

Antipsychotics, Glycemic Disorders, and Life-Threatening Diabetic Events: A Bayesian Data-Mining Analysis of the FDA Adverse Event Reporting System (1968-2004)

WILLIAM DUMOUCHEL, PhD, DAVID FRAM, BA, and XIONGHU YANG, MD, PhD

Lincoln Technologies Inc., Waltham, Massachusetts, USA

RAMY A. MAHMOUD, MD and AMY L. GROGG, PHARM D

Ortho-McNeil Janssen Scientific Affairs, L.L.C., Titusville, NJ, USA

LUELLA ENGELHART, MS

Cordis Corporation, a Johnson & Johnson Company, Warren, NJ, USA

KRISHNAN RAMASWAMY, PhD

Ortho-McNeil Janssen Scientific Affairs, L.L.C., Titusville, NJ, USA

Background. This analysis compared diabetes-related adverse events associated with use of different antipsychotic agents. A disproportionality analysis of the US Food and Drug Administration (FDA) Adverse Event Reporting System (AERS) was performed.

Methods. Data from the FDA postmarketing AERS database (1968 through first quarter 2004) were evaluated. Drugs studied included aripiprazole, clozapine, haloperidol, olanzapine, quetiapine, risperidone, and ziprasidone. Fourteen Medical Dictionary for Regulatory Activities (MedDRA) Primary Terms (MPTs) were chosen to identify diabetes-related adverse events; 3 groupings into higher-level descriptive categories were also studied. Three methods of measuring drug-event associations were used: proportional reporting ratio, the empirical Bayes data-mining algorithm known as the Multi-Item Gamma Poisson Shrinker, and logistic regression (LR) analysis. Quantitative measures of association strength, with corresponding confidence intervals, between drugs and specified adverse events were computed and graphed. Some of the LR analyses were repeated separately for reports from patients under and over 45 years of age. Differences in association strength were declared statistically significant if the corresponding 90% confidence intervals did not overlap.

Results. Association with various glycemic events differed for different drugs. On average, the rankings of association strength agreed with the following ordering: low association, ziprasidone, aripiprazole, haloperidol, and risperidone; medium association, quetiapine; and strong association, clozapine and olanzapine. The median rank correlation between the above ordering and the 17 sets of LR coefficients (1 set for each glycemic event) was 93%. Many of the disproportionality measures were significantly different across drugs, and ratios of disproportionality factors of 5 or more were frequently observed.

Conclusions. There are consistent and substantial differences between atypical antipsychotic drugs in the disproportionality reporting ratios relating to glycemic effects, especially life-threatening events, in the AERS database. The relative associational rankings of drugs are similar in reports from younger and older patients. These results agree

with several other reports in the literature, do not support a “class effect” hypothesis, and provide a strong rationale for further studies to clarify the issue.

Keywords Atypical antipsychotics, Diabetes, Glucose regulation

INTRODUCTION

Abnormal glucose regulation and diabetes appear to occur more frequently in patients with schizophrenia and other psychiatric illnesses than in the general population (1). Furthermore, diabetes and related adverse events are a matter of rising concern for patients taking atypical antipsychotics, as the number of reports of diabetes-related adverse events for patients on these drugs increases (2–5). There was substantial press coverage of this issue in 2003, followed by requests by the US Food and Drug Administration (FDA) to add a “class label” warning to the package inserts for all atypical antipsychotic drugs available in the United States. Previous retrospective analyses of the potential association between antipsychotic use and diabetes vary greatly in statistical methods and in quality of methodology.

Jin et al. (6) reviewed 45 case reports of new-onset diabetes mellitus and diabetic ketoacidosis (DKA) after initiation of atypical antipsychotics. Of these cases, 42% presented as DKA. Twenty patients had received clozapine, 19 received olanzapine, 3 received quetiapine, and 3 received risperidone. The significance of these numbers is difficult to ascertain in the absence of a suitable denominator.

In three separate studies, Koller et al. (7–9) evaluated MedWatch reports and MEDLINE publications documenting glycemic adverse events among patients treated with clozapine (8), olanzapine (9), or risperidone (7). While these studies showed that a greater number of diabetes-related adverse events occurred in patients taking olanzapine and clozapine than risperidone, despite substantially greater patient exposure to risperidone, neither conclusions about causality nor accurate comparison of different drug risks could be made from these descriptive analyses (7–9). Again, comparisons among drugs were difficult to interpret, since no reasonably accurate drug exposure denominator was reported. Another review using the MedWatch drug surveillance system to assess olanzapine and clozapine showed similar results, in which patients aged 13 to 18 years treated with clozapine and olanzapine experienced glycemic adverse events (10). In the studies of Jin et al. (6) and Koller et al. (7–9) the interval between initiation of the antipsychotic and the occurrence of DKA was similar—approximately 3 months for many patients.

Much of the available data on relationships between antipsychotic agents and diabetes-related symptoms are based on case reports, chart reviews, and cross-sectional studies. These types of data do not permit rigorous characterization of degree of risk, and there is considerable debate surrounding whether the observed relationships represent a class effect of all antipsychotics (1–6,11–18). Although head-to-head clinical trials designed

specifically to test this issue are lacking and may never be performed, there is now substantial epidemiologic evidence suggesting that use of certain atypical antipsychotics, particularly olanzapine and clozapine, can increase the risk of diabetes (14,19–23). There is, however, a small body of contradictory evidence from studies that suggest that olanzapine may not be associated with more diabetes-related adverse events than other atypical antipsychotics (12,24). Interpretation of these data remains difficult, since reported excess risk and relative risk are inconsistent. Furthermore, in addition to other methodologic differences, potentially confounding variables such as age, dose, duration of exposure, ethnicity, sex, body mass index, and use of concomitant medications are treated differently in these studies, further increasing the complexity of interpretation.

The database analyses presented in this study suffer from some unavoidable shortcomings of the retrospective and uncontrolled nature of spontaneous reporting. We provide a systematic analysis of more than 2.4 million reports in the US FDA Adverse Event Reporting System (AERS), which until 1997 was called the Spontaneous Reporting System (SRS). Three separate methods of disproportionality analysis are used. They each use statistical models to generate a surrogate denominator for counts of drug-event combinations. This allows the estimation of interpretable disproportionality reporting ratios that measure how much more frequently each drug-event combination occurs in the database than it would if that drug-event pair were not associated in the database. The relation between such reporting ratios and the corresponding relative risks or odds ratios that might be measured by a case-control study is difficult to determine. These methods generate standard errors and confidence intervals that can validate the statistical reliability of the measured associations. As with any statistical associations, the question of whether these associations are causal is a separate one. The presented analyses differ in their complexity and in the ways they attempt to adjust for potential confounding relationships. The purpose of this study is to provide an overview of the strength and statistical reliability of disproportionality measures between 7 different drugs (6 atypical antipsychotics plus haloperidol) and glycemic effects within the FDA’s AERS database of spontaneous reports. Differences among drugs found here should be viewed as generating hypotheses about corresponding differences in the drugs’ risk of metabolic damage.

