How to read a research paper: Reading between and beyond the lines

Chittaranjan Andrade
Department of Psychopharmacology, National Institute of Mental Health and Neurosciences, Bangalore, Karnataka, India

ABSTRACT

Background: Despite peer review, publications in scientific journals are not always well written, sometimes contain errors, and often exhibit deliberate or unintended biases. It is necessary to learn how to identify such limitations. It is also necessary to learn how to read between and beyond the lines of papers no matter how well written they are and no matter how highly ranked the journal is.

Materials and Methods: This paper critically examines an important article in a leading journal with a view to help the reader learn how to place the findings of a study in perspective, understand its limitations, and glean information beyond that actually presented and discussed in the text.

Results: Several issues are examined; these relate to case–control research designs, confounding, propensity matching, absolute risk, confidence intervals, interpretation of findings, real-world relevance, ecological validity, and definition of a cause–effect relationship.

Conclusions: The issues examined in this paper reflect common themes in research, and a reader aware of these themes will more easily identify them in his future readings.

Key words: Absolute risk, case–control study, confidence intervals, confounding, how to read a research paper, pregnancy, selective serotonin reuptake inhibitor, teratogenicity

INTRODUCTION

Published papers vary in standards. Some papers contain obvious errors; in others, the errors are less apparent. Some papers contain biases related to the pharmaceutical industry; these may be deliberate, when the manuscript was published by the industry, or unconscious, when the manuscript was written by a scientist who had earlier received industry funding. Some papers contain biases related to the previous research and views of the authors; again, these may be deliberate or unconscious, such as when the authors steer the reader toward their own pet lines of thought. Peer review helps improve the quality of a submitted manuscript and correct biases. Nevertheless, despite peer review, and despite the best of intentions, even well-written papers in highly ranked journals may yield conclusions that are not immediately apparent, or suffer from limitations that are not obvious to the casual reader.

In the present paper, an important article selected from a leading journal is critically examined with a view to help the reader learn how to place the findings of a study in perspective, understand its limitations, and derive information beyond the findings that were actually presented and discussed in the article. Emphasis is laid on themes that recur in the medical literature in order to arm the reader with tools that will enhance his future reading skills. This paper is part of a series for postgraduate students, academicians, and readers who wish to develop their skills in reading and writing scientific articles. [1-6]
MATERIALS AND METHODS

The use of antidepressant medications during pregnancy is a subject of considerable importance. In a potentially influential selective serotonin reuptake inhibitor (SSRI)-related malformations study published in a leading journal,[7] the authors concluded as follows:

“Fluoxetine use is associated with an increased risk of isolated ventricular septal defects and paroxetine is associated with right ventricular outflow tract defects. The absolute risk for these specific cardiac anomalies is small but should guide clinicians not to consider fluoxetine or paroxetine the first option when prescribing selective serotonin reuptake inhibitors to women planning pregnancy. Special attention should be given to alcohol use in pregnant women using selective serotonin reuptake inhibitors.”[7]

The authors were circumspect in their wording, but the emotional impact of their caveats can be large in the minds of readers. How can readers better understand the findings of the study, place the findings in perspective and, perhaps, derive additional messages from the text? Some suggestions are offered in the next section of this paper.

Whereas the reader is encouraged to examine the full text of the SSRI malformation study[7] before continuing with this paper, a summary of the study[7] is provided below.

Malm et al.[7] used linked database information on 635,583 mother–child pairs to identify and compare offspring with and without exposure to SSRIs during the first trimester of pregnancy. There were 6881 SSRI-exposed women (mean age, 30 years) with 6976 offspring. There were 618,727 SSRI-unexposed control women (mean age, 29 years) with 628,607 offspring. Malformation ascertainment included data from early termination of pregnancy. SSRI use was assumed from records on the purchase of SSRI drugs. Analyses were adjusted for important confounders: pregnancy year, maternal age, parity, marital status, smoking, use of other psychotropics, and prepregnancy diabetes.

Important findings reported in the SSRI malformation study[7] are summarized in Table 1. In a nutshell, the authors found that:
1. SSRI exposure during the first trimester of pregnancy did not increase the overall risk of major congenital malformations
2. The above notwithstanding, specific SSRIs increased the risk of specific malformations.

