Supporting the PROTECT Initiative

Josh Lefton  
*Purdue University*

Jackson Murray  
*Purdue University*

Ahmed Thabet  
*Purdue University*

See next page for additional authors

Follow this and additional works at: https://docs.lib.purdue.edu/pjsl

Part of the Data Science Commons, Medicinal Chemistry and Pharmaceutics Commons, Pharmacology Commons, and the Toxicology Commons

**Recommended Citation**

Lefton, Josh; Murray, Jackson; Thabet, Ahmed; Baireddy, Sriram; Shukla, Prakash; Gupta, Mridul; Becker, Reagan; Ertle, Julie; Doan, Tony; and Yang, Aerin (2022) "Supporting the PROTECT Initiative," *Purdue Journal of Service-Learning and International Engagement*: Vol. 9 : Iss. 1, Article 10.  
DOI: https://doi.org/10.5703/1288284317396  
Available at: https://docs.lib.purdue.edu/pjsl/vol9/iss1/10

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

This is an Open Access journal. This means that it uses a funding model that does not charge readers or their institutions for access. Readers may freely read, download, copy, distribute, print, search, or link to the full texts of articles. This journal is covered under the CC BY-NC-ND license.
Supporting the PROTECT Initiative

Cover Page Footnote
We would like to thank our corporate partner and member of the PROTECT taskforce, Dr. Deborah Spoerner, for introducing this project to the Data Mine and for supporting our work throughout the process. She is performing amazing work alongside the CDC's PROTECT initiative to make medication safer for children. We would also like to thank the entire Data Mine staff for enabling us to work on this project. We would specifically like to acknowledge Dr. Mark Daniel Ward for his passionate leadership of the Data Mine, Kevin Amstutz for lending his time and data science expertise to our work, and Maggie Betz and Heather Goodwin for facilitating Data Mine projects such as ours. Finally, we would like to acknowledge Loan Cao, Karthik Menon, Jack Munson, Margaret Wang, Thomas Zimbelmann, and Patricia Casaca for their contributions through the first semester of this project.

Authors
Josh Lefton, Jackson Murray, Ahmed Thabet, Sriram Baireddy, Prakash Shukla, Mridul Gupta, Reagan Becker, Julie Ertle, Tony Doan, and Aerin Yang

This research with reflection is available in Purdue Journal of Service-Learning and International Engagement: https://docs.lib.purdue.edu/pjsl/vol9/iss1/10
STUDENT AUTHOR BIO SKETCHES

Josh Lefton is an undergraduate senior studying Data Science and Computer Science in the College of Science. He is interested in quantitative modeling and natural language processing.

Jackson Murray is a junior studying UX Design, minoring in Psychology, and is pursuing Collaborative Leadership and Entrepreneurial certificates at Purdue University. In the future, he hopes to discover more learning opportunities and grow within his work positions.

Ahmed Thabet is a graduate student in the ABE department (BIRS program) and a quality assurance intern at Cook Research Incorporated. His research areas of interest include regulatory science and quality assurance, quality system management, and data analysis. His career passion is to be a quality assurance specialist in pharmaceutical/medical device companies.

Sriram Baireddy is a PhD student in Electrical Engineering at Purdue University. He earned his BS and MS degrees in Electrical Engineering at Purdue in 2018 and 2021, respectively, with minors in economics, math, and physics. He currently investigates the application of machine learning techniques to signals, images, and videos for forensic and agricultural research.

Prakash Shukla is a graduate student at Purdue Polytechnic Institute. His research interest areas include human computer interaction and cognition.

Mridul Gupta received his BTech degree in Civil Engineering from the Indian Institute of Technology Roorkee in 2017. He is currently a PhD student in Electrical Engineering at Purdue University. His research focuses on applying deep learning and machine learning tools to problems in computer vision and image processing.

Reagan Becker is a first-year undergraduate student majoring in Data Science and Computer Science. She is interested in studying machine intelligence.
SUPPORTING THE PROTECT INITIATIVE

Julie Ertle is an undergraduate freshman majoring in Data Science with a focus in Computer Science and minoring in Spanish. Her interests include statistics.

Tony Doan is an undergraduate senior in Business Analytics.

Aerin Yang is an undergraduate senior in Applied Math and Data Science. Her research interests include predictive analytics and numerical methods.

ABSTRACT

Recently, medication dosage errors have received more political and media attention. Dosage errors are the most common medical errors, affecting about 1.5 million people annually.

Furthermore, U.S. poison-control centers reported more than 200,000 cases per year of medication errors. These cases result in medical costs of around $3.5 billion, and children under 6 years old constitute approximately 30% of these cases.

The PROTECT Initiative (Preventing Overdoses and Treatment Errors in Children Taskforce) was launched in 2008 as a collaborative effort between public health agencies and patient advocates to minimize dosage errors.

