MINI-SENTINEL WHITE PAPER

METHODS TO EVALUATE THE IMPACT OF FDA REGULATORY ACTIONS

Prepared by: Becky A. Briesacher[1,3], PhD; Stephen B. Soumerai[2], ScD; Fang Zhang[2], PhD; Darren Toh[2], ScD; Susan Andrade[1,3], ScD; Joann L. Wagner[3], MSW; Jerry H. Gurwitz[1,3], MD

Author Affiliations: [1]Division of Geriatric Medicine, University of Massachusetts Medical School, Worcester, MA; [2]Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA; [3]Meyers Primary Care Institute, a joint endeavor of Fallon Community Health Plan, University of Massachusetts Medical School, and Reliant Medical Group, Worcester, MA.

Updated November 2012

Mini-Sentinel is a pilot project sponsored by the U.S. Food and Drug Administration (FDA) to inform and facilitate development of a fully operational active surveillance system, the Sentinel System, for monitoring the safety of FDA-regulated medical products. Mini-Sentinel is one piece of the Sentinel Initiative, a multifaceted effort by the FDA to develop a national electronic system that will complement existing methods of safety surveillance. Mini-Sentinel Collaborators include Data and Academic Partners that provide access to healthcare data and ongoing scientific, technical, methodological, and organizational expertise. The Mini-Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223200910006I.

Mini-Sentinel White Paper

Methods To Evaluate The Impact Of FDA Regulatory Actions

Table of Contents

I. ABBREVIATIONS

DESI	Drug Efficacy Study Implementation
FDA	Food and Drug Administration
REMS	Risk Evaluation and Mitigation Strategies

II. INTRODUCTION

The Mini-Sentinel White Paper on Methods to Evaluate the Impact of FDA Regulatory Actions core workgroup met weekly over a 7-month period to answer two questions about evaluations of impacts of U.S. Food and Drug Administration (FDA) regulatory actions: 1) what methods have been applied to date; and 2) what methods would represent best practices for conducting evaluations of FDA regulatory actions? This white paper presents the findings from our literature review to address the first question and from our expert panel consensus for the second question.

A. BACKGROUND

Few studies have assessed the impact of FDA regulatory actions on state-of-the-art research methods that may be useful for evaluating FDA regulatory actions. Quasi-experimental designs such as interrupted time series and regression discontinuity share many elements of experiments including control groups, pre-test post-test measures, intention-to-treat analyses, and explicit mechanisms for inferring counterfactual scenarios---i.e., what would the outcome have been without the FDA action.[1] Novel statistical methods have also been developed to address special circumstances such as sequential analyses when the data on outcomes are still accumulating or extended Cox models, when the hazard ratios of risk will vary with time.[2,3]

B. STUDY OBJECTIVES

The goal of this study was to understand how state-of-the-art research methods may be used to evaluate the impact of FDA regulatory actions. Our specific objectives were:

1. To characterize prior research approaches used to evaluate FDA regulatory actions, particularly the strengths and weaknesses of previous methodologies.

2. To recommend research designs and statistical approaches that may be useful for evaluating FDA regulatory actions and descriptions of the advantages and disadvantages of each approach.

3. To produce a white paper that summarizes our findings and recommendations and to produce a manuscript suitable for publication.

C. GENERAL APPROACH

We conducted this study in two phases. In the first phase, we characterized prior research approaches by searching MEDLINE for all published research studies on FDA regulatory actions. We reviewed these studies according to selected Cochrane Collaboration criteria for systematic reviews of nonrandomized studies. These included: (a) the weaknesses of the designs to control for threats to validity, (b) the execution of the studies through a careful assessment of their risk of bias, (c) the potential for selection bias and confounding, and (d) the potential for reporting biases, including selective reporting of outcomes. Explicit criteria were used to identify those study designs that would provide reasonably valid effect estimates in our final summary of the published literature.

We also reviewed unpublished FDA documentation of evaluations presented at internal FDA advisory committee meetings. This documentation covered a range of various medical products and regulatory actions and was identified by the FDA.

In the second phase of this project, we identified preferred methodologies for evaluating FDA regulatory actions based on our extensive prior work in policy evaluation and previous evaluations in the literature. These methodologies were selected based on their ability to control for threats to internal validity, statistical performance as well as practical considerations. Sources for prior approaches included seminal well-controlled evaluations of relevant non-FDA policies. Preference was given to methods that could be used with administrative health plan data, such as found in the Mini-Sentinel Distributed Database.

In the end, we produced a comprehensive assessment of previous research methods used to evaluate FDA regulatory actions and describe appropriate alternatives. This white paper summarizes our process and results.

III. METHODS

A. LITERATURE REVIEW

1. Search Strategy

We conducted a literature search utilizing Ovid MEDLINE (see Appendix 1 for search strategy and results). We searched for literature published between 1948 and the first week of August, 2011, which included the term "United States Food and Drug Administration" and an FDA regulatory action (e.g., labeling change (label$), warning (warn$), or REMS); one hundred ninety-five published articles were identified. Additionally, the FDA workgroup members identified additional seminal literature (published and internal FDA literature) for inclusion in the review. Core workgroup members also manually searched the bibliographies of reviewed papers to identify any additional relevant literature (there was some overlap in the results of these search strategies) resulting in a total of 243 published papers and internal FDA documents for review.

2. Selection Criteria

The core workgroup (see page 46 for a list and description of members) drafted selection criteria for abstract and full-text manuscript review to identify relevant studies; the criteria were discussed with the FDA workgroup and revised based upon input of workgroup members. This process was iterative until the entire project workgroup agreed on the selection criteria and had no further feedback. Selection criteria are outlined in Appendix 2 and are an adaptation of the Cochrane steps for defining study selection criteria.[4]

3. Review of Literature

The review of the literature was a three-step process. First, all available abstracts were reviewed by at least one core workgroup member (BAB, SA, FZ). Second, all literature that met inclusion criteria or that could not be excluded on the basis of available abstract information underwent a full review; full-text articles were obtained and each was reviewed by two core workgroup members (BAB, SA, FZ). The third step was employed if there were discrepancies between the two core workgroup members' full reviews; when workgroup members did not agree on the inclusion of an article, it was reviewed by a third core workgroup member and discussed until team consensus was reached. Appendix 3 summarizes the reasons for exclusion.

B. CONSENSUS PANEL

1. Selection Criteria and Conduct of Interviews

To identify methods that may be suitable for evaluating the impact of FDA regulatory actions, we conducted a series of interviews with our panel of experts. The expert panel consisted of six investigators with extensive experience in drug policy research and evaluation methodology. Our main methodology was a generalized version of the Delphi method, which helps to systematically prioritize and synthesize the experiences and knowledge of a panel of experts to achieve group consensus.[5] First, we established several criteria for eligible methods. We agreed that all nominated methods

should use research designs and analytic methods that produce strong internal validity, as defined by Shadish, Cook, and Campbell (2002).[1] Thus, we selected methods that incorporated control groups and pretest measures to support a counterfactual inference about what would have happened in the absence of the FDA regulatory action. We also agreed that the recommended methods should offer pragmatic utility. Therefore, we did not select methods requiring random assignment or primary data collection since both options seemed unlikely in the Mini-Sentinel. As a result, our list of suitable methods to conduct evaluations of FDA regulatory actions assumes that only secondary data are available for the evaluation, in either aggregate-level or person-level data formats. We also allowed for methods that address special circumstances, particularly for the following cases: 1) when post-policy data are still accumulating, 2) the policy is transient, and 3) the data capture is limited to only a few time points. These special circumstance methods are acceptable as only preliminary evaluations until more rigorous evaluations are available. Lastly, we considered several state-of-the-art methods that theoretically have potential for unbiased regulatory evaluations, although in practice these methods require strong assumptions. These methods are acceptable only in sensitivity analyses within the context of a rigorous study design.

Given these criteria, we polled all the panel members for contributions and then created a list of candidate methods. We discussed each of the nominated methods and filtered out redundancies and irrelevant content through an iterative process. We assembled a list of the most suitable methods and then characterized each method by necessary data elements, strengths, and weaknesses. We also identified applied examples published in the literature that could serve as models. Lastly, we presented brief overviews of these methods to the FDA for comment and questions.

IV. RESULTS

A. LITERATURE REVIEW

The literature review produced a total of 18 independent studies that evaluated the impact of FDA regulatory actions and met the criteria for inclusion[6-23]. The key characteristics of these articles are summarized in Appendix 4. In summary, these studies were published over a period of two decades (1987 to 2011), and the investigations ranged in size from a small study[10] with only 1,308 individuals to a large population-level study[12] of 61 million individuals. The studies evaluated a wide range of patient populations, including children and adolescents in 78%[6-10,12-14,16,17,19-21,23] of the studies and adults over the age of 65 in 50%[6-10,17,20,22,23] of the studies.

1. FDA Policies and Medical Products

Most of the studies evaluated the impact of FDA boxed warnings (50%)[8-10,12,15,17,19,21,23], although other FDA policies were also evaluated including labeling changes[8,10,11,19,20,22], advisory/safety warnings[12-18], dear doctor/healthcare professional letters[9,11,20,22], and product withdrawals[6,7,11]. The studies examined policies relating to a variety of medical products, although antidepressants was the most common product (44%)[12-17,19,21]. Other medical products include antipsychotic agents[20,22], antidiabetic drugs[11,18], cisapride[9], Drug Efficacy Study Implementation (DESI) drugs[6], pemoline[10], propoxyphene[23], terfenadine[8], and zomepirac[7].

2. Data and Unit of Analysis

The most common source of data for the FDA policy evaluations was privately-owned proprietary administrative databases available for purchase (61%) [9,10,12-17,19,21,23]. The next most common sources were publicly available Medicaid/ Veteran's Administrative/TRICARE data (33%)[6,7,11,18,20,22] and private health insurance/managed care claims data (6%)[8,10]. Just over a quarter of the evaluations used individual-level data (28%)[10,11,19,21,22]. Assessments of data completeness or quality were usually not well described but were at least mentioned in most (89%)[6,7,9,10,12-23] of the papers.

3. Primary Endpoints

The outcomes measured in the evaluations varied depending on the regulatory action of interest, and a few of the measures related directly to patient health. In evaluations of boxed warnings or withdrawn product policies (n=12), the studies usually assessed changes in the use of substitute medical products and services (92%)[6-11,15,17,19,21,23]. In evaluations of safety warnings[6-23], some of the studies examined changes in the occurrence of recommended laboratory monitoring (17%)[10,11,19]. Adverse events were measured in two studies (11%)[12,23], and contraindicated medical drugs/conditions in two other studies.[8,9]

4. Research Design and Analytic Methods

Overall, the research design of the studies fell into the single broad category of the quasi-experimental design of interrupted time-series/times series, although several studies also used pre-post designs for some analyses. Control groups were used in less than half of the studies (39%)[7,14,16,17,20,22,23]. The rigor and sophistication of the analytic methods varied greatly, ranging from segmented regressions in all of

the studies[6-23] to unadjusted pre-post tests of differences (6%).[8] Only two studies[20,22] described explicit sensitivity analyses to address potential confounders.

