

Effective Health Care Research Reports

Number 12

Developing Evidence-Based Research Priorities for Off-Label Drug Use

Surrey M. Walton, Ph.D.

Glen T. Schumock, Pharm.D., M.B.A.

Ky-Van Lee, Ph.D.

G. Caleb Alexander, M.D., M.S.

David Meltzer, M.D., Ph.D.

Randall S. Stafford, M.D., Ph.D.

Research from the Developing Evidence to Inform Decisions about Effectiveness (DEClDE) Network



Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

May 2009

The DEcIDE (Developing Evidence to Inform Decisions about Effectiveness) network is part of AHRQ's Effective Health Care program. It is a collaborative network of research centers that support the rapid development of new scientific information and analytic tools. The DEcIDE network assists health care providers, patients, and policymakers seeking unbiased information about the outcomes, clinical effectiveness, safety, and appropriateness of health care items and services, particularly prescription medications and medical devices.

This report is based on research conducted by the University of Illinois at Chicago DEcIDE (Developing Evidence to Inform Decisions about Effectiveness) Center under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. HHS290200500381 TO3). The AHRQ Task Order Officer for this project was Anne Trontell, M.D., M.P.H.

The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ or any other Federal agency. Therefore, no statement in this report should be construed as an official position of AHRQ or any other agency of the U.S. Department of Health and Human Services.

None of the authors has a financial interest in any of the products discussed in this report.

This report has also been published in edited form: Walton SM, Schumock GT, Lee K-V, et al. Prioritizing future research on off-label prescribing: results of a quantitative evaluation. *Pharmacotherapy* 2008 Dec;28(12):1443–52.

Suggested citation:

Walton SM, Schumock GT, Lee K-V, et al. Developing evidence-based research priorities for off-label drug use. Effective Health Care Research Report No. 12. (Prepared by the University of Illinois at Chicago DEcIDE Center Under Contract No. HHS290200500381 T03.) Rockville, MD: Agency for Healthcare Research and Quality. May 2009. Available at: effectivehealthcare.ahrq.gov/reports/final.cfm.

Contents

Introduction.....	1
Methods.....	1
Results.....	5
Discussion.....	6
Conclusion.....	9
Acknowledgments.....	10
References.....	10
Figures and Tables.....	12
Appendix.....	18

Author affiliations:

Surrey M. Walton, Ph.D.^a

Glen T. Schumock, Pharm.D., M.B.A.^b

Ky-Van Lee, Ph.D.^c

G. Caleb Alexander, M.D., M.S.^{d,e}

David Meltzer, MD, Ph.D.^{d,f}

Randall S. Stafford, MD, Ph.D.^c

^aDepartment of Pharmacy Administration and Center for Pharmacoeconomic Research, College of Pharmacy, University of Illinois at Chicago.

^bDepartment of Pharmacy Practice and Center for Pharmacoeconomic Research, College of Pharmacy, University of Illinois at Chicago.

^cProgram on Prevention Outcomes and Practices, School of Medicine, Stanford University.

^dSection of General Internal Medicine, Department of Medicine, and Center for Health and Social Sciences, University of Chicago.

^eMacLean Center for Clinical Medical Ethics, University of Chicago.

^fDepartment of Economics, University of Chicago.

Abstract

Background. Drug use for indications not approved by the Food and Drug Administration exceeds 20% of outpatient prescriptions. Available compendia indicate that a minority of off-label uses are well supported by evidence. Policy makers, however, lack information to identify where systematic review of the evidence or other research would be most valuable.

Methods. Nationally representative prescribing data from the IMS Health (Plymouth Meeting PA) National Disease and Therapeutic Index were used to estimate the number of off-label drug uses by indication from 1/2005 through 6/2007. Off-label uses were categorized according to adequacy of scientific support based on the MICROMEDEX DRUGDEX compendium. Our model also incorporated FDA-identified safety concerns, mean prescription prices, marketing expenditures, and recent market entry.

Results. Top priority drugs were quetiapine, warfarin, escitalopram, risperidone, montelukast, bupropion, sertraline, venlafaxine, celecoxib, lisinopril, duloxetine, trazodone, olanzapine, and epoetin alfa.

Conclusion. Future research into off-label drug use should focus on drugs used frequently with inadequate supporting evidence, particularly if further concerns are raised by known safety issues, high drug cost, recent market entry, and extensive marketing. Our quantitative analysis identified particular concerns with the off-label use of antipsychotic and antidepressant medications. Targeted research and policy activities on our list of prioritized drugs have high potential value.

Introduction

The United States Food and Drug Administration (FDA) drug approval process requires clinical trials that demonstrate efficacy for a specific indication prior to market availability. Once approval for an indication is given, however, the FDA has a minimal role in determining whether drugs are prescribed for their approved uses or for other “off-label” indications not part of the FDA-approved product labeling.

Prior studies of off-label prescribing have typically been conducted in narrowly defined clinical contexts, including psychiatric disorders,^{1,2} bacterial infections,³ in children^{2,4} as well as in “orphan” clinical conditions not well served by FDA approved drugs.^{5,6} These studies have demonstrated that off-label use is frequent in these clinical areas, that such uses arise from diverse circumstances, and that there is substantial variation in the available evidence supporting particular off-label uses. Notably, while raising some concerns about off-label use, the studies also identify disease- and patient-specific situations where off-label use may be beneficial.

Three prior studies have examined off-label use among broader, more representative patient populations in the United States. First, in 1985, Strom et al. examined the 100 most common uses of marketed medications and found 31 indications that were not initially approved by the FDA.⁷ Second, in 2003, the Knight-Ridder news services used Verispan (Yardley PA) physician-reported prescribing data to estimate general rates of off-label use.⁸ An estimated 21% of prescription drug use in 2002-2003 and 22% in 1997-1998 was off-label by indication. Rates of off-label use varied widely across drug classes and were highest among anticonvulsants (74%), antipsychotics (60%), and antibiotics (41%). Third, Radley and colleagues examined the frequency and clinical circumstances of off-label prescribing for common medications as a function of the strength of scientific support for those practices.⁹ Of the estimated 21% of drugs uses that were off-label, 73% of these did not have good scientific support. Although several drug classes, including anticonvulsants, antipsychotics, and antibiotics, were associated with a higher frequency of off-label use, other drug characteristics such as manufacturer, and annual sales levels did not predict the likelihood of off-label prescribing. While this study highlighted the magnitude of off-label drug use in the absence of evidence regarding efficacy, it made no attempt to define the level of concern associated with specific drugs.

We sought to build upon these studies to develop a systematic process for prioritizing future federally sponsored research and policy initiatives regarding off-label drug use based on the concept of value of information.¹⁰ Our primary purpose was to consider multiple dimensions of off-label use (extent of use, quality of existing information for that use, drug cost, extent of marketing, and time on market) for a large list of commonly used drugs in order to identify areas where future initiatives would be likely to have the greatest value.

Methods

Framework for Prioritizing Off-Label Drug Uses

We first convened a technical expert panel (see below) representing various stakeholder groups, to help guide our methods. The project developed based on three guiding principles for prioritizing initiatives related to medications used off-label. First, priority would be based on the potential for future initiatives to generate the greatest societal return. This placed the volume of drug use with inadequate evidence (see below) as our central focus. Second, the process would

be driven by data, which meant that all measures in the prioritization model(s) were based on quantitative results from existing data. Third, that the focus would be in terms of individual drugs, rather than drug classes, consistent with FDA regulatory policy.

