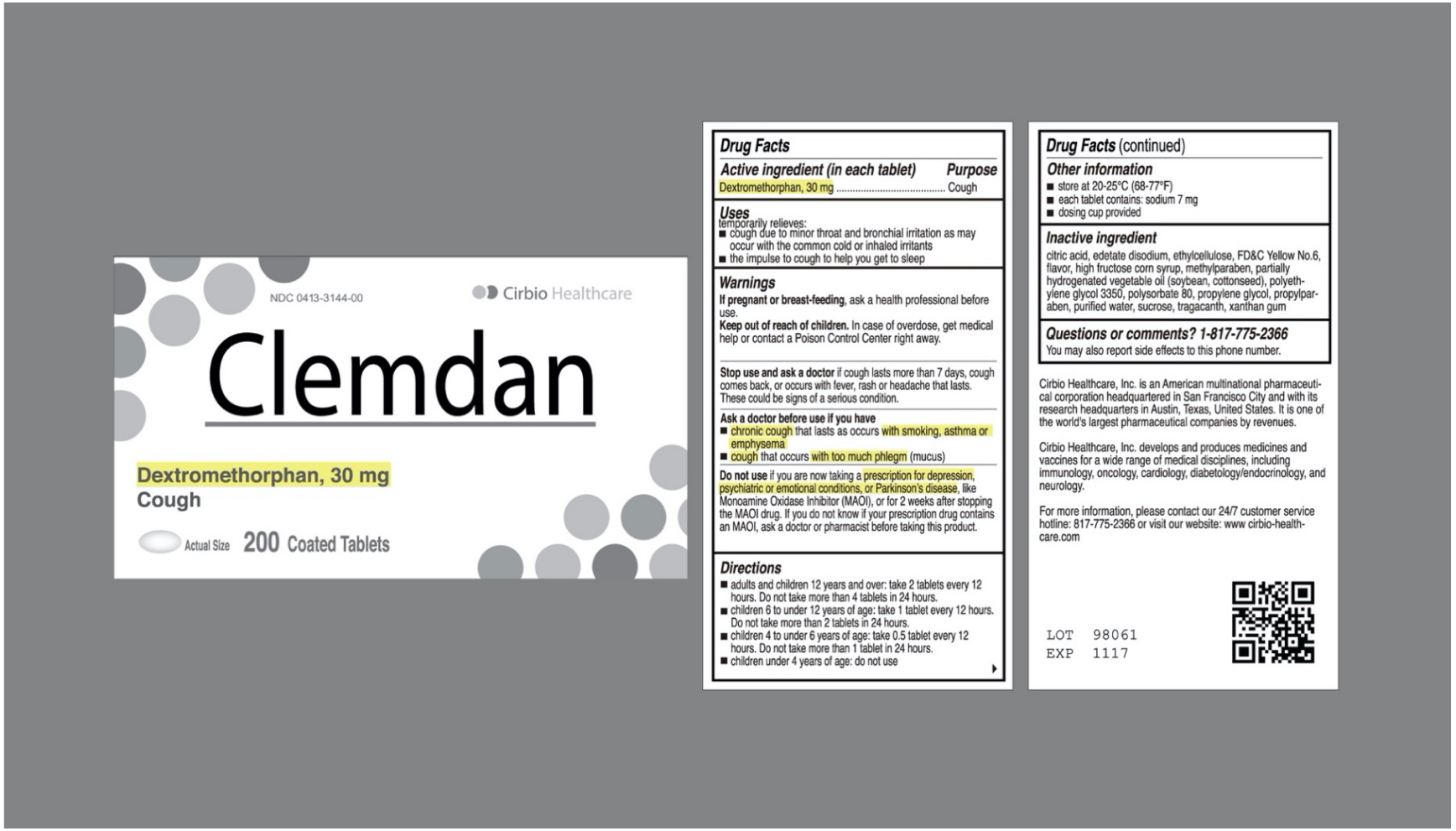


Figure C.14 Example Label for Dextromethorphan, enhanced with the Highlight Treatment

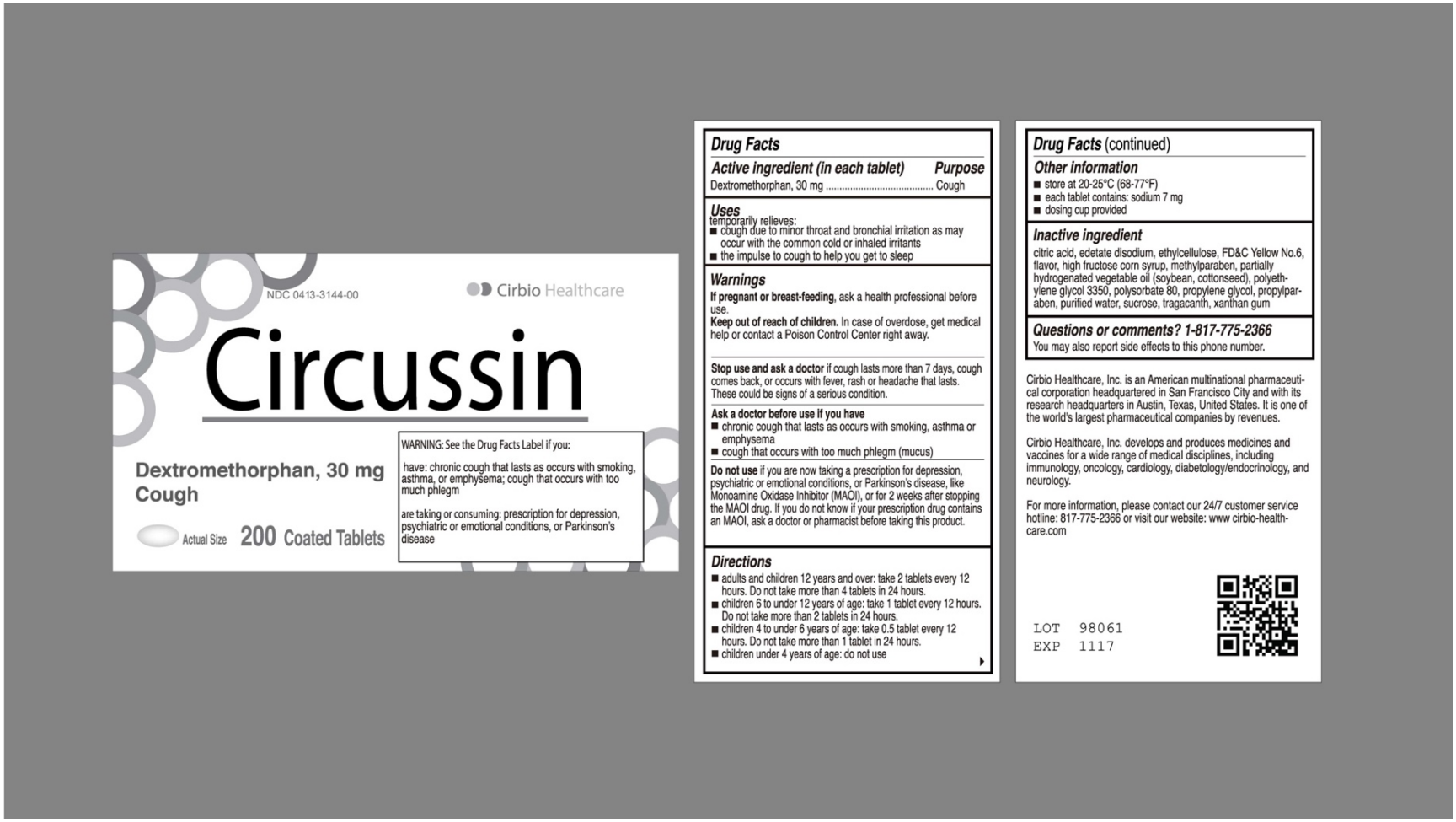


Figure C.15 Example Label for Dextromethorphan, enhanced with the FOP Treatment

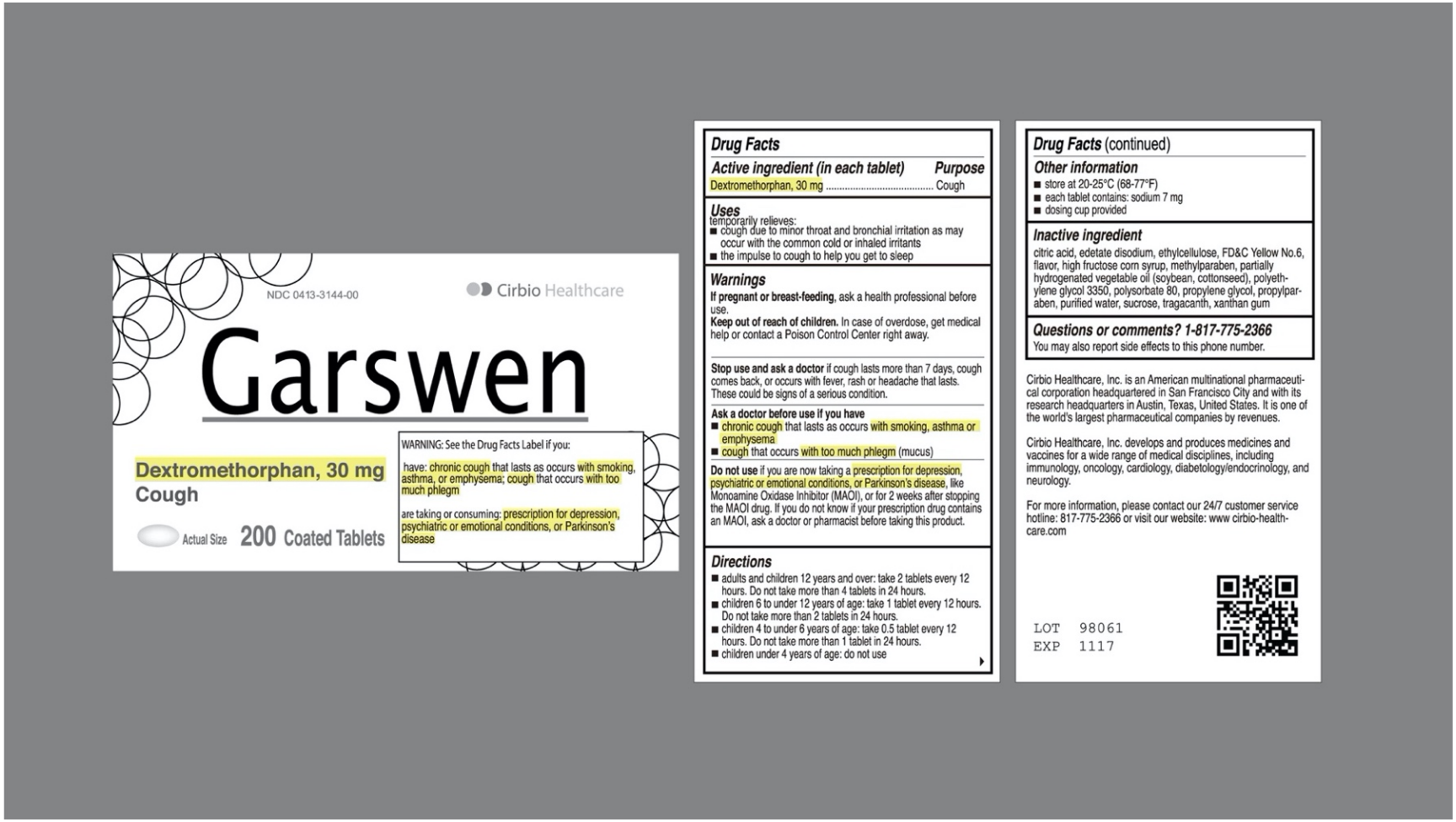


Figure C.16 Example Label for Dextromethorphan, enhanced with both the Highlight and FOP Treatments

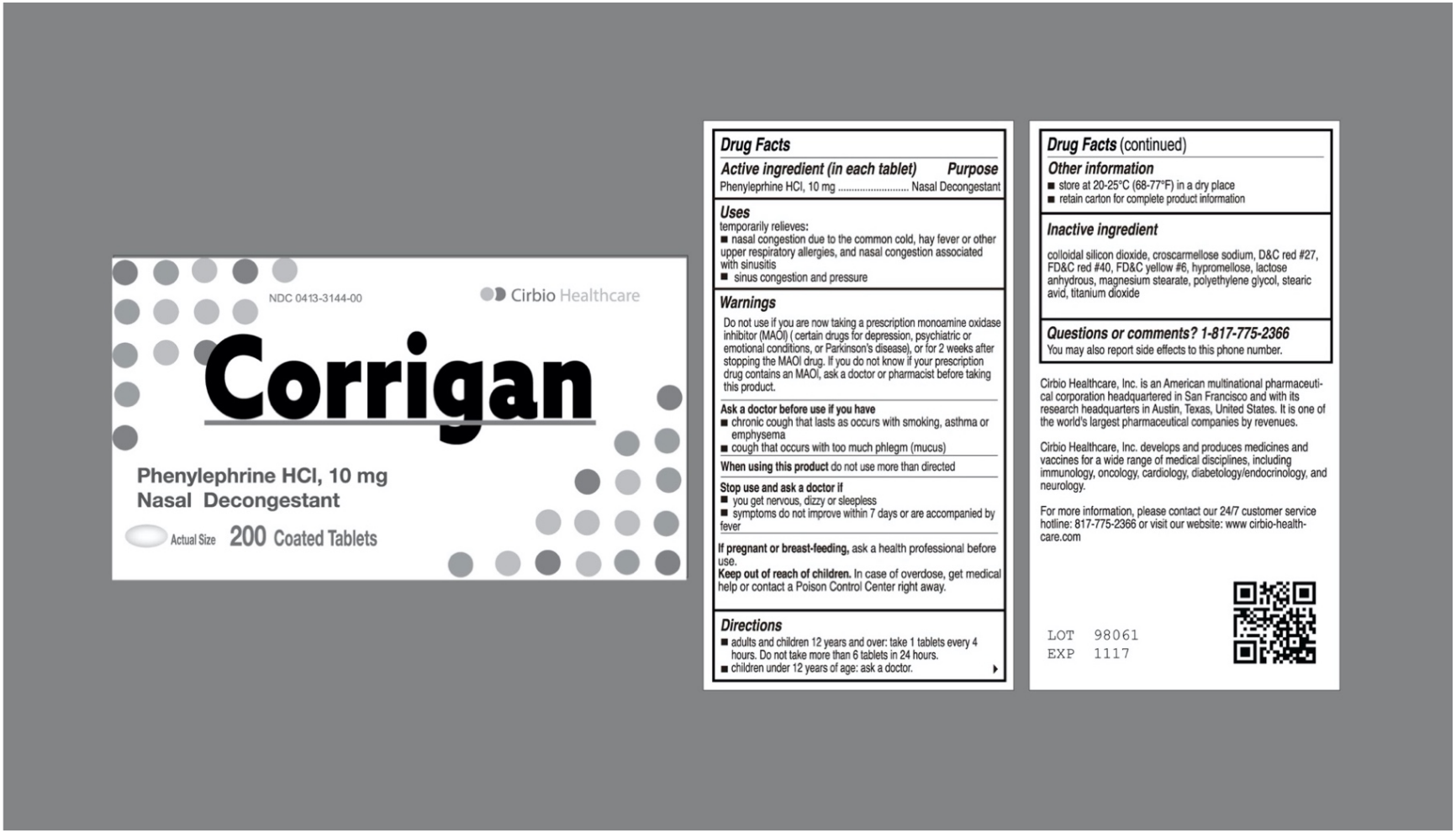


Figure C.17 Example Label for Phenylephrine, without the Highlight or FOP Treatments

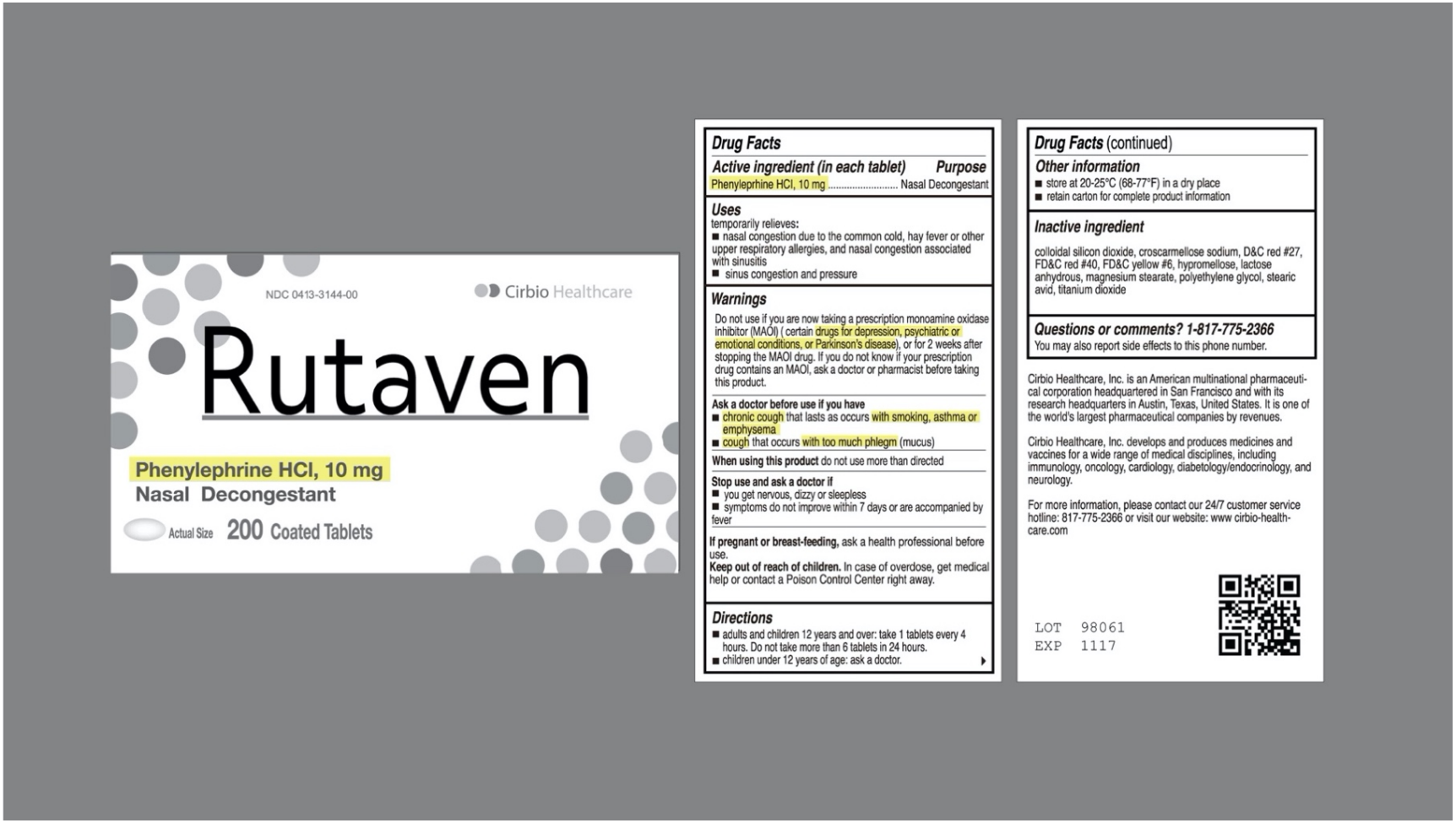


Figure C.18 Example Label for Phenylephrine, enhanced with the Highlight Treatment

