Adverse Drug Effect Detection for Clinical Decision Support

By

Joshua Carl Smith

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biomedical Informatics

May, 2016

Nashville, Tennessee

Approved:

Randolph A. Miller, M.D.

Qingxia Chen, Ph.D.

Joshua C. Denny, M.D.

Kevin B. Johnson, M.D.

Dan M. Roden, M.D.C.M

To Jody, my loving and supportive wife.

ACKNOWLEDGEMENTS

I could not have completed my graduate work without the help of others. First and foremost, I would like to thank my advisor, Dr. Randy Miller. Over the past six years, he has guided my research and shared his belief in the promise of Biomedical Informatics. He has always encouraged me and pushed me to do better work, and for that I am grateful. I would also like to thank my PhD committee – Cindy Chen, Josh Denny, Kevin Johnson, and Dan Roden – for their invaluable assistance towards my PhD and my future career.

This work would not have been possible without financial support provided by the National Library of Medicine (T15LM007450, R01LM007995, R01LM010828) and the Agency for Healthcare Research and Quality (R36HS023485). This dissertation research also benefited from resources provided by the Vanderbilt Institute for Clinical and Translational Research (VICTR), supported by CTSA award UL1TR000445 from the National Center for Advancing Translational Sciences.

I am very grateful to the Faculty of the Department of Biomedical Informatics (DBMI). There are too many individuals who helped me over the years to name them all, but I especially want to thank Cindy Gadd for her guidance throughout my time at Vanderbilt. I want to thank all of the DBMI staff, particularly Joyce Green and Rischelle Jenkins, without whom I would have been lost. I also want to thank my fellow students, past and present, for their friendship and all of our insightful discussions over the years.

For their help in evaluating the DEB2 knowledgebase, I would like to thank Robert Cronin, Dara Eckerle Mize, Jacob Eichenberger, Travis Osterman, Matthew Rioth, and Michael Temple. I want to acknowledge contribution of all the Vanderbilt pharmacists who helped review the ADER alerting subset – Mark Sullivan, Terry Bosen, Fred Hargrove, David DiPersio, and Amy Myers. I am also grateful to Peter Speltz for his help in accessing the Portfolio data necessary for the ADER pilot study.

I am grateful to many people in the Vanderbilt University Medical Center Informatics Center for their help in implementing the ADER system. Dario Giuse, Yun Wang, and Jun Kunavut were instrumental in integrating ADER into StarPanel and the post-implementation evaluation. Leslie Mackowiak, Wing Liu, Joe Warner, and Zhao Zuo were essential in helping to integrate ADER alerts into the CMR form through the Vanderbilt

CPOE system. I would also like to thank Dr. John McPherson, the director of the Internal Medicine Residency Program, and Dr. Thomas Atwater, Chief Resident in Internal Medicine, who championed the ADER system and was instrumental in the successful completion of the ADER pilot study.

Last, but certainly not least, I want to thank my family. My parents, Richard and Nancy Smith, have supported me in all things. Without their love and support, I would not be the person I am today. And finally, I am beyond grateful for the love and support of my wife, Jody. Throughout my graduate work, she has unfailingly encouraged and inspired me. Simply put, I couldn't have done it without her and I wouldn't have wanted to.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADE ..... Adverse Drug Effect
DEB, DEB1, DEB2 ..... Drug Evidence Base
ADER ..... Adverse Drug Effect Recognizer
EMR ..... Electronic Medical Record
WHO ..... World Health Organization
ADR ..... Adverse Drug Reaction
FDA ..... US Food and Drug Administration
NDA ..... New Drug Application
SPL ..... Structured Product Label
XML ..... Extensible Markup Language
SRS ..... Spontaneous Reporting System
FAERS ..... FDA Adverse Event Reporting System
NSAID ..... Non-Steroidal Anti-Inflammatory Drug
VIGOR ..... Vioxx Gastrointestinal Outcomes Research Study
CDC ..... US Centers for Disease Control and Prevention
JAMA ..... Journal of the American Medical Association
FDB ..... First Databank
NLM ..... US National Library of Medicine
NDF-RT ..... National Drug File-Reference Terminology
MeSH ..... Medical Subject Headings
RxCUI ..... RxNorm Concept Unique Identifier
UMLS ..... Unified Medical Language System
NLP ..... Natural Language Processing
MRCOC ..... MEDLINE Co-Occurrence of Concepts
SIDER ..... Side Effects Resource
MedDRA ..... Medical Dictionary for Regulatory Activities
HELP ..... Health Evaluation through Logical Processing
CPOE ..... Computerized Provider Order Entry
ICD ..... International Classification of Disease

CM ..... Clinical Manifestation

H&P ..... History and Physical Exam

SNOMED-CT ..... Systematized Nomenclature of Medicine – Clinical Terms

KMCI ..... KnowledgeMap Concept Identifier

SD ..... Synthetic Derivative

USAN ..... United States Adopted Name

CUI ..... Concept Unique Identifier

SCR ..... Supplementary Concept Records

UNII ..... Unique Ingredient Identifier

ATC ..... Anatomical Therapeutic Chemical

CAS ..... Chemical Abstract Service

KB ..... Knowledgebase

AAS ..... ADER Alerting Subset

RR ..... Relative Risk

CMR ..... Continuing Medications and Results

SCP ..... Secure Copy

CHAPTER I

INTRODUCTION

Adverse drug effects (ADEs) comprise a serious healthcare problem. They cause substantial morbidity and mortality, generate preventable emergency department visits and hospital admissions, prolong hospital stays, increase healthcare costs, and negatively affect patients' quality of life. Recognizing ADEs can be difficult due to under-reporting by providers and patients, as well as the subjectivity in assessing causality between an event and specific drug therapy. In recent years, more Americans have begun taking medication on a regular basis. As more drugs are approved and more ADEs are discovered, it becomes even more difficult for healthcare providers to be aware of all potential side effects.

Addressing previously unrecognized ADEs has the potential to reduce costs and improve patient care. This dissertation partially addresses this expansive problem through the development and evaluation of two biomedical informatics applications: the Drug Evidence Base (DEB2), a knowledgebase of known medication indications and ADEs automatically extracted from reliable public sources, and the Adverse Drug Effect Recognizer (ADER), an automated system to detect newly admitted inpatient's symptomatic ADEs from clinical notes.

Chapter II defines the nature and scope of ADEs in the US, as well as methodologies used to identify new ADEs. The chapter also describes previous research on drug knowledgebases and ADE detection systems. It sets the stage for this dissertation project, and the tools, methods, and materials it uses. Chapter III details the design, construction, and validation of the DEB2 drug knowledgebase. It explains how each drug knowledge source contributes data to DEB2, individually and in combination. The chapter reports results of a preliminary appraisal of the knowledgebase. Chapter IV discusses the algorithmic extraction and manual curation of a subset (for 59 common antihypertensive medications) of the DEB2 ADE data. This subset enabled a pilot ADER system evaluation. The derivation process involved clinical pharmacists manually reviewing the ADE subset. Chapter V details the design and software development of the ADER system. It details how the ADER system works and how it interfaced with the Vanderbilt electronic medical

record (EMR) system. Chapter VI discusses the results of the ADER pilot implementation study performed at Vanderbilt University Hospital. Finally, Chapter VII summarizes this doctoral research project, discusses related potential future work, and presents the author's conclusions.

CHAPTER II

BACKGROUND

Introduction

This dissertation addresses a critical clinical problem – adverse drug effects – through development of a drug information knowledge base and an automated system for recognizing and alerting about adverse drug effects. To address this extensive topic, one must first understand the scope of the problem, as well as the previous research.

Adverse Drug Effects

A therapeutic medication effect is the desirable or beneficial intended result of drug administration for treatment.[1] Adverse drug effects (ADEs) refer to harmful, unintended consequences of medication administration. The World Health Organization[2] (WHO) specifically defined an ADE as "a response to a drug that is noxious and unintended and occurs at doses normally used in man for the prophylaxis, diagnosis or therapy of disease, or for modification of physiological function." More recently, Edwards and Aronson[1] defined an ADE as "an appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product, which predicts hazard from future administration and warrants prevention or specific treatment, or alteration of the dosage regimen, or withdrawal of the product."

The terms adverse drug *effect* and adverse drug *reaction* (ADR) refer to the same concept; whereas an ADE is from the point of view of the drug, an ADR is from the point of view of the patient. And while an ADE is causally attributable to a drug, the more general "adverse event" is an adverse outcome or injury that is not necessarily attributable to a drug.[1] A toxic, or overdose, effect is an ADE resulting from an amplification of the normal therapeutic effect, usually occurring a greater than normal doses. An allergy is an ADE mediated by an immune response. The term "side effect," although commonly used, is ambiguous and should generally be avoided.[3]

To ensure both efficacy and safety of medications marketed in the United States, the US Food and Drug Administration (FDA) requires a drug to pass multi-phase clinical

trials.[4,5] Before drugs can proceed to human trials, they are first tested in laboratory animals to determine potential toxicity. If the drug has an acceptable safety profile, human trials take place. Phase 1 clinical trials determine the pharmacologic action of the drug, as well as how the drug is metabolized and excreted in humans. They generally include between 20 and 80 healthy volunteers.[4,5] The Phase 1 studies determine a safe dosage range for the drug and identify any acute ADEs associated with various dosages. In Phase 2 trials, researchers test the candidate drug on patients who have the disease or condition that the drug is to treat and determine an appropriate dose. Phase 2 studies involve larger groups of volunteers, usually 100 to 300 individuals, and vary in duration from several months to two years.[6] In addition to further evaluating drug safety profiles and determining common ADEs, Phase 2 trials evaluate the efficacy of the drug for treating one or more targeted conditions. If the trial determines that risks seem acceptable with respect to the severity of the disease and the observed efficacy, Phase 3 trials take place. Phase 3 trials involve between 1000 and 3000 patient volunteers with the target condition.[4,5] Phase 2 and Phase 3 trials usually include a control group receiving either placebo or the current standard treatment for the condition. Phase 3 trials ideally determine a drug's risks versus its benefits, reveal less common ADEs, and look for ADEs that occur only after longer term use; the duration of Phase 3 trials is typically between one and four years.[7]

After successful Phase 1-3 studies, the research sponsor, usually a pharmaceutical company, can submit a New Drug Application (NDA) to the FDA.[8] Based on the results of the trials included in the NDA, the FDA will either approve or deny the application. For each approved medication or biological product, the FDA requires that a drug product label written by the manufacturer be reviewed by the FDA.[9] Information on the label comes from data acquired during the multi-phase clinical trials. Labeling requirements include the known information relevant to the safe and effective use of the medication. These include prescribing information like dosage, frequency, and duration, all FDA-approved indications, contraindications for use, and known or suspected ADEs.

The FDA Structured Product Labeling standard governs both label content and formatting.[10] Structured Product Labels (or SPLs) use extensible markup language (XML) to encode drug information and delineate the sections of the label. While some of these

sections contain structured data, much of the information, including indications and ADEs, is presented as unstructured free-text.

The SPLs are generally considered the gold standard primary source for information on a new drug. Nevertheless, not all ADEs listed in an SPL have definitive proof. According to official FDA documentation[11] governing the content of the SPL:

> "The ADVERSE REACTIONS section should present those adverse reactions that occurred below the specified rate for inclusion in the common adverse reactions listing, but for which there is some basis to believe there is a causal relationship between the drug and the event… Typical reasons to suspect causality for an event include (1) timing of onset or termination with respect to drug use, (2) plausibility in light of the drug's known pharmacology, (3) occurrence at a frequency above that expected in the treated population, and (4) occurrence of an event typical of drug-induced adverse reactions."

This guidance may reduce the inclusion of some number of spurious, false-positive ADEs reports, but many drug labels[12,13] list numerous ADEs with the caveat "adverse reactions are shown without attribution of causality," or some similar variation. While "on-target" ADEs (those due to exaggeration of the desired pharmacologic action) might likely be due to the drug, these permissive criteria result in the presence of numerous "off-target" non-causally linked ADEs in the majority of SPLs.

Pharmacovigilance & Post-marketing Surveillance

For a variety of reasons, many ADEs are not recognized during pre-market clinical trials.[14–16] Clinical trials usually include only a small number of relatively healthy patients (usually fewer than 3000) and trials often have restrictive inclusion criteria in terms of age, race, gender, and health status. Rare ADEs, occurring in fewer than 1 in 10,000 people, are only detectable once the drug reaches the market – when thousands to millions of people use the drug regularly over long time intervals. Additionally, small trials cannot represent the exact conditions under which the population will use new medications. Individuals will take the new drug along with other medications with which it was never tested. These individuals will also have other diseases and comorbid conditions that were not studied during the trials. Finally, these individuals will take the medications over a longer period of time than could be studied during the relatively short clinical trials. Sometimes, the FDA requires Phase 4 clinical trials[4] after a drug has been approved. Phase 4 trials, often called

Post Marketing Surveillance Trials, have several objectives, including comparing a drug with other drugs already on the market, monitoring a drug's long-term effectiveness and impact on a patient's quality of life, and studying the cost-effectiveness of a drug therapy relative to other therapies.[7] Nevertheless, most medications are only monitored after market approval through pharmacovigilance efforts.

The WHO defines pharmacovigilance as the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug-related problem.[17] It includes pre-marketing risk assessment, ongoing risk minimization, and post-marketing surveillance.[18] Coordinated pharmacovigilance efforts began in the 1960s following widespread and severe birth defects caused by the drug thalidomide. The FDA, WHO, and others began to track medications using *spontaneous reporting*, the process of healthcare professionals systematically reporting suspected ADEs to a national agency or drug manufacturer. Today, most countries collect post-marketing surveillance information in databases known as Spontaneous Reporting Systems (SRS). These include the FDA Adverse Event Reporting System (FAERS)[19] in the United States and the European Medicines Agency's SRS known as EudraVigilance[20,21]. Internationally, the WHO's Programme for International Drug Monitoring pools SRS data from over 120 countries.[17]

In the United States, drug manufacturers are required by law to submit reports of any suspected ADEs to the FDA. Patients, hospitals, and providers can voluntarily submit spontaneous reports through the FDA's MedWatch program.[22] Introduced in 1993, MedWatch encouraged healthcare professionals to regard ADE reporting as "a fundamental professional and public health responsibility" and simplified the process by which professionals could submit reports to the FDA.[23] Along with mandatory reports from drug manufacturers, reports submitted through MedWatch are recorded in the FAERS database. Using statistical and computational techniques, clinical reviewers monitor FAERS to detect signals suggesting that a medication is associated with a particular ADE.[24] In 2008, the FDA began the Sentinel Initiative[25] to complement existing ADE tracking systems. The Sentinel System enables the FDA to query healthcare data from sources other than FAERS, including administrative and insurance claims databases, as well as electronic health records.

One example of a serious ADE discovered through post-marketing surveillance is the drug rofecoxib. Sold under the trade name Vioxx, rofecoxib is a COX-2 selective non-

steroidal anti-inflammatory drug (NSAID) marketed by the drug manufacturer Merck to treat arthritis, acute pain in adults, and dysmenorrhea.[26] Merck submitted a request for approval of rofecoxib to the FDA in November 1998 and it was approved on May 21, 1999.[27] In January of 1999, Merck began the Vioxx Gastrointestinal Outcomes Research study (VIGOR), a Phase 4 clinical trial designed to test whether rofecoxib was safer than naproxen for the digestive system. The VIGOR study results were published in the New England Journal of Medicine in November 2000,[28] leading to questions about the rate of heart complications in patients on rofecoxib. It was later revealed that the published study did not accurately represent the results. Rofecoxib actually put patients at greater risk of myocardial infarction than had been indicated.[29] After numerous other studies showed an increased risk of rofecoxib-related cardiovascular events,[30] Merck withdrew the drug from the market on September 30, 2004. An increased risk of myocardial infarction and stroke usually required 18 months of use to appear.[31] While no longer sold in the United States, rofecoxib was marketed for over 5 years before discovery of its serious adverse effects. It was estimated that "between 88,000 and 140,000 excess cases of serious coronary heart disease probably occurred in the USA over the market-life of rofecoxib," 44% of which were estimated to be fatal.[32]

ADE Prevalence

Adverse drug effects comprise a serious healthcare problem, causing substantial morbidity and mortality.[33–39] They generate preventable emergency department visits and hospital admissions, prolong hospital stays, and increase healthcare costs. The US Centers for Disease Control and Prevention (CDC) notes that 82% of American adults take at least one medication, and 29% take five or more.[33] A 1998 meta-analysis published in the Journal of the American Medical Association (JAMA) estimated a 6.7% incidence of serious ADRs in hospitalized patients, with a fatality rate of 0.32%.[34] Historical studies suggest ADR prevalence has increased over time. A 2004 British study of 20,000 inpatients found 6.5% of admissions were associated with ADRs, 80% of which directly led to the admission.[35] A 2007 review estimated that ADRs accounted for between 4.2-30% of hospital admissions in the US and Canada and 2.5-10.6% of admissions in Europe.[36] A 2008 systematic review of 25 studies found that on average 5.3% of hospital *admissions* were associated with ADEs.[37]

The CDC estimates that ADEs result in 700,000 emergency department visits and 120,00 US hospitalizations in the US annually.[33]

While ADEs result in a substantial number of hospitalizations, they likely affect even more outpatients. In 2013, a study of approximately 5000 individuals across multiple care settings found that 12% of individuals suffered an ADE in the preceding three-month study period.[39] That study also found a considerable burden of ADEs from many commonly used drugs. In 2014, studies found that recognized, coded ADEs accounted for 3.2% of admissions in England, 4.2% of admission in Spain, 4.8% of admission in Germany, and 5.6% of admissions in the United States.[40,41] Since many ADEs go unrecognized, the true rate of ADEs is likely higher.

In addition to patient safety concerns, ADEs generate substantial costs within the healthcare system.[3,42–45] A 1997 JAMA study reported that an ADE occurring in a hospitalized patient increased average length of stay by nearly 2 days and increased average hospitalization costs by $2262.[42] Post-admission ADEs may cost up to $5.6 million per year per hospital.[43] However, this estimate does not include costs for admissions directly caused by ADEs, ADE-related malpractice and litigation costs, or the costs of injuries suffered by patients not admitted to hospitals. In 2006, the Institute of Medicine estimated $3.5 billion is spent annually on extra medical costs of ADEs in the United States and at least 40% of the costs of non-hospital ADEs are preventable.[33] A more recent study estimated the impact and costs of management of ADEs in the United States may indirectly account for up to $30 billion annually.[44]

Drug Knowledge Sources

Despite available public data and a wide variety of existing drug knowledge resources, no perfect, comprehensive, reliable source of indications and ADEs in a computable format yet exists.[46–48] Popular commercial resources such as Micromedex, First Databank (FDB), ePocrates, and UpToDate provide detailed drug information for use by healthcare professionals.[49–52] However, commercial medication databases are expensive and often lack published validations. They can vary in scope, content, and reliability and are often not available in a locally adaptable, computable format. The frequency of updates of specific entries within commercial systems is generally unknown.[46,47,53,54]

Non-commercial drug information is also available from frequently updated public and governmental sources.[55] As already mentioned, these include the FDA's drug product labels, available through the National Library of Medicine (NLM) DailyMed[56] website, and the FAERS database. Other public sources include MEDLINE, MedlinePlus, RxNorm, and the National Drug File Reference Terminology (NDF-RT).

The NLM's MEDLINE is a bibliographic database that topically indexes millions of journal articles using MeSH (Medical Subject Headings).[57] As of 2016, it contained more than 22 million references to journal articles mostly focused on biomedicine. The MEDLINE database is searchable through the PubMed[58] website and available for download through the MEDLINE Baseline Repository.[59]

MedlinePlus is an NLM-sponsored consumer health information web site.[60]. All information is manually curated and selected from a variety of authoritative resources. Available in free text format, the resource includes drug monographs answering such questions as: Why is this medication prescribed? What are other uses of the medicine? What side effects might this medication cause?

RxNorm is a normalized naming system for both generic and branded drugs, as well as a tool linking different drug terminologies.[61] It contains information on both prescription and over-the-counter medications, represented by an RxNorm concept unique identifier (RxCUI), linked together using numerous relationships. RxNorm is produced by the NLM and included with each release of the Unified Medical Language System (UMLS) Metathesaurus[62] – a large, multi-purpose compendium of clinical ontologies and vocabularies. Along with other information, the UMLS links synonymous concepts across and within vocabularies and represents relationships between concepts.

The NDF-RT is a formal computational drug representation system.[63] It includes information such as drug ingredients, dose forms, physiologic effects, mechanisms of action, pharmacokinetics, and 25 distinct relationship types. The NDF-RT is updated monthly and included in the UMLS and RxNorm distributions.

Derived Drug Knowledge

Substantial efforts have attempted to derive drug-related knowledge from many sources, including FAERS, MEDLINE, and the FDA SPLs, among others. Academic

researchers have developed methods to extract drug knowledge using natural language processing (NLP), text mining, machine learning, and crowd-sourcing approaches. As a result, a number of drug knowledgebases exist.

Researchers including James Cimino, Carol Friedman, and others demonstrated that drug knowledge could be extracted from MEDLINE using co-occurrences of MeSH terms, text mining, and other automated methods.[64–68] In 1998, Zeng and Cimino extracted relationships from the MEDLINE co-occurrence of concepts file (MRCOC), a compilation of counts of MeSH concept that co-occur in MEDLINE articles that was formerly included with each release of the UMLS Metathesaurus.[64] They found that relationships extracted from the MRCOC had good sensitivity, particularly for drug-disease relationships. In 2008, Chen, et al., showed that text-mining of the medical literature (using MEDLINE) combined with NLP of clinical notes could recognize drug-treatment relationships.[65] In 2011, Shetty and Dalal used disproportionality analysis of MeSH-indexed articles to show that drug-ADE signals are often detectable in the literature before they are officially recognized.[66] In 2013, Xu and Wang showed that a pattern-learning algorithm applied to MEDLINE abstracts could accurately extract drug-indication relationships.[67] Also in 2013, Avillach, et al., showed that it was possible to extract known drug-ADE relationships using MeSH concepts (with appropriate subheadings) that co-occurred in at least three MeSH-indexed MEDLINE articles.[68]

Researchers have also extracted drug data from other sources, including FDA product labels, publicly available drug databases, and EMR records. The public domain Side Effects Resource (SIDER) database contains both medication indications and ADEs.[69,70] Developed by Kuhn, et al., the original SIDER database used simple text mining methods to extract indications and ADEs from the FDA SPLs and represent them using the COSTART vocabulary. SIDER has been updated four times over the last six years and currently uses the MedDRA vocabulary. MedDRA (Medical Dictionary for Regulatory Activities) is a terminology specifically designed to support pharmacovigilance, public health monitoring, data analysis, and data exchange between pharmaceutical regulatory authorities.[71] DrugBank is a manually curated database combining chemical, pharmacological and pharmaceutical data with comprehensive drug target information.[72–74] It includes indication information manually curated from the FDA, PubMed, the Kyoto Encyclopedia of

Genes and Genomes (KEGG), the Therapeutic Target Database (TTD), and others. It also includes ADE and toxicity information manually curated from sources such as FDA, ToxNet, and the American Society of Health-System Pharmacists (ASHP).[75] McCoy, et al., used data from an EMR system where prescribers were required to connect the medications to one of the patient's specific problems.[76] They created an accurate indication knowledgebase by compiling those medication-indication links that occurred frequently in the EMR.

To address specific clinical informatics goals, research projects have also compiled drug knowledgebases from multiple sources, including many of those mentioned above. In 2010, Wang, et al.,[53] compiled drug indication information for use in automated pharmacovigilance and decision support systems from the FAERS, NDF-RT, and SemMed – a database generated using NLP on MEDLINE abstracts to identify semantic relationships between concepts. For a set of 20 drugs, they extracted indication knowledge comparable to a manually curated gold standard. In 2011, Li, et al.,[47] combined information from Micromedex, NDF-RT, and FAERS to infer the reasons for prescriptions mentioned in EHR discharge summaries. It revealed promising results, but the study only focused on a limited sample of six drugs. In 2013, Wei, et al.,[77] developed MEDI, a medication indication resource linking data from RxNorm, SIDER, MedlinePlus, and Wikipedia, and representing indications using ICD-9 codes. They further refined their work to include a higher-precision subset of indications retrieved from RxNorm or at least two out of the three sources.

As detailed in the author's 2012 Master's thesis research and published in 2013, the original Drug Evidence Base (or DEB1) was a prototype knowledgebase of medication indications and known ADEs.[78,79] The DEB1 was algorithmically derived from data mining of several publicly available drug knowledge sources. The DEB algorithms extracted drug-indication and drug-ADE pairs from the UMLS MRCOC, the NDF-RT, and through NLP on the *Indications* and *Adverse Reactions* sections of FDA SPLs. Through the process of creating the DEB1 and comparing it to SIDER and MEDI, the authors discovered many problems with both the DEB prototype and many of its "parent" drug knowledgebases. For example, the process of combining diverse drug information resources uncovered idiosyncratic uses of multiple, disparate UMLS concepts within and across the individual drug resources. The disparate, often partially overlapping terms represented what technically were single

specific indications for a drug, or the ADEs of a drug. For example, one resource listed *Acidosis (C0001122)* as an ADE of a medication, while another resource listed *Lactic Acidosis (C0001125)* as an ADE – neither listed the term used by the other. While human readers understand how these terms overlap, the UMLS concepts associated with each are distinct. In addition, DEB1 development revealed that existing drug resource knowledge representation formats may be too limited to capture fine nuances among many drug-indication and drug-ADE relationships. Many such relationships are logically and/or temporally complex. For example, the drug metoprolol is indicated as prophylaxis against ventricular fibrillation, but *only in patients with previous* myocardial infarction. One should not give metoprolol to a patient in the midst of ventricular fibrillation. A primary goal of this dissertation project is to address the many problems uncovered during construction and evaluation of the original DEB1.

A 2015 review of knowledgebases detailing medication indications agreed with many of the conclusions of the earlier DEB1 evaluation.[79,80] The 2015 review determined drug knowledgebases varied in terms of scope, including their coverage of on-label and off-label indications and whether they included contraindications and drug-drug interactions along with single medication indications and ADEs. Both evaluations found that it is difficult to compare content between, and even within, knowledgebases due to non-overlapping similar concept representations used in most clinical vocabularies. The reviewers verified that most, if not all drug knowledgebases, have problems representing complex indications – specifically those medications that that should only be prescribed along with other drugs, those that should only be prescribed in specific cases meeting certain criteria, and those having specified durations of drug therapy.

ADE Detection & Reporting Systems

In addition to obtaining drug knowledge, methods have been developed to address the burden of ADEs in various patient populations. Informatics tools can help to recognize ADEs for both reporting and clinical care purposes. During the early 1990s, David C. Classen, Stanley L. Pestonik, R. Scott Evans, John P. Burke, and others developed and implemented a computerized adverse drug event monitoring system at LDS Hospital, affiliated with the University of Utah.[81–83] The ADE monitoring system was integrated with the HELP (Health Evaluation through Logical Processing) Hospital Information system. The monitoring system identified potential ADEs through both voluntary and automated reporting.

A new application program supported voluntary reporting by allowing physicians, nurses, and pharmacists to report suspected ADEs. Through an interface with HELP, the system was linked to an integrated patient database drawn from multiple sources including pharmacy, laboratory, surgery and radiology.[81] The system monitored the database using various rules specifically developed to detect ADEs. The triggers included sudden discontinuation of medications, decreases in dosages, ordering of antidotes, and certain abnormal lab values.[82] Each day, the list of suspected ADEs were sent to a pharmacist who reviewed the list and verified any true ADEs. The pharmacist then categorized the ADE as mild, moderate, or severe, as well as dose-dependent, predictable, idiosyncratic, or allergic.[81] Over the course of 18 months, there were 731 verified ADEs identified in 648 patients, 701 were characterized as moderate or severe, 92 were voluntarily reported, and 631 were detected automatically by the system. During this time, traditional methods used at LDS Hospital identified only 9 ADEs.

The researchers went on to use the system for the prevention of ADEs. Using the events captured previously, the team designed methods to prevent ADEs.[83] By using computer alerts of known drug allergies at the time the order was placed, ADEs due to allergies were significantly reduced. Additionally, by identifying other ADEs quickly, there was a significant drop in the number of severe ADEs.

During the mid 1990s and early 2000s, David W. Bates and colleagues at Harvard Medical School and Brigham and Women's Hospital investigated how to identify and prevent ADEs using information technology. In a 1995 study, Bates, et al.,[84] used

demographics, diagnostic test results, and current medications to determine that 53% of adverse events could be identified by automated means, and that 5% were preventable. By including automated review of all physician orders, 58% of adverse events were judged identifiable and 13% preventable. By adding additional clinical data, such as automated problem lists, 89% of events were judged identifiable and 23% preventable. Similar research indicated that 28% of ADEs were preventable, 42% of life-threatening ADEs were preventable, and most ADEs resulted from errors at the ordering stage.[85]

In 1998, Bates, et al.,[86] studied the effects of Computerized Provider Order Entry (CPOE) systems on the prevention of serious medical errors. Use of CPOE led to a significant decline in ADEs and potential ADEs. The researchers also found that adding a dedicated team, including a pharmacist, did not significantly improve the results. Bates and colleagues went on to show that automated ADE monitors using diagnosis codes, allergy rules, and text searching were effective and practical in both outpatient and inpatient settings, though not as thorough as manual chart review.[87–89] They also illustrated that evidence for ADEs was identifiable in discharge summaries[90] and that combining laboratory results with medication records had the potential to prevent potential harm from ADEs.[91,92]

Nevertheless, the majority of early detection systems looked for post-facto, reactive signals such as sudden discontinuation of medications, unexpected decreases in medication dosage, administration of known antidotes, and billing codes indicating that ADEs had been recognized.[93,94] Many of the systems focused on medication errors and not necessarily idiosyncratic ADEs that occur at normal dosage during normal use. While useful for documentation purposes, these signals often involved already recognized ADEs that had presumably been addressed by providers. Researchers have in the past used International Classification of Disease (ICD) codes to recognize ADEs,[95,96] but several studies have found that as many as 30% of ICD-9 codes have no supporting evidence in corresponding clinical notes and that administrative data may be of limited use in detecting ADEs.[94,96,97]

As mentioned briefly above, clinical text notes may serve as potential sources of ADE information. In 2001, Bates and colleagues found that text-searching revealed significantly more potential ADEs than laboratory signals or diagnosis codes, but that the positive predictive value of their method was very low.[88] Since the early 2000s, however,

NLP techniques have greatly improved. Researchers have used NLP to identify mentions of ADEs from discharge summaries and other sources with moderate success, but most studies have focused on a small number of serious ADEs not necessarily linked to a causal drug.[98–102] For example, Haerian, et al.,[102] used NLP to study two serious ADEs in 2012 – rhabdomyolysis and agranulocytosis. They were able to differentiate cases in which specific medications were suspected of causing an ADE versus cases where patient's underlying disease processes were likely responsible. In 2013, Eriksson, et al.,[101] developed a dictionary of possible ADE terms from parsing Danish drug labels in 2013. They used this dictionary to detect possible ADEs after the fact in a corpus of clinical notes, recognizing terms such as *anxiety, sedation, pain, anger, unrest, psychosis, paranoia,* and *depression*. However, these were just common terms from drug labels and not actually linked to causal drugs in the clinical notes; rather than indicating presence of an ADE, the findings could have been related to underlying disease processes. Using NLP to identify ADEs shows great promise, but the above-mentioned difficulties must first be addressed.

Dissertation Overview

Identifying ADEs is difficult because of under-reporting by both providers and patients, as well as the subjectivity in assessing causality between an event and specific drug therapy.[1,88] In the early 2000s, Bates and colleagues showed that clinical text could be useful in recognizing ADEs. Since then, researchers at Columbia University have illustrated that NLP applied to discharge summaries can detect many potentially unrecognized ADEs.[102–104] However, those studies also indicated the necessity of using drug knowledge or computational methods to distinguish ADEs from indications and confounders. Researchers at Stanford University have successfully utilized ADE information extracted using NLP of clinical notes, alone and in concert with other data, for pharmacovigilance studies.[105–109] The current FDA Sentinel Initiative aims to implement an electronic system that uses healthcare data, including EMRs, to monitor the safety of drugs, biologics, and medical devices by correlating reports of patient findings with their medications.[110,111]

While there are many sources of drug knowledge, no perfect, comprehensive source of indications and ADEs in a computable format yet exists. A knowledgebase of accurate drug-indication and drug-ADE relationships would benefit pharmacovigilance, drug

repurposing, clinical data mining, phenotyping, and decision support systems, among others. [47,53,67,79,105]

This dissertation research builds on previous work to address both the problems of representing and storing drug knowledge, as well as the difficulties of recognizing ADEs in patients. The project team, referenced throughout this work, consisted primarily of myself, Joshua C. Smith, and my advisor, Professor Randolph A. Miller. It also consisted of my PhD Dissertation Committee – Professors Qingxia Chen, Joshua C. Denny, Kevin B. Johnson, and Dan M. Roden.

First, this dissertation describes the development of the improved Drug Evidence Base (DEB2), based upon previously described work on DEB1.[78,79] The revised DEB2 is algorithmically extracted and compiled from existing, publicly available sources. Because the derivation is nearly fully automated, DEB2 can undergo frequent updates. Unlike DEB1, the improved version draws information from five sources and requires corroboration from at least two sources. Other common drug knowledgebase deficiencies are addressed in a variety of ways and are described in Chapter II.

Second, this dissertation describes the development of the Adverse Drug Effect Recognizer (ADER), a real-time ADE alerting system to detect adult inpatients' previously unrecognized, symptomatic pre-admission ADEs. At the time of admission, the system uses NLP to extract each patient's medications and clinical manifestations (CMs – diseases, symptoms, findings, etc.) from the admission history and physical examination (H&P) notes. The ADER system then cross-references the extracted drug and CM concepts against known medication-ADE associations derived from DEB2 and alerts appropriate providers about their patient's potential ADEs. The development of the system, as well as a pilot implementation study performed at Vanderbilt University Hospital, is described in Chapters V and VI, respectively.

Tools, Materials, and Sources used in this work

This research employed a number of the knowledge sources and tools mentioned above. The project used RxNorm to represent and link medication concepts and the UMLS to represent and connect non-medication concepts. More specifically, clinical manifestation concepts were represented using SNOMED-CT (Systematized Nomenclature of Medicine –

Clinical Terms) – a comprehensive clinical terminology included in the UMLS. The DEB2 knowledgebase used medication information from the NDF-RT, MEDLINE, MedlinePlus, SIDER2, and DrugBank, all described above.

This dissertation also utilized three Vanderbilt-developed NLP tools: SecTag, the KnowledgeMap Concept Identifier (KMCI), and MedEx. SecTag, developed by Denny, et al., identifies section headers in clinical documents.[112] Using a locally developed lexicon of clinical note section header terms and heuristics, the SecTag algorithm identifies sections in H&P notes such as "History of Present Illness," "Chief Complaint," or "Medications." SecTag identifies major headings and subheadings within sections, as well implied section headers using a naive Bayes classifier combined with terminology-based rules. Denny, et al., also developed KMCI, an NLP tool that recognizes UMLS concepts in clinical text.[113,114] Originally developed for use on medical education documents, KMCI has been expanded for clinical use. It performs term normalization and variant generation, combined with concept co-occurrence data from MEDLINE, to identify the ambiguous or under-specified concepts commonly mentioned in clinical notes.[115] Developed by Xu, et al., MedEx is an NLP tool for extracting medications and medication-related information from natural language clinical notes.[116] MedEx identifies medications' generic and trade names, as well as signature information such as strength, dose, route, frequency, duration, and necessity, if they are present.

The primary source of clinical information used to evaluate the project's ADE detection system is Internal Medicine admission H&P notes. Previous research at Vanderbilt confirmed Internal Medicine service H&P notes had a positive predictive value of 96% for medications, compared to a gold standard established through independent review of patients' medications conducted by clinical pharmacists.[117] The project also used clinical notes from the Vanderbilt Synthetic Derivative (SD),[118] a de-identified copy of Vanderbilt EMR system used for research purposes, and the Portfolio database. Portfolio is a system designed to store and evaluate clinical notes written by interns and residents at Vanderbilt University Medical Center.[119,120]

Conclusion

Adverse drug effects comprise a serious healthcare problem. They cause substantial morbidity and mortality, generate preventable emergency department visits and hospital admissions, prolong hospital stays, and increase healthcare costs. Approaches to address this problem involve ADE discovery – finding new, previously unrecognized ADEs – and ADE detection – recognizing known ADEs in patients so that offending medications can be stopped and patients can receive appropriate care. Both of these approaches require systematic drug knowledge. Despite available public data and a wide variety of existing drug knowledge resources, there does not yet exist a complete and reliable source of indications and ADEs in a machine-understandable format. This dissertation builds on previous work to address both the problem of representing and storing drug knowledge, through development of the Drug Evidence Base (DEB2), and the difficulties of recognizing ADEs in patients by developing and evaluating the Adverse Drug Effect Recognizer (ADER).

CHAPTER III

THE DRUG EVIDENCE BASE (DEB)

Introduction

As discussed in Chapter II, medication-related information resides in many disparate sources with widely varying representation formats. Drug knowledge sources vary in reliability and often disagree.[46,47,53,54] Ready availability of accurate medication indication and ADE data would benefit biomedical informatics research, including pharmacovigilance, clinical data-mining, clinical phenotyping, and decision support systems, among others.[47,53,79,105] To that end, this dissertation project developed the Drug Evidence Base (DEB), a verifiably accurate, machine-processable drug knowledge base automatically derived and updated from reliable public sources.

The DEB knowledge base consists of two concept types (drugs and clinical manifestations) and two allowed relationships (indications and adverse drug effects) between these concepts. Clinical manifestations (CMs) include diseases, syndromes, and findings. Findings include patients' historical items, symptoms, signs discovered on physical examination, and results of laboratory and imaging studies. The DEB classifies each drug-CM pair as either an *indication*, defined as "drug treats or prevents CM," or an adverse drug effect (ADE), defined as "drug causes or exacerbates CM." The version of the DEB detailed below derives from lessons learned during the development of a prototypic DEB during 2010-2012 Master's Thesis research (herein called DEB1).[78,79] This new version, called DEB2, addresses many of the shortcomings of DEB1.

Materials

This work utilized a MacBook Pro with a 2.6 GHz Intel Core i7 processor and 16 GB of RAM and a Linux server with forty-eight 2.2 GHz AMD Opteron cores and 256 GB RAM. All data processing scripts used Perl. 5.10.0. The project used MySQL version 5.5.16 for the DEB2 and to store intermediate results.

All work employed the NLM's UMLS 2013AB release, downloaded and installed into a MySQL database on 7 February 2014. That UMLS version included the 3 February2014

release of RxNorm. The 2013AB database mistakenly contained duplicate records in the MRSAT table; they were deleted. The project used data from the MEDLINE 2014 Baseline Distribution, released on 25 November 2013 and downloaded on 6 February 2014.[59]

Methods

Overview

The methods below detail the development of the DEB2. Steps included: defining medication concepts, CM concepts, and drug-CM pairs; extracting data from each of the component drug knowledge sources of the DEB2; and developing the techniques used to combine the data from each source and to evaluate the accuracy of the resulting knowledge base.

Medication Concepts

All drug concepts in DEB2 are single-ingredient medications represented using RxNorm RxCUIs (RxNorm Concept Unique Identifiers).[61] Since the project uses single ingredient medications, drug representations are RxNorm "ingredient" concepts. In RxNorm, ingredients are defined as "A compound or moiety that gives the drug its distinctive clinical properties. Ingredients generally use the United States Adopted Name (USAN)."

The construction of DEB2 began by selecting all RxNorm concepts with the semantic type "clinical drug" from the RxNorm Prescribable Subset.[121] The RxNorm Prescribable Subset, contained within the RxNorm RXNCONSO database table, is a subset of the full RxNorm database containing all medications that practitioners can prescribe in the United States.