METHODS

Data were garnered from the FDA postmarketing AERS database. Reports to this database are submitted by manufacturers

(as mandated by the FDA) and by healthcare providers and patients (voluntarily through the MedWatch program). Raw data for the period 1968 through first quarter 2004 were obtained in ASCII format on CD-ROM from the National Technical Information Service (25). As previously reported, substantial preprocessing of the data was performed (recoding, standardization of nomenclature, elimination of duplicate cases) to make it appropriate for data-mining analysis (26). All reported adverse event terms were mapped to the Medical Dictionary for Regulatory Activities (MedDRA) Version 7.1 Preferred Terms (MPTs). MedDRA is an internationally harmonized, fine-grained medical dictionary for drug safety reporting that now contains more than 15,000 preferred terms representing distinct medical concepts. The resulting cleaned database (hereafter referred to as AERS) represented more than 2.4 million reports, which were mined for this study.

Drugs selected for the study included all six atypical antipsychotics for which meaningful population exposure was expected (aripiprazole, clozapine, olanzapine, quetiapine, risperidone, and ziprasidone) plus haloperidol, an older conventional antipsychotic not generally thought to be associated with diabetes-related risk. Two of the drugs, aripiprazole and ziprasidone, are relatively new to the market, and thus have somewhat sparse representation in the AERS database, which limits the statistical reliability of estimates relating to these drugs.

Fourteen MPTs related to diabetes were chosen to identify diabetes-related adverse events. From the original group of 14 distinct MPTs, 13 of them were divided into 3 groups of highly interrelated, clinically similar terms that would be unlikely to be accurately and systematically differentiated in clinical practice and in coding of reports (i.e., on different occasions, different reporters might select different terms from within these categories

to represent the same medical condition). Analysis of a grouped event consists of treating the event as having been reported if one or more of the MPTs in the group are mentioned in the report. Analyses of grouped events have the advantage of being based on larger, and thus more reliable, reporting frequencies. The analyses by individual MPTs were therefore augmented by analyses of the three groupings of MPTs, resulting in a total of 17 glycemic events for which measures of association strength were computed for each of the seven drugs. The first column of Table 1 shows the MPTs that were selected as well as the three defined groups of MPTs. One of the MPTs, Diabetes mellitus inadequate control, was not grouped with any others.

Certain analyses were repeated for separate subsets of reports from younger (aged ≤ 45 years) and older (aged > 45 years) patients. These subset analyses are presented only for the grouped events, both for brevity and to ensure a sufficient sample size for reliable interpretation of results.

The first measure of association used is the proportional reporting ratio (PRR). Within the database, for each drug-event combination being studied, the 2 × 2 table of counts is defined as:

| Count of reports | Mentioning Drug of Interest | Omitting Drug of Interest | Total |
|------------------------------|-----------------------------|---------------------------|---------|
| Mentioning Event of Interest | a | b | a+b |
| Omitting Event of Interest | c | d | c+d |
| Total | a+c | b+d | a+b+c+d |

Then $PRR = [a/(a+c)] / [b/(b+d)]$. The quantity PRR has the advantage of being simple to compute and simple to interpret. It has the disadvantages of being highly variable when the counts *a*, *b*, and *c* are very small, and it can be biased by the

Table 1 Counts of Reported Drug-Event Combinations in the FDA AERS Database Through 2004 Quarter 1

| Event | Aripiprazole | Clozapine | Haloperidol | Olanzapine | Quetiapine | Risperidone | Ziprasidone | All Reports |
|---|--------------|-----------|-------------|------------|------------|-------------|-------------|-------------|
| Blood glucose abnormal | 1 | 4 | 5 | 10 | 6 | 4 | 0 | 400 |
| Blood glucose increased | 15 | 80 | 48 | 209 | 44 | 59 | 14 | 6,629 |
| Glucose tolerance decreased | 0 | 2 | 3 | 1 | 0 | 3 | 0 | 276 |
| Glucose tolerance impaired | 1 | 8 | 3 | 6 | 1 | 4 | 0 | 150 |
| Glycosylated hemoglobin increased | 1 | 20 | 9 | 36 | 12 | 11 | 2 | 567 |
| Hyperglycemia | 3 | 281 | 97 | 293 | 39 | 125 | 3 | 20,637 |
| Group: Blood Glucose Abnormal | 21 | 380 | 156 | 527 | 94 | 198 | 18 | 28,282 |
| Diabetes mellitus | 5 | 404 | 75 | 241 | 63 | 150 | 17 | 8,027 |
| Diabetes mellitus insulin-dependent | 1 | 43 | 5 | 23 | 6 | 8 | 0 | 317 |
| Diabetes mellitus non-insulin-dependent | 0 | 80 | 13 | 41 | 13 | 11 | 2 | 530 |
| Group: Diabetes Mellitus | 6 | 527 | 93 | 305 | 82 | 169 | 19 | 8,863 |
| Diabetes mellitus inadequate control | 0 | 18 | 6 | 31 | 9 | 15 | 0 | 2,678 |
| Diabetic coma | 0 | 28 | 9 | 29 | 8 | 6 | 0 | 453 |
| Diabetic hyperglycemic coma | 0 | 8 | 0 | 6 | 1 | 1 | 1 | 49 |
| Diabetic hyperosmolar coma | 0 | 4 | 5 | 20 | 6 | 4 | 0 | 121 |
| Diabetic ketoacidosis | 2 | 88 | 29 | 170 | 38 | 40 | 3 | 1,196 |
| Group: Diabetic Life-Threatening Events | 2 | 116 | 39 | 216 | 50 | 47 | 4 | 1,774 |
| All Reports | 1,735 | 25,186 | 14,999 | 13,751 | 4,681 | 18,621 | 2,338 | 2,421,347 |

Note: The events labeled "Group" are counted as happening if any of the MPTs in the same section of the table occur in a report. Drugs are counted if declared either suspect or concomitant in the report.

presence of confounding variables such as age, sex, or report year, if such variables are associated with both the drug of interest and the event of interest. In addition, confidence intervals for PRR are not commonly computed, making it difficult to determine whether 2 PRRs are significantly different.

The second disproportionality technique used is the empirical Bayes data-mining algorithm known as the Multi-Item Gamma Poisson Shrinker (MGPS) (27). For each drug-event combination, the algorithm computes the empirical Bayes geometric mean (EBGM) along with its lower 5% (EB05) and upper 95% (EB95) confidence limits. EBGM is a stable estimate of the *relative reporting ratio*—the count *observed* for the given drug-event combination divided by the count that would be *expected* if drugs and events were independently distributed in the database. In terms of the 2×2 table shown above, the relative reporting ratio (RR) is defined as $RR = a/e$, where $e = (a+b)(a+c)/(a+b+c+d)$.