RESULTS

This section presents a critical examination of the SSRI malformation study.[7] For the sake of simplicity, it is assumed here that the SSRIs would have been prescribed for their most common indication, depression, and that the controls, who did not receive SSRIs, would be much less likely to have depression. The arguments expressed here will not change if the SSRIs had also been prescribed for other psychiatric conditions.

The research design of the SSRI malformation study[7] illustrates confounding by indication with depression (as the indication for SSRI prescription) possibly predisposing to the adverse outcome rather than the SSRI itself. Depression, of course, does not cause malformations, but its behavioral and related accompaniments might. For example, depression is associated with eating disorders and poor nutrition, vulnerability to infection, poor adherence to medical regimens necessitated by the pregnancy, medical comorbidities and the treatments thereof, risk-taking behavior, self-harm, and so on; any of these could result in fetal harm and even fetal malformations. In contrast, the control group would be much less likely to display these vulnerabilities. Therefore, to compare the SSRI and non-SSRI groups would be akin to comparing apples and oranges, fruit which are different by definition and hence non-comparable because the definition already establishes the difference.

Confounding by indication arose from the study design because SSRI-exposed depressed women were compared with unexposed non-depressed women; that is, SSRI use and depression were entangled, and non-use and absence of depression were likewise entangled as the grouping variable. In this study, as well as in other studies that use similar designs, there is no way of knowing to what extent some or all of the results are confounded by indication. An accurate estimate of risk can only come from randomized controlled trial (RCT) data where depression and SSRI use are unentangled; that is, where depressed women are randomized to SSRI or placebo. However, such a study would not be ethical in pregnancy and will almost certainly never be conducted. The next best alternative is a propensity matching design (this is discussed in the next section), but propensity matching is a poor cousin of the RCT.

From Table 1, it is seen that there were several different ways in which women who used SSRI drugs differed from those who did not. Many of the ways in which the groups were different (such as smoking, presence of chronic diseases, and use of other psychotropic and non-psychotropic drugs) had the potential to influence the outcome variable, fetal malformations. Importantly, even if the groups had not differed in these regards, the apples and oranges situation would still exist because the depressed patients would probably have been fundamentally different from those who were not depressed in ways in which the study did not record (examples have already been provided earlier in this section). In statistical analysis which attempts to control
However, the data reported by the authors[7] showed that fivefold increase in right ventricular outflow tract defects. Under discussion,[7] paroxetine was associated with a nearly 10-fold increase in cardiovascular anomalies (OR, 1.40; 95% CI, 1.01–1.95), including isolated ventricular septal defects (OR, 2.03; 95% CI, 1.28–3.21); paroxetine with an increased risk of right ventricular outflow tract defects (OR, 4.68; 95% CI 1.48–14.74); and citalopram with neural tube defects (OR, 2.46; 95% CI, 1.20–5.07). Other organ system risks and specific malformation risks did not differ significantly between SSRI-exposed and unexposed offspring.

From Table 1, we observe that certain SSRIs did increase the risk of major malformations associated with SSRI use for an unbiased estimate of the effect of the grouping variable (SSRI use) on the outcome variable (fetal malformations). Moving to the next finding in Table 1, the upper bound of the 95% confidence interval (CI) was 1.22 for malformations associated with any first trimester SSRI exposure. What additional information can we glean from this datum? Even if SSRIs do increase the risk of major malformations, we can be 95% certain that the risk is elevated by 22% and not more. Data from the full text of the paper indicate that the highest upper bound of the 95% CI for individual SSRIs was 1.64 with paroxetine. This means that we are 95% certain that the malformation risk after first trimester paroxetine exposure is elevated by 64% or less. So, if the malformation risk in the general population is 2%, we are 95% confident that it is less than 2.44% with SSRIs as a group (22% elevated), and less than 3.28% with paroxetine, in particular (64% elevated).

From Table 1, we observe that certain SSRIs did increase the risk of specific malformations. However, the absolute risk of these malformations is low because these malformations are rare in the general population. For example, in the study under discussion,[7] paroxetine was associated with a nearly fivefold increase in right ventricular outflow tract defects. However, the data reported by the authors[7] showed that only 3 of 968 paroxetine exposures resulted in such defects; that is, the absolute risk with paroxetine was just 0.3%.

