© 2011 Adis Data Information BV. All rights reserved.

# Strategies for Quantifying the Relationship between Medications and Suicidal Behaviour What has been Learned?

*Robert D. Gibbons*<sup>1</sup> and *J. John Mann*<sup>2,3</sup>

- 1 Departments of Medicine and Health Studies, and the Center for Health Statistics, University of Chicago, Chicago, Illinois, USA
- 2 Department of Molecular Imaging and Neuropathology, New York State Psychiatric Institute, New York, New York, USA
- 3 Department of Psychiatry, Columbia University College of Physicians & Surgeons, New York, New York, USA

# Abstract

In recent years there has been considerable concern that certain classes of drugs, for example antidepressants, may increase the risk of suicide. In this current opinion article, we examine the literature on methodological and statistical approaches to the design and analysis of suicidal event studies. Experimental, ecological and observational studies of the relationship between drugs and suicidal events (thoughts, attempts and completion) are discussed. Areas considered include analysis of spontaneous reporting system data, ecological trends in national and/or small area (e.g. county) suicide rates, meta-analyses of randomized clinical trials, and large-scale medical claims data. New statistical and experimental strategies for investigating possible associations between drugs and suicide are highlighted, and we suggest directions for future statistical/methodological research. To put this into context, we then review the most recent literature on the relationship between drugs (antidepressants, anti-epileptics, varenicline, montelukast and antipsychotics) and suicidal events.

Overall, there appears to be little evidence that drugs increase the risk of suicide and related behaviour. Numerous lines of evidence in adults clearly demonstrate that inadequate treatment of depression (pharmacotherapy and/ or psychotherapy) is associated with increased risk of suicidal behaviour. In children, the results are less clear and further study is required to better de-lineate which children benefit from treatment and who may be at increased risk as a consequence of treatment. From a statistical and methodological perspective, the field of pharmacoepidemiology is a fertile area for statistical research, both in theory and in application. In general, methods have been adopted from other areas such as general epidemiology, despite the singular nature of many of the problems that are unique to drug safety in general, in particular the study of rare events. Finally, there is considerable debate concerning the communication of risk. For suicide, regulatory action has been taken largely on the basis of evidence suggesting increased risk of suicidal

thoughts. However, suicidal thoughts are quite common, particularly among patients with depression, and may have little relationship to suicidal behaviour and/or completion.

The enormous human cost of suicide makes research and prevention of suicide a national priority.<sup>[1]</sup> Worldwide, there are about 1 million suicides annually. In the last 25 years, approximately 750 000 people committed suicide in the US, and suicides outnumber homicides by at least a 3:2 ratio. Deaths from suicide exceeded deaths from AIDS by 200 000 in the past 20 years, and four times as many Americans died as a result of suicide than in the Vietnam war during the same time period.<sup>[1]</sup> The estimated cost to the US in lost income in 1998 alone was \$U\$11.8 billion. Nonetheless, suicide is a rare event with an annual rate in the US of 12/100000, making it a difficult phenomenon to study using conventional approaches.<sup>[1,2]</sup> The study of suicide spans the concepts of psychobiology and genetics (and other aetiological factors), to the epidemiology and pharmacoepidemiology of suicide, to prevention, treatment, prognostic studies, and ultimately methodological, measurement and statistical issues. In this current opinion article, we focus largely on these latter issues related to the design and analysis of studies of suicidal thoughts, behaviour and completion. We cover a variety of approaches that are in use to study suicide-related events both from experimental and statistical perspectives. We then review the applications of these methods to the study of suicidal thoughts and behaviours (including completed suicide) and provide a general summary of the literature on the relationship between suicidal events and drugs of potential concern. In the following section, we begin with a review of the early literature on the possible relationship between antidepressants and suicide to place the discussion in the proper historical context.

#### 1. Overview of Earlier Literature

Questions regarding a possible relationship between antidepressants and suicide emerged in 1990 with the publication of a series of case reports in which the then newly introduced selective serotonin reuptake inhibitors (SSRIs) were associated with the apparent emergence of suicidal thoughts and behaviour.<sup>[3]</sup> These early observations led to US FDA hearings in 1991 that did not find evidence of an increased risk of suicidal acts associated with antidepressants. These early case studies also set the stage for the development of new approaches to the analysis of pharmacovigilance data in general and with respect to suicide in particular. To provide a foundation for more recent developments, we now review this earlier methodological work.

One early approach to the study of suicide involved large-scale naturalistic studies of at-risk populations such as those with mood disorders. Lithium treatment of bipolar illness appears to reduce suicide rates,<sup>[4-9]</sup> although results vary across studies.<sup>[10-12]</sup> Another broad ecological approach involves analyses of national population datasets. The observed decrease in suicide rates over time correlates with increased antidepressant use in Europe,<sup>[13-17]</sup> Scandinavia,<sup>[18]</sup> the US<sup>[19]</sup> and Australia.<sup>[20]</sup> Doubling of prescriptions for SSRIs correlated with a 25% decrease in the suicide rate in the same time period in Sweden.<sup>[18]</sup> A study in Italy found a 36% rise in prescription rates to correlate with an 18% decline in suicide rates in females only.<sup>[21]</sup> An educational intervention study in a province in Sweden targeting primary care physicians' recognition and treatment of depression had a disproportionate benefit for female suicide rates because of better help-seeking behaviour in women.<sup>[22]</sup> In another study, antidepressant prescription rates rose faster and the fall in suicide rates was greater in women in the US during the period 1985-99.<sup>[23]</sup> Suicide and suicide attempts in depressed patients are associated with no treatment, or inadequate prescription or consumption of antidepressants.<sup>[24-28]</sup> Population-based changes in suicide rates may

have explanations other than increased use of antidepressant medications.<sup>[1]</sup> Unadjusted suicide rates have risen in Japan in the last 4 years and so too has use of SSRIs.<sup>[29]</sup> However, Nakagawa and colleagues<sup>[30]</sup> showed that this was an artefact of Simpson's paradox and that, within age cohorts, suicide rates in Japan have been decreasing proportionately with increases in antidepressant prescriptions. Iceland has had no meaningful change in suicide rates despite a 4-fold increase in antidepressant prescriptions.<sup>[31]</sup>

Small area estimation offers improvements over large-scale ecological modelling. Gibbons et al.<sup>[32]</sup> obtained Centers for Disease Control and Prevention (CDC) small or local municipality area level data on suicide rates for 1996–8 in the US and corresponding antidepressant prescription rates. Increases in SSRI and serotonin-noradrenaline (norepinephrine) reuptake inhibitor (SNRI) prescriptions were associated with decreases in suicide rates both between counties and within counties over time, whereas tricyclic antidepressants (TCAs) were associated with higher suicide rates. These results have been replicated in children and young adolescents aged 5–14 years.<sup>[19,33]</sup>

In 2003–4, the relationship between antidepressant use and suicidal behaviour in children and adolescents suddenly became of concern as regulatory agencies in the UK, Europe and the US observed that the rate of suicidal ideation and perhaps nonfatal attempts appeared to be greater in the active treatment group compared with placebo in some randomized clinical trials (RCTs) of antidepressants in the paediatric population<sup>[34]</sup> (also see section 4.1). Suicide is the third leading cause of death in younger adolescents in the US (aged 10–14 years) and the leading cause of death in other countries, including China, Sweden, Ireland, Australia and New Zealand. Over 90% of older youth suicides (aged 15-18 years) in the US are associated with psychiatric illness, with rates of approximately 60% for 10- to 14-yearolds.<sup>[1,35,36]</sup> However, only approximately 2% of youths committing suicide appear to be on medication at the time of suicide.[37,38] In a study of 49 adolescents committing suicide in Utah, 24% had been prescribed antidepressants, but none tested positive for SSRIs at the time of their death.<sup>[39]</sup>

In a postmortem study of 66 youth suicides in New York City from 1993 through 1998,<sup>[37]</sup> 54 (81.8%) had serum toxicological analysis for antidepressants and an injury-death interval  $\leq 3$  days. Two cases had imipramine detected and another two had fluoxetine (total 10%). None of the other 90% of cases had antidepressants detected.

Each year in the US, depression occurs in 3-5% of youth and 60% of suicides occur in youth and adults with depression. Although a suicide death is rare in younger children (under 1/100 000 per year in 5- to 14-year-olds<sup>[1]</sup>), it is more common after mid-adolescence where rates for 15- to 19-year-olds are 3/100 000 for girls and 15/100 000 for boys.<sup>[40]</sup> Suicidal thinking and attempts are relatively common in youth. Yearly, 19% of teenagers (aged 15-19 years) in the general population report suicidal ideation and 9% make a suicide attempt.<sup>[41]</sup> Rates are even higher in depressed children, where 35-50% attempt suicide when followed through adulthood,<sup>[42-44]</sup> and between 2% and 8% of depressed children commit suicide over the course of a decade.<sup>[42,43,45]</sup> These differences in rates between suicidal ideation, behaviour (i.e. attempts) and completion highlight the inherent difficulty in studying suicide in general and the limitation of inferring risk of suicide data from surrogates such as suicidal ideation.

The observation that almost all youths committing suicide are not receiving antidepressants at the time of death, even when the individual is thought to be depressed, suggests that lack of treatment and/or low treatment adherence contributes to suicide risk, and that more widespread antidepressant treatment may reduce suicide rates. Conversely, concerns have been raised about the effect of antidepressants on risk of suicide in mood disorders, especially in youth. Antidepressants appear to be associated with rates of adverse event reports regarding suicide attempts or ideation that are 2% higher than placebo.<sup>[34]</sup> Because suicidal youth are routinely excluded from antidepressant trials, there is a paucity of RCTs evaluating the safety and efficacy of any antidepressant medication in depressed youth at risk for suicide because of a history of suicidal behaviour. In youth, demonstrated efficacy of antidepressants historically has been limited to fluoxetine,

which is long-acting,<sup>[46-49]</sup> although recent metaanalyses have demonstrated efficacy of newer antidepressants taken as a whole,<sup>[50]</sup> and the FDA recently approved citalopram and escitalopram for the treatment of major depression in youth.

Given the small number of youth who have been enrolled in randomized trials of antidepressants, the rarity of suicide itself, the exclusion of those actively suicidal from such trials<sup>[2]</sup> and the existence of reporting bias that is likely to undercount suicidality in the placebo condition,<sup>[51]</sup> there are very few RCT data that can be used to examine the impact on suicidal behaviour. Alternative approaches are needed. Valuck et al.<sup>[52]</sup> examined data from 24 4119 adolescents with a first diagnosis of major depression and at least 6 months of follow-up data. Treatment with SSRIs, other antidepressants or combinations of antidepressants resulted in no increased risk in suicide attempts.

Simon et al.,<sup>[53]</sup> studied both depressed children and adults and found that suicide attempt rates are highest in the month prior to initiating treatment. Olfson et al.[54] conducted a case-control study of antidepressant treatment in depressed children and adults and found no association with suicide or suicide attempts in adults, but a significant association for both suicide attempts (n=263) and completion (n=8) in children. The association with suicide attempts was not significant for SSRIs, and with only eight completed suicides, results are difficult to interpret. Furthermore, it may be that more severely ill patients were treated with antidepressants; hence, the appearance of an association. Tiihonen et al.[55] found current antidepressant use to be associated with increased risk of suicide attempts but decreased risk of suicide completion in the subgroup of 10- to 19-year-old patients who had ever used an antidepressant.