Further details of this method have been published (27). Note that when a is small compared with b and c , and when these are in turn small compared with d , which is usually true for events and drugs that are mentioned in a small proportion of reports, the above formulas lead one to expect that PRR will be close to RR. However, this approximate equality can be disturbed by the provision within MGPS for adjustment of expected counts to account for potentially confounding covariates. The adjustment consists of stratifying the database according to all combinations of covariate values and computing a and e separately for each stratum, whereupon RR is defined as the ratio of the sum of the a 's to the sum of the e 's. Stratifying will guard against concluding that a drug and a reaction are associated simply because the drug and the reaction both occur preferentially in some subpopulation of the database. In the present analysis, data were stratified according to 3 available major demographic variables likely to show correlation with drug and/or diabetes reporting: sex, age group, and year of FDA receipt of the report.

One phenomenon that might produce reporting rates that vary as a function of the specific drug-event combination would be media coverage of a newsworthy safety issue, or the "publicity effect." The MGPS method tolerates overall variations in reporting rates (by drug and by event), but produces biased results if reporting of a particular drug-event combination is unusually high or low. To check for this, the progression of MGPS signal scores over time was examined retrospectively, to see whether there were any surprising secular changes (e.g., a large spike in a signal score related perhaps to a publicity effect). This was done by creating a series of yearly cumulative subsets of the database (e.g., 1968–1990, 1968–1991, 1968–1992) and carrying out separate MGPS analyses on each subset.

In making comparisons of EBGM disproportionality scores, the criterion used for declaring that 2 drugs have different association strengths with a given adverse event term is that the corresponding 2 (EB05, EB95) confidence intervals do not overlap.

In comparison with PRR, EBGM has the advantages that it is less variable for small counts, that it can be adjusted for confounding covariates, and that it has a corresponding confidence interval methodology built into its computer algorithm. However, the adjustment for confounders by stratification does not scale up well if there are very many potentially confounding variables. For example, the use of just the 3 covariates of age group, sex, and year of report leads to more than 900 strata in the MGPS analysis. However, the presence or absence of any concomitant drug could be viewed as a potentially confounding circumstance. If any drug that has a strong adverse event risk is often coprescribed with the drug of interest, some part of the drug-event association due to the concomitant drug will be transferred to the drug of interest, leading to what is sometimes referred to as an *innocent bystander* effect, or *signal leakage*. Since antipsychotics are often coadministered, there might be innocent bystander effects that serve to bias the results against certain drugs. Both suspect and concomitant medications were included in the data-mining analysis. Causality roles assigned by the adverse event reporters were disregarded so that suspect and concomitant medications were treated equally. Both PRR and EBGM are subject to biases caused by the presence of concomitant medication.

The third analysis method, logistic regression (LR), attempts to simultaneously adjust the measure of association for each drug-event combination of interest for the presence in the reports of those other drugs that seem to have the strongest associations with the glycemetic events being studied. Multiple LR is the standard statistical technique for modeling the probability of occurrence of a *response event* as a function of many other variables. In this case, there are 17 different response events, namely the presence or absence, in an AERS report, of each of the 17 glycemetic events described above. Each event required a separate LR estimation. The variables used to predict each event are the age group, sex, and year of report covariates mentioned above, plus the presence/absence of each of the 7 antipsychotic drugs of interest, plus the presence/absence of each of 100 other drugs that seem, according to preliminary analyses, to be most associated with, on average, the 14 MPTs chosen as primary glycemetic reactions. Each regression model required a total of 154 coefficients to estimate: 107 drugs, plus 36 degrees of freedom (*df*) for report year, 8 *df* for age groups, 2 *df* for sex status (male, female, and unknown) and an intercept. Table 2 lists the 100 drugs that were used as covariates in each of the LR models. They were chosen from among the 2206 generic drugs (single ingredients) in the AERS 2004 quarter 1 data that were present in at least 25 reports. The preliminary analysis used approximate methods to estimate LR coefficients for all (14 MPTs) \times (2206 drugs) combinations and then averaged these coefficients across the 14 MPTs to determine the 100 largest average coefficients and thus the 100 drugs to use as covariates. The 100 drugs listed in Table 2 can be thought of as determining a 100-dimensional descriptor of each report to help equalize the comparisons of the seven antipsychotic drugs.

Table 2 List of 100 Drugs Included as Additional Covariates in the LR Model*

| | | | | | |
|----|----------------------|------------------------|--------------------|-------------------------|--------------------|
| 1 | Acebutolol | “Clinical Study” | Glipizide | Methylprednisolone | Somatropin |
| 2 | Ambroxol | Clobetasol | Glyburide | Metolazone | Stavudine |
| 3 | Amiloride | “Corticosteroid” | Gm-Csf Unspecified | Miglitol | “Steroid” |
| 4 | Amisulpride | Cortisone | Indapamide | Milnacipran | Sulfamethoxazole |
| 5 | “Antidiabetic” | Cyclothiazide | Indinavir | Mizoribine | Tacrolimus |
| 6 | “Antihypoglycemic” | Dapsone | Insulin | Moexipril | Teprenone |
| 7 | Asparaginase | Dehydroepiandrosterone | Interferon-Alfa-N3 | Nateglinide | Terbutaline |
| 8 | Atorvastatin | Delavirdine | Interferon-Beta | Nifedipine Slow Release | Thyroid |
| 9 | Betamethasone | Desmopressin | Lamivudine | Nilvadipine | Ticrynafen |
| 10 | Bethanechol | Dexamethasone | Lecithin | Oxaliplatin | Tofisopam |
| 11 | Bezafibrate | Diagnostic Device | Letrozole | Pioglitazone | Tolazamide |
| 12 | Biperiden | Dorzolamide | Leuprolide | Procaterol | Tolbutamide |
| 13 | Blood | Emtricitabine | Levamisole | Procyclidine | Tolcapone |
| 14 | Calcium Lactate | Flu Vaccine | Lithium | Proheparum | Triamcinolone |
| 15 | Calcium Levofolinate | Fludrocortisone | Manidipine | Quinapril | Triamterene |
| 16 | Capecitabine | Flutamide | Megestrol | Reserpine | Trichlormethiazide |
| 17 | Captopril | Garlic | Metformin | Ritodrine | Tropisetron |
| 18 | Chlorpropamide | Gatifloxacin | Methyldigoxin | Salicylate | Troxipide |
| 19 | Chlorthalidone | Gliclazide | Methyclothiazide | Saquinavir | Vinorelbine |
| 20 | Chromium | Glimepiride | Methylcellulose | Sargramostim | Vitamin E |

*The LR model also included the 7 drugs under study (see Table 1), plus variables for age (9 age groups), sex (male, female, unknown), and year of report (37 years, 1968–2004). There are 154 degrees of freedom in the model, including the intercept term.