Figure C.19 Example Label for Phenylephrine, enhanced with the FOP Treatment

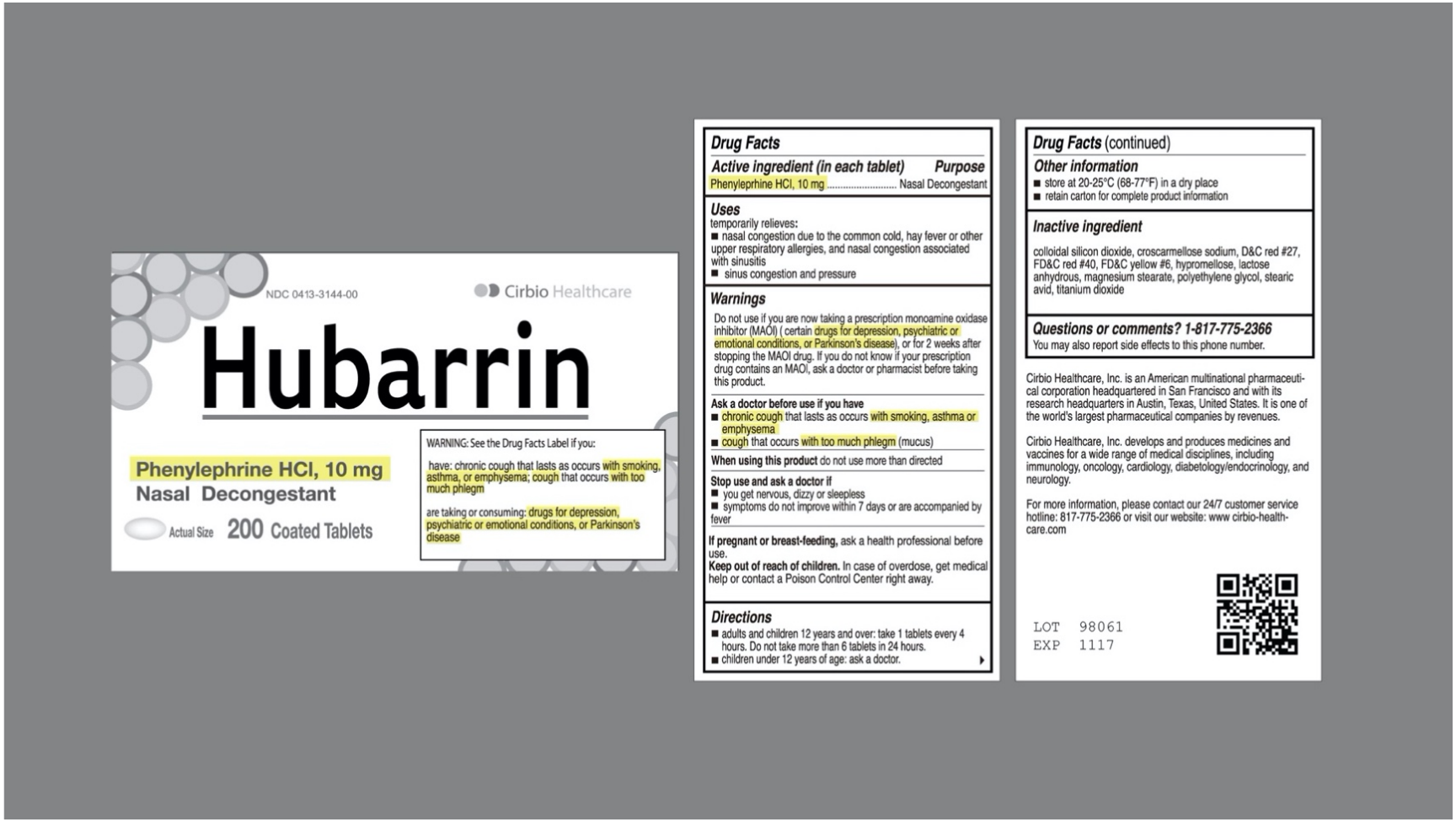


Figure C.20 Example Label for Phenylephrine, enhanced with both the FOP or Highlight Treatment

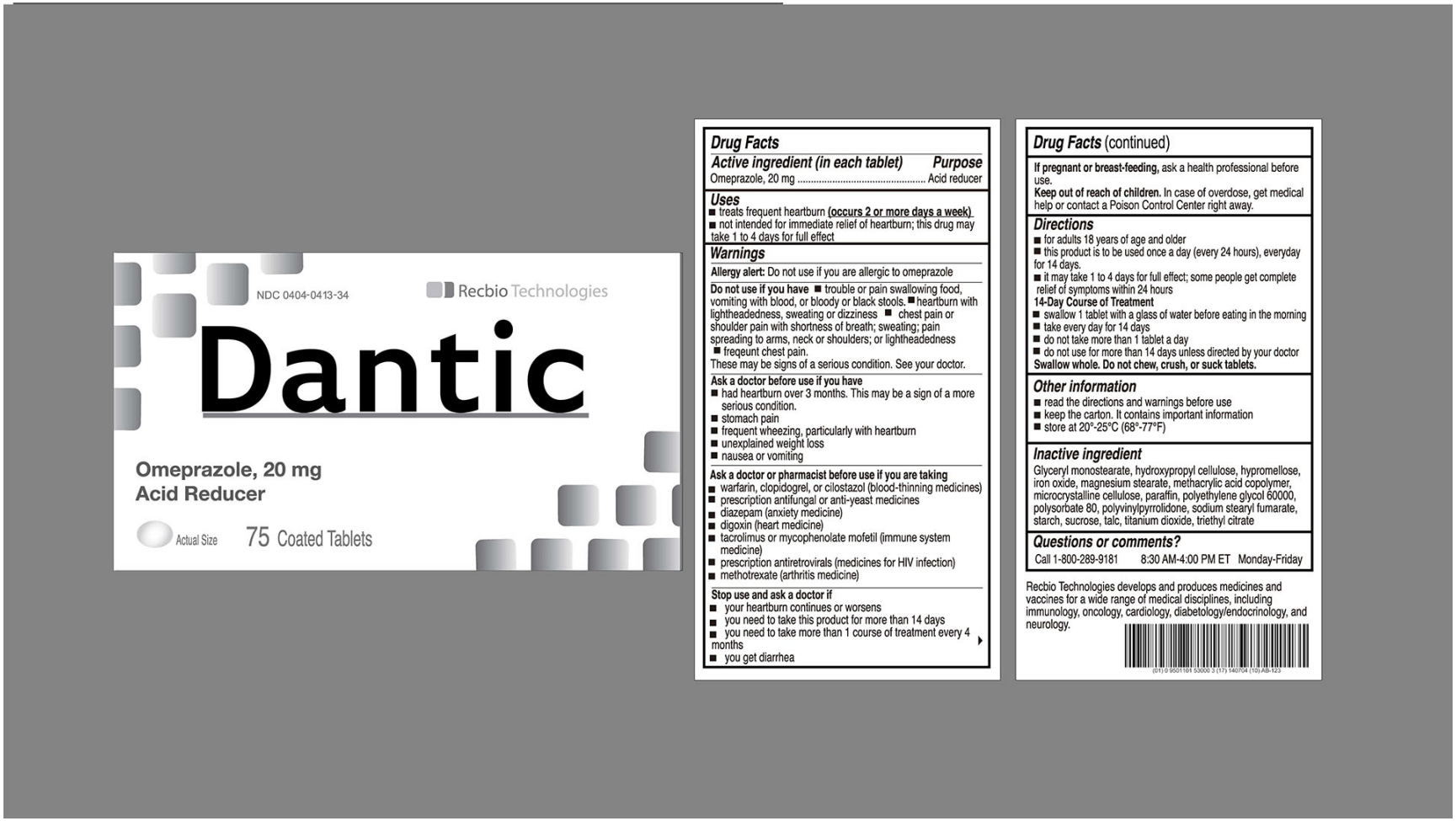


Figure C.21 Example Label for Omeprazole, without the Highlight or FOP Warning Treatment

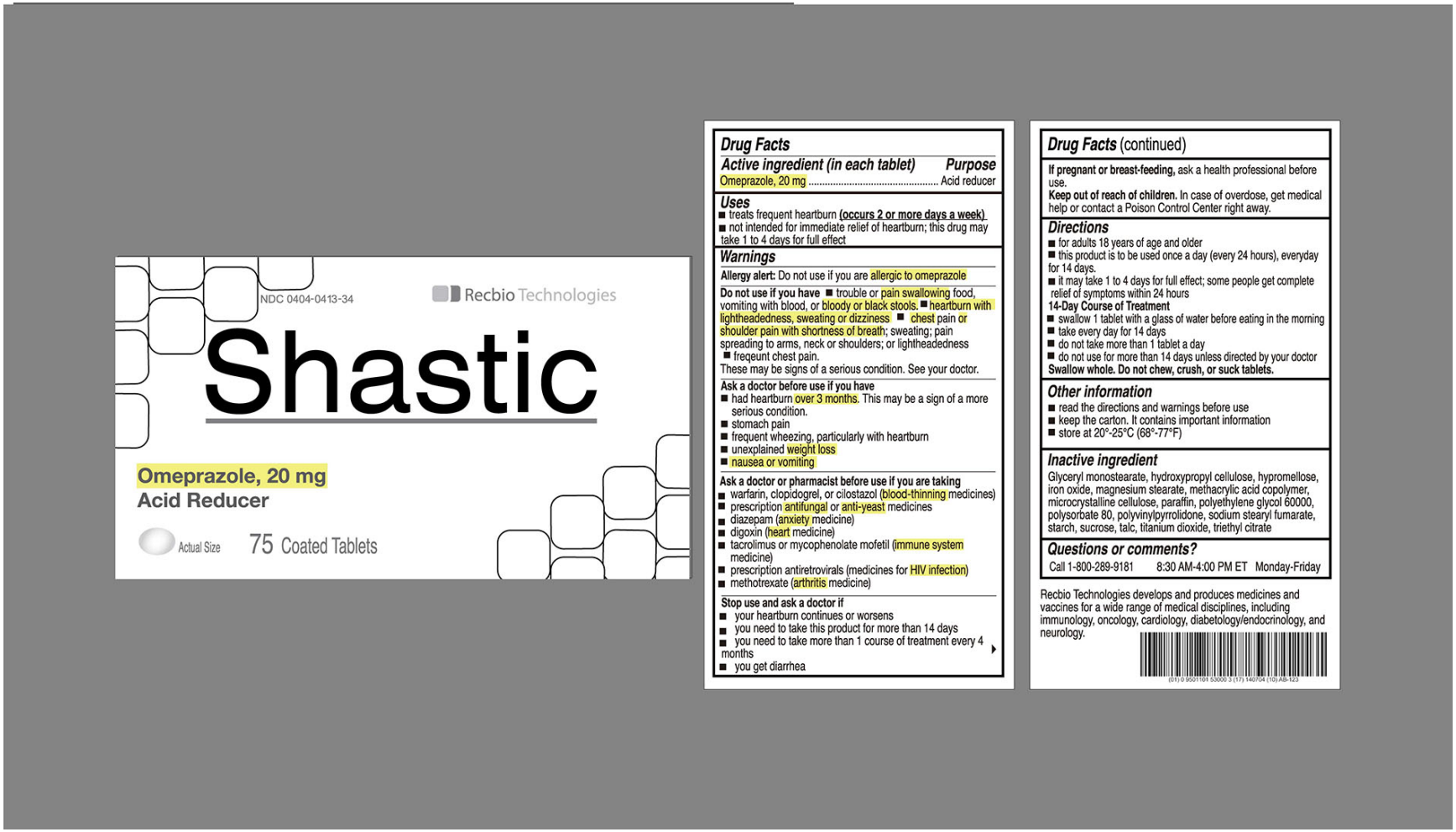


Figure C.22 Example Label for Omeprazole, enhanced with the Highlight Treatment

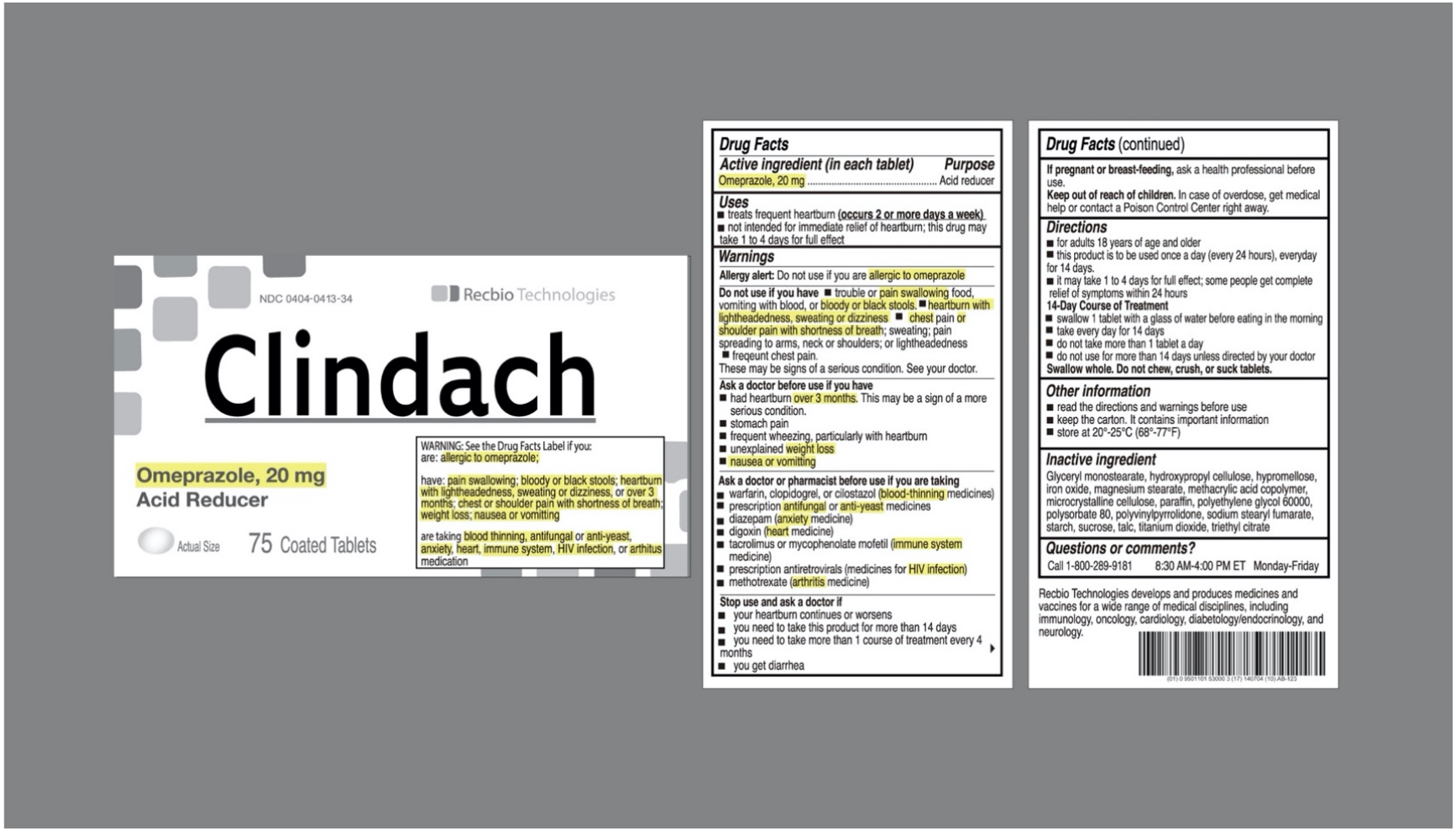


Figure C.23 Example Label for Omeprazole, enhanced with both the Highlight and the FOP Treatment