At a population level, a decline in the suicide rate amongst youth (ages 15–24 years), averaging about 33%, has occurred across 15 countries over the past 14 years.<sup>[56]</sup> The reduction followed three decades of increases. The greatest reductions were in Australia (52%) and Switzerland (50%), and the lowest (14%) was in Japan. In 12 of the 15 countries, the start of the decline occurred after the

introduction of SSRIs, suggesting a relationship between prescription of SSRIs and suicide rates, although other major youth suicide prevention programmes were being implemented as well.<sup>[56]</sup> The decline in the youth suicide rate observed in epidemiological studies cannot be explained by a reduction in exposure to illegal drugs and alcohol<sup>[23]</sup> or better firearms control.<sup>[57]</sup> Nevertheless, it is important for future research to better delineate the effects of change in SSRI use from other social factors that may also influence suicide rates.

Antidepressants are not the only class of drugs that have been associated with suicidal events. On 31 January 2008, the US FDA issued an alert regarding increased risk of suicidal thoughts and behaviour with antiepileptic drugs (AEDs). On 10 July 2008, an FDA scientific advisory committee voted 'yes' that there was a significant positive association between AEDs and suicidality but voted against placing a black-box warning on AEDs for suicidality. Other drugs such as the smoking cessation drug varenicline, the acne medicine accutane and the allergy medicine montelukast have all been linked to increases in numbers of spontaneous reports of suicidality. In the following sections we provide an overview of some of the different possible methodological approaches, explore their strengths and weaknesses, and then attempt to summarize the more recent literature on the relationship between pharmacological agents and suicide.

#### 2. Literature Search Strategy

We used Google Scholar and PubMed to conduct a search (through 2009) for statistical analysis methods and suicidality (thoughts, attempts and completions) induced by drugs of interest. We searched using the phrases 'statistical analysis' (with all of the words) and 'spontaneous reporting system' (exact phrase) in Engineering, Computer Science and Mathematics collections, which resulted in identification of 78 articles. We used the references in these articles to identify additional articles. Additional searches using the keywords 'pharmacoepidemiology' and 'statistical analysis' combined were also conducted, which identified an additional 151 articles. A general

search using the keywords 'statistical analysis' and 'suicide' combined identified 2220 articles that were reviewed for relevant content. In addition, we reviewed all articles published in the journals Drug Safety and Pharmacoepidemiology and Drug Safety for suicidality (regardless of drug) and found 71 and 27 articles, respectively. In order to find suicidality associated with drugs (search result) we used Google Scholar with the following keywords in the title of the extracted papers: 'suicide' combined with 'antidepressant' (145 articles), 'antiepileptic' (12) 'bipolar' (112) and 'antipsychotics' (10). For montelukast and varenicline we searched for keywords anywhere in the articles and found 200 and 192 articles, respectively. We also used a PubMed search with the keywords 'suicide' and 'adverse events' combined with 'antidepressant' (35) 'antiepileptic' (8) 'bipolar' (9) 'montelukast' (90) 'varenicline' (1) and 'antipsychotics' (8). The overlapping findings were omitted.

### Experimental Design and Statistical Methods for Suicidal Event Data

In this section, we review the various types of data, designs and statistical methods that are useful for pharmacoepidemiological studies in general, and studies of suicide and drugs in particular. A summary of these methodologies and their advantages and disadvantages is presented in table I.

#### 3.1 Spontaneous Reporting Systems

Most reports of adverse drug reactions (ADRs) are the product of clinician observation or personal self-report. The WHO global individual case safety report database, VigiBase, is the largest database of case reports of spontaneous reports of ADRs. In the US, the FDA introduced the MedWatch programme in June 1993 to expand reporting of suspected ADRs in the Adverse Event Reporting System (AERS), which now contains over 2 million reports of suspected ADRs. Similar databases exist in various European and other countries, including India, China, Taiwan and Iran. ADR reports from these databases are generally also included in VigiBase.

Proportional reporting ratio (PRR) is the simplest method available for signal detection. It is the ratio of the number of reports of a specific adverse event (AE) to all AEs for a particular drug compared with the same ratio for a large set of other drugs – often all other drugs in the dataset. Of concern is that large numbers of AE reports of a particular kind effectively inflate the denominator for that drug and thereby reduce sensitivity for detecting other signals associated with that drug. PRRs have large numbers of false positive signals because they provide no adjustment for multiple comparisons.

Empirical Bayes screening<sup>[58,59]</sup> computes the baseline (expected) frequency under a row (drug) and column (event) independence assumption for multiple two-way tables. If the drug and event are independent, the proportional representation of that event for a specified drug should be the same as the proportional representation of that event in the entire database.

The cumulative sum (CUSUM) method is based on the cumulative sum of differences between observations and their expected values. A signal is detected if the signal statistics exceed a threshold value.<sup>[60]</sup> The threshold is determined by average run length (ARL) based on the mean and variance of the background incidence. The method requires a background comparison time interval and may therefore limit the timely identification of safety problems.

Random-effect Poisson regression<sup>[61]</sup> is a denominator-based method that requires some estimate of the population at risk (e.g. national prescription rates). Parameters are estimated using marginal maximum likelihood, and individual ADR rate ratios are estimated using either empirical Bayes or parametric or nonparametric full Bayes methods. Confidence (posterior) intervals are used to identify safety signals.

There are numerous limitations of spontaneous reporting system (SRS) data. These include (i) confounding by indication (i.e. patients taking a drug may have a disease that is itself associated with a higher incidence of the AEs); (ii) underreporting; (iii) questionable representativeness