One advantage of LR is that the estimated coefficients can be interpreted as natural logarithms of odds ratios in a 2 × 2 table relating the predictor and the response. Estimates and confidence intervals for this odds ratio, denoted OR (OR05, OR95) were computed for each of the 7 × 17 combinations of antipsychotic drug and glycemic events being considered. The interpretation of the numerical value of OR is almost the same as the interpretation of PRR and EBGM, keeping in mind that OR has been adjusted for the potential confounding effect of the 100 drugs listed in Table 2, in addition to the age–sex–year covariates. Finally, the LRs were repeated on the two subsets of the data where the age of the patient was known to be either under or over 45 years of age, respectively.

RESULTS

Drug-Event Counts and PRR

Table 1 shows the counts of reported drug-event combinations for the 17 defined glycemic events and the seven antipsychotic drugs being studied. As can be seen from Table 1, the drugs aripiprazole, quetiapine, and ziprasidone have fewer reports than the other four drugs. The addition of the “All Reports” column and row on the border of Table 1 allows the disproportionality measure PRR to be computed. For each drug-event combination, the number in the body of the table is *a*, the number in the corresponding “All Reports” column is (*a* + *b*), the number in the corresponding “All Reports” row is (*a* + *c*), while the total number of reports is given in the cell at the lower right of the table as (*a* + *b* + *c* + *d*) = 2,421,347. Note that the counts of the grouped events are usually somewhat less

than the sum of the counts of the events being grouped, because some reports mention more than one event in a group.

To save space, the computed PRR values are not presented here, but Tables 3 and 4 present results for the 3 events based on grouped MPTs, separately for reports based on younger (age ≤ 45 years) and older (age > 45 years) patients. Reports in which age is unknown are not represented in Tables 3 and 4. Table 3 presents the counts while Table 4 presents the PRR values computed from Table 3. It is apparent from Table 4 that the vast majority of PRR values are greater than the null hypothesis value of 1. However, without confidence intervals or *p* values, it is difficult to know how significant this is. Also, since the PRR values are not adjusted for covariates, biases due to confounding covariates may be partly responsible for the large values.

Olanzapine has the largest values in all six rows of Table 4, especially for the grouped event Diabetic Life-threatening Event, where PRR = 20.66 and 15.06 for the younger and older patients, respectively. Among the other drugs, clozapine and quetiapine seem to have generally higher values than the remaining four drugs, but, again, it is difficult to judge the significance of these differences without confidence intervals. Note that the large value of PRR = 8.67 for aripiprazole (Diabetic Life-threatening Events, age > 45 years), is due to a count of two reports, as can be seen from Table 3. There is quite good correspondence between the PRR values in part (a) of Table 4, and those in part (b). As a rough measure of correspondence between the PRRs for the two age groups, the Pearson correlation coefficient between the 21 pairs of numbers is 81%. If the pair for aripiprazole, Life-Threatening Events, which has very discrepant values based on small counts of 0 and 2, is omitted, the remaining 20 pairs are correlated at 94%. This agreement is

Table 3 Counts of Reported Drug-Event Combinations for Grouped Events, by Reported Age (Reports with Unknown Age Omitted)

| Event | Drug | | | | | | | All Reports |
|---|--------------|-----------|-------------|------------|------------|-------------|-------------|-------------|
| | Aripiprazole | Clozapine | Haloperidol | Olanzapine | Quetiapine | Risperidone | Ziprasidone | |
| (a) Age ≤ 45 years | | | | | | | | |
| Group: Blood Glucose Abnormal | 7 | 248 | 64 | 251 | 43 | 113 | 9 | 6,703 |
| Group: Diabetes Mellitus | 2 | 331 | 49 | 160 | 50 | 98 | 8 | 2,344 |
| Group: Diabetic Life-Threatening Events | 0 | 74 | 22 | 130 | 32 | 33 | 3 | 816 |
| All Reports | 768 | 14,522 | 6,194 | 6,462 | 2,164 | 9,094 | 981 | 710,878 |
| (b) Age > 45 years | | | | | | | | |
| Group: Blood Glucose Abnormal | 6 | 99 | 74 | 181 | 39 | 67 | 3 | 14,131 |
| Group: Diabetes Mellitus | 2 | 142 | 39 | 101 | 20 | 54 | 5 | 4,802 |
| Group: Diabetic Life-Threatening Events | 2 | 34 | 16 | 51 | 14 | 11 | 1 | 686 |
| All Reports | 321 | 8,179 | 6,456 | 5,051 | 1,670 | 6,276 | 421 | 952,152 |

Table 4 PRR Statistics Based on Data in Table 3

| Event | Drug | | | | | | |
|---|--------------|-----------|-------------|------------|------------|-------------|-------------|
| | Aripiprazole | Clozapine | Haloperidol | Olanzapine | Quetiapine | Risperidone | Ziprasidone |
| (a) Age ≤ 45 years | | | | | | | |
| Group: Blood Glucose Abnormal | 0.97 | 1.84 | 1.10 | 4.24 | 2.11 | 1.32 | 0.97 |
| Group: Diabetes Mellitus | 0.79 | 7.88 | 2.43 | 7.99 | 7.14 | 3.37 | 2.48 |
| Group: Diabetic Life-Threatening Events | 0.00 | 4.78 | 3.15 | 20.66 | 13.37 | 3.25 | 2.67 |
| (b) Age > 45 years | | | | | | | |
| Group: Blood Glucose Abnormal | 1.26 | 0.81 | 0.77 | 2.43 | 1.58 | 0.72 | 0.48 |
| Group: Diabetes Mellitus | 1.24 | 3.52 | 1.20 | 4.03 | 2.38 | 1.71 | 2.36 |
| Group: Diabetic Life-Threatening Events | 8.67 | 6.02 | 3.50 | 15.06 | 11.86 | 2.46 | 3.30 |

encouraging, since if prescribing habits tend to differ according to patient age, this might have produced anomalies in the PRR results for some drugs.

MGPS Analysis Results

Figure 1 shows the results of the estimation and confidence interval computation for the EBGM disproportionality measure for the three grouped events. The graphs for the grouped events are easier to interpret because the larger frequencies associated with grouped events result in tighter confidence intervals. The corresponding graphs for the individual events within each group (not shown) all followed approximately the same patterns as the graphs of their respective group in Figure 1, with the greater variation and wider confidence intervals to be expected because of the smaller sample sizes.