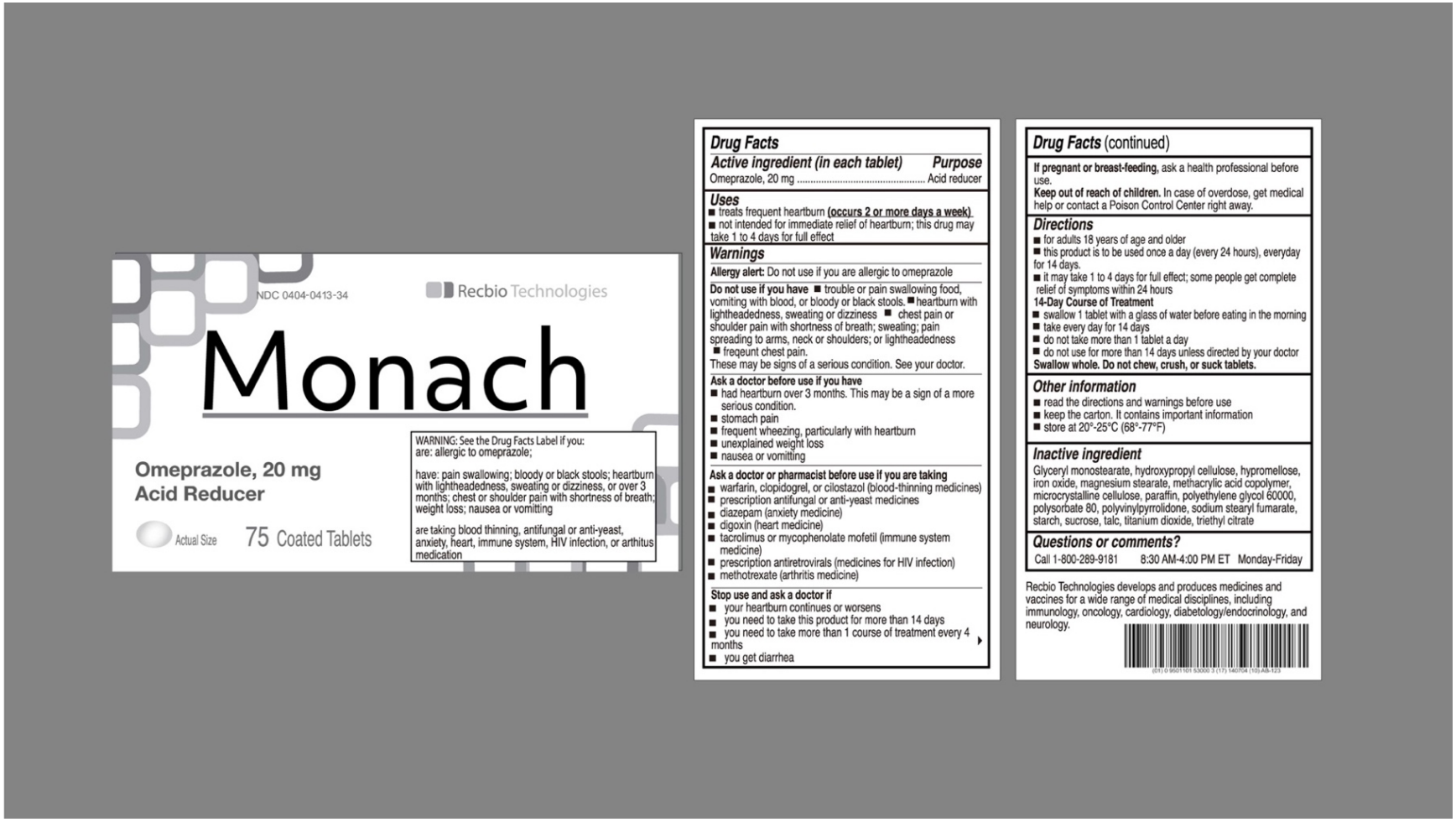


Figure C.24 Example Label for Omeprazole, enhanced with the FOP Treatment

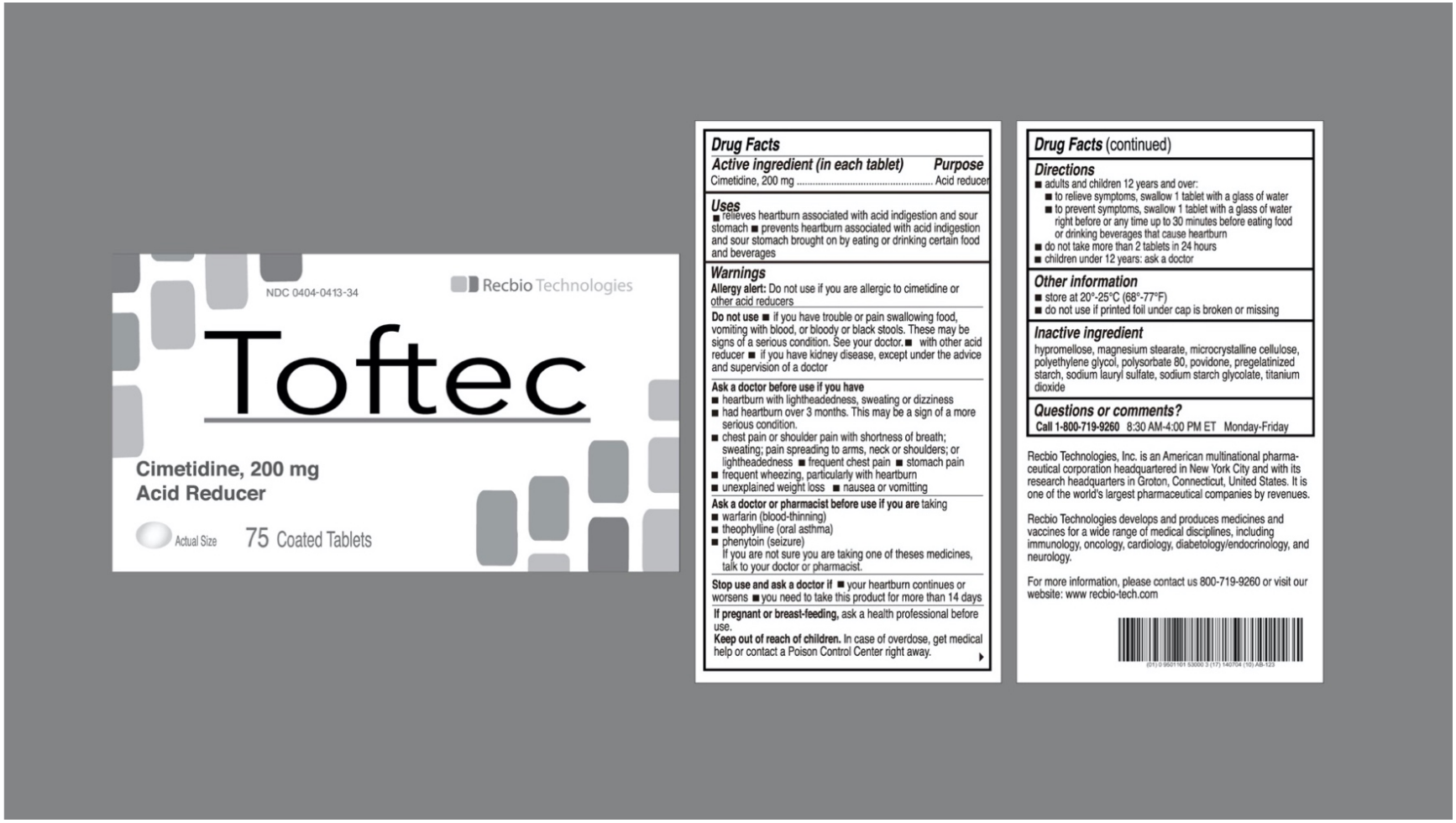


Figure C.25 Example Label for Cimetidine, without the Highlight or FOP Warning Treatment

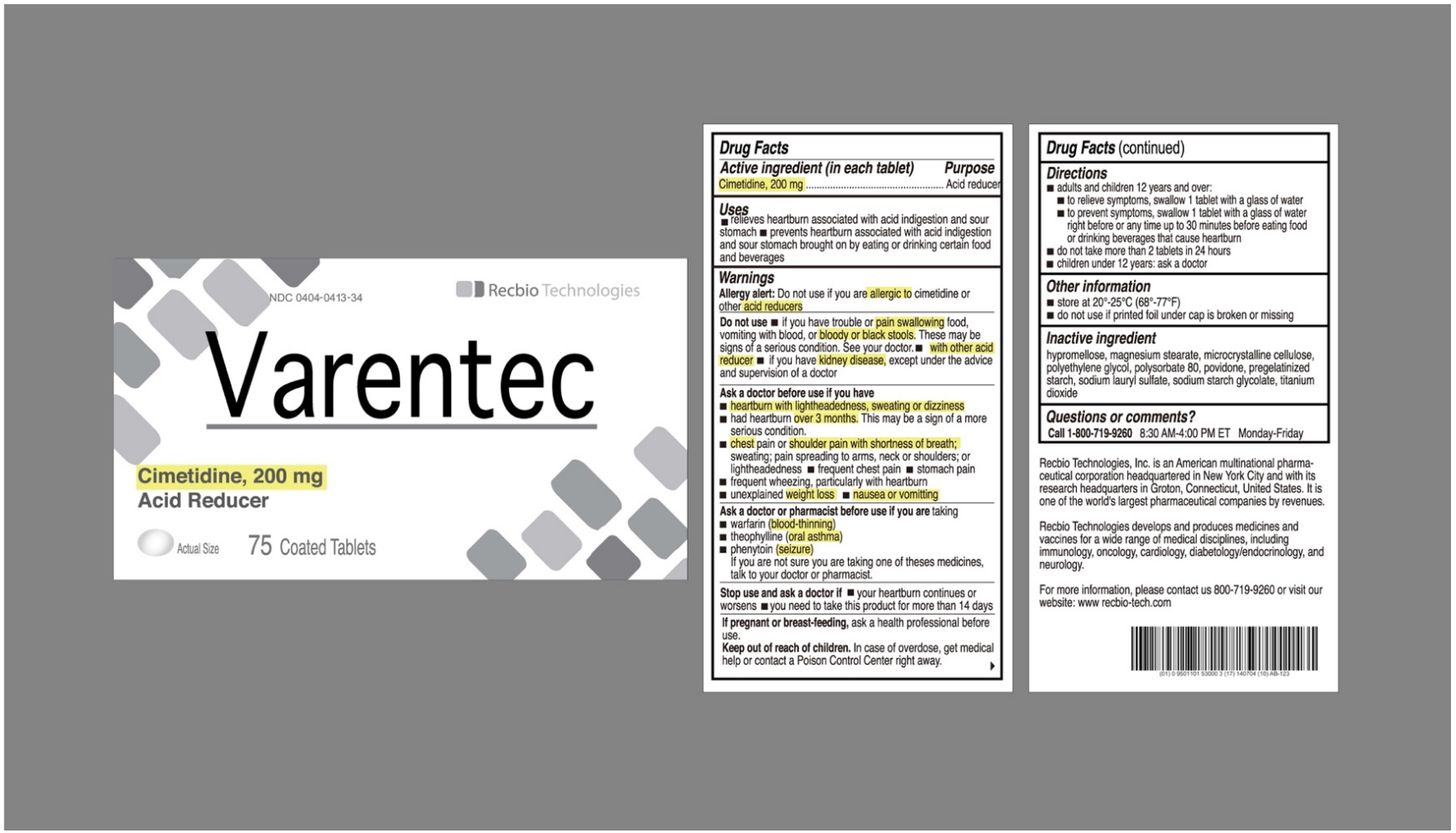


Figure C.26 Example Label for Cimetidine, enhanced with the Highlight Treatment

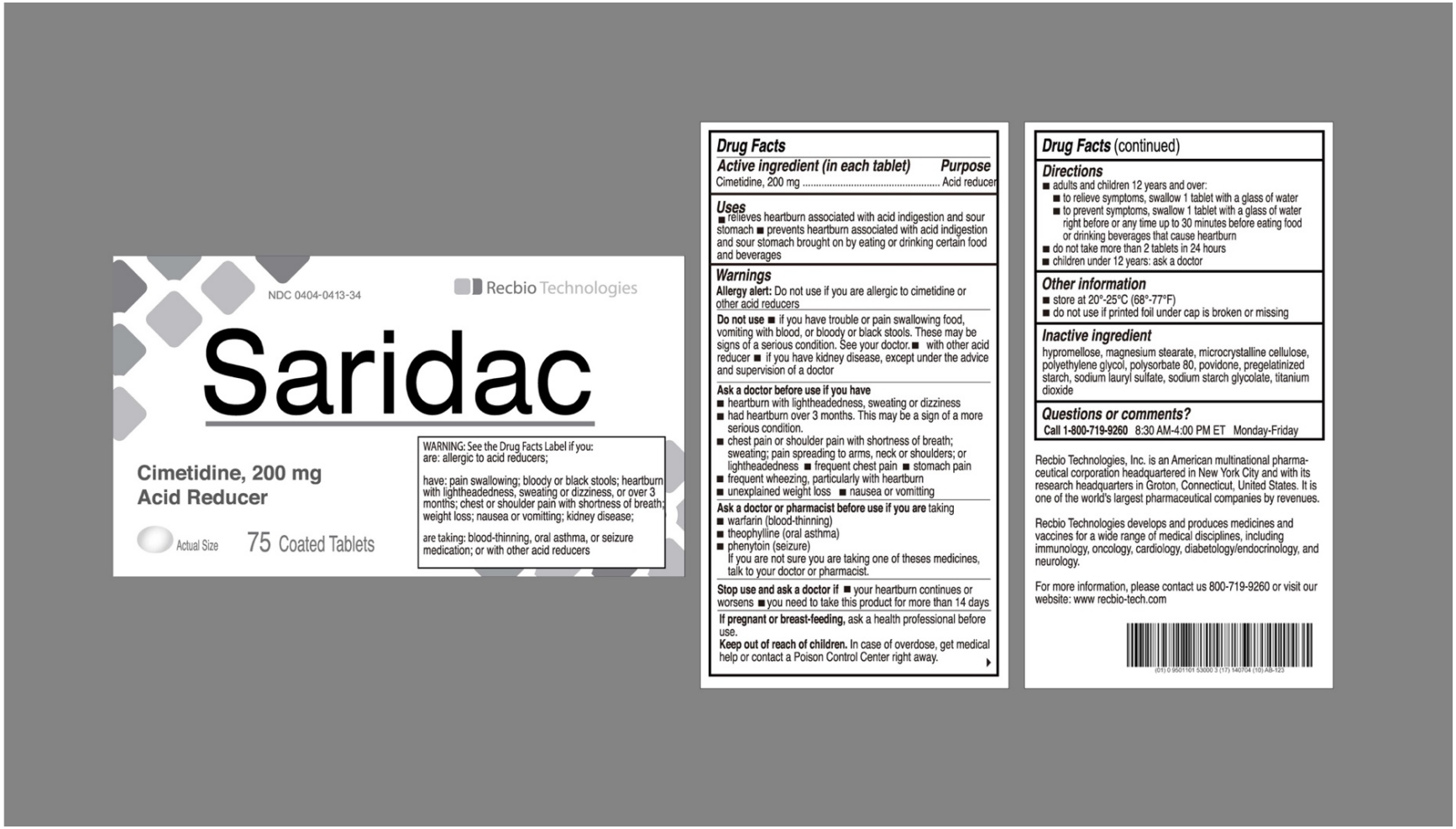


Figure C.27 Example Label for Cimetidine, enhanced with the FOP Treatment

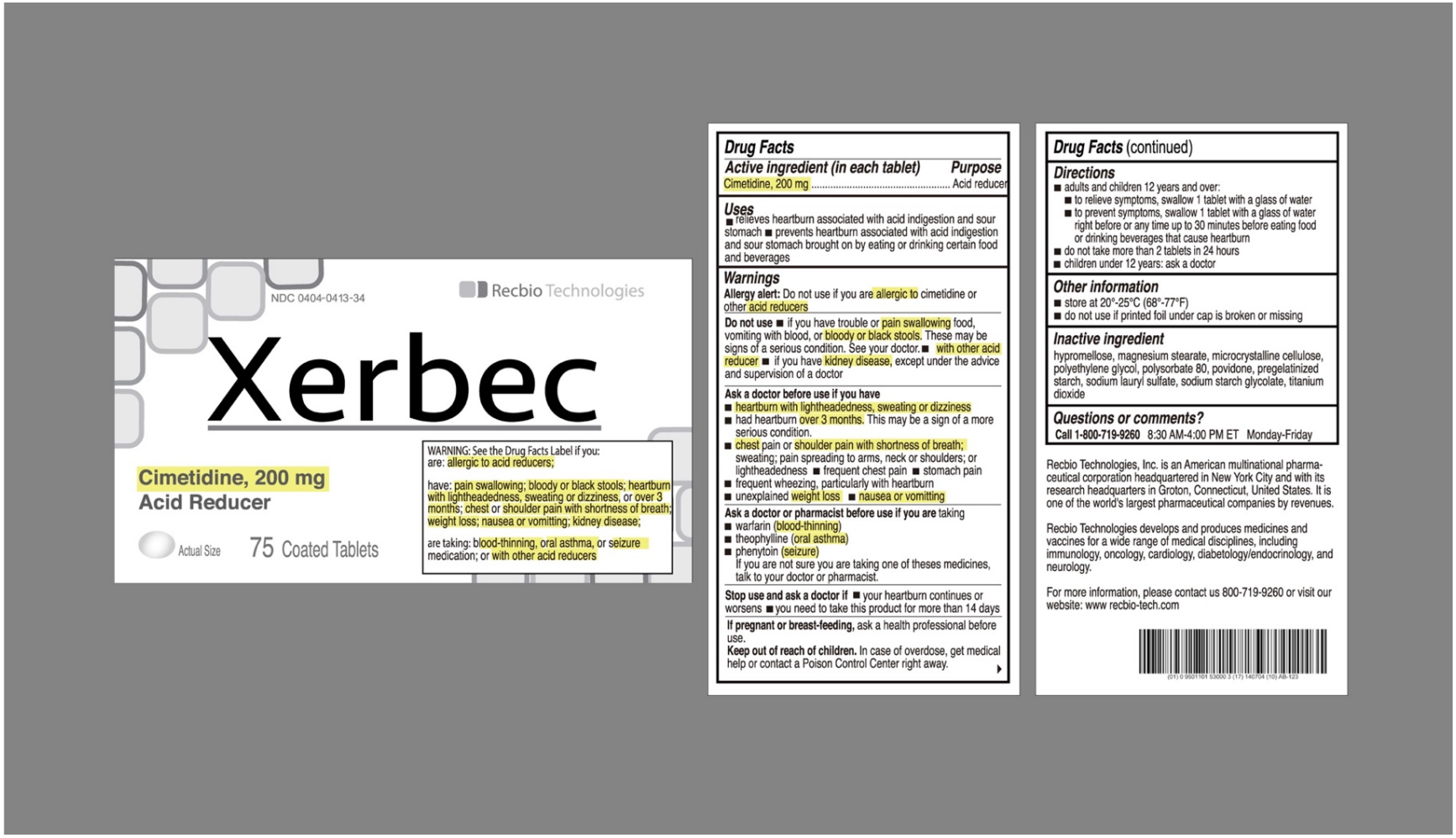


Figure C.28 Example Label for Cimetidine, enhanced with both the FOP or Highlight Treatment