See Appendix 4 for summarized information on all included studies and Appendix 5 for details of each study included.

B. NOVEL METHODS

The consensus panel identified eight research designs and analytic methods (Appendix 6). Under the category of research designs and analytic *methods with strong internal validity*, we identified two quasi-experimental designs --- the interrupted times series and the regression discontinuity design --- and the extended Cox model statistical method. All three approaches offer methodological advantages in their strong ability to control for many potential threats to validity.

Research Design. In the case of the interrupted time series design[1], the pre-intervention data series is used to decrease uncertainty about whether a policy intervention is associated with true change in outcomes. **The key assumption of this design is that extrapolating the pre-intervention level and trend correctly reflects the (counterfactual) outcome that would have occurred had the intervention not happened.** Interrupted times series may be conducted with person-level or aggregate-level longitudinal data. Data series may be as short as 8 points (4 pre and 4 post) but the design becomes stronger with additional time points to establish trends. This study design is among the most robust of the quasi-experimental designs, but it requires a priori knowledge of the diffusion process, especially in case of gradual interventions or delayed causation. Our literature review identified 18 studies with an interrupted time series design[6-23].

Research Design. In the case of the regression discontinuity design[1], changes in the pattern of the outcome at a pre-specified cutoff in an assignment variable are used to indicate a policy impact. This approach is most appropriate for a policy that applies to a patient population defined by a continuous measure with a fixed threshold (e.g., age >=65). The regression discontinuity design is a reasonably robust quasi-experimental design and may be especially useful when pre-exposure data are lacking or limited. This approach also requires knowledge of the correct functional form between the outcome and assignment variable, and it is less well known in the medical literature. Our literature review did not identify any study with a regression discontinuity design to assess FDA policies.

Analytic Method. The extended Cox model[24] uses heaviside functions to yield constant hazard ratios for different time intervals. This statistical method offers the analytic advantages of maximum likelihood (e.g., good large-sample properties) and can be used in various longitudinal study designs. This model has the ability to produce different hazard ratios of risk that are specific to the pre- and post-intervention periods. Extended Cox models do not require the proportional hazard assumption and it is robust to right censored data. Our literature review did not identify any study that used an extended Cox model to assess FDA policies.

All three of these novel methods can incorporate experimental design elements including intention-to-treat analyses and control groups.

Under the category of *methods suitable for special circumstances*, we identified three approaches: sequential analysis of gradually accruing data[2], difference-in-difference-in differences[25] and self-controlled case series[26,27].

> Research Design. Sequential analysis of gradually accruing data offers methodological advantages under the special circumstance when post-policy data are still accumulating. In this approach, the counterfactuals can be obtained from data from the pre-policy period (i.e., "historical controls"), or data from a concurrent control group not affected by the policy. A signal is generated if the likelihood ratio – calculated based on the ratio of the number of observed events to the number of expected events (the counterfactuals) – exceeds a predetermined value. The key aspect of this method is that the p-values are adjusted for looking at the data in a continuous fashion, or multiple testing. An appropriate example for this method would be an FDA policy that is associated with a potential safety concern that requires early detection.
>
> Research Design. The difference-in-difference-in differences design may be suitable under the special circumstance of limited time points in the data. A key component of this method is the use of multiple comparisons to isolate the true effect of the intervention. An appropriate example for this method would be an FDA policy that is assessed with only one or two pre- and post-intervention measurements of the outcome, but any observed differences are evaluated against any changes in two relevant comparators.
>
> Research Design. Self-controlled case series may be useful in the case of a transient policy involving an acute outcome. In this method, the study uses data on only the cases to examine the temporal association between a time-varying exposure and outcome. The exact scenario for the best use of this method is unclear; however, it may be suitable for studying the association between a temporal FDA policy (e.g., temporary suspension of a medical product) and an acute event, in which the risk periods are short.

All of these methods are likely to produce valid inferences especially if they explicitly control for pre-intervention trends.

Under the category of *state-of-the art methods with theoretical potential if assumptions can be met*, we identified confounder summary scores (e.g., propensity scores). This analytic method is best used with caution in sensitivity analyses until the plausibility of assumptions are better established. It must be clear that the score is created with a model that is correctly specified and not biased by important omitted variables. Because of the relative weaknesses of this approach the study team does not endorse its use except in unusual circumstances.

V. CONCLUSIONS

Our literature review produced the following conclusions. First, rigorous evaluations of the impact of FDA regulatory actions have been infrequent, especially relative to the large number of actions implemented by the FDA. Second, when FDA regulatory actions were evaluated minimal research design standards were often not employed. For instance, less than a quarter used a control group. Assessments of data completeness or quality were usually not well described but were at least mentioned in most (86%) of the papers.

Third, many of the assessments were limited in scope and examined only changes in the use of the targeted medical product. Studies that included a broader array of measures such as unintended impacts (e.g., increases in the use of substitute products and services) were uncommon. Studies that included outcomes measures relating directly to patient health and adverse events were even rarer. As a result, many evaluations of FDA regulatory actions used suboptimal research designs and analytic methods, making the results limited and susceptible to biased findings. Overall, this review revealed considerable gaps in the evidence-base making any assessment of the impact of FDA policies premature at this time.[28]

Our expert panel review produced the following conclusions. First, there are novel methods, especially quasi-experimental research designs, which may be useful for evaluating the impact of FDA regulatory actions. The methods identified by this panel can address many of the limitations of previously published evaluations. Second, the careful and consistent application of these novel methods can provide a valuable opportunity to achieve the goal of less biased assessments of FDA regulatory actions.

VI. RECOMMENDATIONS

To adequately evaluate the impact of FDA regulatory actions, researchers should use research designs and analytic methods that produce strong internal validity (e.g., include suitable control groups and pre-test measures to support a counterfactual inference about what would have happened in the absence of the FDA regulatory action). Accordingly, the panel recommends the following methods (details in Appendix 6):

A. METHODS WITH STRONG INTERNAL VALIDITY

Research design of interrupted time series: the pre-intervention data series is used to decrease uncertainty about whether a policy intervention is associated with true change in outcomes. Change is generally measured as a change in the level or in the slope of the trend. The key assumption of this design is that extrapolating the pre-intervention level and trend correctly reflects the (counterfactual) outcome that would have occurred had the intervention not happened. Interrupted times series may be conducted with person-level or aggregate-level longitudinal data. Data series may be as short as 8 points but the design becomes stronger with additional time points to establish trends. This study design is among the most robust of the quasi-experimental designs, but it requires a priori knowledge of the diffusion process, especially in case of gradual interventions or delayed causation. Our literature review identified 18 studies with an interrupted time series design[6-23].

Research design of regression discontinuity: changes in the pattern of the outcome at a pre-specified cutoff in an assignment variable are used to indicate a policy impact. This approach is most appropriate for a policy that applies to a patient population defined by a continuous measure with a fixed threshold (e.g., age >=65). The regression discontinuity design is a reasonably robust quasi-experimental design and may be especially useful when pre-exposure data are lacking or limited. This approach also requires knowledge of the correct functional form between the outcome and assignment variable, and it is less well known in the medical literature. Our literature review did not identify any study with a regression discontinuity design to assess FDA policies.

Analytic method of extended Cox models: heaviside functions are used to yield constant hazard ratios for different time intervals. This analytic approach offers the statistical advantage of maximum likelihood (e.g., good large-sample properties) as well as the ability to produce different hazard ratios of risk that are specific to the pre- and post-intervention periods. Extended Cox models do not require the proportional hazard assumption and it is robust to right censored data. Our literature review did not identify any study that used an extended Cox model to assess FDA policies.

B. METHODS SUITABLE FOR SPECIAL CIRCUMSTANCES

Research design of sequential analysis of gradually accruing data: offers methodological advantages under the special circumstance of when post-policy data are still accumulating. In this approach, the counterfactuals can be obtained from data from the pre-policy period (i.e., "historical controls"), or data from a concurrent control group not affected by the policy. A signal is generated if the likelihood ratio – calculated based on the ratio of the number of

observed events to the number of expected events (the counterfactuals) – exceeds a predetermined value. The key aspect of this method is that the p-values are adjusted for looking at the data in a continuous fashion, or multiple testing. An appropriate example for this method would be an FDA policy that is associated with a potential safety concern that requires early detection.

Research design of self-controlled case series: uses data on only the cases to examine the temporal association between a time-varying exposure and outcome. This method is best suited for studying association between temporary FDA policy and an acute event type of outcome, in which the risk periods are short.

Research design of difference-in-difference-in-differences: may be suitable under the special circumstance of limited time points in the data. A key component of this method is the use of multiple comparisons to isolate the true effect of the intervention. An appropriate example for this method would be an FDA policy that is assessed with only one or two pre- and post-intervention measurements of the outcome, but any observed differences are evaluated against any changes in two relevant comparators.

C. METHODS WITH POTENTIAL IF ASSUMPTIONS CAN BE MET

Analytic method of confounder summary scores (e.g., propensity scores): best used with caution in sensitivity analyses until the plausibility of assumptions are better established. It must be clear that the score is created with a model that is correctly specified and not biased by important omitted variables.

To effectively adopt these recommendations, the panel suggests the following action plan:

Test the performance of recommended methods in Mini-Sentinel data

Many of the methods identified in this review have not been applied to the Mini-sentinel data. A necessary first step is to evaluate the actual performance of these methods on selected case-study FDA regulatory actions and applied to Mini-Sentinel data. The challenges and benefits of this exercise should be carefully documented and shared with the Mini-Sentinel and FDA communities.

Establish Guidelines

Stakeholders in the FDA and Mini-Sentinel should work with academic institutions and professional organizations (e.g., AcademyHealth) to establish clear guidelines on appropriate research designs and analytic methods for policy evaluations (including guidance for specific types of situations, data sources).

Promote Educational Efforts

Stakeholders in the FDA and Mini-Sentinel should work with academic institutions and professional organizations to provide educational programs and materials for appropriate research designs and methods.

Centralize Information

The FDA and Mini-Sentinel should create a central repository for guidelines, educational programming, and materials on appropriate research designs and analytic methods.