Advisory Panel

To obtain a broad range of input on our methods, we assembled a panel of key stakeholders and experts with complementary perspectives. First, we selected four FDA staff to represent the FDA and to provide technical input on ICD-9-code matching, drug safety, and regulatory policy. Other key stakeholders included practicing specialist physicians, large group purchasers, the health insurance industry, the pharmaceutical industry, and academia. This nine-member advisory panel provided input throughout project development, implementation, and analysis (see acknowledgments). Participation was greatest during the selection of key drug characteristics, the relative weighting of these factors in the priority list modeling, and the specifications of the sensitivity analyses.

Prescribing Data

For its large size (approximately 200,000 contact records per year), national scope, timely availability, and reporting of the medications used to treat specific clinical conditions, we used the IMS Health National Disease and Therapeutic Index (NDTI) to quantify the volume of prescribing for off-label uses of specific drugs.¹¹ The NDTI is a physician survey that provides nationally representative diagnostic and prescribing information on patients treated by US office-based physicians. For each contact record, physicians record the specific medications used to treat any applicable diagnosis. Drug reporting reflects the physician's best knowledge of new or continuing prescription or non-prescription medications. While most contacts comprise physician office visits, NDTI also includes physician visits made in hospitals and nursing homes. The NDTI has been cross-validated against other national sources of information on outpatient practice.^{12,13}

We constructed national estimates of prescription use using the NDTI. Using the data and information on labeled indications available from the FDA, prescription uses were counted as off-label if the indication ICD-9 code did not match with an FDA approved indication for the drug. Information on approved indications was obtained from MICROMEDEX, as well as the FDA.^{14,15,16} We took a conservative approach in that if any indication was on labeled as of June of 2007 it was considered as a labeled indication throughout the time period examined. For example, duloxetine was approved for use in generalized anxiety (2/27/07) and atorvastatin for use in coronary artery disease (3/8/07).¹⁶

We also used the IMS Health National Prescription Audit to insure that the most common drugs dispensed through pharmacies matched those reported by physicians at patient encounters in NDTI.¹⁷ These data provided national estimates of the volume of prescription orders received by pharmacies, physical prescriptions dispensed by pharmacies, and mean cost per dispensed prescription based on a national sample of approximately 25,000 retail, discount, and mail-order pharmacies.

Scope of Prescribing Practices Examined

To capture recent prescribing patterns, we analyzed 30 months of prescribing data (January 2005 through June 2007). This time period restriction was applied to all databases used

in the study. We excluded children (age < 18 years) because of the distinctive nature of pediatric off-label use. We excluded vaccines, non-drug products (e.g., liquid nitrogen), and any medications available over-the-counter (whether or not a prescription formulation also was available, e.g., ibuprofen, omeprazole). We also excluded chemotherapeutic drugs with specific anti-neoplastic activity because it is a subset of off label drug use in adults that has already received a relatively high degree of attention. Finally, drug names were aggregated, so that generic and brand names for the same chemical entity administered by the same route were considered together.

Data on Evidence To Support Off-label Drug Uses

For determining levels of existing evidence, we used DRUGDEX (Micromedex, Greenwood Village CO) which contains readily available summaries of evidence-based information on both on- and off-label indications of specific drugs.¹⁸ DRUGDEX provides an assessment in three important evaluation dimensions: (1) efficacy (in four categories: “effective,” “evidence favors efficacy,” “evidence is inconclusive,” and “ineffective”), (2) strength of recommendation (in four categories: “Class I-recommended,” “Class IIa-recommended in most cases,” “Class IIb-recommended in some cases,” and “Class III-not recommended,” and (3) strength of evidence [in three graded categories: “A,” (good RCT evidence) “B,” (less consistent RCT evidence) and “C” (non-RCT forms of evidence)]. We used these three DRUGDEX evaluation dimensions to construct three ordinal levels of aggregate support for specific off-label indications. These were a) evidence-based off-label use, b) uncertain evidence base for off-label use, and c) inadequate evidence for off-label use (Figure 1). While similar to the algorithm of Radley,⁹ we sought to explicitly consider those indications where ambiguous evidence is available and therefore added the “uncertain evidence” category.

We used the U.S. Pharmacopeia Drug Information (USP-DI) as an adjunct to DRUGDEX.¹⁹ USP-DI lists FDA approved indications, as well as “accepted” off-label uses. In general, there were only rare instances in which a drug identified as lacking strong evidence in DRUGDEX was an accepted USP-DI indication. The only notable exception was the use of warfarin in coronary artery disease. The USP-DI assessment, however, is based on expert consensus, rather than clinical trial evidence.

In using the DRUGDEX compendium, a particular challenge was matching the ICD-9 coding of indications reported by physicians in NDTI with the diagnostic labels used by DRUGDEX (including those used by the FDA in describing approved indications). Given limitations of both ICD-9 and DRUGDEX diagnostic labels, there were many instances in which there was an inexact correspondence between labels. Using clinical judgment, we developed several rules of thumb to aid in this matching process. In general, these guidelines sought to provide some degree of latitude to NDTI physicians in coding diagnoses, particularly when non-specific codes were present. For example, ICD-9 490 “bronchitis, not specified as acute or chronic” was considered acute bronchitis for antibiotics and chronic bronchitis for traditional bronchodilators. For these two instances, the most frequent specific codes suggested this interpretation. Similarly, non-specific codes for bipolar disorder were interpreted as chronic maintenance because specific codes for manic, mixed or depressive acute episode were infrequent.

There were some ICD-9 codes representing clinical conditions where no match was found among the clinical indications listed in DRUGDEX. We considered these “not found”

diagnoses as uses with inadequate evidence in the base analyses and conducted a sensitivity analysis (see below) that dropped these uses to measure the impact of this assumption.

Based on the number of off label uses with inadequate evidence, a primary index representing the relative magnitude of off-label use with inadequate evidence (varying from 0 to 1) was constructed as the ratio of a particular drug's number of off-label uses with inadequate evidence to the drug with the greatest number of off-label uses with inadequate evidence.

Additional Drug Characteristics for Prioritization

Based on our review of past research and input from our technical expert panel, we identified four drug characteristics to include in the analyses. These were chosen based on being important to prioritizing future systematic evidence reviews or other research in terms of value of information, and also for being measurable across all drugs. The first was the presence of serious side-effects and/or drug-drug interactions. The second was high drug cost. The third was recent market entry, and the fourth was extensive drug promotion to consumers and physicians. The cost, market entry, and drug promotion factors were combined to indicate a heightened concern about uses with inadequate evidence occurring for the correlated features of recent drug approval, aggressive marketing and higher costs.

For each factor, we created a quantitative relative index measure. We used FDA safety warnings as a proxy for serious side-effects and drug interactions. Specifically, a drug with an FDA Black Box warning was scored as one for most concerning, those with other published safety alerts identified via the FDA Center for Drug Evaluation and Review website were scored as 0.5 for less concerning, and other drugs scored as zero for this index.^{15,16}

Cost concerns around off-label drug use were represented by the mean total cost of a single prescription of the medication.¹⁷ The ratio of a particular drug's prescription cost relative to the drug with the greatest cost formed the cost index. Second, we measured recency of market entry by grouping drugs by years since first FDA approval (source: FDA website Orange Book, and FOIA report).^{20,21} Specifically, drugs approved within the past three years (based from 12/31/06) were coded as 1.0, drugs approved between 4-13 years as 0.5, and drugs approved over 13 year ago or available as generics coded as 0.^{20,21} Third, concerns about the intensity of drug promotion were represented by total expenditures on physician promotion plus expenditures on direct to consumer advertising. The ratio of a particular drug's total marketing expenditures relative to the drug with the greatest combined marketing expenditures formed the marketing index.^{22,23} These three measures were then combined into a single index by taking an equal weighted average of the three components

Developing Priority Scores and Listings

Priority scores and associated listings of drugs were developed from the three primary components described above representing volume of use with inadequate evidence, safety concerns, and the cost and market based characteristics. Note that each of these three primary index scores is bounded between zero and 1 and that any weighted average of these scores is also bounded between zero and one.