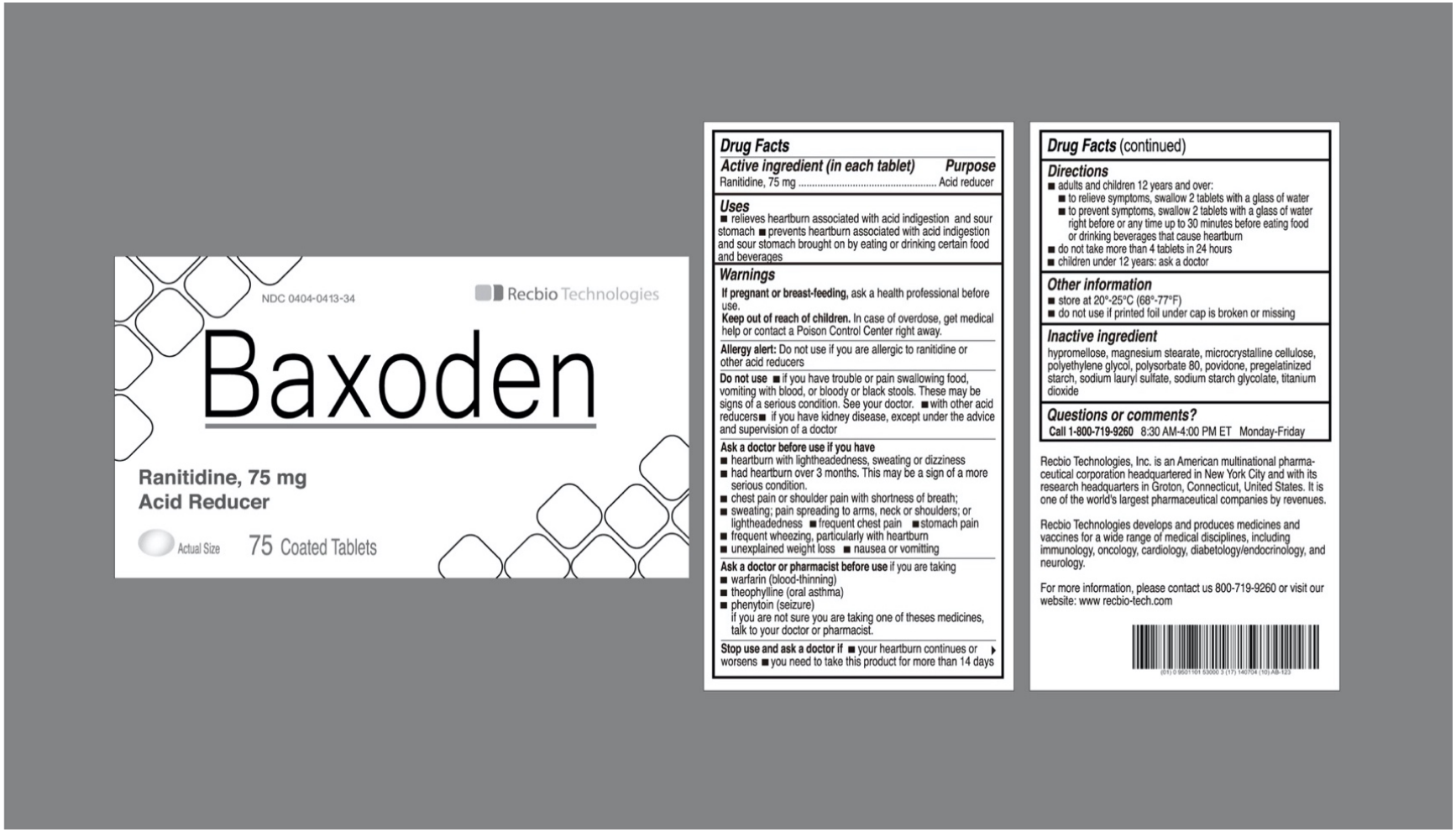


Figure C.29 Example Label for Ranitidine, without the Highlight or FOP Warning Treatment

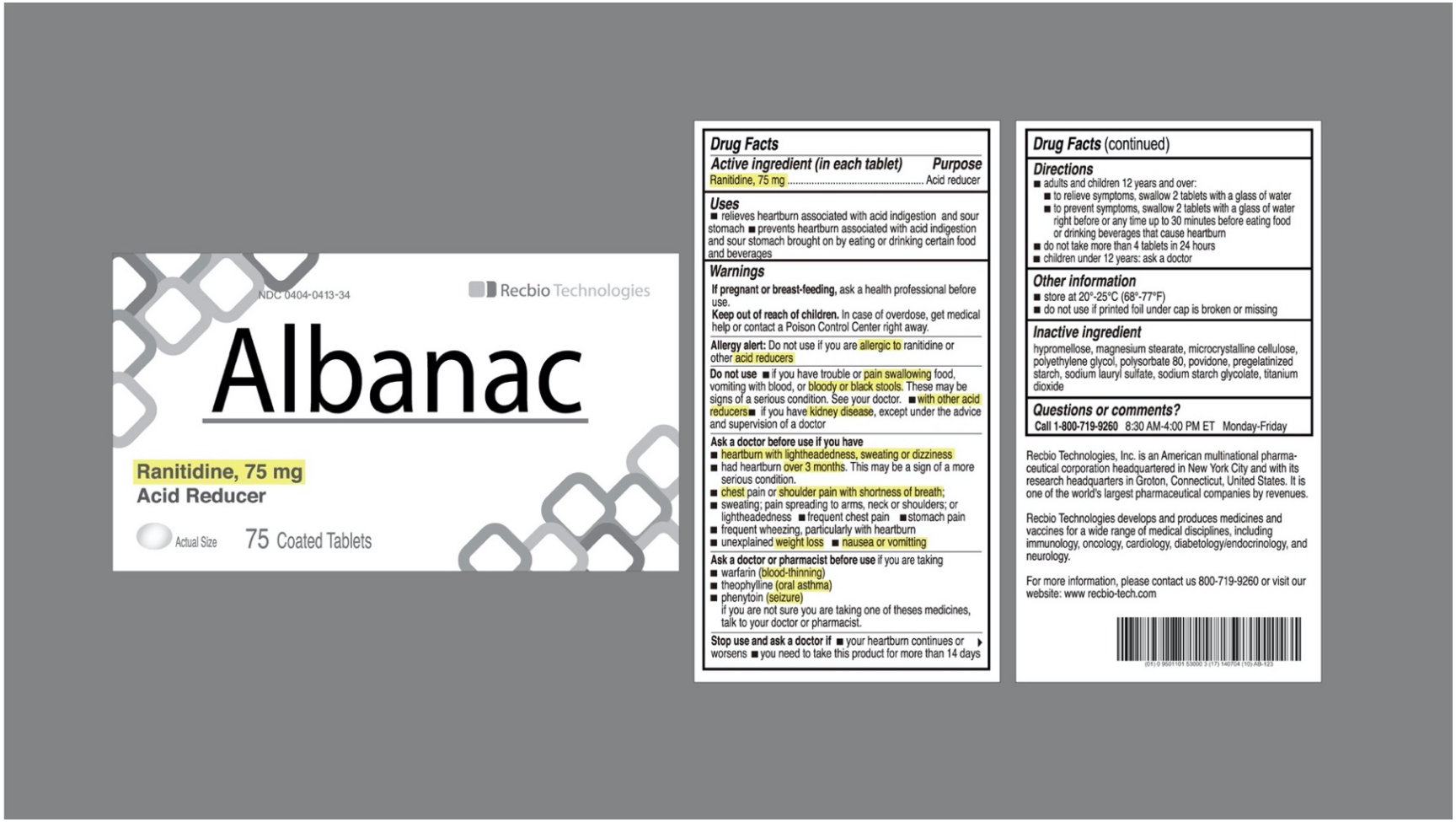


Figure C.30 Example Label for Ranitidine, enhanced with the Highlight Treatment

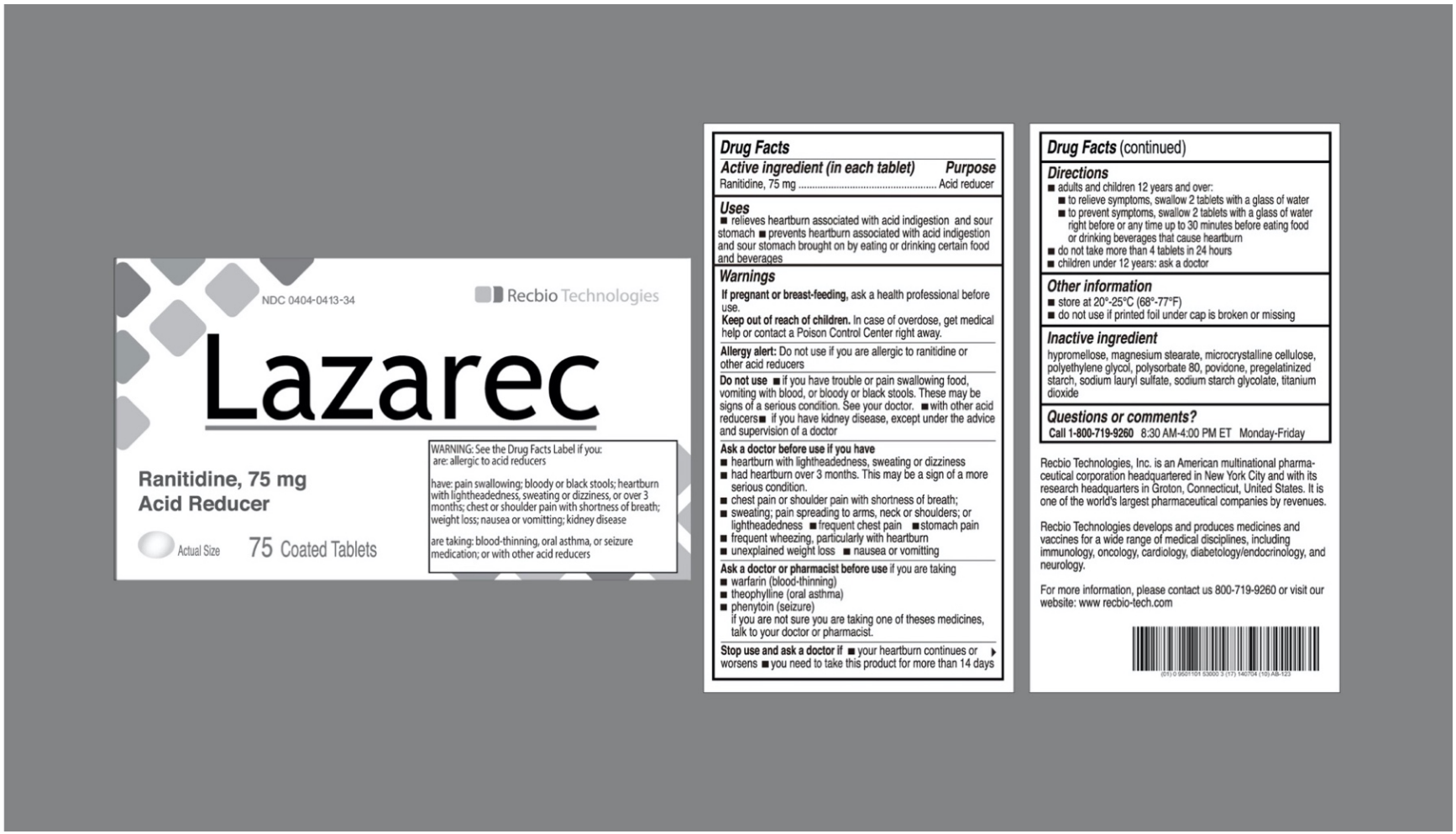


Figure C.31 Example Label for Ranitidine, enhanced with the FOP Treatment

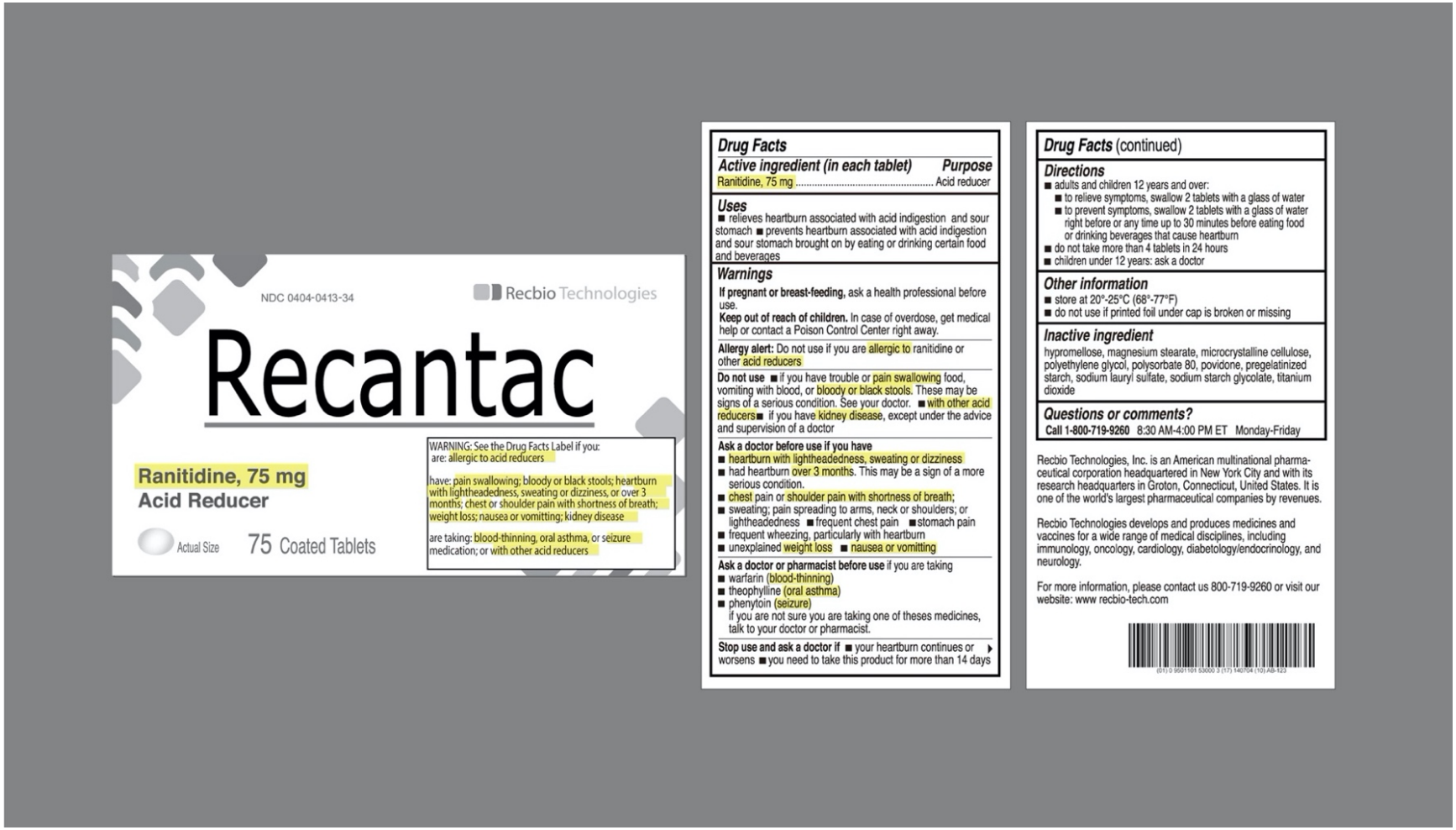


Figure C.32 Example Label for Ranitidine, enhanced with both the FOP or Highlight Treatment

APPENDIX D:

Visual presentation of the distribution of familiarity with OTC Active Ingredients versus Brand Names