VII. APPENDIX

A. APPENDIX 1: LITERATURE SEARCH STRATEGY AND RESULTS

Database: Ovid MEDLINE(R) 1948 to Week 1 of August, 2011

#	Search	Result
1	(United States Food and Drug Administration).mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier]	21208
2	warn$.mp.	17617
3	label$.mp. or Drug Labeling/ or REMS or withdrawal	496653
4	2 or 3	513011
5	1 and 4	2816
6	limit 5 to (clinical trial, all or clinical trial or comparative study or controlled clinical trial or evaluation studies or journal article or multicenter study or randomized controlled trial or "research support, american recovery and reinvestment act" or research support, nih, extramural or research support, nih, intramural or research support, non us gov't or research support, us gov't, non phs or research support, us gov't, phs)	2076
7	limit 6 to (english language and full text and humans)	262
8	from 7 keep 1-5, 7-13, 15-38, 40, 42, 44...	195

B. APPENDIX 2: SELECTION CRITERIA

Cochrane steps for defining study selection criteria	Selection criteria for abstract reviews and full manuscript reviews
5.1 Questions and eligibility criteria	Original research and aggregate or person-level data from administrative/claims data source or electronic medical records.
5.2 Defining types of participants: which people and populations?	Any appropriate U.S. patient population (e.g., no a priori exclusions by age or gender).
5.3 Defining types of interventions: which comparisons to make?	FDA policies: black box warnings, withdrawals, Risk Evaluation and Mitigation Strategies (REMS), labeling change.
5.4 Defining types of outcomes: which outcome measures are most important?	Only studies with primary endpoints meeting BOTH of the following two criteria should be included: 1) changes in utilization (of the medical product in question, substitution, etc), AND 2) changes in the rate of adverse outcomes associated with the medical product. Exceptions: 1) in the case of product withdrawal from the market, outcomes can include use of substitution products or services (regardless of whether other adverse outcomes are assessed; 2) in the case of labeling changes related to drug-drug interactions, outcomes can include changes in rates of interacting drugs; 3) in the case of labeling changes related to monitoring, outcomes could be changes in the rates of laboratory testing.
5.5 Defining types of study	Research designs include: randomized, quasi-experimental, and stronger observational designs with explicit approaches for controlling for bias. Exclude descriptive audits without statistical tests, cross-sectional or single pre-post comparisons without comparison group.

C. APPENDIX 3: SELECTION CRITERIA REVIEW RESULTS

	N
Reviewed Articles	**243**
Excluded Articles by Criteria	
Review/descriptive/opinion/audits/pre-post	89
FDA regulation not focus of study	80
No outcomes endpoint	29
Not medical product	15
Prospective data collection (surveys)	7
Not US data	5
Final Articles	**18**

D. APPENDIX 4: LITERATURE REVIEW RESULTS

Data			
	years	range	1974-2008
	Unit of analysis		
	Person-level		28%
	Aggregate		72%
	Data types		
	Medicaid/VA/TRICARE data		33%
	Private health insurance/MCO claims		6%
	Proprietary administrative databases		61%
	Registries/outcomes databases		17%
	Electronic Medical Record/chart reviews		0%
	Data evaluated for completeness/missing		89%
Sample			
	Sample size	range	1,308-61 million
	Special Populations		
	Including pediatric		78%
	Including geriatric		50%
FDA Policy (including multi-policy evaluations)			
	Boxed warnings		50%
	Labeling change		33%
	FDA Advisory/safety warnings		39%
	Dear Healthcare Professional Letters		22%
	Withdrawals		17%
Primary Endpoints			
	Use of targeted medical product		89%
	Substitute product(s) or services		61%
	Lab monitoring		17%
	Adverse event(s)		11%
	Contraindicated product(s)/condition(s)		11%
Research Design (including multiple designs)			
	Interrupted Time Series/Time Series, including control		39%
	Interrupted Time Series/Time Series, no control		61%
	Pre-post, including control		17%

Analytic Methods (including multiple analyses)			
	Generalized linear models (including GEE and GLMM) and multiple regressions		17%
	Survival analysis (including segmented survival)		11%
	Segmented regression analysis (including autoregressive models or more complicated models such as ARIMA)		61%
	Pre-post differences (t-tests, chi-squares)		17%
	Piecewise regression		6%
Confounder Analyses			
	Propensity score	n	1
Detailed Lists			
Targeted Medical Products			
	antidepressants	n	8
	troglitazone/rosiglitazone	n	2
	atypical antipsychotics	n	2
	cisapride	n	1
	zomepirac	n	1
	propoxyphene	n	1
	DESI drugs	n	1
	terfenadine	n	1
	pemoline	n	1
Data Used			
	MarketScan databases	n	1
	Medicaid	n	4
	IMS health database	n	2
	United Health Group administrative claims	n	1
	U.S. Veterans Affairs claims/registries	n	1
	PharMetrics databases	n	4
	Unnamed MCO/insurer claims	n	2
	CDC mortality databases	n	1
	Drug Abuse Warning Network	n	1
	TRICARE	n	1
	i3 INNOVUS databases	n	1
	Ingenix database	n	1
	Verispan	n	1

NOTES: Statistics are reported in percentages unless otherwise indicated. Percentages may sum to greater than 100% as categories are not all mutually exclusive.

E. APPENDIX 5: DETAILS OF LITERATURE INCLUDED IN REVIEW (BY YEAR OF PUBLICATION)

Soumerai SB, Avorn J, Gortmaker S, Hawley S. Effect of government and commercial warnings on reducing prescription misuse: the case of propoxyphene. Am J Public Health 1987;77:1518-23.

FDA Regulatory Action: Boxed Warning

Targeted Product: Propoxyphene

Data Source(s): IMS America's National Prescription Audit (NPA); Drug Abuse Warning Network (DAWN); IMS National Disease and Therapeutic Index (NDTI); Finkle Study Data

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 1974-1983

Study Population: All Ages

Sample Size: N/A

Primary Endpoint (Utilization): Propoxyphene Prescriptions

Primary Endpoint (Change in Rate): Overdose Deaths

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Substitute Product (NSAIDs)

Research Design: Interrupted Time Series

Control Group: Yes

Analytic Methods: Segmented Regression Analysis

Confounder Analyses: No

Abstract: We analyzed trends in prescribing and overdose deaths related to propoxyphene (e.g., Darvon) before and after a 1978-80 informational campaign carried out by the US Food and Drug Administration and the drug's manufacturer through mailed warnings, face-to-face education of prescribers, press releases, and labeling changes. The goals included a reduction in propoxyphene use with alcohol or other CNS depressants, reduced prescribing of refills, and cessation of prescribing for patients at risk of abuse and misuse (suicide). We conducted time-series analyses of nationwide propoxyphene use data 1974-83 and analyzed data on drug overdose death rates covering a combined population of about 83 million. Segmented regression methods were used to determine if the informational program was associated with changes in trends of prescribing or overdose deaths. Comparison drug series were analyzed to control for other secular trends in prescribing. Nationwide propoxyphene use during the warnings continued a pre-existing decline of about 8 per cent per year, but this decline halted after the warnings. The no-refill recommendation had no impact on refill rates. The risk of overdose death per propoxyphene prescription filled has remained about constant since 1979. Sharper declines in misuse of such drugs will require stronger, more sustained regulatory or educational measures.

Soumerai SB, Ross-Degnan D, Gortmaker S, Avorn J. Withdrawing payment for nonscientific drug therapy. Intended and unexpected effects of a large-scale natural experiment. JAMA 1990;263:831-9.

FDA Regulatory Action: Withdrawal

Targeted Product: DESI Drugs

Data Source(s): Medicaid

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 1980-1983

Study Population: All Ages

Sample Size: 390,465

Primary Endpoint (Utilization): 12 Categories of DESI Drug Prescriptions

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Substitute Product (66 therapeutic categories)

Research Design: Interrupted Time Series

Control Group: No

Analytic Methods: Segmented Regression Analysis

Confounder Analyses: No

Abstract: Little is known about the effect on clinical decision making of nonreimbursement for ineffective medical technologies. Using a time-series design, we studied the effects of cessation of government payment for 12 categories of drugs of questionable efficacy (Drug Efficacy Study Implementation drugs) in a random sample of the New Jersey Medicaid population (N=390 465) and in four cohorts of regular users of these products. We measured changes in the overall levels of prescriptions, expenditures, and physicians' use of substitute drugs. Although withdrawn drugs accounted for 7% of prescriptions in the base year, there was no measurable reduction in overall drug use or expenditures after the regulation; prescription rates actually rose from 0.86 to 1.00 monthly prescriptions per enrollee throughout the 42-month study. Controlling for preexisting trends, an estimated drop in the use of study drugs of 21.7 prescriptions per 1000 enrollees per month was offset by an increase in the use of substitute drugs of 33.7 prescriptions. Both desirable and unimproved therapeutic substitutions were observed. Used alone, curtailment of reimbursement for marginally effective therapies results in both desirable and unintended clinical substitutions and may not reduce costs. Supplementing such restrictions with education may be necessary to promote practices that are more therapeutically and economically appropriate.

Ross-Degnan D, Soumerai SB, Fortess EE, Gurwitz JH. Examining product risk in context. Market withdrawal of zomepirac as a case study. JAMA 1993;270:1937-42.

FDA Regulatory Action: Withdrawal

Targeted Product: Zomepirac

Data Source(s): Medicaid

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 1980-1983

Study Population: All Ages

Sample Size: 173,726

Primary Endpoint (Utilization): Zomepirac Prescriptions

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Substitute Product (other analgesics)

Research Design: Interrupted Time Series

Control Group: Yes

Analytic Methods: Segmented Regression Analysis

Confounder Analyses: No

Abstract: OBJECTIVE: To examine changes in the prescribing of analgesics after the market entry and subsequent withdrawal of zomepirac sodium, a nonsteroidal anti-inflammatory drug (NSAID), following repeated reports of zomepirac-related deaths. DESIGN: To evaluate this natural quasi experiment, we conducted time-series analyses to compare prescribing in two cohorts of primary care physicians from July 1980 through September 1983. SETTING: Study physicians provided outpatient pharmaceutical care to patients enrolled in the New Jersey Medicaid program. PARTICIPANTS: We identified 260 primary care physicians who provided 10 or more prescriptions for zomepirac (zomepirac prescribers) and 308 who provided 10 or more prescriptions for NSAIDs other than zomepirac (other-NSAID prescribers) in Medicaid during the study period. MAIN OUTCOME MEASURES: Monthly rates of prescribing for zomepirac and several categories of substitute analgesics among Medicaid patients seen by study physicians. MAIN RESULTS: Zomepirac accounted for a stable 11.0% of analgesic prescribing among the zomepirac-prescriber cohort; label changes and manufacturer product-risk warnings 11 months before the product's withdrawal from the market had no impact on use. After market entry, zomepirac prescribers reduced use of other NSAIDs and propoxyphene (hydrochloride or napsylate) in comparison with other-NSAID prescribers (-8.1% and -2.8% of total analgesic prescribing, respectively; $P < .001$). After the product's withdrawal from the market, zomepirac prescribers showed significant increases in relative prescribing of other NSAIDs (+6.8%; $P < .001$), propoxyphene (+2.1%; $P < .05$), and analgesics containing barbiturates (+2.7%; $P < .001$). CONCLUSIONS: The sudden withdrawal of zomepirac from the market resulted in substitutions not only of other NSAIDs, but also of alternative analgesics that carry risks of habituation and adverse effects. Apparent gains in patient safety resulting from market withdrawal of medications must be evaluated in comparison with risks of medications likely to be substituted.