As a base analysis, we used a weighted average of the three scores with volume weighted at 0.5, safety at 0.25 and the combined cost and marketing factor at 0.25. We weighted the volume component with a higher weight because extent of use is a fundamental factor in considering calculations of value of research. The fundamental underlying premise to ranking the drugs is that the value of research is proportional to the index score so that drugs with a

higher score are ranked higher in priority. Given that this was a ranking process, the relative levels of the scores are more relevant than the absolute levels for any drug.

Sensitivity Analyses

To assess the robustness of the rankings, we varied some of the fundamental assumptions in the volume and evidence categorization. The most important variation to volume and evidence was to model the volume of off-label uses by alternatively considering, 1) uses where the evidence was inadequate plus one half times the number of uses with uncertain evidence, and 2) considering the sum of uses with inadequate and uncertain evidence. We also examined the impact of dropping the “not found” uses where an NDTI ICD-9 code could not be matched to a DRUGDEX indication, rather than counting them as “inadequate evidence” uses.

Further, we developed and examined a variety of priority lists based on alternative weighted combinations of the three primary indices described above. We also conducted a separate prioritization based on the volume index multiplied by the cost and marketing index. This multiplied index examines the impact of changing the functional form of the index, and can also be interpreted as prioritization based on total dollars spent on the drug.

In addition, alternative methods of characterizing safety, cost, promotional efforts, and recent market entry were investigated, but did not alter our results. Finally, because bipolar disorder was found to be prevalent in many of the top priority drugs, we investigated an additional set of models that considered bipolar disorder to be a uniformly approved indication for antipsychotic and anticonvulsant medication, although current FDA labeling is limited to a subset of patients with bipolar disorder for many of these drugs.

Results

There were 56,110 drug indication pairs, and a total of 1.8 billion reported drug uses with a mean number of 18 indications per drug in the NDTI data. Prescribing volume information from the IMS National Prescription Audit data verified that the NDTI data included all commonly prescribed medications. Table 1 shows the top 25 drugs in terms of total combined off-label uses with uncertain or inadequate evidence and summarizes the number of on- and off-label uses as well as the number of off label uses in each evidence category for those drugs. To be as explicit as possible about the final lists, Table 1 also shows the same information for seven additional drugs (with corresponding overall ranking out of the top 100) to include all that subsequently appear in Table 3 on at least one of the priority lists. Table 2 displays the index scores for each of the individual measures used in forming the priority lists for the top 25 drugs again. It also, for the same reason as stated for Table 1, includes the seven additional drugs that subsequently appear in Table 3. The pair-wise correlations between the three primary components was 0.09, -0.13, and 0.06 for volume and safety, volume and the combined cost and market index, and safety and the combined index respectively. Among the three sub-factors in the combined index, the correlations were somewhat higher, ranging from 0.16 to 0.45.

The drugs listed in Tables 1 and 2 represent a broad range of drug classes and areas of clinical practice. Most prominent among the drugs in the top 25 are antidepressants (escitalopram, trazodone, sertraline, bupropion, amitriptyline, and venlafaxine), antipsychotics (quetiapine and risperidone), and anxiolytics/sedatives (zolpidem, lorazepam, and clonazepam). Not shown as it was outside of the model, the top 5 drugs in terms of percentage off label were

as follows: promethazine (92.9%), clonazepam (76.6%), quetiapine (75.4%) nortriptyline (59.6%), and vancomycin (56.1%).

Table 3 displays a base model and seven additional selected priority lists of 15 drugs each based on different weightings of the individual priority factors. Though the lists are different, with 30 drugs appearing on at least one of the lists, many of the top drugs are consistent across the lists. For example, at least 10 of the 15 drugs in the base model appear on any one of the other lists. Further, despite very different modeling assumptions, the top drug on each list was quetiapine and thirteen of the remaining fourteen drugs in the base model, warfarin, escitalopram, risperidone, montelukast, bupropion, sertraline, venlafaxine, celecoxib, lisinopril, duloxetine, trazodone, olanzopine, and epeotin alfa appear in the top 15 in the majority of lists. (Aripiprazole appears in only two of the variant models.) Further sensitivity analyses varying coding methods, not included in the table, did not lead to significant changes in the lists. Hence, we take these 14 drugs as likely candidates for where future research exploring the evidence base for off label use can generate significant value to society.

Each of the drugs that appear in the base model are shown in Table 4 along with their top on-label use and their top three off-label indications (with uncertain or inadequate evidence). Table 4 provides explicit information on the indications driving our evidence ranking and identifies conditions that have prompted off-label use particularly bipolar disorder and depression. In several instances among atypical antipsychotics, the most common off-label use is more frequent than the leading on-label use. Note that generally the top 1 to 3 indications represented the vast majority of the overall off-label uses of the drug. For example, for quetiapine, the on-label uses and the top three off-label uses account for 77% of all uses, while this proportion is 98% for escitalopram.

Sensitivity Analysis on Coding

We examined the impact of changing the coding assumptions in two key areas of our analysis. First, the base analyses dropped indications that could not be found in DRUGDEX, where a match was not possible between NDTI ICD-9 codes and the DRUGDEX indications. An alternative is to code those as uses with inadequate evidence. When we included “not found” uses coded as inadequate evidence uses, the main effect was to increase the priority of escitalopram and warfarin because their reported uses respectively for bipolar disorder and hypertensive heart disease were not found. All other high priority drugs remained high priority.

Finally, we examined changes when the use of any antipsychotics and anticonvulsants was uniformly coded as approved for all diagnoses related to bipolar disorder. Our base analysis coded bipolar disorder in the absence of decompensation as an off-label use for quetiapine and risperidone because FDA labeling does not identify chronic maintenance as an approved use. With this modification the top drugs remained the same: quetiapine, escitalopram, warfarin, and risperidone.

Discussion

Our findings, which combine unique nationally representative prescribing data identified a high volume of off-label prescribing in the absence of good evidence for a substantial number of drugs, particularly antidepressants, antipsychotics, and anxiolytics/sedatives. Accounting for other drug characteristics pertinent to determining where research into evidence gaps would most

likely be of value, yielded robust results that placed priority on antidepressants and antipsychotics.

These rankings correspond in many respects with past assessments of off-label use, although some important differences were noted. Our assessment has extended past studies by examining drug characteristics that can impact the level of concern associated with specific drugs. Several drugs and drug classes identified by past researchers as having prevalent off-label use,^{1,3,8,9} particularly anticonvulsants and antibiotics, were not as prominent within our listing of priority drugs.

Our list is formed from a general framework and limited to available quantitative measures that span the spectrum of medications. Specific outcomes associated with any particular drug or class could impact the overall value for research in that area. For example, among the 32 drugs found in any of the top 15, there is only one antibiotic (metronidazole). The generally better safety profile of antibiotics, their lower cost, older age, and limited marketing resulted in these drugs generally being ranked lower than antipsychotics and antidepressants. However, consideration of microbial antibiotic resistance could be an argument for increasing the priority of antibiotics beyond the criteria used in our analysis. Nonetheless, we maintain that a primary consideration in a general ranking is volume of use without adequate evidence, and our results were robust to several additional considerations.

Off-label use of the atypical antipsychotics has long been a source of concern. Radley,⁹ as well as the Knight-Ridder⁸ news series, identified members of this drug class as particularly noteworthy. We identified four antipsychotics (quetiapine, risperidone, olanzapine and aripiprazole) among our high priority drugs. The most common conditions prompting their use were bipolar disorder, depression, and dementia. This may result from the lack of entirely acceptable and effective treatments, as well as their overlapping definitions. Further, conditions with subjective endpoints may be more prone to bias from anecdotal reports.