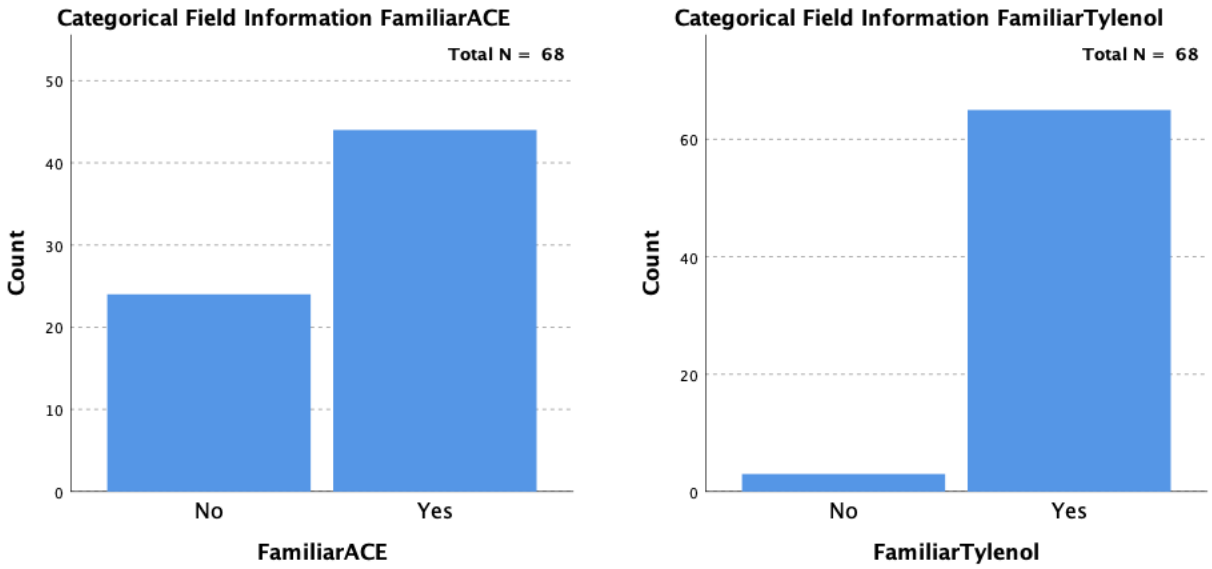


Figure D.1 Frequency counts of familiarity with acetaminophen versus Tylenol

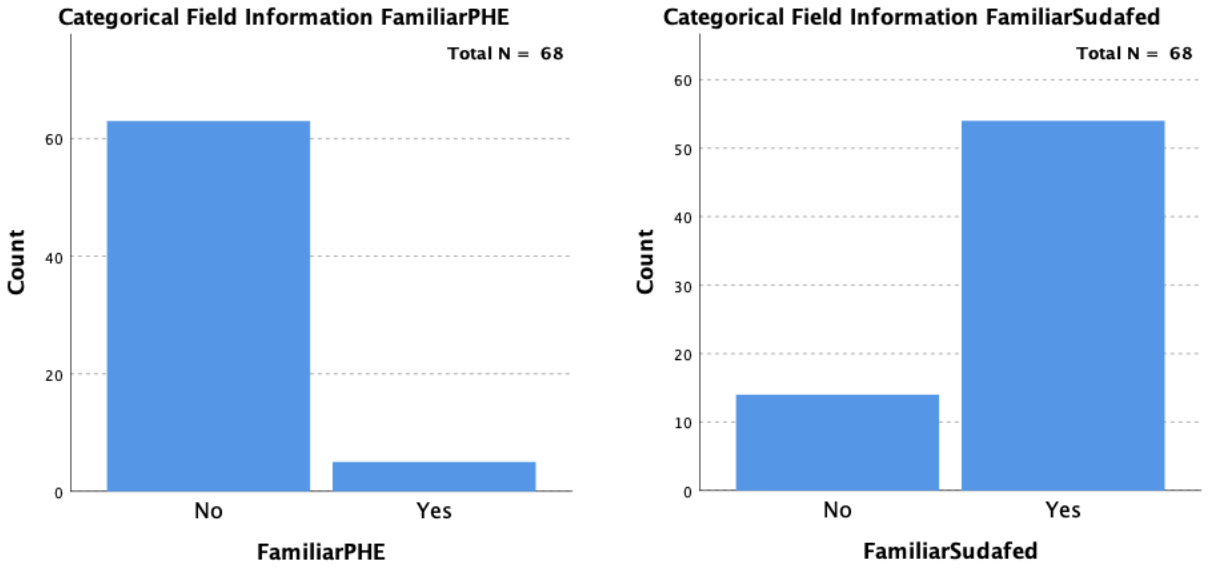


Figure D.2 Frequency counts of familiarity with phenylephrine versus Sudafed

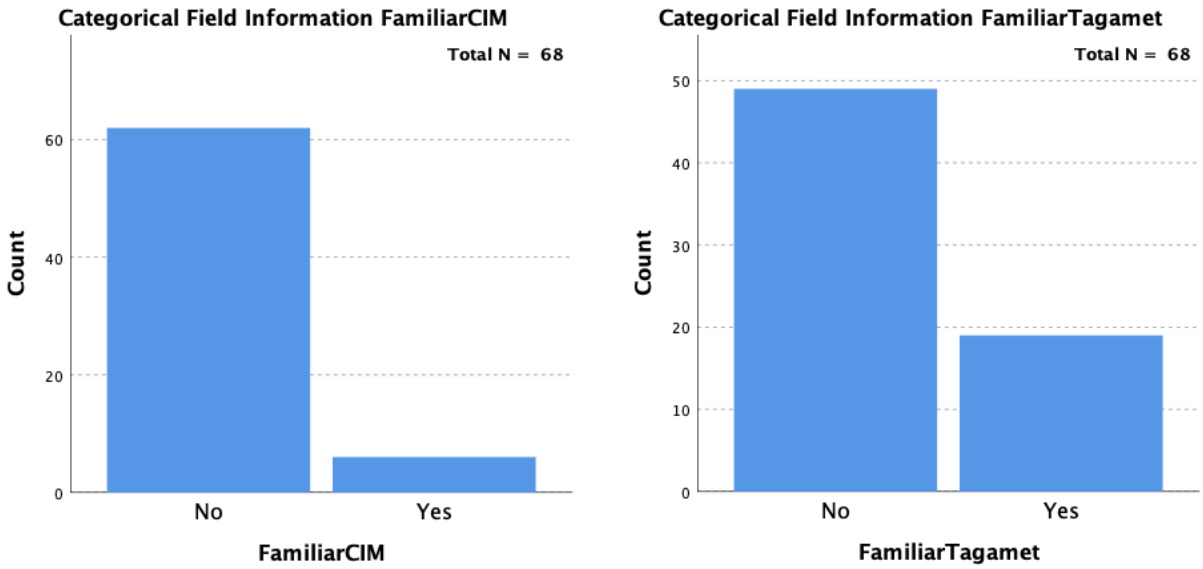


Figure D.3 Frequency counts of familiarity with cimetidine versus Tagamet

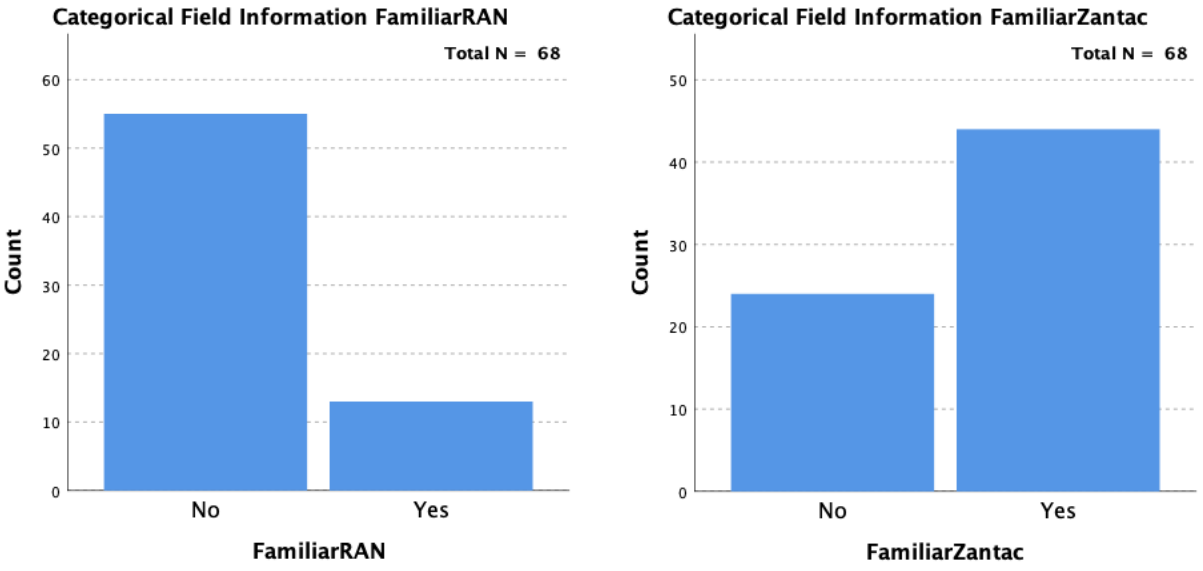


Figure D.4 Frequency counts of familiarity with Ranitidine versus Zantac

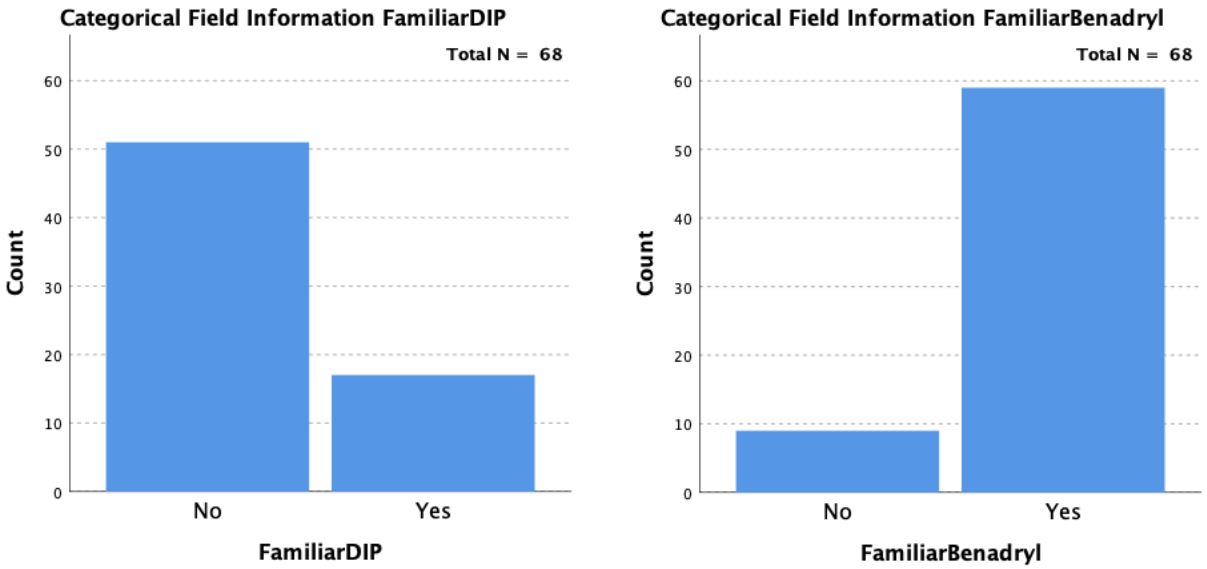


Figure D.5 Frequency counts of familiarity with Diphenhydramine versus Benadryl

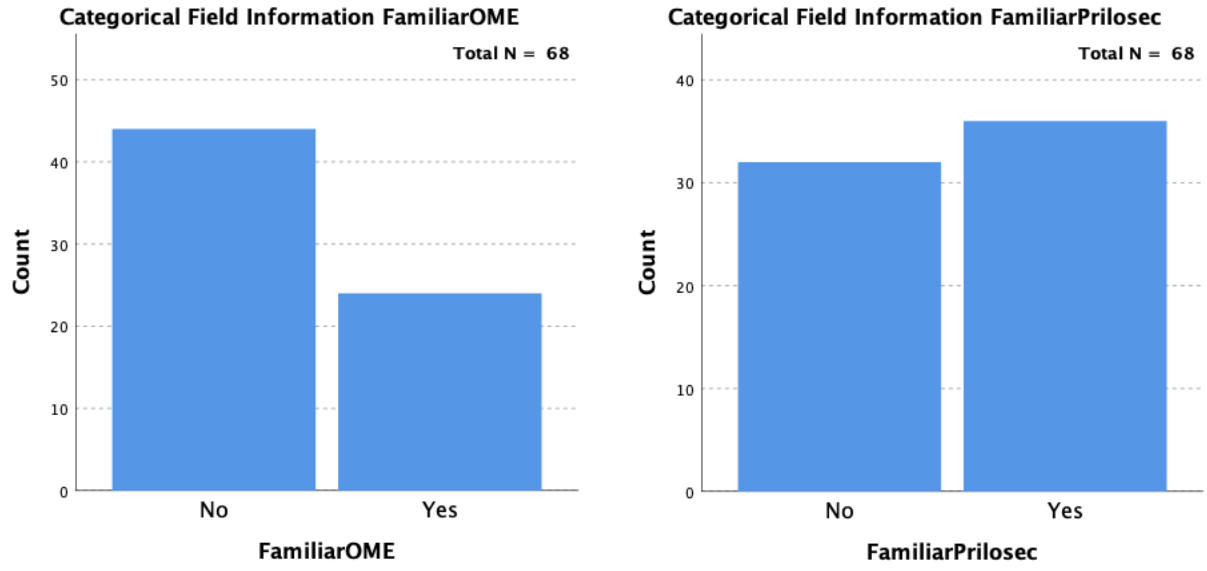


Figure D.6 Frequency counts of familiarity with Omeprazole versus Prilosec

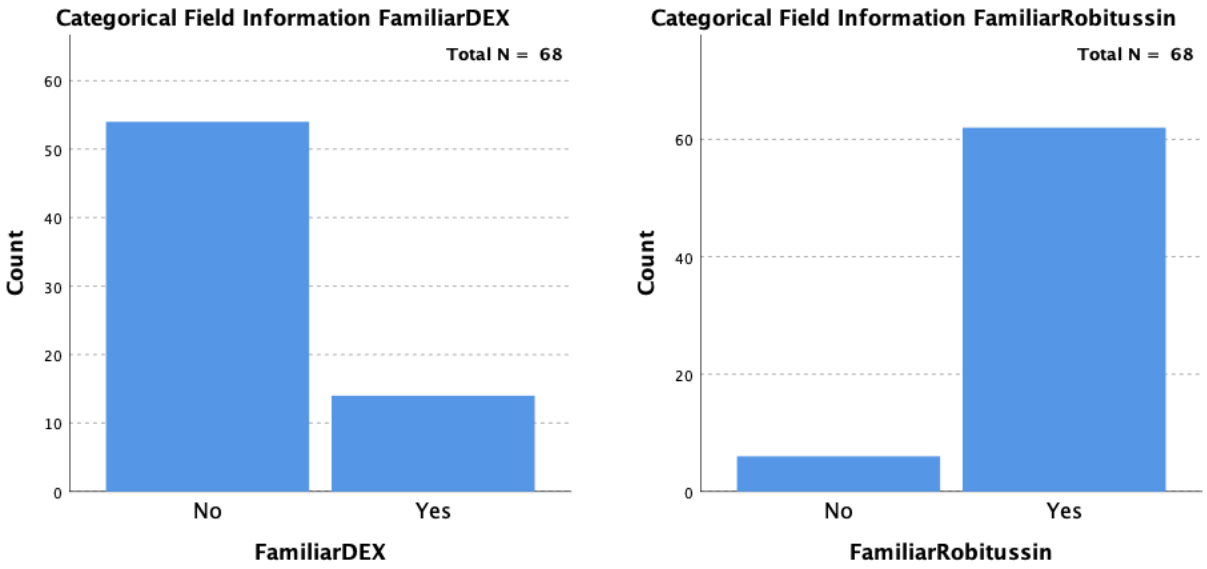


Figure D.7 Frequency counts of familiarity with dextromethorphan versus Robitussin