Thompson D, Oster G. Use of terfenadine and contraindicated drugs. JAMA 1996;275:1339-41.

FDA Regulatory Action: Boxed Warning; Labeling Change

Targeted Product: Terfenadine

Data Source(s): Pharmacy Claims (from 1 Insurer)

Data Quality/Completeness or Validity: No

Unit Level of Data: Aggregate

Years of Data: 1990-1994

Study Population: All Ages

Sample Size: N/A

Primary Endpoint (Utilization): Terfenadine Prescriptions

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Contraindicated Drugs

Research Design: Interrupted Time Series

Control Group: No

Analytic Methods: T-test of Post-policy Contraindication Rates >0

Confounder Analyses: No

Abstract: OBJECTIVE: To assess changes in concurrent use of products containing terfenadine and contraindicated macrolide antibiotics (erythromycin, clarithromycin, troleandomycin) and imidazole antifungals (ketoconazole, itraconazole) following reports of serious drug-drug interactions and changes in product labeling. DESIGN: Retrospective review of computerized pharmacy claims. SETTING: A large health insurer in New England. PATIENTS: Health plan members with 1 or more paid pharmacy claims for products containing terfenadine between January 1990 and June 1994. MAIN OUTCOME MEASURES: Among persons with paid claims for terfenadine in any given month, percentage with a prescription for any contraindicated drug that alternatively was dispensed on the same day as ("same-day dispensing") or had therapy days that overlapped those of ("overlapping use") a prescription for terfenadine. RESULTS: Concurrent use of terfenadine and contraindicated drugs declined over the study period. The rate of same-day dispensing declined by 84% from an average of 2.5 per 100 persons receiving terfenadine in 1990 to 0.4 per 100 persons during the first 6 months of 1994, while the rate of overlapping use declined by 57% (from 5.4 to 2.3 per 100 persons). Most cases involved erythromycin. CONCLUSIONS: Despite substantial declines following reports of serious drug-drug interactions and changes in product labeling, concurrent use of terfenadine and contraindicated macrolide antibiotics and imidazole antifungals continues to occur.

Weatherby LB, Nordstrom BL, Fife D, Walker AM. The impact of wording in "Dear doctor" letters and in black box labels. Clin Pharmacol Ther 2002;72:735-42.

FDA Regulatory Action: Boxed Warning; Dear Healthcare Professional Letters

Targeted Product: Cisapride

Data Source(s): Ingenix Database

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 1995-2000

Study Population: All Ages

Sample Size: N/A

Primary Endpoint (Utilization): Cisapride Prescriptions

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Contraindicated Drugs

Research Design: Interrupted Time Series

Control Group: No

Analytic Methods: Logistic Regression

Confounder Analyses: No

Abstract: OBJECTIVES: The Food and Drug Administration and pharmaceutical manufacturers use "Dear doctor" letters to alert physicians about drug safety. To determine how such warnings may be improved, we retrospectively examined how variations in the wording of one series of "Dear doctor" letters affected their impact on concomitant dispensing of cisapride (Propulsid; Janssen Pharmaceutica, Titusville, NJ) and several medications contraindicated for concomitant use. METHODS: Concomitant dispensing was defined as dispensing cisapride and a contraindicated medication on dates when the intended duration of the two dispensings overlapped on at least 1 day. Using outpatient pharmacy claims from a New England health insurer, we calculated a concomitant dispensing rate for each calendar month as the number of concomitant cisapride dispensings divided by the total number of cisapride dispensings. We grouped drugs contraindicated for concomitant use with cisapride as (1) explicitly named in the warnings, (2) only mentioned as examples of a drug class, or (3) only implied as drug class members. We used multivariate analysis to relate temporal changes in concomitant dispensing rates to type of warning (explicit, example, or implied), patient demographic characteristics, season, calendar year, and temporal relationship to the "Dear doctor" warnings. RESULTS: A highly publicized letter sent in June 1998 was associated with a notable decline (58%) in the concomitant dispensing rate with explicitly contraindicated drugs but not in the concomitant dispensing of cisapride with the example or implied drugs. An earlier letter, which had been explicit but was accompanied by less publicity, had no measurable effect on this study's measure of coprescription, nor did a later letter that emphasized comorbidities. CONCLUSIONS: Explicit, well-publicized drug warnings can change prescriber behavior.

Willy ME, Manda B, Shatin D, Drinkard CR, Graham DJ. A study of compliance with FDA recommendations for pemoline (Cylert). J Am Acad Child Adolesc Psychiatry 2002;41:785-90.

FDA Regulatory Action: Boxed Warning; Labeling Change

Targeted Product: Pemoline

Data Source(s): United Health Group Administrative Claims

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Person

Years of Data: 1998-2000/1998-2001

Study Population: All Ages

Sample Size: 1,308

Primary Endpoint (Utilization): Pemoline Users

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Monitoring (Liver Enzyme Testing)

Research Design: Interrupted Time Series; Pre-Post Comparison

Control Group: No

Analytic Methods: Logistic Rregression

Confounder Analyses: No

Abstract: OBJECTIVE: To assess compliance with product labeling recommendations to use pemoline as second-line therapy for attention-deficit/hyperactivity disorder (ADHD) and to obtain baseline and biweekly liver enzyme tests. METHOD: Retrospective cohort study using administrative claims data to identify first-line therapies and liver enzyme tests among pemoline users between January 1, 1998, and March 31, 2000. Prescriptions for first-line therapy were searched for 90 days prior to the first pemoline claim. Liver enzyme testing (baseline and follow-up) was compared between two groups (the prerecommendation cohort October 1,1998, to March 31, 1999, and the postrecommendation cohort October 1,1999, to March 31,2000). RESULTS: 1,308 patients received at least one pemoline prescription during the study period; 76% of patients < or = 20 years were male. ADHD was the claims-identified indication for 688 patients (52%). Despite the labeling recommendation for use as second-line therapy, only 237 ADHD patients (34%) received a first-line therapy prior to pemoline. Only 12% and 11% of the pre- and post-cohort patients, respectively, received baseline liver enzyme tests; 9% in the pre- and 12% in the post-cohort received at least one liver enzyme follow-up test. CONCLUSIONS: Compliance with product labeling was low for both recommendations. Understanding the reasons for this finding could help improve risk management strategies.

Cluxton RJ, Jr., Li Z, Heaton PC, et al. Impact of regulatory labeling for troglitazone and rosiglitazone on hepatic enzyme monitoring compliance: findings from the state of Ohio medicaid program. Pharmacoepidemiol Drug Saf 2005;14:1-9.

FDA Regulatory Action: Dear Healthcare Professional Letters; Labeling Change; Withdrawal

Targeted Product: Troglitazone

Data Source(s): Medicaid

Data Quality/Completeness or Validity: No

Unit Level of Data: Person

Years of Data: 1997-2000

Study Population: Adults

Sample Size: 7,226 Troglitazone Users and 1,480 Rosiglitazone Users

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Monitoring (Liver Enzyme Testing)

Research Design: Interrupted Time Series; Pre-Post Comparison; Independent Post-test Only Sample

Control Group: Post-test-Only Rosiglitazone Users

Analytic Methods: Survival Analysis

Confounder Analyses: Adjustment for Age, Sex, and Race

Abstract: PURPOSE: Troglitazone, the first drug of the thiazolidinediones class for type II diabetes, was first marketed in March 1997 and was removed from the U.S. market 36 months later after 90 cases of liver failure were reported despite multiple warnings containing liver enzyme monitoring recommendations. Rosiglitazone has been available since June 1999 and is still on the market. The purpose of this study was to evaluate the impact of labeled hepatic enzyme monitoring for troglitazone and rosiglitazone. METHODS: Drug cohorts were assembled, using population-based fee-for-service Medicaid claims, for patients between 18 and 65 years of age who had received at least one troglitazone (n = 7226) or rosiglitazone (n = 1480) prescription between 1 April, 1997, and 21 March, 2000. The outcome of interest was the percentage of patients, based on their first treatment episode, who had baseline and post-baseline liver enzyme testing. RESULTS: Overall baseline testing was under 9% before regulatory actions, increased to 14% after the first two 'Dear Doctor' letters issued by the FDA in October and December 1997, and peaked to about 26% afterwards. Coincident with the marketing of rosiglitazone and the fourth 'Dear Doctor' letter issued in June 1999, baseline testing dropped to 18%. Baseline testing increased 2.5-fold (race-sex-age adjusted) after regulatory action. Achieving 50% post-baseline testing took approximately 6 months for both drugs. CONCLUSION: Regulatory actions had only modest effects on the incidence of liver monitoring. More effective and timely communication strategies, health provider prescribing interventions and modification of health provider behaviors to enhance compliance with recommended risk management measures need to be identified, evaluated and implemented.

Gibbons RD, Brown CH, Hur K, et al. Early evidence on the effects of regulators' suicidality warnings on SSRI prescriptions and suicide in children and adolescents. Am J Psychiatry 2007;164:1356-63.

FDA Regulatory Action: Boxed Warning; FDA Advisory

Targeted Product: Antidepressants

Data Source(s): IMS Health Database; CDC Mortality Reports and Mortality Database; Netherlands Prescription and Mortality Data

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 2003-2005

Study Population: Pediatric

Sample Size: 61,000,000

Primary Endpoint (Utilization): Antidepressant Prescriptions

Primary Endpoint (Change in Rate): Suicide Rates

Research Design: Interrupted Time Series

Control Group: No

Analytic Methods: Poisson Regressions

Confounder Analyses: No

Abstract: OBJECTIVE: In 2003 and 2004, U.S. and European regulators issued public health warnings about a possible association between antidepressants and suicidal thinking and behavior. The authors assessed whether these warnings discouraged use of antidepressants in children and adolescents and whether they led to increases in suicide rates as a result of untreated depression. METHOD: The authors examined U.S. and Dutch data on prescription rates for selective serotonin reuptake inhibitors (SSRIs) from 2003 to 2005 in children and adolescents (patients up to age 19), as well as suicide rates for children and adolescents, using available data (through 2004 in the United States and through 2005 in the Netherlands). They used Poisson regression analyses to determine the overall association between antidepressant prescription rates and suicide rates, adjusted for sex and age, during the periods preceding and immediately following the public health warnings. RESULTS: SSRI prescriptions for youths decreased by approximately 22% in both the United States and the Netherlands after the warnings were issued. In the Netherlands, the youth suicide rate increased by 49% between 2003 and 2005 and shows a significant inverse association with SSRI prescriptions. In the United States, youth suicide rates increased by 14% between 2003 and 2004, which is the largest year-to-year change in suicide rates in this population since the Centers for Disease Control and Prevention began systematically collecting suicide data in 1979. CONCLUSIONS: In both the United States and the Netherlands, SSRI prescriptions for children and adolescents decreased after U.S. and European regulatory agencies issued warnings about a possible suicide risk with antidepressant use in pediatric patients, and these decreases were associated with increases in suicide rates in children and adolescents.