From Table 1, the 10-fold increased risk of fetal alcohol spectrum disorders associated with SSRI exposure does not mean that SSRI exposure predisposes to alcohol abuse and/ or fetal alcohol spectrum disorders. Rather, the underlying psychiatric condition (depression) was likely responsible for both SSRI prescription and the alcohol use which led to the fetal alcohol spectrum disorder in the offspring.

In this study, SSRI use was assumed from purchases of medication. There is no assurance, however, that women who purchased a drug actually took it during pregnancy. As it is now general knowledge that first trimester use of medication can harm the fetus, many women may have exercised caution and refrained from actually consuming the medication. In other words, as explained earlier, the identified findings may have been due to the underlying condition (depression and its accompaniments) rather than the drugs (SSRIs).

This critique notwithstanding, the data do provide real-world estimates of the risks associated with SSRI use. For example, the unadjusted risk of major malformations after SSRI exposure was statistically significant [odds ratio (OR), 1.24; 95% CI, 1.10–1.39]. This means that a depressed woman who uses an SSRI during the first trimester of pregnancy is 24% more likely, on average, to bear a child with a major malformation relative to a woman who does not use an SSRI. The malformation may be due to the SSRI, to a fetotoxic accompaniment of the depression, or to an interaction between these.

Lastly, but still importantly, depressed women probably behave differently in different parts of the world. For example, women in India are less likely to drink and smoke and more likely to receive psychological and material support during pregnancy than women in the country in which the SSRI malformation study was conducted. Therefore, as in all real-world studies, the findings of the SSRI malformation study[7] are ecologically valid only for the population from which the data were obtained.

**DISCUSSION**

**Propensity matching** is a procedure that is sometimes used in case–control studies to address the apples and oranges problem in lieu of an RCT solution. In propensity-matched studies, cases and controls are matched on illness variables

| Table 1: Important findings reported in the SSRI malformations study[7] |
|-------------------------------------------------|
| - Relative to women who did not purchase SSRIs during pregnancy, women who purchased these drugs were less likely to be married, more likely to smoke, more likely to have chronic diseases, and more likely to use non-SSRI psychotropic drugs as well as non-psychotropic drugs |
| - The risk of major congenital anomalies was 4.3% versus 3.6% in SSRI users versus non-users. This dropped to 3.9% versus 3.2% after eliminating offspring with identified chromosomal anomalies (which could not have resulted from drug exposure), who comprised 0.4% of each group. The 0.7% higher risk of major malformations associated with SSRI use was not statistically significant after adjusting for confounding variables (OR, 1.08; 95% CI, 0.96–1.22) |
| - Among individual SSRIs, none was associated with a significantly increased risk of major congenital anomalies |
| - When the risk of specific malformations was examined, fluoxetine was associated with an increased risk of cardiovascular anomalies (OR, 1.40; 95% CI, 1.01–1.95), including isolated ventricular septal defects (OR, 2.03; 95% CI, 1.28–3.21); paroxetine with an increased risk of right ventricular outflow tract defects (OR, 4.68; 95% CI 1.48–14.74); and citalopram with neural tube defects (OR, 2.46; 95% CI, 1.20–5.07). Other organ system risks and specific malformation risks did not differ significantly between SSRI-exposed and unexposed offspring |
| - The risk of fetal alcohol spectrum disorders was nearly 10-fold higher in SSRI-exposed cases (OR, 9.6; 95% CI, 4.6–20.0). |
that predispose to the grouping variable (antidepressant prescription, in the present study).[7] Thus, for example, for each SSRI case, a non-SSRI control woman could be selected who resembles the case on variables such as past history of depressive illness, duration of current depressive episode, severity of current depressive episode, and so on. It is hoped that through such matching, residual confounding would be addressed. Whereas propensity matching is an improvement on a simple case–control design, it comes at the cost of sample size (because it is not easy to find matching pairs) and anyway does not have the same validity as an RCT. Propensity matching designs have already been used in the antidepressant-pregnancy field of research.[8]

This critical appraisal should not be construed to imply that SSRIs are safe in pregnancy. There is a reasonable body of evidence which indicates that SSRIs increase the risk of spontaneous abortion, neonatal withdrawal symptoms, and other adverse outcomes.[9]

