Supplementary Online Content


eAppendix 1. Statistical Analysis
eAppendix 2. Inclusion and Exclusion Criteria
eAppendix 3. Algorithm for Outcome Ascertainment: Sensitivity Analyses
eAppendix 4. Unobserved Confounding: Sensitivity Analysis
eAppendix 5. Lead Time Bias: Sensitivity Analysis
eAppendix 6. Lead Time Bias: Partial Adjustment for Bias
eTable. Propensity Score Covariates for Balancing
eReferences

This supplementary material has been provided by the authors to give readers additional information about their work.
**eAppendix 1. Statistical Analysis**

*Month-level pseudotrial design*

As the main text explains, to model the effect of antipsychotic initiation on incident diabetes, we recast the traditional view of an observational cohort with time varying exposures into one of a sequence of monthly non-randomized trials (termed pseudotrials), mirroring an approach described by Hernan and colleagues. In a two-step approach, we first modeled medication exposure as a function of baseline covariates and then modeled time to onset of diabetes in a discrete time failure paradigm. Details on the rationale for our implementation, alternative approaches, and choices of the modeling steps appear in a prior publication. We summarize the steps below.

The first of 50 pseudotrials began in November 2003, after the end of a 10-month look-back period to identify eligibility criteria; the last pseudotrial began in December 2007. At the start of each pseudotrial, eligibility was determined using the inclusion and exclusion criteria described in the main text and in eMethods 2. The analysis compared eligible children who initiated antipsychotics at the month of pseudotrial entry, determined by a first filled prescription in that month, to eligible children who did not initiate antipsychotics in that month.

The analysis proceeded in two steps: an initial set of propensity-score models to balance covariates between exposed and unexposed groups, and a second set of response models to estimate the association of exposure and diabetes.

*Propensity-score based weights*

In the propensity score model for each pseudotrial, the probability of second-generation antipsychotic (antipsychotic) treatment assignment at the month of pseudotrial entry was modeled using multivariable logistic regression. Children in the exposed group received a weight of 1.0. Children in the unexposed group received a weight of \( \frac{p_j}{1-p_j} \), the odds of exposure, where \( p_j \) is the estimated probability of exposure from the propensity score models for the \( i \)th child in the \( j \)th pseudo-trial.

In the case of slight discrepancies in summed weights among the non-initiators of antipsychotics, all weights assigned to non-initiators were adjusted proportionally to ensure equality in summed weights between initiators and non-initiators. Balance checks of clinical and demographic characteristics at baseline between initiators and non-initiators, using the propensity-score based weights, were conducted within each pseudotrial and in the pooled set of all pseudotrials, using the resulting weights.

We calculated new propensity-score-based weights for each contrast of medication class, in order to achieve balance of baseline covariates at each pseudo-trial for the contrast of interest. For the analysis of risk in different medications within drug class, we again calculated new propensity-score-based weights, in order to achieve balance of baseline covariates across initiators of all five antipsychotic medications. In each case, weights standardized to the characteristics of the children in the reference group.

*Response models*

The response model, a discrete time failure model, applied weighted logistic regression to the stacked set of 50 pseudo-trials to estimate the association of antipsychotic initiation and the risk of developing diabetes during each month of follow-up. Discrete time failure models are an especially appropriate alternative to the more common Cox proportional hazards model for time to event analyses when, as in our case, time is inherently lumped into a finite interval and event times are tied, though both models handle time-varying exposures. For our implementation, all response models were weighted to balance the covariates and to achieve proper standardization, thus adjusting for observed confounders. As the weighting scheme standardizes the unexposed group to the characteristics of the exposed, the response model estimates could then be interpreted as the effect of exposure among the exposed.

In secondary analyses, we examined the risk of incident diabetes associated with multi-drug regimens with antipsychotics and other medications (e.g., stimulants). For each contrast, we estimated the association of the
medication(s) of interest and incident type 2 diabetes, standardized to the socio-demographic and clinical characteristics of a standardizing population.

The first set of analyses examined the simultaneous use of antipsychotics and stimulants. We compared the risk of diabetes in a standardizing population of simultaneous users of antipsychotics and stimulants to three other groups. 1) those who received neither antipsychotic nor stimulant use; 2) those receiving only antipsychotics; and 3) those receiving only stimulants. These three contrasts were mirrored for antidepressant modeling.