Data acurotia       Experimental strategies       Statistical methodiogies       Stratistical methodiogies       Stratistical methodiogies       Methodes       Confounding ty indication         Specimenous reports       Prasive reporting system - VA       Bayesian neural networks       Experimental strategies       Systematic underreporting         Active reporting system - VA       Bayesian neural networks       Experimental strategies       Confounding ty indication         Active reporting system - VA       Bayesian neural networks       Experimental strategies       Systematic underreporting         Active reporting system - VA       Bayesian neural networks       Experimental strategies       Confounding ty indication         Ecological methods       Natural experiments       Bayesian neural networks       Experimental strategies         Ecological methods       Natural experiments       Experimental strategies       Confounding ty indication         Ecological methods       Natural experiments       Experimental strategies       Experimental strategies         Ecological methods       Natural experiments       Experimental strategies       Experimental strategies         Ecological methods       Natural experiments       Experimental strategies       Experimental strategies         Smail anea estimation       Experimental strategi	Iable I. Summary of av				
Sportaneous reports       Pressive raporting system - FDA       Proportional reporting remounds in measurer and weres       Conducting by indication measurer and weres         Active reporting system - VA       Bayes and measurer and weres       Systematic underreporting weres         Active reporting system - VA       Bayes and measurer and weres       Systematic underreporting weres         Active reporting system - VA       Bayes and measurer and weres       Developmental ergenerativements         Multi-Fine Reporting system - VA       Bayes and weres       Developmental ergenerativements         Multi-Fine Reporting system       Cumulative sum       Developmental ergenerativements         Multi-Fine Reporting reportend reportend reporting reporting reporting reportend reporting	Data source	Experimental strategies	Statistical methodologies	Strengths	Weaknesses
Adrise reporting system - VA       Bayesian neural networks       Systematic underreporting constructions       Systematic underreporting multi-ond mu	Spontaneous reports	Passive reporting system – FDA	Proportional reporting ratio	Large enough numbers to measure rare adverse events	Confounding by indication
Expirated Multi-term Garme Poisson   Cuantable serienting   Cuantable representativene     Multi-term Garme Poisson   Publicity bias     Strinker   Cumulative sun     Cumulative sun   Cumulative sun     Cumulative sun   Cumulative sun     Ecological methods   Natoral rates     Natoral rates   Time series methods     Natural experiments   Change-point analyses     Natural experiments   Change-point analyses     Natural experiments   Change-point analyses     Small area estimation   Mised-effects Poisson regression     Wete-analysis   Synthesis of PCI's     Freet-effects Poisson regression   Freet-effects Poisson regression     Modecal claims   Case-control studies     Modecal claims   Case-control studies     Modecal claims   Case-control studies     Modecal claims   Person-level     Contor studies   Proce-effects logistic regression     Modecal claims   Case-control studies     Modecal claims   Person-level     Control studies   Proce-effects logistic regression     Modecal claims   Case-control studies     Modecal claims   Case-control studies     Modecal claims   Person-level     Case-control studies   Proce-effects logistic regression <td></td> <td>Active reporting system – VA</td> <td>Bayesian neural networks</td> <td></td> <td>Systematic underreporting</td>		Active reporting system – VA	Bayesian neural networks		Systematic underreporting
Multi term Gamma Polsson   Publicity bias     Ecological methods   National rates   Extreme duplication     Ecological methods   National rates   Time series methods   Large samples or entire   Do not know if person experiants     Ecological methods   Natural experiments   Time series methods   Large samples or entire   Do not know if person experiants     Rutural experiments   Time series methods   Time series methods   Large samples or entire   Do not know if person experiants     Multural experiments   Change-point analyses   Extreme duplication   Extreme duplication     Multural experiments   Change-point analyses   Entire benean-stratum   Subject to ecological failes/     Multi area estimation   Multural experiments   Change-point analyses   Extreme duplication     Multi area estimation   Change-point analyses   Entire benean-stratum   Subject to ecological failes/     Multi area estimation   Change-faileds   Preson-level   Exclusion of zero-event studies     Multi area estimation   Suprimetion   Exclusion of zero-event studies   Prencipation     Multi area estimation   Suprimetion   Large samples   Confounding by indication     Multi area estimation   Case-control studies   Multi encipation   Prencipation     Multi area estimation   Case-control studi			Empirical Bayes screening		Questionable representativeness
Ecological methods   Nation-reflects Poisson   Extreme duplication     Ecological methods   National rates   Time series methods   Time series methods   Do not know if person experience     Ecological methods   Natural experiments   Charge-point analyses   Large samples or entire   Do not know if person experience     Meta-analysis   Synthesis of RCTs   Charge-point analyses   Permits between-stratum   Subject to ecological tallexy     Meta-analysis   Synthesis of RCTs   Fixed-effects Poisson regression   Hypothesis generation   Geographic variability in raport     Meta-analysis   Synthesis of RCTs   Fixed-effects logistic regression   Hypothesis generation   Geographic variability in raport     Metcalidies   Synthesis of RCTs   Fixed-effects logistic regression   Hypothesis generation   Zero-event studies     Metcalidies   Meta-effects logistic regression   Hypothesis generation   Lange samples   Conforming by indication     Methol   Read-fiftical claims   Control studies   Mutiliter of models   Lange samples   Control of biss     Methol   Case-control studies   Mutiliter of models   Person-level   Control of biss   Ascertament biss     Methol   Control of studies   Mutiliter of models   Control of dignoess   Control of opticatidin     Methol   Pers			Multi-item Gamma Poisson shrinker		Publicity bias
Ecological methods   National rates   Time series methods   Large samples or entire greation   Do not know if person experian gequations can be populations can be the AE actually took the drug studied     Kological methods   Natural experiments   Change-point analyses   Pointis between-stratum   Subject to ecological fallexy compatisons     Meta-analysis   Synthesis of RCTs   Fixed-effects Poisson regression   Hypothesis generation   Geographic variability in report studied     Meta-analysis   Synthesis of RCTs   Fixed-effects nodels (MH model)   Randomization   Geographic variability in report studied     Meta-analysis   Synthesis of RCTs   Fixed-effects logistic regression   Hypothesis generation   Geographic variability in report studies     Meta-analysis   Synthesis of CTs   Fixed-effects logistic regression   Hypothesis generation   Geographic variability in report studies     Meta-analysis   Synthesis of CTs   Random-effect models   Person-level   Ecological fallexy constration     Meta-analysis   Synthesis of CTs   Random-effect models   Person-level   Limited generalizability in report     Meta-analysis   Synthesis of CTs   Randomization   Limited generalizability in report   Limited generalizability in report     Meta-analysis   Synthesis of CTs   Mixed-effects logistic regression   Limited generalizability in report     Meta-ana			Cumulative sum		Extreme duplication
Ecological methods       National rates       Time series methods populations can be studied       Do not know if person experian populations can be not accounted (DL model)       Pandomization       Do not know if person experian populations can be not acco-event studies         Meta analysis       Synthesis of RCTs       Fixed-effects logistic regression       Hypothesis generation of zero-event studies         Meta analysis       Synthesis of observational studies       Randomization       Hererogenety Hererogenety Multilevel mixture models       Hererogenety Hererogenety Multilevel mixture models         Medical claims       Case-control studies       Fixed-effects logistic and Poisson       Lange samples       Confounding by indication Protonal studies         Medical claims       Case-control studies       Fixed-effects logistic and Poisson       Lange samples       Confounding by indication Protonal studies <td></td> <td></td> <td>Random-effects Poisson regression</td> <td></td> <td>Unknown population at risk</td>			Random-effects Poisson regression		Unknown population at risk
Natural experiments   Change-point analyses   Permits between-stratum   Subject to ecological fallacy comparisons     Meta-analysis   Small area estimation   Mixed-effects Poisson regression   Hypothesis generation   Geographic variability interpoint     Meta-analysis   Synthesis of RCTs   Fixed-effects Poisson regression   Hypothesis generation   Geographic variability interpoint     Meta-analysis   Synthesis of RCTs   Fixed-effects Poisson regression   Hypothesis generation   Geographic variability interpoint     Meta-analysis   Synthesis of RCTs   Fixed-effects logistic regression   Handomization   Limited generalizability erson-level     Meta-analysis   Synthesis of RCTs   Randomization   Person-level   Limited generalizability erson-level     Mitelecter   Mixed-effects logistic regression   Mixed-effects logistic regression   Large samples     Medical claims   Case-control studies   Mixed-effects logistic and Poisson   Large samples     Medical claims   Case-control studies   Mixed-effects logistic and Poisson   Confounding by indication     Metion   Conto studies   Mixed-effects logistic and Poisson   Large samples   Confounding by indication     Metion   Mixed-effects logistic and Poisson   Large samples   Confounding by indication   Secretariment bias     Metion   Mitheole <td< td=""><td>Ecological methods</td><td>National rates</td><td>Time series methods</td><td>Large samples or entire populations can be studied</td><td>Do not know if person experiencing the AE actually took the drug</td></td<>	Ecological methods	National rates	Time series methods	Large samples or entire populations can be studied	Do not know if person experiencing the AE actually took the drug
Small area estimation   Mixed-effects Poisson regression   Hypothesis generation   Geographic variability in report     Meta-analysis   Synthesis of RCTs   Fixed-effects models (MH model)   Randomization   Limited generalizability     Meta-analysis   Synthesis of RCTs   Fixed-effects models (MH model)   Randomization   Limited generalizability     Synthesis of RCTs   Fixed-effects logistic regression   Person-level   Exclusion of zero-event studies     Mixed-effects logistic regression   Multilevel mixture models   Heterogeneity   Publication bias     Medical claims   Case-control studies   Mixed-effects logistic and Poisson   Large samples   Confounding by indication     Medical claims   Case-control studies   Mixed-effects logistic and Poisson   Large samples   Confounding by indication     Medical claims   Case-control studies   Mixed-effects logistic and Poisson   Large samples   Confounding by time of treatment bias     Medical claims   Case-control studies   Mixed-effects logistic and Poisson   Large samples   Confounding by time of treatment bias     Medical claims   Case-control studies   Mixed-effects logistic and Poisson   Large samples   Confounding by time of treatment bias     Methinsubject designs   Propensity scores   Concomitant medications   Unmeasured confoundiers     Matching		Natural experiments	Change-point analyses	Permits between-stratum comparisons	Subject to ecological fallacy
Meta-analysis   Synthesis of RCTs   Fixed-effects models (MH model)   Randomization   Limited generalizability     Synthesis of   Random-effects models (MH model)   Person-level   Exclusion of zero-event studies     Synthesis of   Mixed-effects models (DL model)   Person-level   Exclusion of zero-event studies     Mixed-effects logistic regression   Mixed-effects logistic regression   Heterogeneity     Mutilevel mixture models   Mutilevel mixture models   Ascertainment bias     Medical claims   Case-control studies   Mixed-effects logistic and Polsson   Large samples     Medical claims   Case-control studies   Mixed-effects logistic and Polsson   Large samples   Confounding by indication     Mithin-subject designs   Person-level   Confounding by indication   Unmeasured confounders     Matching   Person-level   Concomitant medications   Unmeasured confounders     Matching   Person-level   Concomitant medications   Unmeasured confounders     Matching   Propensity scores   Pronortion of Actions		Small area estimation	Mixed-effects Poisson regression	Hypothesis generation	Geographic variability in reporting
Meta-analysis   Synthesis of RCTs   Fixed-effects models (MH model)   Randomization   Limited generalizability     Synthesis of an extract effects   Random-effect model (DL model)   Person-level   Exclusion of zero-event studies     Synthesis of an extract effects   Mixed-effects logistic regression   Person-level   Exclusion of zero-event studies     Multilevel mixture   Multilevel mixture models   Mixed-effects logistic regression   Heterogeneity     Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples   Confounding by indication     Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples   Confounding by indication     Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples   Confounding by indication     Mithin-subject designs   Person-level   Confounding by indication   Unmeasured confounders     Matching   Matching   Confounding by indications   Unmeasured confounders     Matching   Person-level   Confounding by indications   Unmeasured confounders     Mithin-subject designs   Propensity scores   Confounding by indications   Unmeasured confounders     Matching   Person-level   Confounding by indications   Unmeasured confounders     Matching   Person-level <td></td> <td></td> <td></td> <td></td> <td></td>					
Synthesis of observational studies   Random-effect model (DL model)   Person-level   Exclusion of zero-event studies     Mixed-effects logistic regression   Mixed-effects logistic regression   Heterogeneity     Multilevel mixture models   Mixed-effects logistic regression   Heterogeneity     Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples     Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples     Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples     Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples     Mithin-subject designs   Person-level   Confounding by time of treatme     Reten-subject designs   Person-time logistic models   Concomitant medications     Matching   Ascentainment bias   Unmeasured confounders     Matching   Person-time logistic models   Confounding by time of treatme     Reten-subject designs   Person-time logistic models   Confounding by time of treatme     Matching   Person-time logistic models   Confounding by time of treatme     Mithin-subject designs   Person-time logistic models   Confounding by time of treatme     Matching   Person-time logistic models   Confounding by time of treatme  <	Meta-analysis	Synthesis of RCTs	Fixed-effects models (MH model)	Randomization	Limited generalizability
Mixed-effects logistic regression   Heterogeneity     Multilevel mixture models   Heterogeneity     Multilevel mixture models   Publication bias     Multilevel mixture models   Publication bias     Multilevel mixture models   Confounding by indication     Multilevel mixture models   Eixed-effects logistic and Poisson   Large samples     Cohort studies   Fixed-effects logistic and Poisson   Large samples     Cohort studies   Mixed-effects logistic and Poisson   Confounding by indication     Within-subject designs   Person-level   Confounding by time of treatme     Mutching   Prison-level   Concomitant medications   Unmeasured confounders     Mutching   Propensity scores   Co-morbid diagnoses   Unsystematic diagnostic criterit     Mutching   Prior history of relevant   Based on filled prescriptions or AEs     Differential effects   Canadita diagnoses   Linited dosage and duration diagnose		Synthesis of observational studies	Random-effect model (DL model)	Person-level	Exclusion of zero-event studies
Multilevel mixture models   Publication bias     Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples   Confounding by indication     Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples   Confounding by indication     Medical claims   Case-control studies   Mixed-effects logistic and Poisson   Large samples   Confounding by indication     Mithin-subject designs   Mixed-effects logistic and Poisson   Large samples   Confounding by indication     Mithin-subject designs   Person-level   Confounding by indications   Unmeasured confounders     Mithin-subject designs   Person-line logistic models   Concomitant medications   Unmeasured confounders     Matching   Propensity scores   Co-morbid diagnoses   Unsystematic diagnostic criterit     Matching   Afes   Linited dosage and duration diagnostic   Linited dosage and duration diagnostic			Mixed-effects logistic regression		Heterogeneity
Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples   Confounding by indication     Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples   Confounding by indication     Cohort studies   Mixed-effects logistic and Poisson   Large samples   Confounding by indication     Within-subject designs   Mixed-effects logistic and   Person-level   Confounding by time of treatme     Matching   Person-level   Concomitant medications   Unmeasured confounders     Matching   Propensity scores   Co-morbid diagnoses   Unsystematic diagnostic criterit     Differential effects   Differential effects   Correction   Linited dosage and duration			Multilevel mixture models		Publication bias
Medical claims   Case-control studies   Fixed-effects logistic and Poisson   Large samples   Confounding by indication     Cohort studies   Mixed-effects logistic and   Person-level   Confounding by time of treatme     Cohort studies   Mixed-effects logistic and   Person-level   Confounding by time of treatme     Within-subject designs   Person-time logistic models   Concomitant medications   Unmeasured confounders     Matching   Person-time logistic models   Concomitant medications   Unsystematic diagnostic criterit     Matching   Propensity scores   Co-morbid diagnoses   Unsystematic diagnostic criterit     Differential effects   Differential effects   Generalizable   Limited dosage and duration					Ascertainment bias
Cohort studies   Mixed-effects logistic and Poisson   Person-level   Confounding by time of treatme     Within-subject designs   Person-time logistic models   Concomitant medications   Unmeasured confounders     Between-subject designs   Propensity scores   Co-morbid diagnoses   Unsystematic diagnostic criteri     Matching   Action   Prior history of relevant   Based on filled prescriptions or AEs   Limited dosage and duration duration duration duration duration duration duration	Medical claims	Case-control studies	Fixed-effects logistic and Poisson	Large samples	Confounding by indication
Within-subject designs   Person-time logistic models   Concomitant medications   Unmeasured confounders     Between-subject designs   Propensity scores   Co-morbid diagnoses   Unsystematic diagnostic criteria     Matching   Prior nistory of relevant   Based on filled prescriptions or AEs   Differential effects		Cohort studies	Mixed-effects logistic and Poisson	Person-level	Confounding by time of treatment
Between-subject designs Propensity scores Co-morbid diagnoses Unsystematic diagnostic criterit Matching Ased on filled prescriptions on AEs Limited dosage and duration da		Within-subject designs	Person-time logistic models	Concomitant medications	Unmeasured confounders
Matching Prior history of relevant Based on filled prescriptions on AE AE Differential effects Generalizable Limited dosage and duration da		Between-subject designs	Propensity scores	Co-morbid diagnoses	Unsystematic diagnostic criteria
Differential effects Generalizable Limited dosage and duration de		Matching		Prior history of relevant AEs	Based on filled prescriptions only
		Differential effects		Generalizable	Limited dosage and duration data

Drug Saf 2011; 34 (5)

of patients; (iv) effects of publicity in the media on numbers of reports; (v) extreme duplication of reports; (vi) attribution of the event to a single drug when patients may be exposed to multiple drugs; and (vii) failure to account for the population at risk. These limitations do not characterize all SRS data and some may be eliminated by use of better statistical approaches. For example, when information on indication and concomitant medications is available, it may be included in the analysis to adjust for these factors using many of the previously described methods. Furthermore, the lack of representativeness SRS data, which may somewhat limit generalizability, is generally not a statistical concern since the methods for analysing these data do not relate to the (unknown) exposed population.