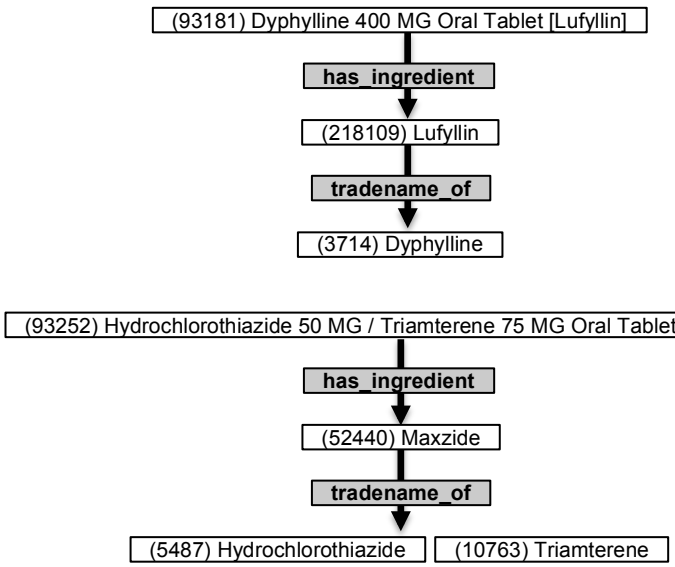
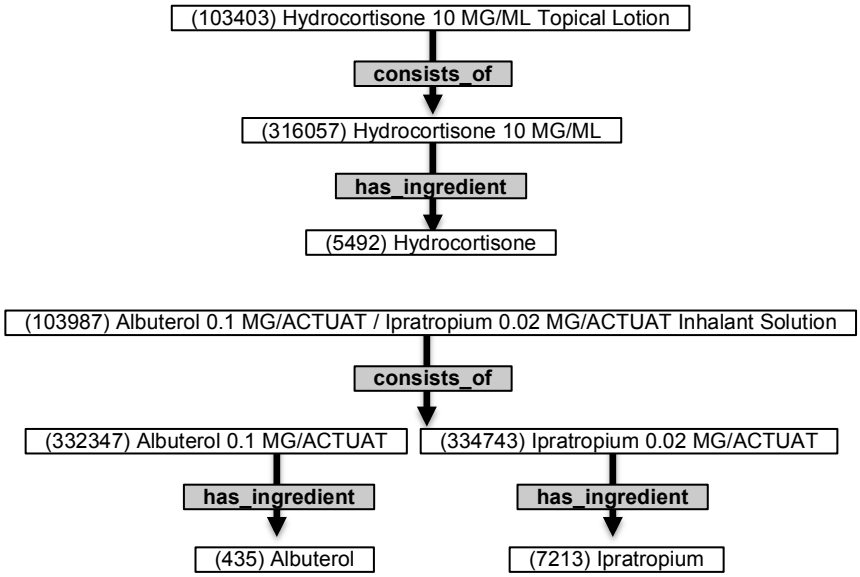
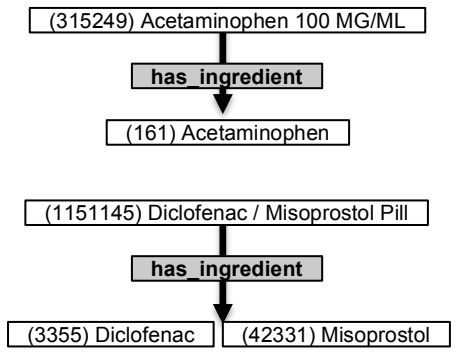
Next, the prescribable clinical drug concepts were mapped to their generic ingredients. Within RxNorm, "clinical drug" concepts are a combination of one or more ingredients, possible strengths, and dose forms. Each clinical drug also has a term type to indicate generic and branded drug names at different levels of specificity. Definitions and examples of each of the term types (TTY) of these clinical drugs are included in Table 1.

Table 1. Term Types of clinical drugs in the RxNorm prescribable subset.

TTY	Name	Description	Example
SBD	Semantic Branded Drug	Ingredient + Strength + Dose Form + Brand Name	Fluoxetine 4 MG/ML Oral Solution [Prozac]
SBDC	Semantic Branded Drug Component	Ingredient + Strength + Brand Name	Fluoxetine 4 MG/ML [Prozac]
SBDF	Semantic Branded Drug Form	Ingredient + Dose Form + Brand Name	Fluoxetine Oral Solution [Prozac]
SBDG	Semantic Branded Dose Form Group	Brand Name + Dose Form Group	Prozac Pill
SCD	Semantic Clinical Drug	Ingredient + Strength + Dose Form	Fluoxetine 4 MG/ML Oral Solution
SCDC	Semantic Clinical Drug Component	Ingredient + Strength	Fluoxetine 4 MG/ML
SCDF	Semantic Clinical Drug Form	Ingredient + Dose Form	Fluoxetine Oral Solution
SCDG	Semantic Clinical Dose Form Group	Ingredient + Dose Form Group	Fluoxetine Oral Product
SY	Synonym	Synonym of another TTY, given for clarity.	Prozac 4 MG/ML Oral Solution
TMSY	Tall Man Lettering Synonym	Tall Man Lettering synonym of another TTY, given to distinguish between commonly confused drugs.	FLUoxetine 10 MG Oral Capsule [PROzac]

Relationships included in RxNorm allow users to map clinical drug concepts to their constituent components. The project used these relationships to map each clinical drug to its generic ingredients. The specific relationships used to map each clinical drug to the appropriate generic ingredients were a function of the term type. The project did not map concepts with a synonym term type (SY, TMSY) since they are by definition included with another term type in the set. Examples illustrating which relationships were used to map each term type to the appropriate generic ingredient are shown in Table 2.

Table 2. Mapping each Term Type to ingredient drugs to identify single-ingredient medications and discard multi-ingredient medications.

TTY	Mapping from TTY to Generic Ingredients
SBD SBDC SBDF SBDG	
SCD	
SCDC SCDF SCDG	

The project removed those drugs with multiple active ingredients from the RxNorm Prescribable Subset. The project restricted the set of medications to common small-molecule drugs, removing those drug concepts that represented dental materials, cell lines (e.g., non-hormonal proteins), food extracts used to test for allergies, antivenins, and reagents, among others. The project accomplished this by removing drug concepts with unwanted semantic types and drug names indicating they were not of the type desired. Concepts with the following RxNorm semantic types were excluded:

- Biomedical or Dental Material
- Bacterium
- Cell
- Enzyme
- Eicosanoid
- Food
- Fungus
- Indicator, Reagent, or Diagnostic Aid
- Immunologic Factor
- Inorganic Chemical

The project also removed drug concepts with any of the following words in the drug name string:

- Antivenin
- Venom
- Allergenic
- Extract
- Vaccine
- Oil

The remaining set of RxNorm drug ingredient concepts (the “DEB2 Medication Subset”) was employed in the remainder of this work.

Clinical Manifestation Concepts

As previously noted, this work defined a CM as a condition that can be present in a patient, such as a disease, syndrome, symptom, or finding. More specifically, the goal was to include any CM that can be treated or prevented by a medication, as well as those that can be caused or exacerbated as medication ADEs. To this end, the project restricted the CMs in DEB2 to the SNOMED-CT vocabulary, represented by UMLS Concept Unique Identifiers (CUIs),[62] with any of the following UMLS semantic types:

- Anatomical Abnormality
- Injury or Poisoning
- Congenital Abnormality
- Finding
- Sign or Symptom
- Acquired Abnormality
- Clinical Attribute
- Disease or Syndrome
- Mental or Behavioral Dysfunction
- Neoplastic Process
- Pathologic Function

The project did not pre-specify the range of all possible CM concepts. Instead, the process involved extracting medication and CM concepts from each DEB2 source and then discarding those CM concepts that did not meet project inclusion criteria. The process by which CM concepts were constrained differs slightly for each source, as described below.

DEB2 Sources

The construction of DEB2 involved two steps. First, custom Perl scripts extracted drug-CM pairs from each of the five constituent sources. The DEB2 construction algorithms stored the drug-CM pairs from each of the five sources in five separate tables in a MySQL database. Second, the scripts combined data from each of the source tables into two combined tables: the *DEB2_full* table, containing all drug-CM pairs from each source, and

the *DEB2_final* table, containing only those drug-CM pairs that appeared in at least two sources. Details of this process follow below.

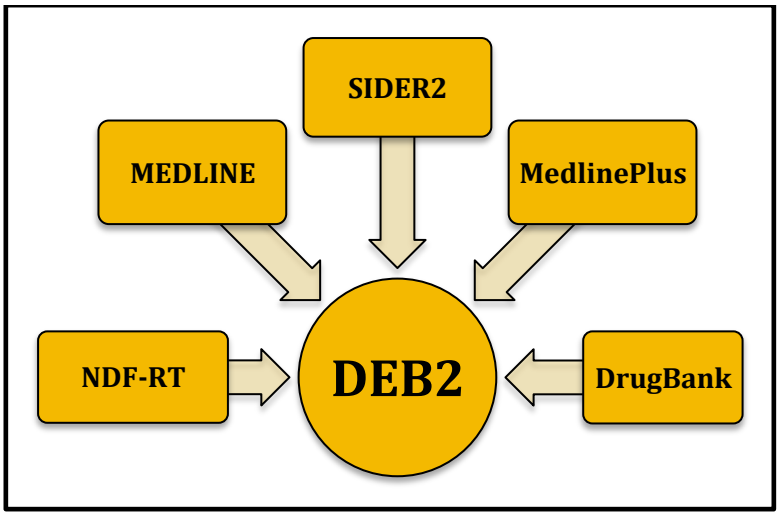


Figure 1. Illustration of DEB2 sources.

Source: NDF-RT

As discussed in Chapter II, the National Drug File-Reference Terminology (NDF-RT) is a formal computational drug representation.[63] It includes information such as drug ingredients, dose forms, physiologic effects, mechanisms of action, and 25 distinct relationships. The terminology is included in the UMLS and RxNorm distributions.

Extraction of the NDF-RT component of DEB2 involved selection of all NDF-RT entries in the UMLS containing the "may-treat," "may-prevent," or "induces" relationships. The DEB2 construction algorithms loaded all such entries into a MySQL database. In the manner outlined above, the algorithms mapped the drug component of each NDF-RT relationship to its RxNorm generic ingredient concept. The algorithms removed those NDF-RT entries in which the drug concept was not present in the DEB2 Medication Subset, as well as any entries in which the CM concepts were not present in SNOMED-CT or did not have a previously mentioned CM semantic type. Finally, based on NDF-RT relationships, the algorithms classified the remaining distinct drug-CM pairs as either indications ("may-treat" or "may-prevent") or ADEs ("induces"). If a pair was listed as both indication and ADE, it was recorded as such and addressed later. These drug-CM pairs, and the identified relationships, are stored in a MySQL database table as the NDF-RT component of DEB2.

Source: MEDLINE

Extraction of the MEDLINE component of DEB2 involved loading the MEDLINE 2014 Baseline Distribution into a MySQL database using a schema modified from Diane E. Oliver and a Perl script modified from the BioText Project at UC Berkeley.[122,123] This distribution contains 22,376,811 records from MEDLINE.

Each MEDLINE entry contains MeSH terms representing the major topics of a given published article. These MeSH terms are sometimes tagged with MeSH Subheadings that indicate the context in which the term is mentioned in the article. Some entries also contain Supplementary Concept Records (SCRs). The latter are similar to MeSH headings, but used to index some chemicals and drugs; other chemicals and drugs have actual MeSH terms and are not included in the SCR. Only MeSH terms have subheadings.

To remove animal studies from consideration, the project eliminated any MeSH-indexed articles without the "Humans" MeSH heading. Using UMLS semantic types, the DEB2 construction algorithm next identified those MeSH headings that represented CMs and those that represented specific drug concepts. When the entry had SCR data, the algorithm mapped drug concepts in the article's SCR to the article's MeSH headings using UMLS "is a" relationships. This allowed the algorithm to recognize if any of the drug concepts in the SCR had corresponding MeSH subheadings. For example, if the drug *Lisinopril* appeared in the SCR, the corresponding MeSH heading would be its drug class – *ACE-inhibitors*. The SCR drug concepts representations in the MEDLINE Baseline Distribution utilized Unique Ingredient Identifiers (UNII).[124] The algorithm mapped the UNII identifiers to RxCUIs using the RxNorm RXNCONSO table.

After identifying all potential MeSH CM and drug concepts as described above, the DEB2 construction algorithm then removed those drug concepts not in the DEB2 Medication Subset, as well as those remaining CM concepts not present in SNOMED-CT or with unwanted semantic types. From these remaining drug and CM concepts, the algorithm removed all those without relevant MeSH subheading qualifiers. For each article, this results in a set of MeSH drug concepts (with the subheading "Therapeutic Use," "Administration and Dosage," "Adverse Effects," "Poisoning," or "Toxicity") and a set of MeSH CM concepts (with the subheading "Drug Therapy," "Etiology," or "Chemically Induced"). At this point, results included sets of drug and CM concepts (each set being from

individual articles from which the concepts were drawn). Since each MEDLINE record contained only lists of SCR concepts (drugs) and major MeSH concepts (CMs) in the corresponding article and no relationships between the sets, the DEB2 algorithm generated all possible pairwise combinations of drugs and CMs as shown in Figure 2.

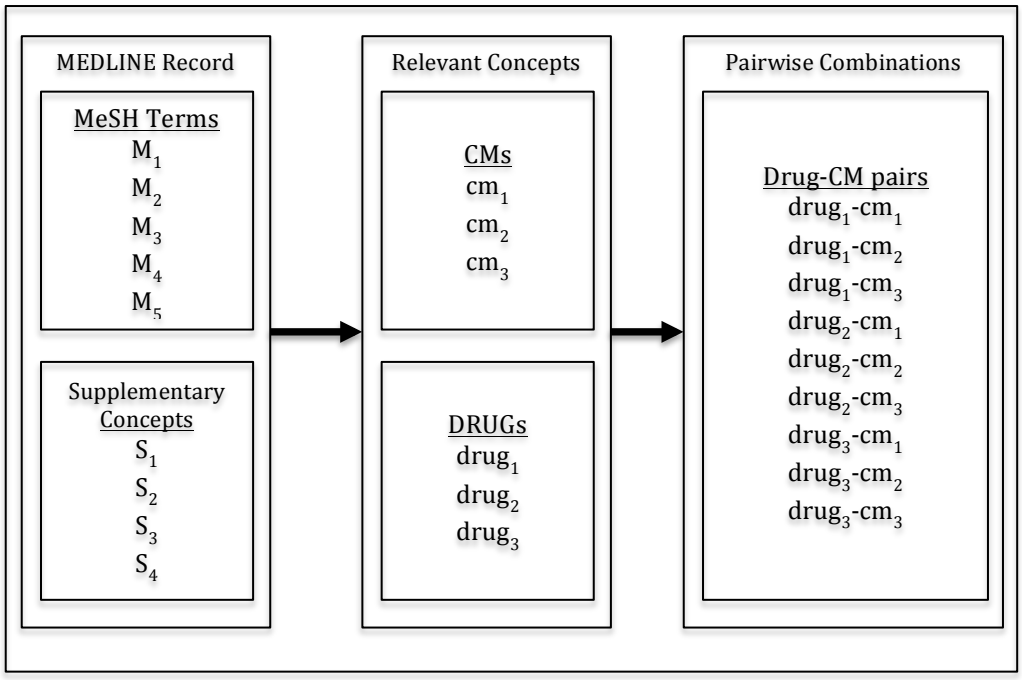


Figure 2. Illustration showing conversion from MeSH terms and supplementary concepts in the MEDLINE record to concepts to all pairwise drug-CM pairs.

Not all of these drug-CM pairs are true indications or valid ADEs. Nevertheless, the project team theorized that when one sums the counts of each drug-CM pair across all articles in MEDLINE, those with valid associations would appear more frequently and therefore have higher counts, as evidenced in prior work.[68] In other words, drug-CM pairs that appeared less frequently were less likely to represent true drug-CM relationships. Thus, the algorithm discarded those drug-CM pairs that were found in too few MEDLINE articles. Based on work by Avillach, et al.,[68] the project team initially set that threshold to occurrence in at least four articles.

The DEB2 construction algorithm next determined whether these drug-CM pairs represented an indication or ADE based on the MeSH subheadings, as illustrated in Figure

3. If the drug concept was qualified by the "Therapeutic Use" or "Administration & Dosage" subheading and the CM concept was qualified by "Drug Therapy," the algorithm classified the pair as an indication. If the drug concept was qualified by "Adverse Effects," "Poisoning," or "Toxicity" and the CM concept was qualified by "Etiology" or "Chemically Induced," the algorithm classified the pair as an ADE. Any drug-CM pairs with other combinations of qualifiers were discarded.

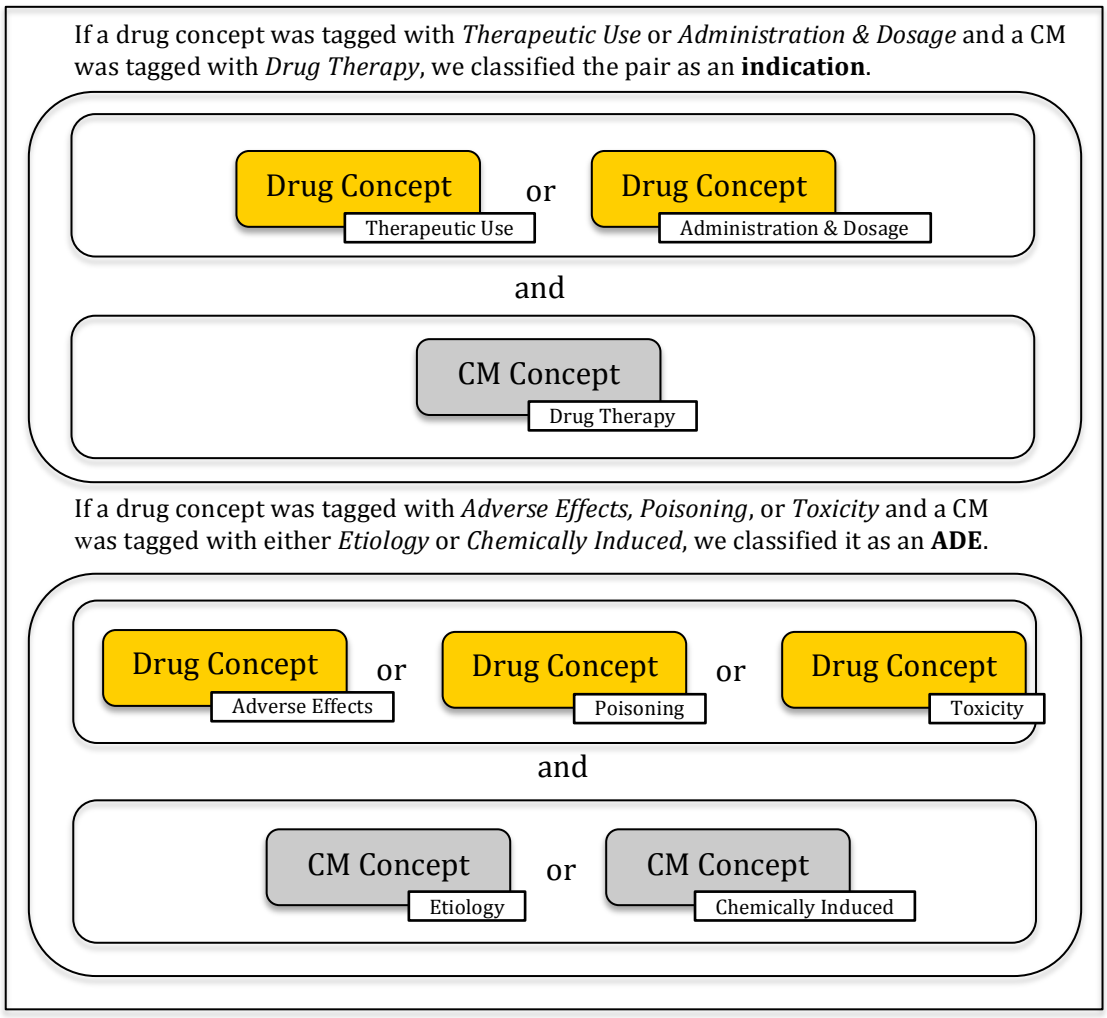


Figure 3. Illustration of how MeSH concepts and subheadings were combined to determine the relationship between the concepts.

Since the determination of relationship type was based solely on MeSH subheadings, this meant that a certain drug-CM pair could be classified as an indication based on one

MEDLINE entry and an ADE based on content from another. To address this, the algorithm summed the counts of each classification; that is, each drug-CM pair had an ADE-count and an indication-count. The final determination of relationship type within the MEDLINE component of the DEB2 was based on which count was higher.

After a preliminary analysis of the results from the MEDLINE component by project team members, the team found that drug-CM pairs were more likely to be related and represent true indication or ADE relationships if both the drug and CM concept terms were mentioned in the article abstract (as shown in previous work).[67] The DEB2 construction algorithm thus proceeds by searching the abstract (for those article entries that included an abstract in the MEDLINE database) for mentions of the corresponding drug and CM concept using a variety of methods. First, it uses simple regular expressions and string matching to look for the relevant MeSH terms in each abstract. For all the abstracts where the first method failed to find the terms, it employs the KMCI NLP tool[115] to process each abstract to find variations of the term. Finally, for all abstracts where the first and second method failed to find the terms, the DEB2 construction algorithm expands that term into all of its possible synonyms using a set of ~15,000 manually curated clinical synonyms created by Dr. Randolph A. Miller.[125] The algorithm then searches the abstracts for each of those newly generated synonyms.

The DEB2 construction algorithm double counts those drug-CM pairs where the related terms were present in the abstract by any of the three above-described methods. This provides more weight to those drug-CM pairs where both concepts were mentioned in the abstract. The project team also doubled the articled count threshold, requiring a pair to be present in at least eight articles before the algorithm classifies the drug-CM pair relationship as valid. These drug-CM pairs, the identified relationships, and the modified article counts are stored in a MySQL database table as the MEDLINE component of DEB2.

Source: SIDER2

As outlined in Chapter II, SIDER2 is a database of indications and ADEs extracted using text mining on FDA Structured Product Labels.[70] Extraction of the DEB2 SIDER2 component first involved downloading the SIDER2 data from the SIDER website on 30 July 2013.[69] The downloadable SIDER2 data consists of the following files: label_mappings.tsv,

adverse_effects_raw.tsv and indications_raw.tsv. For each label, the label_mappings file contains the generic and brand names of the medications, STITCH IDs from PubChem[126] for each drug, and a unique identifier for each label. Since DEB2 uses RxNorm RxCUIs instead of STITCH IDs, the algorithm mapped each drug product label to the appropriate RxCUIs by matching the drug name strings to RxNorm drug names in the DEB2 Medication Subset. The algorithm then discarded all unmapped labels.

The adverse_effects_raw.tsv and indications_raw.tsv are tab-separated files containing CM concepts, represented using MedDRA,[71]extracted from the "Adverse Reactions" and "Indications & Usage" sections of the labels, respectively. The algorithm loaded all concepts and discarded those that did not map to SNOMED-CT or had unwanted semantic types. The algorithm combined CMs extracted from each of the two sections with the corresponding drugs and classified the drug-CM pairs as indications and ADEs. Those CMs mentioned in both sections of a label were classified as indications since the project team observed that mentions of indications often appeared in the text of the Adverse Effects sections of the SPLs, as in "When treating for *high blood pressure*, adverse effects may include...". This technique is used by SIDER2 and DEB1. The extracted drug-CM pairs and the identified relationships are stored in a MySQL database table as the SIDER2 component of DEB2.

Source: MedlinePlus

MedlinePlus is a consumer health information website from NLM and NIH.[60] All information is manually curated from a variety of authoritative resources and is available as free text. Among other information, the site contains drug monographs answering such questions as: Why is this medication prescribed? What are other uses of the medicine? What side effects might this medication cause?

To extract the MedlinePlus component of DEB2, the DEB2 construction algorithms employed the MedlinePlus Connect Web Service.[127] The algorithm connected to the web service and searched by RxCUI for each drug from the DEB2 Medication Subset. If a MedlinePlus entry existed for a drug, the service returned the URL for the MedlinePlus drug monograph. The DEB2 algorithm then parsed the returned HTML monograph into relevant sections denoting common ADEs, serious ADEs, overdose effects, black box warnings, and

indications. It then processed each of these sections using the KMCI NLP tool, identifying all affirmatively mentioned CM concepts and, as before, discarding concepts that were not present in SNOMED-CT or had unwanted semantic types. The algorithm classified CMs extracted from the indications sections as *indications*; CMs extracted from the other sections mentioned above were all classified as *ADEs*. Any drug-CM pairs found to be both an indication and an ADE were reclassified as an indication, similar to the process used with the SIDER2 component. The drug-CM pairs and identified relationships are stored in a MySQL database table as the MedlinePlus component of DEB2.

Source: DrugBank

As discussed in Chapter II, DrugBank is a manually curated database combining chemical, pharmacological and pharmaceutical data with comprehensive drug target information.[72–74] Extraction of the DrugBank component of DEB2 first involved downloading the full XML-formatted DrugBank database on 28 July 2014. For each drug entry in the data, the DEB2 construction algorithms parsed the XML to extract the associated "Indication" and "Toxicity" fields. Using the ATC (Anatomical Therapeutic Chemical) code and CAS (Chemical Abstract Service) number from the XML dataset, the algorithm mapped each drug to the appropriate RxNorm RxCUI using the RxNorm RXCONSO table, discarding data from those drugs not present in the DEB2 Medication Subset.

For each remaining drug, the algorithm used the KMCI NLP tool to identify all affirmatively mentioned CM concepts in the "Indication" and "Toxicity" sections. As before, it discarded all concepts that were not present in SNOMED-CT or had unwanted semantic types. Since the "Toxicity" section contains not only ADEs found in humans, but also toxicity and ADE information found in animal studies, the algorithm first used regular expressions to remove those sentences in the "Toxicity" field in which animal subjects were mentioned. The algorithm classified CMs found in the "Indication" field as such and those found in the "Toxicity" field as ADEs. The drug-CM pairs and identified relationships are stored in a MySQL database table as the DrugBank component of DEB2.

Combining Drug-CM Pairs from All Sources

After extracting indication and ADE data from each of the five sources discussed above and loading it into five separate source-specific MySQL tables, the DEB2 construction algorithms merged the data into two combined tables: the *DEB2_full* table, which contains all drug-CM pairs from all sources, and the *DEB2_final* table, which contains all drug-CM pairs extracted from at least two sources.

To accomplish this, the algorithm first performed a union of the drug-CM pairs from each table and stored them in a new table. For each drug-CM pair, the new table contained a field listing the number of sources in which the drug-CM pair was found, as well as one field for each source table. These fields were populated with either "IND" or "ADE," denoting whether the source identified that drug-CM pair as being an indication or an ADE. A final field with the majority determination (indication or ADE) was added. That is, if there were more indication sources than ADE sources, the algorithm determined that the drug-CM pair was an indication; if there were more ADE determinations, the drug-CM pair was classified as an ADE. If there were an equal number of indication and ADE sources, it was noted in the determination field that there was a tie. This combined table contained all drug-CM pairs from every source and was called the *DEB2_full* table.

As noted above, the CMs were limited to the SNOMED-CT vocabulary and restricted their semantic types to ensure that DEB2 indications and ADEs were useful and specific concepts. However, the project team determined some overly broad or vague CMs, such as groups of diseases or generic drug allergies, remained. Therefore, as a final quality assurance step, the DEB2 construction algorithms removed any drug-CM pairs from a copy of the *deb2_full table* where the CM represented one of these CMs. Table 3 details those 57 CMs identified by the project team that were removed.

Lastly, the algorithm created the *DEB2_final* MySQL table from the above-mentioned copy of the *DEB2_full* table. To improve accuracy by ensuring that any indications or ADEs in DEB2 had corroboration in multiple sources, the algorithm next removed all drug-CM pairs with fewer than two sources. Lastly, the algorithm removed any drug-CM pairs where conflicting information resulted in uncertainty of the true nature of those drug-CM pair – that is, those pairs that were asserted to be both indications and ADEs by an equal number

of sources (the "ties"). This table, *DEB2_final*, is the final version of the DEB2 knowledgebase and contains only those pairs from at least two sources.

Table 3. CMs removed from DEB2 because they represented vague concepts such as disease classes and generic drug allergies.

UMLS CUI	Clinical Manifestation	UMLS CUI	Clinical Manifestation
C0003838	Arterial Occlusive Diseases	C0020517	Hypersensitivity
C0004364	Autoimmune Diseases	C0876973	Infectious disease of lung
C0005940	Bone Diseases	C0021390	Inflammatory Bowel Diseases
C0006111	Brain Diseases	C0021832	Intestinal Diseases, Parasitic
C0006145	Breast Diseases	C0022658	Kidney Diseases
C0016034	Breast Fibrocystic Disease	C0023895	Liver diseases
C0007222	Cardiovascular Diseases	C0024115	Lung diseases
C0007273	Carotid Artery Diseases	C0206062	Lung Diseases, Interstitial
C0007760	Cerebellar Diseases	C0027868	Neuromuscular Diseases
C0009326	Collagen Diseases	C0029928	Ovarian Diseases
C0009450	Communicable Diseases	C0030846	Penile Diseases
C0009759	Conjunctival Diseases	C0031090	Periodontal Diseases
C0012634	Disease	C0085096	Peripheral Vascular Diseases
C0595961	Dis-ease	C0032226	Pleural Diseases
C0013182	Drug Allergy	C0751438	Posterior pituitary disease
C0011609	Drug Eruptions	C0035309	Retinal Diseases
C0029944	Drug Overdose	C0036916	Sexually Transmitted Diseases
C0013221	Drug toxicity	C0162627	Skin Diseases, Bacterial
C0013386	Dyskinesia, Drug-Induced	C0085932	Skin Diseases, Bullous
C0014130	Endocrine System Diseases	C0037278	Skin Diseases, Infectious
C0014852	Esophageal Diseases	C0037928	Spinal Cord Diseases
C0015423	Eyelid Diseases	C0038354	Stomach Diseases
C0017411	Female Genital Diseases	C0039584	Testicular Diseases
C0016977	Gall Bladder Diseases	C0040128	Thyroid Diseases
C0017178	Gastrointestinal Diseases	C0005686	Urinary Bladder Diseases
C0017412	Genital Diseases, Male	C0042131	Uterine Diseases
C0018799	Heart Diseases	C0042373	Vascular Diseases
C0018824	Heart valve disease	C0042769	Virus Diseases
C0018939	Hematological Disease		

Comparison Between DEB1 and DEB2

The project team members compared the DEB2 knowledgebase against a random sample of previously reviewed drug-CM pairs from the original DEB1 study.[78,79] Unlike

DEB2, the DEB1 knowledgebase used UMLS CUIs to represent both drugs and CMs, did not restrict CM concepts to a single vocabulary, and did not restrict drug concepts to prescribable clinical drugs. Additionally, DEB1 only required drug-CM pairs to be present in one source.

To compare DEB2 to the reviewed sample from DEB1, a custom Perl script first mapped the RxCUI drug concepts in DEB2 to UMLS CUIs using the RXNCONSO table from RxNorm. Next, the script compared DEB2 against the DEB1 review set, containing drug-CM pairs rated by reviewers as indications, ADEs, both, or neither (implying the pair was nonsensical, vague, or simply wrong). We examined the overlap between DEB2 and DEB1, as well as whether DEB2 pairs agreed with the reviewers' earlier determinations (indications versus ADEs). We also examined those drug and CM concepts present in DEB1 but missing from DEB2 (due to DEB2's stricter criteria for those concept types).

Evaluation Methods

The project team chose to evaluate the DEB2 knowledgebase, more specifically the DEB2_final table, using manual review. To calculate the minimum sample size required for this evaluation, the team needed a rough initial estimate of DEB2's accuracy. Project team members reviewed a small random sample and estimated accuracy at 85%. From this, the team calculated the minimum sample size to separately determine accuracy of indications and ADEs in DEB2, plus or minus 5%. Most drug-CM pairs were present only in two of the five sources, but some were present in more than two sources. Team members hypothesized that the more sources in which a drug-CM pair were found, the more likely that it was true. Team members stratified the data to determine the accuracy of indication and ADE relationships extracted from two, three, four, and all five sources separately. The team added the necessary number of drug-CM pairs to the review set to allow us to determine that accuracy plus or minus 10%. A custom Perl script randomly selected drug-CM pairs to be included in the review set based on class (that is, indication or ADE from two, three, four, or five sources).

The review set consisted of equal numbers of indications and ADEs. Six physicians with specialty certification in Internal Medicine, Pediatrics, Hematology, Oncology, and/or Endocrinology reviewed this random sample from DEB2 to assess validity. Each physician

reviewed 100 indications and 100 ADEs, and two physicians were assigned to review each drug-CM pair. If the two reviewers agreed that the DEB2 determination was correct, the evaluation rated that drug-CM pair as true. If they agreed that the DEB2 determination was incorrect, it was rated as false. If the two reviewers disagreed with one another, a third reviewer was employed to break the tie and determine if DEB2's asserted relationship was true or false.

Reviewers each received a spreadsheet containing the drug-CM pairs they were to rate. It contained DEB2's asserted relationship for the pair and links with automated predetermined search strategies to search PUBMED[58] and Google[128] for the drug and CM combination. The physicians rated whether the asserted DEB2 relationship was true or false, noted if the CM was too vague or overly broad, and optionally provided any comments on the relationship. They used both their own knowledge and any information sources they wished to determine validity of DEB2's asserted relationships.

Results

Medication Concepts

The DEB2 construction algorithms first extracted 76,212 "clinical drugs" from the RxNorm prescribable subset. Using the techniques described in the Methods section, the algorithm normalized these clinical drugs to 3059 distinct single ingredient concepts. Next it removed those drugs with unwanted semantic types and terms resulting in 1844 single-ingredient medications. Examples of those concepts removed due to semantic type and terms are listed in Table 4. These 1844 single-ingredient medications comprised the DEB2 Medication Subset.

Table 4. Examples of those drug concepts removed from the DEB2 Medication Subset due to semantic types (*cell*, *enzyme*, *food*, *inorganic chemical*, etc.) and medication name strings.

Concepts removed due to semantic type		**Concepts removed due to name string**	
RxCUI	**Drug Name**	**RxCUI**	**Drug Name**
5499	Hydrogen Peroxide	1010965	Oregano allergenic extract
4537	Formic acid	851874	Hard maple pollen extract
17621	Aluminum sulfate	899420	Ginger allergenic extract
8793	Propyliodone	6459	Liver Extract
9884	Sodium Iodide	979384	North American coral snake antivenin
24902	Prussian blue	5880	Interferon Alfa
10938	Turpentine	833079	Smallpox vaccine live vaccinia virus
10323	Talc	283805	Bean Pod Extract
1299884	Ammonia	885187	Crotalus atrox antivenin
6579	Magnesium Chloride	900148	Giant wild rye pollen extract
11295	Water	852174	Pussy willow pollen extract
168	Acetic Acid	899576	Paper birch pollen extract
29261	Manganese chloride	1427022	Influenza B virus vaccine
34322	Potassium phosphate	6972	Mineral Oil
36696	Sodium metabisulfite	852419	Beech pollen extract
3210	Deoxyribonucleases	9949	Soybean Oil
2219	Cellulase	466570	Pumpkin seed oil extract
11423	Zinc Oxide	901291	Vanilla bean allergenic extract
91603	Tetanus immune globulin	899883	Pear allergenic extract
11431	Zinc	1309239	Coconut Oil
4716	Gelatin	885185	Crotalus adamanteus antivenin
2714	Collagen	259266	Hops extract
5496	Hydrofluoric Acid	2669	Cod Liver Oil
9863	Sodium Chloride	319780	Capsicum extract

Source: NDF-RT

The DEB2 construction algorithms extracted a total of 55,758 rows from NDF-RT with relevant relationships; this consisted of 722 “induces” relationships, 48,922 “may_treat” relationships, and 6114 “may_prevent” relationships. The extract included 9596 distinct drug concepts and 1030 distinct CM concepts. When normalized, the drug concepts mapped to 4133 distinct single-ingredient drugs. Of these, 1153 were in the DEB2 Medication Subset; the rest were not in the RxNorm prescribable subset or had unwanted RxNorm semantic types (as detailed above). Of the 1030 CM concepts, 958 CMs had representation in the SNOMED-CT vocabulary. After removing the unwanted semantic types, 831 distinct CMs remained. This resulted in the NDF-RT component of DEB2 containing a total of 4055 indication relationships and 78 ADE relationships (Figure 4).

NDF-RT Relationships	
Induces	722
May prevent	6114
May treat	48,922
Total	55,758

Extracted Drugs		Extracted CMs	
Drugs	9596	CMs	1030
Normalized Drugs	4133	CMs in SNOMED-CT	958
Drugs in Med. Subset	1153	CMs with wanted semantic types	831

Drug-CM Pairs Extracted	
Indications	4055
ADEs	78
Total	4133

Figure 4. Extracting the NDF-RT component of DEB2.

Source: MEDLINE

The MEDLINE 2014 Baseline Distribution contained 22,376,811 articles, 12,944,044 of which were tagged with the "Humans" MeSH heading. Among these articles, there were approximately 50 million DEB2-relevant MeSH heading tags. After the DEB2 construction algorithms removed those articles lacking both drug and CM MeSH concepts, 3,119,346 articles remained. After mapping SCR drug concepts to their associated MeSH terms and removing those MeSH concepts without subheading qualifiers, 3,088,272 articles remained. Each pairwise combination of drug and CM concepts in each article was checked for the relevant MeSH subheadings in the appropriate combinations. The algorithm retained 812,742 MEDLINE references with both drug and CM concepts and the appropriate MeSH subheadings consisting of 1,875,962 total drug-CM pairs. Next, the algorithm removed drug concepts not in the DEB2 Medication Subset, removed CMs without SNOMED-CT representation, and dropped drug-CM pairs occurring in less than four MEDLINE articles. A total 31,200 indications and 9650 ADEs comprised of 2522 distinct CMs and 1157 distinct medications remained.

To remove a number of incorrect associations as described above, the project team decided to double-count those articles where the drug concept and CM concept in question were mentioned in the article abstract and double the inclusion threshold to eight articles. Out of the 812,742 MEDLINE articles with both drug and CM concepts and the appropriate MeSH subheadings, 622,016 had full text abstracts available in the MEDLINE 2014 Baseline Distribution. The algorithm used three methods to identify mentions of the drug or CM terms in the abstracts. First, it used simple regular expression matching to search for the terms, finding 187,242 drug and CM concepts mentioned in abstracts. Next, it processed the abstracts using KMCI to search for the drug or CM concepts not found with the first

method, bringing the total of identified concepts to 359,713. Lastly, the algorithm generated variants of the search terms using Dr. Miller's manually curated set of ~15,000 clinical synonyms. This increased the total to 434,011 drug and CM concepts specifically mentioned in abstracts (a 20% increase over simple regular expression matching and NLP).

Doubling the minimum article count to eight and removing any drug-CM pairs that were found in fewer than eight articles, as described above, enabled construction of the final form of the MEDLINE component of the DEB2. It contains a total of 22,732 indications and 6331 ADEs comprised of 1121 distinct medications and 2253 distinct CMs (Figure 5).

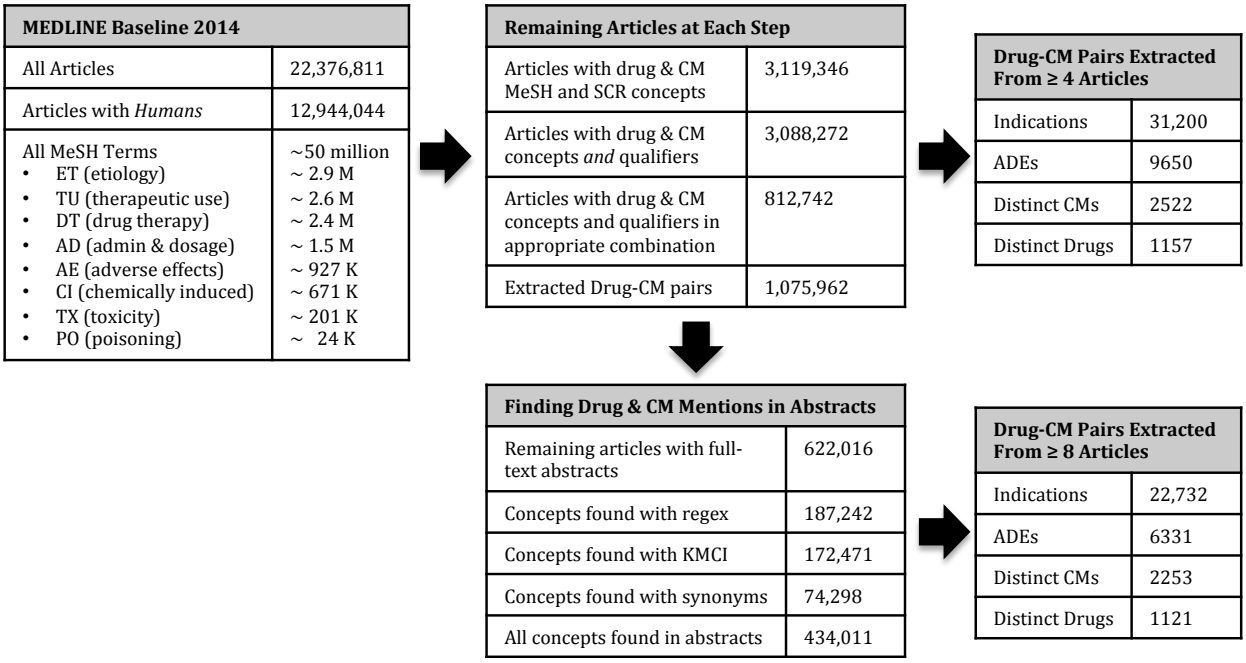


Figure 5. Extracting the MEDLINE component of DEB2.

Source: SIDER2

The DEB2 construction algorithms mapped 32,140 drug labels from the SIDER2 database to 20,507 distinct clinical drug concepts; normalizing these clinical drugs using RxNorm relationships found 931 distinct single ingredients in the DEB2 Medication Subset. Of the 5166 distinct CM concepts in SIDER2, 1351 concepts were not represented in SNOMED-CT and were discarded; 3815 distinct CM concepts remained. The algorithm found 3170 drug-CM relationships that were present in both the "Indications & Usage"

section and the "Adverse Reactions" section and classified them as indications; this resulted in the SIDER2 component of DEB2 containing 9646 indications and 83,956 ADEs (Figure 6).