In fact, although it is often assumed²⁴ that all antipsychotics are approved for chronic maintenance in bipolar disorder, the FDA-approval for quetiapine and risperidone is limited to manic or depressive presentations. Indeed, available evidence for these drugs focuses on short-term studies of acute decompensated bipolar disorder. DRUGDEX lists these two antipsychotics as being inadequately supported by evidence for unspecified bipolar disorder. Our sensitivity analysis, however, suggested that the high ranking of quetiapine and risperidone did not depend on their use in bipolar disorder. Most atypical antipsychotics are used commonly for depression in the absence of strong scientific evidence. Similarly, support is lacking for their use in dementia where safety concerns have been raised about cardiovascular adverse events and included in an FDA Black Box warning for all atypical antipsychotics.

Gabapentin has been highlighted recently in discussions of off label use.^{1,8,9} Although relatively prominent in our analysis, it was not a top priority because it lacks a black box warning, has limited ongoing promotion, and has been on the market for some time. Furthermore, it has already received a good deal of attention, which we argue would imply a relatively greater emphasis on the drugs highlighted in our list for future research priority.

Among the top priority drugs, six were antidepressants, including two serotonin reuptake inhibitors (escitalopram and sertraline), two serotonin- and norepinephrine-reuptake inhibitors (venlafaxine and duloxetine), and two other class (bupropion and trazodone). Bipolar disorder was the most common condition prompting off-label use. There is only limited support for the use of these drugs in this condition, particularly as maintenance therapy. Even as an adjunct to

other therapies in depressive episodes, antidepressants do not appear to be effective in bipolar disorder.²⁵ Hence, we view these medications as valuable focal points for future research.

In terms of how research should proceed, our framework suggests that research efforts are best focused on a small set of high priority off-label uses where further concentrated evaluation can be expected to provide the greatest value. Such evaluations would begin with a comprehensive review of all available information on the use of a drug to assess the extent and quality of the available evidence to date. Such a review may subsequently include a range of analyses from retrospective database queries to prospective randomized trials where appropriate. In addition, policy response may include a requirement for comparative trials, greater scrutiny of marketing for a drug, efforts by payers to require physician justification, and a range of post-market pharmacovigilance activities. Our top priority drugs, quetiapine, warfarin, escitalopram, risperidone, montelukast, bupropion, sertraline, venlafaxine, celecoxib, lisinopril, duloxetine, trazodone, olanzapine, and epoetin alfa are strong candidates for further investigation.

In terms of efficiency, a relevant consideration is whether to focus research on individual medications, drug classes, or clinical indications. Our analysis focused on individual medications, but also highlighted antipsychotic and antidepressant drug classes and the importance of bipolar disorder and depression in prompting off-label use. It may be beneficial to focus research on drug classes or clinical indications if this provides efficiencies in acquiring data or allows synergy in the evaluation of multiple drugs.

Several interesting research and policy issues remain. A general area of interest is that off-label use often arises through clinical proximity of an indication to approved indications. Patterns of use for drugs found to be safe and efficacious for acute indications often expand to include related chronic conditions. It is not at all clear, however, that evidence of efficacy in a clinically proximate indication is sufficient to support common use for the other indication. Similarly, some off-label use arises when a drug is used for an indication proven for another member of the same drug class. Whether this assumption of a “class effect” is appropriate requires further research, and is almost certain to depend on the class and indication in question.

In general, more concentrated evaluation of high priority off-label uses should be used to inform policy strategies in health care settings not only to dissuade physicians from inappropriate off-label prescribing, but also to encourage selective off-label use when appropriate. How future research efforts are organized, funded, and conducted are crucial issues. Only well structured analyses following good scientific techniques can help build useful evidence for or against particular uses of medication. The overall goal should be to promote good evidence based medicine and to discourage use that is found to be harmful or wasteful.

Quality of the Data To Examine Off-label Prescribing

A fundamental issue in the future assessment of off-label utilization, and indeed in assessment of pharmaceutical utilization generally, is the availability of evidence that relates a prescribed drug to an indication. This question is directly related to existing policy surrounding the writing of prescriptions. Currently, physicians are not required to document the indication for which a drug is prescribed. Further, existing systems of coding (e.g., ICD-9) do not capture a variety of important nuances of clinical indication from the perspective of optimal care. Policy surrounding data collection, coding, and the prescription mechanism, need to be updated, particularly in the current environment where many health systems are moving to electronic health records. Future recording of a diagnosis for each drug prescribed is likely to be increasingly useful and feasible.

A related concern is that there is a pressing need for investment in a rigorous and complete national compendium that would provide a level of quality exceeding that of DRUGDEX and other existing compendia. Current compendia are simply unable to keep up with new evidence in order to maintain a timely, comprehensive source of information for practitioners. Providing prescribers with current evidence can only improve the quality of medication use, and more resources devoted to this area are needed.

Limitations

A number of limitations associated with our analysis should be kept in mind. We assessed off-label use involving drugs used for indications not approved by the FDA. We did not evaluate other forms of off-label use (e.g., prescribing to children and other special population, or drug use in dosages beyond those approved). Due to this study's scope, we did not attempt to conduct a systematic review of the clinical literature on specific drugs, because such a review was one of the potential types of research that might be influenced by our study findings. We instead relied on the synthesis available from the DRUGDEX compendium despite its potential limitations. Further, in coding off-label uses, our strategy was to apply our methods systematically to the available data on individual drugs. We did not consider whether another drug in the same class was approved nor did we give special consideration as to where a drug's use may be miscoded, as when an unusual label (e.g., hypertensive heart disease) might be used for a common related, but diagnostically distinct, condition (e.g., atrial fibrillation).

Although NDTI is designed to reflect outpatient physician visits in the US, the sample of physicians is not strictly random, as past respondents may continue their participation from year to year. Clinical information available in NDTI is limited and additional details would have been useful for ambiguous diagnoses (e.g., bronchitis) or where past drug therapy history was important. Although NDTI links a drug's use to a clinical condition, this feature may not account for instances where a cluster of conditions or symptoms prompt use of a particular drug.

Matching the ICD-9 codes from NDTI with the indications listed by the FDA and DRUGDEX is challenging because of differences in terminology, degree of specification, and clinically imprecise diagnoses. DRUGDEX may not always provide an optimal synthesis of available evidence. Not only are recent reports often not included,¹² but DRUGDEX often refers to unpublished studies cited in package inserts. Alternative compendia, however, often lack information on common off-label indications altogether.

In extending past research, our process of prioritizing individual drugs for further research necessarily involves several assumptions. Although we received substantial input from our technical expert panel, the framework we developed included subjective elements. However, our sensitivity analysis indicated that the framework was quite robust to substantial changes in model assumptions. We also were limited in the availability of reliable metrics for drug characteristics to include in our model. For example, Black Box warnings may not always reflect an equal magnitude of concern about drug safety and drug interactions. In prioritizing a list of drugs for future attention, we recognize that different forms of research and evaluation may be needed for different drugs on our list.

Conclusion

Future research into off label drug use should focus on drugs currently being prescribed widely with little or no evidence, particularly if the drugs are costly, known to have safety issues,

and are being marketed heavily. We have identified 14 drugs commonly used for off-label indications without strong evidence that merit priority in future research. While past research has highlighted the substantial frequency of off-label use without good evidence, we have identified and prioritized specific medications warranting attention. Future research to evaluate the efficacy and safety of off-label use should focus on these medications. Such research would initially take the form of meta-analyses to accumulate and assess past clinical trial research on these drugs. Subsequent research might involve observational studies of benefits or harms of off-label drug use. These in turn might prompt RCTs where the value of information appears to warrant it. Understanding how non-supported off-label use diffuses into widespread practice may help gauge the adequacy of current regulatory policies, as well as suggest new strategies for disseminating evidence-based comparisons of drug effectiveness and safety. Large scale attention to off-label use is also likely to require new tools, including: 1) new, clinically coherent methods for classifying clinical conditions, 2) a consolidated, continuously updated, and authoritative data-source on the evidence supporting off-label use, and 3) routine collection of clinical indications in the process of medication prescribing. The significant volume and cost of off-label drug use with inadequate or uncertain scientific evidence warrants substantial future efforts to address this issue.