Because the counts in the corresponding rows of Table 1 are large, the confidence intervals in these graphs are narrow, with the exception of the recently marketed drugs ziprasidone and aripiprazole. Every graph has a vertical grid line at the value $EBGM = 1$, which represents the “null hypothesis” of no association in the database between the drug-event pair under consideration. Confidence intervals that intersect this grid line are consistent with the null hypothesis. Thirteen of the 21

confidence intervals in the three graphs lie wholly to the right of the line at $EBGM = 1$. In all three graphs, the three drugs quetiapine, clozapine, and olanzapine have the largest values of EBGM and the largest values of EB05. Haloperidol and risperidone have smaller values of these association measures, falling a bit to the left of $EBGM = 1$ for the blood glucose abnormal group, near $EBGM = 1.5$ for the group diabetes mellitus, and near $EBGM = 2.5$ for the group diabetic life-threatening terms. The latter two sets of associations are significantly greater than the null hypothesis value, but they are significantly less than most of the estimates for the drugs quetiapine, clozapine, and olanzapine, because most of the confidence intervals for these latter drugs do not overlap those for haloperidol and risperidone. The newer drugs, ziprasidone and aripiprazole, have generally lower estimates of EBGM and wider confidence intervals than those for the other drugs in all three of the graphs describing the grouped events. The non-overlapping confidence intervals show that these two drugs have significantly lower database associations than all five other drugs for both grouped events diabetes mellitus and diabetic life-threatening terms. Overall, the values of EBGM for these three grouped events tend to be somewhat smaller than the values of PRR in Table 4. The former have a median of 1.6 and a mean of 2.7, while the PRRs have a median of 2.4 and a mean of 4.0. The EBGM values may tend to be smaller

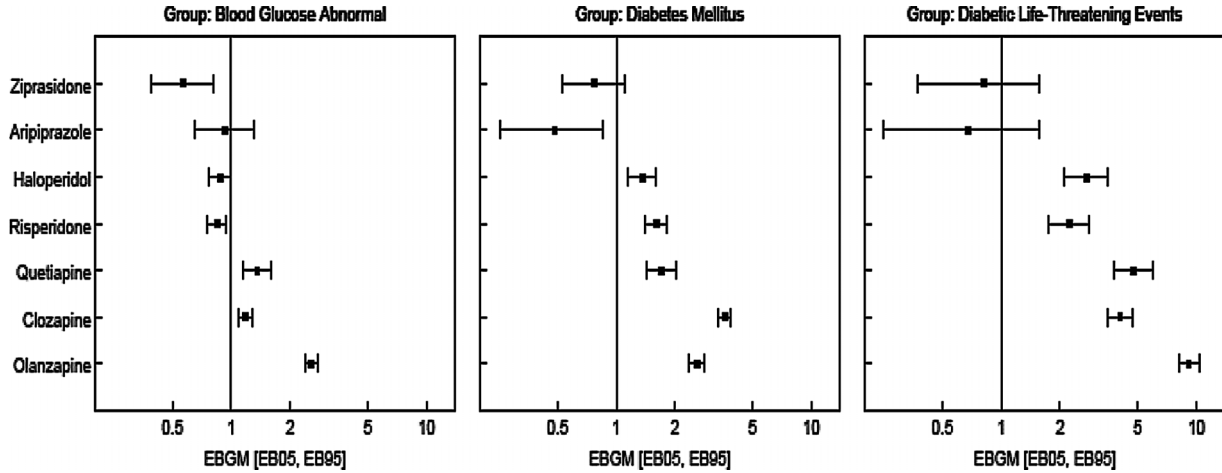


Figure 1 Estimates and confidence intervals for EBGM.

because they are adjusted for the covariates age, sex, and report year, eliminating spurious associations due to these confounding variables. The values of EBGM for the ungrouped MPT, Diabetes mellitus inadequate control, followed a different pattern from all other events studied, in that every EBGM was less than 1, meaning that this MPT occurred somewhat less frequently in AERS in combination with all seven drugs than expected by chance.

LR Results

Figure 2 displays the estimates and confidence intervals for the three grouped events based on LR analysis. The odds ratios (OR) are the exponentials of the coefficients from the estimated model equation. The numeric values of OR have about the same interpretation as PRR or EBGM, namely a factor by which the frequency of the drug-event combination in the

database seems to be multiplied, compared with that expected if there were no association between the drug and the event. The advantage of the LR analysis, in principle, is its ability to adjust the drug-event associations of interest for very many concomitant factors, namely, the presence or absence of the 100 drugs listed in Table 2, plus the age, sex, and report year covariates that the MGPS analysis handled via the stratification strategy.

For the grouped event blood glucose abnormal, the LR estimates are very similar to those of MGPS. The top four drugs in the graph show no increased frequency of the event, unlike the drugs quetiapine, clozapine, and olanzapine. The numerical values of OR are somewhat larger than the values of EBGM for these drug-event combinations. In fact, the graphs for the other two grouped events in Figure 2 show OR > EBGM consistently across all three events for those three drugs, whereas OR tends to be a little closer to 1 than EBGM for the other four drugs. It is possible that

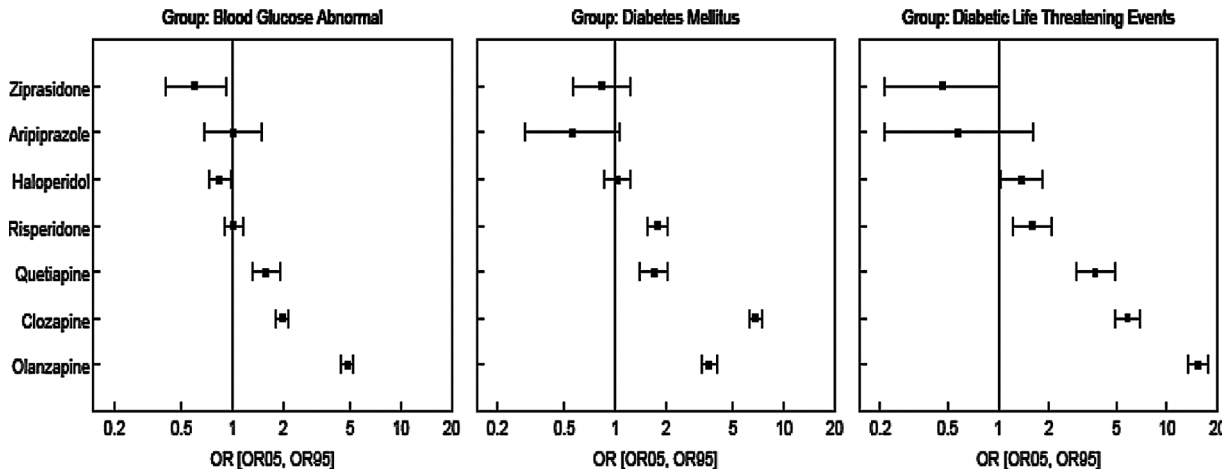


Figure 2 LR estimates and confidence intervals for odds ratios.

these differences are due to different concomitant drug patterns between the last three drugs versus the first four drugs, so that adjustment tends to increase the estimates for some drugs but decrease them for others.