Libby AM, Brent DA, Morrato EH, Orton HD, Allen R, Valuck RJ. Decline in treatment of Pediatric depression after FDA advisory on risk of suicidality with SSRIs. Am J Psychiatry 2007;164:884-91.

FDA Regulatory Action: FDA Advisory

Targeted Product: Antidepressants (SSRIs)

Data Source(s): PharMetrics Database

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 1998-2005

Study Population: Pediatric

Sample Size: 65,349 (individuals)

Primary Endpoint (Utilization): Antidepressant Users

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Substitute Product (Psychotherapy; Antipsychotics; Anxiolytics)

Research Design: Interrupted Time Series

Control Group: No

Analytic Methods: Segmented Regressions; t-test of Predicted vs. Observed

Confounder Analyses: No

Abstract: OBJECTIVE: In October 2003, the U.S. Food and Drug Administration (FDA) issued a public health advisory about the risk of suicidality in pediatric patients taking selective serotonin reuptake inhibitors (SSRIs) for depression. This study used data from a large national pediatric cohort to examine patterns of diagnosis of depression, prescription of antidepressants, prescription of pharmacological alternatives to antidepressants, and use of psychosocial care before and after the FDA advisory was issued. METHOD: A large pediatric cohort with newly diagnosed episodes of depression was created from a national integrated claims database of managed care plans from October 1998 to September 2005 (N=65,349). Time-series models were used to compare diagnosing and prescribing trends during the 2 years after the FDA advisory and the expected trends based on data from the 5-year period preceding the advisory. RESULTS: From 1999 to 2004, pediatric diagnoses of depression increased from 3 to 5 per 1,000. After the FDA advisory was issued, the national rate decreased to 1999 levels, a significant deviation from the historical trend. Pediatricians and nonpediatrician primary care physicians accounted for the largest reductions in new diagnoses. Among patients with depression, the proportion receiving no antidepressant increased to three times the rate predicted by the preadvisory trend, and SSRI prescription fills were 58% lower than predicted by the trend. There was no evidence of a significant increase in use of treatment alternatives (psychotherapy, atypical antipsychotics, and anxiolytics). CONCLUSIONS: The FDA advisory was associated with significant reductions in aggregate rates of diagnosis and treatment of pediatric depression.

Nemeroff CB, Kalali A, Keller MB, et al. Impact of publicity concerning Pediatric suicidality data on physician practice patterns in the United States. Arch Gen Psychiatry 2007;64:466-72.

FDA Regulatory Action: FDA Advisory

Targeted Product: Antidepressants

Data Source(s): Verispan (retail pharmacy and physician audit data)

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 2000-2005

Study Population: Pediatric; Young Adults

Sample Size: N/A

Primary Endpoint (Utilization): Antidepressant Prescriptions

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Substitute Product (Antipsychotics; Benzodiazepines/Anxiolytics); Substitute Service (Psychiatric Specialty Care)

Research Design: Interrupted Time Series

Control Group: Yes

Analytic Methods: Piecewise Regression

Confounder Analyses: No

Abstract: CONTEXT: IMS Health Inc data presented by the Food and Drug Administration (FDA) on September 13 and 14, 2004, at a joint meeting of the Center for Drug Evaluation and Research's Psychopharmacologic Drugs Advisory Committee and the FDA's Pediatric Advisory Committee suggested that the number of children and teenagers who were prescribed antidepressants continued to increase in 2004, despite widespread publicity surrounding 2 FDA advisories regarding the potential for pediatric suicidality with selective serotonin reuptake inhibitor use. These results are contradictory to findings from the Medco Health Solutions, Inc, March 2004 analysis of pharmacy benefit claims and a separate subsequent analysis conducted by NDC Health using dispensing data from March 31, 2004, through June 30, 2005. OBJECTIVES: To investigate the contradictory findings and provide additional analyses on the prescribing trends of antidepressants across age groups and physician specialties in the United States. DESIGN: Retail pharmacy prescription data and physician audit data were obtained from Verispan, a joint venture between Quintiles Transnational and McKesson. In addition to examining prescribing trends, a joinpoint regression analysis was conducted to identify the timing for significant changes in prescription use. RESULTS: The analyses suggest that the number of children and teenagers who were prescribed antidepressants has decreased significantly (P = .02) in the wake of widespread publicity surrounding the FDA public health advisories. Another impact of the advisories seems to be a shift in care from "generalists" to psychiatric specialists when it comes to prescribing antidepressants to patients younger than 18 years. Finally, the analyses highlight a slight shift in prescribing toward the non-selective serotonin reuptake inhibitor bupropion hydrochloride, even though it carries the same FDA "black box" warning as the selective serotonin reuptake inhibitors. CONCLUSIONS: The effect on antidepressant prescribing volume observed in our analysis of the Verispan data parallels earlier findings reported by Medco Health Solutions, Inc, and NDC Health that the FDA actions have had a significant effect on the prescribing of antidepressants to children and adolescents. Together, these findings underline the importance of presenting a fair balance within the media due to the significant reach of this channel among prescribing physicians.

Valuck RJ, Libby AM, Orton HD, Morrato EH, Allen R, Baldessarini RJ. Spillover effects on treatment of adult depression in primary care after FDA advisory on risk of Pediatric suicidality with SSRIs. Am J Psychiatry 2007;164:1198-205.

FDA Regulatory Action: Boxed Warning; FDA Advisory

Targeted Product: Antidepressants

Data Source(s): PharMetrics Databases

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 1998-2005

Study Population: Adults

Sample Size: 40,011

Primary Endpoint (Utilization): Antidepressant Prescriptions

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Psychotherapy Visits; Depression Diagnoses

Research Design: Interrupted Time Series

Control Group: No

Analytic Methods: Segmented Time Series

Confounder Analyses: No

Abstract: OBJECTIVE: In 2003, the U.S. Food and Drug Administration (FDA) issued a public health advisory about the risk of suicidality in pediatric patients taking selective serotonin reuptake inhibitors (SSRIs) for depression, and in 2005, the agency mandated a black box warning and medication guide indicating that pediatric and adult patients may be at risk. The authors examine the effects of this pediatric policy on treatment of adult depression in the community. METHOD: An adult cohort with newly diagnosed episodes of depression was created from a large national integrated claims database of managed care plans from October 1998 to September 2005 (N=475,838 unique episodes). Time-series analyses were used to compare the post-FDA advisory trends to the trends during the 5 years preceding the advisory. RESULTS: The rate of diagnosed depression was significantly lower after the advisory than would have been expected on the basis of the preadvisory historical trend. The average percentage of adults with new (versus recurrent) depressive episodes was 88.6% in the preadvisory period (declining at an annual rate of 1.69%), and it decreased significantly to 77.5% (declining more rapidly, at an annual rate of 7.70%). The percentage of adults with depression who did not receive an antidepressant increased from an average of 20% (declining at 0.45% annually) before the policy action to an average of 30% (increasing at an annual rate of 20.6%). The data did not show any compensatory increases in psychotherapy or prescription of atypical antipsychotics or anxiolytics. CONCLUSIONS: The FDA advisory had a significant spillover effect into community treatment for adults with depression, despite the focus of the policy on pediatric patients.

Morrato EH, Libby AM, Orton HD, et al. Frequency of provider contact after FDA advisory on risk of pediatric suicidality with SSRIs. Am J Psychiatry 2008;165:42-50.

FDA Regulatory Action: FDA Advisory

Targeted Product: Antidepressants (SSRI)

Data Source(s): PharMetrics Databases

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 1998-2005

Study Population: Pediatric

Sample Size: 27,370 (children); 193,151 (adults)

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Monitoring (HEDIS Quality Indicators of Health Care Contacts)

Research Design: Interrupted Time Series

Control Group: Adults

Analytic Methods: Segmented Regression Analysis

Confounder Analyses: No

Abstract: OBJECTIVE: The Food and Drug Administration (FDA) issued a public health advisory in October 2003 on the risk of suicide in pediatric patients taking antidepressants and advised maintaining "close supervision" of such patients. In this study, the authors compared trends in the frequency of provider contacts for patients with depression before and after the advisory was issued. METHOD: Retrospective cohorts of children (N=27,370) and adults (N=193,151) with new episodes of depression treated with antidepressants were created from a national claims database of managed care plans (1998-2005). Two standards were used in measuring patient monitoring: the Health Plan Employer Data and Information Set (HEDIS) quality-of-care criterion calling for three contacts in 3 months and the FDA-recommended contact schedule totaling seven visits in 3 months. Time-series models compared postadvisory trends to the expected trend based on preadvisory measures. RESULTS: Less than 5% of all patients met FDA contact recommendations before the advisory, and the rate did not change after the advisory. A greater proportion of patients met the HEDIS contact criterion before the advisory (60% for children and 40% for adults), and the rate did not change after the advisory. A greater proportion of pediatric patients seen by a psychiatrist (80%) met the HEDIS criterion than those seen by a pediatrician (60%) or a non-pediatrician primary care physician (54%), and than adults seen by a psychiatrist (65%) or a primary care physician (37%). The proportions of pediatric patients who met the FDA recommendations did not differ by specialty. CONCLUSIONS: Contrary to expectations, the frequency of visits by patients with new episodes of depression treated with antidepressants did not increase after the October 2003 FDA advisory was issued.

Libby AM, Orton HD, Valuck RJ. Persisting decline in depression treatment after FDA warnings. Arch Gen Psychiatry 2009;66:633-9.