SIDER2 Raw Data	
Label IDs	32,140
Clinical Drugs	20,507
Distinct Drugs from DEB2 Med. Subset	931

SIDER2 CM Concepts	
CMs in SNOMED-CT	3815
CMs not in SNOMED-CT	3170
CM Concepts by Section	
Indications & Usage	9646
Adverse Reactions	87,126

Drug-CM Pairs Extracted	
Indications	9646
ADEs	83,956
Total	93,602

Figure 6. Extracting the SIDER2 component of DEB2.

Source: MedlinePlus

The DEB2 construction algorithm queried the MedlinePlus Connect Web Service for all 1844 RxCUIs in the DEB2 Medication Subset and found 955 with drug monographs. After extracting the target sections and processing them using KMCI, the algorithm found 5325 drug-indication pairs from the "Indications" sections, 18,699 drug-ADE pairs from the "Side Effects" sections (9640 that were noted as serious side effects and 8734 that were noted as common), 3190 drug-ADE pairs from the "Overdose Effects" section, and 5004 from the "Boxed Warning" section. This resulted in the MedlinePlus component of the DEB2 containing 5325 total indications and 23,444 total ADEs (Figure 7).

MedlinePlus Drugs	
Drugs from DEB2 Medication Subset	1844
Drugs Found to Have a MedlinePlus Drug Monograph	955

CM Concepts by Section	
Indications	5325
Side Effects • Serious • Common	18,699 9640 8734
Overdose Effects	3190
Boxed Warning	5004

Drug-CM Pairs Extracted	
Indications	5325
ADEs	23,444
Total	28,769

Figure 7. Extracting the MedlinePlus component of DEB2.

Source: DrugBank

Of the 1844 medications in the DEB2 Medication Subset, the DEB2 construction algorithms found 972 in the DrugBank database. Using KMCI, the algorithm extracted 1263 distinct indication concepts and 774 distinct ADE concepts (1772 distinct CM concepts

total). This resulted in the DrugBank component of DEB2 containing a total of 3369 indications and 4788 ADEs (Figure 8).

DrugBank Drugs	
Drugs from DEB2 Medication Subset	1844
Drugs Found in DrugBank Database	972

Distinct CM Concepts by Section	
Indications	1263
Toxicity	774
Distinct Total	1778

Drug-CM Pairs Extracted	
Indications	3369
ADEs	4788
Total	8157

Figure 8. Extracting the MedlinePlus component of DEB2.

Combining Drug-CM Pairs from All Sources

Finally, the DEB2 construction algorithms combined data from each source table for a total of 137,023 drug-CM pairs from all sources. After removing the 57 vague CMs representing disease classes and generic drug allergies, 131,718 drug-CM pairs remained. These pairs are stored in the *DEB2_full* table. It contains 98,745 ADEs, 31,203 indications, and 1770 ties (drug-CM pairs that were found to be both indications and ADEs in an equal number of sources). The algorithm next removed those ties and drug-CM relationships that were found in only one of the five sources. This resulted in a total of 17,515 pairs consisting of 6183 indications and 11,332 ADEs stored in the *DEB2_final* table. The overall counts are summarized in Table 5. The detailed counts from each source are illustrated in Table 6 for the DEB2_full table and in Table 7 for the DEB2_final table (Figure 9).

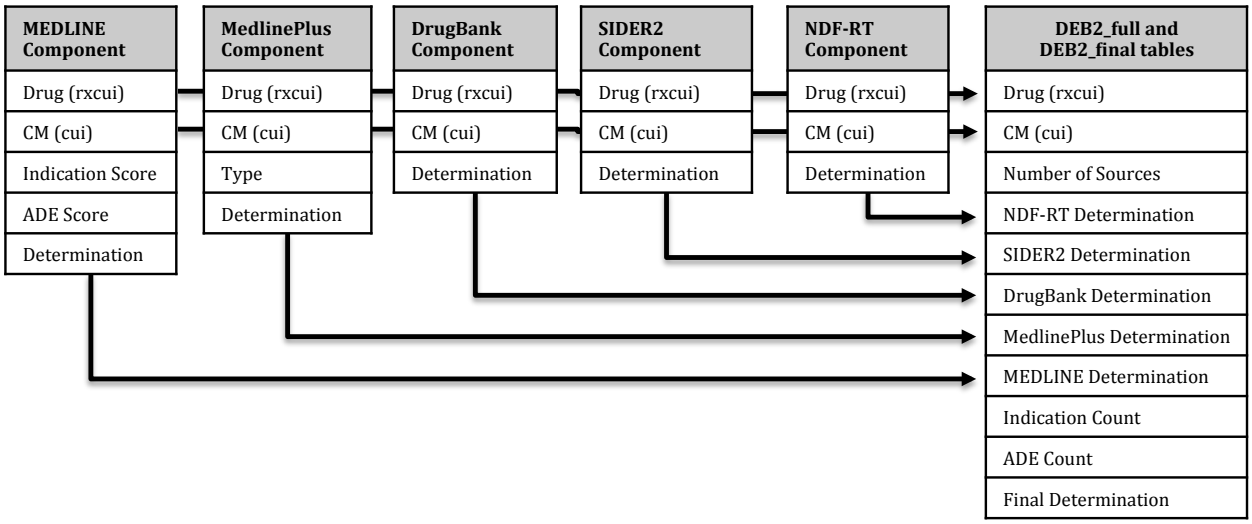


Figure 9. DEB2 Components tables merged to form the DEB2_full and DEB2_final tables.

Table 5. Summary of counts from the DEB2_full and DEB2_final tables.

	DEB2_full	DEB2_final
Unique pairs	131,718	17,515
Indications	31,203	6183
ADEs	98,745	11,332
Ties	1770	-
Unique drugs	1554	1162
Unique CMs	5661	1546

Table 6. Number of pairs from all sources in the DEB2_full table.

	NDF-RT	MEDLINE	MedlinePlus	DrugBank	SIDER2	Unique Total
IND	3926	21,803	4590	3165	9123	31,203
ADE	77	5625	22,448	4062	81,389	98,745
Total	4003	27,428	27,038	7227	90,512	131,718 (1770 ties)

Table 7. Number of pairs from all sources from the DEB2_final table.

	NDF-RT	MEDLINE	MedlinePlus	DrugBank	SIDER2	Unique Total
IND	2644	5054	2155	1882	4062	6183
ADE	20	2718	8722	1933	11,005	11,332
Total	2664	7772	10,877	3815	15,067	17,515

Of the 17,515 DEB2 drug-CM pairs from multiple sources, the vast majority were mentioned in only two sources. Specifically, 77% of pairs in DEB2 were listed in two sources, 17% were listed in three sources, 4.4% were listed in four sources, and 1.3% were listed in all five sources.

The project team also calculated the percentage of DEB2 drug-CM pairs in a given source that were also present in another source. A total of 68% of NDF-RT pairs, 54% of DrugBank pairs, 43% of MedlinePlus pairs, 33% of MEDLINE pairs, and 18% of SIDER2 pairs were corroborated by at least one other source. For those drug-CM pairs present in multiple sources, the percentage of drug-CM relationships that agreed with overall DEB2 consensus are shown in Table 8.

Table 8. Percentage of drug-CM pairs from each source that agree with the consensus of the other sources (when present in at least two source).

DEB2_full (ties included)		**DEB2_final (ties excluded)**	
MEDLINE	84.7%	**MEDLINE**	97.8%
SIDER2	88.3%	**SIDER2**	98.5%
MedlinePlus	93.0%	**MedlinePlus**	98.8%
DrugBank	95.8%	**NDF-RT**	99.5%
NDF-RT	96.8%	**DrugBank**	99.6%

DEB1 and DEB2 Comparison Results

The original DEB1 knowledgebase contained 137,194 drug-finding pairs consisting of 3242 distinct drugs and 8266 distinct CM concepts. There were 132,629 pairs (97%) listed by only one source, 4086 pairs in two sources, and 479 pairs in all three original sources. Since DEB1 did not require pairs to be in multiple sources, it is more comparable to the DEB2_full table, containing all drug-CM pairs. The DEB1 classified 79,284 pairs as ADEs and 57,854 pairs as INDs. The DEB2_full table contains a similar numbers of drug-CM pairs as DEB1, however DEB2_full contains only 36% as many unique drugs and 68% as many unique CMs. The DEB2_full table also contains more drug-CM pairs from multiple sources; only 3% of DEB1 was from multiple sources while 13% of the DEB2_full table is from multiple sources.

A total of 60% of the pairs in DEB2_full are present in DEB1; 58% of the pairs agree with the indication versus ADE determination of DEB1. A total of 87% of the pairs in DEB2_final are also in DEB1; of pairs in DEB2_final, 84% of pairs agree with the indication versus ADE determination of DEB1. Full results for the comparison are shown in Table 9.

The project team also compared DEB2 against a random sample of 659 previously physician-reviewed drug-CM pairs from the original DEB1 knowledgebase.[78] In the original analysis, reviewers agreed with DEB1 on 61% of drug-CM pairs. They disagreed with approximately 10% of DEB1 determinations, meaning drug-CM pairs classified as indications and were rated by reviewers as ADEs, and vice versa. The original reviewers also found that as many as 26% of drug-CM pairs in DEB1 were neither indications nor ADEs, mostly due to vague or inappropriate drug and CM concepts.

Table 9. Drug-CM pair overlap and agreement for DEB1, DEB2_full, & DEB2_final.

Drug-CM Pairs in DEB1 or DEB2	# Pairs	Percent of Source
Pairs in DEB2_final	17,516	
Pairs in both DEB2_final and DEB1	15,228	87% of DEB2_final
Pairs in DEB2_final and not in DEB1	2288	13% of DEB2_final
Pairs in DEB1 and not in DEB2_final	121,966	89% of DEB1
Pairs in DEB2_final and DEB1 and agree	14,796	84% of DEB2_final
Pairs in DEB2_full	131,718	
Pairs in both DEB2_full and DEB1	79,136	60% of DEB2_full
Pairs in DEB2_full and not in DEB1	52,627	40% of DEB2_full
Pairs in DEB1 and not in DEB2_full	58,058	42% of DEB1
INDs & ADEs in DEB2_full and DEB1:	76,076	58% of DEB2_full

Of the 659 drug-CM pairs in the previously reviewed sample from DEB1, 317 pairs were present in the DEB2_full table and 121 in the DEB2_final table. A total of 538 reviewed pairs were not present in DEB2_final. Of those 538 pairs, 196 had only one source and the remaining pairs were not in DEB2 for a variety of other reasons: 129 pairs did not have the drug concept in the DEB2 Medication Subset; 47 had CM concepts that were not in SNOMED-CT; 17 were missing both the drug from the DEB2 Medication Subset and the CM concept in SNOMED-CT; and 149 had the drug concept and CM concepts in DEB2, but the *pairs* were not found in the DEB2 sources. Detailed counts are summarized in Table 10.

Table 10. Drug-CM pairs in DEB2 and the DEB1 review set.

Type	Count
Drug-CM Pairs in Reviewed Set	659
Drug-CM Pairs in DEB2_full	317
Drug-CM Pair in DEB2_final	121
Drug-CM Pairs not in DEB2_final	**538**
Only one source	196
Drug not Present	129
CM not present	47
Drug & CM not present	17
Pair not present	149

Of the 121 drug-CM pairs present in DEB2_final, 90% agreed with the reviewers' determination of indication or ADE. Of the 133 DEB1 pairs rated as "neither indication nor

ADE" by reviewers previously, only two remained in DEB2 – a reduction of 98% of these inappropriate pairs between DEB1 and DEB2. Of the drug-CM pairs not in DEB2 but present in the DEB1 review set, project team members found that the majority contained drug class concepts, instead of individual drugs, and either vague or overly specific CM concepts and were correctly not included in DEB2. Examples of drug and CM concepts from DEB1, but not present in DEB2, are shown in Table 11 and Table 12.

Table 11. Sample medication concepts from DEB1 missing from DEB2.

DEB1 Medication Concepts
Immunologic Adjuvants
Polymyxin B Sulfate
Aminolevulinic Acid
Analgesics, Opioid
Anesthetics, Dissociative
ACE Inhibitors
Anti-rheumatic Agents
Anti-Infective Agents, Local
Drugs, Chinese Herbal
Estrogens

Table 12. Sample CM concepts from DEB1 missing from DEB2.

DEB1 CM Concepts
Reaction neurotic
Foot Diseases
Poisoning by anticholinesterase agents
Foreign Bodies
Eye symptom
Radiation Injuries, Experimental
Substance-Related Disorders
Chemotherapy-induced nausea and vomiting
Cardiac Event
Drug toxicity
Dog Diseases

Evaluation Results: Sample Size Determination

We defined a drug-CM pair "class" as a combination of relationship type (indication or ADE) and the number of sources in which the drug-CM pair was found; for example,

indications from two sources, indications from three sources, ADEs from 2 sources, etc. The minimum sample size needed to determine the overall accuracy (plus or minus approximately 5%) of indications and ADEs stratified only by relationship type is shown in Table 13. The actual sample size reviewed, shown in Table 14, included more drug-CM pairs so the experts could determine the accuracy (plus or minus approximately 10%) of each class of drug-CM relationship.

Table 13. Minimum sample size required to evaluate DEB2 and maintain the appropriate ratio of indications, ADEs, and number of sources.

	IND	**ADE**
2 sources	118.31	166.62
3 sources	47.14	24.04
4 sources	18.92	2.34
5 sources	6.63	0
Total	**191**	**193**

Table 14. Sample size reviewed by DEB2 physician reviewers and used to calculate the accuracy of each relationship class.

	IND	**ADE**
2 sources	140	180
3 sources	60	70
4 sources	50	50
5 sources	50	0
Total	**300**	**300**

DEB2 Evaluation Results: Inter-Rater Agreement

The 600 drug-CM pairs randomly selected for review were split into three groups of 100 indications and three groups of 100 ADEs each. Each group contained an equivalent number of pairs from two, three, four, or five sources. Each of the six DEB2 physician reviewers were given one set of indications and one set of ADEs; they reported taking between one and two hours to evaluate the pairs. For each set of 100 indications and 100 ADEs, the project team calculated Cohen's Kappa to determine pairwise inter-rater agreement. The kappa value and percent agreement between pairs of reviewers are shown in Table 15.

Table 15. Inter-rater agreement between reviewers.

Group	Kappa	Agreement	% Agreement
Indications-A	0.457	Moderate	92%
Indications-B	0.237	Fair	90%
Indications-C	0.408	Moderate	90%
Indications-Avg.	0.372	Fair	91%
ADE-A	0.038	Poor	82%
ADE-B	0.021	Poor	81%
ADE-C	0.315	Fair	96%
ADE-Avg.	0.054	Poor	86%

Evaluation Results: Expert Review

Results of the manual review are shown in Table 16. Of the 600 reviewed drug-CM pairs, the physician reviewers disagreed on a total of 69 pairs, or approximately 12% of the review set (28 indications and 41 ADEs). The third reviewer adjudicated the results. He determined that 29 were true (meaning DEB2 was correct) and 13 were false. He also found that 27 of the drug-CM pairs contained a CM that was too vague or that the relationship was too complex to determine whether the drug-CM pair represented a true indication or ADE. The full set of these indeterminate drug-CM relationships, along with a summary of reviewer comments, is given in Table 17. A total of 27 drug-CM pairs (18 ADEs and 9 indications representing less than 5% of all reviewed pairs) were classified in this manner as vague, complex, or unknown.

Table 16. Summary of DEB2 manual review results.

Initial review set (total of 600 drug-CM pairs)			
Indications	300	**ADEs**	300
Initial reviewers' ratings (where reviewers agreed)			
True	262	**True**	256
False	10	**False**	3
Disagreed	28	**Disagreed**	41
Adjudication results (when initial reviewers disagreed)			
True	11	**True**	18
False	8	**False**	5
Unknown	9	**Unknown**	18

Table 17. Drug-CM pairs rated as vague, complex, or unknown during adjudication. The relationships (IND or ADE) and number of sources are listed in the *Type* column.

Type(#)	Drug	CM	Comments
IND (2)	Triamcinolone	Colitis	**Too vague**; this steroid is used in ulcerative colitis but not for other forms of colitis. Also, there are now preferred oral GI steroids and triamcinolone is normally a topical preparation.
IND (2)	Dextroamphetamine	Mental disorders	**Too vague**; this drug is only given for narcolepsy or ADHD, not all mental disorders.
IND (2)	Clomiphene	Hyperprolactinemia	**Complex**; this drug is used in combination to induce ovulation in women with hyperprolactinemia. It does not really treat the *disease*, but is used to treat the *effects* of the disease, which is infertility (or more accurately anovulation).
IND (3)	Busulfan	Leukemia, Myelocytic, Acute	**Complex**; in combination with cyclophosphamide, Busulfan is used prior to bone marrow transplant in AML patients to improve survival. It is also used in *chronic* myeloid leukemia (CML).
ADE (2)	Ciprofloxacin	Muscle Weakness	**Complex**; Ciprofloxacin can exacerbate muscle weakness in patients with myasthenia gravis.
ADE (3)	Cyclophosphamide	Dyspnea	**Complex**; in patients with scleroderma, this drug treats interstitial lung disease, which causes dyspnea. Conversely, some case reports suggest this is an ADE, while others indicate possible cardiac and pulmonary causes of dyspnea when used in combination with other common chemotherapeutics.
ADE (3)	Sirolimus	Lymphoma	**Complex**; lymphoma is a reported complication of immunosuppression by sirolimus, but sirolimus is also indicated in lymphoma patients post bone marrow transplant (to prevent graft-vs-host disease).
ADE (2)	Tretinoin	Ulcer	**Vague**; known ADEs include blistering of the skin, but the term "ulcer" is too broad here. Tretinoin can be used to treat diabetic foot ulcers and scrotal ulcers in leukemia patients.
IND (2)	Tranexamic Acid	Subarachnoid Hemorrhage	**Questionable**; was previously considered an indication, now some studies urge caution and others show benefit.
ADE (2)	Testosterone	Paresthesia	**Questionable**; it is listed by FDA as a side effect as "numbness and tingling," but without proof. One MEDLINE case report of a testosterone injection into a nerve caused this symptom.
ADE (3)	Celecoxib	Somnolence	**Questionable;** Literature has many studies comparing ADEs of opiates to NSAIDs which list somnolence as side effect of opiates only; might confuse DEB algorithm.
ADE (3)	Risperidone	Agitation	**Questionable;** good evidence that agitation is both an indication and ADE of risperidone.
ADE (2)	Amphotericin B	Arthralgia	**Questionable;** Used to treat fungal causes of joint infections.
ADE (3)	Cyclophosphamide	Urticaria	**Questionable;** Rare reports of urticaria due to drug, many more reports of drug treating chronic autoimmune urticaria; also general allergy concern applies here
ADE (3)	Dexamethasone	Pseudotumor Cerebri	**Questionable;** Used in distant past as therapy but might induce it; no longer used as a prescribed indication. I think this is more of a contra-indication for this disorder as opposed to an ADE.
ADE (2)	Ketamine	Seizures	**Questionable;** ketamine is used to control some types of seizures; FDA warnings say it can induce seizures but literature on this hard to find (one study found ketamine less effective than sleep at inducing seizures). Not clear if it is more IND or ADE.
IND (2)	Prednisolone	Erythema	**Vague;** can be used against certain causes of erythema but not others.
IND (2)	Mechlorethamine	Carcinoma	**Vague;** Mechlorethamine is a chemotherapy drug used in hematologic malignancy (not carcinoma) but has not been used for decades.
IND (2)	Clonazepam	Neuralgia	**Vague;** Clonazepam is effective in 65% of patients with trigeminal neuralgia; little evidence to support the use of anticonvulsants as first-line treatment for neuropathic pain.
IND (4)	Promethazine	Pain	**Vague;** frequently given with pain medicine to help with anxiety, nausea, etc. Alone it would be inadequate for most pains.
ADE (2)	Ibuprofen	Asthenia	**Vague;** Weakness may be more commonly used
ADE (2)	Aripiprazole	Confusion	**Vague;** Used to treat schizophrenia and confused states; more likely an indication than an ADE.
ADE (2)	Ceftizoxime	Exanthema	**Vague;** this is again part of the allergic reaction spectrum
ADE (3)	Dicloxacillin	Pruritus	**Vague;** common form of allergic reaction.
ADE (3)	Albendazole	Fever	**Vague;** fever is part of drug allergy syndrome. Also, it is likely that the fever indicated that the underlying infection is not being treated, not necessarily an ADE.

Table 18. Combined results of manual DEB2 review and adjudication (unweighted).

	Indications		ADEs	
	Sum	**Percentage**	**Sum**	**Percentage**
TRUE	273	91%	274	91%
FALSE	18	6%	8	3%
Other	9	3%	18	6%

When considering all reviewed pairs, reviewers agreed with 91% of the DEB2 indications and 91% of the DEB2 ADEs, and disagreed with 6% of the indications and 3% of ADEs (Table 18). However, those numbers do not take into account the ratio of indications and ADEs found in each class. The weighted average of the rates, calculated for indications and ADEs in two, three, four or five sources, is shown in Table 19. To incorporate the variability of sampling from each cluster, project team members used the conditional variance formula and the multivariate delta method.

Table 19. Estimated DEB2 accuracy and 95% confidence interval, stratified by relationship type (indication or ADE) and number of sources.

Class		Estimate	Variance	95% Confidence Interval		
				Lower Bound	**Upper Bound**	**Width/2**
BOTH	Overall	**86%**	0.0002	83%	89%	3.1%
IND	All	**88%**	0.0004	84%	92%	3.9%
ADE	All	**84%**	0.0002	81%	87%	3.1%
IND	2 sources	**79%**	0.0012	72%	86%	6.9%
IND	3 sources	**92%**	0.0014	85%	99%	7.2%
IND	4 sources	**96%**	0.0012	89%	100%	5.3%
IND	5 sources	**98%**	0.0035	87%	100%	6.5%
ADE	2 sources	**84%**	0.0008	79%	90%	5.4%
ADE	3 sources	**83%**	0.0021	74%	92%	9.1%
ADE	4 sources	**94%**	0.0073	79%	100%	11%

The analysis found that drug-CM pairs in DEB2 are 86% accurate overall, plus or minus 3.1%. Indications are 88% accurate and ADEs are 84%. Accurate. Drug-CM pairs found in more sources generally have higher accuracy, however confidence intervals are wider due to smaller sample size and higher variance.

Discussion

Summary

The DEB2 knowledgebase combines medication data from five different publicly available sources. Each of these sources is frequently updated and, since DEB2 is constructed algorithmically with little manual intervention, it can be reconstituted periodically to take advantage of these updates. By requiring each drug-CM pair to be present in at least two sources, not all true relationships are captured but the accuracy of those indication and ADE relationships in DEB2 is greatly improved over DEB1.

DEB2 Sources

Most drug-CM relationships found in the five sources were found only in one source. The final version of DEB2, the aforementioned DEB2_final table, requires some amount of corroborating evidence by requiring drug-CM relationships to be present in at least two sources. This reduces the number of pairs in DEB2 from 131,718 to 17,515, but project team members believe these are of higher quality (see Table 5). This is not to say that the approximately 115,000 pairs that are present in only one source are incorrect, just that the pairs from at least two sources are more likely to represent true relationships.

Each of the data sources used in DEB2 focus on different types of information. The NDF-RT and MEDLINE component both contain more indication relationships. The MedlinePlus, DrugBank, and SIDER2 components all contain more ADE relationships. The SIDER2 component contains nearly ten times more ADEs than it does indications (Table 6). This is to be expected, as the SPLs used as the source for SIDER2 generally contain very few indications for a given drug but list all known ADEs.

Of the 17,515 DEB2 drug-CM pairs from multiple sources, most came from only two sources. When we calculated the percentage of DEB2 drug-CM pairs in a given source that were also present in another source, we found that 68% of NDF-RT pairs, 54% of DrugBank pairs, 43% of MedlinePlus pairs, 33% of MEDLINE pairs, and 18% of SIDER2 pairs were corroborated by at least one other source. Since NDF-RT is high-quality, relatively small compared to the other sources, and we did not need to use NLP to extract the concepts, it makes sense that more of its drug-CM relationships would be present in other sources.

Similarly, since SIDER2's data comes from SPLs, it contains many possible and unverified ADEs for each drug. Thus, it makes sense that fewer of SIDER2's relationships are corroborated by other sources.

Since many regard the SPL as a definitive source of drug information, it is not surprising that 11,005 of the 11,332 ADEs in DEB2 (97%) were included in SIDER2. SIDER2 also contained 4062 of the 6183 indications in DEB2 (65%). The 35% of indications not found in the SPL are likely a combination of off-label indications, synonymous terminology used in other sources but not recognized as such, and, possibly, NLP mismatching. Off-label indications would be more likely to be found in MEDLINE, MedlinePlus, and DrugBank.

Combined DEB2

As expected, indication relationships were generally more likely to be rated as true by external reviewers. This is likely due to medication indications being more widely agreed upon than ADEs. While there are only a limited number of approved indications for a medication, and a relatively small number of accepted off-label uses, medications are often suspected of causing a wide variety of ADEs. Many ADEs mentioned in the FDA product labels often have no proven causal relationship to the associated medication. Additionally, studies showing new ADEs are often not reproduced.

Also as expected, indication and ADE relationships were usually more likely to be rated by external reviewers as true when found in more sources. Indications found in three, four, and five sources were 92%, 96%, and 98% accurate respectively, as opposed to 79% accuracy when found in only two sources. On average, accuracy of ADEs found in three sources was very similar to ADEs found in only two sources (83% and 84% respectively). ADEs found in four sources were found to be 94% accurate. There were no ADEs found in five sources, most likely because the NDF-RT component contained very few ADEs

The project team found that between 85% and 97% of drug-CM pairs agree with the consensus of the other sources when the pair is present in multiple sources (Table 8). After excluding ties in the DEB2_final table, between 98% and 100% of drug-CM pairs agree with the consensus. Therefore the drug-CM relationships in DEB2_final are in broad agreement across sources.

Comparison Between DEB1 and DEB2

In the original DEB1 analysis, reviewers found that 61% of drug-CM pairs in DEB1 were true indications and ADEs. Reviewers disagreed with 10% of those drug-CM pairs, meaning they determined the pairs were actually indications when DEB1 classified them as ADEs, and vice versa. Reviewers classified 3% of DEB1 pairs as "both indications and ADEs," and 26% of DEB1 pairs as "neither indications nor ADEs." The improved methodology of DEB2 resulted in an 86% of DEB2 pairs being rated as true, less than 10% rated as false, and only 5% rated as vague, complex, "both," or "neither" drug-CM pairs. Those changes represent 30% increase in accuracy and a 50% decrease in incorrect drug-CM pairs (that is, 29% false or "neither" in DEB1 to 15% false and vague combined in DEB2).

A total of 121 drug-CM pairs from the DEB2_final table were present in the available DEB1 expert review set. The determination of DEB2 agreed with reviewers 90% of the time (92% for ADEs and 89% for indications) – a substantial improvement over the 61% agreement of DEB1 with those reviewers' determinations.

In the DEB1 review set, experts classified 26% of pairs as "neither indication nor ADE." These "neither" pairs contained both drug and CM concepts that were too vague, overly specific, or often related to non-human medications and CMs.[78] Vague drug concepts were often due to generic drug classes, such as *Opioid Analgesics*, *ACE Inhibitors*, or *Chinese Herbal Drugs*. The "neither" CM concepts were often groups of diseases or non-specific symptoms, such as *Sheep Diseases*, *Eye Symptom*, *Drug Toxicity* or, *Cardiac Event*. Of the 133 pairs in the DEB1 review set classified as "neither," only two were present in DEB2; based on this small sample, that is a 98% reduction in false positive associations in DEB2. This is likely due to the restrictions places upon drug and CM concepts – that is, requiring CMs to be in SNOMED-CT and limiting medications to the DEB2 Medication Subset. Restricting CM and medication concepts improved both relevance and corroboration among sources from DEB1 to DEB2.

DEB2 Evaluation: Inter-Rater Reliability

Landis and Koch provided one of the most common interpretations of the Kappa statistic,[129] but it is not universally accepted.[130] They reported that a Kappa less than or

equal to zero indicates poor agreement; a Kappa of 0.01-0.20, slight agreement; 0.21- 0.40, fair agreement; 0.41-0.60, moderate agreement; 0.61-0.80, substantial agreement; and almost perfect agreement at 0.81-1.00.

While agreement among reviewers for indications was fair to moderate, agreement among ADEs was only fair to poor. As mentioned above, while indications for most medications are limited and widely agreed upon, the causal relationship between medications and ADEs is not always agreed upon. While the SPL provides a litany of possible ADEs, they are not necessarily causal. Many ADEs are widely agreed upon with a well-understood mechanism of action, but others are only suspected and different sources require different levels of evidence. Still, the percent agreement among the expert reviewers is greater than 80% for all indications and ADE pairs.

DEB2 Evaluation: Expert Review

After adjusting for the number of pairs present from two, three, four, and five sources, the analysis found that approximately 86% of both indications and ADEs in DEB2 are true, plus or minus 3%. Of the remainder, reviewers found that less than 5% involved concepts that were too vague or relationships that were too complex. This is a dramatic decrease from the original DEB1 knowledgebase.

As shown in Table 17, rating indications and ADEs is difficult and sometimes subjective. For example, reviewers classified the DEB2 ADE *sirolimus induced Lymphoma* as false (more specifically, "complex"). It is listed as an ADE in multiple sources but some reviewers felt that, since lymphoma is a reported complication of immunosuppression (the intention of sirolimus administration), it is not necessarily an ADE of the drug. Similarly, reviewers rated the DEB2 indication *tranexamic acid for subarachnoid hemorrhage* as false ("questionable") because, while it was once a commonly accepted indication, newer studies urge caution. Another example, *seizures* as an ADE of *ketamine,* was rated as false ("questionable") because ketamine is used to control some types of seizures.

The appropriate level of specificity needed in a drug knowledgebase varies depending on the intended use of the knowledgebase. For example, one application might require broader information while another would require more specific information. Additionally, different reviewers required different levels of specificity when rating

indication or ADE relationships as true or false. For example, one reviewer, an oncologist, demanded a higher level of specificity for indications of antineoplastic drugs than another reviewer, an internal medicine physician.

DEB2 Compared to Other Publicly Available Drug Knowledgebases

As discussed in Chapter II, Salmasian, et al.,[80] systematically reviewed a number of medication indication knowledgebases (KBs). They included the NDF-RT, SIDER, MEDI, and works by McCoy, et al., and others.[63,69,76,77] Like the majority of those indication KBs, DEB2 is derived and updated using semi-automated methods. DEB2 normalizes medication concepts using RxNorm relationships, something only two out of the seven reviewed KBs included. DEB2 contains 1162 normalized medications and 6183 indications, greater than the average of the reviewed KBs. While DEB2 has fewer distinct CM concepts, many of the other KBs have problems with redundant CM representation due to their choice of vocabularies. Additionally, DEB2 only includes indications from at least two of its five sources to improve precision; in that regard, the MEDI high-precision subset is similar. DEB2 is the only KB in the set that represents CMs using SNOMED-CT. And while DEB2 does not include any formal methods for normalizing CM concepts, the NLP methods used tended to match concepts to the most appropriate general SNOMED-CT concept.

The project team did not find a similar formal comparison of ADE knowledgebases. Previous work comparing DEB1 to SIDER[79] found that since SIDER is derived from the SPL, most included ADEs are not corroborated elsewhere. Since DEB2 includes a newer version of SIDER but also requires corroboration from other sources, the precision in DEB2 is likely higher while the recall is lower. Similar to other ADE and indication knowledgebases, DEB2 is stored in a tabulated format.[80] Like most of the reviewed KBs, it does not contain multi-ingredient drugs, contraindications, or differentiate between different subtypes of indications (prevention or treatment) or ADEs (Common ADEs, serious ADEs, “Black Box” warnings, etc.), but current methodologies can facilitate inclusion of that additional information in the future.

Limitations

Reviewers found that the vast majority of indications and ADEs in DEB2 are accurate, but there is room for improvement. Drug knowledgebases must take into account complex indication information.[79,80] Reviewers classified approximately 10% of indications and ADEs as "vague, complex, or unknown." To correct these problems, future versions of DEB will require the inclusion of additional relationship types and criteria. The DEB2 knowledgebase does not include contraindications or differentiate between treatment indications versus prevention indications, or between common ADEs, serious ADEs, or overdose effects.

The analysis of complex drug-CM pair relationships (Table 17) revealed that some of the indications found in SIDER2 and MEDLINE are also out of date. Future versions should take into account the published date of both biomedical research articles from MEDLINE and drug label information from the SPLs; newer versions of product labels and research articles should likely be weighted more heavily than older ones. The analysis also revealed that some medications are indicated only along with other drugs (co-medications) or to treat a confluence of symptoms. These other criteria should be included in future versions of DEB.

The DEB2 requirement that drug-CM pairs appear in more than one source improved the precision of the knowledgebase. However, this resulted in many true indications and ADEs being excluded due to lack of corroboration. This was frequently the result of CMs being matched to different SNOMED-CT concepts despite their similarity. By using some form of concept normalization, it should be possible to match similar CMs and recognize corroboration across sources, even when disparate terminologies are used in different sources.

Finally, DEB2 only includes single-ingredient drugs. To improve the usefulness of the knowledgebase, future versions should include indications and ADEs for multi-ingredient drugs. Similarly, future versions of DEB2 should include drug-drug interactions to address those ADEs due to a co-administration of medications, not just those from idiosyncratic effects of single drugs.

Conclusion

The analysis found that the DEB2 knowledgebase is 86% accurate overall, with indications being slightly more accurate and ADEs slightly less accurate. The project team expected that ADEs would be less accurate due to their somewhat subjective nature. While DEB2 does not contain all possible ADEs, the two-source requirement ensures that the vast majority of indications and ADEs are correct and corroborated across sources. The improvements in DEB2 address many of the shortcomings found in DEB1. The project team believes the DEB2 knowledgebase can be useful for biomedical informatics research in the areas of pharmacovigilance, clinical data mining, clinical phenotyping, and clinical decision support systems, among others. The remainder of this dissertation discusses an application of structured drug knowledge to the area of clinical decision support.

CHAPTER IV

DEVELOPING THE ANTIHYPERTENSIVE ADER ALERTING SUBSET

Overview

The goal of the Adverse Drug Effect Recognizer (ADER) is to identify potential ADEs that patients have experienced recently prior to or at the time of hospital admission. It requires as input both admission H&P notes and a database of known ADEs for the system to detect. The DEB2 knowledge base is a verifiably accurate source of ADEs; still, not all ADEs in DEB2 are valid and not all true ADEs are included. Also, DEB2 does not contain class-based ADEs. Additionally, alerting physicians about unlikely or inconsequential ADEs may be counterproductive.

A successful ADE detection system should focus on ADEs of clinical significance. For the pilot evaluation, the project included only the ADEs of antihypertensive medications. Antihypertensive medications are frequently used in the practice of internal medicine and have a wide variety of ADEs. To develop this set, project team members started with an initial set derived algorithmically from DEB2. Next, the team worked with expert clinical pharmacists to manually transform and curate that set to create the antihypertensive ADER Alerting Subset (AAS). This chapter details the design and creation of the AAS, which was used in the ADER pilot evaluation detailed in Chapters V and VI.

Materials

This work was performed on a MacBook Pro with a 2.6 GHz Intel Core i7 processor and 16 GB of RAM and a Linux server with forty-eight 2.2 GHz AMD Opteron cores and 256 GB RAM. All data processing scripts were written in Perl 5.10.0. Development of DEB2 was described in Chapter III.

Methods

The project team decided to evaluate ADER using its ability to detect side effects from antihypertensive medications – drugs used to treat high blood pressure. The team selected antihypertensives because patients using antihypertensive medications are very

common in the adult Internal Medicine population. Additionally, the drugs included in the set of antihypertensive medications have an adequate variety of well-established side effects.

To create the AAS, the project team first determined the set of antihypertensive drugs. Next, the team extracted individual ADEs from DEB2 for those drugs and algorithmically determined drug class ADEs for antihypertensive drug classes. Lastly, expert clinical pharmacists manually curated the final set. The methods used are described below.

Use of H&P Test Corpus to Determine Baseline Rate of ADEs

The project team analyzed approximately 350,000 Vanderbilt University Hospital de-identified retrospective H&P notes from the SD repository to discover the prevalence of antihypertensive medication usage and the detectable rate of related ADEs in the patient population at Vanderbilt. The project team used the KMCI, SecTag, and MedEx NLP tools to identify medications and affirmatively mentioned CMs in the relevant sections from each H&P note. The methodology used to classify CMs, drugs, and relevant H&P sections was identical to that detailed in Chapters III and IV.

Determining Antihypertensive Medications and DEB2 ADEs

Using multiple sources, including the authoritative reports of the Joint National Committee on the Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC7, JNC8),[131–133] the project team determined the medications and drug classes that belong to the broad superset commonly referred to as "antihypertensive drugs."

Using the H&P test corpus, the algorithm identified the number of notes containing mentions of each medication. The algorithm also calculated the number of notes containing drugs from each drug class, as well as the within-class prevalence of each drug's usage. For each antihypertensive medication, a Perl script extracted all individual drug-ADE pairs listed in DEB2. In epidemiology, relative risk (RR) is the ratio of the probability of an event occurring in an exposed group to the probability of an event occurring in a comparable non-exposed group. For each extracted drug-CM ADE pair, the algorithm calculated the relative risk of the drug being associated with the CM in the H&P test corpus. In this case,

the "event" was the presence of a documented potential ADE in a given note and the "exposure" was mention of a given drug in the same note. The baseline risk was the rate of the CM in all patients in the population.

The project team empirically decided that, for the initial set of potential antihypertensive ADEs, only drug-ADE pairs with a relative risk score greater than or equal to 1.2 were to be included in the set of antihypertensive ADEs. Requiring this minimum relative risk score ensured that the ADE in question was more likely caused by the drug than by chance.

Deriving Drug Class ADEs

For each of the 16 drug classes determined to be within the set of antihypertensives, the project team attempted to identify class-specific side effects using DEB2. For each ADE associated with a drug in a given class, the team calculated the percentage of drugs in that class which listed that ADE.

For classes containing at least three drugs, the algorithm categorized an ADE a *class side effect* if at least 85% of the drugs in that class individually had the ADE listed or if the relative risk of a drug-ADE pair in our test corpus was greater than 1.2 for at least 50% of the drugs in the class. For those classes with only two drugs, both drugs were required to have the ADE listed with a relative risk greater than 1.0 for the ADE to be considered a class side effect.

ADEs of All Antihypertensives

The project team also identified any ADEs common to all antihypertensive drugs using a similar method to that which identified class side effects. After reviewing different possible inclusion thresholds, the project team empirically concluded that an ADE was a side effect of all antihypertensives if the relative risk of a drug-ADE pair (in the H&P test corpus) was greater than 1.2 for at least 35% of all antihypertensive drugs.

Compiling the Preliminary AAS

Next, the algorithm combined the individual ADEs, the class ADEs, and the antihypertensive-wide ADEs extracted from DEB2. If a drug-ADE pair was included as an

antihypertensives-wide ADE, it was removed from both the class ADE and individual ADE sets in which it previously appeared. Similarly, if a drug-ADE pair was included in a class ADE, it was removed from the listings where it previously appeared.

After reviewing the AAS ADEs, the project team manually removed four CMs considered either vague or non-specific. The removed CMs were: *C0027497 – Nausea, C0004093 – Asthenia, C0015672 – Fatigue*, and *C0342579 – Electrolyte imbalance*. The team recorded the overall AAS ADE set in a Microsoft Excel spreadsheet and presented it to expert pharmacists for review. The spreadsheet included the RxNorm RxCUI, drug name, drug class, CM CUI, CM name, and the relative risk score of the drug causing the ADE in the H&P test corpus.

Expert Review and Curation of the AAS

The project team provided the list of selected antihypertensives medications and the spreadsheet containing the initial, algorithmically created AAS to four reviewers. All reviewers were Doctors of Pharmacy; the review team included an Internal Medicine Service Clinical pharmacist, the Medication Safety Program Director, the Informatics Pharmacy Manager, and the Director of Operations for the Vanderbilt University Hospital Pharmacy. They reviewed the initial version of the AAS to determine if the ADEs were in their opinions valid, and expressed their preferences as to whether each drug-ADE pair should be used for the ADER pilot implementation. The reviewers determined if any of the class or individual effects should be removed or if any other known ADEs should be added to the AAS.

Final Version of the AAS

The project team reviewed the results of the AAS analysis completed by the pharmacists. In consultation with the expert reviewers, the project team created the final version of the AAS. The project team compared the final, manually reviewed version of the AAS against the algorithmically created initial set to evaluate the methodology used to create the initial set used for review.

Results

Antihypertensive Medications

Using the methods explained above, the project team determined the medications and drug classes that belong to the set of "antihypertensive drugs." The set consisted of 59 medications and 16 drug classes (including two separate classes for cardioselective and non-cardioselective beta blockers). The drugs, their classes, and the prevalence of the drugs in the H&P test corpus appear in Table 20.

Table 20. List of antihypertensive medications, including drug class and prevalence, in the H&P test corpus of ~350,000 notes. Medications are shown with only one drug class, but may actually belong to multiple classes.

