Psychological Medicine (2011), **41**, 1127–1131. © Cambridge University Press 2010
doi:10.1017/S0033291710001595

ORIGINAL ARTICLE

The use of well controls: an unhealthy practice in psychiatric research

S. Schwartz[1]* and E. Susser[1,2]

[1] *Mailman School of Public Health, Columbia University, New York, NY, USA*
[2] *New York State Psychiatric Institute, New York, NY, USA*

Background. Studies comparing cases with controls to uncover the causes of psychiatric disorders are common in biological research. The validity of these studies depends upon adherence to the methodological principles underlying the case-control design. However, these principles are often violated. One common practice that violates these principles is the use of well controls. In this paper we describe the bias that it can cause and discuss why the use of well controls leads to invalidity in case-control studies.

Method. Using hypothetical numerical examples we illustrate the consequences of using well controls.

Results. The results illustrate that the use of well controls can cause substantial bias. In no instance does the use of well controls improve validity.

Conclusions. We conclude that the use of well controls is an unhealthy practice in psychiatric research.

Received 4 January 2010; Revised 9 June 2010; Accepted 17 July 2010; First published online 1 September 2010

Key words: Case-control studies, control selection, well controls.

Introduction

Comparisons between cases and controls are ubiquitous in biological psychiatry. When properly applied, this approach is scientifically valid and efficient. The differences between cases and controls can lead to discoveries about the causes and the nature of the illness under study.

Yet this approach is not always properly applied. In biological psychiatry, investigations of causes that compare cases and controls are not always perceived or labeled as 'case-control' studies and their conduct often does not conform to the principles developed for valid case-control studies (Lee *et al*. 2007). Because the comparison between cases and non-cases is so intuitive, it may seem unnecessary to have an elaborated methodology. But our intuitions sometimes lead us astray. Nowhere is this more apparent than in the use of well controls in psychiatric research.

By 'well controls' (also referred to as supernormal or hypernormal controls) we mean controls who are accepted into a case-control study with exclusion criteria that are different from, and stricter than, those applied to cases. What makes them 'well' controls is that the exclusion criteria include the presence of disorders other than the one under investigation. For example, in a study of major depressive disorder, a well control group would be one that excludes individuals with panic disorder from the control group but not the case group. We do this to distinguish carefully 'cases', individuals who are ill, from 'controls', individuals who are normal.

Our intuitions about well controls are so strong that, despite several papers and textbook chapters that have demonstrated the bias that can result from this practice (Schwartz & Link, 1989; Kendler, 1990; Susser *et al*. 2006), the use of well controls continues. There are many methodological problems that plague case-control studies, but well controls are particularly troubling because the practice is not only widespread but also often recommended as an appropriate method to reduce bias (e.g. Adami *et al*. 2002; Schechter & Levobitch, 2005; Talati *et al*. 2008). Unfortunately, however, the use of well controls does not improve validity, is costly and has the potential to create significant bias.

To address this issue we use numeric examples to demonstrate how the use of well controls may be an important source of artifact in biological studies. The recognition and remedy of this practice could help to advance this field of research.

* Address for correspondence: S. Schwartz, Ph.D., Mailman School of Public Health, Columbia University, 722 West 168th Street, Room 720b, New York, NY 10032, USA.
(Email: sbs5@columbia.edu)

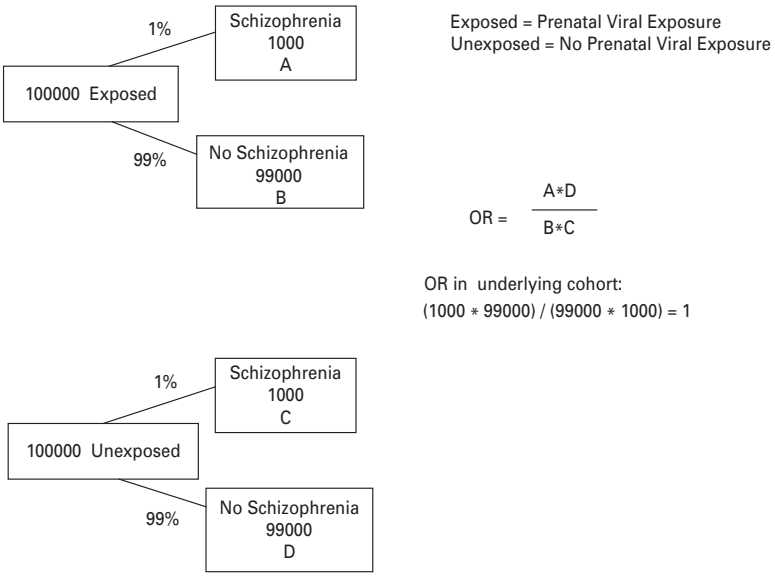

Fig. 1. Odds ratios in underlying cohort, case-control study: no effect of exposure on the outcome. (Adapted with permission of Oxford University Press from *Psychiatric Epidemiology* by Susser *et al*. 2006, p. 251.)