#### 3.2 Ecological Methods for Rare Adverse Events

For very rare events (e.g. death by suicide), there may be few options for routine drug surveillance. One approach is to use ecological data that relate changes in drug prescription rates to AE rates. These more global associations do not support causal inferences, but the availability of large denominators and close to complete enumeration of events such as suicides can be hypothesis-generating and help support inferences drawn from other studies. In some cases, natural experiments, such as black-box warnings,<sup>[62]</sup> provide an opportunity to evaluate the positive or negative consequences of decreased access to the drug on the event of interest.

Analysis is often based on log-linear or Poisson regression of suicide rates over time, using exposure based on prescription rates during the same time period. Serial correlation can be accommodated using Huber-White robust standard errors, which allow for an arbitrary autocorrelation pattern.<sup>[63-65]</sup> Where data from multiple countries are combined, both fixed-effects<sup>[66]</sup> and randomeffects models<sup>[67]</sup> can be used to allow each country to have its own linear time trend. Alternatively, simultaneous county-level analyses can be performed<sup>[11]</sup> in which AE rates are stratified by demographic characteristics such as age, race and sex within counties, and a mixed-effects Poisson regression model is used to analyse the data, treating the county as the unit of analysis (using population as an offset). County-level prescription rates can be added to the model. When longitudinal data are available, between-county effects and within-county effects can be uniquely estimated.<sup>[32,67]</sup>

#### 3.3 Meta-Analysis of Randomized Controlled Trials

Prior to drug approval and/or release of a new drug, a series of RCTs is conducted that include obtaining information on AEs. While an individual study typically has insufficient power to detect a potential drug-AE relationship, meta-analysis of data from several trials may provide a more powerful statistical inference. Meta-analysis is an observational study of studies and does not in and of itself provide a causal inference. Studies are often combined using multiple drugs used for multiple indications, and this inherent heterogeneity is often ignored in conclusions drawn regarding the effect of the class of drugs as a whole on the AE of interest. The studies that are available for a meta-analysis may also suffer from selection bias, in that they may be restricted to only those studies in which the drug was shown to be efficacious.

The Mantel-Haenszel method, a moment-based estimator, assumes that treatment is a fixed effect, and combines studies using an inverse variance to determine the weight given to each study. The method assumes that the log odds ratio is equal across all studies (i.e. there is no treatment variability).

DerSimonian and Laird,<sup>[68]</sup> developed a momentbased estimator in which they assume that treatment is a random effect and can vary over studies. However, a number of simulation studies<sup>[69-71]</sup> have shown that the heterogeneity estimate has a large negative bias, leading to a biased estimate of the pooled treatment effect as well. In addition, the Q statistic that is used to test heterogeneity has low power to detect departure from homogeneity.<sup>[72,73]</sup>

Mixed-effects logistic regression models (likelihood-based estimators), take heterogeneity into account while combining results across trials.<sup>[67]</sup> These methods do not suffer from the same potential statistical estimation bias inherent in the previously described moment-based estimators. In addition to providing estimates of treatment effects and treatment and background incidence heterogeneity (for randomized studies), it also allows trial-level co-variates in the analysis. Unlike the previous two moment-based methods, studies with zero events can be included in the analysis and there is no need to add a constant to studies that have a single arm with no events (i.e. continuity correction). These methods should be used in place of the moment-based estimators that have been used by the FDA and others in performing meta-analyses of RCT data. To the extent that there is heterogeneity of the treatment effect across studies, different results between the moment- and likelihood-based approaches to meta-analysis can be obtained.

While a detailed review of the statistical and methodological properties of meta-analysis is beyond the scope of this review, a few general points are important to guide further work in this area. First, it is critically important to make every attempt to include all studies, both published and unpublished, when attempting to synthesize research in a particular area. Second, rather than assuming homogeneity, it would seem more prudent to assume heterogeneity of treatment effects and background incidence rates, and only fit simpler models when there is convincing evidence to indicate that such heterogeneity is not present. It is unlikely that traditional moment-based estimators will provide such information. Third, whenever possible, all of the original data should be obtained (including longitudinal data when the original studies involved repeated observations) and these data should be re-analysed preserving the nesting of repeated observations within subjects and subjects within studies, rather than relying upon meta-analytic procedures for pooling effect sizes or odds ratios.

#### 3.4 Medical Claims Data

Medical claims data have several advantages over spontaneous reports for drug surveillance. First, they represent person-level data, similar to RCTs and spontaneous reports, but unlike spontaneous reports, we know the population at risk. Second, several medical claims databases such as the Veterans Administration (VA) or PharMetrics databases contain longitudinal information on AEs, concomitant medications and co-morbid diagnoses both before and after the drug exposure. Third, the populations that can be sampled are often large enough to study even the rarest of events, such as suicide. Their primary limitation is that they are observational, and any association identified and related inferences may be biased due to unmeasured confounding.

There are several different types of studies that can be conducted using medical claims data. The simplest is a *case-control study*, where 'cases' are defined as patients who have experienced the AE of interest, and 'controls' are similar to the cases but have not experienced the AE. The goal of the analysis is then to compare the rate of drug exposure between cases and controls. If a significant difference is identified, then there is evidence of an association between the drug and the AE. In some cases, propensity score matching (PSM)<sup>[74]</sup> can be used to identify controls that are matched in probability on a large number of potential confounders to the cases. A major limitation of case-control studies in drug surveillance is that the available potential confounders are often inadequate for matching the cases and controls in terms of severity of illness. As such, the resulting comparison may still represent confounding by selection (i.e. sicker patients are more likely to be treated and exhibit the AE). Gibbons et al.<sup>[75]</sup> have shown that despite balancing observed covariates, pre-treatment suicide attempt rates remained unbalanced despite PSM in a large sample of bipolar patients treated with an AED.

A *cohort study* identifies a sample from a well defined population based on predetermined criteria. The cohort can be defined in terms of an illness (e.g. major depressive disorder [MDD]) or based on an exposure (e.g. all patients taking an SSRI), both within a given timeframe. Analysis can be restricted to 'new cases' who have not been diagnosed or treated for some fixed period of time. The cohort study can be designed to have a

fixed time window before and after the indication (either diagnostic date or first treatment date) so that a pre-treatment base-rate can be estimated.

Within-subject cohort studies are those in which the same patients are repeatedly measured over time, typically before and after initiating drug treatment. The basic idea is to compare the rate of a non-fatal AE (e.g. suicide attempt) before and after exposure to the drug. The strength of the design is that it is restricted to only those patients who ultimately take the drug, thereby minimizing selection effects. However, in the case of depression and suicidal behaviour or ideation, a limitation of the design is that the natural course of the disease (e.g. decrease in the severity of depressive symptomatology over time) can become confounded with the pre-post nature of the design. In some cases, the emergence of the AE may even lead to treatment. For example, a suicide attempt may lead to the identification of the depressive disorder that may in turn lead to treatment. By the 'regression to the mean' effect alone, we would expect the AE, on average, to decrease, and this decrease could incorrectly be attributed to a protective effect of the drug. Person-time logistic regression models can be used to adjust for such confounding.<sup>[76]</sup>

Between-subject cohort studies involve comparison of patients who took the drug versus those who did not. In certain cases, it may be useful to compare monotherapy versus no therapy, at least with respect to the other drugs within the class. In other cases in which polytherapy is the norm, concomitant drug therapy can be statistically adjusted for. The importance of considering monotherapy is that patients using multiple drugs within the same class may be of greater initial severity and/or treatment resistant and may, in general, have higher rates of AEs. If monotherapy only is considered, we must consider the possibility that some monotherapies may require additional treatments whilst others do not, which can lead to selection bias. The primary limitation of between-subject designs is that they may be subject to confounding by indication. In general, more severely ill patients will be treated with pharmacotherapy; therefore, we would expect them to have a greater incidence of AEs that are related to the severity of illness. PSM can be helpful for reducing bias, but only to the extent that the potential confounders are known and measurable.

#### 3.5 New Statistical Approaches

*Differential effects*<sup>[77]</sup> attempts to provide an unbiased comparison of two treatments or treatment versus no treatment by stratifying on concomitant treatment. Consider a comparison of SSRIs versus TCAs in terms of suicidal attempts. First, rates of suicide attempts are compared in depressed patients who received either an SSRI or a TCA, but not both. Second, suicide attempt rates are compared for patients who received SSRIs plus psychotherapy or TCAs plus psychotherapy, but not both. If SSRIs are stimulating suicidality, when they are present in the drug mix, we would expect to see an excess of suicide attempts, otherwise this pattern is compatible with confounding by indication.

*Person-time logistic regression*<sup>[78,79]</sup> allows us to use drug exposure as a time-varying co-variate in estimating the hazard rate of an AE on a month-by-month basis. This analysis can combine patients who did not take the drug with non-medication months for other patients who did take the drug, and compares them with active treatment months. The model adjusts for month, which allows determination of whether the risk of the AE decreases (or increases) over time (e.g. regression effect). Inclusion of a treatment by time interaction permits non-proportional hazard rates, an example of which is when the effect of the drug on suicidal behaviour can change over time.

#### 4. Recent Literature on Drugs and Suicide

#### 4.1 Antidepressants and Suicide

Recent attention regarding antidepressants and suicide led to a US black-box warning for children aged <18 years in October 2004. The evidence supporting the warning was a meta-analysis conducted by the FDA,<sup>[34,62]</sup> which combined spontaneous reports of suicidal thoughts and behaviours from 25 placebo-controlled paediatric RCTs of SSRIs and SNRIs. These spontaneous reports were made by patients during the course of their participation in RCTs and are quite different from spontaneous reports that arise in clinical practice settings. The primary outcome was suicidal ideation and behaviour, and a higher rate was found for children treated with antidepressants versus placebo (OR 1.78; 95% CI 1.14, 2.77). The suicidal behaviour and ideation measure was based on spontaneous reports of adverse events recorded by clinicians during the course of the studies. The FDA also presented results of an analysis of prospective data (suicidal ideation or behaviour rating-scale item), which showed no effect for emergence (OR 0.93; 95% CI 0.75, 1.15), or emergence and/or worsening of suicidal thoughts and behaviour (OR 0.92; 95% CI 0.76, 1.11). The disconnection between the prospective clinician ratings and spontaneous patient reports has never been adequately explained. The difference may be due to ascertainment bias in which children treated with an active medication develop more adverse effects in general than children receiving placebo, which leads to increased contact with medical staff and more opportunity for reporting suicidal ideation.<sup>[80]</sup> It has also been suggested that patients attempting suicide by overdosing on their study medication will be more likely to have the suicidal behaviour detected if they are on active medication relative to placebo.<sup>[81]</sup>

In January 2006, the FDA conducted a second meta-analysis<sup>[82]</sup> of 372 RCTs of antidepressants (SSRIs and SNRIs) in an adult population (about 100 000). While the overall analysis revealed no evidence of an association, stratification by age revealed that for the primary endpoint of suicidal ideation or behaviour, 18- to 24-year-olds had an increased risk on medication relative to placebo approaching significance (OR 1.62; 95% CI 0.97, 2.71); adults aged 25-64 years had a significantly decreased risk (OR 0.79; 95% CI 0.64, 0.98), and geriatric patients had a markedly significantly decreased risk (OR 0.37; 95% CI 0.18, 0.76) on antidepressants relative to placebo. On the basis of these results, the FDA extended the black-box warning to 18- to 24-year-olds. In this analysis, the FDA did not provide results for the prospective clinician ratings. The majority of events in these studies were suicidal ideation.