FDA Regulatory Action: Boxed Warning; FDA Advisory

Targeted Product: Antidepressants

Data Source(s): PharMetrics Database

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 1999-2007

Study Population: 5-18; 19-24; 25-89

Sample Size: 643,313 (individuals)

Primary Endpoint (Utilization): Antidepressant Users

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Substitute Product (Psychotherapy; Antipsychotics; Anxiolytics)

Research Design: Interrupted Time Series

Control Group: Yes

Analytic Methods: Segmented Segressions, t-test of Predicted vs. Observed

Confounder Analyses: No

Abstract: CONTEXT: In October 2003 the Food and Drug Administration (FDA) issued a Public Health Advisory about the risk of suicidality for pediatric patients taking antidepressants; a boxed warning, package insert, and medication guide were implemented in February 2005. The warning was extended to young adults aged 18 to 24 years in May 2007. Immediately following the 2003 advisory, unintended declines in case finding and non-selective serotonin reuptake inhibitor substitute treatment were shown for pediatric patients, and spillover effects were seen in adult patients, who were not targeted by the warnings. OBJECTIVE: To determine whether the unintended declines in depression care persisted for pediatric, young adult, and adult patients. DESIGN: Time series analyses. SETTING: Ambulatory care settings nationally. Patients Pediatric, young adult, and adult cohorts of patients with new episodes of depression (n = 91 748, 70 311, and 630 748 episodes, respectively). INTERVENTIONS: Post-FDA advisory trends were compared with expected trends based on preadvisory patterns using a national integrated managed care claims database from July 1999 through June 2007. MAIN OUTCOME MEASURES: Depression diagnosis; antidepressant, antipsychotic, and anxiolytic prescriptions; and psychotherapy visits. RESULTS: Changes in pediatric depression care were similar to changes for adults. National diagnosis rates of depression returned to 1999 levels for pediatric patients and below 2004 levels for adults. Primary care providers continued significant reductions in new diagnoses of depression (44% lower for pediatric, 37% lower for young adults, 29% for adults); diagnoses by mental health providers who were not psychiatrists increased. Numbers of prescriptions of anxiolytic and atypical antipsychotic medications did not significantly change from preadvisory trends. Psychotherapy increased significantly for adult, though not pediatric, cases. Selective serotonin reuptake inhibitor use decreased in all cohorts; serotonin-norepinephrine reuptake inhibitor increased for adults. CONCLUSIONS: Diagnosing decreases persist. Substitute care did not compensate in pediatric and young adult groups, and spillover to adults continued, suggesting that unintended effects are nontransitory, substantial, and diffuse in a large national population. Policy actions are required to counter the unintended consequences of reduced depression treatment.

Stewart KA, Natzke BM, Williams T, Granger E, Casscells SW, Croghan TW. Temporal trends in anti-diabetes drug use in TRICARE following safety warnings in 2007 about rosiglitazone. Pharmacoepidemiol Drug Saf 2009;18:1048-52.

FDA Regulatory Action: FDA Advisory

Targeted Product: Rosiglitazone

Data Source(s): TRICARE

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 2006-2008

Study Population: Adults

Sample Size: 214,480

Primary Endpoint (Utilization): Rosiglitazone Prescriptions

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Alternate Diabetes Drugs

Research Design: Interrupted Time Series

Control Group: No

Analytic Methods: Segmented Regression Analysis

Confounder Analyses: No

Abstract: PURPOSE: To describe utilization patterns for anti-diabetes medications among a cohort of diabetes patients in the Military Health System (MHS) before and after warnings about rosiglitazone issued in May 2007. METHODS: We used segmented regression analysis to compare changes in the level and trend of rosiglitazone utilization and use of other anti-diabetes therapies in the period prior to the drug warnings (between April 2006 and May 2007) and the period after the warnings were issued (between October 2007 and May 2008). RESULTS: The level and trend of rosiglitazone use changed after the highly publicized warnings. The number of prescriptions filled fell by almost 7000 after the warning ($p < 0.001$). The number of prescriptions filled for pioglitazone, sulfonylureas, and other diabetes drugs increased significantly after the warnings ($p < 0.05$ in all models). Overall, the level and trend of filled prescriptions per month for all anti-diabetic drugs did not significantly change after the warnings. CONCLUSIONS: Utilization patterns changed in response to warnings about rosiglitazone. While overall utilization of anti-diabetic drugs did not change, further study is needed to determine the associated health outcomes.

Busch SH, Frank RG, Leslie DL, et al. Antidepressants and suicide risk: how did specific information in FDA safety warnings affect treatment patterns? Psychiatric Serv 2010;61:11-6.

FDA Regulatory Action: Boxed Warning; Labeling Change

Targeted Product: Antidepressants

Data Source(s): MarketScan Databases

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Person

Years of Data: 2001-2005

Study Population: Pediatric

Sample Size: 22,689 (new episodes of treatment)

Primary Endpoint (Utilization): Antidepressant Prescriptions

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Monitoring (Outpatient Visits)

Research Design: Interrupted Time Series

Control Group: No

Analytic Methods: Segmented Logistic Regression; Cox Proportional Hazards Models: follow-up testing

Confounder Analyses: Logistic Analyses with time period dummies; Controlled for Age, Gender, Geographic Region (Season for Some Analyses)

Abstract: OBJECTIVE: From June 2003 through October 2004, the U.S. Food and Drug Administration (FDA) released five safety warnings related to antidepressant use and the increased risk of suicidality for children. Although researchers have documented a decline in antidepressant use among children over this period, less is known about whether specific safety information conveyed in individual warnings was reflected in treatment patterns. METHODS: Thomson Reuters MarketScan claims data (2001-2005) for a national sample of privately insured children were used to construct treatment episodes (N=22,689). For each new episode of major depressive disorder, it was determined whether treatment followed specific recommendations included in warnings released by the FDA. Treatment recommendations pertained to the use of the antidepressants paroxetine and fluoxetine and to patient monitoring. Treatment patterns were expected to change as the risk information conveyed by the FDA changed over time. RESULTS: The timing of FDA recommendations was associated with trends in the use of paroxetine and fluoxetine by children with major depressive disorder who were initiating antidepressant treatment. However, no evidence of increases in outpatient visits (indicative of monitoring) among depressed children initiating antidepressant use was found. CONCLUSIONS: Release of specific risk and benefit information by the FDA was associated with changes in prescribing but not in outpatient follow-up. These results suggest that the FDA plays an important role in communicating information to the public and providers. Yet, although public health safety warnings were associated with changes in some practice patterns, not all recommendations conveyed in warnings were followed.

Morrato EH, Druss B, Hartung DM, et al. Metabolic testing rates in 3 state Medicaid programs after FDA warnings and ADA/APA recommendations for second-generation antipsychotic drugs. Arch Gen Psychiatry 2010;67:17-24.

FDA Regulatory Action: Dear Healthcare Professional Letter; Labeling Change

Targeted Product: Atypical Antipsychotics

Data Source(s): Medicaid

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Aggregate

Years of Data: 2002-2005

Study Population: All Ages

Sample Size: 109,451 (antipsychotics); 203,527 (non-antipsychotics)

Primary Endpoint (Utilization): Atypical Antipsychotic Prescriptions

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Monitoring (Glucose and Lipid Testing)

Research Design: Interrupted Time Series

Control Group: Yes

Analytic Methods: Segmented Regression Analysis

Confounder Analyses: Propensity Matched Control

Abstract: CONTEXT: In 2003, the Food and Drug Administration (FDA) required a warning on diabetes risk for second-generation antipsychotic (SGA) drugs. The American Diabetes Association (ADA) and American Psychiatric Association (APA) recommended glucose and lipid testing for all patients starting to receive SGA drugs. OBJECTIVE: To characterize associations between the combined warnings and recommendations and baseline metabolic testing and SGA drug selection. DESIGN: Interrupted time-series analysis. SETTING: California, Missouri, and Oregon. Patients: A total of 109 451 individuals receiving Medicaid who began taking SGA medication and a control cohort of 203 527 patients who began taking albuterol but did not receive antipsychotic medication. INTERVENTIONS: Prewarning and postwarning trends in metabolic testing were compared using laboratory claims for the cohort collected January 1, 2002, through December 31, 2005. Changes in SGA prescribing practices were similarly evaluated. MAIN OUTCOME MEASURES: Monthly rates of baseline serum glucose and lipid testing for SGA-treated and propensity-matched albuterol-treated patients and monthly share of new prescriptions for each SGA drug. RESULTS: Initial testing rates for SGA-treated patients were low (glucose, 27%; lipids, 10%). The warning was not associated with an increase in glucose testing among SGA-treated patients and was associated with only a marginal increase in lipid testing rates (1.7%; P = .02). Testing rates and trends in SGA-treated patients were not different from background rates observed in the albuterol control group. New prescriptions of olanzapine (higher metabolic risk) declined during the warning period (annual share decline, 19.9%; P < .001). New prescriptions of aripiprazole (lower metabolic risk) increased during the warning period (share increase, 12.1%; P < .001) but may be attributable to the elimination of prior authorization in California during the same time frame. Quetiapine, risperidone, and ziprasidone use were not associated with the warning. CONCLUSIONS: In a Medicaid-receiving population, baseline glucose and lipid testing for SGA-treated patients was infrequent and showed little change following the diabetes warning and monitoring recommendations. A change in SGA drug selection consistent with intentions to reduce metabolic risk was observed.

Valluri S, Zito JM, Safer DJ, Zuckerman IH, Mullins CD, Korelitz JJ. Impact of the 2004 Food and Drug Administration Pediatric suicidality warning on antidepressant and psychotherapy treatment for new-onset depression. Med Care 2010;48:947-54.

FDA Regulatory Action: Boxed Warning

Targeted Product: Antidepressants

Data Source(s): i3 INNOVUS Administrative Claims Databases

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Person

Years of Data: 2003-2006

Study Population: Pediatric

Sample Size: 40,309

Primary Endpoint (Utilization): Antidepressant Users

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Substitute Product (Psychotherapy)

Research Design: Interrupted Time Series; Pre-Post Comparison

Control Group: No

Analytic Methods: GEE Logistic Model

Confounder Analyses: No

Abstract: OBJECTIVE: To assess the national impact of the March 2004 Food and Drug Administration (FDA) antidepressant suicidality warning on the outpatient treatment of new-onset depression in youth. METHOD: A repeated measures, longitudinal design in a cohort of youth diagnosed with new-onset depression was used to assess pre- and post-FDA warning effects. US commercial insurance enrollees in the i3 INNOVUS database from January 2003 through December 2006 were examined. The study population included youth 2- to 17-years old with a new-onset depression diagnosis from July 2003 through June 2006 (N = 40,309). The main independent variables were the warning period (post- vs. pre-FDA warning) and age group (children vs. adolescents). The main outcome measures were youth with antidepressant dispensings and psychotherapy visits measured in 30-day intervals across 36 months following a new-onset diagnosis of any depressive disorder (N = 40,309) and specifically major depressive disorder (MDD) (N = 11,532). RESULTS: Compared to youth with a new-onset diagnosis of depression in the pre-FDA warning period, youth with new-onset diagnosis of depression during the postwarning period had (1) A significantly lower likelihood of antidepressant use: (odds ratio [OR] = 0.85 [0.81-0.89]); When youth with the diagnosis of depression were separated into those with MDD and those with less severe depression diagnoses, only the latter had a significant postwarning antidepressant decline. (2) A significant increase in the odds of a psychotherapy visit (children, OR = 1.31 [1.23-1.40]; adolescents OR = 1.19 [1.15-1.24]). CONCLUSIONS: The FDA suicidality warning was associated with an overall decrease in antidepressant treatment for youth with a clinician-reported diagnosis of depression, but not for those with MDD. Also, following the warning, psychotherapy without medication increased.