Developing a valid case-control study

A modern understanding of a case-control study conceptualizes it as an efficient way to sample an underlying cohort of exposed and unexposed people, some of whom develop the disease of interest. This is most clearly seen in a nested case-control study, a case-control study developed from an underlying cohort that is known and delineated.

This process is illustrated in an hypothetical study of prenatal viral exposure as a cause of schizophrenia (Fig. 1), an example that we build on throughout this paper. Imagine a population of 200 000 infants, all of whom were classified as having experienced prenatal viral exposure (exposed) or not (unexposed) based on serologic tests of maternal blood samples. This population yielded an underlying cohort (a 'nest') of 100 000 exposed and 100 000 unexposed infants. Imagine further that we observe these infants through the age of risk for schizophrenia with no loss to follow-up. The disease status of the cohort members is indicated in the boxes in the second column. In this instance 1% of the exposed individuals develop schizophrenia (box A) and the remaining 99% do not (box B). Similarly, among the unexposed, 1% develop schizophrenia (box C) and 99% do not (box D). The odds ratio calculated from this underlying cohort (as shown in the right-hand column) is 1, indicating that there is no association between prenatal viral exposure and schizophrenia in this sample.

Although this study provides a correct estimate of the effect of prenatal viral exposure on schizophrenia in this population, it does so very inefficiently. The study includes a very large number of non-diseased individuals (boxes B and D).

A more efficient method to arrive at the same answer would be to conduct a case-control study derived from this underlying cohort. A researcher could attempt to identify all of the cases (through a psychiatric case registry for example) and then randomly select, some proportion, for instance 20%, of the non-diseased as controls. This process is shown in Fig. 2. The cases are the exposed and unexposed individuals who become diseased (cells a and c depicted in the grey box with solid borders). The controls are a random 20% sample of the exposed and unexposed individuals who did not develop disease (cells b and d depicted in the grey box with dashed borders). The odds ratio based on this case-control study (shown in the right column) is also 1. The case-control study yielded the same odds ratio as the underlying cohort, but with far fewer people. We have used large sample sizes for illustrative purposes, but we would recreate the odds ratio of the underlying cohort, within sampling error, no matter what proportion of the non-diseased we selected as controls as long as they were selected independent of exposure status.

Results using well controls

We now contrast the results in Fig. 2 with what happens when we use well controls. Suppose a researcher was concerned that the null results obtained in the study in Fig. 2 were due to a failure to adequately screen controls for other psychiatric disorders, for

https://doi.org/10.1017/S0033291710001595 Published online by Cambridge University Press

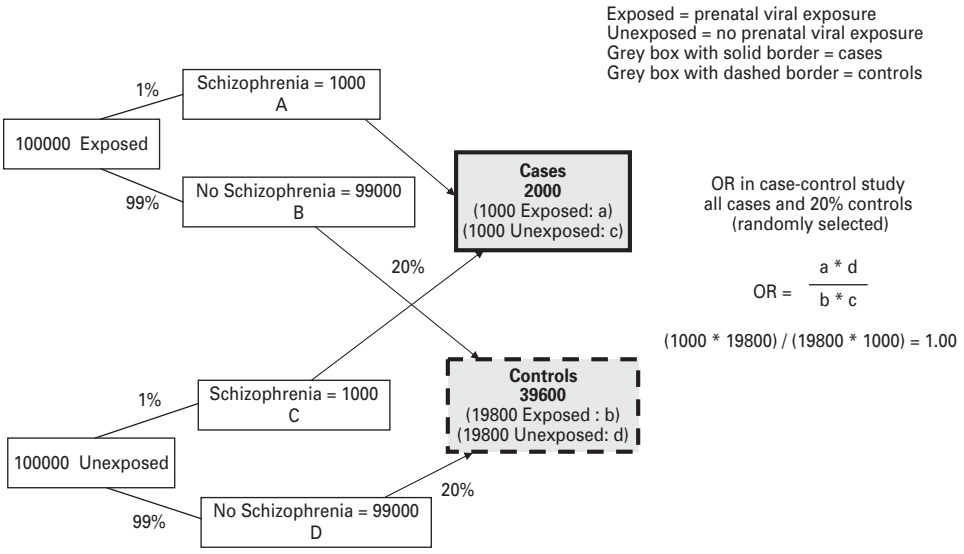

Fig. 2. Developing a case-control study from an underlying cohort. (Adapted with permission of Oxford University Press from *Psychiatric Epidemiology* by Susser *et al*. 2006, p. 251.)