In alignment with the PROTECT Initiative effort, this project aims to highlight medication dose error causes and recommendations for improvement by downloading, parsing, and analyzing OTC (over-the-counter) medication labels from the DailyMed historical medication library. We found that over 70% of labels surveyed contained measurement labels in “teaspoons”; in addition, only 10% of manufacturers are using “syringe” as an optimum dosing delivery device. Therefore, health care providers must ensure that the appropriate drug tools, information, and dosages are prescribed to children (especially neonates) because of their differences in response to drugs compared with adults. With this initiative, we can work to minimize the impact of the dosage error problem (for over-the-counter pediatric orally administered liquid medications) on users.

INTRODUCTION

Medication dosing error is responsible for up to 17.8% of hospitalizations in children (Hoyle Jr. et al., 2012). It often leads to morbidity and mortality in hospitalized patients. The vulnerability of children to dosing error increases because the amount of dose is calculated based on the bodyweight of children. Caregivers often refer to health professionals in the absence of specific dosing directions on the product labels. It has been found that more than 80% percent of parents have made at least one dosing error when administering medicine to their young children (Neville et al., 2015).

Dosing error is defined as incorrect dosage quantity or frequency of administration. Some researchers have defined it to be more than 20% deviation from the weight-appropriate dose for the child (Hoyle Jr. et al., 2012). In the past, researchers have reported various reasons for these errors. They can occur at multiple levels. One potential cause is error in knowledge of prescribed dose. Another is error in observed dose measurement. Volumetric dosing errors and the use of incorrect dosing delivery devices are the most common preventable errors for orally administered liquid medication.

To avoid these scenarios, researchers at the U.S. Centers for Disease Control and Prevention (CDC) have given many recommendations and guidelines (Neville et al., 2015). Some of these relate to the instructions that need to be displayed on the medicine label. These include removing abbreviation inconsistencies in the dosing units and using milliliter-based dosing, for example, using mL (small m and capitalized L) instead of ml, ML, and cc for milliliter (Yin et al., 2014). Furthermore, the CDC advises manufacturers to be more careful with decimal points and avoid dosing amounts in hundredths of a milliliter. Using a preceding zero before the decimal point and avoiding a trailing zero after the decimal point prevents tenfold error, in which a medication is unintentionally administered at 10 times the recommended amount or frequency. Parents are also recommended to administer doses using metric marking devices such as syringes instead of tablespoons.

This project is the result of a partnership between the Data Mine’s Corporate Partners Program at Purdue University, a data science learning community, and the CDC PROTECT Initiative. The authors are members of a Data Mine team that was active throughout the 2021–2022 academic year.

We began this project with the goal of analyzing medicine labels to determine the frequency of features known to be
SUPPORTING THE PROTECT INITIATIVE

The accidental overdosing problem being researched is primarily seen with these medications. First, to determine if the medication is a nontopical liquid, the program parses each label for the keywords “liquid” and “topical.” If a label contains “liquid” but not “topical,” it is searched further, otherwise the program adds it to the total files tally and moves on to the next label. Then, the program searches the purpose section of the label for words containing “otis”/“otic” and “ophthalmic” (meaning ear and eye) and the phrases “external use only” and “mouthwash.” If the label does not contain any of these, it is considered a relevant liquid file and is added to that tally. The parser examines the dosage and usage sections of each liquid file and identifies instances of dosing tools and measurement units using keywords such as “ml,” “tsp,” “cup,” “syringe,” “pipette,” “dropper,” and “spoon.”

RESULTS

We extracted 42,726 OTC and 9,590 prescription liquid XML files from the DailyMed website to test our method described in the previous section. Figure 2 shows the percentage of medicine labels using recommended metric units such as milliliters and discouraged non-metric units like teaspoons over the analyzed XML files. We can observe that over 5% of the labels use teaspoons as a measurement unit, which often causes dosing errors.

We took a subset of our testing set to determine the usage of measurement tools. In Figure 3, we show the variation in suggested measurement tools for these 135 drugs. Less than a quarter of these labels suggest measuring with a syringe, despite it being the tool recommended by the CDC.

METHODOLOGY

DailyMed is a searchable database maintained by the National Library of Medicine (NLM) that contains labeling for both prescription (RX) and nonprescription or over-the-counter (OTC) drugs for human and animal use, as well as other products such as cosmetics and dietary supplements. The database is updated daily, so it contains the most recent labeling submitted by companies to the Food and Drug Administration (FDA). The labeling for nonprescription, or over-the-counter drugs is called Drug Facts and consists mostly of information like the drug’s purpose, warnings, and usage directions. Prescription drugs have similar components on their labels in addition to some more specific information due to their controlled status. The information for each drug is stored in the DailyMed dataset and is made available to users in several formats like PDF and XML.