Acknowledgments

The authors gratefully acknowledge constructive input from Anne Trontell as well as the members of the Technical Expert Panel: Laurie Burke (FDA), Daniel J. Carlat (Private Practice Psychiatrist, Newburyport MA), Marc Berger (Merck and Company), David Hopkins (Pacific Business Group on Health), Sigalit Kaplan (FDA), Diane Kennedy (FDA), Karl Matuszeski (University HealthSystem Consortium), David C. Radley (Dartmouth University), and Robert Temple (FDA). The authors also acknowledge the helpful administrative assistance of Katherine Clarke, Rachel Johns, and Hannah Park.

References

1. Chen H, Deshpande AD, Jiang R, Martin BC. An epidemiological investigation of off-label anticonvulsant drug use in the Georgia Medicaid population. *Pharmacoepidemiol Drug Saf* 2005;14:629-38.
2. Ma J, Lee K-V, Stafford RS. Depression treatment during outpatient visits by U.S. children and adolescents. *J Adolesc Health* 2005;37(6):434-442.
3. Linder JA, Huang ES, Steinman MA, Gonzales R, Stafford RS. Fluoroquinolone prescribing in the United States, 1995-2000. *Am J Med*, 2005;118:259-68.
4. McIntyre J, Conroy S, Avery A, Corns H, Choonara I. Unlicensed and off-label prescribing of drugs in general practice. *Arch Dis Child* 2000;83:498-501.
5. Loder EW, Biondi DM. Off-label prescribing of drugs in specialty headache practice. *Headache* 2004;44:636-41.
6. Donofrio PD, Busis NA. Regulatory and reimbursement issues in treating patients with immune-mediated neuropathies. *Neurology* 2002;59:S41-5.
7. Strom BL, Melmon KL, Miettinen OS. Post-marketing studies of drug efficacy: why? *Am J Med* 1985;78:475-80.
8. Adams C, Young A. Risky Rx: A Knight Ridder Investigation. Newspaper article series, Nov 2-4, 2003. Available at: <http://www.anderson.ucla.edu/documents/areas/adm/loeb/04h17riskyrx.pdf>. Accessed September 18, 2008.
9. Radley D, Finkelstein S, Stafford RS. Off-label prescribing among office-based

Effective Health Care Research Report Number 12

- physicians. *Arch Intern Med* 2006;166:1021-1026.
10. Claxton K. Bayesian approaches to the value of information: implications for the regulation of new pharmaceuticals. *Health Econ* 1999; May;8(3):269-74.
 11. IMS Health. National Disease and Therapeutic Index, Medication Reference File, 2001. Plymouth Meeting, PA: IMS Health; 2007.
 12. Zell ER, McCaig LF, Kupronis BA, Besser RE, Schuchat A. A comparison of the National Disease and Therapeutic Index and the National Ambulatory Medical Care Survey to evaluate antibiotic usage. Proceedings of the Section on Survey Research Methods American Statistical Association. Alexandria, VA;2000:840-845.
 13. Stafford RS, Radley DC. The underutilization of cardiac medications of proven benefit, 1990 to 2002. *J Am Coll Cardiol* 2003;41:56-61.
 14. International Classification of Diseases, Ninth Revision, Clinical Modification. Washington DC; Public Health Service, US Department of Health and Human Services; 1988.
 15. PDR[®] Electronic Library[™] [Internet database]. Greenwood Village, CO: Thomson Micromedex. Updated periodically.
 16. Food and Drug Administration, Center for Drug Evaluation and Review. Available at: <http://www.fda.gov/cder>. Accessed April 26, 2007.
 17. IMS National Prescription Audit Data, IMS Health (Plymouth Meeting PA).
 18. DRUGDEX[®] System [Internet database]. Greenwood Village, CO: Thomson Micromedex. Updated periodically.
 19. USP DI[®] Volume I: Drug Information for the Health Care Professional [Intranet database]. Version 5.1. Greenwood Village, Colo: Thomson Micromedex.
 20. FDA Orange Book Accessed April 24, 2007. Available at: <http://www.fda.gov/cder/ob/>.
 21. Electronic report obtained by Randall S. Stafford under a Freedom of Information Act request that the FDA provide information on all drugs approved prior to 1982.
 22. IMS Health Office Promotion Report, IMS Health (Plymouth Meeting, PA).
 23. TNS Media Intelligence. Ad\$ponder Survey. TNS, New York, NY: 2007.
 24. Shekelle P, Maglione M, Bagley S, et al. Comparative Effectiveness of Off-Label Use of Atypical Antipsychotics. Comparative Effectiveness Review No. 6. (Prepared by the Southern California/RAND Evidence-based Practice Center under Contract No. 290-02-0003.) Rockville, MD: Agency for Healthcare Research and Quality. January 2007.
 25. Sachs GS, Nierenberg AA, Calabrese JR, et al. Effectiveness of adjunctive antidepressant treatment for bipolar depression. *N Engl J Med* 2007;356:1711-22.

Figure 1. Summary of relationship between DRUGDEX evaluation parameters and the categories of off-label use developed for this project

DRUGDEX RECOMMENDATION	DRUGDEX LEVEL OF EFFICACY					
	Effective <i>or</i> Favors Efficacy			Inconclusive <i>or</i> Ineffective		
I Recommended <i>or</i> IIa Recommended for most patients	A	B	C	A	B	C
IIb Recommended for selected patients	A	B	C	A	B	C
III Not Recommended	A	B	C	A	B	C

DRUGDEX STRENGTH OF EVIDENCE	
A	Good and/or consistent Randomized Controlled Trials (RCTs)
B	Less adequate or inconsistent RCTs
C	Non-RCT evidence

RESULTING CATEGORY OF OFF-LABEL USE FOR THIS PROJECT	
	Good Evidence
	Uncertain Evidence
	Inadequate Evidence
	Note: No Observations for this Combination

Table 1. Number of drug uses, by off-label status and level of supporting evidence, ranked by decreasing order of off-label uses with inadequate evidence, United States, July 2005 through June 2007, IMS Health National Disease and Therapeutic Index

Rank	Generic Name	DRUG USES IN THOUSANDS				
		On-Label	Total Off-Label	Off Label Adequate Evidence ^a	Off Label Uncertain Evidence ^b	Off Label Inadequate Evidence ^c
1	quetiapine	2120	6507	0	0	6507
2	warfarin	13000	5325	0	0	5325
3	clonazepam	1297	4235	0	0	4235
4	escitalopram	21376	3580	0	0	3580
5	gabapentin	446	6334	84	3423	2827
6	promethazine	208	2732	0	0	2732
7	risperidone	2366	4034	0	1465	2569
8	digoxin	4701	2561	0	0	2561
9	lorazepam	3841	2490	0	0	2490
10	lisinopril	32140	2601	227	0	2374
11	prednisolone	6024	2359	0	0	2359
12	trazodone	2858	2166	0	0	2166
13	zolpidem	5114	2085	0	0	2085
14	sertraline	16617	2071	0	0	2071
15	bupropion	12196	1929	0	0	1929
16	acetamino/tramadol	2925	1802	0	0	1802
17	dexamethasone	784	2455	0	803	1652
18	metronidazole	2644	1648	0	0	1648
19	amitriptyline	848	1918	70	308	1540
20	albuterol	11749	8373	6881	0	1492
21	venlafaxine	11475	1472	0	0	1472
22	vancomycin	1089	1394	0	0	1394
23	esomeprazole	10237	1385	0	0	1385
24	montelukast	8201	1353	0	0	1353
25	celecoxib	6998	1296	0	0	1296
27	levothyroxine	27689	1143	0	0	1143
32	olanzapine	3721	1197	0	168	1029
38	divalproex	1416	6112	406	4838	868
45	lamotrigine	5756	1776	1112	0	664
46	duloxetine	7759	654	0	0	654
75	aripiprazole	2915	293	0	0	293
78	epoetin alfa	199	3402	0	3079	274

^a Off label is defined by indication based on NDTI survey data.