	Drug Class	RxCUI	Generic name	Count	Class Count	Class %
1	Thiazide diuretics	5487	hydrochlorothiazide	24165	26766	90%
2	Thiazide diuretics	6916	metolazone	1680	26766	6%
3	Thiazide diuretics	5764	indapamide	620	26766	2%
4	Thiazide diuretics	2396	chlorothiazide	285	26766	1%
5	Thiazide diuretics	2409	chlorthalidone	155	26766	1%
6	Loop diuretics	4603	furosemide	35569	38773	92%
7	Loop diuretics	1808	bumetanide	1828	38773	5%
8	Loop diuretics	38413	torsemide	1692	38773	4%
9	Loop diuretics	62349	ethacrynic acid	90	38773	0%
10	Potassium-sparing diuretics	9997	spironolactone	9041	9688	93%
11	Potassium-sparing diuretics	10763	triamterene	471	9688	5%
12	Potassium-sparing diuretics	644	amiloride	192	9688	2%
13	Aldosterone-receptor blockers	298869	eplerenone	286	286	100%
14	Beta blockers (cardioselective)	6918	metoprolol	24716	37017	67%
15	Beta blockers (cardioselective)	1202	atenolol	11807	37017	32%
16	Beta blockers (cardioselective)	31555	nebivolol	206	37017	1%
17	Beta blockers (cardioselective)	19484	Bisoprolol	203	37017	1%
18	Beta blockers (cardioselective)	1520	Betaxolol	171	37017	0%
19	Beta blockers (cardioselective)	149	Acebutolol	109	37017	0%
20	Beta blockers (nonselective)	8787	propranolol	2822	7082	40%
21	Beta blockers (nonselective)	10600	timolol	1845	7082	26%
22	Beta blockers (nonselective)	9947	sotalol	1768	7082	25%
23	Beta blockers (nonselective)	7226	nadolol	609	7082	9%
24	Beta blockers (nonselective)	8332	pindolol	73	7082	1%
25	Alpha-1 blockers	49276	doxazosin	1747	3648	48%
26	Alpha-1 blockers	37798	terazosin	1678	3648	46%
27	Alpha-1 blockers	8629	prazosin	228	3648	6%
28	Alpha and beta blockers	20352	carvedilol	7249	8656	84%
29	Alpha and beta blockers	6185	labetalol	1433	8656	17%
30	Centrally acting agents	2599	clonidine	9149	9556	96%
31	Centrally acting agents	6876	methyldopa	442	9556	5%
32	Peripheral nerve-acting agents	9260	reserpine	23	26	88%
33	Peripheral nerve-acting agents	5036	guanethidine	3	26	12%
34	Direct-acting vasodilators	5470	hydralazine	2318	3954	59%
35	Direct-acting vasodilators	6984	minoxidil	1705	3954	43%

Table 20 (continued). List of antihypertensive medications, including drug class and prevalence, in the H&P test corpus of ~350,000 notes. Medications are shown with only one drug class, but may actually belong to multiple classes.

	Drug Class	RxCUI	Generic name	Count	Class Count	Class %
36	Calcium-channel blockers	17767	amlodipine	15014	33697	45%
37	Calcium-channel blockers	3443	diltiazem	9098	33697	27%
38	Calcium-channel blockers	7417	nifedipine	5311	33697	16%
39	Calcium-channel blockers	11170	verapamil	3848	33697	11%
40	Calcium-channel blockers	4316	felodipine	613	33697	2%
41	Calcium-channel blockers	7396	nicardipine	270	33697	1%
42	Calcium-channel blockers	7435	nisoldipine	258	33697	1%
43	Calcium-channel blockers	33910	isradipine	161	33697	0%
44	ACE inhibitors	29046	lisinopril	22597	40105	56%
45	ACE inhibitors	3827	enalapril	5980	40105	15%
46	ACE inhibitors	35296	ramipril	4369	40105	11%
47	ACE inhibitors	35208	quinapril	3139	40105	8%
48	ACE inhibitors	18867	benazepril	1801	40105	4%
49	ACE inhibitors	1998	captopril	1623	40105	4%
50	ACE inhibitors	50166	fosinopril	841	40105	2%
51	Angiotensin-receptor blockers	69749	valsartan	5432	15348	35%
52	Angiotensin-receptor blockers	52175	losartan	4759	15348	31%
53	Angiotensin-receptor blockers	83818	irbesartan	1885	15348	12%
54	Angiotensin-receptor blockers	321064	olmesartan	1525	15348	10%
55	Angiotensin-receptor blockers	73494	telmisartan	992	15348	6%
56	Angiotensin-receptor blockers	214354	candesartan	816	15348	5%
57	Angiotensin-receptor blockers	83515	eprosartan	35	15348	0%
58	Renin Inhibitors	325646	Aliskiren	69	69	100%
59	Endothelin receptor blockers	75207	Bosentan	232	232	100%

Individual Antihypertensive ADEs in DEB2

Of the 59 drugs determined to be "antihypertensives," DEB2 contained 56. The drugs *ethacrynic acid* (from the *loop diuretics*), *guanethidine* (from the *peripheral nerve-acting agents*), and *isradipine* (from the *calcium-channel blockers*) were not present in DEB2. Collectively, those three medications were only present in 254 notes out of the 350,000 notes in the H&P test corpus – significantly fewer than 1% of available notes.

For the remaining 56 antihypertensive drugs, DEB2 contained 708 total ADEs involving 118 distinct CMs. The four aforementioned vague CMs were removed from consideration. After requiring the relative risk score of each drug-ADE pair to be greater than 1.2, only 352 ADEs remained. Those 352 ADEs were made up of 55 distinct drugs and 80 distinct CMs (no DEB2 side effects of *Hydralazine* had a relative risk score in the H&P test corpus greater than 1.09 out of 2318 notes mentioning hydralazine).

Deriving Drug Class ADEs

As described in Methods, the project team attempted to algorithmically determine common drug-class side effects based on data in DEB2 and the relative risk scores of drug-ADE pairs from the H&P test corpus. Of the 16 drug classes determined to make up the antihypertensives, the algorithms identified 25 class ADEs (17 distinct CMs) from 13 classes. Table 21 lists the drug classes and their extrapolated side effects, as well as the justification of why each was included.

Table 21. List of extrapolated drug class side effects for inclusion in AAS.

Class	CM String	Inclusion Criteria
Thiazide diuretics	Hypokalemia	In DEB2 for 60% of class; RR of 2.12.
Thiazide diuretics	Exanthema	In DEB2 for 100% of class.
Thiazide diuretics	Muscle Cramp	In DEB2 for 100% of class; RR of 1.21.
Alpha-1 blockers	Cardiac Arrhythmia	In DEB2 for 67% of class; RR of 1.93.
Calcium-channel blockers	Headache	In DEB2 for 88% of class.
Endothelin receptor blockers	Headache	Only 1 drug in class; RR of 1.22.
Endothelin receptor blockers	Pruritus	Only 1 drug in class; RR of 1.46.
Endothelin receptor blockers	Cardiac Arrhythmia	Only 1 drug in class; RR of 1.57.
Endothelin receptor blockers	Flushing	Only 1 drug in class; RR of 11.52.
Beta blockers (cardioselective)	Bradycardia	In DEB2 for 100% of class; RR of 2.39.
Loop diuretics	Hypokalemia	In DEB2 for 50% of class; RR of 1.66.
Loop diuretics	Hypovolemia	In DEB2 for 50% of class; RR of 1.96.
ACE inhibitors	Headache	In DEB2 for 86% of class.
ACE inhibitors	Coughing	In DEB2 for 100% of class.
ACE inhibitors	Fever	In DEB2 for 86% of class.
Centrally acting agents	Headache	In DEB2 for 100% of class; RR of 1.3.
Centrally acting agents	Bradycardia	In DEB2 for 100% of class; RR of 1.86.
Centrally acting agents	Orthostatic Hypotension	In DEB2 for 100% of class; RR of 1.31.
Angiotensin-receptor blockers	Angioedema	In DEB2 for 71% of class; RR of 2.71.
Alpha and beta blocker	Bradycardia	In DEB2 for 100% of class; RR of 2.24.
Renin Inhibitors	Headache	Only 1 drug in class; RR of 1.52.
Renin Inhibitors	Back Pain	Only 1 drug in class; RR of 1.65.
Renin Inhibitors	Seizures	Only 1 drug in class; RR of 1.25.
Beta blockers (nonselective)	Bradycardia	In DEB2 for 100% of class; RR of 2.61.
Beta blockers (nonselective)	Lightheadedness	In DEB2 for 60% of class; RR of 1.76.

Deriving Antihypertensive-wide ADEs

The project team identified any ADEs common to all antihypertensive drugs using the methods described above. As with the class ADEs, the team required any antihypertensive-wide ADEs to have a relative risk score in the H&P test corpus greater than 1.2. The top ten candidates for all-antihypertensives ADEs are listed in Table 22. Since the team set the inclusion threshold at 35%, only the top four were included. The table lists two percentages – the first is the percentage of antihypertensive drugs in DEB2 where the ADE in question had a relative risk score greater than 1.2; the second is the percentage of drugs that mentioned the ADE in DEB2 regardless of relative risk score. The relative risk score included in the table is the risk of any drug in the class causing the ADE.

Table 22. Top ten ADEs of all antihypertensive medications, sorted by percentage of drugs with the ADE in DEB2 and a relative risk score > 1.2.

Percent of Drugs (with RR > 1.2)	Percent of All Drugs	CM CUI	ADE	Relative Risk	Included
66% (39/59)	81% (48/59)	C0012833	Dizziness	1.72	Yes
53% (31/59)	66% (39/59)	C0013404	Dyspnea	1.91	Yes
41% (24/59)	58% (34/59)	C0039070	Syncope	1.5	Yes
37% (22/59)	46% (27/59)	C0020649	Hypotension	1.74	Yes
34% (20/59)	36% (21/59)	C0428977	Bradycardia	2.55	No
34% (20/59)	54% (32/59)	C0015672	Fatigue	1.35	No
27% (16/59)	32% (19/59)	C0004093	Asthenia	1.65	No
19% (11/59)	71% (42/59)	C0018681	Headache	0.94	No
14% (8/59)	19% (11/59)	C0008031	Chest Pain	1.94	No
12% (7/59)	15% (9/59)	C0220870	Lightheadedness	1.69	No

The most obvious ADE of antihypertensive medications is *C0020649 – Hypotension*. Despite this, it was only found in DEB2 as an ADE for 46% of the medications. The project team determined that this was due to most of the sources either omitting it (since lowering blood pressure is the *intended* effect of the medication) or instead mentioning the symptoms of low blood pressure: *headache*, *dizziness*, *lightheadedness*, *syncope,* etc.

The inclusion of hypotension was a major motivating factor for the project team setting the initial inclusion threshold at 35%. That is, the project team concluded that an ADE was a side effect of all antihypertensives if the relative risk of a drug-ADE pair (in the

H&P test corpus) was greater than 1.2 for at least 35% of all antihypertensive drugs. As all antihypertensives can reduce blood pressure, all ADEs determined to be antihypertensive-wide class side effects were the result of excessively low blood pressure.

Compiling Preliminary AAS

The project team combined the extracted individual, class, and all-antihypertensives ADEs to create the preliminary AAS. The quasi-automated review process created a total of four antihypertensive-wide ADEs, 25 class ADEs, and 143 individual drug ADEs. When split into individual drug-CM pairs, the set contained 472 drug-ADE pairs consisting of 58 distinct medications and 80 distinct CMs.

Expert Curation of AAS

The expert pharmacist reviewers made a number of changes to the drug classes in the initial algorithmic set. The calcium-channel blockers (CCBs) drug class contains three distinct sub-classes: Dihydropyridines, Benzothiazepines, and Phenylalkylamines. At the suggestion of the pharmacist reviewers, the project team split the CCBs into two sub-classes for the purpose of class-based side effects: the Dihydropyridines and the non-Dihydropyridines. The pharmacist reviewers added an additional class for "all diuretic medications" due to their similar side effects. Similarly, they added a class for "alpha and beta blockers," to encompass the side effects those classes have in common.

The pharmacist reviewers removed the three single-medication drug classes from the initial set: the aldosterone-receptor blockers (containing the drug *eplerenone*), the renin inhibitors (containing the drug *aliskiren*), and the endothelin receptor blockers (containing the drug *bosentan*). Each of the drugs are rarely used as antihypertensives and were present in less than 1% of notes from the H&P test corpus. Bosentan is a treatment for pulmonary hypertension, a different kind of hypertension than systemic hypertension. Similarly, the pharmacists removed the beta-blockers sotalol and timolol from the set; sotalol is used to treat cardiac arrhythmias and timolol is primarily used as a treatment for glaucoma.

The reviewers removed *Dyspnea* as an ADE of all antihypertensive medications and added *Lightheadedness* as an ADE for all-antihypertensives. They also added 15 new class

ADEs, mostly by changing individual ADEs from the algorithmic set to class ADEs. A total of 33 out of 40 class ADEs (83%) were present as one or more individual ADEs in the initial preliminary set. The pharmacist reviewers reported that starting with the initial algorithmically created set made the process of identifying antihypertensive ADEs substantially easier than compiling a set from primary sources. Due to these changes, the number of individual ADEs dropped from 143 down to 52.

Final Version of AAS

The initial algorithmically derived AAS contained 472 drug-CM ADE pairs and the final AAS contained 496 pairs. The final set contains 65% of the pairs in the original set and consists of four all-antihypertensive ADEs, 40 class ADEs (from 13 drug classes), and 52 individual ADEs. The AAS, made up of 54 drugs and 61 distinct CMs, is shown in Table 23.

Table 23. The final version of the ADER Alerting Subset (AAS).

Class	Type	Drug	CM
All antihypertensives	ALL	-	Dizziness
All antihypertensives	ALL	-	Hypotension
All antihypertensives	ALL	-	Lightheadedness
All antihypertensives	ALL	-	Syncope
ACE inhibitors	CLASS	-	Acute Kidney Injury
ACE inhibitors	CLASS	-	Coughing
ACE inhibitors	CLASS	-	Hyperkalemia
ACE inhibitors	CLASS	-	Kidney Failure
ACE inhibitors	DRUG	Benazepril	Orthostatic Hypotension
ACE inhibitors	DRUG	Captopril	Cholestasis
ACE inhibitors	DRUG	Captopril	Pemphigus
ACE inhibitors	DRUG	Enalapril	Hyponatremia
ACE inhibitors	DRUG	Enalapril	Pancreatitis
ACE inhibitors	DRUG	Lisinopril	Pancreatitis
All diuretics	CLASS	-	Dehydration
All diuretics	CLASS	-	Frequent Urination
All diuretics	CLASS	-	Hypovolemia
All diuretics	CLASS	-	Muscle Cramp
All diuretics	CLASS	-	Orthostatic Hypotension
All diuretics	CLASS	-	Polyuria
Alpha and beta blocker	CLASS	-	Bradycardia
Alpha and beta blocker	DRUG	Labetalol	Headache
Alpha-1 blockers	CLASS	-	Orthostatic Hypotension
Alpha-1 blockers	DRUG	Doxazosin	Impaired Vision
Alpha-1 blockers	DRUG	Doxazosin	Myalgia

Table 23 (continued). The final version of the ADER Alerting Subset (AAS).

Class	Type	Drug	CM
Angiotensin-receptor blockers (ARB)	CLASS	-	Acute Kidney Injury
Angiotensin-receptor blockers (ARB)	CLASS	-	Angioedema
Angiotensin-receptor blockers (ARB)	CLASS	-	Hyperkalemia
Angiotensin-receptor blockers (ARB)	CLASS	-	Kidney Failure
Beta blockers (cardioselective)	CLASS	-	Bradycardia
Beta blockers (cardioselective)	DRUG	Atenolol	Confusion
Beta blockers (cardioselective)	DRUG	Atenolol	Headache
Beta blockers (cardioselective)	DRUG	Atenolol	Mental Depression
Beta blockers (cardioselective)	DRUG	Metoprolol	Cardiogenic Shock
Beta blockers (cardioselective)	DRUG	Metoprolol	Heart Block
Beta blockers (cardioselective)	DRUG	Metoprolol	Mental Depression
Beta blockers (nonselective)	CLASS	-	Bradycardia
Beta blockers (nonselective)	CLASS	-	Erectile dysfunction
Beta blockers (nonselective)	DRUG	Pindolol	Arthralgia
Beta blockers (nonselective)	DRUG	Pindolol	Heartburn
Beta blockers (nonselective)	DRUG	Propranolol	Hallucinations
Beta blockers (nonselective)	DRUG	Propranolol	Mental Depression
Beta blockers (nonselective)	DRUG	Propranolol	Raynaud Disease
Calcium-channel blockers (DHP)	CLASS	-	Constipation
Calcium-channel blockers (DHP)	CLASS	-	Edema
Calcium-channel blockers (DHP)	CLASS	-	Gingival Hyperplasia
Calcium-channel blockers (DHP)	CLASS	-	Peripheral edema
Calcium-channel blockers (non-DHP)	CLASS	-	Bradycardia
Calcium-channel blockers (non-DHP)	CLASS	-	Constipation
Calcium-channel blockers (non-DHP)	CLASS	-	Gingival Hyperplasia
Calcium-channel blockers (non-DHP)	CLASS	-	Headache
Calcium-channel blockers (non-DHP)	CLASS	-	Heart Block
Centrally acting agents	CLASS	-	Bradycardia
Centrally acting agents	CLASS	-	Headache
Centrally acting agents	CLASS	-	Orthostatic Hypotension
Centrally acting agents	DRUG	Clonidine	Confusion
Centrally acting agents	DRUG	Clonidine	Constipation
Centrally acting agents	DRUG	Clonidine	Erectile dysfunction
Centrally acting agents	DRUG	Clonidine	Mental Depression
Centrally acting agents	DRUG	Clonidine	Somnolence
Centrally acting agents	DRUG	Clonidine	Xerostomia
Centrally acting agents	DRUG	Methyldopa	Agranulocytosis
Direct-acting vasodilators	DRUG	Minoxidil	Edema
Direct-acting vasodilators	DRUG	Minoxidil	Headache
Direct-acting vasodilators	DRUG	Minoxidil	Hypertrichosis
Direct-acting vasodilators	DRUG	Minoxidil	Pericardial effusion
Direct-acting vasodilators	DRUG	Minoxidil	Recent weight gain
Direct-acting vasodilators	DRUG	Minoxidil	Vomiting

Table 23 (continued). The final version of the ADER Alerting Subset (AAS).

Class	Type	Drug	CM
Loop diuretics	CLASS	-	Hypokalemia
Loop diuretics	DRUG	Bumetanide	Exanthema
Loop diuretics	DRUG	Furosemide	Alkalosis
Loop diuretics	DRUG	Furosemide	Bullous pemphigoid
Loop diuretics	DRUG	Furosemide	Deafness
Loop diuretics	DRUG	Furosemide	hearing impairment
Loop diuretics	DRUG	Furosemide	Interstitial Nephritis
Loop diuretics	DRUG	Furosemide	Nephrocalcinosis
Loop diuretics	DRUG	Furosemide	Pancreatitis
Loop diuretics	DRUG	Furosemide	Tinnitus
Peripheral nerve-acting agents	DRUG	Reserpine	Anxiety
Peripheral nerve-acting agents	DRUG	Reserpine	Bradycardia
Peripheral nerve-acting agents	DRUG	Reserpine	Mental Depression
Peripheral nerve-acting agents	DRUG	Reserpine	Psychotic Disorders
Potassium-sparing diuretics	CLASS	-	Hyperkalemia
Potassium-sparing diuretics	DRUG	Spironolactone	Agranulocytosis
Potassium-sparing diuretics	DRUG	Spironolactone	Exanthema
Potassium-sparing diuretics	DRUG	Spironolactone	Gynecomastia
Potassium-sparing diuretics	DRUG	Triamterene	Kidney Calculi
Thiazide diuretics	CLASS	-	Anorexia
Thiazide diuretics	CLASS	-	Exanthema
Thiazide diuretics	CLASS	-	Gout
Thiazide diuretics	CLASS	-	Hyperuricemia
Thiazide diuretics	CLASS	-	Hypokalemia
Thiazide diuretics	CLASS	-	Hyponatremia
Thiazide diuretics	CLASS	-	Pancreatitis
Thiazide diuretics	DRUG	Hydrochlorothiazide	Hyperglycemia
Thiazide diuretics	DRUG	Metolazone	Pruritus

Discussion

The algorithmic methods used to identify antihypertensive ADEs provided a useful tool for compiling drug information for decision support systems. By providing an initial set for manual review, expert pharmacists were able to quickly and efficiently compile a list of antihypertensive ADEs for the ADER system. The project team and reviewing pharmacists considered both prevalence and severity, using clinical expertise to determine which ADEs would be most useful to detect and alert patients' physicians.

The AAS provides an initial test set for the ADER system. By using antihypertensive medications, the project team ensured a large number of patients on the target medications for the pilot implementation test. Additionally, antihypertensives have a wide variety of well-established side effects. By using this set of ADEs, the ADER pilot implementation can

prove that the systems methodology can detect ADEs for any number of medications and drug classes.

It is difficult to associate ADEs with specific medications in a systematic manner. Reported incidence of ADEs can vary widely. For example, the patient may not mention the symptoms due to perceived inconsequentiality, or the physician might not recognize the symptoms could be due to an ADE. Also, individual patients experience ADEs at a rate that is different from the general population. Suspected ADEs can also be the result of patients' various underlying medical conditions. When attempting to detect ADEs at admission, it is important to consider both severe ADEs that may result in admission as well as ADEs that may go unrecognized for a variety of reasons.

CHAPTER V

THE ADVERSE DRUG EFFECT RECOGNIZER (ADER)

Introduction

As discussed in Chapter II, unrecognized adverse drug effects (ADEs) pose serious clinical problems. They can cause preventable hospitalizations, increase healthcare costs, and worsen health outcomes.[33–39] Outpatient medications used to manage chronic conditions can cause many unrecognized ADEs. Upon admission to a medical facility, a detailed physical exam is documented in the admission H&P note. To potentially improve quality of care, the project team developed and evaluated the Adverse Drug Effect Recognizer (ADER), a novel system to detect adult inpatients' previously unrecognized symptomatic ADEs from H&P notes and alert appropriate care providers. Using automated natural language processing of clinician-generated electronic admission H&P notes, ADER identifies patients' current medications and clinical manifestations (CMs). The system then compares the medications and CMs against a set of known ADEs, generating appropriate patient-specific alerts in the Vanderbilt EMR system, StarPanel, and the Vanderbilt CPOE system, Horizon Expert Orders. Each alert identifies the offending medication, the suspected nature of the ADE, and evidence supporting the alert.

This implementation of the ADER system was designed for a pilot study targeting interns and residents in the Department of Internal Medicine at Vanderbilt University Hospital as subjects for decision support advice. The pilot study was limited to ADEs from antihypertensive medications, as described in Chapter IV. This chapter details the desiderata for a successful ADE detection system, the basic design of the ADER system, and the implementation-specific details required to interface ADER with the Vanderbilt clinical systems. Chapter VI details the results of the pilot study.

Materials

This work was performed on a MacBook Pro with a 2.6 GHz Intel Core i7 processor and 16 GB of RAM, a Linux server with forty-eight 2.2 GHz AMD Opteron cores and 256 GB RAM, referred to as the ADER server, and a database server with two 8-core Intel Xeon

processors and 6 GB of RAM called the database server. All data processing scripts were written in Perl. 5.10.0. The project used MySQL version 5.5.16 on the ADER server and MySQL version 5.0.95 on the database server. Derivation of the ADER AAS is described in Chapter IV.

Desiderata for an ADE Detection System

The project team designed the inpatient-based ADER system to detect symptomatic ADEs potentially due to pre-admission outpatient medications. Determining causality between a medication and a CM is difficult. The goal of ADER is to alert clinicians to potential ADEs, letting clinicians decide what action to take, if any. ADER compares patients' current medications and CMs against a database of known ADEs. These drug-CM pairs represent potential ADEs – CMs that may either directly result from the medication or stem from some other cause but that are potentially exacerbated by a current medication. The presence of the CM may also be coincidental, not directly related to a medication. This section enumerates the requirements for a successful ADE detection system.

Access to a Primary Source for Patient Information

The first requirement of a successful ADE detection system involves gaining reliable and timely information about each patient. When patients are admitted to a hospital, clinicians record a detailed patient history and perform a physical examination. The care team documents this information in the admission history and physical exam (H&P) note; the H&P contains several sections such as *Chief Complaint*, *History of Present Illness*, *Past Medical History*, *Medications*, and *Family Medical History*. Since the H&P is written at admission, any potential ADEs mentioned in the note would likely be due to outpatient medications instead of treatment received in the hospital.

Ability to Identify Current CMs

As stated above, the project required a rich source of information about each patient. The H&P note provides a rich source of not only the patient's current CMs, but also CMs recorded as being absent in the patient. The H&P note also contains sections listing the disease history of family members. A successful ADE detection system must be able to differentiate between those CMs present in the patient and those CMs mentioned as absent or present in others (i.e., *Family Medical History*). Other potentially relevant information in the H&P includes vital signs, such as blood pressure, body temperature, and heart rate, as well as abnormal laboratory test results that may or may not be mentioned in the text of the admission H&P note.

Ability to Identify Current Medications

A successful ADE detection system requires an accurate list of all the patient's pre-admission medications. The admission H&P note contains a curated list of current medications. Previous research at Vanderbilt has shown that the H&P is an accurate catalog of recent outpatient medications.[117] Often, the *Plan* section of the H&P lists medications that the care team has decided to use to treat the patient's current conditions. Since the goal of the project is to detect ADEs due to outpatient medications, it is important for the system to differentiate those medications currently being taken versus those medications that will be given post-admission in the course of treatment.

Ability to Identify Potential Confounders

As mentioned above, just because a patient is taking a given drug and experiencing a CM that *can* be an ADE of that drug, it does not prove that the drug is causing the CM. Some other medical condition might be the primary cause of the CM. For example, a patient taking the drug *Lisinopril* might be experiencing a cough, a known side effect. However, if the patient also suffers from *Chronic Obstructive Pulmonary Disease*, and the onset of the cough antedates use of the medication, it is more likely the disease is causing the cough rather than the drug. However, it is also possible that the medication might be exacerbating the symptom. The project team decided a successful ADE detection system should also

identify potential confounders – that is, diseases that may be causing suspected potential ADEs.

A Reliable and Readily Available Source of Known ADE Information

Alert fatigue, the phenomenon where users exposed to frequent borderline useful alerts become desensitized to them, adversely affects many clinical decision support systems.[134,135] To avoid this problem, a successful ADE detection system must be limited as best possible to true ADEs of clinical significance. The project team members believe that experts should prospectively and carefully review the catalog of ADEs that an alerting system attempts to detect. The ADER Alerting Subset (AAS), the set of known antihypertensive ADEs to be used by the system, is described in the previous chapter. Conversely, users must be able to provide feedback on the alerts that they receive. This feedback should be monitored to ensure that any ADE alerts that are bothersome, redundant, or incorrect can be revised and potentially removed from the system.

Non-Trivial, Non-Disruptive ADE Alerts

The project team, in consultation with practicing physicians, studied a variety of options for issuing alerts when ADEs are detected. Past research has shown that similar alerts for drug-drug interactions should be non-interruptive.[136] Due to concerns of alert fatigue, and the experimental nature of the project, the team decided to use non-interruptive alerts. A successful ADE detection system should present alerts to the care team at an appropriate time – that is, when the physicians are entering orders or documenting clinical notes for the patient in question, or at a time when the care team is discussing the patient's care, such as during rounds.

ADER Design Overview

Using automated natural language processing of clinician-generated electronic admission history and physical exam (H&P) notes, ADER detects potential ADEs in patients where a drug and related potential ADE CMs are present in the patient. The system then compares each patient's recognized medications and CMs against the aforementioned known list of ADEs, generating appropriate patient-specific alerts for potential adverse

effects related to the patient's current medications. Each alert will identify the offending medication, the suspected ADE, and any evidence supporting the assertion. Alerts also list potential confounding diseases that could be the primary cause of any suspected ADEs. The basic ADER design is illustrated in Figure 10. This section details the general design of the ADER system; the next section details the specific software design, including those design decisions affected by the pilot implementation design.

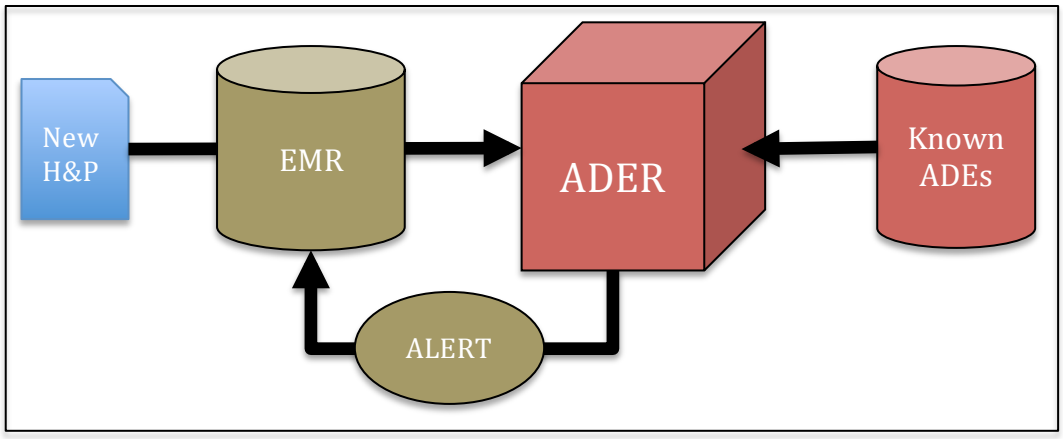


Figure 10. Basic ADER System Design.

Getting H&P notes and Labs

Portfolio is a system designed to store and evaluate clinical notes written by interns and residents at Vanderbilt University Medical Center. When new notes are authored in StarPanel, they are immediately copied to the Portfolio database server. Since Internal Medicine interns and residents are the target audience for the ADER pilot study, ADER retrieves new H&P notes from the Portfolio database server.

Identifying CMs

To identify CMs, ADER scans the EMR-based H&P documentation using NLP software, specifically KMCI, and CMs are represented using UMLS CUIs.[62] Similar to the methods used to create DEB2 in Chapter III, the project restricted the CMs identified by ADER to the SNOMED-CT vocabulary, represented by UMLS CUIs with any of the following semantic types:

- Anatomical Abnormality
- Injury or Poisoning
- Congenital Abnormality
- Finding
- Sign or Symptom
- Acquired Abnormality
- Clinical Attribute
- Disease or Syndrome
- Mental or Behavioral Dysfunction
- Neoplastic Process
- Pathologic Function

As mentioned above, the H&P contains some sections, such as *Plan* and *Family History* that are unlikely to mention potential ADEs; the ADER system used the SecTag NLP tool to ensure that candidate CMs were extracted only from relevant sections.

Identifying concepts using NLP is not perfect; there are often many false-positive matches. Since false-positive matches in a clinical decision support system would result in many unnecessary and inaccurate alerts, the team manually refined the text strings that could be used to detect relevant UMLS CM concepts to reduce those false positives. To do this for all possible CM terms from SNOMED-CT would be intractable; however, once a set of known ADEs was selected for the alerting system (described in Chapter IV), the project team improved accuracy by refining the text matching for those specific ADE CMs targeted by the system.

Identifying Lab Results

Based on the specific set of targeted ADEs in the AAS, the project team members determined which laboratory test results would be useful in identifying potential ADEs. After determining appropriate out-of-range values, the team mapped those laboratory values to their corresponding CMs. For example, any high serum potassium laboratory test results were mapped to the SNOMED-CT concept *Hyperkalemia (C0020461)*. The project team decided to independently determine out-of-range values for the laboratory tests and

not necessarily use the same range values used in the laboratory system. By using a wider range for "normal," the project ensured that false positive alerts due to borderline high or low lab results were eliminated. To determine those specific ranges, the project team analyzed ranges of laboratory results from de-identified patient data in the Vanderbilt Synthetic Derivative (SD).

Identifying Medications

The MedEx NLP tool can identify medications mentioned in plain-text clinical notes, including the drug name, dose, route, and form. To identify current medications, relevant medication sections in the H&P are first extracted using the SecTag NLP tool. Since the goal of ADER is to identify ADEs due to outpatient medications, the selected relevant sections are those that list outpatient medications, not those containing medications administered in the hospital or those planned to be administered to address current symptoms. Next, the extracted sections are scanned using MedEx. Identified medications are then normalized to their generic ingredient using RxNorm relationships, and represented using RxNorm RxCUIs.

Identifying Confounders

Some potential ADEs may not truly be ADEs; the CMs could instead be symptoms of the patient's other conditions. The project team members decided that, since the drug may still be exacerbating those symptoms, ADER should still generate alerts for suspected ADEs that could have another clinical cause. However, the project team determined that is important to notify the patient's care team about the possible non-medication causes of the symptoms, as well. To identify those potential clinical causes and inform the care teams, the project team analyzed a large corpus of ~350,000 historical H&P notes from the SD to identify symptoms that frequently co-occur with individual diseases, indicating the likely symptoms of a given disease. The analysis calculated the relative risk score of each disease causing each potential symptom to identify likely confounders. The ADER system includes these possible confounding co-morbid conditions in all ADE alerts where they are detected.

List of Target ADEs

Since the NLP components of ADER represent CMs using UMLS SNOMED-CT CUIs and drugs using RxNorm RxCUIs, the system also represents ADEs as RxCUI-CUI pairs (called drug-CM pairs). The specific set of ADEs used for alerts, the AAS, was described in Chapter IV. Due to system requirements, each class ADE in the AAS was represented within ADER as individual drug-CM pairs.

ADER Alerts

The project team decided to present ADER alerts at multiple points in the patient care workflow, both electronically and on paper. After a patient is admitted and the final version of the admission H&P note is saved, ADER processes the note to identify potential ADEs, as described above. If potential ADEs are detected, alerts are placed in three different locations within the Vanderbilt clinical systems.

First, ADER places alerts in the Team Summary, an information view provided in Vanderbilt's StarPanel EMR. The team summary is available to multiple care teams and includes a patient synopsis, problem lists, plans by system, recent laboratory test results, and patient code status. To display relevant information to different care teams, different Team Summary views exist for each type of care team, for example Cardiology, Trauma, Orthopedics, Pediatrics, Surgery, and General Medicine, among others. If ADER detects a potential ADE, it creates an alert panel detailing the relevant information in the team summary.

Second, ADER inserts a demarcated "alert box" in the patient's StarPanel progress note form. Resident and attending MDs create daily progress notes for each patient after admission. The electronic forms for creating the daily progress note in StarPanel contain data entry frames and snapshots of information from the team summary. In the case of a detected potential ADE, ADER places the same alert box used in the team summary in the progress note form and inserts an ADE alert warning at the top of the form.

Third, ADER inserts a text version of the alert in the Continuing Medications and Results (CMR) form, a patient-specific paper-based rounding report generated by the CPOE system; it lists information such as current medications, recent lab results, radiology reports, and known drug allergies. The project team included ADER alerts in the CMR so

care teams can discuss the potential ADEs during rounds and decide the appropriate course of action.

The primary design goal of both the electronic and paper-versions of the alerts was to include a succinct description of the detected potential ADEs. The design included the specific information from the H&P notes and labs that triggered the alerts, as well as potential confounding diseases that could be causing the symptoms in question. The electronic version contained links to open popup windows in the EMR with a more detailed description of the suspected ADEs and general information about ADER. Since the primary purpose of the pilot study was the evaluation of the ADER system, the electronic alerts also contained survey forms to ascertain the accuracy and usefulness of each alert. During the pilot, the project team redesigned these survey forms to improve response rate from the care teams (described below).

Specific Software Design

The ADER system comprises over 10,000 lines of code spread across multiple platforms and component scripts. The system retrieves notes through the Portfolio Database server using the ADER Database Monitor component. It continually checks the database for newly saved admission H&P and discharge notes. The other components are all installed on the ADER server. The ADER Processor component uses the KMCI, SecTag, and MedEx NLP tools to extract medications and CMs from appropriate sections in the notes. The ADER Detector component identifies potential ADEs and potential confounding diseases and sends ADE alerts to the StarPanel EMR. The ADER Status Monitor components ensure the system is running normally and allows the project team to monitor preliminary study results of the pilot implementation. The system stores results in the ADER MySQL database, installed on the ADER server. The component scripts are written in Perl; a Perl module, *ader.pm*, contains functions and procedures shared among the components. The ADER post-alert survey component, a Perl CGI script, processes the results of survey questionnaires and stores the results in the ADER database. All data are stored behind the Vanderbilt University Medical Center firewall. Any data sent between servers is transferred using encrypted secure copy (SCP).

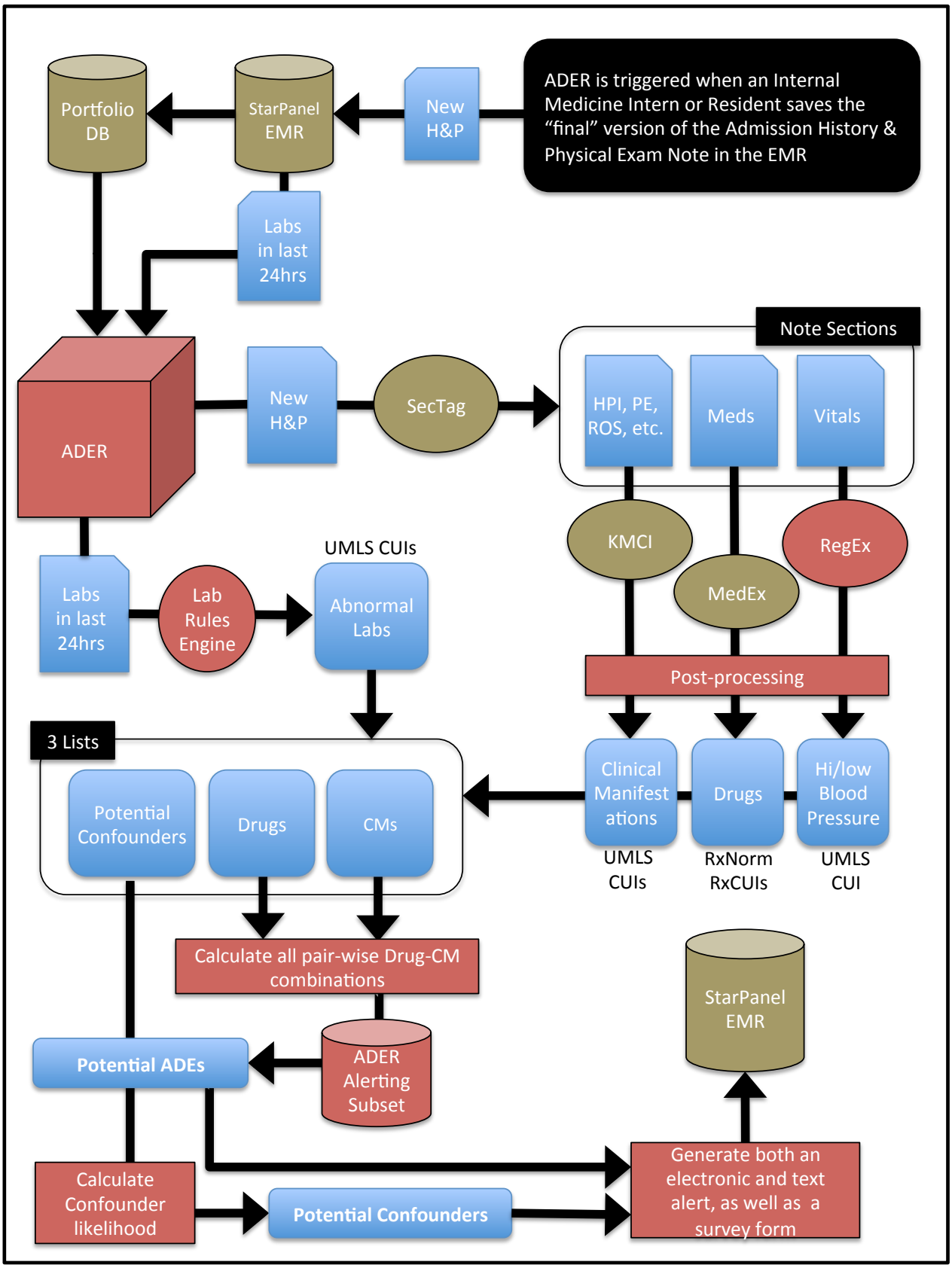


Figure 11. ADER workflow.

An illustration representing the ADER system workflow is shown in Figure 11. The complete system software design, including accessing the H&P notes, detecting potential ADEs, issuing alerts, and capturing the electronic responses, is described below. This section details the design of the software, alerts, and survey forms, as well as the survey form redesign performed during the early ADER pilot study. Chapter VI describes the results of the pilot study.

Integration into Vanderbilt Systems

Since the pilot study targets admission notes authored by Internal Medicine residents and interns, ADER uses the Portfolio database server as its source for H&P notes. The system monitors the database and retrieves any new H&Ps saved to the database. After retrieving the note, ADER queries a StarPanel API to retrieves any laboratory test results for the patient in the 24 hours prior to admission. Since treatment usually begins at admission, the results of laboratory tests performed shortly before admission can be used to identify ADEs due to outpatient medication.