Alternatively, the Bayesian prior distribution that is estimated and used in the MGPS analysis could produce a different pattern of “shrinkage” (modification of estimates toward $EBGM = 1$) than the relatively weak prior distribution involved in the LR analysis. Whatever the reason, the LR analysis tends to estimate a somewhat greater separation in association strength with the three grouped events between quetiapine and particularly clozapine and olanzapine compared with ziprasidone, aripiprazole, haloperidol, and risperidone. The LR estimates for the individual MPT events (not shown to save space) are qualitatively quite similar to the corresponding results for EBGM, and to the LR results for their corresponding groups in Figure 2. Once again, the results for the ungrouped MPT Diabetes mellitus inadequate control showed a pattern of lower association values than any other event studied. The only drug showing a weakly significant LR odds ratio greater than 1 with this event was olanzapine, having a 90% confidence interval of 1.1, 2.0.

Figure 3 shows the result of re-estimating the LR model on subsets of the data defined by age ≤ 45 years and age > 45 years. Reports for which age is missing are excluded from both re-estimations. Within each pair of adjacent confidence intervals in Figure 3, the upper interval with the filled center symbol represents the younger age ≤ 45 years patients, while the lower interval with the open center symbol represents the older age > 45 years patients. As was the case with the PRR analyses presented in Table 4, the two age subsets show very similar patterns of LR associations with the three grouped events and do not suggest substantial effects of age on major analysis findings. In 15 of the 21 drug-grouped event combinations represented in Figure 3 (including all drugs combined with the

Diabetic Life Threatening Event group), the 90% confidence intervals for two age groups overlap. For the two groups Blood Glucose Abnormal and Diabetes Mellitus, there seems to be a consistent and significantly greater disproportionality odds ratio with the two drugs clozapine and olanzapine in the reports of younger patients than older patients.

Summary of MGPS and LR Results

In summary, there are large, consistent, and statistically significant differences in association strength with these glycemic events across these seven antipsychotic agents. The totality of results displayed in Figures 1–3 and of the more detailed results not shown here suggest that association strength for glycemic adverse events, including the life-threatening events, is different for different drugs, with clozapine and olanzapine exhibiting the greatest association strength.

Subset Analyses

Analogous to the re-analysis by age groups presented in Table 4 and Figure 3, separate MGPS analyses by sex (data not shown) did not yield interesting differences in patterns of association strength. Separate MGPS analyses over successively increasing time intervals were also performed (data not shown), and in these analyses the report counts and EBGM values seemed to progress steadily and smoothly over time. Although no formal statistical test to detect a sudden surge in reports for particular drug-event combinations was performed, subjective assessment of graphs of the trends in EBGM over time gave no evidence of “publicity effects” or other artifacts affecting the scores.

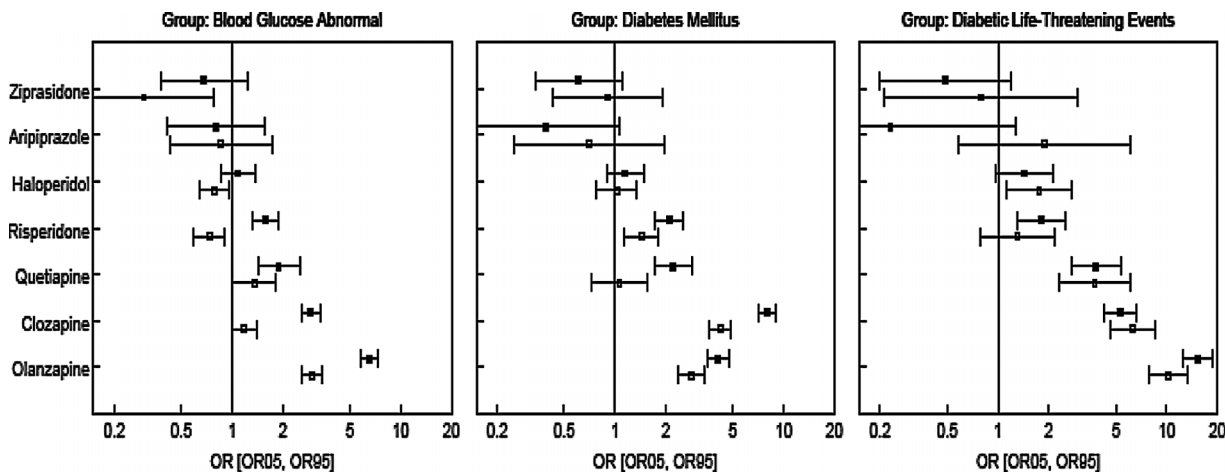


Figure 3 Estimates of odds ratios and confidence intervals from LR, separately for age ≤ 45 years (top graphs and filled squares) and for age > 45 years (bottom graphs and open squares).

DISCUSSION

Study Limitations

The method used in this study for collecting data (spontaneous reporting) provides information only about the relative reporting of adverse events in the postmarket setting. This analysis does not provide estimates of the absolute incidence of adverse events. Although disproportionality methods have been partially validated, in that they often identify labeled drug-event combinations, there has been no validation as of yet that differences in database association strength among drugs are reliable indicators of corresponding differences in metabolic risk. The differences reported here should be viewed as generating rather than confirming hypotheses.

The overall patient exposure to the agents studied varies substantially. Drugs with more exposure tend to contribute more to the database and have more precisely measured associations, as reflected by the varying confidence interval widths in the figures. The two most recently introduced drugs, aripiprazole and ziprasidone, have far fewer reports than the other five drugs, and measurements of their associations with these glycemic events are most tentative. Moreover, if the event of concern develops only after a patient has been on the drug for a longer duration, then there will be less evidence of an association with newer drugs. Another distinction among the drugs is their variation in typical patient populations. The use of clozapine in particular for a systematically different population (refractory schizophrenia) than other more broadly used antipsychotics must also be considered. Other sources of potential bias in these drug comparisons are potential differences in adherence rates among drugs as well as potential differences in prescribing rates by patient race. Note that non-overlapping confidence intervals in the figures presented here imply that observed differences in disproportional reporting ratios are greater than can be expected by random sampling error, but do not allow for differences due to other sources of potential bias.

It is estimated that only 1% to 10% of adverse events experienced are ever reported to the AERS database (8). It is unknown if this rate varies significantly from drug to drug and from event to event. However, the disproportionality measures presented here are not biased by under-reporting alone, provided that the actual reporting rates vary essentially only as a function of the drug or of the event—and not as a function of drug-event combination. This robustness to variable reporting rates has been described (28) for odds ratios, and it applies also to PRR, EBG, and LR results when both the event and the drug of interest show up in a very small proportion of reports (on the order of a few percent or less).