Rather than interpret the contents of this paper as a critical appraisal of a specific article, the reader is encouraged to apply the same lines of thought when appraising other studies too. Thus, the points that emerged in the appraisal of the SSRI malformation study[7] can serve as tools that will help the reader glean information and draw conclusions beyond those that may actually appear in the presented study. As a simple example, a non-randomized, naturalistic comparison of depressed patients who receive antidepressant drugs and depressed patients who receive electroconvulsive therapy (ECT) may find greater cognitive impairment in the latter patients. This prompts the question of whether at least a part of the greater cognitive impairment in the latter patients arises from greater severity of illness, given that patients with more severe illness would be more likely to receive ECT. For a more dramatic and amusing example, the reader is referred to an earlier article on confounding.[10]

With regard to the issue of ecological validity, readers should question whether the findings of real-world studies such as CATIE[10] and STAR*D[11] are relevant to third-world countries such as India. Both these studies would certainly have been confounded by higher environmental stress, lower social and family support, poorer medication supervision, and other variables, all of which had the potential to alter clinical outcomes. All these confounding variables were part of the sociocultural milieu in which CATIE and STAR*D were conducted, and all these variables are quite different in third-world countries.

A final, though general, point is that in an observational study such as that of Malm et al.,[7] if a statistically significant relationship between an independent (e.g. SSRI use) and a dependent (e.g. teratogenicity) variable does not necessarily mean that the former causes the latter, when would a causal relationship be considered likely? Possible criteria are listed in Table 2. The larger the number of criteria met, the greater the likelihood of a cause–effect relationship.[12] The best evidence, of course, is that which comes from RCTs. For example, individuals who are physically and intellectually active are known to be at lower risk of Alzheimer's disease. Two explanations are possible, which are that physical and intellectual activity protects against Alzheimer's disease and that individuals whose brains are less affected by pre-Alzheimer's pathology are more likely to remain physically and intellectually active. The only way to resolve the issue is to randomize elderly individuals to high and low levels of physical and intellectual activity and to ascertain the incidence of Alzheimer's disease in the different subgroups after a reasonable period of time.

REFERENCES

1. Andrade C. Confounding. Indian J Psychiatry 2007;49:129-31.
2. Andrade C, Radhakrishnan R. Prayer and healing: A medical and scientific perspective on randomized controlled trials. Indian J Psychiatry 2009;51:247-53.
Andrade: How to read a research paper

3. Andrade C, Mendhekar DN. Lithium, trifluperazine, and idiopathic leucopenia: Author and reviewer perspectives on how to write a good case report. Indian J Psychiatry 2010;52:187-90.

4. Andrade C. A 6-week, multicentre, randomized controlled clinical trial to evaluate the safety and efficacy of placeboetine hydrochloride in the treatment of major depressive disorder in an Indian setting. Indian J Psychiatry 2011;53:69-72.

5. Andrade C. Lithium, trifluperazine, and idiopathic leucopenia: Author and reviewer perspectives on how to write a good case report. Indian J Psychiatry 2010;52:187-90.

6. Andrade C. A 6-week, multicentre, randomized controlled clinical trial to evaluate the safety and efficacy of placeboetine hydrochloride in the treatment of major depressive disorder in an Indian setting. Indian J Psychiatry 2011;53:69-72.

7. Malm H, Artama M, Gissler M, Ritvanen A. Selective serotonin reuptake inhibitors and risk for major congenital anomalies. Obstet Gynecol 2011;118:111-20.

8. Oberlender TF, Warburton W, Misri S, Aghajanian J, Hertzman C. Neonatal outcomes after prenatal exposure to selective serotonin reuptake inhibitor antidepressants and maternal depression using population-based linked health data. Arch Gen Psychiatry 2006;63:898-906.

9. Andrade C. Postgraduate corner: Continuing medical education: SSRIs and pregnancy. Indian J Psychiatry 2010;52:83-6.

10. Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, et al. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. New Engl J Med 2005;353:1209-23.

11. Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. for the STAR*D Study Team. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. Am J Psychiatry 2006;163:28-40.

12. Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet 2002;358:248-52.

Source of Support: Nil, Conflict of Interest: None declared