^b Off label uses in these columns are divided into evidence categories based on Drugdex categories of evidence and recommendation for use (see Figure 1).

Table 2. Indices for volume of off-label use and other drug characteristics, United States, July 2005 through June 2007, IMS health national therapeutic and disease index

Rank	Generic Name	PARAMETER INDICES (VARY FROM 0.00 to 1.00)						
		Volume Inadequate Evidence ^a	Volume Inadequate & Uncertain Evidence ^b	Drug Safety ^c	Drug Cost ^d	Corporate Promotion ^e	Market Entry ^f	Cost + Promotion + Entry ^g
1	quetiapine	1.00	1.00	1.00	0.24	0.19	0.50	0.31
2	warfarin	0.82	0.82	1.00	0.03	0.01	0.00	0.01
3	clonazepam	0.65	0.65	0.00	0.03	0.01	0.00	0.01
4	escitalopram	0.55	0.55	1.00	0.09	0.47	1.00	0.52
5	gabapentin	0.43	0.96	0.00	0.12	0.01	0.50	0.21
6	promethazine	0.42	0.42	0.00	0.02	0.00	0.00	0.01
7	risperidone	0.39	0.62	1.00	0.25	0.12	0.50	0.29
8	digoxin	0.39	0.39	0.00	0.01	0.00	0.00	0.00
9	lorazepam	0.38	0.38	0.00	0.03	0.00	0.00	0.01
10	lisinopril	0.36	0.36	1.00	0.03	0.00	0.00	0.01
11	prednisilone	0.36	0.36	0.00	0.01	0.00	0.00	0.00
12	trazodone	0.33	0.33	1.00	0.02	0.00	0.00	0.01
13	zolpidem	0.32	0.32	0.00	0.10	0.73	0.50	0.45
14	sertraline	0.32	0.32	1.00	0.11	0.44	0.00	0.18
15	bupropion	0.30	0.30	1.00	0.13	0.57	0.00	0.23
16	acetamino/tramadol	0.28	0.28	0.00	0.05	0.06	0.50	0.20
17	dexamethasone	0.25	0.38	0.00	0.02	0.00	0.00	0.01
18	metronidazole	0.25	0.25	1.00	0.04	0.00	0.00	0.01
19	amitriptyline	0.24	0.28	1.00	0.01	0.00	0.00	0.00
20	albuterol	0.23	0.23	0.00	0.04	0.05	0.00	0.03
21	venlafaxine	0.23	0.23	1.00	0.15	0.36	0.50	0.34
22	vancomycin	0.21	0.21	0.00	0.36	0.00	0.00	0.12
23	esomeprazole	0.21	0.21	0.00	0.19	1.00	0.50	0.56
24	montelukast	0.21	0.21	1.00	0.12	0.74	0.50	0.45
25	celecoxib	0.20	0.20	1.00	0.14	0.49	0.50	0.38

Table 2. Indices for volume of off-label use and other drug characteristics, United States, July 2005 through June 2007, IMS health national therapeutic and disease index—con.

Rank	Generic Name	PARAMETER INDICES (VARY FROM 0.00 to 1.00)						
		Volume Inadequate Evidence ^a	Volume Inadequate & Uncertain Evidence ^b	Drug Safety ^c	Drug Cost ^d	Corporate Promotion ^e	Market Entry ^f	Cost + Promotion + Entry ^g
27	levothyroxine	0.18	0.18	1.00	0.02	0.08	0.00	0.03
32	olanzapine	0.16	0.18	1.00	0.41	0.12	0.50	0.34
38	divalproex	0.13	0.88	1.00	0.16	0.09	0.00	0.08
45	lamotrigine	0.10	0.10	1.00	0.28	0.09	0.50	0.29
46	duloxetine	0.10	0.10	1.00	0.14	0.43	1.00	0.52
75	aripiprazole	0.05	0.05	1.00	0.43	0.15	1.00	0.53
78	epoetin alfa	0.05	0.52	1.00	1.00	0.13	0.50	0.54

- ^a The index value for volume with inadequate evidence for each drug is defined as the number of off label uses with inadequate evidence divided by the maximum number of off label uses with inadequate evidence found in any drug in the NDTI data set.
- ^b The index value for Volume with Inadequate & Uncertain Evidence is analogous to that for Volume with Inadequate evidence, but is based on the sum of uses with inadequate evidence and uses with uncertain evidence (see Figure 1).
- ^c The Drug Safety index value is coded as 1 for drugs with a black box warning, 0.5 for drugs that had an official FDA warning letter, and zero otherwise.
- ^d The Drug Cost Index value is the mean prescription cost of the drug divided by the maximum mean prescription cost for any drug on the list.
- ^e The Corporate Promotion index value is equal to the total dollar amount spent on direct to consumer advertising added with the total dollar amount spent on physician advertising, divided by the maximum amount among drugs on the list.
- ^f The Market Entry index value is coded as 1 for drugs that entered within the past 3 years, 0.5 for drugs that entered 4 to 13 years ago, and zero for drugs that entered more than 13 years ago and zero for any generic drug regardless of entry year.
- ^g The Combined Cost and Marketing index value is the equal weighted average of the Drug Cost, Corporate Promotion, and Market Entry index values.

Table 3. Sensitivity analysis of rankings. Base model and variant models. United States, July 2005 through June 2007, IMS Health National Disease and Therapeutic Index

Model	Base Model ^a	Weighted by Evidence ^b	Fully Include Uncertain ^c	Emphasize Safety/Cost	Exclude Not Found Uses ^d	Volume + Safety No Cost	Inadequate Volume Only	Aggregate Cost ^e
Inadequate Evidence Uses	.50	.333	.25	.333	.50	.50	1.00	1.00
Uncertain Evidence Uses		.167	.25					
FDA Safety Concerns	.25	.25	.25	.333	.25	.50		
Cost + Promotion + Entry	.25	.25	.25	.333	.25			1.00
Rank								
1	quetiapine	quetiapine	quetiapine	quetiapine	quetiapine	quetiapine	quetiapine	quetiapine
2	warfarin	warfarin	divalproex	escitalopram	risperidone	warfarin	warfarin	risperidone
3	escitalopram	escitalopram	warfarin	warfarin	bupropion	escitalopram	clonazepam	vancomycin
4	risperidone	risperidone	escitalopram	risperidone	clonazepam	risperidone	escitalopram	olanzapine
5	montelukast	epoetin alfa	epoetin alfa	montelukast	metronidazole	lisinopril	gabapentin	gabapentin
6	bupropion	divalproex	risperidone	duloxetine	lisinopril	trazodone	promethazine	escitalopram
7	sertraline	montelukast	gabapentin	epoetin alfa	epoetin alfa	sertraline	risperidone	epoetin alfa
8	venlafaxine	bupropion	montelukast	celecoxib	levothyroxine	bupropion	digoxin	esomeprazole
9	celecoxib	sertraline	bupropion	aripiprazole	aripiprazole	metronidazole	lorazepam	bupropion
10	lisinopril	venlafaxine	sertraline	venlafaxine	duloxetine	amitriptyline	lisinopril	sertraline
11	duloxetine	celecoxib	venlafaxine	bupropion	escitalopram	venlafaxine	prednisilone	venlafaxine
12	trazodone	lisinopril	celecoxib	sertraline	trazodone	montelukast	trazodone	zolpidem
13	olanzapine	duloxetine	lisinopril	olanzapine	meloxicam	celecoxib	zolpidem	lamotrigine
14	epoetin alfa	olanzapine	duloxetine	olmesartan	lamotrigine	levothyroxine	sertraline	celecoxib
15	aripiprazole	trazodone	olanzapine	meloxicam	montelukast	olanzapine	bupropion	montelukast

^a The base analyses ranks the drugs based on the weighted sum of the Volume with Inadequate Evidence Index, the Safety Index, and the Combined Cost and Marketing Index, with weights of 0.5, 0.25 and 0.25 respectively.