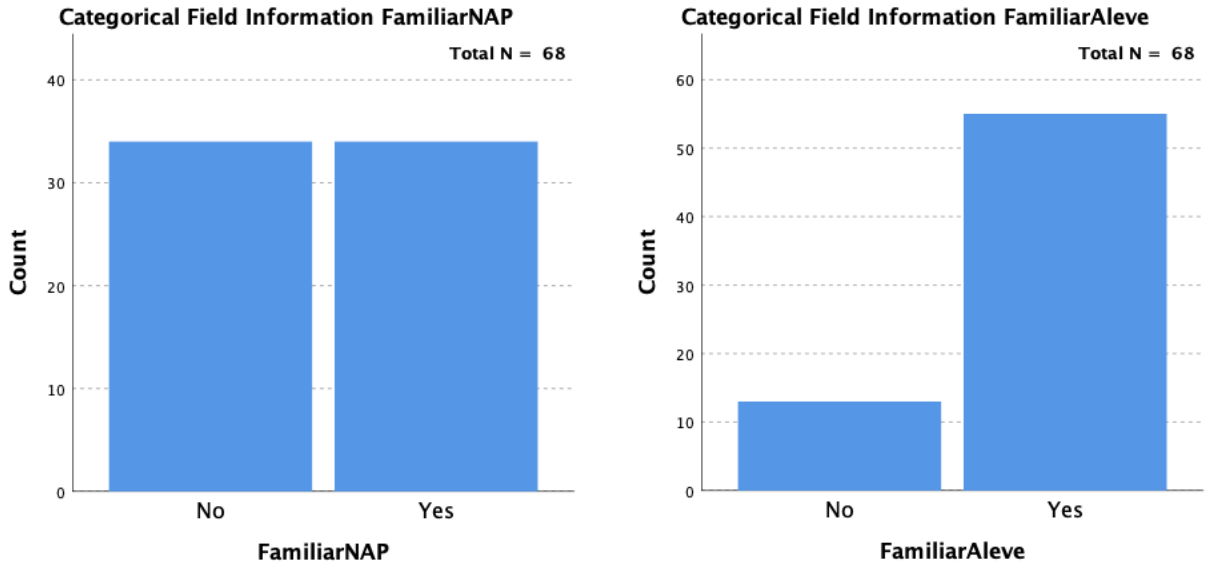


Figure D.8 Frequency counts of familiarity with naproxen versus Aleve

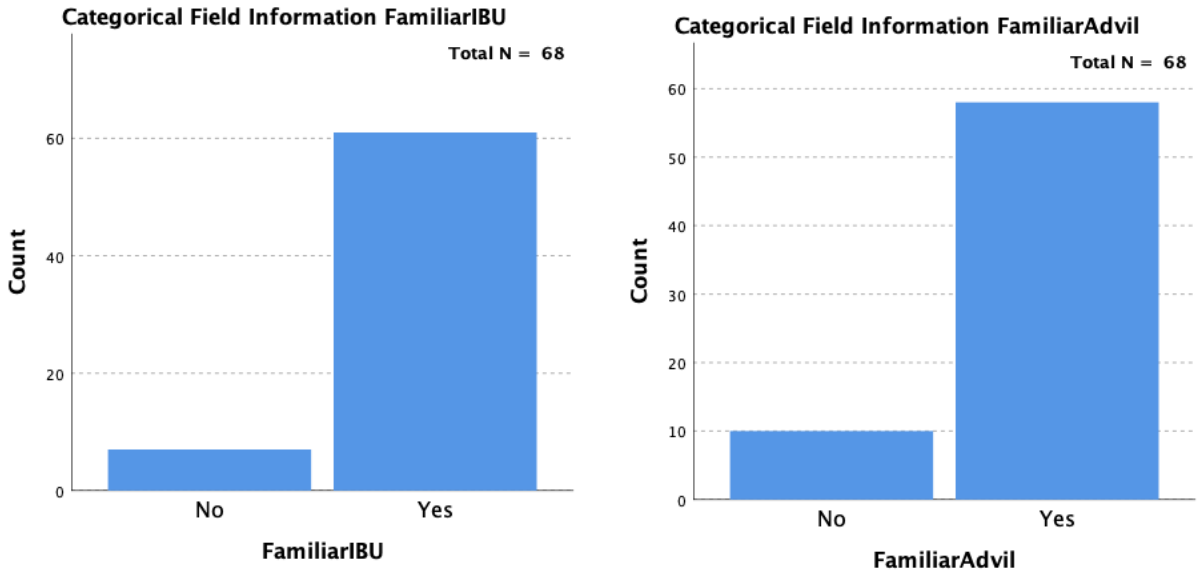


Figure D.9 Frequency counts of familiarity with ibuprofen versus Advil

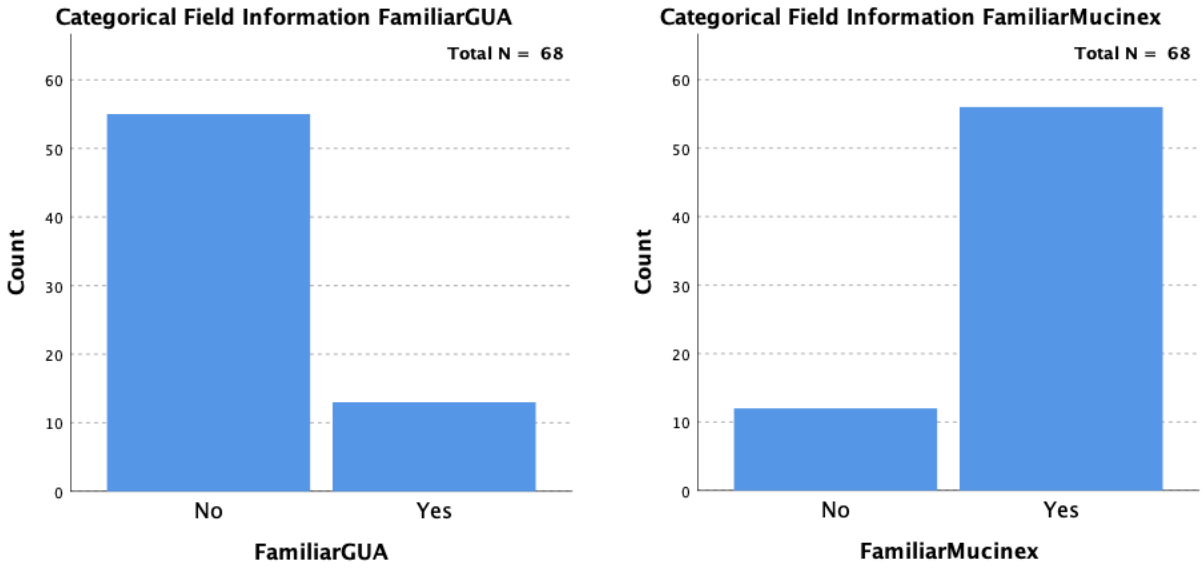


Figure D.10 Frequency counts of familiarity with guaifenesin versus Mucinex

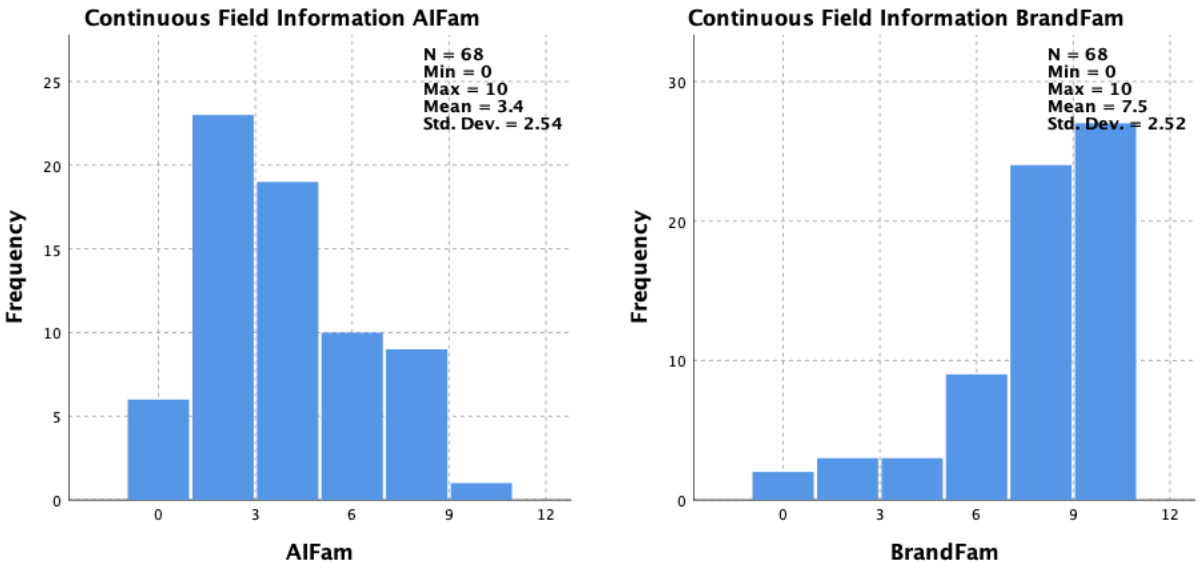


Figure D.11 Distributions of participants overall Active Ingredient familiarity versus overall brand familiarity

APPENDIX E:

Questions used in the Forced Choice Tasks

Table E.1 Questions Used in the Yes/No Forced Choice Task	
Question Type	Question Text
AI	Does one tablet of this medication contain 200mg of active ingredient?
AI	Does one tablet of this medication contain 20mg of active ingredient?
AI	Does this contain Acetaminophen?
AI	Does this contain Cimetidine?
AI	Does this contain Dextromethorphan?
AI	Does this contain Ibuprofen?
AI	Does this contain Naproxen?
AI	Does this contain Omeprazole?
AI	Does this contain Phenylephrine?
AI	Does this contain Ranitidine?
AI	In each tablet, is there 200mg of active ingredient?
AI	In one tablet of this medication, is there 75mg of the active ingredient?
AI	Is there 10mg of active ingredient in each tablet?
AI	Is there 220mg of active ingredient in each tablet?
AI	Is there 30mg of active ingredient in each tablet?
DD	Is there 325mg of active ingredient in one tablet?
DD	Is the chance of stomach bleeding when taking this medication higher if you are age 60 or older?
DD	Is the risk of stomach bleeding greater if you are taking a blood thinning drug?
DD	Should some avoid this medication if allergic to acetaminophen?
DD	Should someone ask a doctor before taking if they are also using antifungal medication?
DD	Should someone avoid consuming 3 or more alcoholic drinks while taking this medication?
DD	Should someone avoid this product if they have trouble swallowing food?
DD	Should someone avoid this product right before or after heart surgery?
DD	Should someone consult their doctor if they have a chronic cough with too much phlegm?
DD	Should this medication be avoided by someone using a prescription for Parkinson's disease?
DD	Should this medication be avoided by someone using certain drugs for depression?
DD	Should you ask your doctor before taking this medication if you have a chronic cough?
DD	Should you consult a doctor before taking this medication if you suffer from chest pain and shortness of breath?
DD	Should you consult a doctor before taking this medication if you suffer from nausea or vomitting?
DD	Should you consult your doctor before taking this medication if you suffer from kidney disease?
DD	Should you contact a doctor if you have unexplained headaches or nausea?
DD	Should you contact a doctor if you have unexplained weight loss while on this medication?
DD	Is the risk of stomach bleeding greater if you are taking a blood thinning drug?
Distraction	Does this medicine prevent heartburn due to eating certain foods?

Table E.1 (cont'd)	
Distraction	Does this medicine relieve heart burn associated with sour stomach?
Distraction	Does this medicine temporarily reduce a fever?
Distraction	Does this medicine temporarily relieve minor aches and pains due to backache?
Distraction	Does this medicine temporarily relieve pain associated with headache?
Distraction	Does this medicine temporarily relieve sinus pressure?
Distraction	Does this medicine temporarily relieve the impulse to cough?
Distraction	Does this medicine treat frequent heartburn occurring 2 or more days in a week?

Table E.2 Questions for the Cross-Product Comparison Task	
Question Type	Question Text
AI	Which medication contains Acetaminophen?
AI	Which medication contains Cimetidine?
AI	Which medication contains Dextromethorphan?
AI	Which medication contains Ibuprofen?
AI	Which medication contains Naproxen?
AI	Which medication contains Omeprazole?
AI	Which medication contains Phenylephrine?
AI	Which medication contains Ranitidine?
AI	Which medication contains 20mg of active ingredient?
AI	Which medication contains 200mg of active ingredient?
AI	Which medication contains 200mg of active ingredient?
AI	Which medication contains 75mg of active ingredient?
AI	Which medication contains 10mg of active ingredient?
AI	Which medication contains 220mg of active ingredient?
AI	Which medication contains 30mg of active ingredient?
AI	Which medication contains 325mg of active ingredient?
DD	Which medication should someone avoid if allergic to acetaminophen?
DD	Which medication should someone ask a doctor before taking if they are also using antifungal medication?
DD	Which medication should someone avoid consuming 3 or more alcoholic drinks while taking?
DD	Which medication should someone avoid if they have trouble swallowing food?
DD	Which medication should someone avoid right before or after heart surgery?
DD	Which medication should someone consult their doctor about if they have a chronic cough with too much phlegm?
DD	Which medication should be avoided by someone using a prescription for Parkinson's disease?
DD	Which medication should be avoided by someone using certain drugs for depression?

Table E.2 (cont'd)	
DD	Which medication should you ask your doctor about before taking if you have a chronic cough?
DD	Which medication should you consult a doctor before taking if you suffer from chest pain and shortness of breath?
DD	Which medication should you consult a doctor about before taking if you suffer from nausea or vomiting?
DD	Which medication should you consult your doctor about before taking if you suffer from kidney disease?
DD	Which medication should you contact a doctor about if you have unexplained headaches or nausea?
DD	Which medication should you contact a doctor about if you have unexplained weight loss?
DD	Which has a higher chance of stomach bleeding when taking the medication, if you are age 60 or older?
DD	Which medication has a higher risk of stomach bleeding if you are taking a blood thinning drug?
Distraction	Which of these medicines relieves heartburn associated with sour stomach?
Distraction	Which of these medicines relieves heartburn due to eating certain foods?
Distraction	Which of these medicines temporarily reduces a fever?
Distraction	Which of these medicines temporarily relieves minor aches and pains due to backache?
Distraction	Which of these medicines temporarily relieves pain associated with headache?
Distraction	Which of these medicines temporarily relieves sinus pressure?
Distraction	Which of these medicines temporarily relieves the impulse to cough?
Distraction	Which of these medicines treats frequent heartburn occurring 2 or more days in a week?

REFERENCES

REFERENCES

Aker, J., Beck, M., Travis, S., & Harris, J. (2014). Consumer Navigation and Selection Behaviors for OTC Products in a Retail Setting, (September), 35. Retrieved from https://www.chpa.org/PDF/connav.aspx

Albert, S. M., Bix, L., Bridgeman, M. M., Carstensen, L. L., Dyer-Chamberlain, M., Neafsey, P. J., & Wolf, M. S. (2014). Promoting safe and effective use of OTC medications: CHPA-GSA national summit. *Gerontologist*, *54*(6), 909–918. https://doi.org/10.1093/geront/gnu034

Amoako, E. P., Richardson-Campbell, L., & Kennedy-Malone, L. (2003). Self-Medication with Over-the-Counter Drugs Among Elderly Adults. *Journal of Gerontological Nursing*, 10–16.

Asseray, N., Ballereau, F., Trombert-Paviot, B., Bouget, J., Foucher, N., Renaud, B., … Queneau, P. (2013). Frequency and severity of adverse drug reactions due to self-medication: A cross-sectional multicentre survey in emergency departments. *Drug Safety*, *36*, 1159–1168. https://doi.org/10.1007/s40264-013-0114-y

Atkin, P., Veitch, P., Veitch, E., & Ogle, S. (1999). The Epidemiology of Serious Adverse Drug Reactions Among the Elderly: EBSCOhost. *Drugs & Aging*, *14*(2), 141–152.

Austrailian Government Department of Health. Guideline for the labeling of medicines (2011).

Austrailian Government Department of Health and Aging. (2011). *TGA Medicine Labelling and Packaging Review*.

Ayanoglu, H., Duarte, E., Noriega, P., Teixeira, L., & Rebelo, F. (2012). The importance of integrating perceived affordances and hazard perception in package design. *Advances in Usability Evaluation Part I*, 627–636.

Bailey, S. C., Agarwal, N., Sleath, B., Gumusoglu, S., & Wolf, M. S. (2011). Improving drug labeling and counseling for limited english proficient adults; 22080698. *Journal of Health Care for the Poor and Underserved*, *22*(4), 1131–1143. https://doi.org/10.1353/hpu.2011.0145

Bailey, S. C., Navaratnam, P., Black, H., Russell, A. L., & Wolf, M. S. (2015). Advancing Best Practices for Prescription Drug Labeling. *Annals of Pharmacotherapy*, *49*(11), 1222–1236. https://doi.org/10.1177/1060028015602272

Bass, P. F., Wilson, J. F., & Griffith, C. H. (2003). A Shortened Instrument for Literacy Screening. *Journal of General Internal Medicine*, *18*, 1036–1038. https://doi.org/10.1111/j.1525-1497.2003.10651.x

Becker, M. W., Bello, N. M., Sundar, R. P., Peltier, C., & Bix, L. (2015). Front of pack labels enhance attention to nutrition information in novel and commercial brands. *Food Policy*, *56*, 76–86. https://doi.org/10.1016/j.foodpol.2015.08.001

Becker, M. W., Prashant Sundar, R., Bello, N., Alzahabi, R., Weatherspoon, L., & Bix, L. (2016). Assessing Attentional Prioritization of Front-of-Pack Nutrition Labels using Change Detection. *Applied Ergonomics*, (54), 90–99. https://doi.org/10.1016/j.apergo.2015.11.014

Bendall, R. C. A., & Thompson, C. (2015). Emotion has no impact on attention in a change detection flicker task. *Frontiers in Psychology*, *6*(OCT), 1–9. https://doi.org/10.3389/fpsyg.2015.01592

Bhargava, S., Buzzell, C., Charm, T., Das, R., Fradin, M., Grimmelt, A., … Seid, C. (2020). A global view of how consumer behavior is changing amid COVID-19. *McKinsey & Company*, 1–15. Retrieved from https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/a-global-view-of-how-consumer-behavior-is-changing-amid-covid-19

Bialkova, S., Grunert, K. G., Juhl, H. J., Wasowicz-Kirylo, G., Stysko-Kunkowska, M., & van Trijp, H. C. M. (2014). Attention mediates the effect of nutrition label information on consumers' choice: Evidence from a choice experiment involving eye-tracking. *Appetite*, *76*, 66–75. https://doi.org/10.1016/j.appet.2013.11.021

Bialkova, S., Grunert, K. G., & van Trijp, H. (2013). Standing out in the crowd: The effect of information clutter on consumer attention for front-of-pack nutrition labels. *Food Policy*, *41*, 65–74. https://doi.org/10.1016/j.foodpol.2013.04.010

Bialkova, S., & van Trijp, H. (2010). What determines consumer attention to nutrition labels? *Food Quality and Preference*, *21*(8), 1042–1051. https://doi.org/10.1016/j.foodqual.2010.07.001

Bix, L., Bello, N. M., Auras, R., Ranger, J., & Lapinski, M. K. (2009). Examining the conspicuousness and prominence of two required warnings on OTC pain relievers. *Proceedings of the National Academy of Sciences*, *106*(16), 6550–6555. https://doi.org/10.1073/pnas.0810665106

Bix, L., Kosugi, W., Bello, N., Sundar, R., & Becker, M. (2010). The Use of Change Detection as a Method of Objectively Evaluating Labels. *Packaging Technology and Science*, (23), 393–401.

Bix, L., Seo, D. C., Ladoni, M., Brunk, E., & Becker, M. W. (2016). Evaluating varied label designs for use with medical devices: Optimized labels outperform existing labels in the correct selection of devices and time to select. *PLoS ONE*, *11*(11), 1–11. https://doi.org/10.1371/journal.pone.0165002

Bix, L., Sundar, R. P., Bello, N. M., Peltier, C., Weatherspoon, L. J., & Becker, M. W. (2015). To see or not to see: Do front of pack nutrition labels affect attention to overall nutrition information? *PLOS ONE*, 1–20. https://doi.org/10.1371/journal.pone.0139732

Bojka, A., Gaddy, C., Lew, G., Quinn, A., & Israelski, E. (2005). Evaluation of Drug Label Designs Using Eye Tracking. *Human Factors*, 1033–1037.