In addition, we conducted within-medication-class contrasts by specific antipsychotic medication, namely olanzapine, clozapine, risperidone, quetiapine, ziprasidone and aripiprazole. Based on a literature review, we hypothesized that olanzapine and clozapine would have the highest risk, followed by risperidone and quetiapine, and with ziprasidone and aripiprazole having the lowest risk.5-12
eAppendix 2. Inclusion and Exclusion Criteria

1) Fee-for-service Medicaid payment arrangements

In order to identify children in fee-for-service Medicaid payment arrangements, we used the following variables in the Medicaid Analytic eXtract (MAX) data: EL_PHP_TYPE_1_1 - EL_PHP_TYPE_1_12; EL_PHP_TYPE_2_1 - EL_PHP_TYPE_2_12; EL_PHP_TYPE_3_1 - EL_PHP_TYPE_3_12; and EL_PHP_TYPE_4_1 - EL_PHP_TYPE_4_12. For each month of a child’s enrollment in Medicaid, these variables identify the types (up to 4) of pre-paid plans for which the child was eligible. A child was considered to be enrolled in a Medicaid managed care payment arrangement if he or she had one of the following three values for one of these variables: 01 (“Eligible is enrolled in a medical or comprehensive managed care plan this month (e.g. HMO”), 03 (“Eligible is enrolled in a behavioral managed care plan this month”), or 08 (“Eligible is enrolled in another managed care plan this month”). Otherwise, the child was considered to be in fee-for-service.

2) Mental health diagnoses

For ascertainment of mental health diagnosis-related inclusion and exclusion criteria at pseudotrial entry, ten diagnostic categories were identified in the Diagnostic and Statistical Manual of Mental Disorders IV-TR and coded using the International Classification of Diseases, Ninth Revision (ICD-9) classification: schizophrenia (295), bipolar disorder (296.00–296.10, 296.36–296.89), depression (296.20–296.35, and 311), anxiety disorder (300.00–300.29 and 301.4), conduct disorder (312.00–313.89), autism (299), attention deficit disorder (314), intellectual disability (317–319), developmental delay (315), and a composite variable of miscellaneous mental health diagnoses, inclusive of the following: mental disorders due to conditions classified elsewhere (293, 294); delusional disorders (297); other nonorganic psychoses (298); dissociative and somatoform disorders (300.10–300.39, 300.99); personality disorders (301.10–301.99); special symptoms or syndromes, not elsewhere classified (307); acute reaction to stress (308); adjustment reaction (309); and disturbance of emotions specific to childhood and adolescence (313.90–313.99).

For the purpose of coding mental health diagnoses longitudinally, we considered autism, developmental delay, and intellectual disability to be time-invariant; thus, if a child had one of these diagnoses at any point during her observation window, she was considered to have this diagnosis for her entire observation window. The remaining seven diagnoses (schizophrenia, bipolar disorder, depression, anxiety disorder, conduct disorder, attention deficit disorder, and miscellaneous mental health diagnoses) were coded forward from the first identification of diagnosis in the dataset.

3) Polycystic ovary syndrome

Polycystic ovary syndrome was identified in the dataset using the ICD-9 code 256.4.

4) Definition of chronic steroid use

Chronic use of inhaled corticosteroid or oral steroid was defined as at least 14 days of use per month (identified using filled prescriptions) for at least two calendar months in the look-back period.
eAppendix 3. Algorithm for Outcome Ascertainment: Sensitivity Analyses

We conducted two sensitivity analyses of our algorithm for ascertainment of the outcome of type 2 diabetes. In the first sensitivity analysis, we used a wider time window for a given child to satisfy the diagnosis and prescription requirements of the algorithm (6 months, instead of 4). For this purpose, we used a sample of children with enrollment in a fee-for-service Medicaid payment arrangement to compare the incidence of diabetes using the two definitions. On average, using the 4-month time window for identification of diabetes, 684 incident diabetes cases were identified per year. Using a wider window of 6 months, an average of 36 additional cases per year (5% additional cases) would have been identified.

The second sensitivity analysis recoded all children with a type 1 diabetes diagnosis, who did not have a prescription for insulin at any time, as having type 2 diabetes. In this case we replicated the primary analysis, using this new outcome definition, in order to observe the difference in results. The resulting odds ratio was 1.52 (95% CI, 1.38 to 1.68). The odds ratio in the main analysis was 1.51 (95% CI, 1.35 to 1.69).
eAppendix 4. Unobserved Confounding: Sensitivity Analysis

The method used by Lin and colleagues is a commonly used method to assess the sensitivity of regression results to unmeasured confounders in observational studies. The method allows us to determine, for a given unobserved confounder, the required differential between exposure groups in 1) prevalence of the confounder and 2) association between the confounder and the outcome, in order for our findings to lose statistical significance.
eAppendix 5. Lead Time Bias: Sensitivity Analysis

Lead time of diagnosis of diabetes could bias results if children on medication were selectively more likely to be monitored for diabetes and, thus, were likely to be diagnosed earlier.\textsuperscript{14,15} Lead time could, at least temporarily, result in an inflated risk of diabetes among those children who initiated medication and were being screened, even in the absence of any actual elevated risk of diabetes. Over time, lead time bias should attenuate, but only after extended follow-up. Because of a lack of good information about the latency period between actual disease onset and the clinical manifestation of diabetes, our sensitivity analysis assumed that the effects of lead time would persist beyond the duration of observation in this study. Below we outline a sensitivity analysis of the possible observed association of initiation of medication and development of diabetes under the null hypothesis of no association, i.e., of no effect of medication on diabetes risk.