^b The ranking in this column is similar to the base ranking except that the volume of uses with uncertain evidence are considered along with the volume of uses with inadequate evidence with uses with uncertain evidence receiving half the weight of uses with inadequate evidence.

^c The ranking in this column is similar to the base ranking, however uses with uncertain evidence are considered with the same weight as uses with inadequate evidence.

^d The rankings in this column are similar to the base ranking except that uses that were not found in the Drugdex compendia were dropped rather than counted as uses with inadequate evidence.

^e The rankings in this column are based on an index formed by multiplying, for each drug, the volume with inadequate evidence index by the cost and marketing index.

Table 4. Clinical indications associated with largest number of uses for approved and off-label uses with inadequate evidence, United States, July 2005 through June 2007, IMS health national therapeutic and disease index.

Rank	Generic Name	Most Common On-Label Use	Most Common Off-label Indication	Second Most Common Off-label Indication	Third Most Common Off-label Indication
1	quetiapine	Schizophrenia	Bipolar, maintenance ^{a,c}	Depression	Anxiety
2	warfarin	Atrial fibrillation	<i>Hypertensive heart dis</i>	Coronary heart disease	<i>Ill-defined acute CVA</i>
3	escitalopram	Depression	Bipolar ^b	OCD	Adjustment reaction
4	risperidone	Schizophrenia	Bipolar, maintenance ^a	Depression	Psychosis
5	montelukast	Asthma	COPD	Sinusitis	
6	bupropion	Depression	Bipolar ^b	Unsp affective psychoses	
7	sertraline	Depression	Bipolar ^b	Adjustment reaction	
8	venlafaxine	Depression	Bipolar ^b	Schizophrenia	
9	celecoxib	Joint Sprain/ Strain	Fibromatosis	Enthesopathy	
10	lisinopril	Hypertension	Coronary artery disease	Diabetes	
11	duloxetine	Depression	Anxiety	Bipolar ^b	
12	trazodone	Depression	Sleep disturbance	Anxiety	Bipolar ^b
13	olanzapine	Schizophrenia	Depression	Dementia	Anxiety
14	epoetin alfa	Chron renal failure	Anemia of chronic disease	Myelodysplastic syndrome	

Footnotes:

Diagnoses that were not listed in Micromedex Drugdex (and thereby consider to have inadequate evidence are listed in *Italics*)

^a Excludes those diagnoses of Bipolar Affective Disorder (Manic-Depressive Disorder) listed as including Mania, Depression, or Mixed Presentation.

^b Includes all diagnoses listing Bipolar Affective Disorder.

^c The number of uses for this off-label drug indication exceeded that for its most common on-label indication.

Appendix

The purpose of this appendix is to augment the detailed study description above. The appendix includes extended descriptions of the TEP, the role of the TEP, and extended descriptions of particular methods in the selection and coding of variables. The appendix also includes an extended list of conclusions and a section on translation of findings.

Further Description of the TEP

Given the complexity of weighing diverse considerations in prioritizing drugs for future action related to off-label use, we formed a technical expert panel to provide input to the researchers. The following criteria were used for selection:

- 1) Based on their experience, potential panel members were familiar with the issue of off-label prescribing and its implications.
- 2) Members either represented specific key stakeholders (see below) and/or provided technical expertise beyond that available among the research team.
- 3) The choice of panel members was acceptable to all members of the research team.

Particular stakeholders that we identified as being critical included practicing generalist physicians, practicing specialist physicians, large group purchasers, the health insurance industry, pharmaceutical benefits managers, the pharmaceutical industry, and drug regulators from the FDA. We sought outside expertise that was relevant to the analytic issues faced in investigating off-label use, expertise in psychiatry, knowledge of drug characteristics and their potential role as complementing our quantitative methods in the prioritization process. Members of the research team provided expertise in pharmacy management, medical ethics, health economics, data analysis of prescribing patterns, and research methods, as well as representing generalist physicians and academic research institutions. The research team (Table A1) and the technical expert panel (Table A2) are listed, as well as their affiliations, stakeholder representation, and areas of expertise.

Further Description of Methods and Coding

In the process of making use of the DRUGDEX compendium, a particular challenge was matching the ICD-9 coding of indications reported by physicians in NDTI with the diagnostic labels used by DRUGDEX (including those used by the FDA in describing approved indications).²⁴ Given limitations of both ICD-9 and DRUGDEX diagnostic labels, there were many instances in which there was an inexact correspondence between labels. We developed several rules of thumb to aid in this matching process between ICD-9 disease diagnoses and the DRUGDEX indication labels:

- 1) We matched labels when both used different terminology to refer to the same condition (e.g., ischemic heart disease and coronary atherosclerosis, seizures and epilepsy).
- 2) We matched labels when the DRUGDEX label represented a prominent symptom associated with a particular ICD-9 diagnosis (e.g., pain as an indication was matched to a post-operative status as a diagnosis).

3) Likewise, we matched labels when the ICD-9 diagnosis represented a prominent symptom or clinical sign associated with a particular indication (e.g., a diagnosis of chest pain was matched with an indication of angina or rash with dermatitis).

4) We also matched instances where the diagnosis and label were pathologically related (e.g., we matched a diagnosis of herpes zoster to an indication of post-herpetic neuralgia).

5) We matched labels when a more specific ICD-9 diagnosis was provided that fit within a more general indication (e.g., hypertensive heart disease was matched to an indication of hypertension, a diagnosis of ischemic heart disease was matched to an indication for atherosclerosis).

6) We also matched labels when a diagnosis was more broad than the indication, but where a majority of the patients described by the general term were likely to have the specific condition given the use of a specific drug (e.g., a diagnosis of urinary tract infection was matched to an indication of cystitis for antibiotics and a diagnosis of bronchitis not specified as acute or chronic was matched to an indication of chronic obstructive pulmonary disease when an inhaler was being used).

7) We did not match indication and diagnosis where a non-specific diagnosis was used and the indication was quite specific. We did not match if far less than half of patients with the non-specific diagnosis would be expected to have the specific indication. For example, we did not match a diagnosis of bipolar affective disorder, unspecified with an indication of bipolar disorder with mania because most outpatients treated for bipolar are unlikely to be experiencing mania.

We realize that these guidelines may lead to some misclassification. In particular, uses which may in fact be off-label may have been considered on-label using these guidelines.

Further Description of Technical Expert Panel Input

Additional Drug Characteristics for Prioritization

Based on both our review of past research and the input from the TEP, we identified a series of drug characteristics that might be expected to alter the priority of a drug independent of its overall volume of off-label use without good evidence. This process involved two steps: 1) conceptualization of the features of particular drugs that might make their off-label use more or less concerning compared to other drugs, and 2) development of specific metrics aimed at allowing measurement of the most important concepts.