XML files are designed to enable easy-to-read access for both humans and machines. This is achieved using “tags” that define the structure of the document. The file’s data is stored in plain text throughout the file, and the tags are used to identify what the data describes. For example, in the DailyMed XML files, a tag could be ‘ProductName’, while the associated plain text would be the actual product name of the drug. Keeping this in mind, we developed a script to parse the thousands of XML files to compile the relevant data in a single file. We extracted product names, company names, purpose information, warnings, usage directions, and ingredients. With this information we were able to separate generic drugs from nongeneric drugs by comparing primary ingredients to product names. To identify drugs with problematic dosing instructions, we used a multistep process.

To filter the database and find the needed information, we created a parser to search the XML label files for dosing units and tools (Figure 1). The parser outputs data about the number of labels using each type of measurement unit and mentioning each type of dosing tools in a table. The parser separates the orally administered liquid medication labels from the rest of the dataset using keywords, because the accidental overdosing problem being researched is primarily seen with these medications. First, to determine if the medication is a nontopical liquid, the program parses each label for the keywords “liquid” and “topical.” If a label contains “liquid” but not “topical,” it is searched further, otherwise the program adds it to the total files tally and moves on to the next label. Then, the program searches the purpose section of the label for words containing “otis”/“otic” and “ophthalmic” (meaning ear and eye) and the phrases “external use only” and “mouthwash.” If the label does not contain any of these, it is considered a relevant liquid file and is added to that tally. The parser examines the dosage and usage sections of each liquid file and identifies instances of dosing tools and measurement units using keywords such as “ml,” “tsp,” “cup,” “syringe,” “pipette,” “dropper,” and “spoon.”

Figure 1. Filter methodology to parse and separate the orally administered liquid medication labels from the rest of the dataset using keywords.

52 PURDUE JOURNAL OF SERVICE-LEARNING AND INTERNATIONAL ENGAGEMENT: VOLUME 9, FALL 2022
A total of 90 participants, 45 English speakers and 45 Spanish speakers, were divided into 3 groups of 30 patients each. Group 1 participants received a prescription and verbal instructions. Participants in Group 2 received a prescription, a syringe, and the correct dosage. Group 3 participants received a prescription, a syringe with a line marked at the correct dosage, and a demonstration. The results indicated that only 37% of Group 1 received the correct dose (11 out of 30), 83% of Group 2 received the correct dose (25 out of 30), and 100% of Group 3 received the correct dose (30 out of 30).

McMahon et al.’s study aligns with our project outcome. A low percentage of labels recommend a standard dosing measurement and an optimum dosing tool, but failure to do so may increase the likelihood of medication error and adverse results. Therefore, adjusting medication labels could result in a substantial impact on pediatric medication safety.

Our contributions to the CDC center around strengthening awareness of and adherence to CDC medication label guidelines. Specifically, we support the concern regarding the current situation for poorly designed over-the-counter labels using inaccurate dosing measurement units and inappropriate dosing delivery devices for pediatric orally administered liquid medications by providing an analysis of the current state of standardization on medication labels. Additionally, PROTECT members can use our work to advise pharmaceutical manufacturers to use only the universal standard measurement unit, milliliter (mL), for orally administered liquid medications. We also support PROTECT in their efforts to help the regulatory body, represented by the FDA, take more actions to standardize the labeling of OTC products by developing guidelines to promote the recommended universal standard measurement unit (milliliters, mL) and dosing tool (syringe) for pediatric orally administered liquid medications. Finally, our work will help the CDC to develop a national awareness program for medication error. This program could target all children’s hospitals and all health care providers nationwide.

In an effort to share our work, we gave a presentation to 90+ participants at the CDC’s 14th Annual PROTECT Initiative meeting (Figure 4). It is usually held in person; however it was held virtually this year due to the pandemic. The purpose of the meeting is to share progress on PROTECT projects. It consists of many five-minute presentations from PROTECT members and affiliated groups such as ours. The audience is encouraged to ask questions and share comments after each presentation. Our presentation received questions.

**COMMUNITY IMPACT**

Errors, such as inaccurate doses, volumetric dosing errors, or using inappropriate dosing delivery devices, may cause unintended toxicity or failure to cure the underlying disease. According to the U.S. Poison Control Center’s records, more than 200,000 cases per year were reported as medication errors—30% of reported cases were of children under 6 years old (American Academy of Pediatrics, n.d.).

Moreover, medication errors increase the economic burden in the United States, resulting in an additional $3.5 billion annually from extra medical costs. Therefore, it is crucial to take actions to minimize these errors.