While sex, age, and FDA report year were accounted for in the MGPS analysis, and in addition the concomitant drugs in Table 2 were accounted for in the LR analysis, covariates not taken into account by any of these analyses include basal body mass index, weight gain, dose of antipsychotic, other diabetogenic agents used, and family predisposition. For these and

other reasons, a relatively high reporting rate does not represent proof of a causal relationship between the drug in question and the diabetes-related adverse event.

There have been many reports of diabetes-related adverse events occurring in patients taking antipsychotics. (1–9,11–23) Some of these reports are based on patient records and MEDLINE searches, others on claims and prescription data. In many of the previous analyses, (6–9,15,20–22) patients treated with risperidone experienced fewer diabetes-related adverse events, including life-threatening events such as DKA, or lower incidence of diabetes, or lower risk of diabetes when compared with olanzapine or clozapine. Results of the current study are consistent with the trends shown in these other published reports. Some of these studies (12,24) found that all atypical antipsychotics as a class had a higher risk of diabetes-related adverse events, perhaps because the comparator group was a “general population” rather than patients with major mental illness. In the current study, the comparisons were among atypical and conventional antipsychotics, reducing the likely risks or magnitude of confounding due to indication. Some studies also found no significant differences between the atypical antipsychotics (12). The studies based on claims data gathered from patients with a diagnostic claim for diabetes (20,21) and/or from patients who had prescription claims for antidiabetic medication (12,24).

It is not clear whether the MedWatch system is more or less sensitive than claims data for detecting hyperglycemia or diabetes, but it is clear that each would be less reliable than a prospectively designed epidemiology study. Serious and life-threatening diabetes-related conditions are less likely to be under-reported, but their rarity makes them less likely to appear in a prospective study of even a fairly large population. By using the AERS database and 17 definitions of diabetes-related events, the current study attempts to include more cases of diabetes-related adverse events, particularly life-threatening events, than in previous studies.

In an analysis of 45 case reports of new-onset diabetes mellitus and DKA after initiation of atypical antipsychotics, Jin et al. (6) reported that 20 patients had received clozapine and 19 had received olanzapine. Three cases in each group were related to patients who had taken quetiapine and risperidone, suggesting that clozapine and olanzapine may carry greater risk for the development of new-onset diabetes mellitus or DKA (6). Jin et al.’s results are in keeping with those obtained by Koller et al. (7–9), namely that clozapine and olanzapine appear to have an association with life-threatening diabetes-related events. The results of the current disproportionality analyses support this finding.

In three separate studies, Koller et al. (7–9) evaluated clozapine, olanzapine, and risperidone in similar manners, using the MedWatch drug surveillance system as the data source. Among patients treated with clozapine, there were 384 reports of hyperglycemia, 80 cases of DKA, and 25 deaths (8). Among patients treated with olanzapine, there were 237 reports of hyperglycemia, 80 cases of DKA, and 15 deaths (9). Among

patients treated with risperidone, there were 131 reports of hyperglycemia, six cases of DKA, and four deaths (7). The Koller et al. studies were conducted separately and did not make head-to-head comparisons between the antipsychotics. However, the results of the present study support the trends observed by Koller et al. and allow for comparison across the groups. Moreover, the methods used in this study provide estimates of the relative reporting ratio, both individually and comparatively among drugs.

Hedenmalm et al. (29) searched the World Health Organization (WHO) Collaborating Centre for International Drug Monitoring database and identified reports for clozapine, olanzapine, and risperidone that were associated with diagnoses of glucose intolerance. Strengths of the associations over time between glucose intolerance and the individual drugs were analyzed using the Bayesian Confidence Propagation Neural Network (BCPNN) technique, which is similar in principle to MGPS but uses a fixed "prior distribution." (The MGPS technique uses an empirical Bayesian approach of fitting the prior distribution to the available data.) From 1968 to December 2000, the WHO received 480 reports of glucose intolerance associated with clozapine, 253 with olanzapine, and 138 with risperidone. The authors report that the strengths of association suggest a positive quantitative association between glucose intolerance and all three atypical antipsychotics, but the authors did not attempt to compare the magnitude of association or overlap between individual atypical antipsychotics as was done in the present study. Unlike in the LR analyses of the current study, no adjustment for concomitant drugs was reported (29).

The disproportionality methods employed in the current study provided a surrogate denominator and allowed meaningful quantitative comparisons to be made. Among the three methods employed here, we give greatest credence to the LR analyses because they include adjustments for concomitant drugs and should be most impervious to the signal leakage caused by polytherapy and adjusting for the different patient populations taking each drug. The LR estimation method leads to greater average estimates of reporting ratio than MGPS for quetiapine, clozapine, and olanzapine. The PRR estimates seem larger, on average, than those of the other two methods, perhaps because they are not adjusted for patient covariates or concomitant drugs. It is noteworthy that the three methods (PRR, MGPS, and LR) all showed roughly the same pattern across the 17 different events and 7 drugs. As a coarse summary, quetiapine, clozapine, and olanzapine show significantly greater associations (nonoverlapping confidence intervals) than the other 4 drugs with the grouped events ("Blood glucose abnormal," "Diabetes mellitus," and "Diabetic life-threatening events").

Another potential bias in a database disproportionality analysis is *dilution* of a signal for one drug-event combination caused by a great number of reports of another event with the same drug. For example, some of these drugs have many reports of movement disorders, resulting in a lowered proportion of diabetes-related events. Simple arithmetic shows that

dilution cannot reduce a disproportionality ratio by a factor of 2 unless the other excessively reported events show up in half or more of the reports. Even then, comparisons among drugs would only be affected if this excess varied significantly between the drugs being compared. The most common MedDRA System Organ Class (SOC) among these drugs is the Nervous System SOC, which contains the preferred terms involving movement disorders. For the four most frequent drugs in our study, the percentage of reports containing any MPT within the Nervous System SOC are: clozapine—26%, haloperidol—40%, olanzapine—31%, and risperidone—34%. These differentials are not great enough to affect the ordering of the glycemic reporting ratios among the drugs.

CONCLUSIONS

The results of the current study raise serious doubt as to whether diabetes and related adverse events are a class effect of all atypical antipsychotics, and suggest a set of drug-specific relationships and patterns that require explanation. Although these patterns could be merely artifacts of the spontaneous reporting data collection process, the strength and statistical significance of the patterns, and the consistency across glycemic events and methods of analysis, are suggestive of important differences in atypical antipsychotic risk profiles with regard to diabetes-related adverse events, particularly life-threatening adverse events. Only three of the drugs, clozapine, olanzapine, and, to a lesser extent, quetiapine, have consistently and significantly larger reporting ratios than the conventional antipsychotic haloperidol. This supports the conclusions of the American Diabetes Association/American Psychiatric Association expert consensus report, which suggested that risk for diabetes-related adverse events may differ substantially among different antipsychotic drugs. The differential reporting ratios found in this study should inform hypothesis generation for the future studies required to fully understand this issue.