To interface ADER with StarPanel, the project team members worked closely with members of the Vanderbilt University Medical Center (VUMC) Informatics Center, the organization responsible for maintaining the information systems used throughout the medical center. Informatics Center personnel modified the CMR paper form, the Team Summary, the Progress Note form, and the StarPanel API to work with the ADER system.

Transmitting Alerts to the EMR System

The system transmits the alert information to StarPanel using HTTPS. The system queries the StarPanel API for a specific server address and then sends an HTTPS Post containing the patient MRN, the CMR alert text, the StarPanel alert HTML, and an "action" field. The action field contains two possible commands: *update* and *delete.* When the update command is used, StarPanel either saves the new ADER alert variables into the patient records or overwrites any previous ADER alert variables. The *delete* command is used to remove the ADER alert variables from the patient record after the survey has been answered, the patient has been discharged, or the alert has expired. The ADER pilot was IRB-approved as an experimental study (Appendix A). Since experimental ADER alerts are

not part of normal clinical activity and not necessarily accurate, they are not considered clinical data. Therefore, ADER alerts are not permanently recorded in StarPanel.

If the transmission to StarPanel is successful, the ADER alert variables are stored in the patient records. The alert is then visible from the Team Summary or Progress Note form. The CPOE system retrieves the CMR alert text from StarPanel when the CMR is accessed. All alerts are recorded in the ADER database and system log (external to StarPanel). If for some reason the transmission is not successful, ADER stores the alert information in separate error log so that it can be transmitted at a later time. Information about all processed H&P notes, including any CMs, medications, and potential ADEs found by the system, are saved in both the ADER database and an HTML output log for analysis after the pilot implementation.

ADER Database Monitor Component

The ADER Database Monitor component script monitors the Portfolio database for the creation of admission or discharge notes. It automatically retrieves these entries from the Portfolio database upon arrival. First, ADER saves the relevant metadata regarding the admission from the database to a text file. The content of the metadata files include: a unique patient identification number; a unique note identification number; a medical record number (MRN), used internally by Vanderbilt clinical systems; the date and time the note was created; and whether or not the note amends a previous note. The system then uses the note creation date to find the patient's recent laboratory tests and to calculate the expiration date of the alert; ADER cancels any alerts not acknowledged within five days.

The metadata also contains the file path to the H&P note in a StarPanel-specific XML format. The system parses the XML using regular expressions and saves the H&P as a plain text file.

Finally, the ADER Database Monitor component retrieves the patient's recent laboratory test results. The system queries a StarPanel API to retrieve any laboratory test results for the patient in the 24 hours prior to admission. The system saves the laboratory test results to a plain text file, as shown below (Figure 12).

Date	Time	Lab	Value	Units	Normal Range
2015/09/15	16:34	PCV	37	%	36-43
2015/09/15	16:34	Ca	9.8	mg/dL	8.4-10.5
2015/09/15	16:34	Gluc	120	mg/dL	70-99
2015/09/15	16:34	WBC	13.0	x10(3)/mcL	3.9-10.7
2015/09/15	16:34	Plt-Ct	218	x10(3)/mcL	135-371
2015/09/15	16:34	Cl	95	mmol/L	98-107
2015/09/15	16:34	BUN	16	mg/dL	8-26
2015/09/15	16:34	Na	132	mmol/L	136-144
2015/09/15	16:34	K	3.9	mmol/L	3.3-4.8
2015/09/15	16:34	MCV	96	fL	81-98
2015/09/15	16:34	Creat	4.56	mg/dL	0.57-1.11
2015/09/15	16:34	CO2	26	mmol/L	23-31
2015/09/15	16:34	AlkP	142	unit/L	40-150

Figure 12. Sample de-identified pre-admission lab data from StarPanel API.

After the H&P, metadata, and lab files are saved, the system transmits them, using an encrypted secure copy (SCP), to the server running the other ADER software components.

Similarly, the system monitors the Portfolio database for discharge notes. Using a record of the admission H&P notes transmitted to ADER, the system monitors for the creation of corresponding discharge notes for the same patients. Once discharge notes are found, the system transmits the full text of the discharge note, the extracted medications section from the note, and the metadata related to the discharge note to the ADER server.

ADER Processor Component

The ADER Processor is a component Perl script installed on the ADER server. When the Database Monitor transmits new admission H&P notes, the ADER Processor script first confirms the three required files have been successfully received. If the note is an original admission H&P (or an amended note created less than 12 hours after the original), the system moves the files to appropriate directories and launches the NLP extraction tools used by ADER. First, the system runs KMCI and SecTag to tag the relevant sections of the H&P and index the UMLS concepts mentioned in the note. Next, the system moves the NLP results files to appropriate directories for processing. The system extracts the medication sections from the SecTag output files to use as input for MedEx. The system then runs the MedEx tool and, once complete, moves the output files to the appropriate directory for processing.

When amended H&P notes are received, the system processes the note as above only if the original H&P was created less than 12 hours earlier; older amended notes are discarded. After processing an amended H&P, the system updates the ADER database entry to link it to the new amended note. Amended H&P notes only add information to what is already present in the original; therefore, the new note can necessarily only contain more potential ADEs, not fewer. If new ADEs are found in the amended note, a new combined ADER alert will be sent to StarPanel to replace the previous one.

When new *discharge* notes are received for processing, the ADER Processor first confirms the three required files have been successfully transmitted to the server. The system then moves the files to the appropriate directories. The full discharge note is moved to a directory for post-implementation analysis to determine changes to suspected ADE-causing medications. The discharge metadata file is used to link the discharge note to the admission note in the ADER database. The system also runs the MedEx tool on the discharge medications file, saving the output for the post-implementation analysis.

When a new discharge note is received for processing, the system also sends a message to the StarPanel API to delete and thus inactivate the ADER alert for the patient. The system then updates the entry in the ADER *notes* database table to indicate that the patient has been discharged and the alert has been successfully deleted from StarPanel.

The final function of the ADER Processor is to delete expired alerts from StarPanel. As stated earlier, ADER alerts expire if they go unanswered for five days. The system monitors the ADER database and, if an alert is still active and has not yet been answered when the expiration time is reached, the system sends a delete message to StarPanel and updates the database. As with the other component scripts, the ADER Processor creates regular status messages for the ADER Status Monitor to ensure the system is functioning normally.

ADER Detector Component

The ADER Detector is the primary alerting component of the ADER system. The component script analyzes the NLP results from the ADER Processor, identifies any potential ADEs, and formats and transmits any necessary alerts to StarPanel.

Once KMCI, SecTag, and MedEx have completed processing a new H&P note, the detector component first scans the SecTag and KMCI results, constructing the *CM list* from the relevant affirmatively mentioned CMs in the note. The system checks to see if each CUI is negated using KMCI's implementation of the Negex[137] algorithm; examples of potential negation include, "*no* fever," or "The patient *does not have* a headache." Next, the system checks that the CUI has a CM semantic type and appears in a relevant section of the H&P; excluded sections include *Family History, Plan, Medications, Lab Results,* and *Vaccination* sections. Finally, the system checks the actual text string found by KMCI in the note and runs it though a regular expression filter designed to improve accuracy for AAS CMs. If the CM is not negated, from a relevant section, and the specialized AAS text string filter determines the match is accurate, the CM is added to the case's CM list.

Next, the system loads the SecTag results file and extracts the *Vitals* section. Using regular expressions, the system identifies the blood pressure measurement to determine if the patient has hypotension, a potential ADE of all antihypertensive medications. The patient is considered *hypertensive* if the systolic pressure is greater than 150 and the diastolic pressure is greater than or equal to 95. If the systolic pressure is less than or equal to 90, the patient is considered *hypotensive* and the appropriate CM CUI is added to the CM list.

The last element of the CM list derives from abnormal laboratory results. To identify such results, the system loads the lab results file that was sent to the server by the ADER Database Monitor. Using rules created by the project team, the system maps a specific set of potential abnormal lab values to the appropriate AAS CMs and adds them to the CM list. The rules use a wider, more liberal normal range than typically used by laboratory systems to ensure that borderline lab results do not generate bothersome alerts. The laboratory rules are shown in Table 24.

Table 24. Abnormal laboratory results rules.

Lab Test	Code	Result	Value	CM CUI	CM String
Packed Cell Volume	PCV	Low	< 36	C0002871	Anemia
Hemoglobin	HGB	Low	< 12	C0002871	Anemia
Serum Potassium	K	Low	< 3.2	C0020621	Hypokalemia
Serum Potassium	K	High	> 5.5	C0020461	Hyperkalemia
Serum Sodium	Na	Low	< 132	C0020625	Hyponatremia
Creatinine	Creat	High	> 2	C0035078	Renal insufficiency
Blood Urea Nitrogen	BUN	High	> 40	C0035078	Renal insufficiency
Glucose	Gluc	High	> 200	C0020456	Hyperglycemia
Total CO2 (HCO3)	CO2	High	> 28	C0002063	Alkalosis
pH	pH	High	> 7.5	C0002063	Alkalosis
Uric Acid	UricA	High	> 8	C0740394	Hyperuricemia

Next, the ADER Detector component creates the *Confounder List* – a list of potentially confounding diseases that might also cause the patient's ADEs. All potential confounders come from a set of approximately 500 disease concepts that match the above criteria and commonly cause some ADE CMs. The project team constructed this list from the H&P test corpus and created a regular expression string-matching filter similar to what was done for the AAS. Similar to creating the CM list, the system checks each CUI for negation, H&P section, semantic type, and that the matching text string is accurate. The required semantic types are a subset of CM semantic types and are listed in Table 25. The required sections for potential confounders are *Chief Complaint, History of Present Illness, Past Medical History,* or *Problem List.*

Table 25. Semantic Types for potential confounders.

Anatomical Abnormality	Injury or Poisoning
Acquired Abnormality	Mental or Behavioral Dysfunction
Congenital Abnormality	Neoplastic Process
Disease or Syndrome	Pathologic Function

Next, the ADER Detector component creates the *Medication List* – a list of all the patient's current medications in the form of RxNorm ingredient RxCUIs. For each clinical drug concept (see Chapter III) identified by MedEx, ADER uses RxNorm relationships to map the drug to its generic ingredient concept; multiple-ingredient drugs were not

considered in the ADER pilot study. The system adds all generic ingredient concepts with signature information (strength, dose, route, etc.) to the Medication List in the form of RxNorm RxCUIs.

Using the complete Medication List and CM List, the system can check for potential ADEs. All combinations of medications and CMs in the form of drug-CM pairs are checked against the AAS. Any drug-CM pairs that match those in the AAS are considered potential ADEs. To identify potential confounders – that is, diseases or syndromes that may be causing or contributing to the CM – the system checks the potential confounding diseases in the Confounder List.

To identify whether the drug or a possible confounder are the more likely cause of the ADE, the project team members pre-computed the relative risk score of any confounder causing the CM in question. The project team members calculated the relative risk scores for the drug causing the CM or the confounder causing the CM from the H&P test corpus (described Chapter IV). If any of the potential confounders are more likely to be causing the potential ADE than the suspected drug, that information is reported in the alert along with the ADEs. That is, if the relative risk score for the confounder causing the CM was at least 1.5 and it was greater than the relative risk of the drug causing the CM, then the potential confounder was considered a *likely* confounder. The most likely confounder for each ADE, that with the highest relative risk score, is included with the alert.

At this point, the ADER detector component has a list of potential ADEs, as well as a list of likely confounders. If potential ADEs were identified, the system generates and transmits the alert to StarPanel. First, the system generates the alert text for the CMR. Figure 13 shows an example of the CMR alert text. Each potential ADE is shown on a separate line along with the drug or drugs suspected as causes. If the ADE string is not an exact match to the wording found in the H&P, the alert also lists the string found in the note. If an ADE warning was derived from laboratory results, the pertinent lab results are listed; if the ADE is *Hypotension* found in the vitals section, the blood pressure reading is listed.

```
2015/09/23: ALERT!! Antihypertensive Adverse Drug Effect Recognizer (ADEr)
Based on the H&P and recent labs, the patient may be experiencing the following side effects:
 == SPIRONOLACTONE --> HYPERKALEMIA (K=6.2)
 == HYDROCHLOROTHIAZIDE --> HYPONATREMIA (Na=127)
 == HYDROCHLOROTHIAZIDE --> ANOREXIA (loss appetite)
 == FUROSEMIDE, HYDROCHLOROTHIAZIDE, SPIRONOLACTONE --> HYPOTENSION  (BP:88/63)
See the TeamSummary in StarPanel for more information. Evaluate the situation and consider
changing medications if indicated.
```

```
2015/09/20: ALERT!! Antihypertensive Adverse Drug Effect Recognizer (ADEr)
Based on the H&P, the patient may be experiencing the following side effects:
== FUROSEMIDE, LOSARTAN --> DIZZINESS
== FUROSEMIDE, LOSARTAN --> LIGHTHEADEDNESS
See the TeamSummary in StarPanel for more information. Evaluate the situation and consider
changing medications if indicated.
```

Figure 13. Examples CMR alert text.

Next, the system generates the HTML for the StarPanel alert. Figure 14 shows examples of the original format StarPanel alert for the Team Summary and the Progress Note. The StarPanel alert and survey questionnaire were redesigned during the pilot evaluation; the changes that led to the final format alert are described below. As with the team summary, the ADEs are listed at the top of the alert. The original format StarPanel alert also lists any potential confounders as shown in last example in Figure 14.

The last line of the original format StarPanel alert contained a link to acknowledge the alert. When a user clicked the link, it opened an HTML popup window containing an information panel with a full description of the alert, as well as an ADER evaluation survey questionnaire about the alert. An example of the popup illustrating the information panel and the survey questions are shown in Figure 15. Each popup also listed contact information for the project team in case of any problems.

There were five types of questions in the original format of the post-alert survey questionnaire. The first two questions concerned the accuracy of the NLP. The user was to confirm the patient was experiencing the CM and taking the medication for each suspected drug and potential ADE. The next question asked, for each ADE, if the survey respondent knew that the drug could cause the ADE in question. The next question asked whether the physician believed the drug was causing the ADE in this instance. The final question asked what course of action the care team planned to take. There was also an optional field for comments. After the HTML for the complete alert description and survey form was generated, the system saved the HTML file in the web server directory on the ADER server. As stated above, a link to this popup was included with the alert sent to StarPanel.

Possible Side Effects (from ADEr)

ALERT: The experimental ADEr system detected potential ADEs for this patient!

2015/07/21: Based on the H&P, the patient may be experiencing the following side effect:

- **Carvedilol, Lisinopril, Spironolactone, Triamterene --> Hypotension** (BP:86/55)

Consider also that the patient's underlying medical conditions may cause some of these symptoms. Alerts should in no way substitute for the patient care team's clinical judgment clinical judgment.

Click here for more information and to acknowledge this alert.

Possible Side Effects (from ADEr)

ALERT: The experimental ADEr system detected potential ADEs for this patient!

2015/07/23: Based on the H&P and recent labs, the patient may be experiencing the following side effect:

- **Lisinopril ----> Hyperkalemia** (K=6.4)

Consider also that the patient's underlying medical conditions may cause some of these symptoms. Alerts should in no way substitute for the patient care team's clinical judgment clinical judgment.

Click here for more information and to acknowledge this alert.

Possible Side Effects (from ADEr)

ALERT: The experimental ADEr system detected potential ADEs for this patient!

2015/07/22: Based on the H&P, the patient may be experiencing the following side effect:

- **Furosemide ----> Muscle Cramp** (associated severe r thigh cramps)

Consider also that the patient's underlying medical conditions may cause some of these symptoms. Alerts should in no way substitute for the patient care team's clinical judgment clinical judgment.

Click here for more information and to acknowledge this alert.

Possible Side Effects (from ADEr)

ALERT: The experimental ADEr system detected potential ADEs for this patient!

2015/07/21: Based on the H&P, the patient may be experiencing the following side effect:

- **Lisinopril ----> Renal insufficiency** (Creat=6.38)
- **Lisinopril ----> Coughing**
- **Amlodipine ----> Edema**

Consider also that the patient's underlying medical conditions (e.g., chronic kidney failure, COPD, and heart failure) may cause some of these symptoms. Alerts should in no way substitute for the patient care team's clinical judgment clinical judgment.

Click here for more information and to acknowledge this alert.

Figure 14. Examples of the original format StarPanel alert.

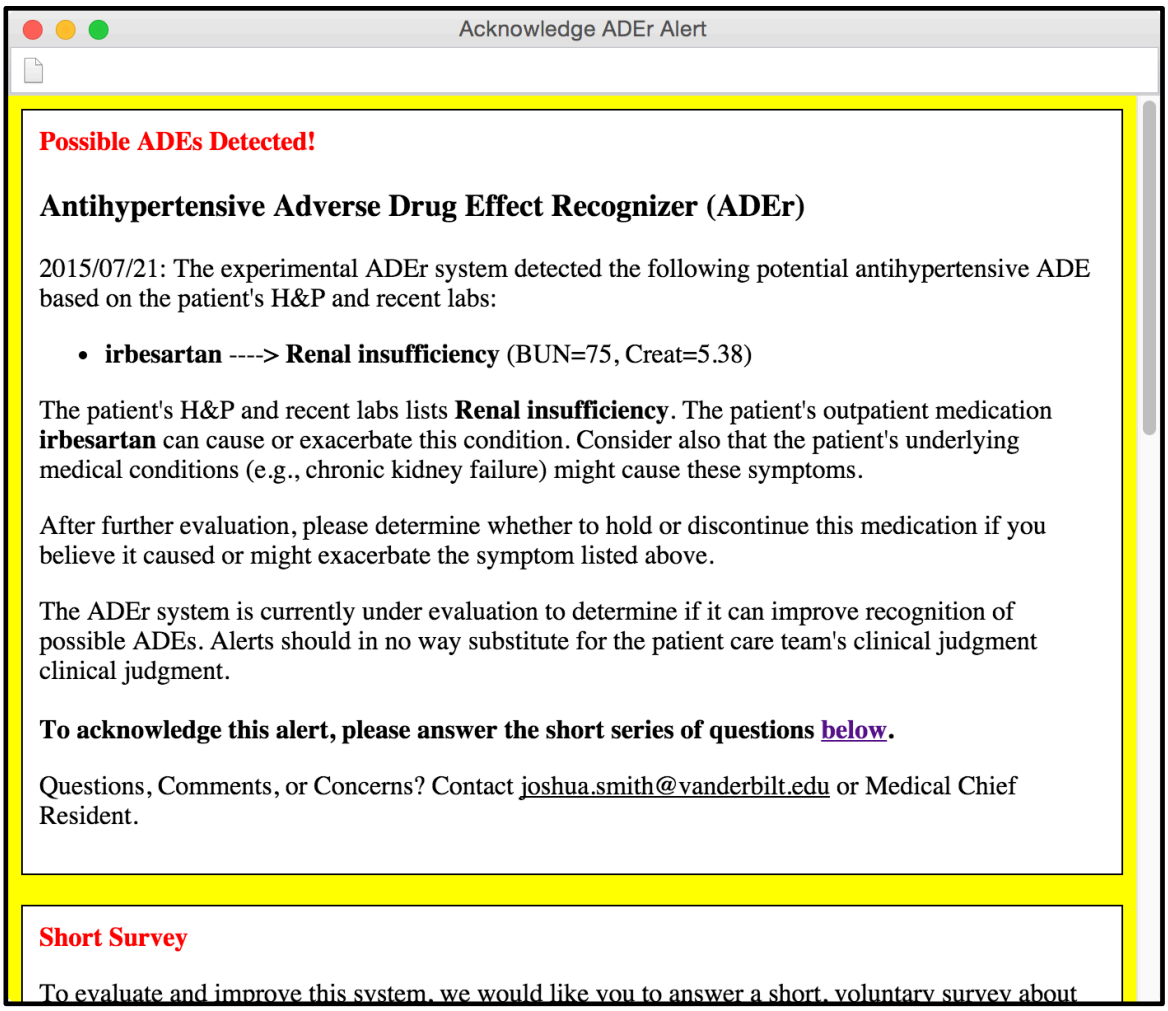


Figure 15 (a). Original format alert popup, showing the complete alert description.

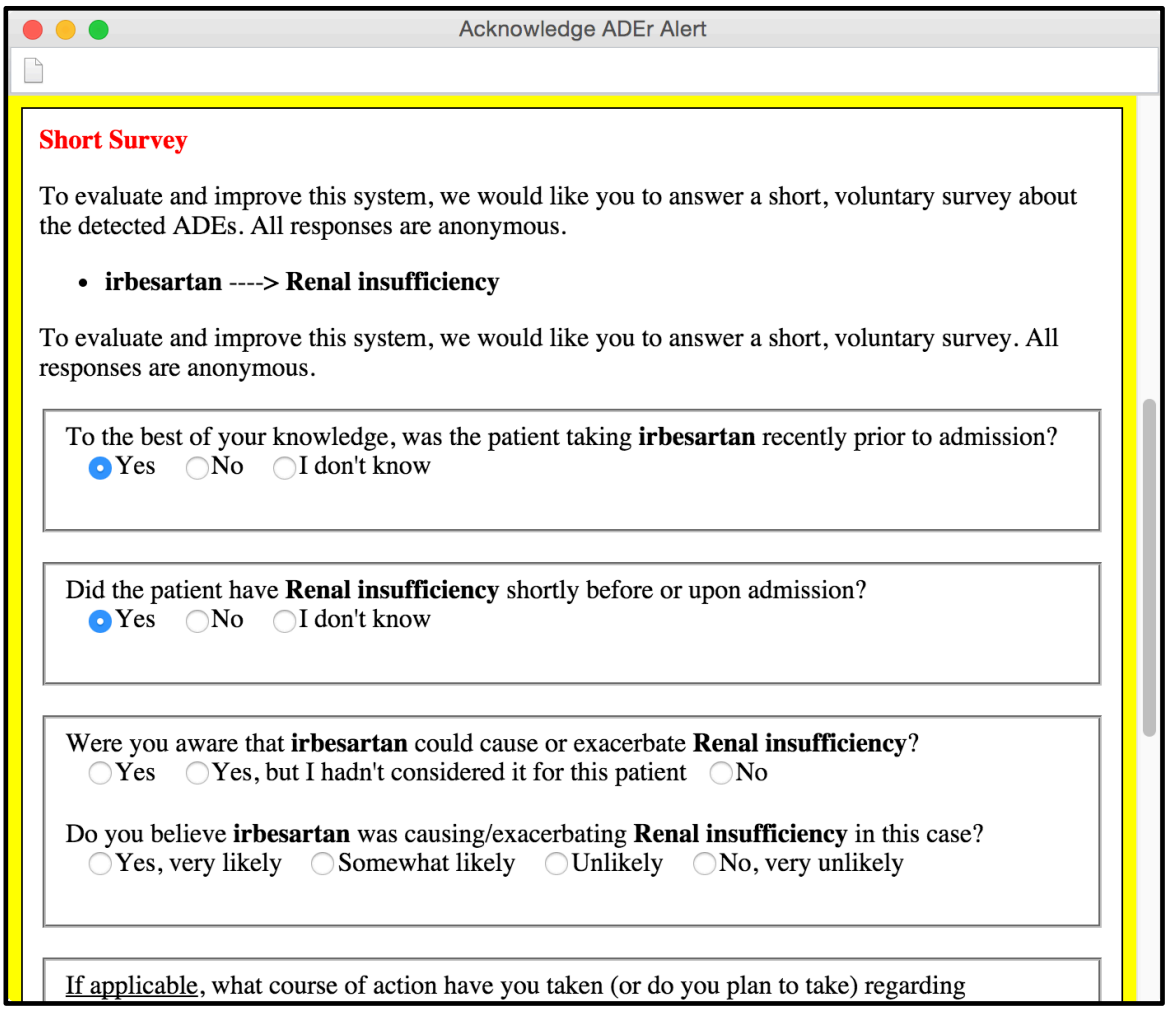


Figure 15 (b). Original format alert popup (continued), showing the first half of the survey.

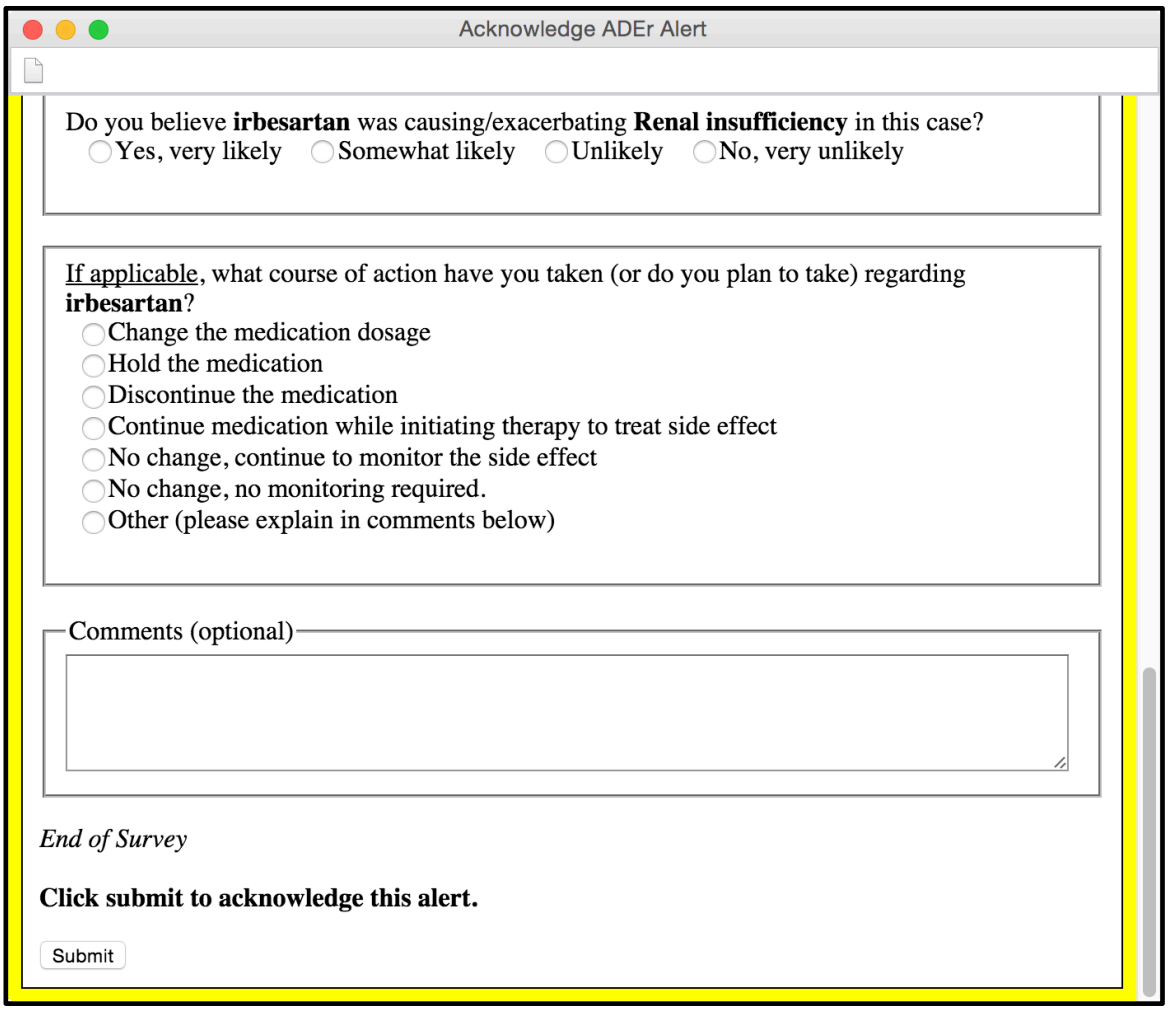


Figure 15 (c). Original format alert popup (continued), showing the second half of the survey.

The number of questions in this original format grew with the number of ADEs detected in the note. For alerts with more than one or two ADEs, the number of questions proved unwieldy. The number of questions, along with having to click the link to see the questions, resulted in very few end-user survey responses. The project team recognized this problem during the early days of the live pilot study, and, in consultation with the Internal Medicine Chief Resident, redesigned the questionnaire to have fewer questions. The redesigned StarPanel alert and survey questionnaire are shown in Figure 16. The new alerts presented the information in a more compact form. The redesigned survey comprised three short multiple-choice questions and an optional comments field. The questions were as follows:

1. *Did any of these alerts merit at least passing consideration for this patient?* This question assessed the validity of the alert, if it was appropriate, and if the care team thought it was potentially relevant clinically for the patient in question.
2. *Were any of these alerts helpful in managing this patient?* This question assessed whether the care team actively considered the alert in determining patient care decisions.
3. *Will you (or did you already) change patient's therapy due to these possible ADEs?* This question assessed whether the care team actually changed the patient's therapy based on any of the included ADEs, even if they had already made the changes before seeing the alert. This included, but was not limited to, medication holds, dosage changes, discontinuing medications, adding new medications to treat side effects, etc.

As with the original survey form, there is optional field for comments. The redesigned alert has a link to a popup of the detailed description of the alert, similar to the information panel in the original alert. The complete description popup is saved as an HTML file in the web server directory; it is shown in Figure 16.

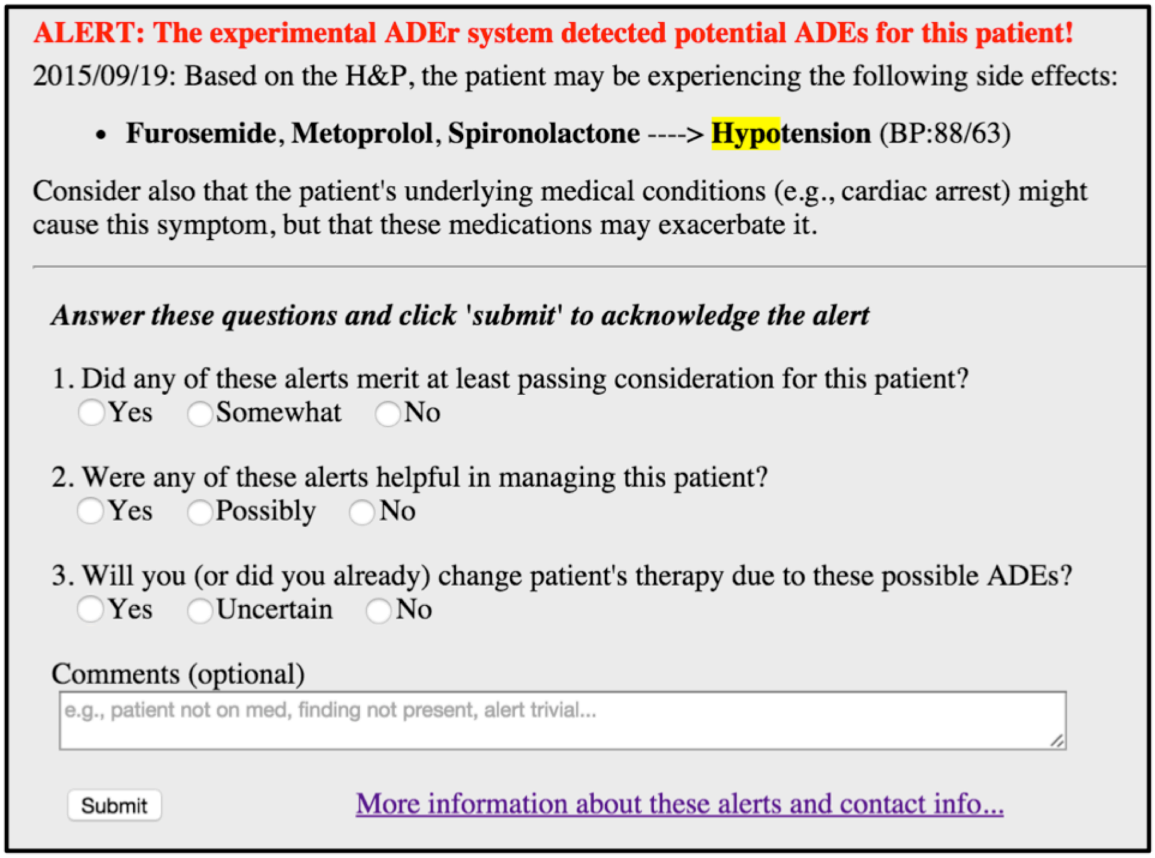


Figure 16. Redesigned alert, with the survey form included in the alert panel.

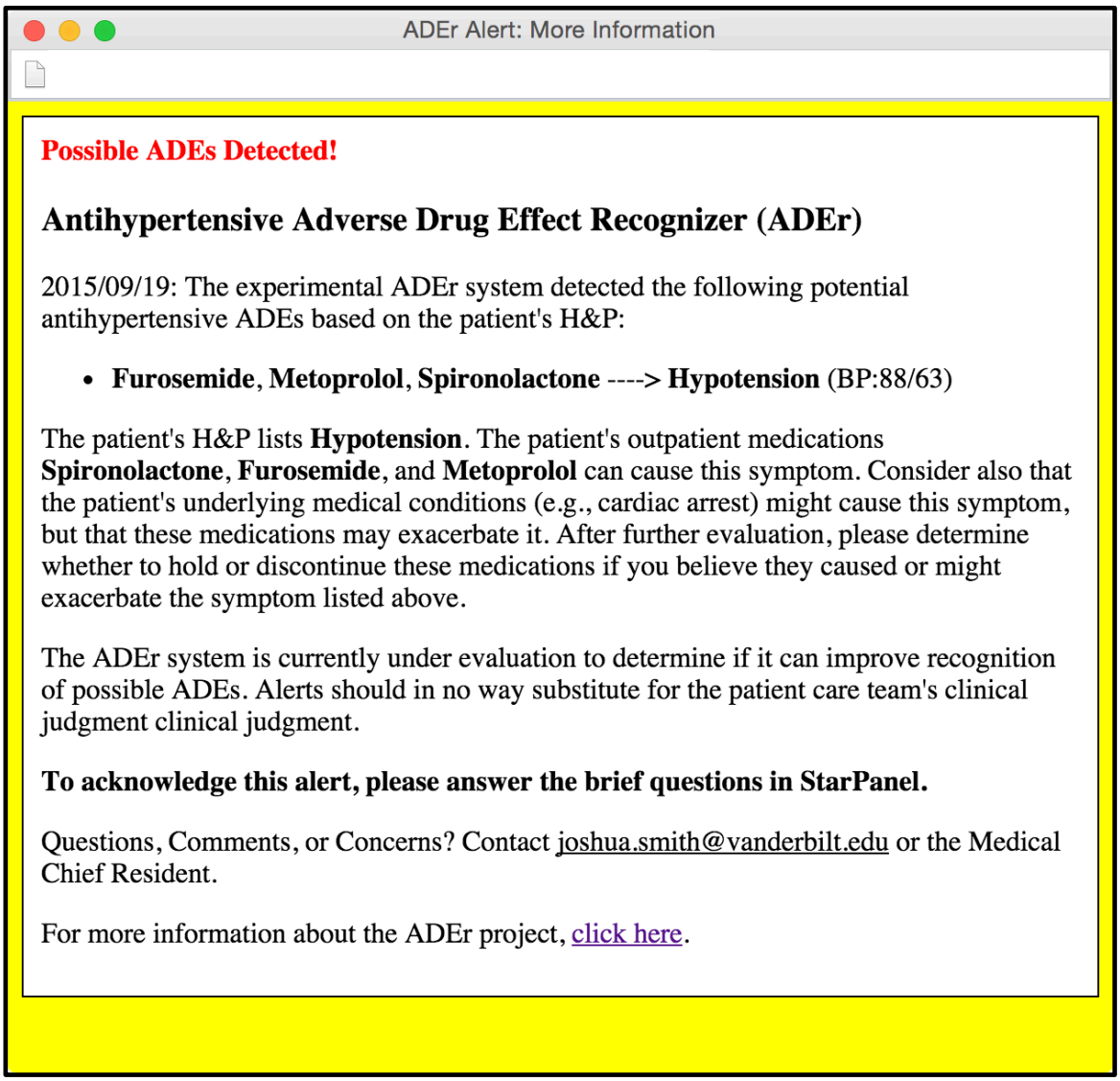


Figure 17. Redesigned "complete description popup" giving full information on the ADE in question, as well as information on the ADER system.

Information about all processed H&P notes, including any CMs, medications, and potential ADEs found by the system, was saved in both the ADER database and an HTML output log for analysis after the pilot implementation. When the system finished processing the H&P note, all input files were moved to a directory for completed notes and the system resumed monitoring for new input from the ADER processor. As with the other components, the ADER Detector regularly updated its status for the ADER Status Monitor.

ADER Survey Processor

The ADER Survey Processor component processed and saved the survey responses from the ADER alerts in StarPanel using a Perl CGI script. Responses were sent from StarPanel using a secure HTML Post to the CGI script on the ADER server. The system saved the responses to the ADER database and, after receiving the response, transmitted the *delete* message to StarPanel, removing the alert from the patient record. If for any reason the database was down, the script saved the responses to an error log so they could be saved to the database at a later time.

ADER Status Monitor

The ADER Status Monitor component continually monitors the status updates from each of the ADER components. If the system was functioning normally, the status monitor sent regular emails to the project team every twelve hours with information such as system uptime, the time of the last H&P notes transmitted by the ADER Database Monitor, the number of alerts and detected ADEs, and the current number of survey responses. If any system component failed, the status monitor transmitted email messages to the project team. The status emails contained information on which component failed, the time of the failure, and other information included in the regular status updates. Figure 18 illustrates sample status update email messages from the system. After receiving a failure status message, the project team could repair or restart the ADER system within the hour. During any downtime, the system did not miss any new H&P notes; as they are saved in the Portfolio database, project team members simply reset the system and indexed it to the time of failure. This resulted in any alerts being delayed by only a few minutes.

```
Subject: ADER: STATUS:OK


CURRENT TIME:          Wed Oct 21 08:05:33 CDT 2015  Time    Status
-------------------------------------------------------------------
ader_db_monitor        Wed Oct 21 08:02:55 CDT 2015    2     OK
ader_processor         Wed Oct 21 08:04:03 CDT 2015    1     OK
ader_main              Wed Oct 21 08:04:46 CDT 2015    0     OK


ader_db_monitor:       LastSendTime:2015-10-21 06:03:10

ader_main:
notes processed:       3368
alerts:                1028
ADE Rate:              0.305
ades found:            2449
ades per ADE_note:     2.382
lab ades found:        402
text ades found:       2047

last note saved to log: 2015-10-21 06:08:15
surveys answered:      295
active alerts:         29

Status: OK Everything seems to be running.

EOM
```

```
Subject: ADER: STATUS:ERROR


1/3
CURRENT TIME:          Wed Oct 21 15:25:33 CDT 2015  Time    Status
-------------------------------------------------------------------
ader_db_monitor        Wed Oct 21 15:01:30 CDT 2015    24    DOWN!
ader_processor         Wed Oct 21 15:24:23 CDT 2015    1     OK
ader_main              Wed Oct 21 15:24:45 CDT 2015    0     OK


ader_db_monitor:       LastSendTime:2015-10-21 14:40:31

ader_main:
notes processed:       3374
alerts:                1030
ADE Rate:              0.305
ades found:            2454
ades per ADE_note:     2.382
lab ades found:        402
text ades found:       2052

last note saved to log: 2015-10-21 14:34:32
surveys answered:      297
active alerts:         25

These scripts don't seem to be running: ader_db_monitor

EOM
```

Figure 18. Sample ADER status and error email messages.

The monitoring system component also contained a number of system-generated HTML pages for project team members to use for later analysis. This included the HTML

log, showing all past activity by the ADER system, and the Status Monitor & Control Panel page, showing the current status of the system and the survey results.

The ADER HTML log showed each alert by ID and H&P submit time. The detected CMs, drugs, laboratory CMs, possible ADEs, and possible confounders, as well as the generated CMR alert text and StarPanel alert HTML, are listed for each log entry. Sample log entries are shown in Figure 19.