Bridge and colleagues<sup>[50]</sup> analysed an expanded set (27 studies) of paediatric RCTs of antidepressant treatment and suicidality, originally analysed by the FDA. They performed meta-analyses of both suicidality and efficacy. The association between receiving antidepressant treatment and suicidality was reduced and significant efficacy was found for all indications.

Gibbons et al.<sup>[81]</sup> studied a cohort of 226866 veterans who had a new diagnosis of MDD without a history of MDD or antidepressant treatment in the previous 2 years. The sample considered only patients aged >18 years and was weighted towards patients >25 years. Comparisons of suicide attempt rates were made both within patients (before and after initiation of treatment) and between individuals (i.e. comparison of those taking and not taking antidepressant medication). The suicide attempt rate was significantly lower for patients treated with an SSRI only (monotherapy) compared with those without antidepressant treatment (123/100000 for SSRIs vs 335/100000 for no antidepressant; OR 0.37; p < 0.0001). In patients treated with an SSRI only, the rate of suicide attempts was significantly lower after treatment (123/100000) than before treatment (221/100000) [relative risk 0.56; p<0.0001].

Analyses stratified by age did not confirm the FDA's findings of increased suicidality for 18- to 24-year-olds. Comparison of suicide attempt rates for depressed patients not treated with antidepressants, versus those patients treated with SSRIs only, yielded consistent estimates of decreased risk with treatment for patients aged 18–25 years (OR 0.35; 95% CI 0.14, 0.85; p < 0.021), 26–45 years (OR 0.44; 95% CI 0.29, 0.65; p < 0.0001), 46–65 years (OR 0.42; 95% CI 0.30, 0.59; p < 0.0001) and >65 years (OR 0.38; 95%CI 0.16, 0.91; p < 0.036). Differences in suicide attempt rates before and after treatment initiation were also unrelated to age.<sup>[81]</sup>

The difference between the VA data and the FDA data in young adults is not due to differences in active treatment suicide attempt rates (477/100 000 VA vs 551/100 000 FDA), but rather the difference in suicide attempt rates in untreated patients versus placebo controls (1368/100 000 VA vs 268/100 000 FDA). Lack of treatment in

an observational study, and receiving placebo in an RCT appear to be associated with quite different suicide attempt rates. There are several possible explanations. First, it may be that patients in RCTs are not representative of the patients seen in routine practice, and that patients selected for RCTs are less suicidal. By contrast, in routine practice, the most suicidal patients may not be offered antidepressant treatment for fear of exacerbating their suicidal tendencies based on the recent regulatory attention to this issue. Second, patients receiving placebo in RCTs may still receive supportive clinical contact, which decreases their suicidal tendencies, whereas the patients in the VA who do not receive medication may not be receiving any other form of supportive care. Any of these potential sources of selection can lead to bias in observational studies.

The VA data have been reanalysed using persontime logistic regression,<sup>[76]</sup> revealing a significant decrease in suicide attempt rate during months with SSRI (monotherapy) treatment (hazard ratio [HR] 0.17; 95% CI 0.10, 0.28; p=0.0001). This compares favourably with the observed data, for which the monthly suicide attempt rate was 207/805525 (0.026%) for untreated months and 17/328 648 (0.005%) for treated months, yielding a raw HR of 0.19. Overall, the suicide attempt rate decreased with time from the index episode (see figure 1a). Figure 1b provides the estimated hazard functions for the non-proportional hazards model (i.e. risk difference varies over time). Figure 1b reveals that the difference in hazard rates is largest early in treatment (favouring SSRI treatment) but hazard rates are essentially equivalent by approximately 9 months following the index episode.

Recent ecological studies conducted following the black-box warning revealed that there may have been unintended consequences of the warning. Several authors (including the FDA<sup>[83]</sup>) have now shown that antidepressant prescription rates precipitously dropped following the warning.<sup>[51,83-85]</sup> Both Gibbons et al.<sup>[51]</sup> and the CDC<sup>[86]</sup> documented increased child and adolescent (5–19 years of age) suicide rates (a 14% increase) following the decreases in antidepressant prescriptions. Libby et al.<sup>[87,88]</sup> found significant decreases in the diagnosis of new cases of child and young adult depression (5–18 years of age) among general practitioners following the black-box warning (44% reduction in paediatric patients and 37% for young adults). Similar studies in the UK did not identify increases in suicide despite decreases in antidepressant prescriptions; however, the use of medication for the treatment of childhood depression in the UK is less than in the US.<sup>[89]</sup>

In summary, the data suggest that there may be an association between antidepressant treatment and suicidal ideation in children. However, large-scale observational studies of suicide attempts have generally failed to replicate these findings. In adults, the evidence is in the opposite direction, where most studies find significant decreases in suicidality among treated patients.

#### 4.2 Antiepileptic Drugs and Suicide

In addition to antidepressants, other classes of drugs (e.g. AEDs) have been suspected of having a relationship with suicide. Epilepsy carries an elevated risk of suicide and many AEDs are used as mood stabilizers in bipolar disorder or may have antidepressant properties, so the effect of AEDs on suicidal behaviour is of importance. The FDA conducted a meta-analysis of 199 placebo-controlled trials including 43892 patients (27863 in drug treatment groups and 16029 in placebo groups) for the AEDs gabapentin, divalproex, felbamate, lamotrigine, levetiracetam, oxcarbazepine, pregabalin, tiagabine, topiramate, zonisamide and carbamazepine. Suicidal ideation or behaviour was reported by 0.43% of treated patients versus 0.22% of patients receiving placebo. Although the FDA concluded that the risk was 'generally consistent' across the 11 AEDs, examination of the data indicates that this does not seem to be the case. Topiramate (27 events) and lamotrigine (40 events) had 61% of all of the events, but only represented 38% of the data. Individually, these two drugs showed significant association with suicidality (lamotrigine: OR 2.08; 95% CI 1.03, 4.40 and topiramate: OR 2.53; 95% CI 1.21, 5.86), whereas none of the other AEDs did (OR 1.13; 95% CI 0.65, 1.95.<sup>[90]</sup>



Fig. 1. Comparison of selective serotonin reuptake inhibitor (SSRI) monotherapy vs no antidepressant for suicide attempts over time. (a) proportional hazards model; (b) non-proportional hazards model.<sup>[76]</sup>

Goodwin et al.<sup>[91]</sup> studied 20638 bipolar disorder patients from two large integrated health plans from California and Washington. Suicide attempt rates were significantly higher for divalproex versus lithium (31.3/1000 vs 10.8/1000 person-years [PY]). Completed suicide rates were also higher for divalproex versus lithium (1.7/1000 vs 0.7/1000). There was no adjustment for previous

suicide attempts or comparison with untreated controls.

Collins and McFarland<sup>[92]</sup> compared suicide completion and attempt rates between lithium, gabapentin, divalproex and carbamazepine in a cohort of 12 662 bipolar patients from an Oregon Medicaid medical claims database. There were 11 suicides and 79 attempts. Relative to lithium, divalproex had a higher suicide attempt rate, and gabapentin had a higher rate of suicide completion. In the absence of untreated controls, it is unclear whether lithium is protective or the AEDs are harmful. There may also be confounding with lack of treatment response and/or pain, leading to use of an AED.

Gibbons et al.<sup>[75]</sup> studied a cohort of 47918 patients with bipolar disorder with a minimum of a 1-year window of information before and after the index date of their illness. The primary analysis was restricted to patients receiving monotherapy with 1 of the 11 AEDs or lithium. Overall, there was no significant difference in suicide attempt rates between AED treatment (13/1000 PY) versus no AED (or lithium) treatment (13/1000 PY). In AED-treated subjects, the rate of suicide attempts was significantly higher prior to treatment (72/1000 PY) than after treatment (13/1000 PY). In patients receiving no concomitant treatment with an antidepressant, other AED or antipsychotic, AEDs had a significantly lower rate of suicide attempts relative to the untreated controls (3/1000 PY vs 15/1000 PY). A person-time logistic regression adjusting for age, sex, previous suicide attempt, month, concomitant antidepressants, antipsychotics and other AEDs found significantly decreased suicide attempt risk with AED treatment (OR 0.59; 95% CI 0.47, 0.75; p < 0.0001), which was even larger in the 662 patients with prior attempts (OR 0.35; 95% CI 0.17, 0.74; p < 0.005). These results have recently been replicated for both suicide attempts and completions,<sup>[93]</sup> and in a cohort of over 130 000 patients taking gabapentin.<sup>[94]</sup>

Patorno et al.<sup>[95]</sup> compared the 11 AEDs for risk of suicide attempts, completion and violent deaths. In addition to co-variate adjustment, the authors used pair-wise 1:1 high-dimensional PSM for each of the ten AEDs relative to a reference AED (topiramate). The key assumption is that PSM will produce drug groups that are equivalent on all measured and unmeasured confounders, such that residual differences represent differential drug effects. This is a very difficult assumption to verify. While we can determine success of the algorithm for measured confounders, hidden bias in terms of unmeasured confounders cannot be verified.<sup>[96]</sup>

Their analysis revealed that gabapentin, lamotrigine, oxcarbazepine, tiagabine and valproate had higher overall rates (primarily attempts) than topiramate. However, in the absence of an untreated comparison group, the between-drug differences may simply represent differences in the degree of possible protective effects of these drugs.<sup>[75]</sup>

There are large differences among these 11 AEDs in rates of suicidal behaviour prior to treatment.<sup>[75]</sup> PSM was used to match AED-treated and untreated patients on a large number of potential confounders.<sup>[75]</sup> While PSM brought about balance for the measured confounders, pre-treatment suicide attempt rate remained twice as high in patients who were ultimately treated with an AED (71/1000 PY) versus those who were not (38/1000 PY; p<0.004).<sup>[75]</sup> Had this difference occurred after treatment, it would have been interpreted as increased risk of suicidal behaviour associated with AED treatment, when in fact it simply represents failure of PSM to eliminate bias.

Finally, matched samples are generally not representative, and large differences in incidence can be found depending on which drug is selected to be the reference. Patorno et al.<sup>[95]</sup> used topiramate as a primary reference drug and carbamazepine as a secondary reference drug. When topiramate was compared with carbamazepine (reference drug) in patients with epilespsy, the suicide attempt rate for carbamazepine was 41/1000 PY, which was higher than the rate for topiramate (27/1000 PY). However, when carbamazepine was compared with topiramate (reference drug), the rate for carbamazepine was 4/1000 PY, which was lower than the rate for topiramate (6/1000 PY). Even comparisons between the same two drugs can yield large differences in incidence and even direction, depending on which drug is selected as the reference, because quite different subsets of patients are included.

Olesen et al.<sup>[97]</sup> conducted a case-crossover study in which the rate of AED treatment was compared during 30 days prior to death versus two control periods (60-90 days and 90-120 days prior to suicide). The authors concluded that clonazepam, valproate, lamotrigine and phenobarbitone (phenobarbital) may increase the risk of suicide relatively shortly after the initiation of treatment, whereas none of the other AEDs conferred such risk. It is unclear whether the increased risk was due to increased severity of illness during the 30 days prior to suicide (which coincided with increased treatment intensity) or to exposure to the AED. These authors also conducted a cohort study in which they compared suicide rates between the 11 AEDs versus carbamazepine as a reference. Clonazepam, valproate, lamotrigine, phenobarbitone and levetiracetam were all significantly associated with higher suicide rates when compared with carbamazepine. This cohort study is similar to the drug-to-drug comparison study conducted by Patorno et al.<sup>[95]</sup> and also suffers from the absence of an untreated control comparison group. In addition, there is little overlap between the drugs found to have elevated suicidality rates in the two studies.