ACKNOWLEDGMENTS

Supported by funding from Ortho-McNeil Janssen Scientific Affairs, L.L.C. A preliminary version of this work was presented at the 2003 Annual Meeting of the International College of Geriatric Psychoneuropharmacology, San Juan, Puerto Rico, December 12–14, 2003.

REFERENCES

1. Dixon L, Weiden P, Delahanty J, Goldberg R, Postrado L, Lucksted A, Lehman A. Prevalence and correlates of diabetes in national schizophrenia samples. *Schizophr Bull.* 2000;26:903–912.
2. Liebbeit KA, Markowitz JS, Caley CF. New onset diabetes and atypical antipsychotics. *Eur Neuropsychopharmacol.* 2001; 11:25–32.

3. Lindenmayer JP, Nathan AM, Smith RC. Hyperglycemia associated with the use of atypical antipsychotics. *J Clin Psychiatry*. 2001;62(suppl 23):30–38.
4. Mir S, Taylor D. Atypical antipsychotics and hyperglycaemia. *Int Clin Psychopharmacol*. 2001;16:63–73.
5. Wirshing DA, Spellberg BJ, Erhart SM, Marder SR, Wirshing WC. Novel antipsychotics and new onset diabetes. *Biol Psychiatry*. 1998;44:778–783.
6. Jin H, Meyer JM, Jeste DV. Phenomenology of and risk factors for new-onset diabetes mellitus and diabetic ketoacidosis associated with atypical antipsychotics: an analysis of 45 published cases. *Ann Clin Psychiatry*. 2002;14:59–64.
7. Koller EA, Cross JT, Doraiswamy PM, Schneider BS. Risperidone-associated diabetes mellitus: a pharmacovigilance study. *Pharmacotherapy*. 2003;23:735–744.
8. Koller E, Schneider B, Bennett K, Dubitsky G. Clozapine-associated diabetes. *Am J Med*. 2001;111:716–723.
9. Koller EA, Doraiswamy PM. Olanzapine-associated diabetes mellitus. *Pharmacotherapy*. 2002;22:841–852.
10. Koller E, Malozowski S, Doraiswamy PM. Atypical antipsychotic drugs and hyperglycemia in adolescents. *JAMA*. 2001;286:2547–2548.
11. Allison DB, Mentore JL, Heo M, Chandler LP, Cappelleri JC, Infante MC, Weiden PJ. Antipsychotic-induced weight gain: a comprehensive research synthesis. *Am J Psychiatry*. 1999;156:1686–1696.
12. Buse JB, Cavazzoni P, Hornbuckle K, Hutchins D, Breier A, Jovanovic L. A retrospective cohort study of diabetes mellitus and antipsychotic treatment in the United States. *J Clin Epidemiol*. 2003;56:164–170.
13. Citrome LL, Jaffe AB. Relationship of atypical antipsychotics with development of diabetes mellitus. *Ann Pharmacother*. 2003;37:1849–1857.
14. Fuller MA, Shermock KM, Secic M, Grogg AL. Comparative study of the development of diabetes mellitus in patients taking risperidone and olanzapine. *Pharmacotherapy*. 2003;23:1037–1043.
15. Gianfrancesco F, White R, Wang RH, Nasrallah HA. Antipsychotic-induced type 2 diabetes: evidence from a large health plan database. *J Clin Psychopharmacol*. 2003;23:328–335.
16. Henderson DC, Cagliero E, Gray C, Nasrallah RA, Hayden DL, Schoenfeld DA, Goff DC. Clozapine, diabetes mellitus, weight gain, and lipid abnormalities: a five-year naturalistic study. *Am J Psychiatry*. 2000;157:975–981.
17. Meyer JM. Effects of atypical antipsychotics on weight and serum lipid levels. *J Clin Psychiatry*. 2001;62(suppl 27):27–34.
18. Wirshing DA, Pierre JM, Eyeler J, Weinbach J, Wirshing WC. Risperidone-associated new-onset diabetes. *Biol Psychiatry*. 2001;50:148–149.
19. Caro JJ, Ward A, Levinton C, Robinson K. The risk of diabetes during olanzapine use compared with risperidone use: a retrospective database analysis. *J Clin Psychiatry*. 2002;63:1135–1139.
20. Gianfrancesco F, Grogg A, Mahmoud R, Wang RH, Meletiche D. Differential effects of antipsychotic agents on the risk of development of type 2 diabetes mellitus in patients with mood disorders. *Clin Ther*. 2003;25:1150–1171.
21. Gianfrancesco FD, Grogg AL, Mahmoud RA, Wang RH, Nasrallah HA. Differential effects of risperidone, olanzapine, clozapine, and conventional antipsychotics on type 2 diabetes: findings from a large health plan database. *J Clin Psychiatry*. 2002;63:920–930.
22. Koro CE, Fedder DO, L'Italien GJ, Weiss SS, Magder LS, Kreyenbuhl J, Revicki DA, Buchanan RW. Assessment of independent effect of olanzapine and risperidone on risk of diabetes among patients with schizophrenia: population based nested case-control study. *BMJ*. 2002;325:243–247.
23. Sernyak MJ, Leslie DL, Alarcon RD, Losonczy MF, Rosenheck R. Association of diabetes mellitus with use of atypical neuroleptics in the treatment of schizophrenia. *Am J Psychiatry*. 2002;159:561–566.
24. Feldman PD, Hay LK, Deberdt W, Kennedy JS, Hutchins DS, Hay DP, Hardy TA, Hoffmann VP, Hornbuckle K, Breier A. Retrospective cohort study of diabetes mellitus and antipsychotic treatment in a geriatric population in the United States. *J Am Med Dir Assoc*. 2004;5:38–46.
25. FDA Quarterly Data Extract from the Adverse Event Reporting System (AERS) [database on CD-ROM]. Washington: Food and Drug Administration; 2002.
26. Fram DM, Almenoff JS, DuMouchel W. Empirical Bayesian data mining for discovering patterns in post-marketing drug safety. Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press;2003:359–368.
27. DuMouchel W, Pregibon D. Empirical Bayes screening for multi-item associations. Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press;2001:67–76.
28. van der Heijden PG, van Puijenbroek EP, van Buuren S, van der Hofstede JW. On the assessment of adverse drug reactions from spontaneous reporting systems: the influence of under-reporting on odds ratios. *Stat Med*. 2002;21:2027–2044.
29. Hedenmalm K, Hagg S, Stahl M, Mortimer O, Spigset O. Glucose intolerance with atypical antipsychotics. *Drug Saf*. 2002;25:1107–1116.