Conceptual Drug Characteristics

Table A3 below shows the initial criteria. We used a questionnaire to survey the TEP as to their ratings (from 1 to 5 where 5 was the highest level of importance in terms of prioritizing future research) for each criterion

From the questionnaire, only those items where more than 50% of the respondents rated the item as a 4 or a 5 in importance were kept. The survey results suggested consideration of 1) serious side effects and serious drug-drug interactions, 2) high daily cost, 3) high relative daily cost, 4) drug is less than 5 years old, 5) drug is extensively marketed to consumers, and 5) drug is extensively marketed to physicians.

The technical expert panel (TEP) provided input throughout the development of this project's analytic plan. As described above, they helped select the drug characteristics to include in the model and helped determine the relative weighting.

Once the initial modeling process was completed, the TEP was asked to re-rank all drugs that had appeared within the top 25 on any of the nine priority lists. The TEP was asked to complete a second survey asking them to place the drugs in the final lists into four tiers:

- 1) Drugs most urgently in need of future research regarding their off-label use.
- 2) Drugs important to examine for future research regarding their off-label use but not of the highest degree of urgency.
- 3) Drugs likely to be of interest for future research regarding their off-label use though not as important as 1 and 2 above.
- 4) Drugs not important for future research into off-label use.

Drugs were kept in the final top tier if the majority of the responding TEP placed them there. Those drugs that were listed by an individual in the top tier, but did not have a majority of the TEPs' support were considered to have been listed on that individual's second tier. Drugs were kept in the second tier if the majority of the responding TEP placed them there (including the drugs shifted from the top tier). Similarly, for the other two tiers. Drugs that did not have a majority for any of the first three tiers were put in tier four. The TEP also provided input into an assessment of the initial models.

There were five respondents to the final TEP survey. The majority of TEP members favored the use of the fully specified model 9, which included all drug characteristics. Other members felt that drug characteristics related to cost, marketing and market entry should not be included. Based on the responding TEPs' survey input, a set of nine drugs were identified as Tier 1, or top priority: quetiapine, warfarin, risperidone, escitalopram, sertraline, venlafaxine, olanzapine, bupropion and valproate.

Unfortunately, further TEP follow-up on final lists was not possible. However, many of the drugs that the TEP helped identify can be found on the final list.

Extended Conclusions

- 1) Our strategy for ranking medications used for off-label indications proved successful in generating a priority list that accounted for key factors, including the volume of off-label use without good evidence, drug safety concerns, and relevant drug characteristics.
- 2) Future research to evaluate the efficacy and safety of off-label use should focus on the most commonly prescribed antipsychotic and antidepressant medications.
- 3) In concert with a focus on specific drugs, future research efforts also should focus on ascertaining the current range of therapeutic approaches for bipolar disorder.
- 4) An important next research step will be to conduct a series of meta-analyses to accumulate and assess past clinical trial research support for the drugs identified in this report, including an assessment of where the largest gaps exist.
- 5) Subsequent research should seek to provide definitive clinical trial data on the safety and efficacy of those indications and drugs identified as priorities in the drug specific meta-analyses.

6) New tools are needed to aid future assessments of off-label use. In particular, new methods for classifying clinical conditions are necessary so that it is feasible to match physician reported clinical diagnoses with the diagnostic labels used by the FDA and drug compendia.

7) Given the shortcomings of current compendia, a consolidated, frequently updated, and authoritative data-source on the level of evidence supporting common off-label use is needed, including data organized or readily searchable by clinical diagnosis.

8) Future research also should focus on greater understanding of the factors which affect the diffusion of off-label use without good evidence. Such an analysis may suggest new strategies for acquiring and utilizing new evidence on effectiveness and safety as well as educational and translational methods for making this evidence readily available and understandable.

Additional Report Section: Translation of Findings

We measured counts of off-label utilization based on indication, with off-label uses further characterized as having good, uncertain, or inadequate scientific support based on an established drug compendia (DRUGDEX). We developed eight different prioritized listings based on the volume of inadequate and uncertain evidence off-label uses and available information on FDA-identified safety concerns, drug price, the extent of direct-to-consumer and physician office-based promotion, and time since market entry.

Our top priority drugs, quetiapine, warfarin, escitalopram, risperidone, montelukast, bupropion, sertraline, venlafaxine, celecoxib, lisinopril, duloxetine, trazodone, olanzapine, and epoetin alfa are strong candidates for further investigation.

It is important to note that our findings are not sufficient to determine the desirability or outcomes of off-label use. However, the findings do suggest significant and numerous gaps in evidence for medications that are being used frequently. Hence, our findings suggest that there is value in research aimed at evaluating the safety and outcomes associated with off-label use of the priority drugs. The primary goal of that research should be to distinguish off-label uses that add clinical value from those that do not.

Effective Health Care Research Report Number 12

Table A1. Background information regarding the research team

Name	Affiliation	Expertise/Stakeholder
Surrey Walton, PhD	University of Illinois, Chicago	Health Economics
Randall Stafford, MD, PhD	Stanford University	Pharmacoepidemiology/Generalist Physicians
Ky-Van Lee, PhD	Stanford University	Analysis of Prescribing Data
Glen Schumock, PharmD, MBA	University of Illinois, Chicago	Pharmacy Management
G. Caleb Alexander, MD, MS	University of Chicago	Chronic Disease, Medical Ethics/Generalist Physicians
David Meltzer, MD, PhD	University of Chicago	Health Economics/Generalist Physicians

Table A2. Background information regarding the Technical Expert Panel

Name	Affiliation	Stakeholder/Expertise
David Hopkins, Ph D	Pacific Business Group on Health	Business Community/ Quality of Care
Karl Matuszeski, PharmD	University HealthSystem Consortium	Academic Medical Centers/Pharmacologist
David C. Radley, MPH	Dartmouth University	Expertise in Off-label Use
Robert Temple, MD	Food and Drug Administration	Regulator/FDA Medical Policy
Sigalit Kaplan, PhD, RPh	Food and Drug Administration	Regulator/Pharmacist/Drug Safety
Diane Kennedy, RPh, MPH	Food and Drug Administration	Regulator/Drug Analysis
Laurie Burke, RPh, MPH	Food and Drug Administration	Regulator/Drug Labeling and Marketing
Daniel J. Carlat, MD	Practicing Psychiatrist	Specialist Physicians/ Psychopharmacotherapy
Marc Berger, MD	Merck & Company	Pharmaceutical Industry/ Outcomes Research

Table A3. TEP survey questions regarding the relative importance of potential factors for prioritizing off label research

1) The drug has serious side effects.**
2) The drug has frequent side effects.
3) The drug has serious drug-drug interactions.**
4) The drug has frequent drug-drug interactions.*
5) The drug is approved in other countries for its primary off-label indication(s).
6) Another drug in the same class is FDA-approved for the drug's primary off-label indication(s).
7) The drug's primary off-label indication(s) is for condition(s) with non specific symptoms such as depression or chronic pain.
8) The drug, when used for its primary off-label indication(s), acts to modify symptoms rather than modify disease progression.
9) The daily treatment cost of the drug is high.*
10) The cost of the drug is high relative to an approved substitute treatment.*
11) The drug is new (i.e., less than five years).*
12) The drug is being used for its primary off-label indication(s) mainly as a second-line agent.
13) The drug's primary off-labeled indication(s) leads to short-term rather than long-term use.
14) The drug is extensively marketed directly to consumers.*
15) The drug is extensively marketed to physicians.*
16) The primary off-label indication(s) is of special relevance to the Medicare or Medicaid populations.
17) The primary off-label indication(s) is of special relevance to other vulnerable populations.

** Denotes Factors where more the 50% of the TEP rated the importance of the factor as a 5 on a 1 to 5 scale.

* Denotes factors where more than 50% of the TEP rated the importance of the factor as a 4 or a 5 on a 1 to 5 scale.