VanCott et al.<sup>[98]</sup> conducted a case-control study of an elderly VA population who received new prescriptions for AED monotherapy (n= 112096). Cases were defined as exhibiting suicidal ideation or behaviour. Twelve controls were selected for each case and were matched in terms of history of suicide-related behaviours prior to AED treatment, the first year of AED treatment and diagnosis of epilepsy. The case-control sample consisted of 832 individuals (64 cases with suicidality and 768 controls). In this study, gabapentin was used as the reference medication. Results of the analysis revealed that the strongest predictor of suicide-related behaviour was having an affective disorder. Patients prescribed newer AEDs (levetiracetam or lamotrigine) had significantly increased suicidality risk compared with gabapentin (OR 10.2; 95% CI 1.1, 97.0).

In summary, these studies do not suggest that AEDs as a class are associated with increased risk

of suicidality relative to patients not treated with AEDs. The FDA's meta-analysis appears to have been driven by topiramate and lamotrigine, and their conclusion that the risk was generally similar across the 11 AEDs is not supported by their data. Furthermore, the recent drug-to-drug comparisons<sup>[95,97,98]</sup> all show significant differences between AEDs in suicidality risk, which also contradicts the FDA's conclusions regarding similarity of the association across all AEDs. The most consistent effect appears to be for lamotrigine, which had a significantly higher suicidality rate relative to placebo in the FDA's study and consistently higher than the various active AED reference drugs in the more recent drug-to-drug comparison studies. The absence of untreated controls, however, limits our ability to conclude that these between-drug differences represent increased risk relative to what would be observed without AED treatment. The study by Gibbons et al.<sup>[75]</sup> did include untreated control conditions and failed to identify increased risk of AEDs as a group in bipolar patients who are at the highest risk of suicidality among those treated with these drugs.

#### 4.3 Montelukast and Suicide

Montelukast is the most recent drug to be linked to suicide. The FDA first issued a safety alert in March 2008, following a media report of a 15-year-old in New York who killed himself 17 days after starting to take the drug for allergies.

According to the FDA,<sup>[90]</sup> there was one case of suicidal ideation but no suicide attempts or completions out of 9929 montelukast-treated patients across 41 placebo-controlled trials. There were no events among 7780 placebo patients. No events were found for 7540 zafirlukast-treated patients in 45 placebo-controlled clinical trials, but two events were observed for placebo. No events were found in 1745 zileuton-treated or 1063 placebo-treated patients in 11 clinical trials. The FDA's conclusion from this analysis was that, while these data do not suggest that montelukast, zafirlukast or zileuton are associated with suicide or suicidal behaviour, these clinical trials were not designed specifically to examine neuropsychiatric events.

A separate review of placebo-controlled paediatric studies of montelukast found similar results.<sup>[99]</sup> Among 2751 paediatric patients in four placebo-controlled RCTs there were no neuropsychiatric AEs.

The literature contains few data to support the FDA precaution. After the initial FDA alert was issued, Holbrook and Harik-Khan<sup>[100]</sup> analysed the association between montelukast and depression based on data from three previously published RCTs involving 504 patients exposed to montelukast. Using emotional well-being as a marker for depression, they reported no evidence of a negative effect from montelukast and no cases of psychiatric disturbances, suicide or depressive episodes.

Jick et al.<sup>[101]</sup> conducted a population-based cohort study using data from the UK General Practice Research Database (GPRD). The investigators identified 23 500 patients exposed to one or more prescriptions for montelukast from February 1998 to March 2007, representing 21 050 PY at risk. Only one case of suicide was identified; this was in a woman who had been prescribed a single 28-day course of montelukast approximately 2 years prior to her death.

In summary, there does not appear to be any strong evidence in support of an association between montelukast and suicide. It should be noted, however, that asthma is a risk factor for suicidal ideation, attempts and completion,<sup>[102]</sup> leading to confounding by indication for drugs such as montelukast as well.

#### 4.4 Varenicline and Suicide

While varenicline may be the most effective smoking cessation agent currently available on the market, it has also been suspected of having a link to depression and suicide.<sup>[103,104]</sup> In December 2007, following spontaneous reports of depression and suicidal thoughts, the UK Medicines and Healthcare products Regulatory agency (MHRA) issued a warning concerning increased risk of suicidality. In July 2009, the FDA issued a black-box warning on varenicline and suicidality. By April 2009 the MHRA had received reports of 14 suicides in people taking varenicline. Gunnell et al.<sup>[105]</sup> showed that rates of spontaneous reports of suicide-related events increased dramatically following early regulatory actions. Patients who smoke have two to three times the risk of suicide relative to non-smokers,<sup>[106,107]</sup> making it more difficult to identify an association.

Hughes<sup>[103]</sup> provided a review on smoking and suicide, which included a review of the published trials of varenicline. Among the 2183 smokers exposed to varenicline in published trials, several AEs were more common with varenicline than placebo but suicide and depression were not cited as common AEs.<sup>[108,109]</sup> The two studies that reported on post-cessation mood reported varenicline improved mood compared with placebo.<sup>[110,111]</sup> One trial reported a suicide in the placebo group.<sup>[112]</sup> Another trial reported a suicide in the varenicline group in a participant with a past history of depression.<sup>[113]</sup>

Gunnell et al.<sup>[105]</sup> conducted a cohort study comparing reports of suicidal thoughts, depression, and fatal and non-fatal self-harm in patients taking nicotine replacement products (n = 63265), varenicline (n = 10973) and bupropion (6422). Compared with nicotine replacement products, the HR for self-harm for patients prescribed varenicline was 1.12 (95% CI 0.67, 1.88) and 1.17 (95% CI 0.59, 2.32) for patients prescribed bupropion. Similarly, there was no evidence that varenicline was associated with increased rates of depression (HR 0.88; 95% CI 0.77, 1.00) or suicidal thoughts (HR 1.43; 95% CI 0.53, 3.85). These findings are particularly important given the long-term health benefits of smoking cessation and the documented efficacy of varenicline as an aid to smoking cessation.

Kasliwal et al.<sup>[114]</sup> used prescription event monitoring (PEM; a survey of physicians who prescribed varenicline) to study the adverse event profile for varenicline. This form of postmarketing surveillance is active rather than passive and has the advantage of a denominator indicating the number of patients at risk. The weakness of PEM is that not all physicians who prescribe the drug of interest return the questionnaire. Their cohort consisted of 2682 patients. Two cases of attempted suicide were reported during treatment, both of which were in patients who had a previous history of psychiatric illness and precipitating factors for the suicidal event.

In summary, there does not appear to be any strong evidence in support of an association between varenicline and suicide. It is unclear from our review whether adverse event data from the original RCTs have been analysed with respect to the emergence of suicidality.

#### 4.5 Antipsychotics and Suicide

Approximately 50% of patients with schizophrenia or schizoaffective disorder attempt suicide, and approximately 10% die as a result of suicide.<sup>[115]</sup> Unlike other CNS drugs, much of the evidence for antipsychotics and suicide suggests protective effects. Khan et al.<sup>[116]</sup> found no difference in rates of suicide and attempted suicide between placebo and active treatment in a sample of 10118 subjects from worldwide phase I-III RCTs of risperidone, olanzapine and quetiapine fumarate. Herings and Erkens<sup>[117]</sup> studied 603 patients with schizophrenia in terms of uninterrupted versus interrupted use of two antipsychotics (olanzapine and risperidone), where suicide attempt rates were 20/1000 versus 72/1000 PY for uninterrupted versus interrupted use respectively). These data reveal that patients who do not refill their prescriptions for atypical antipsychotics are at higher risk for suicide attempts. Meltzer et al.<sup>[115]</sup> found significantly decreased risk of suicidal behaviour with clozapine versus olanzapine (HR 0.76; 95% CI 0.58, 0.97) in a large multicentre RCT of schizophrenic patients at high risk for suicide attempt. This is one of the few studies designed to examine differential drug effects on suicidal events. Early studies in mood disorders<sup>[118]</sup> suggest that flupenthixol may reduce suicide risk. In studies of borderline personality disorder, Soloff et al.<sup>[119]</sup> did not find consistent evidence that antipsychotics reduced the rate of suicide attempts, but they also did not find that they increased the risk.

#### 5. Discussion

Based on this review, there appears to be surprisingly little evidence in support of associations between drugs and suicidality. The FDA's metaanalysis of 372 adult studies failed to identify a positive association between antidepressants and suicide in adults over 24 years of age. In fact, significant decreases in suicidality rates were observed, the magnitude of which increased with age. Large-scale cohort studies confirm these results and indicate that the highest period of risk is prior to initiating antidepressant treatment. A signal for AEDs and suicidality (primarily ideation) has been identified by the FDA. Differences between AEDs and suicide attempts and completions have been found,<sup>[95]</sup> which differs from the FDA's conclusion that the effects on suicidality were generally similar across AEDs. Given the different indications for which AEDs are used, it is not surprising to find difference in suicidality rates between AEDs. In contrast, a large cohort study of bipolar patients<sup>[75]</sup> found no evidence of increased suicide attempt rates either within (before and after treatment) or between (treated vs non-treated) patients. These differences could be due to differences in outcomes studied (i.e. suicidal behaviour vs primarily ideation), duration of observation that is longer in the observational studies, and the exclusion of more suicidal patients from RCTs. With respect to the other drugs for which associations with suicide have been suspected (montelukast, varenicline, antipsychotics), the evidence is based largely on small numbers of spontaneous reports and has not been substantiated in either large-scale cohort studies or RCTs.

This leaves us with a possible association between antidepressants and suicidal thoughts and behaviour in children and young adults. For 18- to 24-year-olds, the FDA identified a small positive association but this was not confirmed in a large VA cohort study.<sup>[81]</sup> In children, the FDA identified a positive association from RCT data, which became borderline with the addition of two new studies overall and not significant within individual indications.<sup>[50]</sup> In the FDA analysis, prospective ratings of suicidal thoughts and behaviour did not show evidence of an association for either emergence or worsening of suicidal thoughts and behaviour. A large cohort study of depressed children also failed to identify a risk for antidepressants. Finally, large reductions in antidepressant prescriptions both in the US and around the world did not show any evidence of associated decreases in suicide rates. In fact, suicide rates in children increased following public health advisories in early 2004 that led to significant decreases in antidepressant prescription rates in children. No increase was observed in the population over age 60 years where antidepressant prescriptions continued to increase. If the association identified by the FDA is real, then it is likely restricted to ideation and does not increase the rate of suicide attempts or completion. Further prospective work is required to verify their earlier results.

From a statistical and methodological perspective, the field of pharmacoepidemiology is a fertile area for statistical research, both in theory and in application. This review has shown that, in general, methods have been adopted from other areas such as general epidemiology, despite the singular nature of many of the problems that are unique to drug safety in general and the study of rare events in particular. The use of person-time models – which adjust for the natural course of the disorder and its relationship to an adverse event of interest (e.g. suicidal events), regression effects and complex patterns of utilization - is a new method that appears to be uniquely suited to the analysis of pharmacoepidemiological data. Better methods are needed for large-scale screening of drug-AE interactions, which can then be confirmed using more intensive observational and/or experimental studies. Extending screening beyond the traditional spontaneous reporting system studies to large-scale integrated medical records and claims databases is certainly a fruitful direction for future research and practice. Applying and/or developing new methods for deriving causal inference from observational pharmacoepidemiological data should be a major priority for statistical research in this internationally important field of study.