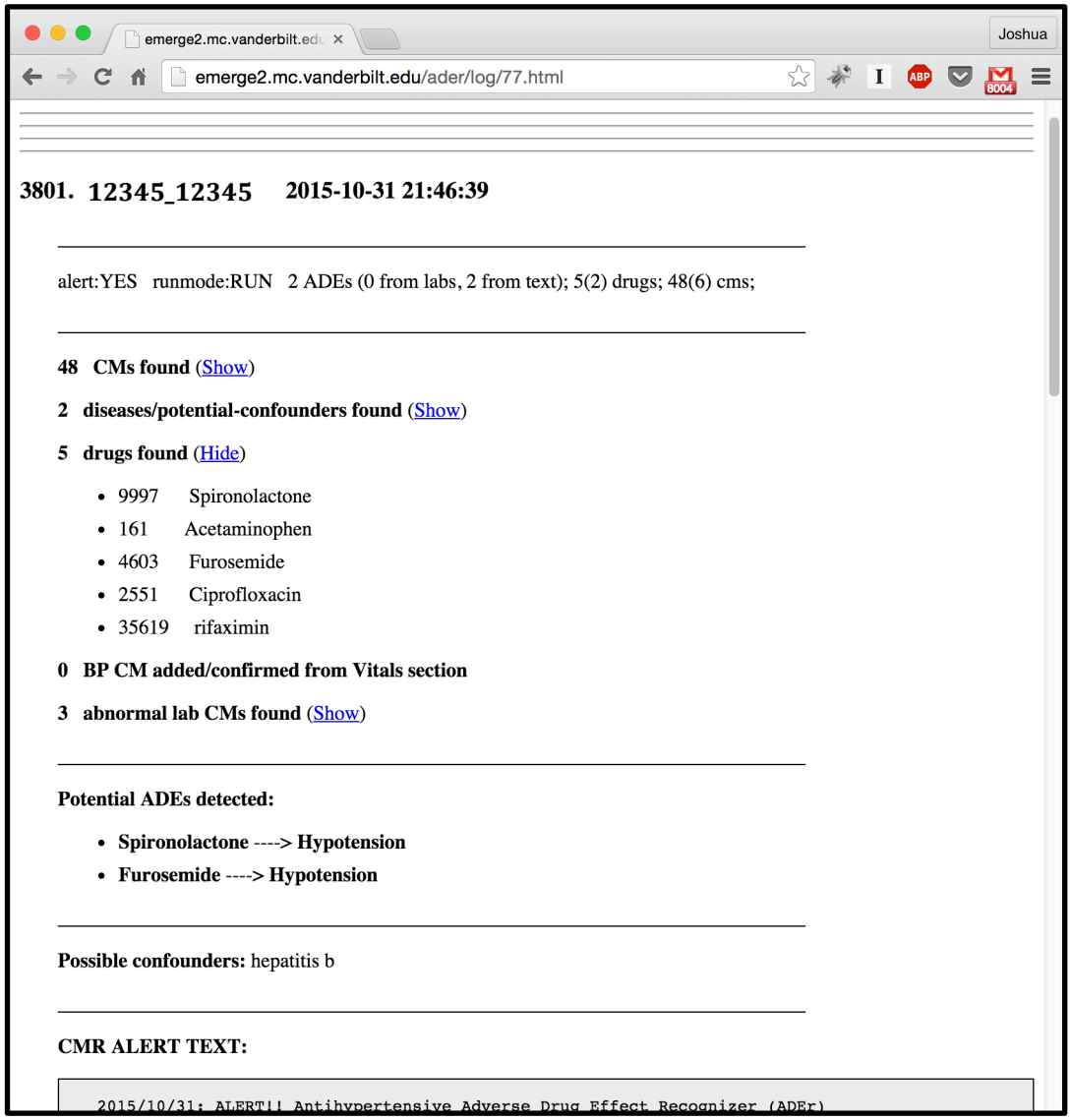


Figure 19 (a). HTML log of ADER activity showing CMs, medications, laboratory CMs, possible ADEs, and possible confounders.

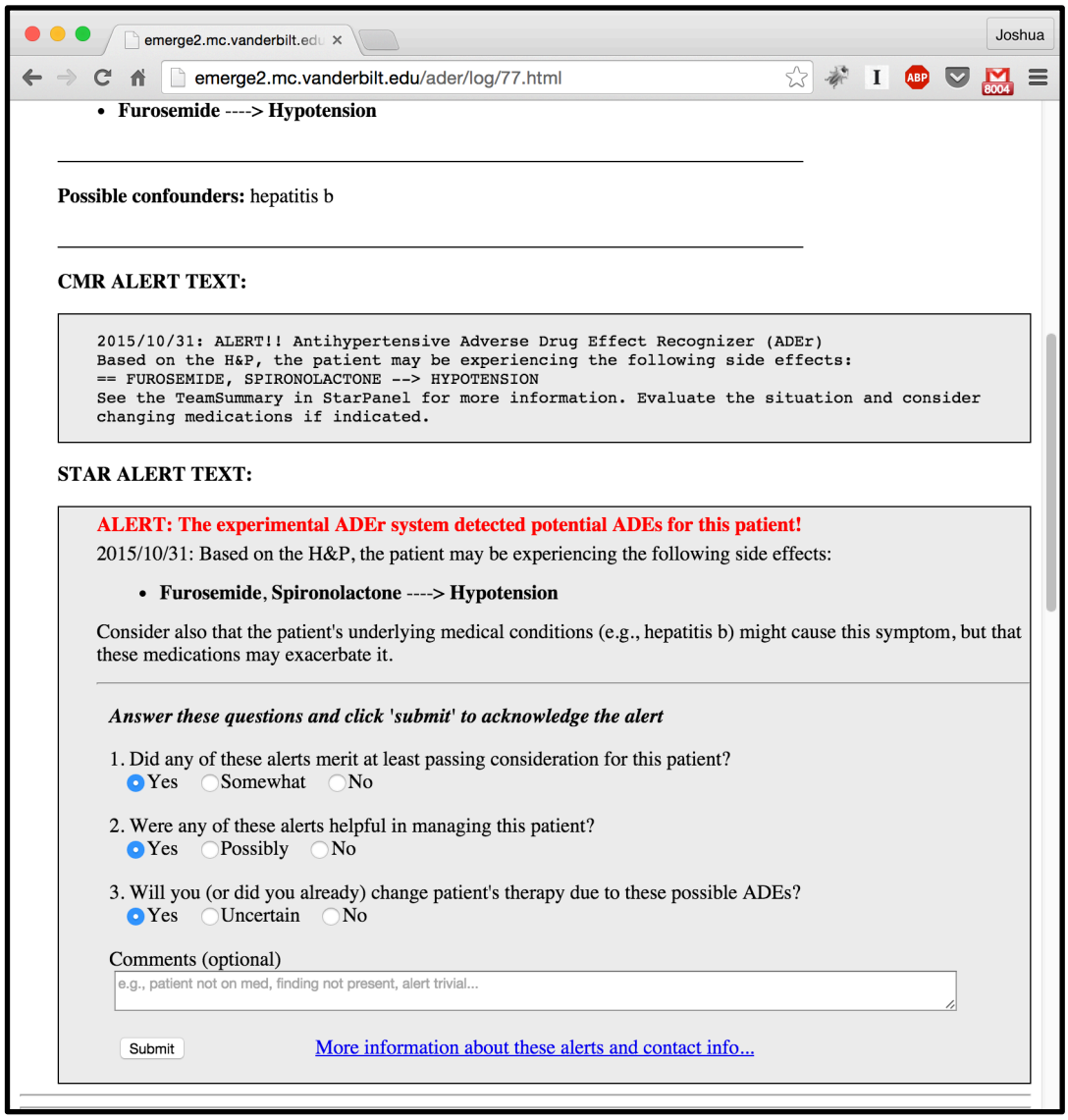


Figure 19 (b). HTML log of ADER activity showing the generated CMR alert text and the StarPanel alert HTML with survey.

The ADER Status Monitor & Control Panel is a CGI web interface. The webpage contains a series of panels used by the project team members to monitor the ADER system status, as well as the status of the pilot evaluation. The interface is shown in Figure 20. The first panel is the ADER Status panel; it provides the same information available in status update emails. The second panel is the Survey results panel, containing the current cumulative results of the redesigned survey format. The panel also contains the CMs of the detected ADEs. For each CM, the interface displays the responses to each question. The project team members used this information to determine if any of the potential ADEs were

causing an inordinate number of false positives or ADE alerts not considered useful. The third panel shows the survey results from the original survey design. The fourth panel shows the status of the ADER component scripts running on the server. The fifth panel shows the last 1000 lines of the ADER system log. In the event of a system component failure, the project team members can use the log to determine the cause and reinitialize the system.

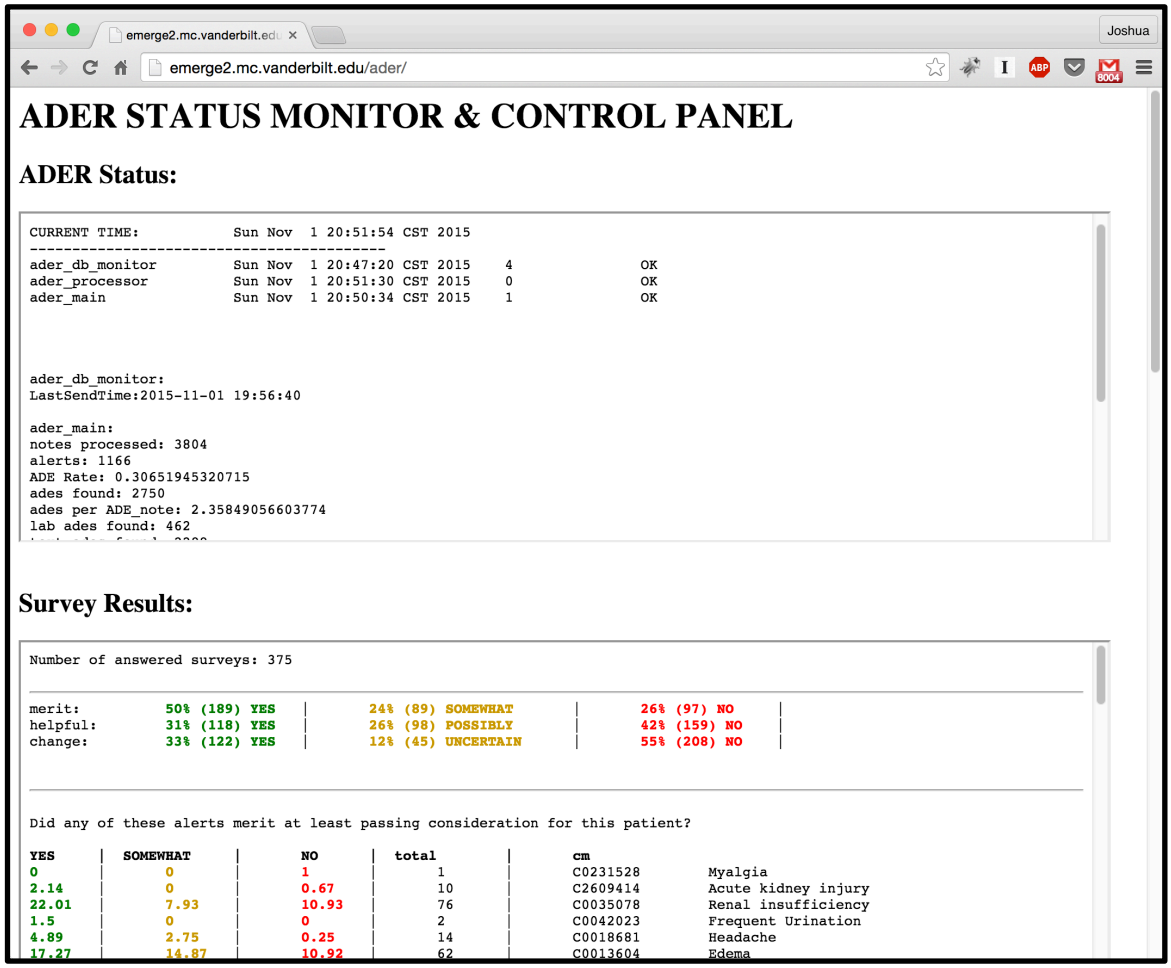


Figure 20 (a). Status Monitor web interface, showing most recent ADER status message and cumulative survey results.

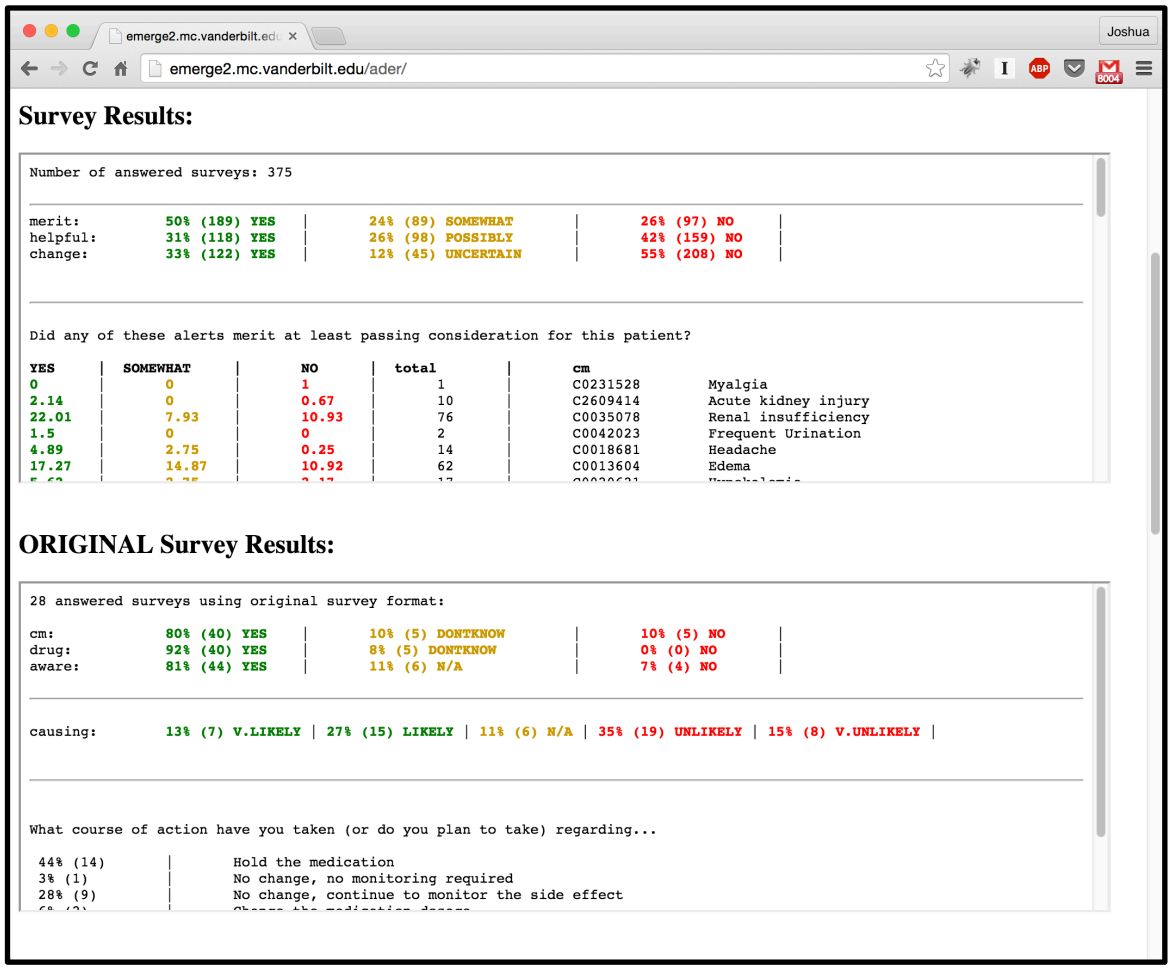


Figure 20 (b). Status Monitor web interface, showing cumulative survey results from the redesigned format, as well as the original survey results.

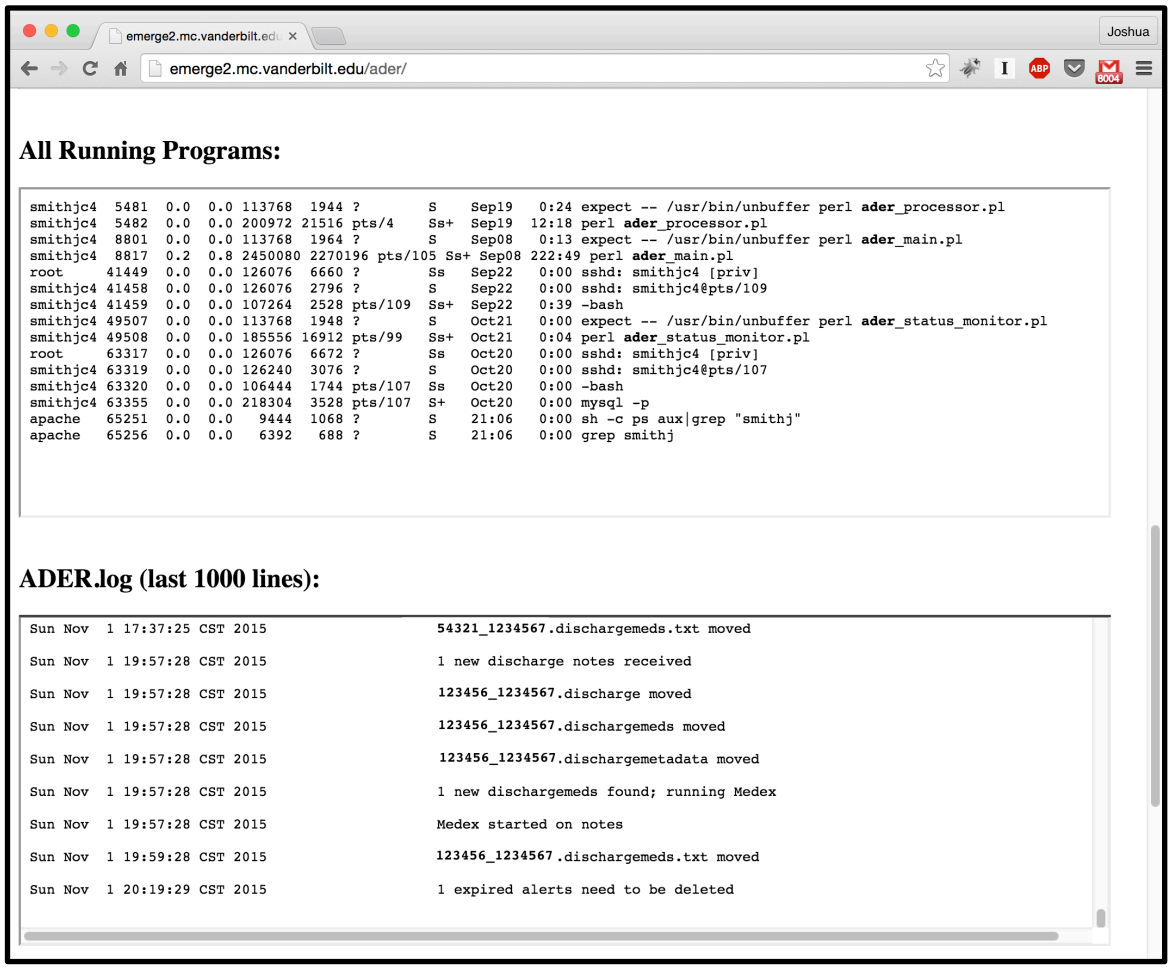


Figure 20 (c). Status Monitor web interface, showing most ADER components running on the ADER server and the last 1000 lines of the ADER log file.

Conclusion

A successful ADE detection system requires an accurate and reliable source of patient information; the ability to identify CMs, medications, and potential confounders; a set of known, relevant ADEs on which to alert; and informative alerts presented to providers at appropriate points in their workflow. The ADER system addresses these criteria. The ADER pilot implementation is a complex, multi-component system consisting of dozens of files and over 10,000 lines of code, not including the NLP tools used by the system. It is connected across two servers and dozens of directories, linked to the Vanderbilt EMR though secure messaging protocols. While the ADER system is complex, extensive logging and real-time monitoring allow the project team to maintain ADER with minimal input. The extensive data collection is useful for monitoring the live system and

post-implementation analysis, as well as process improvement for the current instance and future iterations of the ADER system. The code described above was used for the ADER pilot evaluation described in the next chapter.

CHAPTER VI

ADER IMPLEMENTATION

Introduction

The project team implemented a pilot study of the ADER system at Vanderbilt University Hospital. The team designed the pilot study to evaluate ADER system methodology and functionality. The study tested the hypothesis that ADER can help providers to better recognize and respond to ADEs presenting in patients at admission. The Vanderbilt IRB approved the study (Appendix A). To detect possible ADEs from outpatient medications, the ADER system scanned admission H&P notes written by Internal Medicine interns and residents during the pilot. If ADER detected any ADEs, the system alerted the patient's care team and collected survey responses from clinicians who received alerts. After the implementation period, the project team analyzed the ADER alerts and survey questionnaires, as well as corresponding inpatient medication orders and discharge notes to determine whether clinicians held or changed the alerted suspected ADE-causing medications. The project team limited the ADER pilot to patients admitted to the Internal Medicine service by interns and residents and limited ADEs to a predefined set of common antihypertensive medications, as described in Chapter IV. This chapter details the pilot study design, the results of the ADER pilot, and the significance of those results.

Materials

This work utilized a MacBook Pro with a 2.6 GHz Intel Core i7 processor and 16 GB of RAM, a Linux server with forty-eight 2.2 GHz AMD Opteron cores and 256 GB RAM, referred to as the ADER server, and a database server with two 8-core Intel Xeon processors and 6 GB of RAM referred to as the database server. All data processing scripts were written in Perl. 5.10.0. The project used MySQL version 5.5.16 on the ADER server and MySQL version 5.0.95 on the database server. Derivation of the ADER AAS is described in Chapter V. Study participants interacted with the ADER system through the Vanderbilt electronic medical record system, StarPanel, and received alerts thought StarPanel and the Vanderbilt CPOE system, Horizon Expert Orders.

Pilot Study Design

The project team designed a pilot implementation study to evaluate the ADER system in consultation with VUMC expert staff pharmacists, Internal Medicine residents and physicians, a biostatistician, and the director of the Internal Medicine Residency Program. The project team hypothesized that ADER would measurably improve clinicians' recognition of and response to their patients' ADEs. To evaluate this hypothesis, the system included survey questions with each ADER alert. Additionally, the project team analyzed the rate at which clinicians held or reduced the dose of antihypertensive medications after receiving an ADER alert. The team also ran ADER on a retrospective "control group" of historical Internal Medicine service admission notes to identify instances where alerts would have been issued had the system been live at the time. The analysis determined for those cases how often changes were made to the medications that would have caused alerts. Finally, the analysis compared that rate to the rate of changes in the pilot study for clinicians who received ADER alerts. The project team also believed that ADER notifications might provide a useful education tool for interns and residents; the anonymous survey responses provided data on this.

Procedure and Data Collection

At the time of a patient's admission to the Internal Medicine service, the ADER system used NLP methods to recognize each potentially eligible patient's medications and CMs from admission H&P notes. The ADER system then cross-referenced the extracted drug and CM concepts against a set of known medication-ADE associations (the AAS). Finally, ADER alerted clinicians (i.e., any house staff member of the patient's care team who accessed the team summary or the progress note form) about any potential ADEs through StarPanel and Horizon Expert Orders, the Vanderbilt CPOE system. Alerts listed potential ADEs and the medications suspected of causing them, as well as any potentially confounding diseases mentioned as present in the H&P note.

The pilot study involved running the ADER system for three months, collecting data on alerts generated by the system and the physician responses to the anonymous surveys. The study recorded H&P notes and laboratory data that generated the alerts, as well as

historical H&P notes from before the pilot implementation start date. The study also collected an analyzed corresponding discharge notes and inpatient medication orders.

The ADER system is a combination of multiple Perl scripts and two databases running on two separate servers. Each component must function normally for the system to operate correctly. Due to this complexity, all systems will sometimes fail or "crash." The ADER Status Monitor component (see Chapter V) alerted project team members if any part of the system failed to operate properly due to software malfunction or other technical difficulties.

Post-Alert Survey Questionnaire

In the concluding section of an alert, ADER presented a survey questionnaire related to the potential ADE and the intended response to it. Completion of the electronic survey questionnaire was anonymous and voluntary. The project team originally designed the survey questionnaire to determine if the clinician was aware of the potential ADE, whether the ADE seemed plausible or likely in this case, and whether it required intervention or changes to the patient's therapy. Throughout the pilot, study participants responded to survey questions regarding ADER alerts in StarPanel. The number of questions in the original questionnaire format increased in proportion to the number of ADEs detected in the note. During the initial phase of the ADER pilot study, a low survey response rate occurred. As a result, the project team simplified the questionnaire and made it directly visible as part of the alert (previously, the user had to click a hyperlink to open the questionnaire). The new survey involved only three questions that assessed whether the ADE alert merited consideration, was helpful in managing the patient, and if the respondent had changed or intended to change any of the patient's medications due to the detected ADEs. After the patient was discharged, ADER automatically extracted discharge medication data from the Patient Discharge Summary. At the end of the pilot implementation period, project members collected any discharge notes not automatically identified by ADER, as well as inpatient medication orders for each patient.

Project team members analyzed the survey results after the pilot study ended. The team identified the rate at which suspected ADE-causing medications were held or changed during the inpatient stay, based on inpatient medication orders, and at discharge, based on

discharge summaries. The project used ADER to detect potential ADEs during two time periods: first, during the pilot study and second, during the year prior to the pilot study. The project analyzed H&P notes and medication orders using ADER for that time period, noting the alerts that ADER generated in retrospect (and which were never issued to clinicians at the time). The ADER analysis compared the hold/change rate for alerted medications included in the pilot study against the hold/change rate for what would have been alerted ADE-causing medications detected by ADER in retrospective data (i.e., before ADER alerts were implemented). The analysis evaluated both the effectiveness of the system in improving ADE recognition and its effectiveness in improving clinicians' responses to ADEs.

Pilot Study Enrollment and Training

The pilot study implemented the ADER alerting system for the residents and interns on the general Internal Medicine Service at Vanderbilt University Hospital. Each care team consists of an attending physician, one resident, and one intern. The project team believed the general internal medicine service to provide a good initial test for this system, instead of including specialty wards such as those dealing solely with cardiology or oncology patients.

With the support and permission of the Director of the Internal Medicine residency program, the IRB granted a waiver of documentation of informed consent for the physician study subjects. At a meeting of the entire Department of Medicine house staff prior to the start of the pilot study, project members informed potential subjects about the ADER system, the pilot study, and the nature of the ADE alerts. Subjects learned that responding to the alert or survey information was strictly voluntary, and that the pilot study would collect their survey responses anonymously. Responding to the survey questions constituted agreement to participate in the study. The project team reminded subjects that the system was meant solely to improve their recognition of possible ADEs and their response to ADEs, and that the system should not in any way be a substitute for the care team's clinical judgment. This information was provided both in-person to potential study subjects and in writing. Links to information about the project and contact information for the project team were included in every ADER alert and notification.

Pilot Study Considerations

No patient medications were directly changed by the ADER system; all treatment and medication decisions remained entirely with the care team. The purpose of the system was only to make clinicians aware of potential ADEs, and to prompt them to consider changing medications *when appropriate*. Once the study period began, the ADE detection and alerting system was activated for all eligible internal medicine residents and interns (i.e., those on targeted hospital wards) simultaneously.

The project team and the IRB (approval #141341), determined this study to be low-risk, but project team members continually monitored physician responses to the alerting system and provided the Director of the Internal Medicine Residency Program, the Internal Medicine Chief Resident, and all participants with status updates. If the clinical staff found the alerts to be impeding patient care or causing a disruption, the project team members would have de-activated the system and contacted the IRB immediately. The project team presented preliminary results to the physician study participants after the completion of the ADER Pilot implementation.

Pilot Implementation

Historical Admissions as a Control Group

Before the pilot implementation, the project team retrieved Internal Medicine admission H&P notes written by interns and residents during the year prior to the study start date. The team denoted this set of retrospective Internal Medicine admission notes as the study control group. The project team ran ADER on the retrospective notes in the control group to detect potential ADEs, and recorded any alerts that would have been generated had the system been "live" at the time. The team analyzed medication orders in the control group to analyze the differences in behavior between clinician who received an ADER alert and those in the control period who did not. Specifically, for the control group, the project team identified the rate at which ADER would have detected potential ADEs, as well as the rate at which potential ADE-causing medications were held, changed, or discontinued.

Pilot Software Testing

To debug the ADER system, project team members extensively tested the ADER system on the retrospective notes in the control group and also ran the system silently (without alerts) for nearly a month before activation in the EMR. Team members verified that ADER was correctly identifying medications and CMs, as well as potential disease confounders. Team members also used this data to refine the ADER alert design, confirm the accuracy of ADER's activity log recording, and ensure the ADER system interacted with StarPanel and the various ADER components as expected.

Pilot Implementation

Before beginning the pilot, the project team estimated that the system would receive notes for approximately 150-200 patients per week. The estimate was consistent with the publicly available information[138] stating "Clinical faculty in the Department [of Internal Medicine] admitted over 17,700 patients to the Vanderbilt University Hospital or the on-campus Nashville Veterans Administration Hospital in fiscal year 2015." Previous research suggested that anywhere from 5% to 30% of inpatients could have experienced an ADE prior to admission.[35,36]

Evaluation Methods

The project team hypothesized that the system would measurably improve recognition of and response to ADEs. The team tested this hypothesis using a variety of evaluation methods described below. Figure 21 lists the questions asked by the analysis at each point between admission and discharge. The figure illustrates the path of antihypertensive medications taken by the patient before admission, at admission, throughout the inpatient stay, and at discharge. Each evaluation point is labeled in the figure and referenced throughout this section.

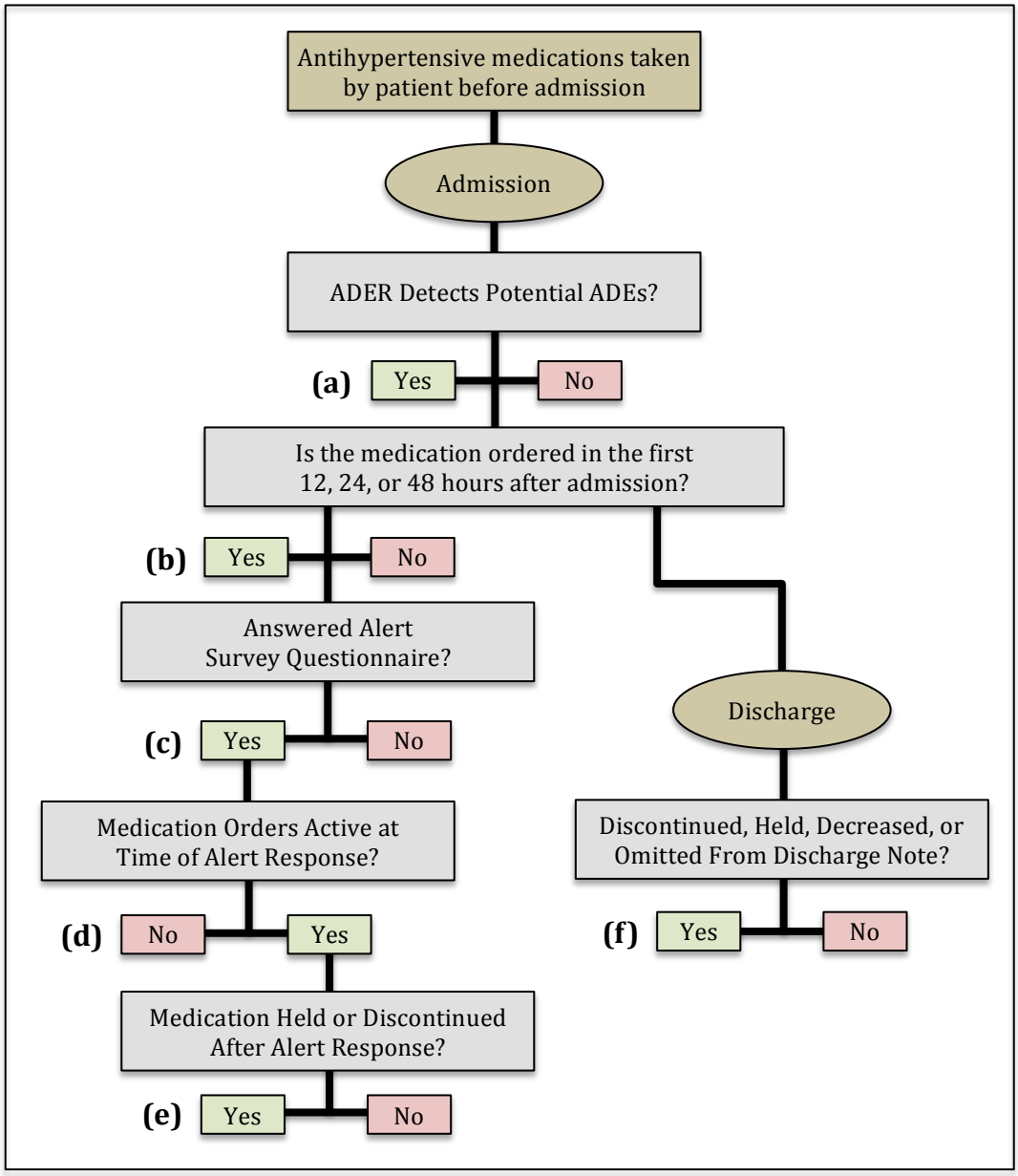


Figure 21. Diagram showing how the ADER Pilot study evaluated antihypertensive medications from the Admission H&P note, throughout the inpatient stay using inpatient medication orders, and at discharge from the Discharge Summary.

Post-Alert Medication Holds and Changes

After ADER scanned admission H&P notes, as shown in Figure 21, for potential ADEs (a), project team members analyzed both inpatient medication orders (b-e) and patient's discharge summaries (f) to determine if providers changed suspected ADE-causing medications during the inpatient stay, or for outpatient status after hospital discharge. The

analyses included only antihypertensive medications. The EHR recorded the inpatient medication orders. Automated review revealed what medications were ordered for the patient during their hospital stay, as well as when medications were given, held, or discontinued. Discharge summaries revealed whether outpatient medications where to be held or discontinued after the patient left the hospital. Additionally, the dosage of suspected ADE-causing medications might have been reduced to address the potential ADEs; the analysis identified dosage decreases in both inpatient medication orders and discharge summaries.

As with admission H&P notes, the project team used the MedEx tool to extract medication information including name, form, strength, dose, route, and frequency from both recorded inpatient medication orders and from discharge summaries. Each medication was normalized to its generic ingredient as described in Chapter III. Using the MedEx-extracted *strength* and *frequency* signature information, the system calculated the normalized daily dose of each medication. The system uses regular expressions to identify the daily frequency (*BID*, *twice a day, twice daily, 2x day*, etc.). For example, a mediation listed as "50 mg twice daily" would have a normalized dose of 100mg; a medication ordered as "250 mg daily" would have a normalized dose of 250mg. Normalized doses were only compared if their *route* (PO, IV, etc.) was the same. However, medication signature information, including strength, frequency, and route, is sometimes missing, especially in admission notes. If a normalized dose could not be calculated because data was missing for either dosage, the system did not compare those doses or attempt recognize a dose decrease for that medication.

For inpatient medication orders, the analysis determined whether the medication was given, held, discontinued, or decreased, as shown in Figure 21(b). Each order has a medication order entry time, start time, and discontinue time; some orders also have a "hold start time" and "hold stop time." To determine whether medications were given after admission, the system processed all medication orders written between 24 hours before admission and 24 hours after discharge. The system then parsed the patient stay into 12-hour periods from admission to discharge and determined those medications with an active medication order during each time period.

For the discharge summaries, the analysis determined whether each admission antihypertensive medication (including those causing alerts) was to be continued, held, discontinued, or decreased for subsequent outpatient status, as shown in Figure 21(f). The analysis also noted if a medication was omitted from the discharge note. A medication was considered "held" if the medication section specifically mentioned the drug was held at discharge. The analysis recognized medications listed in the "Stopped Discharge Mediations" section of the note as "discontinued," or "D/C." Additionally, some admission medications are not mentioned in the discharge summary; it is unclear whether those medications omitted from the discharge summary were continued, held, or discontinued (e.g., the patient may have already purchased a 90-day supply, so no refill prescription was given; or, the patient may have been told not to take the medication, but no record of this was made). As described above, the analysis also compared normalized doses to determine if suspected ADE-causing medications had their dosages reduced at discharge.

The project team members performed the above analyses for both the pilot study group and the retrospective control group that did not receive ADER alerts. The project team compared difference in rates for the pilot group and the control group to determine if ADER alerts affected provider behavior in regards to suspected ADE-causing medications.

System Performance

As stated above, the ADER Status Monitor component alerted project team members if any part of the system went down due to malfunction or technical difficulties. The project team analyzed system performance, including ADER up time and downtime during the pilot study. The team also calculated the time elapsed between when an H&P note was saved in StarPanel and when an alert was issued, as well as the time elapsed between an alert being issued by ADER and study participants responding to the survey questionnaire.

Analysis of ADER Alerts

The project team calculated summary statistics for the ADER pilot implementation to determine the number of alerts generated, the number of unique patients in that set, and the number of ADEs detected. The team analyzed the specific ADEs detected, the medications and CMs causing those ADEs, and whether the ADEs were identified based on

H&P text or lab results. The team performed the same analysis for both the experimental ADER-alerting set and the retrospective, pre-ADER control group.

Analysis of ADER NLP Accuracy

The project team reviewed 100 H&P notes from the study set where ADER detected potential ADEs to determine the accuracy of ADER's NLP. To create the set used for analysis, a script randomly selected at least one note containing ADEs with each distinct CM and medication in the ADER Alerting Subset (AAS); next, additional notes with detected ADEs were randomly included up to a final count of 100 notes. Team members analyzed each note to determine if the CM and medication were actually mentioned affirmatively in the note and, if not, attempted to discern the reason ADER misidentified the CM.

Analysis of ADER Survey Questionnaire Responses

The project team analyzed the collected responses from the survey questionnaires included with each ADER alert for both the original long-form survey and the shorter three-question version, as illustrated in Figure 21(c). The team calculated the margin of error for the survey responses. Since both questionnaire formats allowed for optional comments, the project team analyzed all comments submitted by study participants, when necessary comparing them against the original H&P notes that triggered the alert.

Admission Medications compared to Inpatient Medication Orders

The team evaluated inpatient medication orders corresponding to those patients whose admission H&Ps were included in both the ADER pilot study and retrospective control group. As stated above, the system parsed the inpatient stay into 12-hour periods from admission to discharge and identified those medications with an active medication order during each time period. The team analyzed patterns across all inpatient medication orders, as well as the rates at which medications were not given (held, discontinued, or not ordered) or decreased during the first 12 hours, 24 hours, and 48 hours after admission, as illustrated by Figure 21(b). The analysis also determined the number of suspected ADE-causing medication with active orders at the time the provider responded to the survey questionnaire (Figure 21(d) above). The analysis determined the rates at which those

active medications were held, discontinued, or decreased *after* the provider responded to the alert (Figure 21(e) above). The team compared the difference in hold/change rates for both the pilot and control groups using hypothesis testing for two proportions.

Admission Medications Compared to Discharge Medications

Similarly, the team examined the difference between admission medications and discharge medications in patients with suspected potential ADEs and evaluated the rate at which medications were held, discontinued, or changed in both the experimental study population (where ADER detected potential ADEs and issued alerts) versus the control group (retrospective notes with detected potential ADEs but without ADER alerts). This is shown above in Figure 21(f). To account for the unknown status of medications not mentioned in the discharge summaries, the team analyzed both the rate of medication being held or discontinued, as well as the rate of medications being held, discontinued, or *not included* in the discharge notes. Again, the team compared the difference in hold/change rates for both the pilot and control groups using hypothesis testing for two proportions.

Medication Holds Analyzed by ADE

The project team analyzed all alerts detected during the pilot study. For each CM suspected to be due to an ADE, the team calculated the number of times that any drug suspected of causing that CM was held or the dosage decreased. With this analysis, the project team hoped to identify which ADE CMs were more likely to result in a medication changes or holds.

Suspected ADE-Causing Medications Compared to All Medications

To help determine whether ADER was successfully able to recognize medications that should be held or changed, the team analyzed the hold/change rate for antihypertensive medications suspected of causing ADEs versus those not recognized as causing ADEs. The analysis compared the two rates to determine if there was a significant difference between the rates at which those medications were held or changed.

Survey Responses Compared to Discharge and Inpatient Medication

As noted above, one of the questions included with ADER alerts was "Will you (or did you already) change patient's therapy due to these possible ADEs?" When responses to the short-form survey questionnaire indicated the clinician had made a change to suspected ADE-causing medications, the team attempted to identify whether those changes (medication holds or a decrease in medication dose) actually took place.

Results

The ADER Pilot study ran for 13 weeks (or 92 days) from Saturday, August 1, 2015 through Saturday, October 31, 2015. These results include data from notes scanned and ADER alerts issued during the pilot study period, as well as potential ADEs detected by ADER (without alerts) from the retrospective control group.

ADER System Performance

Over the course of the 13-week ADER pilot study, the system crashed 12 times, mostly in the first half of the study. Each time, the ADER Status Monitor Component alerted the project team. The average downtime was 2.5 hours, with a median downtime of one hour, and a maximum downtime of 7.5 hours. Each time the system was restarted, it was indexed to the time of failure; this resulted in some alerts being delayed, but no admission H&P notes being missed. The early system crashes were due to software "bugs" that were not serious. One system failure was due to the ADER server running low on memory storage due to another, unrelated project temporarily using too much space on the server. The majority of crashes were due to the ADER Database Monitor component. The cause of the problem was discovered late in the implementation – a regular expression would lock the ADER Database Monitor component into an infinite loop due to improper formatting in a small number of discharge notes. The longest time the system was down was from approximately midnight to 7:30am. The ADER Detector and Processor components did not crash after the second week of the pilot. No components crashed after the ADER DB Monitor bug was discovered and corrected.

Considering the 12 system failures and an average downtime of 2.5 hours, ADER was down approximately 30 hours over the course of the 3-month study period. That equates to 1% downtime over the entire study period.

The average time elapsed between when an H&P note was saved in StarPanel and an alert was issued was 11.8 minutes; 75% of alerts were issued less than 9 minutes after the note was saved. The maximum delay was 427 minutes when the system was down overnight for its aforementioned maximum 7.5-hour downtime. This does not include one outlier that took 41 hours between note submission time and the alert time; this was the result of an extended processing time from a single KMCI thread for an unknown reason.