Sensitivity analyses were empirically based on our data. Glucose screenings were identified in the claims data using the following \textit{Current Procedural Terminology (CPT)} procedure codes: a comprehensive metabolic or general health panel (80048, 80050, 80053, or 80054), a glucose test (82947, 82948, 82950, 82951, 81000, 81002, 81005, or 81099), a glycolated hemoglobin (A1C) test (83036), or a home glucose monitoring device (82962).\textsuperscript{16,17} We excluded all tests occurring in the month of the diagnosis or the two months prior.\textsuperscript{18}

For the analyses of glucose testing, we attempted to select a state in each Census division where CPT codes were used reliably to identify procedures. As such, states were excluded from selection if they had either a) low utilization of CPT and Healthcare Common Procedure Coding System (HCPCS) codes in the claims field for procedures or b) a very low percentage of children with a glucose or lipid test. These two criteria largely overlapped. To ensure a large enough sample of eligible children that was consistent across years, we also excluded states with a low percentage of children in fee-for-service Medicaid payment arrangements, or that had large shifts in Medicaid payment arrangements over the time period. Finally, of the remaining states, we randomly selected one state from each Census division.

From our data, we found that follow-up time was about the same in the antipsychotic and non-antipsychotic groups; in the nine states selected, follow-up was, on average, 27 months in the first set of 10 pseudotrials and 22 months in the second set of 10 pseudotrials, with shorter length of follow-up in the remaining 30 pseudotrials. Assuming, conservatively, that follow-up was approximately 25 months, and that the underlying risk of diabetes was 0.0002 per month, regardless of antipsychotic usage, we considered three scenarios to determine the possible impact of increased likelihood of screening on our outcome.

First, we considered the case where all individuals were screened. Our data suggested that among those who were actually screened, the time to the initial screen was shorter for the antipsychotic initiators. From our data, we estimated conservatively that the lead time for screening was 4 months for the antipsychotic users; that is, they were screened 4 months earlier than non-initiators. Assuming that the true rate of diabetes was the same in both groups, and the only reason for an observed increase in risk of diabetes was lead time, we would anticipate the cumulative incidence of diabetes for antipsychotic users at 21 months to be the same as that for non-users at 25 months. Then, the cumulative incidence for antipsychotic initiators would be 0.00499 at 25 months ($=1-\exp(25*ln(1-0.0002))$). For the non-initiators, this incidence would correspond to that for 21 months of follow-up, giving 0.00419 ($=1-\exp(21*ln(1-0.0002))$). Thus, we could expect a lead time bias equivalent to a ratio of about 1.2 ($=0.00499/0.00419$) if all children were being tested, although at different frequencies, assuming the underlying risk of diabetes was the same in both groups. This estimate fell below the bounds of our confidence interval for the association of initiation of diabetes medication and development of diabetes.

In the second scenario we assumed that only a certain percentage of children were tested (a more realistic assumption). Based on our data, approximately 50% of antipsychotic initiators were screened, compared to about 30% of non-initiators. To arrive at estimates using these screening percentages, we used a paradigm of 100,000 children, assigned alternatively to two screening protocols: one for antipsychotic initiators and one for non-initiators. We then estimated the number of children who would have a diagnosis of diabetes under these two alternative screening frequencies. We assumed the underlying rate of diabetes was the same for all children, and for 25 months of follow-up, the cumulative incidence of diabetes was approximately 0.005 ($=25*0.0002$). Among these 100,000
children, 50,000 would never be screened, 30,000 would be screened regardless of treatment assignment, and 20,000 would be screened only if initiating an antipsychotic.