Finally, there is considerable debate concerning the strength of evidence needed to take regulatory action on a potentially lethal ADR. It should be obvious that such evidence need not be as convincing as that required for efficacy; however, exactly where the threshold should lie for such actions remains unclear. For suicide, regulatory action has been taken largely on the basis of evidence suggesting increased risk of suicidal thoughts. However, suicidal thoughts are quite common among patients with depression and may have little relationship to suicidal behaviour and/or completion. The lack of positive impact that the black-box warning has had on suicide completion rates for children and adolescents in the US further suggests a disconnection between suicidal thoughts and completion. Better integration of both experimental and well analysed observational data based on more proximal endpoints (e.g. suicide attempts and completion) will help better articulate the risks of treatment and lead to a better identification of the threshold at which regulatory action should be taken. In addition, the prospective use of new and more comprehensive suicide measurement instruments<sup>[80]</sup> in the context of RCTs will also lead to improved detection of suicide-related ADRs. Further research on the communication of risk is critical.

## 6. Conclusions

Overall, there appears to be little evidence that drugs increase the risk of suicide and related behaviour. Numerous lines of evidence in adults clearly demonstrate that inadequate or no treatment of depression (pharmacotherapy and/or psychotherapy) is associated with increased risk of suicidal behaviour. In children, the results are less clear and further study is required to better delineate which children benefit from treatment and who may be at increased risk as a consequence of treatment. From a statistical and methodological perspective, the field of pharmacoepidemiology is a fertile area for statistical research, both in theory and in application. In general, methods have been adopted from other areas such as general epidemiology, despite the singular nature of many of the problems that are unique to drug safety in general and the study of rare events in particular. Finally, there is considerable debate concerning the communication of risk. For suicide, regulatory action has been taken largely on the basis of evidence suggesting increased risk of suicidal thoughts.

However, suicidal thoughts are quite common among patients with depression, and, especially in youth, may have a more distant relationship to suicidal behaviour and/or completion. The development of new experimental designs and analytical tools will help future studies specifically target suicidal behaviour and completion, and address this very important worldwide public health concern.

#### Acknowledgements

This work was supported by NIMH grants R01 MH8012201 (R.D. Gibbons) and MH062185 (J.J. Mann). Dr Gibbons has served as an expert witness for the US Department of Justice, Wyeth and Pfizer Pharmaceuticals on suicide-related cases. Dr Mann has received research support from GlaxoSmithKline and Novartis for unrelated brain imaging studies.

#### References

- Goldsmith SK, Pellmar TC, Kleinman AM, et al. Reducing suicide: a national imperative. Washington, DC: The National Academies Press, 2002; 1-516
- Brown CH, Wyman PA, Brinales JM, et al. The role or randomized trials in testing interventions for the prevention of youth suicide. Int J Psychiatry 2007; 19: 617-31
- Teicher MH, Glod C, Cole JO. Emergence of intense suicidal preoccupation during fluoxetine treatment. Am J Psychiatry 1990; 147: 207-10
- Baldessarini RJ, Tondo L, Hennen J. Effects of lithium treatment and its discontinuation on suicidal behavior in bipolar manic-depressive disorders. J Clin Psychiatry 1999; 60 Suppl. 2: 77-84
- Goodwin FK, Fireman B, Simon GE, et al. Suicide risk in bipolar disorder during treatment with lithium and divalproex. J Am Med Assoc 2003; 290:1467-73
- Thies-Flechtner K, Muller-Oerlinghausen B, Seibert W, et al. Effect of prophylactic treatment on suicide risk in patients with major affective disorders: data from a randomized prospective trial. Pharmacopsychiatry 1996; 29: 103-7
- Tondo L, Jamison KR, Baldessarini RJ. Effect of lithium maintenance on suicidal behavior in major mood disorders. Ann N Y Acad Sci 1997; 836: 339-51
- Tondo L, Baldessarini RJ, Hennen J, et al. Lithium treatment and risk of suicidal behavior in bipolar disorder patients. J Clin Psychiatry 1998; 59: 405-14
- Tondo L, Hennen J, Baldessarini RJ. Lower suicide risk with long-term lithium treatment in major affective illness: a metaanalysis. Acta Psychiatr Scand 2001; 104: 163-72
- Bowden RJ, Turkington DA. Instrumental variables. Cambridge: Cambridge University Press, 1984
- Brodersen A, Licht RW, Vestergaard P, et al. Sixteen-year mortality in patients with affective disorder commenced on lithium. Br J Psychiatry 2000; 176: 429-33
- Baldessarini RJ, Tondo L. Suicide risk and treatments for patients with bipolar disorder. J Am Med Assoc 2003; 290: 1517-9

- Isacsson G, Bergman U, Rich CL. Epidemiological data suggest antidepressants reduce suicide risk among depressives. J Affect Disord 1996; 41: 1-8
- Markowitz JC. Antidepressants and suicide risk [letter]. Br J Psychiatry 2001; 178: 477
- Ohberg A, Vuori E, Klaukka T, et al. Antidepressants and suicide mortality. J Affect Disord 1994; 50: 225-33
- Rich CL. Relationship between antidepressant treatment and suicide [letter]. J Clin Psychiatry 1999; 60: 340
- Rihmer Z, Rutz W, Pihlgren H, et al. Decreasing tendency of seasonality in suicide may indicate lowering rate of depressive suicides in the population. Psychiatry Res 1998; 81: 233-40
- Isacsson G. Suicide prevention: a medical breakthrough? Acta Psychiatr Scand 2000; 102: 113-7
- Olfson M, Shaffer D, Marcus SC, et al. Relationship between antidepressant medication treatment and suicide in adolescents. Arch Gen Psychiatry 2003; 60: 978-82
- Hall WD, Mant A, Mitchell PB, et al. Association between antidepressant prescribing and suicide in Australia, 1991-2000: trend analysis. BMJ 2003; 326: 1008-12
- Barbui C, Campomori A, D'Avanzo B, et al. Antidepressant drug use in Italy since the introduction of SSRIs: national trends, regional differences and impact on suicide rates. Soc Psychiatry Psychiatr Epidemiol 1999; 34:152-6
- 22. Rutz W. Preventing suicide and premature death by education and treatment. J Affect Disord 2001; 62: 123-9
- Grunebaum MF, Ellis SP, Li S, et al. Antidepressants and suicide risk in the United States, 1985-1999. J Clin Psychiatry 2004; 65 (11): 1456-62
- Isacsson G, Boëthius G, Bergman U. Low level of antidepressant prescription for people who later commit suicide: 15 years of experience from a population-based drug database in Sweden. Acta Psychiatr Scand 1992; 85: 444-8
- Oquendo MA, Malone KM, Ellis SP, et al. Inadequacy of antidepressant treatment for patients with major depression who are at risk for suicidal behavior. Am J Psychiatry 1999; 156: 190-4
- Isometsä E, Henriksson M, Heikkinen M, et al. Suicide and the use of antidepressants: drug treatment of depression is inadequate [letter]. BMJ 1994; 308: 915
- Henriksson S, Boethius G, Isacsson G. Suicides are seldom prescribed antidepressants: findings from a prospective prescription database in Jamtland county, Sweden, 1985-95. Acta Psychiatr Scand 2001; 103: 301-6
- Blazer DG, Hybel CF, Simonsick EM, et al. Marked differences in antidepressant use by race in an elderly community sample: 1986-1996. Am J Psychiatry 2000; 157: 1089-94
- Takahashi Y. Amidst a sharp increase of suicide: suicide in Japan. Tokyo: Kokoro no Kagaku (Nihon Hyoron-sha) 1999; 88: 2-10
- Nakagawa A, Grunebaum MF, Ellis SP, et al. Suicide and antidepressant prescription rates in Japan, 1999-2003. J Clin Psychiatry 2007; 68: 908-16
- Helgason T, Tomasson H, Zoega T. Antidepressants and public health in Iceland: time series analysis of national data. Br J Psychiatry 2004; 184: 157-62
- 32. Gibbons R, Hur K, Bhaumik D, et al. The relationship between antidepressant medication use and rate of suicide. Arch Gen Psychiatry 2005; 62: 165-72

- Gibbons R, Hur K, Bhaumik D, et al. The relationship between antidepressant prescription rates and rate of early adolescent suicide. Am J Psychiatry 2006; 163: 1898-904
- Hammad T. Review and evaluation of clinical data 2004 [online]. Available from URL: http://www.fda.gov/ohrms/ dockets/ac/04/briefing/2004-4065b1-10-TAB08-Hammads-Re view.pdf [Accessed 2011 Mar 1]
- Brent DA, Perper JA, Goldstein CE, et al. Risk factors for adolescent suicide: a comparison of adolescent suicide victims with suicidal inpatients. Arch Gen Psychiatry 1988; 45 (6): 581-8
- Shaffer D. Suicide: risk factors and the public health. Am J Public Health 1993; 83 (2): 171-251
- Leon AC, Marzuk PM, Tardiff K, et al. Paroxetine, other antidepressants, and youth suicide in New York City: 1993 through 1998. J Clin Psychiatry 2004; 65: 915-8
- Isacsson G, Holmgren P, Ahlner J. Selective serotonin reuptake inhibitor antidepressants and the risk of suicide: a controlled forensic database study of 14,857 suicides. Acta Psychiatr Scand 2005; 111: 286-90
- Gray D, Moskos M, Keller T. Utah Youth Suicide Study new findings. Presented at the Annual Meeting of the American Association of Suicidology; 2003 Apr 23-26; Santa Fe
- Anderson RN. Deaths: leading causes for 2000. National Vital Statistics Reports. Hyattsville (MD): National Center for Health Statistics 2002; 50 (16): 1-48
- Grunbaum JA, Kann L, Kinchen SA, et al. Youth risk behavior surveillance: United States, 2001. MMWR Surveill Summ 2002; 51: 1-62
- Fombonne E, Wostear G, Cooper V, et al. The Maudsley long-term follow-up of child and adolescent depression;
   2, suicidality, criminality and social dysfunction in adulthood. Br J Psychiatry 2001; 179: 218-23
- Weissman MM, Wolk S, Goldstein RB, et al. Depressed adolescents grown up. J Am Med Assoc 1999; 281: 1707-13
- Kovacs M, Goldston D, Gatsonis C. Suicidal behaviors and childhood-onset depressive disorders: a longitudinal investigation. J Am Acad Child Adolesc Psychiatry 1993; 32: 8-20
- Rao U, Weissman MM, Martin JA, et al. Childhood depression and risk of suicide: a preliminary report of a longitudinal study. J Am Acad Child Adolesc Psychiatry 1993; 32: 21-7
- Hazell P, O'Connell D, Heathcote D, et al. Tricyclic drugs for depression in children and adolescents. Cochrane Database Syst Rev 2002; (2): CD002317
- 47. Treatment for Adolescents with Depression Study (TADS) Team. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents With Depression Study (TADS) randomized controlled trial. J Am Med Assoc 2004; 292: 807-20
- Mann JJ, Emslie G, Baldessarini RJ, et al. ACNP Task Force report on SSRIs and suicidal behavior in youth. Neuropsychopharmacology 2005; 31: 473-92
- Emslie GJ, Rush AJ, Weinberg WA. A double-blind, randomized, placebo-controlled trial of fluoxetine in children and adolescents with depression. Arch Gen Psychiatry 1997; 54: 1031-7
- 50. Bridge JA, Iyengar S, Salary CB, et al. Clinical response and risk for reported suicidal ideation and suicide at-