The average time elapsed between an alert being issued by ADER and study participants responding to the survey differed slightly between the long-form and the shorter-form questionnaire. The average response time for the original-format questionnaire was 36.6 hours, or about two and half days. The minimum response time was 2.2 hours; the median response time was 20 hours; and the maximum response time was 103 hours (or 4.3 days). The average response time for the short-form questionnaire was 29.4 hours. The minimum, median, and maximum response time was four minutes, 23.5 hours, and 116 hours (or 4.8 days), respectively. As mentioned in Chapter V, alerts expired either five days into the admission or when the patient was discharged.

Pilot Implementation Results

During the study period, the system scanned 3103 notes from 2708 unique patients. A total of 309 patients were admitted more than once during the study period; 247 of whom were admitted twice, 47 who were admitted three times, 10 who were admitted four times; and 5 who were admitted more than four times. Approximately 200 interns and residents authored the notes scanned by the system (there are 150 residents in the Department of Internal Medicine and residents from other specialties rotate through Medicine wards as part of their training). The system generated alerts using the original long-form survey questionnaire prior to August 23, 2015, and used the new short-form survey questionnaire thereafter.

Of the 3103 notes scanned by ADER, 927 notes from 840 unique patients generated alerts due to detected potential ADEs. That is, potential ADEs were detected in 30% of

admission notes. The system identified an average of 1.2 antihypertensive medications and 2.4 AAS CMs per patient note. In those notes with alerted ADEs, ADER identified an average of 2.4 antihypertensive medications and 3.4 AAS CMs per patient. The average length of admission was 114 hours (or 4.8 days).

Each alert can include multiple ADEs; a total of 2191 ADEs were found – an average of 2.4 ADEs per patient when ADEs were detected. Approximately 83% of detected ADEs were found in the text of the H&P and 17% were based upon laboratory results. The minimum, first quartile, median, third quartile, and maximum number of ADEs detected were 1, 1, 2, 3, and 15 ADEs, respectively. The results compared to the control group are shown in Table 29 in the next section.

Table 26 lists the number of ADEs found for each medication and Table 27 lists the number of ADEs detected for each CM. Table 28 shows the most frequent medication-ADE pairs detected by the system – those ADEs detected at least 15 times during the pilot study. The top 43 ADEs in the table account for 65% of all ADEs detected by the system.

Table 26. Medications responsible for ADEs detected by ADER. Listed in decreasing order of frequency. The "count" is the number of ADEs detected with that medication; the "fractional count" is the proportion of ADEs in each alert accounted for by that drug.

Medication	Count	Fractional Count	Medication	Count	Fractional Count
Lisinopril	389	178.3	Quinapril	5	0.9
Furosemide	313	122.5	Fosinopril	5	1.9
Metoprolol	191	84.4	Nadolol	4	2.8
Amlodipine	182	103.7	Nicardipine	4	2.5
Spironolactone	137	40.0	Prazosin	4	1.3
Hydrochlorothiazide	136	62.1	Triamterene	4	0.4
Losartan	134	48.5	Telmisartan	4	1.9
Carvedilol	114	36.4	Verapamil	3	1.3
Nifedipine	97	43.2	Nebivolol	3	1.5
Clonidine	80	30.2	Terazosin	3	1.2
Metolazone	64	25.7	Olmesartan	3	2.2
Atenolol	41	18.5	Felodipine	2	1.5
Diltiazem	36	17.1	Methyldopa	2	0.8
Bumetanide	33	11.0	Eplerenone	2	0.2
Torsemide	32	10.1	Captopril	1	0.5
Irbesartan	21	12.5	Isradipine	1	0.5
Enalapril	18	10.8	candesartan	1	0.3
Valsartan	18	8.3	Acebutolol	0	0.0
Chlorthalidone	16	7.3	Amiloride	0	0.0
Ramipril	14	6.8	Betaxolol	0	0.0
Propranolol	13	5.4	Chlorothiazide	0	0.0
Benazepril	13	5.2	Guanethidine	0	0.0
Labetalol	10	2.8	Nisoldipine	0	0.0
Minoxidil	10	4.5	Pindolol	0	0.0
Hydralazine	8	1.3	Reserpine	0	0.0
Indapamide	8	4.5	Ethacrynate	0	0.0
Bisoprolol	6	2.4	Eprosartan	0	0.0
Doxazosin	6	2.0	**Total:**	**2191**	**927**

Table 27. CMs responsible for ADEs detected by ADER. Listed in decreasing order of frequency. The "count" is the number of ADEs detected with that CM; the "fractional count" is the proportion of ADEs in each alert accounted for by that CM.

CM	Count	Fractional Count	CM	Count	Fractional Count
Lightheadedness	372	109.0	Tinnitus	4	1.5
Hypotension	309	112.6	Vomiting	4	1.7
Dizziness	296	92.0	Pruritus	3	1.3
Syncope	207	61.3	Hyperuricemia	3	1.7
Renal insufficiency	185	101.6	Somnolence	3	1.5
Edema	140	90.8	Heart Block	2	0.6
Coughing	75	50.4	Erectile dysfunction	2	0.4
Hyperkalemia	55	22.2	Mental Depression	1	1.0
Constipation	51	29.8	Gynecomastia	1	0.5
Hypokalemia	45	22.7	Pericardial effusion	1	1.0
Headache	42	21.0	Raynaud Disease	1	1.0
Hypovolemia	42	12.9	Myalgia	1	1.0
Bradycardia	40	24.1	Agranulocytosis	0	0.0
Pancreatitis	37	21.1	Angioedema	0	0.0
Dehydration	35	12.4	Anxiety	0	0.0
Gout	35	21.9	Arthralgia	0	0.0
Exanthema	28	17.5	Cholestasis	0	0.0
Hyponatremia	28	14.7	Gingival Hyperplasia	0	0.0
Alkalosis	26	19.7	Hallucinations	0	0.0
Hyperglycemia	21	11.3	Heartburn	0	0.0
Orthostatic Hypotension	21	4.8	Hypertrichosis	0	0.0
Acute kidney injury	17	5.6	Kidney stones	0	0.0
Peripheral edema	10	7.7	Interstitial Nephritis	0	0.0
Confusion	9	5.2	Nephrocalcinosis	0	0.0
Polyuria	8	3.5	Bullous pemphigoid	0	0.0
Hearing loss	7	5.2	Pemphigus	0	0.0
Muscle Cramp	5	1.5	Psychosis	0	0.0
Cardiogenic Shock	5	2.6	Impaired Vision	0	0.0
Frequent Urination	5	2.7	Dry mouth	0	0.0
hearing impairment	5	3.3	Recent weight gain	0	0.0
Anorexia	4	3.2	**Total:**	**2191**	**927**

Table 28. Most frequently detected ADEs (those occurring at least 15 times) during the pilot study; listed in decreasing order of frequency. The "count" is the total number of ADEs detected; the "fractional count" is the proportion of each ADEs from all alerts.

Drug	CM	Count	Fractional Count
Lisinopril	Renal insufficiency	97	53.6
Amlodipine	Edema	90	64.2
Lisinopril	Coughing	66	44.0
Lisinopril	Lightheadedness	58	17.0
Furosemide	Lightheadedness	55	15.2
Metoprolol	Hypotension	53	25.9
Losartan	Renal insufficiency	50	24.5
Furosemide	Hypotension	49	16.2
Furosemide	Dizziness	47	14.9
Lisinopril	Dizziness	45	13.8
Lisinopril	Hypotension	44	16.6
Metoprolol	Lightheadedness	44	15.3
Metoprolol	Dizziness	41	16.8
Nifedipine	Edema	41	21.4
Furosemide	Syncope	39	12.0
Lisinopril	Syncope	33	11.2
Carvedilol	Lightheadedness	30	7.0
Metoprolol	Syncope	30	13.9
carvedilol	Dizziness	28	7.1
Spironolactone	Hypotension	28	7.4
Furosemide	Alkalosis	26	19.7
Hydrochlorothiazide	Lightheadedness	26	8.2
Carvedilol	Hypotension	23	6.7
Losartan	Lightheadedness	23	5.9
Amlodipine	Constipation	22	14.5
Spironolactone	Dizziness	22	4.6
Amlodipine	Dizziness	21	8.3
Amlodipine	Lightheadedness	21	6.3
Clonidine	Headache	21	9.9
Hydrochlorothiazide	Hyperglycemia	21	11.3
Spironolactone	Hyperkalemia	21	8.9
Lisinopril	Hyperkalemia	20	7.9
Losartan	Dizziness	20	6.4
carvedilol	Syncope	19	4.9
Furosemide	Hypovolemia	19	7.2
Metolazone	Gout	19	12.1
Spironolactone	Lightheadedness	18	3.4
Lisinopril	Pancreatitis	17	10.8
Furosemide	Dehydration	17	7.8
Furosemide	Hypokalemia	17	8.6
Losartan	Hypotension	17	4.7
Metoprolol	Bradycardia	16	8.8
Nifedipine	Constipation	16	8.4
	Total:	**1430**	**613.3**

Control Group Analysis

The project team obtained 9,248 retrospective Internal Medicine admission H&P notes for use as a comparison experimental control; interns and residents authored the notes between May 2014 and April 2015. The set consisted of notes from 7272 unique patients. A total of 1297 patients were admitted more than once during that time; 885 of whom were admitted twice, 259 of whom were admitted three times, 89 of whom were admitted four times; and 64 who were admitted more than four times. The average length of admission was 114 hours – the same as the pilot study group. Table 29 illustrates the similarities between the pilot group and the control group.

Table 29. Pilot study group compared to control group.

	Notes	**Unique patients**		**Admission Notes with ADEs Detected**		**Notes without ADEs**		**Notes with ADEs**	
	Num.	**Num.**	**Pct.**	**Num.**	**Pct.**	**Anti-htn. Drugs**	**ADE CMs**	**Anti-htn. Drugs**	**ADE CMs**
Pilot	3103	2408	78%	927	30%	1.2	2.4	2.4	3.4
Control	9248	7272	79%	2691	29%	1.2	2.2	2.4	3.2

		ADER ADE Source		**Number of ADEs Found per Alert**					
	Average Length of Stay	**H&P Text**	**Labs**	**Min.**	**1Q**	**Median**	**3Q**	**Max.**	**Avg.**
Pilot	114 hours	83%	17%	1	1	2	3	15	2.4
Control	114 hours	85%	15%	1	1	2	3	23	2.4

The team scanned the notes using ADER, which identified 2691 notes with detected potential ADEs. That is, potential ADEs were detected in 29% of admission notes, nearly the same as the pilot group. The average number of ADEs found per note with detected ADEs was 2.4. Approximately 85% of detected ADEs were found in the text of the H&P and 15% were based upon laboratory results. The distribution of the number of ADEs detected per note is nearly identical to the pilot group. As with the pilot group, patients were more likely to be experiencing a potential ADE when they were on more antihypertensive medications and their notes mentioned more AAS CMs. The relative number of ADEs found for each medication, each CM, and the most frequent medication-ADE pairs detected by the system were also similar to the experimental pilot set.

ADER NLP Accuracy

The project team reviewed 100 H&P notes where potential ADEs were detected in the pilot study to assess ADER's NLP accuracy. For each concept, project team members determined whether the CM or medication was mentioned affirmatively in the note. The reviewed notes contained 157 medication concepts and 166 CM concepts (136 CMs identified from text, 30 CMs from lab values).

A total of 98% of the drug concepts identified by ADER were mentioned as current medications. Of the three that were judged inaccurate, the medications were mentioned in the "History of Present Illness" section of the note, but only as a previous medication that had been discontinued.

A total of 86% of the CM concepts identified from text were found to be affirmatively mentioned CM concepts. Of the 20 CMs found to be incorrect, three quarters were misidentified due to problems with negation detection. The Negex algorithm failed to recognize 9 CMs as negated due to abnormal formatting in the note , including two cases where providers included both a *negative* review of systems section written as "-ROS") and *positive* review of systems section ("+ ROS"). In one instance, providers abbreviated the term "without" as "w/o," which was not included in Negex as a negation "trigger word." In four cases, the Negex window size was not long enough to recognize that all the terms after a negation trigger word were negated. In two instances, ADER recognized diseases that were mentioned as part of the possible differential diagnosis. The remaining misidentified CMs were mentioned as past diagnoses that were not relevant; for example, one patient had been diagnosed with *gout* in the past, but was not having a "gout flare" at the time of admission.

Of the lab CMs, 97% were judged to be accurate based on available lab values; the one incorrectly classified laboratory CM in the sample was hyperkalemia. In that case, the serum potassium value was elevated, but the specimen was listed as being "hemolyzed" and therefore unreliable. ADER did not check whether samples were tainted and later retested. Table 30 lists the estimated precision for each concept type.

Table 30. ADER NLP precision from 100 random H&P notes generating alerts.

Type	Total	# Accurate	% Accurate
Medications	157	154	98%
CM (text)	136	116	85%
CM (lab)	30	29	97%

Survey Questionnaire Results

Of those 927 notes with detected ADEs, 221 alerts linked to the original long-form survey and 706 included the newer short-form survey. Only 28 study participants responded to the original, longer survey – a response rate of only 13%. Participants responded to the new shorter questionnaire 377 times, a response rate of 53%. The overall response rate for either questionnaire was 44%. The aggregated responses to the original questionnaire are shown in Table 31. The responses to the short-form questionnaire are shown in Table 32.

According to responses to the original-format survey, 80% of potential ADE CMs identified by ADER were being experienced by the patients shortly on or before admission; only 10% (five ADEs) were marked as definitely not being experienced by the patient. Physicians responded that the patient was taking 92% of the suspected ADE-causing medications identified by ADER. Physician respondents were already aware of 81% of the ADEs and were not aware of 7% of the ADEs. Respondents believed that the drug was "likely" or "very likely" causing or exacerbating the CM in 40% of cases; in 50% of cases, respondents believed that it was "unlikely" or "very unlikely." In 44% of cases, physicians planned to hold the medication.

According to the short-form questionnaire completed for 377 ADER alerts, 50% of alerts merited at least passing consideration for the patient. A further 24% believe the alerts "somewhat" merited consideration. A total of 58% of alerts were considered either "helpful" or "possibly helpful." Nearly one third of respondents planned to (or already did) change the patient's therapy due to the listed possible ADEs. The margin of error for the survey is plus or minus 5%.

Table 31. Responses from the long-form survey questionnaire (n = 28).

Did the patient have CM shortly before or upon admission?		
Yes 80% (40)	**Don't Know** 10% (5)	**No** 10% (5)
To the best of your knowledge, was the patient taking DRUG recently prior to admission?		
Yes 92% (40)	**Don't Know** 8% (5)	**No** 0% (0)
Were you aware that DRUG could cause or exacerbate CM?		
Yes 81% (44)	**Not Applicable** 11% (6)	**No** 7% (4)
Do you believe DRUG was causing/exacerbating CM in this case?		
13% (7)	**Very Likely**	
27% (15)	**Likely**	
11% (6)	**Not applicable**	
35% (19)	**Unlikely**	
15% (8)	**Very Unlikely**	
If applicable, what course of action have you taken (or do you plan to take) regarding Drug?		
44% (14)	**Hold the medication**	
28% (9)	**No change, continue to monitor the side effect**	
6% (2)	**Change the medication dosage**	
6% (2)	**Other (please explain in comments below)**	
6% (2)	**No response.**	
3% (1)	**No change, no monitoring required**	
3% (1)	**Continue medication while initiating therapy to treat side effect**	
3% (1)	**Discontinue the medication**	

Table 32. Responses from the short-form survey questionnaire (n = 377).

Did any of these alerts merit at least passing consideration for this patient?		
Yes 50% (189)	**Somewhat** 24% (91)	**No** 26% (97)
Were any of these alerts helpful in managing this patient?		
Yes 31% (118)	**Possibly** 27% (100)	**No** 42% (159)
Will you (or did you already) change patient's therapy due to these possible ADEs?		
Yes 32% (122)	**Uncertain** 12% (47)	**No** 55% (208)

Both questionnaire formats allowed for optional comments from the study participants. The short form and long-form surveys together generated a total of 52 comments. The project team analyzed each comment, when necessary comparing survey responses against the original H&P notes that triggered the alert.

Approximately 30% of the comments were positive, indicating that ADER was correct in asserting the medication was causing the recognized ADEs. Many of the comments stated that ADER was helpful in recognizing potential ADEs; several comments indicated that providers had recognized the ADE before seeing the alerts. Another 30% of the comments stated that the provider suspected the potential-ADE was not due to a medication, but because of the patient's medical condition. In the majority of those cases, providers indicated that the patients were suffering from end stage renal disease and that the ADER-identified *renal insufficiency* or *acute kidney injury* was likely not the result of medication administration. It is possible, however, that the medications were exacerbating the recognized CMs in some of those cases.

In approximately 20% of the comments, providers indicated that the medications were held, but not because of the ADER-recognized ADEs. In 10% of the comments, providers indicated that the patients were not suffering from the CM. In one of those cases, ADER had recognized "gout" as a possible ADE; while the patient had been previously diagnosed with gout, they were not having an active attack at admission. In one case, ADER incorrectly identified that the patient had a cough, but the sentence from which it was extracted stated "Hurts with movement, cough, sneezing, position changes."

In one instance, the provider commented, "The patient does not have acute kidney injury or hypovolemia, nor hyperkalemia." However, at the time of admission, the note stated that the patient was hypovolemic and lab results indicated a serum potassium of 6.5, a creatinine of 11.9, a BUN of 72. In another comment, a provider indicated that lightheadedness was not the patient's chief complaint, however the admission note listed "significant lightheadedness" in the *History of Present Illness* section. In at least two cases, providers responded that they did not make a change due to the suspected ADEs, but the H&P *Plan* section indicated they had held the medications for the reasons identified by ADER.

Comparing Admission Medications to Inpatient Medication Orders

For the pilot implementation group, the team retrieved approximately 118,000 inpatient medication orders corresponding to the 3103 admissions. After removing all non-antihypertensive medication orders, a total of 10,908 orders remained. The retrospective control group records contained approximately 282,000 medication orders from 9249 admissions, 24,754 of which were for antihypertensive medications. Using the medication order parameters "start time," "discontinue time," "hold start time," and "resume time," the team identified whether there was an active order for each admission medication during every 12-hour period from admission to discharge for each patient. The team identified any medication holds, discontinuations, or dosage reductions during each 12-hour period.

Table 33 illustrates the rate of medication holds, discontinuations, and medications not ordered during the first 12, 24, and 48 hours after admission for both the pilot and control groups. For each time period, there was no significant differences in rate at which suspected ADE-causing medications were given between the control group and the experimental group (Appendix B). There was also no significant difference between groups in the rate at which suspected ADE-causing medications dosages were reduced.

Approximately 10% of suspected ADE-causing medications had "conditional hold" instructions, most of which included instructions not to administer the medication if the systolic blood pressure was less than some threshold. The rate was the same in both the pilot and control groups. Since the analysis could not know whether the medication was given from the inpatient medication orders, conditional holds were not treated as holds.

Table 33. Suspected ADE-causing admission medications compared to inpatient medication orders in the first 12, 24, and 48 hours after admission, for both the pilot and the control group. For all hypothesis tests, the null hypothesis was that the pilot rate was not equal to the control group rate with alpha=0.05.

	ADE-causing Drugs	No active orders for medication after admission for at least…					
		12 hours		24 hours		48 hours	
		Number	Rate	Number	Rate	Number	Rate
Pilot	1517	1054	70%	923	61%	887	59%
Control	4525	3229	71%	2890	64%	2773	61%
Comparison		No significant difference.		No significant difference.		No significant difference.	
	ADE-causing Drugs	**Medication dosage decreased after admission in the first…**					
		12 hours		**24 hours**		**48 hours**	
		Number	**Rate**	**Number**	**Rate**	**Number**	**Rate**
Pilot	1517	45	3.0%	59	3.9%	76	5.0%
Control	4525	142	3.1%	176	3.9%	209	4.6%
Comparison		No significant difference.		No significant difference.		No significant difference.	
	ADE-causing Drugs	**No active orders or medication decreased in first…**					
		12 hours		**24 hours**		**48 hours**	
		Number	**Rate**	**Number**	**Rate**	**Number**	**Rate**
Pilot	1517	1099	72%	982	65%	963	64%
Control	4525	3371	75%	3066	68%	2982	66%
Comparison		No significant difference.		No significant difference.		No significant difference.	

The project team further analyzed patterns in inpatient medication orders. As shown in Table 34, the analysis revealed that the majority of admission medications in both the pilot and control groups followed one of four patterns during the inpatient stay: *never* – the medication was never ordered or was held throughout the stay; *always* – there was an active order for the medication throughout the entire stay; *started then stopped* – the medication was given initially, but was held or discontinued at some point before discharge; and *stopped then started* – the medication was not ordered or held initially, but was started at some point before discharge. That 40% of medications were never ordered and another 15% of medications were not given at the beginning of the admission, may contribute to high rate at which ADE-causing medications were not given above in Table 33.

Table 34. Patterns found in inpatient medication orders for all pre-admission antihypertensive medications.

	Pilot Group		Control Group	
Type	**# Drugs**	**Pct.**	**# Drugs**	**Pct.**
Never Ordered	1400	37%	4419	39%
Always Ordered	1334	35%	4007	36%
Stopped, then Started	593	16%	1618	14%
Started, then Stopped	206	5%	570	5%
Other	282	7%	667	6%
Total	**3815**	**100%**	**11281**	**100%**

While the project team intended for providers to see ADER alerts in the CMR when they were going on rounds, the only time it is certain providers saw the alert is when they responded to the survey questionnaire. Since nearly 40% of antihypertensive medications were not ordered during the entire hospital stay, the project team analyzed the data to determine the number of alerting medications that had active orders at the time of survey response. Table 35 illustrates the percentage of alerting medications in the pilot group with active orders at the time of the survey responses, as well as those suspected ADE-causing medications in the control group. In the pilot group, 262 ADE-causing medications had an active order at the time providers responded to the alert. The rate at which those medications are held or discontinued *after* providers responded to the alert was 17%. Since the control group did not actually receive ADER alerts, the analysis used the median response time from the pilot group (24 hours). The rate at which providers changed medications in the control group after that time was only 10% (18% ≠ 10%, p-value < 0.001, alpha=0.05).

Table 35. Alerting medications with an active order held or discontinued after the provider responded to the ADER survey. For the control group, the median response time *in the pilot group* was used.

All medications with active order discontinued or held after alert response			
	Alerting Drugs Held or D/C	**All Active Medication Orders at Time of Response**	**Pct. Drugs Held or D/C After Response**
Pilot	47	262	18%
Control	163	1620	10%
Compared			Pilot ≠ Control, p<0.001

Comparing Admission Medications to Discharge Medications

Out of the 927 notes that triggered an alert during the pilot study implementation period, the project team was able to retrieve 698 corresponding discharge notes. Team members determined that those that were missing were likely due to patients being admitted by interns and residents, but cared for and discharged by attending physicians and nurse practitioners. From the sample of 9,248 Internal Medicine admission notes processed in the retrospective pre-ADER control group, 2691 generated an alert (1711 of which had corresponding discharge notes). Project team members determined when medications were held, discontinued, or the dosage was decreased at discharge. Table 36 shows the rates of admission medications held, discontinued, or decreased at discharge for both the pilot group and the control group.

Table 36. Admission medications compared to discharge medications in notes with detected ADEs. "Omitted" medications were listed in the admission notes, but missing from the discharge notes; "D/C" means the medication was discontinued.

	Number of admission medications suspected of causing an ADE…			
	Total	**Omitted from Discharge**	**Held or D/C at Discharge**	**Dosage Decreased at Discharge**
Pilot	1141	112 (9.8%)	262 (23%)	55 (4.8%)
Control	2880	742 (26%)	161 (5.6%)	131 (4.5%)
At hospital discharge	**Rate of hold or D/C**	**Rate of dosage decreases**	**Rate of D/C, hold, or decrease**	**Rate of D/C, Hold, Decrease, or Omitted**
Pilot	23%	4.8%	28%	38%
Control	5.6%	4.5%	10%	36%
Comparison	Pilot ≠ Control, $p < 0.001$.	No significant difference.	Pilot ≠ Control, $p < 0.001$.	No significant difference.

The 698 admissions with corresponding discharge notes in the pilot group contained 1141 drugs suspected of causing an ADE, 1029 of which were mentioned in the discharge notes; 262 drugs were specifically mentioned as being held or discontinued at discharge and 55 were continued at discharge at a decreased dose. The rate of suspected ADE-causing drugs with either a reduced dose or specifically listed as being held or stopped was 28%.

The 1711 admissions with corresponding discharge notes in the control group contained 2880 medications suspected of causing an ADE, 2138 of which were mentioned in the discharge note; 161 (or 5.6%) were specifically listed as being held or discontinued at discharge, and 131 (or 4.5%) were continued at discharge at a decreased dose. The rate of suspected ADE-causing drugs with either a reduced dose or specifically held or discontinued was 10%.

The project team compared the rates of medication holds, discontinuations, and dosage decreases in the ADER pilot study population against the rates in the retrospective control group (see Table 36). The analysis used hypothesis testing for two proportions to compare the rates, testing the null hypothesis that the two rates were significantly different from one another. The rate at which suspected ADE-causing drugs have their dosage reduced or are specifically mentioned as being held is significantly greater in the ADER pilot study population compared to the rate in the control group (28%≠ 10%, p-value < 0.001, alpha=0.05) (Appendix B).

As stated above, 112 of 1141 admission medications were missing from the corresponding discharge notes in the pilot group and 742 of 2880 were missing from the control group. If all of those omitted medications were actually continued, the rate changes would be the same as that shown above. However in the unlikely event that all omitted medications were actually discontinued, the rate would increase to 38% in the pilot group and 36% in the control group. The rate at which suspected ADE-causing drugs have their dosage reduced, are specifically mentioned as being held, *or are not included in the discharge note* is not significantly different in the two groups (see Table 36 and Appendix B). In summary, there is statistically significant difference between rates of medications specifically listed as being held, discontinued, or decreased when ADER issued warnings; in the unlikely event that all admission medications omitted from the discharge notes were actually discontinued, there would be no significant difference between rates.

Medication Holds Analyzed by ADE

The project team analyzed the rate at which suspected ADE-causing medications were held based upon the suspected ADE in question. For each ADE CM, the team calculated the rate at which the suspected ADE-causing medications were held or

discontinued in discharge notes and the rate at which they were not given in the inpatient medication orders. Table 37 shows all ADE CMs in decreasing order of frequency.

Table 37. ADEs detected during the pilot study and their hold rates.

ADE Clinical Manifestation	Number of alerts with ADE	Rate of any drug suspected of causing ADE being	
		Not given during first 24 hours inpatient	Held or Discontinued at Discharge
Renal insufficiency	182	72%	31%
Lightheadedness	177	72%	35%
Hypotension	160	90%	39%
Dizziness	147	65%	35%
Edema	138	46%	19%
Syncope	100	62%	37%
Coughing	75	57%	13%
Hyperkalemia	47	83%	46%
Constipation	46	46%	21%
Hypokalemia	38	66%	31%
Headache	38	45%	19%
Bradycardia	37	73%	26%
Gout	35	69%	25%
Pancreatitis	33	55%	19%
Hypovolemia	30	90%	46%
Hyponatremia	28	86%	50%
Exanthema	27	67%	0%
Dehydration	26	73%	17%
Alkalosis	26	23%	11%
Hyperglycemia	21	48%	19%
Acute kidney injury	17	100%	40%
Orthostatic Hypotension	14	50%	33%
Peripheral edema	10	40%	0%
Confusion	9	67%	17%
Hearing loss	7	29%	33%
Polyuria	6	50%	33%
Cardiogenic Shock	5	100%	67%
Hearing impairment	5	60%	25%
Muscle Cramp	5	60%	0%
Frequent Urination	4	100%	33%
Anorexia	4	75%	33%
Tinnitus	4	25%	25%
Vomiting	4	0%	25%
Somnolence	3	100%	0%
Hyperuricemia	3	67%	33%
Pruritus	3	67%	50%
Heart Block	2	100%	100%
Erectile dysfunction	2	50%	0%
Gynecomastia	1	100%	0%
Mental Depression	1	100%	0%
Pericardial effusion	1	100%	0%
Raynaud Disease	1	0%	0%
Myalgia	1	0%	0%

Suspected ADE-Causing Medications Compared to All Medications

Comparing admission medications to both discharge medications and inpatient medication orders revealed there was a significant difference between the hold rate of drugs suspected of causing ADEs by ADER compared to those not suspected of causing ADEs. Table 38 shows this comparison.

When comparing the admission medications to discharge medications in the pilot study group, the team found that 23% of suspected ADE-causing antihypertensive medications were held or discontinued versus 17% of antihypertensive medications not suspected of causing an ADE ($p < 0.001$) (Appendix B). When comparing the rate of suspected ADE-causing medications held, discontinued, or decreased, the difference is greater at 28% versus 20%. Even if one considers admission medications omitted from the discharge note, the team found that 37% of suspected ADE-causing medications were held, *omitted*, or decreased versus 27% of antihypertensive medications that were not suspected of causing an ADE ($p < 0.001$) (Appendix B).

Table 38. Comparing medications suspected of causing an ADE versus those not suspected of causing an ADE in the pilot group. The table illustrates the rate of holds, discontinuations, and dosage decreases in discharge notes and medication orders for inpatients.

Pilot Pre-Admission Medications Compared to Discharge Medications							
	ADE Drugs	**Held or Discontinued at Discharge.**		**Dosage Decreased at Discharge.**		**Held, Discontinued, or Decreased at Discharge.**	
	Num.	**Num.**	**Pct.**	**Num.**	**Pct.**	**Num.**	**Pct.**
ADE	1141	262	23%	55	4.8%	317	28%
Non-ADE	1686	287	17%	43	2.6%	330	20%
Compared			Pilot > Control, p<0.001		Pilot > Control, p<0.002		Pilot > Control, p<0.001
Pilot Pre-Admission Medications Compared to Inpatient Medication Orders							
	ADE Drugs	**Not ordered 12 hours post-admission.**		**Not ordered 24 hours post-admission.**		**Not ordered 48 hours post-admission.**	
	Num.	**Num.**	**Pct.**	**Num.**	**Pct.**	**Num.**	**Pct.**
ADE	1517	1054	70%	923	61%	887	59%
Non-ADE	2299	1356	59%	1151	50%	1101	48%
Compared			Pilot > Control, p<0.001		Pilot > Control, p<0.001		Pilot > Control, p<0.001

When comparing the admission medications to inpatient medication orders in the pilot study group, the team found that in the first 12 hours after admission, 70% of suspected ADE-causing medications were not ordered versus 59% of medications that were not suspected of causing an ADE. In the first 24 hours after admission, 61% of suspected ADE-causing medications were not ordered versus 50% of medications that were not suspected of causing an ADE. Similarly, 59% of suspected ADE-causing medications were not given in the first 48 hours versus 48% of medications that were not suspected of causing an ADE. Similar differences were present in the control group and when considering inpatient dosage decreases. In all cases, the rate of medication holds was greater among medications ADER identified as causing ADEs ($p < 0.001$). (Appendix B)

Survey Responses Compared to Discharge Medications and Inpatient Medication Orders

As noted above, one of the questions included with ADER alerts was "Will you (or did you already) change patient's therapy due to these possible ADEs?" A total of 32% of survey respondents said they made some change due to potential ADEs included in that ADER alert. Of those 122 alerts, 110 had some detectable change made to one or more suspected ADE-causing medications (52 had drugs held or decreased at discharge, 110 had drugs not given or the dosage reduced during the first 24 hours after admission). That is, for those providers who responded that they would change the patient's therapy due to the identified potential ADEs, the analysis identified holds or dosage decreases in 90% of cases. Of the 293 medications included in those 122 alerts, 83% were held or decreased at some point. Of the 47 providers who responded they were uncertain whether they would make a change, the analysis identified 44 alerts (or 94%) where one or more ADE-causing medications were held or changed (or 83% of the 131 ADE-causing medications).

Each ADER alert contained a survey questionnaire. The project team compared the medication hold/change rate for alerts where providers responded to the survey questionnaire versus those where providers did not respond to the questionnaire.
The project team analyzed the hold/change rate of medications suspected of causing ADEs

For the 406 survey responses, 37% had at least one suspected ADE-causing medication held or the dose decreased at discharge. Those alerts consisted of 480 medications, 30% of which were held or decreased at discharge. For the 522 ADER alerts

where providers did not respond to the survey questionnaire, 30% had at least one suspected ADE-causing medication held or decreased at discharge. Those 522 alerts consisted of 661 medications, 24% of which were held or decreased at discharge. In terms of both number of medications held and the number of alerts with at least one medication held, the rate was higher for those patients whose providers responded to the alert ($p < 0.05$).

Discussion

The ADER pilot ran for three months at Vanderbilt University Hospital. During that time, 75% of those who responded to the ADER alert questionnaire indicated the alerts merited or somewhat merited consideration for the patient. Almost 60% of respondents said the alerts were helpful or "possibly helpful" in managing the patient, and 32% of providers said they were going to change (or had already changed) the patient's therapy due to the ADEs detected by ADER. Based on discharge summaries, the rate of suspected ADE-causing medications held or discontinued was 27% when providers received an ADER alert compared to only 9.5% in the control group. The ADER pilot revealed the ADER system accurately detects likely ADEs from EMR notes and can measurably help providers recognize potential ADEs presenting at admission resulting from outpatient medication.

ADER System Performance

It took on average less than twelve minutes from the time when residents and interns saved admission H&P notes to the time when ADER generated alerts and sent them to StarPanel. Over the course of the pilot, the system was down for only 1% of the time and no admission notes were dropped. The median response time to the alert questionnaire (from alert generation to alert response) was 24 hours. The project team believes this was due to the ADER alert location in the daily progress note form; daily progress notes typically are written the day after admission. The system scanned almost 240 notes per week for a total of 3103 notes from 2708 unique patients and detected potential ADEs in approximately 30% of notes. This is within the bounds of previously published estimates that the rate of unrecognized ADEs is as high as 30%,[35] though it is important to recognize

that ADER detects *potential* ADEs, not necessarily ones with definite causality in the patient.

Comparability of Historical Control Group

As is shown in Table 29, the pilot study group and the control group of historical H&Ps were very similar. First, each set consisted solely of internal medicine admission H&Ps written by interns and residents; since internal medicine residents are in residency for three years, many individuals in the pilot are likely also in the control group. In both groups, there is an average length of stay of 114 hours and ADER recognized potential antihypertensive ADEs in approximately 30% of notes. Notes with detected ADEs in both groups were found to have an average 2.4 ADEs, approximately 15% of which were found from laboratory results and 85% from the H&P text. In both groups, it was also the case that when ADEs were detected, the patient was on twice as many antihypertensive medications as when no ADEs were detected. It is well known that when patients are on more medications, they are at higher risk of experiencing an ADE.[139,140]

ADER NLP and Concept Recognition

The project team found that ADER NLP was accurate. The analysis of 100 H&P notes indicated the precision for medications was 97% and the precision for CMs recognized from text was 86%. While only 13% of users responded to the longer original-format survey, the results corroborated that ADER was accurate in identifying antihypertensive medications (92%) and CMs (80%). Those CMs misidentified as present by ADER were mostly due to errors in negation detection.

Physician Awareness of ADEs

While there were few responses to the original longer-format questionnaires, results indicated providers were already aware of over 80% of the ADEs recognized by ADER; this was corroborated by many of the optional survey comments. This is likely because the pilot study focused on antihypertensive medications and most internal medicine physicians are very familiar with that drug class. Additionally, the most frequent ADE CMs were hypotension, lightheadedness, dizziness, and syncope (Table 27).

Lightheadedness, dizziness, and syncope are all symptoms of hypotension. A total of 40% of respondents to the original-format survey indicated they believed that it was likely that the drug in question was causing or exacerbating the CM.

Survey Questionnaire Responses

When the survey format was changed, the response rate increased from 13% to 53% (overall response rate of 44%). The project team believed the combination of fewer questions and the fact that the questions were presented in the alert panel contributed to higher response rate. The project team collaborated closely with the Internal Medicine chief resident during the pilot implementations and believe his actions as a "champion" in promoting the ADER study and encouraging residents to complete the questionnaires contributed to this high response rate.

Survey responses from the new-format survey revealed that 74% of users believed the alerts merited at least some consideration. A third of users felt that the alert was helpful in managing their patients and another third felt the alert was "possibly" helpful. Finally, 32% of respondents indicated that they would change the patient's therapy, or they had already, due to the ADEs listed in the ADER alert. For those providers who responded that they would change the patient's therapy due to the identified potential ADEs, the analysis was able to identify corresponding holds or dosage decreases in 90% of cases.

While there is no evidence that any providers missed the alerts in StarPanel, the project team can be certain that those providers who responded to the questionnaire actually noticed and read the alert. The analysis found that 37% of providers who responded to the alert held or changed suspected ADE-causing medications at discharge versus only 30% for providers who did not respond. Providers might be more likely to respond if they believed the ADEs was true, but approximately 50% or respondents said they were not planning on making a change due to suspected ADEs. The project team believes it is likely the difference is attributable to the fact that they took the time to read and consider the ADER warnings.

The analysis of provider comments submitted with the questionnaires revealed that 30% of those providers who submitted comments agreed with ADER – that is, they believed ADER was correct that the medication was causing the patients systems. Another

30% of commenters believed that the CM was actually being caused by the patient's illnesses. In those cases, most still indicated the alerts merited consideration, as the medications could have been exacerbating the conditions. At least two providers who left comments responded that they did not make a change due to the suspected ADE, but indicated in the *Plan* section that they did. Several comments were useful in identifying the conditions where ADER incorrectly recognized CMs as present in the patient, mostly due to problems with negation detection.

Comparing Admission Medications to Discharge and Inpatient Medications

For both the pilot and control groups, the project team analyzed the rate at which suspected ADE-causing medications were not given or decreased during the inpatient stay. However, the rates were the same across both groups at all time periods. The median response time for the electronic alert was 24 hours, and multiple study participants informed the team they did not notice the ADER alerts in the CMR paper form. Since half of the survey respondents answered the questionnaire after that time, it is reasonable to believe that most providers became aware of the alert only a full day after admission. Additionally, since 44% of admission medications were never given during the inpatient stay. This could likely explain why inpatient medication orders didn't reveal a difference in the rate of holds or changes between the pilot study group and the control group.

However, the analysis did find that for those providers who definitely saw the alert (that is, the ones who responded to the questionnaire), they were more likely to hold or discontinue suspected ADE-causing medications after responding than those in the control group the same amount of time after admission.

The team also evaluated whether admission medications in both groups where held, discontinued, or decreased at discharge. The rate at which suspected ADE-causing drugs have their dosage reduced or are specifically mentioned as being held or discontinued is significantly greater in the ADER study population (28%) versus the control group that did not receive ADER alerts (10%).

A significant number of admission medications were not mentioned in corresponding discharge summaries. If one considers medications omitted from the discharge notes as all being held, the rates between the control group and pilot group are

the same. However, since the medications were not mentioned anywhere in the note, it is not possible to know for certain what, if any, action was taken. Like all clinical notes, discharge summaries can vary in quality and thoroughness. The medications could have been held or they could have been continued at the current dose.

Medication Holds Analyzed by ADE

Among the most frequent ADEs resulting in a high rate of medication holds or decreases were those that were could obviously be exacerbated by antihypertensive medications – hypotension, lightheadedness, syncope, and dizziness. Those medications suspected of causing serious ADEs such as hyperkalemia, hypokalemia, hyponatremia, and hypovolemia were held or discontinued at discharge at rates of 46%, 31%, 50%, and 46%, respectively. ADEs suspected of causing or exacerbating renal insufficiency and acute kidney injury were held or changed at discharge 31% and 40% of the time, respectively. Potentially less serious ADEs, such as cough or headache were held or discontinued at discharge at lower rates: 13% and 19%, respectively.

Suspected ADE-Causing Medications Compared to All Medications

The rate of holds, discontinuations, or decreases for medications identified by ADER as causing or exacerbating ADE CMs was significantly greater than for those medications that were not been recognized as potentially causing or exacerbating ADEs. This is true when looking at discharge summaries and inpatient medication orders for both the pilot group and the control group. These rates illustrate that ADER successfully recognizes antihypertensive medications that are likely to be causing ADEs and those that are more likely to be held or discontinued by providers.

Survey Responses Compared to Discharge and Inpatient Medications

Similarly, when survey respondents said they had changed or planned to change the patient's current therapy due to recognized potential ADEs, the analysis identified some changed being made 90% of the time. This illustrates that ADER alerts can recognize when medication changes should be made. Additionally, it supports the validity of the alert

questionnaire responses when providers indicated that they would change the patient's medications.

Limitations

This pilot study illustrated the ADER system and its associated methodology can successfully detect potential ADEs in recently admitted inpatients. However, there were a number of limitations to the study. First, the pilot study was restricted to internal medicine interns and residents. Patients admitted to the Internal Medicine service at Vanderbilt were an appropriate general test of ADER, but the system should also be tested in other departments and subspecialties. Interns and residents were enrolled in the study both as a convenient sample and to test the ADER system as an educational tool for increasing clinician knowledge and awareness of ADEs, but future versions should be evaluated on all physicians, as well as nurse practitioners.