Bongard, V., Menard-Tache, S., Bagheri, H., Kabiri, K., Lapeyre-Mestre, M., & Montastruc, J. L. (2002). Perception of the risk of adverse drug reactions: differences between health professionals and non health professionals. *British Journal of Clinical Pharmacology*, *54*(June), 433–436. https://doi.org/10.1046/j.1365-2125.2002.01674.x

Bourgeois, F. T., Shannon, M. W., Valim, C., & Mandl, K. D. (2010). Adverse drug events in the outpatient setting: an 11-year national analysis. *Pharmacoepidemiology and Drug Safety*, (19), 901–910. https://doi.org/10.1002/pds.1984

Bown, D., Kisuule, G., Ogasawara, H., Siregar, C., & Williams, G. (2000). WHO guidelines for the regulatory assessment of medicinal products for use in self medication. *WHO Drug Information*, *14*(1), 18–26. https://doi.org/10.1016/S0009-9236(03)00239-X

Brass, E. P., & Weintraub, M. (2003). Label development and the label comprehension study for over-the-counter drugs. *Clinical Pharmacology and Therapeutics*, *74*(5), 406–412. https://doi.org/10.1016/S0009-9236(03)00239-X

Cadigan, D. A., Magaziner, J., & Fedder, D. O. (1989). Polymedicine use among community resident older women: How much of a problem? *American Journal of Public Health*, *79*(11), 1537–1540. https://doi.org/10.2105/AJPH.79.11.1537

Calamusa, A., Di Marzio, A., Cristofani, R., Arrighetti, P., Santaniello, V., Alfani, S., & Carducci, A. (2012). Factors that influence Italian consumers' understanding of over-the-counter medicines and risk perception. *Patient Education and Counseling*, *87*(3), 395–401. https://doi.org/10.1016/j.pec.2011.10.003

Card, S. K., Moran, T. P., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale. NJ: Lawrence Erlbaum Associates, Inc. Publishers.

Carpenter, S. M., & Yoon, C. (2012). Aging and Consumer Decision Making. *Annals of the New York Academy of Sciences*, *1235*(1), 1–15. https://doi.org/10.1111/j.1749-6632.2011.06390.x.Aging

Cheatham, D. B., & Wogalter, M. S. (2002). Reported Likelihood of Reading Over-the-Counter (OTC) Medication Labeling and Contacting a Physician. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 1452–1456).

Consumer Healthcare Products Association. (2010). *CHPA Your Health at Hand : Perceptions of over-the-counter medicine in the U . S .*

Costello, M. C., Madden, D. J., Mitroff, S. R., & Whiting, W. L. (2010). Age-related decline of visual processing components in change detection. *Psychology and Aging*, *25*(2), 356–368. https://doi.org/10.1037/a0017625

Cryer, B., Barnett, M. A., Wagner, J., & Wilcox, C. M. (2016). Overuse and Misperceptions of Nonsteroidal Anti-inflammatory Drugs in the United States. *The American Journal of the Medical Sciences*, *352*(5), 472–480. https://doi.org/10.1016/j.amjms.2016.08.028

Davies, E. A., & O'Mahony, M. S. (2015). Adverse drug reactions in special populations - The elderly. *British Journal of Clinical Pharmacology*, *80*(4), 796–807. https://doi.org/10.1111/bcp.12596

Davis, T. C., Federman, A. D., Bass, P. F., Jackson, R. H., Middlebrooks, M., Parker, R. M., & Wolf, M. S. (2009). Improving patient understanding of prescription drug label instructions. *Journal of*

General Internal Medicine, *24*(1), 57–62. https://doi.org/10.1007/s11606-008-0833-4

Davis, T. C., Wolf, M. S., Bass, P. F., Middlebrooks, M., Kennen, E., Baker, D. W., … Parker, R. M. (2006). Low literacy impairs comprehension of prescription drug warning labels. *Journal of General Internal Medicine*, *21*(8), 847–851. https://doi.org/10.1111/j.1525-1497.2006.00529.x

Davis, T. C., Wolf, M. S., Bass, P. F., Thompson, J. A., Tilson, H. H., Neuberger, M., & Parker, R. M. (2006). Literacy and misunderstanding prescription drug labels. *Annals of Internal Medicine*, *145*(12), 887–894. https://doi.org/10.7326/0003-4819-145-12-200612190-00144

de la Fuente, J. (2013). *Usability of Tabs in Semi-Rigid Packaging*. Michigan State Univeristy.

de la Fuente, J., Gustafson, S., Twomey, C., & Bix, L. (2015). An affordance-based methodology for package design. *Packaging Technology and Science*, *28*(2), 157–171. https://doi.org/10.1002/pts.2087

DeHenau, C., Becker, M. W., Bello, N. M., Liu, S., & Bix, L. (2016). Tallman lettering as a strategy for differentiation in look-alike, sound-alike drug names: The role of familiarity in differentiating drug doppelgangers. *Applied Ergonomics*, *52*, 77–84. https://doi.org/10.1016/j.apergo.2015.06.009

Dejoy, D. M. (1991). Warnings process derived. *Proceedings of the Human Factors Society 35th Annual Meeting*, 1043–1047.

Department of Health Therapeutic Goods Administration. Therapeutic Goods Order No. 92 - Standard for labels of non-prescription medicines, 92 § (2016). Retrieved from https://www.legislation.gov.au/Details/F2016L01287

Dieleman, J. L., Squires, E., Bui, A. L., Campbell, M., Chapin, A., Hamavid, H., … Murray, C. J. L. (2017). Factors associated with increases in US health care spending, 1996-2013. *JAMA - Journal of the American Medical Association*, *318*(17), 1668–1678. https://doi.org/10.1001/jama.2017.15927

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, Supplement*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Edwards, I. R., & Aronson, J. K. (2000). Adverse drug reactions: definitions, diagnosis, and management. *Lancet (London, England)*, *356*(9237), 1255–1259. https://doi.org/10.1016/S0140-6736(00)02799-9

Esfahanian, S. (2020). *A Patient-Centered Approach to Labeling for Over-the-Counter Medications: Using Data to Drive Design Decisions for the Benefit of Older Adults*. Michigan State Univeristy.

Federman, A. D., Sano, M., Wolf, M. S., Siu, A. L., & Halm, E. A. (2009). Health literacy and cognitive performance in older adults. *Journal of the American Geriatrics Society*, *57*(8), 1475–1480. https://doi.org/10.1111/j.1532-5415.2009.02347.x

Fick, D. M., Semla, T. P., Steinman, M., Beizer, J., Brandt, N., Dombrowski, R., … Sandhu, S. (2019).

American Geriatrics Society 2019 Updated AGS Beers Criteria® for Potentially Inappropriate Medication Use in Older Adults. *Journal of the American Geriatrics Society*, *67*(4), 674–694. https://doi.org/10.1111/jgs.15767

Food and Drug Administration. (2006). A History of the FDA and Drug Regulation in the United States. Retrieved from http://www.fda.gov/centennial/history/history.html

Fraeyman, J. (2015). Consumer Choice Between Common Generic and Brand Medicines in a Country with a Small Generic Market. *Journal of Managed Care & Specialty Pharmacy*, *21*(4).

Franceschi, M., Scarcelli, C., Niro, V., Seripa, D., Pazienza, A. M., Pepe, G., … Pilotto, A. (2008). Prevalence, Clinical Features and Avoidability of Adverse Drug Reactions as Cause of Admission to a Geriatric Unit A prospective study of 1756 patients. *Drug Safety*, *31*(6), 545–556. https://doi.org/10.2165/00002018-200831060-00009

Gawasane, A., Bix, L., de la Fuente, J., Sundar, R. P., & Smith, T. J. (2012). Consumer attention to an over-the-counter warning in four different styles of design. *Packaging Technology and Science*, *25*(7), 385–396. https://doi.org/10.1002/pts.979

Gotelli, N. J., & Ellison, a. . (2004). Framing and testing hypotheses. *A Primer of Ecological Statistics*, 79–106.

Goyal, R. K., Rajan, S. S., Essien, E. J., & Sansgiry, S. S. (2012). Effectiveness of FDA's new over-the-counter acetaminophen warning label in improving consumer risk perception of liver damage. *Journal of Clinical Pharmacy and Therapeutics*. https://doi.org/10.1111/j.1365-2710.2012.01371.x

Guthrie, B., Makubate, B., Hernandez-Santiago, V., & Dreischulte, T. (2015). The rising tide of polypharmacy and drug-drug interactions: population database analysis 1995-2010. *BMC Medicine*, *13*(74). https://doi.org/10.1186/s12916-015-0322-7

Halme, M., Linden, Ka., & Kääriä, K. (2009). Patients' Preferences for Generic and Branded Over-the-Counter Medicines. *The Patient: Patient-Centered Outcomes Research*, *2*(4).

Hanoch, Y., Gummerum, M., & Brass, E. P. (2007). American and German Students ' Knowledge , Perceptions , and Behaviors With Respect to Over-the-Counter Pain Relievers. *Health Psychology*, *26*(6), 802–806. https://doi.org/10.1037/0278-6133.26.6.802

Harben, A., Esfahanian, S., & Bix, L. L. (2018). A Patient Centered Approach to Labeling for Over-the-Counter Medications: Using Data to Drive Design Decisions for the Benefit of Older Adults. In *Graduate Academic Conference*. East Lansing.

Hardie, N. A., Kyanko, K., Busch, S., Losasso, A. T., & Levin, R. A. (2011). Health literacy and health care spending and utilization in a consumer-driven health plan. *Journal Health Communication*, *16 Suppl 3*(1087-0415 (Electronic)), 308–321. https://doi.org/10.1080/10810730.2011.604703

Hawley, K. L., Roberto, C. A., Bragg, M. A., Liu, P. J., Schwartz, M. B., & Brownell, K. D. (2013). The science on front-of-package food labels. *Public Health Nutrition*, *16*(3), 430–439. https://doi.org/10.1017/S1368980012000754

Health Canada. (2017). Good Label and Package Practices Guide for Non-prescription Drugs and Natural Health Products. Retrieved from https://www.canada.ca/content/dam/hc-sc/documents/services/drugs-health-products/reports-publications/medeffect-canada/good-label-package-practices-guide-eng.pdf

Hellier, E., Edworthy, J., Derbyshire, N., & Costello, A. (2006). Considering the impact of medicine label design characteristics on patient safety. *Ergonomics*, *49*(5–6), 617–630. https://doi.org/10.1080/00140130600568980

Hersey, J. C., Wohlgenant, K. C., Arsenault, J. E., Kosa, K. M., & Muth, M. K. (2013). Effects of front-of-package and shelf nutrition labeling systems on consumers. *Nutrition Reviews*, *71*(1), 1–14. https://doi.org/10.1111/nure.12000

Hess, C., Linnebur, S. A., Rhyne, D. N., & Valdez, C. A. (2016). Over-the-Counter Drugs to Avoid in Older Adults with Kidney Impairment. *Continuing Nursing Education*, *43*(5).

Hoffman, L., Atchley, P., McDowd, J. M., & Dubinsky, R. (2005). The role of visual attention in predicting driving impairment in older adults. *Psychology and Aging*, *20*(4), 610–622. https://doi.org/10.1037/0882-7974.20.4.610

Holden, R. J., Srinivas, P., Campbell, N. L., Clark, D. O., Bodke, K. S., Hong, Y., … Callahan, C. M. (2018). Understanding older adults' medication decision making and behavior: A study on over-the-counter (OTC) anticholinergic medications. *Research in Social and Administrative Pharmacy*, (January), 0–1. https://doi.org/10.1016/j.sapharm.2018.03.002

Hoy, M. G., & Levenshus, A. B. (2018). A mixed-methods approach to assessing actual risk readership on branded drug websites. *Journal of Risk Research*, *9877*, 1–18. https://doi.org/10.1080/13669877.2016.1223160

IBM Support. (2020). Corrected Model Sums of Squares in UNIANOVA and GLM Multivariate. Retrieved from https://www.ibm.com/support/pages/corrected-model-sums-squares-unianova-and-glm-multivariate

International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. (1994). CLINICAL SAFETY DATA MANAGEMENT: DEFINITIONS AND STANDARDS FOR EXPEDITED REPORTING E2A. In *Efficacy Guidelines* (p. 12). Retrieved from http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf

ISO 9241-11. (2018). INTERNATIONAL STANDARD Usability : Definitions and concepts, *2018*.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10–12), 1489–1506. https://doi.org/10.1016/S0042-6989(99)00163-7

Itti, L., & Koch, C. (2001). Computational Modelling of Visual Attention. *Nature Reviews Neuroscience*,

2(March), 1–10. https://doi.org/10.1038/35058500

Jacobsen, L. a, Kent, M., Lee, M., & Mather, M. (2011). America's aging population. *Population Bulletin*, *66*(1), 1–18.

Jerez-Roig, J., Medeiros, L. F. B., Silva, V. A. B., Bezerra, C. L. P. A. M., Cavalcante, L. A. R., Piuvezam, G., & Souza, D. L. B. (2014). Prevalence of Self-Medication and Associated Factors in an Elderly Population: A Systematic Review. *Drugs and Aging*, *31*(12), 883–896. https://doi.org/10.1007/s40266-014-0217-x

Johnson, M. S., & Drungle, S. C. (2000). Purchasing OTC medication: The influence of age and familiarity. *Experimental Aging Research*, *26*(September), 245–261. https://doi.org/10.1080/036107300404886

Jones, G., & Richardson, M. (2007). An objective examination of consumer perception of nutrition information based on healthiness ratings and eye movements. *Public Health Nutrition*, *10*(3), 238–244. https://doi.org/10.1017/S1368980007258513

Kanter, R., Vanderlee, L., & Vandevijvere, S. (2018). Front-of-package nutrition labelling policy: Global progress and future directions. *Public Health Nutrition*, *21*(8), 1399–1408. https://doi.org/10.1017/S1368980018000010

Katzman, R., Brown, T., Fuld, P., Peck, A., Schechter, R., & Schimmel, H. (1983). Validation of a Short Orientation-Memory-Concentration Test of Cognitive Impairment. *American Journal of Psychiatry*, 734–739.

Kauppinen-Räisänen, H., Owusu, R. A., & Bamfo, B. A. (2012). Brand salience of OTC pharmaceuticals through package appearance. *International Journal of Pharmaceutical and Healthcare Marketing*, *6*(3), 230–249. https://doi.org/10.1108/17506121211259403

Kelly, B., Hughes, C., Chapman, K., Chun Yu Louie, J., Dixon, H., Crawford, J., … Slevin, T. (2009). Consumer testing of the acceptability and effectiveness of front-of-pack food labelling systems for the Australian grocery market. *Health Promotion International*, *24*(2), 120–129. https://doi.org/10.1093/heapro/dap012

Kinchla, R. A., & Wolfe, J. M. (1979). The order of visual processing:"Top-down,""bottom-up," or "middle-out." *Perception & Psychophysics*, *25*(3), 225–231. https://doi.org/10.3758/BF03202991

King, J. P., Davis, T. C., Bailey, S. C., Jacobson, K. L., Hedlund, L. A., Di Francesco, L., … Wolf, M. S. (2011). Developing consumer-centered, nonprescription drug labeling: A study in acetaminophen. *American Journal of Preventive Medicine*, *40*(6), 593–598. https://doi.org/10.1016/j.amepre.2011.02.016

Kobayashi, L. C., Wardle, J., Wolf, M. S., & Von Wagner, C. (2016). Aging and Functional Health Literacy: A Systematic Review and Meta-Analysis. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *71*(3), 445–457. https://doi.org/10.1093/geronb/gbu161

Koenigstorfer, J., Wąsowicz-Kiryło, G., Styśko-Kunkowska, M., & Groeppel-Klein, A. (2014). Behavioural effects of directive cues on front-of-package nutrition information: The combination matters! *Public Health Nutrition*, *17*(9), 2115–2121. https://doi.org/10.1017/S136898001300219X

Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (2000). *To Err Is Human*. https://doi.org/10.17226/9728

Lavan, A. H., & Gallagher, P. (2016). Predicting risk of adverse drug reactions in older adults. *Therapeutic Advances in Drug Safety*, *7*(1), 11–22. https://doi.org/10.1177/2042098615615472

Lee, J., Ladoni, M., Richardson, J., Sundar, R. P., & Bix, L. (2019). Investigating the efficacy of an interactive warning for use in labeling strategies used by us pharmacies. *Pharmacy Practice*, *17*(2), 1–13. https://doi.org/10.18549/PharmPract.2019.2.1463

Levy, D. E., Riis, J., Sonnenberg, L. M., Barraclough, S. J., & Thorndike, A. N. (2012). Food choices of minority and low-income employees: A cafeteria intervention. *American Journal of Preventive Medicine*, *43*(3), 240–248. https://doi.org/10.1016/j.amepre.2012.05.004

Liu, L. (2016). *The Effect of Labeling Content and Prominence on Information Processing among Older Adults during Self-Selection of Over-the-Counter Medications*. Michigan State University.

Llewelyn, H., Ang, H. A., Lewis, K., & Al-Abdullah, A. (2007). Oxford Handbook of Clinical Diagnosis. In *Oxford Handbook of Clinical Diagnosis* (3rd ed., Vol. 76, p. 63). https://doi.org/10.1093/med/9780199679867.001.0001

Mannesse, C. K., Derkx, F. H. M., de Ridder, M. A. J., Man In 'T Veld, A. J., & Van Der Cammen, T. J. M. (2000). Contribution of adverse drug reactions to hospital admission of older patients. *Age and Ageing*, *29*(1), 35–39. https://doi.org/10.1093/ageing/29.1.35

Marcum, Z. A., Amuan, M. E., Hanlon, J. T., Aspinall, S. L., Handler, S. M., Ruby, C. M., & Pugh, M. J. V. (2012). Prevalence of unplanned hospitalizations caused by adverse drug reactions in older veterans. *Journal of the American Geriatrics Society*, *60*(1), 34–41. https://doi.org/10.1111/j.1532-5415.2011.03772.x

Masnoon, N., Shakib, S., Kalisch-Ellett, L., & Caughey, G. E. (2017). What is polypharmacy? A systematic review of definitions. *BMC Geriatrics*, *17*(1), 1–10. https://doi.org/10.1186/s12877-017-0621-2

McCarley, J. S., Vais, M. J., Pringle, H., Kramer, A. F., Irwin, D. E., & Strayer, D. L. (2004). Conversation disrupts change detection in complex traffic scenes. *Human Factors*, *46*(3), 424–436. https://doi.org/10.1518/hfes.46.3.424.50394

McNeil Consumer Healthcare. (2015). Americans Should Pay More Attention to Over-the-Counter (OTC) Medicine Labels According to New Survey. Retrieved from https://www.prnewswire.com/news-releases/americans-should-pay-more-attention-to-over-the-counter-otc-medicine-labels-according-to-new-survey-300160108.html

Milosavljevic, M., Navalpakkam, V., Koch, C., & Rangel, A. (2012). Relative visual saliency differences induce sizable bias in consumer choice. *Journal of Consumer Psychology*, *22*(1), 67–74. https://doi.org/10.1016/j.jcps.2011.10.002

Mintel. (2018). Managing Your Health - US - January 2018.