Kales HC, Zivin K, Kim HM, et al. Trends in antipsychotic use in dementia 1999-2007. Arch Gen Psychiatry 2011;68:190-7.

FDA Regulatory Action: Dear Healthcare Professional Letters, Labeling Change

Targeted Product: Atypical Antipsychotics

Data Source(s): Veterans Affairs Registries

Data Quality/Completeness or Validity: Yes

Unit Level of Data: Person

Years of Data: 1999-2007

Study Population: Geriatric

Sample Size: 254,564

Primary Endpoint (Exceptions: substitute products; interacting products; monitoring): Substitute Product (Conventional Antipsychotics; Psychotropics)

Research Design: Interrupted Time Series

Control Group: Users of Non-antipsychotic Psychotropic Medications

Analytic Methods: Segmented Regression Analysis

Confounder Analyses: Sensitivity Analyses to Assess Prescribing by Age, Type of Dementia, Comorbidity

Primary Endpoint (Utilization): Atypical Antipsychotic Users

Abstract: CONTEXT: Use of atypical antipsychotics for neuropsychiatric symptoms of dementia increased markedly in the 1990s. Concerns about their use began to emerge in 2002, and in 2005, the US Food and Drug Administration warned that use of atypical antipsychotics in dementia was associated with increased mortality. OBJECTIVE: To examine changes in atypical and conventional antipsychotic use in outpatients with dementia from 1999 through 2007. DESIGN: Time-series analyses estimated the effect of the various warnings on atypical and conventional antipsychotic usage using national Veterans Affairs data across 3 periods: no warning (1999-2003), early warning (2003-2005), and black box warning (2005-2007). SUBJECTS: Patients aged 65 years or older with dementia (n = 254 564). MAIN OUTCOME MEASURES: Outpatient antipsychotic use (percentage of patients, percentage of quarterly change, and difference between consecutive study periods). RESULTS: In 1999, 17.7% (95% confidence interval [CI], 17.2-18.1) of patients with dementia were using atypical or conventional antipsychotics. Overall use began to decline during the no-warning period (rate per quarter, -0.12%; 95% CI, -0.16 to -0.07; P < .001). Following the black box warning, the decline continued (rate, -0.26%; 95% CI, -0.34 to -0.18; P < .001), with a significant difference between the early and black box warning periods (P = .006). Use of atypical antipsychotics as a group increased during the no-warning period (rate, 0.23; 95% CI, 0.17-0.30; P < .001), started to decline during the early-warning period (rate, -0.012; 95% CI, -0.14 to 0.11; P = .85), and more sharply declined during the black box warning period (rate, -0.27; 95% CI, -0.36 to -0.18; P < .001). Olanzapine and risperidone showed declining rates and quetiapine showed an increase during the early-warning period, but rates of use for all 3 antipsychotics declined during the black box warning period. In the black box warning period, there was a small but significant increase in anticonvulsant prescriptions (rate, 0.117; 95% CI, 0.08-0.16; P < .001). CONCLUSIONS: Use of atypical antipsychotics began to decline significantly in 2003, and the Food and Drug Administration advisory was temporally associated with a significant acceleration in the decline.

F. APPENDIX 6: NOVEL RESEARCH DESIGNS AND ANALYTIC METHODS RECOMMENDED FOR EVALUATING FDA REGULATORY ACTIONS

1. Methods with Strong Internal Validity

Research Design	Analytic Method	Application Example	Necessary Data Elements	Strengths	Weaknesses
Interrupted Time Series Designs with control group[1]	Segmented regression analysis. Can include time series estimators (PROC AUTOREG, ARIMA in SAS) using aggregate data or GEE/GLMM using individual data. Specification of model is: 1 parameter for level change and 1 parameter for change in slope. Requires correction for autocorrelation.	For aggregate data: 1) Serumaga B, et al. Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study. BMJ 2011;342:d108[29] 2) Soumerai SB, et al. Payment restrictions for prescription drugs under Medicaid. Effects on therapy, cost, and equity. N Engl J Med 1987;317:550-6.[23] For individual data: Kozhimannil KB, et al. New Jersey's efforts to improve postpartum depression care did not change treatment patterns for women on Medicaid. Health Aff 2011;30:239-301.[30]	Can use person-level or aggregate-level longitudinal data. Data must provide multiple pre- and post-intervention time-points. Short times series (more than 8 pre- and post-time point but <100) are useful but the design becomes stronger with additional time points to establish trends.	Appropriate for policy change that occurs at particular time point(s). One of the most robust of all quasi-experimental designs. Especially strong for testing abrupt changes in levels and slopes.	Requires knowledge of diffusion process, especially in case of gradual interventions or delayed causation. Visual inspection analyses without statistical testing are useful but insufficient for establishing effect.

Research Design	Analytic Method	Application Example	Necessary Data Elements	Strengths	Weaknesses
Regression Discontinuity Designs[1]	Various.	Grootendorst PV, et al. On becoming 65 in Ontario. Effects of drug plan eligibility on use of prescription medicines. Med Care 1997;35:386-98.[31]	Person-level data. Requires an assignment variable (preferably a continuous measure) observed on entire sample.	Appropriate for policy that applies to a defined patient population with fixed eligibility criterion based on a threshold across continuous measure (e.g., age >=65, or blood pressure values >= x). A reasonably robust quasi-experimental design. Useful when pre-exposure data are lacking or limited.	Requires knowledge of function form of relationship between outcome and assignment variable. Has less power for detecting small effects. Not well-known in medical literature.

Research Design	Analytic Method	Application Example	Necessary Data Elements	Strengths	Weaknesses
Various	Extended Cox models with heaviside function[24] (segmented survival analyses).	Briesacher BA, et al. Medicare part D's exclusion of benzodiazepines and fracture risk in nursing homes. Arch Intern Med 2010;170:693-8.[32] Zhang Y, et al. Effects of prior authorization on medication discontinuation among Medicaid beneficiaries with bipolar disorder. Psychiatr Serv 2009;60:520-7.[33]	Person-level longitudinal data.	Appropriate for policy with binary outcome that is expected to change over time. Offers advantages of maximum likelihood (e.g., good large-sample properties). Does not require proportional hazard assumption. Allows for multiple estimates of exposure and outcome relationship as a function of time intervals. Robust to right censored data.	Requires accurate specification of time intervals.

2. Methods Suitable for Special Circumstances

Research Design	Analytic Method	Application Example	Necessary Data Elements	Strengths	Weaknesses
Sequential Analysis of gradually accruing data[2]	Poisson or Bernoulli maximized sequential probability ratio test, Flexible Exact Sequential Analysis, group sequential methods.	Yih et al. Active surveillance for adverse events: the experience of the Vaccine Safety Datalink project. Pediatrics 2011;127 Suppl 1:S54-64.[34]	Aggregate-level data. Requires rich historical data and prospective data collection.	Appropriate for a policy with a priori expectation of outcome effect and null hypothesis of no effect. Best suited for early detection of suspected problem as data are accumulating.	Data must be continuously collected at regular intervals over time, rather than at a single point in time.
Self-controlled Case Series		Whitaker H. The self controlled case series method. BMJ 2008;337:a1069.[26]	Person-level longitudinal data.	Best suited for studying association between transient exposure (e.g., temporary policy) and acute event (i.e., risk periods are short).	Highly sensitive to time interval between exposure and the event. Less appropriate for long-term monitoring or variable risk period.
Difference-in-difference-in-differences	Various.	Afendulis CC, et al. The impact of Medicare Part D on hospitalization rates. Health Serv Res 2011;46:1022-38.[35]	Person-level or aggregate-level data.	Improvement over difference-in difference models. Best suited for studies with limited measurement of pre-post observation.	No controls for historical trends.

3. Methods with Potential if Assumptions Can Be Met

Research Design	Analytic Method	Application Example	Necessary Data Elements	Strengths	Weaknesses
Various	Summary Confounding Scores (propensity scores, inverse probability of treatment weighting, disease risk scores).	Morrato EH, et al. Metabolic screening after the American Diabetes Association's consensus statement on antipsychotic drugs and diabetes. Diabetes Care 2009;32:1037-42.[36]	Person-level.	Best suited as sensitivity analyses to augment more rigorous study design.	Does not control for unmeasured confounding.

VIII. DESCRIPTION OF CORE WORKGROUP LEADER AND MEMBERS

Becky A. Briesacher, PhD; Team Leader and Principal Investigator. Dr. Briesacher is a senior investigator at the Meyers Primary Care Institute and a health services researcher with extensive experience in drug policy research and evaluation methodology. She brings over 10 years of experience in assessing drug policy and regulations and using quasi-experimental study designs, most recently in evaluations of Medicare Part D policies.[32,37-39]

Jerry H. Gurwitz, MD; Co-Investigator. Dr. Gurwitz is Executive Director of the Meyers Primary Care Institute, a joint endeavor of Fallon Community Health Plan, the University of Massachusetts Medical School, and Reliant Medical Group. He also serves as Co-Director of the HMORN Cardiovascular Research Network which is funded by the NHLBI. He brings extensive experience in assessing the usefulness of health plan administrative data for identifying patients with various medical conditions.

Susan Andrade, ScD; Co-Investigator. Dr. Andrade is an investigator at the Meyers Primary Care Institute and a lead investigator for the HMO Research Network Food and Drug Administration (FDA) Epidemiology contract site. She is a pharmacoepidemiologist with extensive expertise in the use of automated databases for pharmacoepidemiologic and health services research. She is also an expert in case identification, validation, and adjudication of various health outcomes related to drug safety.[40]

Stephen B. Soumerai, ScD; Co-Investigator. Dr. Soumerai has published over 200 scientific articles in leading journals, and is well known nationally and internationally for his work on the impacts of drug coverage and cost-containment policies, cost-related nonadherence, clinical outcomes among vulnerable populations.[23,41-44] He has conducted multiple evaluations of drug regulatory actions including a current NIMH study of SSRI warnings and suicide among youth.[40,45]

Fang Zhang, PhD; Co-Investigator. Dr. Zhang is a statistician specializing in time series analyses, survival analysis, categorical data analysis, and Bayesian hierarchical models. He has made unique methodological contributions in estimating confidence intervals and developing power calculations for interrupted time series analyses, using bootstrapping. He has extensive experience in longitudinal research using large Medicaid or commercial insurance claims datasets for studies of drug and other health outcomes.[46-48]

Darren Toh, ScD; Co-Investigator. Dr. Toh is a pharmacoepidemiologist. He offers experience in methods for estimating time-varying treatment effects in longitudinal studies with complex exposure patterns. He is the Deputy Director of Scientific Operations in Mini-Sentinel, and will help align this project with Mini-Sentinel's needs and visions.