Assuming a lead time bias of as much as 12 months for antipsychotic initiators who were screened compared to non-initiators who were not screened, we would expect to discover ½ the number of diabetes cases among the non-initiators. We make the severe assumption of a 12 month lead time, because if patients are not being screened at all, then the physician must wait for other clinical signs and symptoms to arrive at a decision to conduct a diagnostic test. Thus, among the 20,000 children who would be screened only if initiating an antipsychotic, there would be 100 (=20,000*0.005) cases among the antipsychotic initiators and 50 cases among the non-initiators. Among the 50,000 who would not be screened, we would expect 250 cases of diabetes (=50,000*0.005), regardless of treatment assignment. Finally, among the 30,000 who would be screened in both groups, we would expect 150 (=30,000*0.005) cases in the antipsychotic group, but perhaps only 125 (=150/1.2) among the non-initiators. This estimate follows from the first set of calculations of lead time bias described above, due to the 4 month lag in screening for screened non-initiators.

Thus, under the null hypothesis of no association between antipsychotic and diabetes, we would have 500 (=100+250+150) diabetes cases among the antipsychotic users and 425 (=50+250+125) among the non-users, yielding a ratio of 1.18 (=500/425). Again, this estimate fell below our confidence bounds for the association of initiation of medication and the risk of diabetes.

The final scenario we considered modified the second scenario and assumed that those assigned to antipsychotics had an increased rate of diagnosis (by virtue of screening only). We again assumed that the accelerated time to diagnosis was 12 months, and there would be a 50% increase in the number of diagnosed cases among those assigned to antipsychotics over the 25-month follow-up period. This increase of the rate of diagnoses attributable only to screening is unlikely. Now, among the 20,000 who would be screened only if assigned to antipsychotics screening practice, there would be 100 (=20,000*0.005) cases among the non-users, and 150 among the antipsychotic users. We would expect to see the same number of cases as in the second scenario for the never-screened and always-screened groups. Then, under the null of no association between antipsychotics and diabetes, there would be 550 (=150+250+150) diabetes cases among the antipsychotic users and 475 (=100+250+125) among the non-users, yielding a ratio of 1.16 (=550/475). Under the extreme assumption that the lead time bias was 25 months (instead of 12) for the 20,000 who would be screened only if assigned to antipsychotics, then there would be 475 (=100+250+125) non-antipsychotics cases, but there would be 600 (=200+250+150) antipsychotics cases, yielding a ratio of 1.26 (=600/475). Thus, even under extreme assumptions of lead time bias, the observed odds ratio under the null would be no more than 1.26, given our estimates of the probability of screening and the frequency of screening given that the child was screened. This estimate also fell below our confidence bounds.
eAppendix 6. Lead Time Bias: Partial Adjustment for Bias

The potential for lead time bias would be offset, or controlled at least in part, if covariates already in our main analysis explained some of the differences in glucose screening. To investigate this potential, we used logistic regression applied to the stacked set of pseudo-trials with screening as the response and adjusting for the clinical and demographic covariates used in the propensity score model of the main analysis.

Without adjustment for covariates, antipsychotic initiators were significantly more likely to have a glucose screening during follow-up than non-initiators (odds ratio, 2.20; 95% confidence interval [CI], 2.15 to 2.24; P<0.001). After adjustment for covariates, the odds ratio was reduced to 1.58 (95% CI, 1.54 to 1.62; P<0.001). These results suggested that screening rates in antipsychotic initiators vs. non-initiators could be at least partially explained by the covariates balanced on in the main analyses, and that our analyses adjusted, at least partially, for lead time bias.
### eTable. Propensity Score Covariates for Balancing

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Additional Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Binary</td>
<td>Age was coded as younger (ages 10-14) or older (ages 15-18) at entry into the pseudotrial</td>
</tr>
<tr>
<td>Age</td>
<td>Binary</td>
<td>Race/ethnicity was coded as Hispanic, black, white, other, or unknown. In cases where the race or ethnicity of a child was coded differently across years in the Medicaid data, the following algorithm was implemented to assign race/ethnicity for the study. First, children who ever had race/ethnicity identified as Hispanic were always coded as “Hispanic.” Otherwise, children whose race was identified as Asian or “other” and never identified as Hispanic were always coded as “other.” Children whose race was identified as black and never identified as Hispanic, Asian, or other were always coded as “black.” Children whose race was identified as white and never identified as Hispanic, Asian, other, or black were always coded as “white.” Otherwise, race was coded as “unknown.”</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td>Categorical</td>
<td>Medicaid eligibility status was coded as eligible based on foster care status, eligible due to Supplemental Security Income (SSI)/disability, or income-eligible through Temporary Assistance for Needy Families (TANF)/other eligibility.</td>
</tr>
<tr>
<td>Medicaid eligibility status</td>
<td>Categorical</td>
<td>The nine Census divisions were: Division 1: New England (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont); Division 2: Middle Atlantic (New Jersey, New York, Pennsylvania); Division 3: East North Central (Indiana, Illinois, Michigan, Ohio, Wisconsin); Division 4: West North Central (Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota); Division 5: South Atlantic (Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia); Division 6: East South Central (Alabama, Kentucky, Mississippi, Tennessee); Division 7: West South Central (Arkansas, Louisiana, Oklahoma, Texas); Division 8: Mountain (Arizona, Colorado, Idaho, New Mexico, Montana, Utah, Nevada, Wyoming); and Division 9: Pacific (Alaska, California, Hawaii, Oregon, Washington).</td>
</tr>
<tr>
<td>Census division</td>
<td>Categorical</td>
<td>Eight binary variables were used to identify a history of each of eight mental health diagnoses at entry into the pseudotrial. These included: bipolar disorder, depression, anxiety disorder, conduct disorder, autism, attention-deficit/hyperactivity disorder (ADHD), intellectual disability, and developmental delay.</td>
</tr>
</tbody>
</table>