Morrell, R. W., Park, D. C., & Poon, L. W. (1989). Quality of instructions on prescription drug labels: Effects on memory and comprehension in young and old adults. *Gerontologist*, *29*(3), 345–354. https://doi.org/10.1093/geront/29.3.345

Mullen, R. J., Curtis, L. M., O'Conor, R., Serper, M., McCarthy, D., Bailey, S. C., … Wolf, M. S. (2018). Visual Acuity, literacy, and unintentional misuse of nonprescription medications. *American Society of Health Systems Pharmacists*, *75*(9), 253–266. https://doi.org/10.1002/pts

Mullen, R. J., Duhig, J., Russell, A., Scarazzini, L., Lievano, F., & Wolf, M. S. (2018). Best-practices for the design and development of prescription medication information: A systematic review. *Patient Education and Counseling*, *101*(8), 1351–1367. https://doi.org/10.1016/j.pec.2018.03.012

Murty, S., & Sansgiry, S. S. (2007). Consumer comprehension of OTC medication labels and the scope for improvement in font size. *Journal of Pharmacy Technology*, *23*, 207–213. https://doi.org/10.1177/875512250702300404

Nair, N. P., Chalmers, L., Peterson, G. M., Bereznicki, B. J., Castelino, R. L., & Bereznicki, L. R. (2016). Hospitalization in older patients due to adverse drug reactions - The need for a prediction tool. *Clinical Interventions in Aging*, *11*, 497–505. https://doi.org/10.2147/CIA.S99097

O'Connor, M. N., Gallagher, P., Byrne, S., & O'Mahony, D. (2012). Adverse drug reactions in older patients during hospitalisation: Are they predictable? *Age and Ageing*, *41*, 771–776. https://doi.org/10.1093/ageing/afs046

Oscanoa, T. J., Lizaraso, F., & Carvajal, A. (2017). Hospital admissions due to adverse drug reactions in the elderly. A meta-analysis. *European Journal of Clinical Pharmacology*, (73), 759–770. https://doi.org/10.1007/s00228-017-2225-3

Pharmacodynamics - an overview. (n.d.). Science Direct. Retrieved January 2, 2021, from https://www.sciencedirect.com/topics/medicine-and-dentistry/pharmacodynamics

Pharmacokinetics - an overview. (n.d.). Retrieved January 2, 2021, from https://www.sciencedirect.com/topics/medicine-and-dentistry/pharmacokinetics

Popescu, G. H. (2014). Economic aspects influencing the rising costs of health care in the United States. *American Journal of Medical Research*, *1*(1), 47–47.

Pringle, H. L., Irwin, D. E., Kramer, A. F., & Atchley, P. (2001). The role of attentional breadth in perceptual change detection. *Psychonomic Bulletin and Review*, *8*(1), 89–95. https://doi.org/10.3758/BF03196143

Qato, D. M., et al., Alexander, G. C., Conti, R. M., Johnson, M., Schumm, P., & Lindau, S. T. (2008). Use of Prescription and Over-the-counter Medications and Dietary Supplements Among Older Adults in the United States. *Jama.*, *300*(24), 2867–2878. https://doi.org/10.1001/jama.2008.892

Qato, D. M., Wilder, J., Schumm, L. P., Gillet, V., & Alexander, G. C. (2016). Changes in prescription and over-the-counter medication and dietary supplement use among older adults in the United States, 2005 vs 2011. JAMA Internal Medicine, 176(4), 473–482. https://doi.org/10.1001/jamainternmed.2015.8581

Raghavan, A., Paliwal, Y., & Slattum, P. W. (2017). Gaps in OTC Labeling: Potentially Inappropriate Medications for Older Adults. *Studies in Higher Education*, *5079*(6), 1–15. https://doi.org/10.1037//0003-066X.46.5.506

Reisenwitz, T. H., & Wimbish, J. (1997). The Purchase Decision Process and Involvement of the Elderly Regarding Nonprescription Products. *Health Marketing Quarterly*, *15*(1), 95–102. https://doi.org/10.1300/J026v15n01

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). TO SEE OR NOT TO SEE : The Need for Attention to Perceive Changes in Scenes. *American Psychological Society*, *8*(5), 368–373.

Roe, B., Levy, A. S., Brenda, M., & Derby, B. M. (1999). The Impact of Health Claims on Consumer Search and Product Evaluation Outcomes: Results from FDA Experimental Data. *Journal of Public Policy & Marketing*, *18*(1), 89–105. https://doi.org/128.210.126.199

Rolita, L., & Freedman, M. (2008). Over-the-Counter medication use in older adults. *Journal of Gerontological Nursing*, *34*(4), 8–17. https://doi.org/10.1097/01.NAJ.0000344046.42145.9d

Rousseau, G. K., Lamson, N., & Rogers, W. A. (1998). Designing warnings to compensate for age-related changes in perceptual and cognitive abilities. *Psychology and Marketing*, *15*(7), 643–662. https://doi.org/10.1002/(SICI)1520-6793(199810)15:7<643::AID-MAR3>3.0.CO;2-F

Routledge, P. A., O'Mahony, M. S., & Woodhouse, K. W. (2004). Adverse drug reactions in elderly patients. *British Journal of Clinical Pharmacology*, *57*(2), 121–126. https://doi.org/10.1046/j.1365-2125.2003.01875.x

Ryan, C. L., & Bauman, K. (2016). Educational Attainment in the United States: 2015. *United States Census Bureau*, *2010*, 20–578. https://doi.org/P20-578

Ryan, M. P., Costa, P. L., & Cruz, A. B. (2017). Under what conditions can a nonprescription drug label serve as refutation text? The role of directed attention and processing strategy. *Health Psychology Open*, *4*(2). https://doi.org/10.1177/2055102917730676

Ryan, M. P., & Costello-White, R. N. (2017). Does the Drug Facts Label for nonprescription drugs meet its design objectives ? A new procedure for assessing label effectiveness. *Health Psychology Open*, 1–19. https://doi.org/10.1177/2055102917720331

Sansgiry, S. S., Cady, P. S., & Shubhada, S. (2001). Consumer Involvement: Effects on Information

Processing from Over-the-Counter Medication Labels. *HEALTH MARKETING QUARTERLY*. https://doi.org/10.1109/TDEI.2009.5211872

Schmiedl, S., Rottenkolber, M., Hasford, J., Rottenkolber, D., Farker, K., Drewelow, B., … Thürmann, P. (2014). Self-medication with over-the-counter and prescribed drugs causing adverse-drug-reaction-related hospital admissions: Results of a prospective, long-term multi-centre study. *Drug Safety*, *37*, 225–235. https://doi.org/10.1007/s40264-014-0141-3

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

Shackel, B. (2009). Usability - Context, framework, definition, design and evaluation. *Interacting with Computers*, *21*(5–6), 339–346. https://doi.org/10.1016/j.intcom.2009.04.007

Shrank, W., Avorn, J., Rolon, C., & Shekelle, P. (2007). Effect of content and format of prescription drug labels on readability, understanding, and medication use: A systematic review. *Annals of Pharmacotherapy*, *41*(5), 783–801. https://doi.org/10.1345/aph.1H582

Shrank, W. H., Agnew-Blais, J., Choudhry, N. K., Wolf, M. S., Kesselheim, A. S., Avorn, J., & Shekelle, P. (2007). The variability and quality of medication container labels. *Archives of Internal Medicine*, *167*(16), 1760–1765. https://doi.org/10.1001/archinte.167.16.1760

Shrank, W. H., & Avorn, J. (2007). Educating patients about their medications: The potential and limitations of written drug information. *Health Affairs*, *26*(3), 731–740. https://doi.org/10.1377/hlthaff.26.3.731

Simonson, I. (1999). The Effect of Product Assortment on Buyer Preferences. *Journal of Retailing*, *75*(3), 347–370. https://doi.org/10.1016/S0022-4359(99)00012-3

Stevenson, F. A., Leontowitsch, M., & Duggan, C. (2008). Over-the-counter medicines: Professional expertise and consumer discourses. *Sociology of Health and Illness*, *30*(6), 913–928. https://doi.org/10.1111/j.1467-9566.2008.01108.x

Sundar, R. P., Becker, M. W., Bello, N. M., & Bix, L. (2012). Quantifying age-related differences in information processing behaviors when viewing prescription drug labels. *PLoS ONE*, *7*(6), 1–8. https://doi.org/10.1371/journal.pone.0038819

The Gerontological Society of America. (2013). Over-the-Counter Medication Behaviors of Older Adults: Research is Needed to Better Understand and Promote Safe and Effective Use (pp. 1–12).

The National Council on Patient Information and Education. (2003). *Uses and Attitudes About Taking Over-the-Counter Medicines :*

Thorndike, A. N., Riis, J., Sonnenberg, L. M., & Levy, D. E. (2014). Traffic-light labels and choice architecture: Promoting healthy food choices. *American Journal of Preventive Medicine*, *46*(2), 143–149. https://doi.org/10.1016/j.amepre.2013.10.002

Tong, V., Raynor, D. K., & Aslani, P. (2014). Design and comprehensibility of over-the-counter product labels and leaflets: a narrative review. International Journal of Clinical Pharmacy, 36, 865–872. https://doi.org/10.1007/s11096-014-9975-0

Tong, V., Raynor, D. K., & Aslani, P. (2015). 'It's all there in black and white' – or is it? Consumer perspectives on the proposed Australian Medicine Information Box over-the-counter label format. *Health Expectations*, *19*, 948–961. https://doi.org/10.1111/hex.12389

Tong, V., Raynor, D. K., & Aslani, P. (2018). Developing alternative over-the-counter medicine label formats: How do they compare when evaluated by consumers? *Research in Social and Administrative Pharmacy*, *14*, 248–261. https://doi.org/10.1016/j.sapharm.2017.03.003

Treue, S. (2003). Visual attention: The where, what, how and why of saliency. *Current Opinion in Neurobiology*. https://doi.org/10.1016/S0959-4388(03)00105-3

Trivedi, H., Trivedi, A., & Hannan, M. F. (2014). Readability and comprehensibility of over-the-counter medication labels. *Renal Failure*, *36*(3), 473–477. https://doi.org/10.3109/0886022X.2013.872571

van Beusekom, M. M., Kerkhoven, A. H., Bos, M. J. W., Guchelaar, H. J., & van den Broek, J. M. (2018). The extent and effects of patient involvement in pictogram design for written drug information: a short systematic review. *Drug Discovery Today*, *23*(6), 1312–1318. https://doi.org/10.1016/j.drudis.2018.05.013

Veale, R., Hafed, Z. M., & Yoshida, M. (2017). How is visual salience computed in the brain? Insights from behaviour, neurobiology and modeling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1714). https://doi.org/10.1098/rstb.2016.0113

Veehof, L. J., Jong, B. M., & Haaijer-Ruskamp, F. (2000). Polypharmacy in the elderly -a literature review. *European Journal of General Practice*, *6*(3), 98–106. https://doi.org/10.3109/13814780009069956

Veiel, L. L., Storandt, M., & Abrams, R. A. (2006). Visual search for change in older adults. *Psychology and Aging*, *21*(4), 754–762. https://doi.org/10.1037/0882-7974.21.4.754

Wall, E. C. (2002). *A Comprehensive Look at the Fair Packaging and Labeling Act of 1966 and the FDA Regulation of Deceptive Labeling and Packaging Practices : 1906 to Today. Harvard Law School Third Year Paper*.

Wawruch, M., Kuzelova, M., Foltanova, T., Ondriasova, E., Luha, J., Dukat, A., … Shah, R. (2013). Characteristics of elderly patients who consider over-the-counter medications as safe. *International Journal of Clinical Pharmacy*, *35*, 121–128. https://doi.org/10.1007/s11096-012-9718-z

Wazaify, M., Shields, E., Hughes, C. M., & McElnay, J. C. (2005). Societal perspectives on over-the-counter (OTC) medicines. *Family Practice*, *22*(2), 170–176. https://doi.org/10.1093/fampra/cmh723

Webb, J., Davis, T. C., Bernadella, P., Clayman, M. L., Parker, R. M., Adler, D., & Wolf, M. S. (2008). Patient-centered approach for improving prescription drug warning labels. *Patient Education and Counseling*, *72*(3), 443–449. https://doi.org/10.1016/j.pec.2008.05.019

Weiss, S. (2009). Compliance packaging for over-the-counter drug products. *Journal of Public Health*, *17*(2), 155–164. https://doi.org/10.1007/s10389-008-0233-6

Wilcox, C. M., Cryer, B., & Triadafilopoulos, G. (2005). Patterns of Use and Public Perception of Over-the- Counter Pain Relievers: Focus on Nonsteroidal Antiinflammatory Drugs. *The Journal of Rheumatology J Rheumatol The Journal of Rheumatology The Journal of Rheumatology*, *3232*(32). Retrieved from http://www.jrheum.org/content/32/11/2218%5Cnhttp://www.jrheum.org/alerts%5Cnhttp://jrheum.com/faq%5Cnhttp://jrheum.com/reprints_permissions%5Cnwww.jrheum.org

Wilson, E. A. H., & Wolf, M. S. (2009). Working memory and the design of health materials: A cognitive factors perspective. *Patient Education and Counseling*, *74*(3), 318–322. https://doi.org/10.1016/j.pec.2008.11.005

Wogalter, M. S., Brelsford, J. W., Desaulniers, D. R., & Laughery, K. R. (1991). Consumer product warnings: The role of hazard perception. *Journal of Safety Research*, *22*(2), 71–82. https://doi.org/10.1016/0022-4375(91)90015-N

Wogalter, M. S., & Vigilante, W. J. (2003). Effects of label format on knowledge acquisition and perceived readability by younger and older adults. *Ergonomics*, *46*(4), 327–344. https://doi.org/10.1080/0014013021000048006

Wold, R. S., Lopez, S. T., Yau, C. L., Butler, L. M., Pareo-Tubbeh, S. L., Waters, D. L., … Baumgartner, R. N. (2005). Increasing trends in elderly persons' use of nonvitamin, nonmineral dietary supplements and concurrent use of medications. *Journal of the American Dietetic Association*, *105*(1), 54–63. https://doi.org/10.1016/j.jada.2004.11.002

Wolf, M. (2017). Improving Prescription Drug Labeling. In *North Carolina Medical Journal* (Vol. 5079, pp. 1–15). https://doi.org/10.1037//0003-066X.46.5.506

Wolf, M. S., Curtis, L. M., Wilson, E. A. H., Revelle, W., Waite, K. R., Smith, S. G., … Baker, D. W. (2012). Literacy, cognitive function, and health: Results of the LitCog study. *Journal of General Internal Medicine*, *27*(10), 1300–1307. https://doi.org/10.1007/s11606-012-2079-4

Wolf, M. S., Davis, T. C., Bass, P. F., Curtis, L. M., Lindquist, L. A., Webb, J. A., … Parker, R. M. (2011). Improving pregnancy drug warnings to promote patient comprehension. *American Journal of Obstetrics and Gynecology*, *204*(4), 318.e1-318.e5. https://doi.org/10.1016/j.ajog.2010.12.040

Wolf, M. S., Davis, T. C., Curtis, L. M., Bailey, S. C., Knox, J. A. P., Bergeron, A., … Wood, A. J. J. (2016). A Patient-Centered Prescription Drug Label to Promote Appropriate Medication Use and Adherence. *Journal of General Internal Medicine*, *31*(12), 1482–1489. https://doi.org/10.1007/s11606-016-3816-x

Wolf, M. S., Davis, T. C., Shrank, W., Rapp, D. N., Bass, P. F., Connor, U. M., … Parker, R. M. (2007).

To err is human: Patient misinterpretations of prescription drug label instructions. *Patient Education and Counseling*, *67*(3 SPEC. ISS.), 293–300. https://doi.org/10.1016/j.pec.2007.03.024

Wolf, M. S., Gazmararian, J. A., & Baker, D. W. (2005). Health literacy and functional health status among older adults. *Arch Intern Med*, *165*. https://doi.org/10.1001/archinte.165.17.1946

Wolf, M. S., Gazmararian, J. A., & Baker, D. W. (2007). Health Literacy and Health Risk Behaviors Among Older Adults. *American Journal of Preventive Medicine*, *32*(1), 19–24. https://doi.org/10.1016/j.amepre.2006.08.024

Wolf, M. S., King, J., Jacobson, K., Di Francesco, L., Bailey, S. C., Mullen, R., … Parker, R. M. (2012). Risk of unintentional overdose with non-prescription acetaminophen products. *Journal of General Internal Medicine*, *27*(12), 1587–1593. https://doi.org/10.1007/s11606-012-2096-3

Yin, H. S., Johnson, M., Mendelsohn, A. L., Abrams, M. A., Sanders, L. M., & Dreyer, B. P. (2009). The health literacy of parents in the United States: a nationally representative study. *Pediatrics*, *124 Suppl*(November), S289–S298. https://doi.org/10.1542/peds.2009-1162E

Yin, H. S., Parker, R. M., Wolf, M. S., Mendelsohn, A. L., Sanders, L. M., Vivar, K. L., … Dreyer, B. P. (2012). Health literacy assessment of labeling of pediatric nonprescription medications: Examination of characteristics that may impair parent understanding. *Academic Pediatrics*, *12*(4), 288–296. https://doi.org/10.1016/j.acap.2012.02.010