McMahon et al. (1997) designed a study to test a randomized convenience sample stratified into three study groups.
regarding our parser, which we were able to explain in more detail during the feedback portion, and the possibility of using time-related data. Following the ideas presented by the audience, we were able to find the publication date of each label in the XML files. Comparing the proportion of nonstandard labels over time will be a useful measure of the effectiveness of PROTECT intervention and standardization efforts overall. We will leave this part of the analysis to next year’s team due to time constraints.

The CDC’s work greatly benefits pharmaceutical consumers, especially parents or guardians of children too young to administer medication themselves. With our findings, educating patients and parents on how to give accurate doses with appropriate dosing tools means that medication users will experience fewer adverse drug events (ADEs) caused by dosing errors (inaccurate dose or inappropriate dosing tool). Alongside PROTECT, we envision our work decreasing communication barriers created by using vague language in over-the-counter labels using milliliters (mL) instead of teaspoon (tsp) and using a syringe to replace unmeasured tools. Finally, our work aligns with PROTECT’s overall goal to improve communication between health agencies, private sector companies, professional organizations, consumer/patient advocates, and academic experts to increase awareness of dosing errors for children and neonates.

STUDENT IMPACT

While we are most proud of the potential impact of our study in the prevention of medication overdoses, we are pleased with the personal and professional growth we experienced during this project. Data Mine teams structure project work using the Agile methodology, which is commonly implemented by professional development teams. Many students had seen this method utilized in previous internships or work experiences. Those who had not will likely see it in the future and will be a step ahead of associates who have never worked under the paradigm before.

The Agile methodology involves maintaining a backlog of future tasks and focusing on a prioritized subset of tasks within two-week iterations known as sprints. In a project with many moving pieces, it was helpful to formally organize the tasks we wanted to complete in each sprint. It is a tool we can envision using in future long-term projects. The methodology also prescribes weekly standup meetings to keep the entire team involved and informed.

Figure 4. CDC’s 14th annual PROTECT Initiative meeting, November 2021 (CDC, 2020).
informed about individuals’ progress. The ability to concisely summarize the work one has completed and connect it to the work of others is a valuable skill. Our standup meetings allowed each of us to practice this skill every week. Students have also observed that communicating progress to others in an organization is a key factor in success. Each week students had the opportunity to practice quickly summarizing their work to teammates, as well as asking and answering questions related to the current tasks. These skills will serve students well in industry and in their current classes, especially project-based courses that require collaboration.

Our team consisted of members across the entire spectrum of data science experience. At the start of the project, some members had never worked on a data science project, while others had spent time in the industry and worked on graduate-level research. One of the greatest aspects of the team was the willingness to teach and the desire to learn from each member. Any time one person hit a roadblock, there were other team members and Data Mine staff ready to help. We all learned new technical skills and best practices throughout the project. Those new to data science were able to see how the process works and learn about the tools used, and the experienced students improved their ability to communicate their findings to others and honed their technical skills. Having a team with a variety of levels and areas of expertise is somewhat uncommon in typical courses but is certainly common in an organization with a wide variety of roles. We all benefited from the diversity in experience by improving our communication and collaboration skills. Going forward, we are all more comfortable explaining topics to and learning skills from others.

CONCLUSION

Amid the COVID-19 pandemic, the CDC has been working endlessly to increase safety in the United States. Adapting to the needs of the United States, previous focuses and projects are moved back in priority to concentrate on the relief of the main issue: COVID-19. Our team’s contribution is still significant to the CDC because we provide direct conclusions from large datasets, advocating for standard dosage measurements and tools for OTC drugs to reduce accidental overdoses among children. Our team’s analysis supports the rationale for CDC recommendations and highlights the necessity of enforcing standardization practices.

Representing the Purdue Data Mine and partnering alongside the CDC throughout this adventure has been an exceptional experience. We are the first team from the Data Mine Corporate Partners Program to collaborate with the CDC, and significant work and findings are still to come. Our team has benefited directly from this experience by developing our professional technical skills, presenting at CDC conferences, and having unique experiences as a team. Purdue’s partnership with the PROTECT Initiative is just beginning.

REFERENCES


ACKNOWLEDGMENTS

We would like to thank our corporate partner and member of the PROTECT taskforce, Dr. Deborah Spoerner, for introducing this project to the Data Mine and for supporting our work throughout the process. She is performing amazing work alongside the CDC’s PROTECT initiative to make medication safer for children. We would also like to thank the entire Data Mine staff for enabling us to work on this project. We would specifically like to acknowledge Dr. Mark Daniel Ward for his passionate leadership of the Data Mine, Kevin Amstutz for lending his time and data science expertise to our work, and Maggie Betz and Heather Goodwin for facilitating Data Mine projects such as ours. Finally, we would like to acknowledge Loan Cao, Karthik Menon, Jack Munson, Margaret Wang, Thomas Zimbelmann, and Patricia Casaca for their contributions through the first semester of this project.