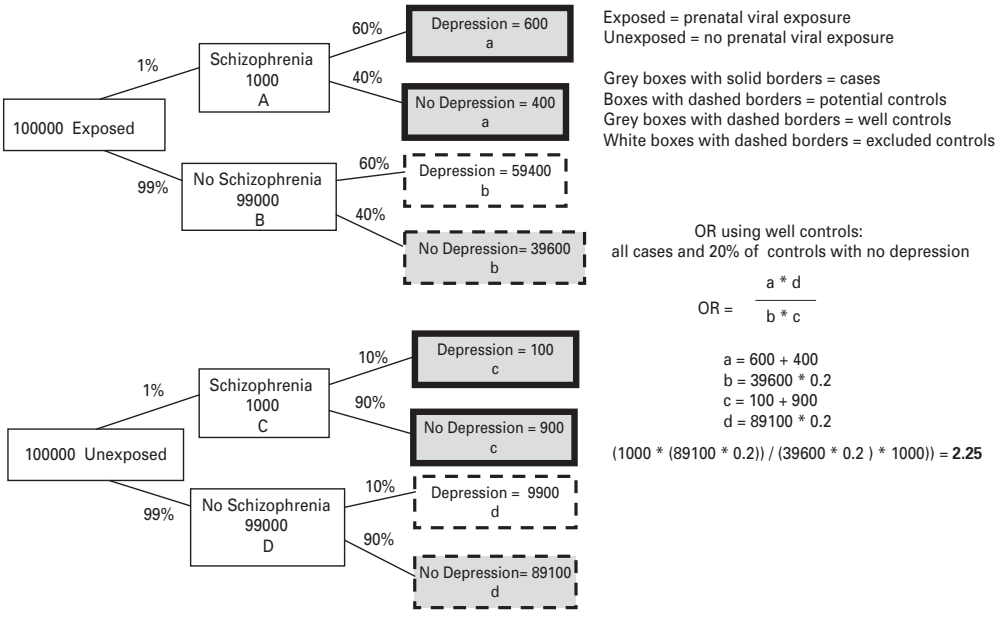

Fig. 3. Odds ratios in case-control study using well controls: no effect of exposure on schizophrenia. Exposure is associated with depression. (Adapted with permission of Oxford University Press from *Psychiatric Epidemiology* by Susser *et al*. 2006, p. 251.)

example depression. For this reason, they suggest that we use as controls only individuals who developed neither schizophrenia nor depression. In this example, let us assume that depression is associated with the exposure, but that schizophrenia is not. In other words, the original estimate of 1 was correct. This scenario is illustrated in Fig. 3.

As in Figs 1 and 2, we posit that the risk of schizophrenia among both the exposed and unexposed is 1%. As the exposure is associated with depression, 60% of the exposed become depressed but only 10% of the unexposed. As schizophrenia is not associated with depression, this is true for both those who do and do not develop schizophrenia.

The cases from this underlying cohort are in the grey boxes with solid borders (cells a and c). We show two a cells and two c cells to represent the exposed cases (a) and the unexposed cases (c) with and without depression. The potential controls (those without schizophrenia) are depicted in the boxes with dashed borders (cells b and d). The grey boxes with dashed borders depict the well controls (those without

https://doi.org/10.1017/S0033291710001595 Published online by Cambridge University Press

depression) and the white boxes with dashed borders, those who would be excluded from the controls (those with depression). In a case-control study using well controls we might select all of those with schizophrenia as the cases and some proportion, say 20%, of those without schizophrenia and without depression as the well controls (i.e. selecting cases and controls from the grey boxes). The odds ratio calculated from this case-control study using well controls is 2.25. The true odds ratio is 1, a reflection of the equal risk for schizophrenia in the exposed and unexposed participants. Using well controls has created an association between the exposure and the schizophrenia that was entirely an artifact of the improper selection of controls. Because depression is associated with the exposure, the controls were not sampled independently of the exposure, violating a key methodological principle of case-control studies.

In our example, we made the assumption that the exposure did not cause the disease and that the disease under study was not co-morbid with the excluded disorder. However, the bias created by the use of well controls is not contingent on these assumptions. The bias will be created whether or not there is a true effect of the exposure, and whether or not the excluded disorder is co-morbid with the disorder under investigation. Any time the excluded disorder is associated with the exposure under study, it will create bias. Under no circumstances does it improve validity (Susser *et al*. 2006).