IX. ACKNOWLEDGEMENTS

Many thanks are due to those who participated in this project and the subsequent development of this report, including:

The U.S. Food and Drug Administration (FDA) and the Mini-Sentinel Initiative

Project Core Workgroup Lead and Members: Becky A. Briesacher (Lead), Jerry H. Gurwitz, Susan Andrade, Stephen B. Soumerai, Fang Zhang, Darren Toh, Yong Chen, Joann L. Wagner

Mini-Sentinel Operations Center Team: Darren Toh, Aarthi Iyer

FDA Workgroup Lead and Members: Azadeh Shoaibi (Lead), Judith Racoosin, Andrea Sutherland, Antonio Paredes, Chris Jones, Danica Marinac-Dabic, David J. Graham, George Neyarapally, George Rochester, Helen Sullivan, Jan Johannessen, Mark Levenson, Marsha Reichman, Melissa Robb, Michael Kieffer, Michie Hunt, Nilsa Loyo-Berrios, Thomas Buttolph

X. REFERENCES

1. Shadish WR, Cook TD, Campbell DT. Experimental and Quasi-experimental Designs for Generalized Causal Inference. Boston, MA: Houghton Mifflin Company; 2002.
2. Li L, Kulldorff M. A conditional maximized sequential probability ratio test for pharmacovigilance. Statistics in medicine 2010;29:284-95.
3. Kulldorff M, Davis R, Kolczak M, Lewis E, Lieu T, Platt R. A maximized sequential probability ratio test for drug and vaccine safety surveillance. Sequential Analysis: Design Methods and Applications 2011;30:58-78.
4. Cochrane Handbook for Systematic Reviews of Interventions. England: The Cochrane Collaboration and John Wiley & Sons Ltd; 2008.
5. NC D. The Delphi method: An experimental study of group opinion. . In: DL R, R L, D S, eds. Studies in the quality of life: Delphi and decision-making. Lexington, MA: Lexington Books; 1972:13-54.
6. Soumerai SB, Ross-Degnan D, Gortmaker S, Avorn J. Withdrawing payment for nonscientific drug therapy. Intended and unexpected effects of a large-scale natural experiment. JAMA 1990;263:831-9.
7. Ross-Degnan D, Soumerai SB, Fortess EE, Gurwitz JH. Examining product risk in context. Market withdrawal of zomepirac as a case study. JAMA 1993;270:1937-42.
8. Thompson D, Oster G. Use of terfenadine and contraindicated drugs. JAMA 1996;275:1339-41.
9. Weatherby LB, Nordstrom BL, Fife D, Walker AM. The impact of wording in "Dear doctor" letters and in black box labels. Clin Pharmacol Ther 2002;72:735-42.
10. Willy ME, Manda B, Shatin D, Drinkard CR, Graham DJ. A study of compliance with FDA recommendations for pemoline (Cylert). J Am Acad Child Adolesc Psychiatry 2002;41:785-90.
11. Cluxton RJ, Jr., Li Z, Heaton PC, et al. Impact of regulatory labeling for troglitazone and rosiglitazone on hepatic enzyme monitoring compliance: findings from the state of Ohio medicaid program. Pharmacoepidemiology and drug safety 2005;14:1-9.
12. Gibbons RD, Brown CH, Hur K, et al. Early evidence on the effects of regulators' suicidality warnings on SSRI prescriptions and suicide in children and adolescents. American Journal of Psychiatry 2007;164:1356-63.
13. Libby AM, Brent DA, Morrato EH, Orton HD, Allen R, Valuck RJ. Decline in treatment of pediatric depression after FDA advisory on risk of suicidality with SSRIs. Am J Psychiatry 2007;164:884-91.
14. Nemeroff CB, Kalali A, Keller MB, et al. Impact of publicity concerning pediatric suicidality data on physician practice patterns in the United States. Arch Gen Psychiatry 2007;64:466-72.
15. Valuck RJ, Libby AM, Orton HD, Morrato EH, Allen R, Baldessarini RJ. Spillover effects on treatment of adult depression in primary care after FDA advisory on risk of pediatric suicidality with SSRIs. American Journal of Psychiatry 2007;164:1198-205.
16. Morrato EH, Libby AM, Orton HD, et al. Frequency of provider contact after FDA advisory on risk of pediatric suicidality with SSRIs. American Journal of Psychiatry 2008;165:42-50.
17. Libby AM, Orton HD, Valuck RJ. Persisting decline in depression treatment after FDA warnings. Archives of General Psychiatry 2009;66:633-9.
18. Stewart KA, Natzke BM, Williams T, Granger E, Casscells SW, Croghan TW. Temporal trends in anti-diabetes drug use in TRICARE following safety warnings in 2007 about rosiglitazone. Pharmacoepidemiology & Drug Safety 2009;18:1048-52.
19. Busch SH, Frank RG, Leslie DL, et al. Antidepressants and suicide risk: how did specific information in FDA safety warnings affect treatment patterns? Psychiatr Serv 2010;61:11-6.

20. Morrato EH, Druss B, Hartung DM, et al. Metabolic testing rates in 3 state Medicaid programs after FDA warnings and ADA/APA recommendations for second-generation antipsychotic drugs. Archives of General Psychiatry 2010;67:17-24.
21. Valluri S, Zito JM, Safer DJ, Zuckerman IH, Mullins CD, Korelitz JJ. Impact of the 2004 Food and Drug Administration pediatric suicidality warning on antidepressant and psychotherapy treatment for new-onset depression. Med Care 2010;48:947-54.
22. Kales HC, Zivin K, Kim HM, et al. Trends in antipsychotic use in dementia 1999-2007. Archives of General Psychiatry 2011;68:190-7.
23. Soumerai SB, Avorn J, Gortmaker S, Hawley S. Effect of government and commercial warnings on reducing prescription misuse: the case of propoxyphene. American Journal of Public Health 1987;77:1518-23.
24. Kleinbaum DG, Klein M. Survival Analysis: A self-Learning Text. Second ed. New York, NY: Springer; 2005.
25. Woolridge J. Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: The MIT Press; 2002.
26. Whitaker H. The self controlled case series method. BMJ 2008;337:a1069.
27. Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: the self-controlled case series method. Statistics in medicine 2006;25:1768-97.
28. Dusetzina SB, Higashi AS, Dorsey ER, et al. Impact of FDA drug risk communications on health care utilization and health behaviors: a systematic review. Med Care 2012;50:466-78.
29. Serumaga B, Ross-Degnan D, Avery AJ, et al. Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study. BMJ 2011;342:d108.
30. Kozhimannil KB, Adams AS, Soumerai SB, Busch AB, Huskamp HA. New Jersey's efforts to improve postpartum depression care did not change treatment patterns for women on medicaid. Health Aff (Millwood) 2011;30:293-301.
31. Grootendorst PV, O'Brien BJ, Anderson GM. On becoming 65 in Ontario. Effects of drug plan eligibility on use of prescription medicines. Med Care 1997;35:386-98.
32. Briesacher BA, Soumerai SB, Field TS, Fouayzi H, Gurwitz JH. Medicare part D's exclusion of benzodiazepines and fracture risk in nursing homes. Arch Intern Med 2010;170:693-8.
33. Zhang Y, Adams AS, Ross-Degnan D, Zhang F, Soumerai SB. Effects of prior authorization on medication discontinuation among Medicaid beneficiaries with bipolar disorder. Psychiatr Serv 2009;60:520-7.
34. Yih WK, Kulldorff M, Fireman BH, et al. Active surveillance for adverse events: the experience of the Vaccine Safety Datalink project. Pediatrics 2011;127 Suppl 1:S54-64.
35. Afendulis CC, He Y, Zaslavsky AM, Chernew ME. The impact of Medicare Part D on hospitalization rates. Health services research 2011;46:1022-38.
36. Morrato EH, Newcomer JW, Kamat S, Baser O, Harnett J, Cuffel B. Metabolic screening after the American Diabetes Association's consensus statement on antipsychotic drugs and diabetes. Diabetes Care 2009;32:1037-42.
37. Briesacher B, Limcangco R, Simoni-Wastila L, Doshi J, Gurwitz J. Evaluation of nationally mandated drug use reviews to improve patient safety in nursing homes: a natural experiment. J Am Geriatr Soc 2005;53:991-6.
38. Briesacher BA, Stuart B, Peluso R. Drug use and prescribing problems in the community-dwelling elderly: a study of three state Medicaid programs. Clin Ther 1999;21:2156-72.

39. Briesacher BA, Zhao Y, Madden JM, et al. Medicare part D and changes in prescription drug use and cost burden: national estimates for the Medicare population, 2000 to 2007. Med Care 2011;49:834-41.
40. Wagner AK, Chan KA, Dashevsky I, et al. FDA drug prescribing warnings: is the black box half empty or half full? Pharmacoepidemiology and drug safety 2006;15:369-86.
41. Soumerai SB, Avorn J, Ross-Degnan D, Gortmaker S. Payment restrictions for prescription drugs under Medicaid. Effects on therapy, cost, and equity. N Engl J Med 1987;317:550-6.
42. Soumerai SB, Ross-Degnan D, Fortess EE, Abelson J. A critical analysis of studies of state drug reimbursement policies: research in need of discipline. Milbank Q 1993;71:217-52.
43. Soumerai SB, McLaughlin TJ, Ross-Degnan D, Casteris CS, Bollini P. Effects of a limit on Medicaid drug-reimbursement benefits on the use of psychotropic agents and acute mental health services by patients with schizophrenia. N Engl J Med 1994;331:650-5.
44. Soumerai S. Unintended outcomes of medicaid drug cost-containment policies on the chronically mentally ill. J Clin Psychiatry 2003;64 Suppl 17:19-22.
45. Briesacher B, Ross-Degnan D, Adams A, Wagner A, Gurwitz J, Soumerai S. A new measure of medication affordability. Social work in public health 2009;24:600-12.
46. Zhang F, Wagner AK, Soumerai SB, Ross-Degnan D. Methods for estimating confidence intervals in interrupted time series analyses of health interventions. Journal of clinical epidemiology 2009;62:143-8.
47. Velentgas P, Bohn RL, Brown JS, et al. A distributed research network model for post-marketing safety studies: the Meningococcal Vaccine Study. Pharmacoepidemiology and drug safety 2008;17:1226-34.
48. Soumerai SB, Zhang F, Ross-Degnan D, et al. Use of atypical antipsychotic drugs for schizophrenia in Maine Medicaid following a policy change. Health Aff (Millwood) 2008;27:w185-95.