tempts in pediatric antidepressant treatment. JAMA 2007; 297: 1683-96

- Gibbons R, Brown CH, Hur K, et al. Early evidence on the effects of the FDA black box warning on SSRI prescriptions and suicide in children and adolescents. Am J Psychiatry 2007; 164: 1356-63
- 52. Valuck RJ, Libby AM, Sills MR, et al. Antidepressant treatment and risk of suicide attempt by adolescents with major depressive disorder: a propensity-adjusted retrospective cohort study. CNS Drugs 2004; 18: 1119-32
- Simon GE, Savarino J, Operskalski B. Suicide risk during antidepressant treatment. Am J Psychiatry 2006; 163: 41-7
- Olfson M, Marcus SC, Shaffer D. Antidepressant drug therapy and suicide in severely depressed children and adults. Arch Gen Psychiatry 2006; 63: 865-72
- Tiihonen J, Lönnqvist J, Wahlbeck K, et al. Antidepressants and the risk of suicide, attempted suicide, and overall mortality in a nationwide cohort. Arch Gen Psychiatry 2006; 63: 1358-67
- 56. World Health Organization. Suicide prevention and special programmes (2003) [online]. Available from URL: http:// www.who.int/mental\_-health/prevention/suicide/ country\_reports/en/ [Accessed 2011 Mar 1]
- De Leo D, Dwyer J, Firman D, et al. Trends in hanging and firearm suicide rates in Australia: substitution of method? Suicide Life Threat Behav 2003; 33: 151-64
- DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA Spontaneous Reporting System. Am Stat 1999; 53: 177-90
- DuMouchel W, Pregibon D. Empirical Bayes screening for multi-item association. Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2001 Aug 26-29; San Francisco (CA), 67-76
- Sonesson C, Bock D. A review and discussion of prospective statistical surveillance in public health. J Royal Stat Soc A 2003; 166: 5-21
- 61. Gibbons RD, Segawa E, Karabatsos G, et al. Mixed-effects Poisson regression analysis of adverse event reports: the relationship between antidepressants and suicide. Stat Med 2008; 27: 1814-33
- Hammad TA, Laughren T, Racoosin J. Suicidality in pediatric patients treated with antidepressant drugs. Arch Gen Psychiatry 2006; 63: 332-9
- 63. Huber PJ. The behavior of maximum likelihood estimates under nonstandard conditions. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability; 1965 Jun 21-Juy 18; Berkeley (CA). Vol 1. Berkeley (CA): University of California Press, 1967: 221-33
- 64. White H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 1980; 48: 817-30
- 65. White H. Maximum likelihood estimation of misspecified models. Econometrica 1982; 50: 1-25
- Ludwig J, Marcotte D. Anti-depressants, suicide and drug regulation. J Policy Anal Manage 2005; 24: 249-72
- 67. Hedeker D, Gibbons RD. Longitudinal data analysis. New York: Wiley, 2006
- DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986; 7: 177-88

- Malzahn U, Bohning D, Holling H. Nonparametric estimation of heterogeneity variance for the standardized difference used in meta-analysis. Biometrika 2000; 87: 619-32
- Sidik K, Jonkman JN. Simple heterogeneity variance estimation for meta-analysis. Appl Stat (J Royal Stat Soc, Series C) 2005; 54: 367-84
- Sidik K, Jonkman JN. A comparison of heterogeneity variance estimators in combining results of studies. Stat Med 2007; 26: 1964-81
- Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. Stat Med 1998; 17: 841-56
- Mittlboeck M, Heinzl H. A simulation study comparing properties of heterogeneity measures in meta-analyses. Stat Med 2006; 25: 4321-33
- Rosenbaum P, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983; 70: 41-50
- Gibbons RD, Hur K, Brown CH, et al. The relationship between antiepileptics and suicide attempts in patients with bipolar disorder. Arch Gen Psychiatry 2009; 66: 1354-60
- Gibbons RD, Amatya AK, Brown CH, et al. Post-approval drug safety surveillance. Ann Rev Pub Health. 2010; 31: 419-37
- Rosenbaum PR. Differential effects and generic biases in observational studies. Biometrika 2006; 93: 573-86
- Efron B. Logistic regression, survival analysis, and the Kaplan Meier curve. J Am Stat Assoc 1988; 83: 414-25
- Gibbons RD, Duan N, Meltzer D, et al. Waiting for organ transplantation: results of an analysis by Institute of Medicine Committee. Biostatistics 2003; 4: 207-22
- Posner K, Oquendo MA, Gould M, et al. Columbia classification algorithm of suicide assessment (C-CASA): classification of suicidal events in the FDA's pediatric suicidal risk analysis of antidepressants. Am J Psychiatry 2007; 164: 1035-43
- Gibbons RD, Brown CH, Hur K, et al. The relationship between antidepressants and suicide: results of analysis of the Veterans Health Administration datasets. Am J Psychiatry 2007; 164: 1044-9
- 82. US Food and Drug Administration. Clinical review: relationship between antidepressant drugs and suicidality in adults [online]. Available from URL: http://www.fda.gov/ ohrms/dockets/ac/06/briefing/2006-4272b1-01-FDA.pdf [Accessed 2011 Mar 25]
- Pamer CA, Hammad TA, Wu YT, et al. Changes in US antidepressant and antipsychotic prescription patterns during a period of FDA actions. Pharmacoepidemiol Drug Saf 2010; 19: 158-74
- Rosack J. New data show declines in antidepressant prescribing. Psychiatr News 2005; 40: 1-6
- 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
- Centers for Disease Control and Prevention. Suicide trends among youths and young adults aged 10–24 years United States, 1990-2004. Morbidity and Mortality Weekly Review 2007; 56: 905-8

- Libby AM, Brent DA, Morrato EH, et al. Decline in treatment of pediatric depression after FDA advisory on risk of suicidality with SSRIs. Am J Psychiatry 2007; 164: 884-91
- Libby AM, Orton H, Valuck RJ. Persisting decline in depression treatment after FDA warnings. Arch Gen Psychiatry 2009; 66: 633-9
- 89. Wheeler BW, Gunnell D, Metcalfe C, et al. The population impact on incidence of suicide and non-fatal self harm of regulatory action against the use of selective serotonin reuptake inhibitors in under 18s in the United Kingdom: ecological study. BMJ 2008; 336: 542-7
- FDA 8/28/2009. Statistical review and evaluation: antiepileptic drugs and suicidality [online]. Available from URL: http://www.fda.gov/Safety/MedWatch/SafetyIn formation/SafetyAlertsforHumanMedicalProducts/ucm 166246.htm [Accessed 2008 May 23]
- Goodwin FK, Fireman B, Simon GE, et al. Suicide risk in bipolar disorder during treatment with lithium and divalproex. JAMA 2003; 290: 1467-73
- Collins JC, McFarland BH. Divalproex, lithium, and suicide among Medicaid patients with bipolar disorder. J Affect Disorder 2008; 107: 23-8
- Arana A, Wentworth CE, Ayuso-Mateos JL, et al. Suiciderelated events in patients treated with antiepileptic drugs. N Engl J Med 2010; 363: 542-51
- Gibbons RD, Hur K, Brown CH, et al. Gabapentin and suicide attempts. Pharmacoepidemiol Drug Saf 2010; 19: 1241-7
- Patorno E, Bohn RL, Wahl PM, et al. Anticonvulsant medications and the risk of suicide, attempted suicide, or violent death. JAMA 2010; 303: 1401-9
- Marcus S. Using omitted variable bias to assess uncertainty in the estimation of an AIDS education treatment effect. J Ed Behav Stat 1997; 22: 193-201
- Olesen JB, Hansen PR, Erdal J, et al. Antiepileptic drugs and risk of suicide: a national study. Pharmacoepidemiol Drug Saf 2010; 19: 518-24
- VanCott AC, Cramer JA, Copeland LA, et al. Suiciderelated behaviors in older patients with new anti-epileptic drug use : data from the VA hospital system. BMC Med 2010; 8: 1-7
- Bisgaard H, Skoner D, Boza ML, et al. Safety and tolerability of montelukast in placebo-controlled pediatric studies and their open-label extensions. Pediatr Pulmonol 2009; 44: 568-79
- Holbrook JT, Harik-Khan R. Montelukast and emotional well-being as a marker for depression: results from 3 randomized, double-masked clinical trials. J Allergy Clin Immunol 2008; 122: 828-9
- Jick H, Hagberg KW, Egger P. Rate of suicide in patients taking montelukast. Pharmacotherapy 2009; 29: 165-6
- Goodwin RD, Eaton WW. Asthma, suicidal ideation, and suicide attempts: findings from the Baltimore epidemiologic catchment area follow-up. Am J Public Health 2005; 95: 717-22
- 103. Hughes JR. Smoking and suicide: a brief overview. Drug Alcohol Depend 2008; 98: 169-78
- Hays JT, Ebert JO. Varenicline for tobacco dependence. N Engl JMed 2008; 359: 2018-24

- 105. Gunnell D, Irvine D, Wise L, et al. Varenicline and suicidal behavior: a cohort study based on data from the General Practice Research Database. BMJ 2009; 339: b3805
- 106. Hemmingsson T, Kriebel D. Smoking at age 18-20 and suicide during 26 years of follow-up: how can the association be explained? Int J Epidemiol 2003; 32: 1000-4
- 107. Miller M, Hemenway D, Rimm E. Cigarettes and suicide: a prospective study of 50 000 men. Am J Public Health 2000; 90: 768-73
- Cahill K, Stead L, Lancaster T. A preliminary benefit-risk assessment of varenicline in smoking cessation. Drug Saf 2009; 32: 119-35
- Nides M, Oncken C, Gonzalez D, et al. Smoking cessation with varenicline, a selective a4b2 nicotinic receptor partial agonist. Arch Intern Med 2006; 166: 1561-8
- Keating GM, Siddiqui MA. Varenicline: a review of its use as an aid to smoking cessation therapy. CNS Drugs 2006; 20: 945-60
- 111. Jorenby DE, Hays JT, Rigotti NA, et al. Efficacy of varenicline, an a4×2 nicotinic acetylcholine receptor partial agonist, vs placebo or sustained-release bupropion for smoking cessation. JAMA 2006; 296 (1): 56-63
- 112. Oncken C, Gonzalez D, Nides M, et al. Efficacy and safety of the novel selective nicotinic acetylcholine receptor partial agonist varenicline, for smoking cessation. Arch Intern Med 2006; 166: 1571-7
- 113. Tonstad S, Tonnesen P, Hajek P, et al. Effect of maintenance therapy with varenicline on smoking cessation: a randomized controlled trial. JAMA 2006; 296: 64-71

- 114. Kasliwal R, Wilton LV, Shakir SAW. Safety and drug utilization profile of varenicline as used in general practice in England. Drug Saf 2009; 32: 499-507
- Meltzer HY, Alphs L, Green AI, et al. Clozapine treatment for suicidality in schizophrenia. Arch Gen Psychiatry 2003; 60: 82-91
- 116. Khan A, Khan SR, Leventhal RM, et al. Symptom reduction and suicide risk among patients treated with placebo in antipsychotic clinical trials: an analysis of the Food and Drug Administration database. Am J Psychiatry 2001; 158: 1449-54
- 117. Herings RM, Erkens JA. Increased suicide attempt rate among patients interrupting use of atypical antipsychotics. Pharmacoepidemiol Drug Saf 2003; 12: 423-4
- 118. Montgomery SA, Montgomery DB, Rani SJ, et al. Maintenance therapy in repeat suicidal behaviour: a placebocontrolled trial. Proceedings of the Xth International Congress for Suicide Prevention and Crisis Intervention; 1979 June 17-20; Ottawa (ON), 227-9
- Soloff PH, Lis JA, Kelly T, et al. Self-mutilation and suicidal behavior in borderline personality disorder. J Personal Disord 1994; 8: 257-67

Correspondence: Professor *Robert D. Gibbons*, Departments of Medicine and Health Studies, and Director of the Center for Health Statistics, University of Chicago, 5841 S. Maryland Avenue, MC 2007 Office W260, Chicago, IL 60637, USA.

E-mail: rdg@uchicago.edu