Second, the ADER pilot was limited to antihypertensive medications. The project team chose to target antihypertensives because they cause a wide variety of ADEs and are commonly used in internal medicine. However, because they are so familiar to internal medicine physicians, study participants reported to the project team that they were aware of most antihypertensive ADEs and frequently recognized the detected ADEs before they received ADER alerts. Many participants suggested that ADER would be most useful for medications that they prescribe less frequently. In the future, ADER should be expanded to additional drugs and drug classes.

Next, the ADER study analyzed three months of pilot data compared against a retrospective control group. While the analysis showed the control group to be comparable to the pilot group in nearly all respects, the groups contained different participants (with some overlap) and the admissions occurred at different times. A randomized controlled trial, where the same physicians are in both the pilot and control group at the same time, would be a more effective test. Also, this study analyzed inpatient medication orders to determine provider actions, but it was sometimes difficult to determine if a medication was actually given at a specific time (i.e., when there was a conditional hold in the medication instructions). Future studies would benefit from analyzing medication administration records to determine more precisely when medications were given.

Finally, the ADER system used as input the results of NLP on clinical notes. Therefore the quality of NLP results significantly influences any outcomes. While analyses showed that ADER was 86% accurate at identifying ADE CMs and 97% accurate at identifying current medications, improvements can still be made. Specifically, the analysis revealed that failures to correctly recognize negation accounted for nearly three quarters of misidentified CM concepts. Improvements to clinical NLP would improve the accuracy of the ADER system.

Conclusion

During the three-month pilot at Vanderbilt University Hospital, the ADER system successfully identified potential ADEs presenting in patients at admission. Based on questionnaires included with alerts, 75% of providers believed alerts merited or somewhat merited consideration for the patient, 60% believed alerts were helpful or "possibly helpful" in managing their patients, and 32% made changes to the patients therapy because of ADEs recognized by the system. Based on discharge summaries, the rate of suspected ADE-causing medications held, discontinued, or decreased was 28% when providers received an ADER alert compared to only 10% in the historical control group that did not receive alerts. Based on inpatient medication orders, providers were more likely to hold or discontinue medications after they responded to alerts than providers in the control group after a similar amount of time. Importantly, the study showed that ADER correctly identifies those medications that are more likely to be held by physicians both during the inpatient stay and at discharge. The ADER system accurately detects likely ADEs from EMR notes and can measurably help providers recognize potential ADEs presenting at admission. This study strongly supports further development and implementation of the ADER system for patient care.

CHAPTER VII

SYNOPSIS & CONCLUSIONS

Synopsis

Unrecognized ADEs pose a serious clinical problem causing substantial morbidity and mortality. This dissertation addresses important aspects of this problem through the development of the Drug Evidence Base and the Adverse Drug Effect Recognizer.

As described previously, DEB2 is an accurate, machine-processable drug knowledgebase, automatically derived and updated from NDF-RT, MEDLINE, MedlinePlus, SIDER2, and Drugbank. The project team determined that DEB2 is 86% accurate overall, with indications 88% accurate and ADEs 84% accurate. Drug knowledge available through DEB2 can support many areas of biomedical research.

The ADER system provides clinical decision support to help providers recognize potential ADEs in recently admitted inpatients. The system uses NLP on admission H&P notes to identify patients' medications and CMs and compares those against a set of known ADEs derived from DEB2. If potential ADEs are detected, ADER alerts providers though the EMR. This dissertation described a three-month pilot study of the ADER system at Vanderbilt University Hospital. Based on responses to survey questions included in ADER alerts, 75% of providers believed alerts merited or somewhat merited consideration for the patient, 60% believed alerts were helpful or "possibly helpful" in managing their patients, and 32% changed their patient's medication therapy because of those ADEs detected by ADER. Compared to a retrospective control group, the study found that providers were more likely to hold or discontinue medications after they responded to ADER alerts and that they held or discontinued medications at discharge at a higher rate. The pilot study results strongly support further development, evaluation, and implementation of the ADER system.

Future Work

While DEB2 is a significant improvement over DEB1, there are still opportunities for future enhancements. As described in Chapter II, drug knowledgebases must take into account complex drug-CM relationships. For example, DEB2 identified the drug Metoprolol as indicated for ventricular fibrillation; this is because metoprolol is indicated for *prevention* of ventricular fibrillation, but *only* in patients who have recently suffered a myocardial infarction. DEB2 also found that lymphoma is an ADE of the drug sirolimus; lymphoma is a reported complication of immunosuppression by sirolimus, but sirolimus is also indicated in lymphoma patients after a bone marrow transplant. Future versions of the DEB should develop ways extract and represent these and other complex drug-CM relationships.

Future version should include new relationship in addition to indications and ADEs. For example, indications could be divided into preventive indications (drug prevents CM) and treatment indications (drug treats CM). Similarly, adverse drug effects could be further divided into classes such as common ADEs, rare ADEs, serious ADEs, or overdose effects. Deriving and cataloging contraindications and potential drug-drug interactions would also be useful. Currently, DEB2 only includes single-ingredient drugs. Future versions should expand the knowledgebase to include multiple-ingredient drugs.

One of the major problems with DEB1 was that similar CM concepts were represented using different UMLS concepts. This resulted in multiple indications and ADEs that appeared to be from only one source, instead of single indications or ADEs corroborated by different sources. Representing CM concepts using the SNOMED-CT vocabulary, instead of the complete UMLS, resolved many instances of this problem. However, future versions of the DEB would benefit from techniques that normalize similar CMs into unified concepts.

The authors plan to publicly release DEB2 for use by other researchers. The knowledgebase will be updated regularly when new versions of its component sources are released. Additionally, the authors intend to create a version of DEB2 using the FDA's preferred MedDRA terminology in addition to SNOMED-CT.

The ADER system was limited in multiple ways for the purposes of the pilot study. For example, the current ADER Alerting Subset consists only of ADEs from

antihypertensive medications. In the future, ADER should be expanded to additional drug classes and use additional laboratory tests for input. Future versions of ADER should take into account other variables known to affect patients' likelihood of experiencing certain ADEs, such as age, race, gender, medication dose, and pharmacogenomic data. After expanding DEB2, ADER could also be enhanced to detect drug-drug interactions in patients who are experiencing the actual effects of that interaction.

The ADER pilot study was restricted to internal medicine interns and residents. Future versions should be evaluated for providing decision support for physicians and nurse practitioners from other departments and subspecialties. Additionally, future evaluation of ADER would benefit from a randomized controlled trial, where the same physicians are in both the pilot and control group at the same time. Future studies should also analyze medication administration records, in addition to inpatient medication orders, to better recognize provider actions post-alert. Finally, the authors believe that the response to ADER alerts would be improved if alerts could be generated more quickly and alert information was included in the CPOE system, in addition to the Team Summary and Progress Note form.

Conclusion

In conclusion, this dissertation developed useful tools and methodologies to address a serious clinical problem. Building on prior work by the authors as well as others, this dissertation developed the DEB2 drug information knowledgebase and the ADER clinical decision support system. The DEB2 knowledgebase can support research in the areas of drug repurposing, pharmacovigilance, clinical data mining, phenotyping, and clinical decision support systems. The ADER system has the potential to improve both recognition and treatment of ADEs in a generalizable manner. This dissertation demonstrated the accuracy of DEB2 and promising results from the ADER pilot implementation at Vanderbilt.

This dissertation project developed innovative methods to compile and represent drug information and demonstrated the potential for real-time ADE detection in patient populations. Since ADER uses standard admission H&P text, as well as widely used diagnostic tests, the system should be generalizable to other hospitals using electronically stored H&Ps and laboratory values. Further development and additional validation of the

ADER system could improve both recognition and treatment of ADEs in a manner applicable across the US. Addressing previously unrecognized ADEs has the potential to reduce costs and improve patient care.

APPENDIX A

IRB APPROVAL LETTER

The following email is the approval letter from the Vanderbilt Institutional Review Board.

Subject: IRB Letter Notification re: IRB#141341 - New Study - Approve
From: Gail.Mayo@vanderbilt.edu <Gail.Mayo@vanderbilt.edu>
Date: Thu, Oct 16, 2014 at 10:32 AM
To: Joshua C. Smith <joshua.smith@vanderbilt.edu>,
Randolph A. Miller <randolph.a.miller@vanderbilt.edu>

STUDY TITLE: The Adverse Drug Effect Recognizer (ADEr) Project

IRB NUMBER: 141341
PI: Joshua Smith
SUBMISSION TYPE: New Study
DATE REQUESTED: 8/21/2014
REVIEW STATUS: Approve

The IRB's determination regarding the above referenced study submission is complete.

Please click the link below to log in to DISCOVR-E and view the letter.

If you have any questions, please contact the IRB at 322-2918.

Thank you.

APPENDIX B

STATISTICAL ANALYSES FROM THE ADER PILOT STUDY

Analysis: Differences in rates at which suspected ADE-causing medications were not given within 24 hours of admission between the control group and the experimental group.

XLSTAT 2015.6.01.25742 - Tests for two proportions - Start time: 2/18/2016 at 4:46:44 PM
Frequency 1: 900
Sample size 1: 1357
Frequency 2: 2858
Sample size 2: 4460
Hypothesized difference (D): 0
Variance: p1q1/n1+p2q2/n2
Continuity correction: Yes
Significance level (%): 5

z-test for two proportions / Two-tailed test:

95% confidence interval on the difference between the proportions:
] -0.506 , 0.551 [

Difference	0.022
z (Observed value)	1.492
z (Critical value)	1.960
p-value (Two-tailed)	0.136
alpha	0.05

Test interpretation:
H0: The difference between the proportions is equal to 0.
Ha: The difference between the proportions is different from 0.
As the computed p-value is greater than the significance level alpha=0.05, one cannot reject the null hypothesis H0.
The risk to reject the null hypothesis H0 while it is true is 13.57%.

Analysis: Differences in rates at which suspected ADE-causing medications were held, discontinued, or decreased at discharge between the control group and the experimental group.

XLSTAT 2015.6.01.25742 - Tests for two proportions - Start time: 2/16/2016 at 11:43:20 PM
Frequency 1: 317
Sample size 1: 1141
Frequency 2: 292
Sample size 2: 2880
Hypothesized difference (D): 0
Variance: p1q1/n1+p2q2/n2
Continuity correction: Yes
Significance level (%): 5

z-test for two proportions / Two-tailed test:

95% confidence interval on the difference between the proportions:
] -0.352 , 0.705 [

Difference	0.176
z (Observed value)	12.207
z (Critical value)	1.960
p-value (Two-tailed)	< 0.0001
alpha	0.05

Test interpretation:
H0: The difference between the proportions is equal to 0.
Ha: The difference between the proportions is different from 0.
As the computed p-value is lower than the
significance level alpha=0.05, one should reject
the null hypothesis H0, and accept the
alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.01%.

Analysis: Differences in rates at which suspected ADE-causing medications were held, discontinued, decreased, or omitted at discharge between the control group and the experimental group.

XLSTAT 2015.6.01.25742 - Tests for two proportions - Start time: 2/16/2016 at 11:45:51 PM
Frequency 1: 429
Sample size 1: 1141
Frequency 2: 1034
Sample size 2: 2880
Hypothesized difference (D): 0
Variance: p1q1/n1+p2q2/n2
Continuity correction: Yes
Significance level (%): 5

z-test for two proportions / Two-tailed test:

95% confidence interval on the difference between the proportions:

] -0.516 , 0.550 [

Difference	0.017
z (Observed value)	0.967
z (Critical value)	1.960
p-value (Two-tailed)	0.333
alpha	0.05

Test interpretation:
H0: The difference between the proportions is equal to 0.
Ha: The difference between the proportions is different from 0.
As the computed p-value is greater than the significance level alpha=0.05, one cannot reject the null hypothesis H0.
The risk to reject the null hypothesis H0 while it is true is 33.34%.

Analysis: Differences in rates at which medications were held, discontinued, or decreased at discharge between alerting mediations and non-alerting medications in the pilot experimental group.

XLSTAT 2015.6.01.25742 - Tests for two proportions - Start time: 2/17/2016 at 12:31:42 AM
Frequency 1: 317
Sample size 1: 1141
Frequency 2: 330
Sample size 2: 1686
Hypothesized difference (D): 0
Variance: p1q1/n1+p2q2/n2
Continuity correction: Yes
Significance level (%): 5

z-test for two proportions / Two-tailed test:

95% confidence interval on the difference between the proportions:
] -0.450 , 0.614 [

Difference	0.082
z (Observed value)	4.959
z (Critical value)	1.960
p-value (Two-tailed)	< 0.0001
alpha	0.05

Test interpretation:
H0: The difference between the proportions is equal to 0.
Ha: The difference between the proportions is different from 0.
As the computed p-value is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.01%.

Analysis: Differences in rates at which medications were held, discontinued, decreased, or omitted at discharge between alerting mediations and non-alerting medications in the pilot experimental group.

XLSTAT 2015.6.01.25742 - Tests for two proportions - Start time: 2/17/2016 at 12:33:05 AM
Frequency 1: 429
Sample size 1: 1141
Frequency 2: 464
Sample size 2: 1686
Hypothesized difference (D): 0
Variance: p1q1/n1+p2q2/n2
Continuity correction: Yes
Significance level (%): 5

z-test for two proportions / Two-tailed test:

95% confidence interval on the difference between the proportions:
] -0.434 , 0.636 [

Difference	0.101
z (Observed value)	5.559
z (Critical value)	1.960
p-value (Two-tailed)	< 0.0001
alpha	0.05

Test interpretation:
H0: The difference between the proportions is equal to 0.
Ha: The difference between the proportions is different from 0.
As the computed p-value is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.01%.

Analysis: Differences in rates at which medications were not given within 24 hours of admission between alerting mediations and non-alerting medications in the pilot experimental group.

XLSTAT 2015.6.01.25742 - Tests for two proportions - Start time: 2/18/2016 at 4:52:35 PM
Frequency 1: 900
Sample size 1: 1357
Frequency 2: 1156
Sample size 2: 2055
Hypothesized difference (D): 0
Variance: p1q1/n1+p2q2/n2
Continuity correction: Yes
Significance level (%): 5

z-test for two proportions / Two-tailed test:

95% confidence interval on the difference between the proportions:
] -0.432 , 0.634 [

Difference	0.101
z (Observed value)	5.935
z (Critical value)	1.960
p-value (Two-tailed)	< 0.0001
alpha	0.05

Test interpretation:
H0: The difference between the proportions is equal to 0.
Ha: The difference between the proportions is different from 0.
As the computed p-value is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.01%.

REFERENCES

1. Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. Lancet. 2000 Oct 7;356(9237):1255–9.

2. International drug monitoring: the role of national centres. Report of a WHO meeting. World Health Organ Tech Rep Ser. 1972;498:1–25.

3. Nebeker JR, Barach P, Samore MH. Clarifying adverse drug events: a clinician's guide to terminology, documentation, and reporting. Ann Intern Med. 2004 May 18;140(10):795–801.

4. Research C for DE and. Information for Consumers (Drugs) - Inside Clinical Trials: Testing Medical Products in People [Internet]. [cited 2016 Mar 3]. Available from: http://www.fda.gov/Drugs/ResourcesForYou/Consumers/ucm143531.htm

5. Understanding Clinical Trials - ClinicalTrials.gov [Internet]. [cited 2012 Mar 23]. Available from: http://clinicaltrials.gov/ct2/info/understand

6. The Drug Development Process > Step 3: Clinical Research [Internet]. [cited 2016 Mar 8]. Available from: http://www.fda.gov/ForPatients/Approvals/Drugs/ucm405622.htm

7. Overview of Clinical Trials | CenterWatch [Internet]. [cited 2016 Mar 8]. Available from: http://www.centerwatch.com/clinical-trials/overview.aspx

8. Research C for DE and. Information for Consumers (Drugs) - The FDA's Drug Review Process: Ensuring Drugs Are Safe and Effective [Internet]. [cited 2016 Mar 3]. Available from: http://www.fda.gov/drugs/resourcesforyou/consumers/ucm143534.htm

9. Introduction to the New Prescription Drug Labeling by the FDA [Internet]. Medscape. [cited 2016 Mar 3]. Available from: http://www.medscape.com/viewarticle/566885

10. Structured Product Labeling Resources [Internet]. [cited 2012 Feb 8]. Available from: http://www.fda.gov/ForIndustry/DataStandards/StructuredProductLabeling/default.htm

11. Guidance for Industry Adverse Reactions Section of Labeling for Human Prescription Drug and Biological Products — Content and Format [Internet]. U.S. Department of Health and Human Services, Food and Drug Administration.; 2006 [cited 2016 Mar 3]. Available from: http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm075057.pdf

12. Physicians Total Care, Inc. SPL - Evista - raloxifene hydrochloride tablet [Internet]. [cited 2016 Mar 3]. Available from: http://dailymed.nlm.nih.gov/dailymed/fda/fdaDrugXsl.cfm?setid=01143c04-f5d7-481d-95bb-d384f2413585

13. Warner Chilcott (US), LLC. SPL -Atelvia - risedronate sodium hemi-pentahydrate and risedronate sodium monohydrate tablet, delayed release [Internet]. Available from: http://dailymed.nlm.nih.gov/dailymed/fda/fdaDrugXsl.cfm?setid=c8b9ab88-1a26-46c3-80ec-4eaa45202021

14. Roden DM. An underrecognized challenge in evaluating postmarketing drug safety. Circulation. 2005 Jan 25;111(3):246–8.

15. Wilson AM, Thabane L, Holbrook A, Wilson AM, Thabane L, Holbrook A. Application of data mining techniques in pharmacovigilance, Application of data mining techniques in pharmacovigilance. British Journal of Clinical Pharmacology, British Journal of Clinical Pharmacology. 2004 Feb 1;57, 57(2, 2):127, 127–34, 134.

16. Amery WK. Why there is a need for pharmacovigilance. Pharmacoepidemiology and Drug Safety. 1999;8(1):61–4.

17. WHO | Pharmacovigilance [Internet]. WHO. [cited 2012 Apr 2]. Available from: http://www.who.int/medicines/areas/quality_safety/safety_efficacy/pharmvigi/en/index.html

18. US Department of Health and Human Services. Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment [Internet]. 2005. Available from: http://www.fda.gov/downloads/regulatoryinformation/guidances/ucm126834.pdf

19. Research C for DE and. Questions and Answers on FDA's Adverse Event Reporting System (FAERS) [Internet]. [cited 2016 Mar 3]. Available from: http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/default.htm

20. EudraVigilance - Pharmacovigilance in EEA [Internet]. [cited 2012 Apr 3]. Available from: http://eudravigilance.ema.europa.eu/human/index.asp

21. Srba J, Descikova V, Vlcek J. Adverse drug reactions: Analysis of spontaneous reporting system in Europe in 2007-2009. European Journal of Clinical Pharmacology [Internet]. 2012 Feb 1 [cited 2012 Apr 9]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/22294060

22. Commissioner O of the. MedWatch: The FDA Safety Information and Adverse Event Reporting Program [Internet]. [cited 2016 Mar 3]. Available from: http://www.fda.gov/Safety/MedWatch/default.htm

23. Kessler DA. Introducing MEDWatch. A new approach to reporting medication and device adverse effects and product problems. Gen Hosp Psychiatry. 1993 Jun;16(2):96–101; discussion 102.

24. Almenoff JS. Innovations for the future of pharmacovigilance. Drug Saf. 2007;30(7):631–3.

25. FDA's Sentinel Initiative [Internet]. [cited 2016 Mar 3]. Available from: http://www.fda.gov/Safety/FDAsSentinelInitiative/ucm2007250.htm

26. Research C for DE and. Postmarket Drug Safety Information for Patients and Providers - Vioxx (rofecoxib) Questions and Answers [Internet]. [cited 2012 Mar 27]. Available from: http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm106290.htm

27. Timeline: The Rise and Fall of Vioxx : NPR [Internet]. [cited 2016 Mar 6]. Available from: http://www.npr.org/templates/story/story.php?storyId=5470430

28. Bombardier C, Laine L, Reicin A, Shapiro D, Burgos-Vargas R, Davis B, et al. Comparison of Upper Gastrointestinal Toxicity of Rofecoxib and Naproxen in Patients with Rheumatoid Arthritis. New England Journal of Medicine. 2000 Nov 23;343(21):1520–8.

29. Curfman GD, Morrissey S, Drazen JM. Expression of Concern Reaffirmed. New England Journal of Medicine. 2006 Mar 16;354(11):1193–1193.

30. Waxman HA. The Lessons of Vioxx — Drug Safety and Sales. New England Journal of Medicine. 2005 Jun 23;352(25):2576–8.

31. Topol EJ. Failing the Public Health — Rofecoxib, Merck, and the FDA. N Engl J Med. 2004;351(17):1707–9.

32. Graham DJ, Campen D, Hui R, Spence M, Cheetham C, Levy G, et al. Risk of acute myocardial infarction and sudden cardiac death in patients treated with cyclo-oxygenase 2 selective and non-selective non-steroidal anti-inflammatory drugs: nested case-control study. The Lancet. 2005 Feb;365(9458):475–81.

33. CDC - Basic - Medication Safety Program [Internet]. [cited 2014 Jan 15]. Available from: http://www.cdc.gov/medicationsafety/basics.html

34. Lazarou J, Pomeranz BH, Corey PN. Incidence of Adverse Drug Reactions in Hospitalized Patients: A Meta-analysis of Prospective Studies. JAMA. 1998 Apr 15;279(15):1200.

35. Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ. 2004 Jul 3;329(7456):15–9.

36. Howard RL, Avery AJ, Slavenburg S, Royal S, Pipe G, Lucassen P, et al. Which drugs cause preventable admissions to hospital? A systematic review. British Journal of Clinical Pharmacology. 2007 Feb;63(2):136–47.

37. Kongkaew C, Noyce PR, Ashcroft DM. Hospital Admissions Associated with Adverse Drug Reactions: A Systematic Review of Prospective Observational Studies. Ann Pharmacother. 2008 Jul 1;42(7/8):1017–25.

38. Atiqi R, van Bommel E, Cleophas TJ, Zwinderman AH. Prevalence of iatrogenic admissions to the Departments of Medicine/Cardiology/ Pulmonology in a 1,250 bed general hospital. Int J Clin Pharmacol Ther. 2010 Aug;48(8):517–24.

39. Hakkarainen KM, Gyllensten H, Jönsson AK, Andersson Sundell K, Petzold M, Hägg S. Prevalence, nature and potential preventability of adverse drug events - A population-based medical record study of 4970 adults. Br J Clin Pharmacol. 2013 Dec 25;

40. Pedrós C, Quintana B, Rebolledo M, Porta N, Vallano A, Arnau JM. Prevalence, risk factors and main features of adverse drug reactions leading to hospital admission. European Journal of Clinical Pharmacology. 2014 Mar;70(3):361–7.

41. Stausberg J. International prevalence of adverse drug events in hospitals: an analysis of routine data from England, Germany, and the USA. BMC Health Serv Res. 2014;14:125.

42. Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality. JAMA. 1997 Jan 22;277(4):301–6.

43. Reducing and Preventing Adverse Drug Events To Decrease Hospital Costs | Agency for Healthcare Research & Quality (AHRQ) [Internet]. [cited 2014 Jan 15]. Available from: http://www.ahrq.gov/research/findings/factsheets/errors-safety/aderia/index.html

44. Sultana J, Cutroneo P, Trifirò G. Clinical and economic burden of adverse drug reactions. Journal of Pharmacology and Pharmacotherapeutics. 2013;4(5):73.

45. Lahue BJ, Pyenson B, Iwasaki K, Blumen HE, Forray S, Rothschild JM. National burden of preventable adverse drug events associated with inpatient injectable medications: healthcare and medical professional liability costs. Am Health Drug Benefits. 2012 Nov;5(7):1–10.

46. Miller RA, Gardner RM, Johnson KB, Hripcsak G. Clinical decision support and electronic prescribing systems: a time for responsible thought and action. J Am Med Inform Assoc. 2005;12(4):403–9.

47. Li Y, Salmasian H, Harpaz R, Chase H, Friedman C. Determining the Reasons for Medication Prescriptions in the EHR using Knowledge and Natural Language Processing. AMIA Annu Symp Proc. 2011;2011:768–76.

48. Wang W, Haerian K, Salmasian H, Harpaz R, Chase H, Friedman C. A Drug-Adverse Event Extraction Algorithm to Support Pharmacovigilance Knowledge Mining from PubMed Citations. AMIA Annu Symp Proc. 2011;2011:1464–70.

49. Micromedex [Internet]. [cited 2012 Apr 23]. Available from: http://www.micromedex.com/

50. Drug Data | FDB (First Databank) [Internet]. [cited 2012 May 4]. Available from: http://www.fdbhealth.com/

51. Point of Care Medical Applications | Epocrates [Internet]. [cited 2013 Mar 19]. Available from: http://www.epocrates.com/

52. UpToDate Inc. [Internet]. [cited 2012 May 4]. Available from: http://www.uptodate.com/index

53. Wang X, Chase HS, Li J, Hripcsak G, Friedman C. Integrating heterogeneous knowledge sources to acquire executable drug-related knowledge. AMIA Annu Symp Proc. 2010;2010:852–6.

54. Schadow G. Assessing the impact of HL7/FDA Structured Product Label (SPL) content for medication knowledge management. AMIA Annu Symp Proc. 2007;2007:646–50.

55. Sharp M, Bodenreider O, Wacholder N. A framework for characterizing drug information sources. AMIA Annu Symp Proc. 2008;2008:662–6.

56. DailyMed [Internet]. [cited 2012 Feb 8]. Available from: http://dailymed.nlm.nih.gov/dailymed/about.cfm

57. MEDLINE Fact Sheet [Internet]. [cited 2016 Mar 3]. Available from: http://www.nlm.nih.gov/pubs/factsheets/medline.html

58. Home - PubMed - NCBI [Internet]. [cited 2013 Apr 3]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/

59. MEDLINE Baseline Repository - Available Files [Internet]. [cited 2016 Mar 4]. Available from: https://mbr.nlm.nih.gov/Download/

60. MedlinePlus - Health Information from the National Library of Medicine [Internet]. [cited 2014 Jan 16]. Available from: http://www.nlm.nih.gov/medlineplus/

61. RxNorm Overview [Internet]. [cited 2016 Mar 4]. Available from: https://www.nlm.nih.gov/research/umls/rxnorm/overview.html

62. Fact SheetUMLS® Metathesaurus® [Internet]. [cited 2012 Apr 10]. Available from: http://www.nlm.nih.gov/pubs/factsheets/umlsmeta.html

63. National Drug File Reference Terminology (NDF-RT) [Internet]. [cited 2016 Mar 4]. Available from: http://www.va.gov/TRM/StandardPage.asp?tid=5221%5E

64. Zeng Q, Cimino JJ. Automated knowledge extraction from the UMLS. Proc AMIA Symp. 1998;1998:568–72.

65. Chen ES, Hripcsak G, Xu H, Markatou M, Friedman C. Automated acquisition of disease drug knowledge from biomedical and clinical documents: an initial study. J Am Med Inform Assoc. 2008 Feb;15(1):87–98.

66. Shetty KD, Dalal SR. Using information mining of the medical literature to improve drug safety. J Am Med Inform Assoc. 2011 May 5;18(5):668–74.

67. Xu R, Wang Q. Large-scale extraction of accurate drug-disease treatment pairs from biomedical literature for drug repurposing. BMC Bioinformatics. 2013;14:181.

68. Avillach P, Dufour J-C, Diallo G, Salvo F, Joubert M, Thiessard F, et al. Design and validation of an automated method to detect known adverse drug reactions in MEDLINE: a contribution from the EU-ADR project. J Am Med Inform Assoc. 2013 May 1;20(3):446–52.

69. SIDER Side Effect Resource [Internet]. [cited 2012 Feb 22]. Available from: http://sideeffects.embl.de/

70. Kuhn M, Campillos M, Letunic I, Jensen LJ, Bork P. A side effect resource to capture phenotypic effects of drugs. Mol Syst Biol. 2010;6:343.

71. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Understanding MedDRA: The Medical Dictionary for Regulatory Activities [Internet]. 2013. Available from: https://www.meddra.org/sites/default/files/page/documents/meddra2013.pdf

72. DrugBank [Internet]. [cited 2016 Mar 5]. Available from: http://www.drugbank.ca/

73. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res. 2006 Jan 1;34(Database issue):D668–72.

74. Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y, et al. DrugBank 4.0: shedding new light on drug metabolism. Nucleic Acids Res. 2014 Jan;42(Database issue):D1091–7.

75. DrugBank: Documentation and Sources [Internet]. [cited 2016 Mar 5]. Available from: http://www.drugbank.ca/documentation#drug-cards

76. McCoy AB, Wright A, Laxmisan A, Ottosen MJ, McCoy JA, Butten D, et al. Development and evaluation of a crowdsourcing methodology for knowledge base construction: identifying relationships between clinical problems and medications. J Am Med Inform Assoc. 2012 Sep 1;19(5):713–8.

77. Wei W-Q, Cronin RM, Xu H, Lasko TA, Bastarache L, Denny JC. Development and evaluation of an ensemble resource linking medications to their indications. J Am Med Inform Assoc [Internet]. 2013 Apr 10 [cited 2013 Apr 11]; Available from: http://jamia.bmj.com/cgi/doi/10.1136/amiajnl-2012-001431

78. Smith JC. Exploring Adverse Drug Effect Discovery from Data Mining of Clinical Notes [Master's Thesis]. [Nashville, Tennessee]: Vanderbilt University; 2012.

79. Smith JC, Denny JC, Chen Q, Nian H, Spickard A, Rosenbloom ST, et al. Lessons Learned from Developing a Drug Evidence Base to Support Pharmacovigilance: Applied Clinical Informatics. 2013;4(4):596–617.

80. Salmasian H, Tran TH, Chase HS, Friedman C. Medication-indication knowledge bases: a systematic review and critical appraisal. Journal of the American Medical Informatics Association. 2015 Sep 2;ocv129.

81. Classen DC, Pestotnik SL, Evans RS, Burke JP. Computerized surveillance of adverse drug events in hospital patients. JAMA. 1991 Nov 27;266(20):2847–51.

82. Evans RS, Pestotnik SL, Classen DC, Bass SB, Menlove RL, Gardner RM, et al. Development of a computerized adverse drug event monitor. Proc Annu Symp Comput Appl Med Care. 1991;23–7.

83. Evans RS, Pestotnik SL, Classen DC, Bass SB, Burke JP. Prevention of adverse drug events through computerized surveillance. Proc Annu Symp Comput Appl Med Care. 1992;437–41.

84. Bates DW, O'Neil a C, Boyle D, Teich J, Chertow GM, Komaroff a L, et al. Potential identifiability and preventability of adverse events using information systems. Journal of the American Medical Informatics Association : JAMIA. 1(5):404–11.

85. Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, Servi D, et al. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group. JAMA. 1995 Jul 5;274(1):29–34.

86. Bates DW, Leape LL, Cullen DJ, Laird N, Petersen LA, Teich JM, et al. Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. JAMA. 1998 Oct 21;280(15):1311–6.

87. Jha AK, Kuperman GJ, Teich JM, Leape L, Shea B, Rittenberg E, et al. Identifying adverse drug events: development of a computer-based monitor and comparison with chart review and stimulated voluntary report. Journal of the American Medical Informatics Association : JAMIA. 1998;5(3):305–14.

88. Honigman B, Lee J, Rothschild J, Light P, Pulling RM, Yu T, et al. Using computerized data to identify adverse drug events in outpatients. J Am Med Inform Assoc. 2001 Jun;8(3):254–66.

89. Bates DW, Evans RS, Murff H, Stetson PD, Pizziferri L, Hripcsak G. Detecting adverse events using information technology. Journal of the American Medical Informatics Association : JAMIA. 2003;10(2):115–28.

90. Murff HJ, Forster AJ, Peterson JF, Fiskio JM, Heiman HL, Bates DW. Electronically screening discharge summaries for adverse medical events. J Am Med Inform Assoc. 2003 Aug;10(4):339–50.

91. Raschke RA, Gollihare B, Wunderlich TA, Guidry JR, Leibowitz AI, Peirce JC, et al. A Computer Alert System to Prevent Injury From Adverse Drug Events: Development and Evaluation in a Community Teaching Hospital. JAMA. 1998 Oct 21;280(15):1317.

92. Peterson JF. Drug-Lab Triggers Have Potential to Prevent Adverse Drug Events in Outpatients. Journal of the American Medical Informatics Association. 2002 Nov 1;9(90061):39S – 40.

93. Handler SM, Altman RL, Perera S, Hanlon JT, Studenski SA, Bost JE, et al. A systematic review of the performance characteristics of clinical event monitor signals used to detect adverse drug events in the hospital setting. Journal of the American Medical Informatics Association : JAMIA. 2007;14(4):451–8.

94. Murff HJ, Patel VL, Hripcsak G, Bates DW. Detecting adverse events for patient safety research: a review of current methodologies. Journal of Biomedical Informatics. 2003 Feb;36(1-2):131–43.

95. Weingart SN, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, et al. Use of administrative data to find substandard care: validation of the complications screening program. Med Care. 2000 Aug;38(8):796–806.

96. Hohl CM, Karpov A, Reddekopp L, Stausberg J. ICD-10 codes used to identify adverse drug events in administrative data: a systematic review. J Am Med Inform Assoc. 2013 Nov 12;

97. Lawthers AG, McCarthy EP, Davis RB, Peterson LE, Palmer RH, Iezzoni LI. Identification of in-hospital complications from claims data. Is it valid? Med Care. 2000 Aug;38(8):785–95.

98. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. Journal of the American Medical Informatics Association : JAMIA. 12(4):448–57.

99. Gysbers M, Reichley R, Kilbridge PM, Noirot L, Nagarajan R, Dunagan WC, et al. Natural language processing to identify adverse drug events. AMIA Annu Symp Proc. 2008;961.

100. Haerian K, Salmasian H, Friedman C. Methods for identifying suicide or suicidal ideation in EHRs. AMIA Annu Symp Proc. 2012;2012:1244–53.

101. Eriksson R, Jensen PB, Frankild S, Jensen LJ, Brunak S. Dictionary construction and identification of possible adverse drug events in Danish clinical narrative text. Journal of the American Medical Informatics Association. 2013 May 23;20(5):947–53.

102. Haerian K, Varn D, Vaidya S, Ena L, Chase HS, Friedman C. Detection of pharmacovigilance-related adverse events using electronic health records and automated methods. Clin Pharmacol Ther. 2012 Aug;92(2):228–34.

103. Wang X, Hripcsak G, Markatou M, Friedman C. Active Computerized Pharmacovigilance Using Natural Language Processing, Statistics, and Electronic Health Records: A Feasibility Study. J Am Med Inform Assoc. 2009 Mar 4;16(3):328–37.

104. Wang X, Chase H, Markatou M, Hripcsak G, Friedman C. Selecting information in electronic health records for knowledge acquisition. J Biomed Inform. 2010 Aug;43(4):595–601.

105. LePendu P, Iyer SV, Bauer-Mehren A, Harpaz R, Mortensen JM, Podchiyska T, et al. Pharmacovigilance Using Clinical Notes. Clin Pharmacol Ther [Internet]. 2013 Mar 4 [cited 2013 Apr 11]; Available from: http://www.nature.com/doifinder/10.1038/clpt.2013.47

106. Harpaz R, Vilar S, Dumouchel W, Salmasian H, Haerian K, Shah NH, et al. Combing signals from spontaneous reports and electronic health records for detection of adverse drug reactions. J Am Med Inform Assoc. 2013 May 1;20(3):413–9.

107. Harpaz R, DuMouchel W, Shah NH, Madigan D, Ryan P, Friedman C. Novel data-mining methodologies for adverse drug event discovery and analysis. Clin Pharmacol Ther. 2012 Jun;91(6):1010–21.

108. Tatonetti NP, Ye PP, Daneshjou R, Altman RB. Data-driven prediction of drug effects and interactions. Sci Transl Med. 2012 Mar 14;4(125):125ra31.

109. Harpaz R, Callahan A, Tamang S, Low Y, Odgers D, Finlayson S, et al. Text mining for adverse drug events: the promise, challenges, and state of the art. Drug Saf. 2014 Oct;37(10):777–90.

110. FDA's Sentinel Initiative [Internet]. [cited 2012 Apr 3]. Available from: http://www.fda.gov/Safety/FDAsSentinelInitiative/default.htm

111. Mini-Sentinel [Internet]. [cited 2013 Aug 23]. Available from: http://www.mini-sentinel.org/default.aspx

112. Denny JC, Miller RA, Johnson KB, Spickard A 3rd. Development and evaluation of a clinical note section header terminology. AMIA Annu Symp Proc. 2008;156–60.

113. Denny JC, Smithers JD, Miller RA, Spickard A. "Understanding" medical school curriculum content using KnowledgeMap. J Am Med Inform Assoc. 2003;10(4):351–62.

114. Denny JC, Peterson JF, Choma NN, Xu H, Miller RA, Bastarache L, et al. Extracting timing and status descriptors for colonoscopy testing from electronic medical records. J Am Med Inform Assoc. 2010 Aug;17(4):383–8.

115. KMCI - KnowledgeMap Concept Indexer | Center for Precision Medicine [Internet]. [cited 2016 Mar 6]. Available from: https://medschool.vanderbilt.edu/cpm/center-precision-medicine-blog/kmci-knowledgemap-concept-indexer

116. Xu H, Stenner SP, Doan S, Johnson KB, Waitman LR, Denny JC. MedEx: a medication information extraction system for clinical narratives. J Am Med Inform Assoc. 2010 Feb;17(1):19–24.

117. Peterson JF, Shi Y, Denny JC, Matheny ME, Schildcrout JS, Waitman LR, et al. Prevalence and Clinical Significance of Discrepancies within Three Computerized Pre-Admission Medication Lists. AMIA Annu Symp Proc. 2010;2010:642–6.

118. Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. Clin Pharmacol Ther. 2008 Sep;84(3):362–9.

119. Denny JC, Spickard A. Automated capture and assessment of medical student clinical experience. AMIA Annu Symp Proc. 2008;926.

120. Spickard A, Ridinger H, Wrenn J, O'brien N, Shpigel A, Wolf M, et al. Automatic scoring of medical students' clinical notes to monitor learning in the workplace. Med Teach. 2014 Jan;36(1):68–72.

121. RxNorm Current Prescribable Content [Internet]. [cited 2014 Jan 28]. Available from: http://www.nlm.nih.gov/research/umls/rxnorm/docs/prescribe.html

122. Oliver DE. SQL Medline Schema [Internet]. [cited 2016 Mar 8]. Available from: http://biotext.berkeley.edu/code/medline-schema/medline-schema-perl-oracle.sql

123. Oliver DE. parsemedline.pl [Internet]. 2004 [cited 2016 Mar 8]. Available from: http://biotext.berkeley.edu/code/medline-schema/parsemedline.pl

124. Substance Registration System - Unique Ingredient Identifier (UNII) [Internet]. [cited 2016 Mar 8]. Available from: http://www.fda.gov/ForIndustry/DataStandards/SubstanceRegistrationSystem-UniqueIngredientIdentifierUNII/default.htm

125. Miller M Randolph A. Synonyms - Final Version. 2015.

126. The PubChem Project [Internet]. [cited 2012 Apr 10]. Available from: http://pubchem.ncbi.nlm.nih.gov/

127. MedlinePlus Connect: Web Service [Internet]. [cited 2016 Mar 8]. Available from: https://www.nlm.nih.gov/medlineplus/connect/service.html

128. Google [Internet]. [cited 2016 Mar 8]. Available from: https://www.google.com/

129. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977 Mar;33(1):159–74.

130. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther. 2005 Mar;85(3):257–68.

131. Medications for treating hypertension - Harvard Health [Internet]. [cited 2016 Mar 8]. Available from: http://www.health.harvard.edu/heart-health/medications-for-treating-hypertension

132. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. JAMA. 2003 May 21;289(19):2560–72.

133. James PA, Oparil S, Carter BL, Cushman WC, Dennison-Himmelfarb C, Handler J, et al. 2014 Evidence-Based Guideline for the Management of High Blood Pressure in Adults: Report From the Panel Members Appointed to the Eighth Joint National Committee (JNC 8). JAMA. 2014 Feb 5;311(5):507.

134. Embi PJ, Leonard AC. Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study. Journal of the American Medical Informatics Association. 2012 Apr 25;19(e1):e145–8.

135. Footracer KG. Alert fatigue in electronic health records. JAAPA. 2015 Jul;28(7):41–2.

136. Phansalkar S, van der Sijs H, Tucker AD, Desai AA, Bell DS, Teich JM, et al. Drug-drug interactions that should be non-interruptive in order to reduce alert fatigue in electronic health records. J Am Med Inform Assoc. 2013 May 1;20(3):489–93.

137. Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform. 2001 Oct;34(5):301–10.

138. Fast Facts | Department of Medicine [Internet]. [cited 2016 Mar 9]. Available from: https://medicine.mc.vanderbilt.edu/fast-facts

139. Hohl CM, Dankoff J, Colacone A, Afilalo M. Polypharmacy, adverse drug-related events, and potential adverse drug interactions in elderly patients presenting to an emergency department. Annals of Emergency Medicine. 2001 Dec;38(6):666–71.

140. Shah BM, Hajjar ER. Polypharmacy, adverse drug reactions, and geriatric syndromes. Clin Geriatr Med. 2012 May;28(2):173–86.