*ADHD was not included as a balancing covariate in analyses of multiple-drug regimens with stimulants.*
<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Additional Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental health diagnoses</td>
<td>Additional Information</td>
<td>Mental health diagnoses were coded as above in eMethods 1.</td>
</tr>
<tr>
<td>Prior diagnosis of lipid or metabolic disorder or hypertension</td>
<td>Binary</td>
<td>A single binary variable was used to capture whether the child had a lipid disorder, a metabolic disorder, or hypertension prior to entry into the pseudotrial. Lipid disorders were defined by the following ICD-9 codes: 272.0 (pure hypercholesterolemia), 272.1 (pure hyperglyceridemia), 272.2 (mixed hyperlipidemia), and 272.4 (other and unspecified hyperlipidemia). Metabolic disorders were defined by the following ICD-9 codes: 277.7 (metabolic syndrome); 278.00, 278.01, or 278.03 (obesity); and 790.2x (abnormal glucose). Hypertension was defined by ICD-9 codes 401.xx.</td>
</tr>
<tr>
<td>Complex chronic conditions</td>
<td>Binary</td>
<td>A single binary variable was used to indicate whether the child had a history of any one of nine types of complex chronic conditions at entry into the pseudotrial. Identification of complex chronic conditions followed the specification provided in Feudtner et al. (2000)²⁰. ICD-9 codes were as follows: Neuromuscular conditions: 740.0-742.9; 318.0-318.2; 330.0–330.9, 334.0–334.2, 335.0–335.9; 343.0-343.9; 359.0-359.3 Cardiovascular conditions: 745.0-747.4; 425.0-425.4, 429.1; 426.0-427.4; 427.6-427.9 Respiratory conditions: 748.0-748.9; 770.7; 277.0 Renal conditions: 753.0-753.9; 585 Gastrointestinal conditions: 750.3, 751.1–751.3, 751.6–751.9; 571.4-571.9; 555.0-556.9 Hematologic or immunologic conditions: 282.5-282.6; 282.0-282.4; 279.00–279.9, 288.1–288.2, 446.1; 0420-0421 Metabolic conditions: 270.0-270.9; 271.0-271.9; 272.0-272.9; 277.3, 277.5; 275.0–275.3, 277.2, 277.4, 277.6, 277.8–277.9 Other congenital or genetic defect: 758.0-758.9; 259.4, 737.3, 756.0–756.5; 553.3, 756.6–756.7; 759.7–759.9 Malignancy: 140.0–208.9, 235.0–239.9</td>
</tr>
<tr>
<td>Variable</td>
<td>Type</td>
<td>Additional Information</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Other psychotropic medications</td>
<td>6 binary variables</td>
<td>Six binary variables were used to indicate the use of each of six other psychotropic medications, aside from antipsychotics, either in the month of entry into the pseudotrial or in the three months prior to pseudotrial entry.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The six other psychotropic medication classes included stimulants,† antidepressants,‡ mood stabilizers, alpha agonists, sedatives/hypnotics, and anxiolytics.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Antidepressants included selective serotonin reuptake inhibitors (SSRIs), tricyclic antidepressants (TCAs), and other antidepressants. Mood stabilizing agents included carbamazepine, valproic acid, gabapentin, lamotrigine, and oxcarbazepine anticonvulsants and lithium. Sedatives/hypnotics excluded antihistamines, which in pediatric practice most often have a non-psychiatric indication for use.</td>
</tr>
</tbody>
</table>

† The use of stimulants was not included as a balancing covariate in analyses of multi-drug regimens with stimulants.
‡ The use of antidepressants was not included as a balancing covariate in analyses of multi-drug regimens with antidepressants.
References


© 2015 American Medical Association. All rights reserved.