Why does using well controls create bias?

Conceptualizing a case-control study as a condensed version of a cohort study is the foundation for the principles guiding the appropriate selection of controls.

In the valid nested case-control study depicted in Figs 1 and 2, people who develop the disease of interest are included as cases in the study and a random sample of those without the disease of interest are selected as controls. Thus the controls represent people who, if they had developed the disorder of interest, would have been included as cases in the study. The exposure experience of these people represents the exposure experience of the non-diseased in the cohort from which the cases arose. Whatever the definition of the disorder under investigation, those who meet criteria for the disorder are cases and those who do not meet criteria are eligible to be controls. All other inclusion and exclusion criteria are the same for cases and controls.

In our example, a traditional case-control study, controls are selected from the underlying cohort at the end of the study period. Controls can also be selected over the course of the cohort study or at the beginning of the study (Rothman *et al*. 2008; Susser *et al*. 2006). The principles we describe here apply to these types of case-control studies as well.

This same logic applies to a non-nested case-control study with the additional complexity that the underlying cohort that gave rise to the cases is not known with certainty and researchers must use a thought experiment, prone to error, to conceptualize this underlying 'nest'. In attempts to do this, investigators sometimes choose controls from a sample of people treated at the same facility as the cases or individuals living in the same geographic area. Which approach is correct depends on the selection processes that lead to the identification of the cases in that particular study. (For a detailed discussion of these issues in the context of psychiatric disorders, see Susser *et al*. 2006.)

Conceptualizing and sampling correctly from the underlying cohort does not ensure a valid effect estimate. It may be that there is confounding in the underlying cohort itself; that is, the exposed and unexposed differ on causes of disease other than the exposure of interest. For example, in a cohort study of stressful life events and depression, we might be concerned that substance use, a factor that may cause both stressful life events and depression, may be a confounder in our study. If so, we would need to adjust for this variable in the cohort study. Any variable that causes confounding in the underlying cohort will also cause confounding in the case-control study derived from this cohort. Therefore, if we conducted a nested case-control study in this cohort, we would need to adjust for these confounding variables. In addition, temporal order may be more opaque in a case-control study and recall bias and other methodological errors can arise.

Conclusion

With all the potential sources of bias in a biologic case-control study, why do we focus on the use of well controls? We do so because the use of well controls is a common, and often recommended, method to select controls (Adami *et al*. 2002; Schechter & Levobitch, 2005; Talati *et al*. 2008). Yet it is time-consuming and expensive, can cause considerable bias and does not improve study results. For these reasons we think it is worthy of attention.

Researchers may use well controls because they are concerned that including people with other disorders in their control group will make it difficult to see the true effects of the exposure on their outcome of interest. That is, they reason that making cases and controls

https://doi.org/10.1017/S0033291710001595 Published online by Cambridge University Press

more distinct would increase validity. In selecting controls, therefore, researchers may exclude not only people with the disorder under investigation but also those with any Axis I disorder. Sometimes they also exclude people with any physical disorder, a high score on a screening scale of psychiatric symptoms, or with a first-degree relative who has an Axis I disorder.

If we consider the principles of the case-control design, we can see the flaws in this approach. The principle that the same inclusion and exclusion criteria should apply to the cases and controls (other than the presence of the disorder under study) argues against such practices. As we noted, this is because controls should represent the people from the underlying cohort who gave rise to the cases. Therefore, if individuals with other disorders are eligible to be cases if they develop the disease under study, they should also be eligible to be controls if they do not develop the disease under study. The inclusion of such individuals does not cause confounding. However, as demonstrated above, their exclusion does cause selection bias.

Often biological researchers are concerned that other conditions present in the controls will mask or distort the biologic measure of the exposure under study. For example, heavy alcohol use might have diffuse effects on the brain. But excluding heavy alcohol users from the controls does not solve this problem and is likely to cause bias. However, if both cases and controls are excluded on this basis, there is no bias. For example, if the researchers are uncertain about the diagnostic criteria, they could exclude those with indistinct presentations from both the case and control groups. Such a case-control study would yield an accurate effect estimate for this study group.

In attempting to develop valid case-control studies, we are always faced with the real-world constraints. We need to balance the possible against the ideal and make compromises in our studies. Fortunately, the decision against using well controls blends the practical and the ideal; it provides a more efficient, less expensive and more valid case-control study.

Acknowledgments

We thank D. Barnes, D. Freedman and the reviewers for helpful comments on earlier versions of this paper.

Declaration of Interest

None.

References

Adami H, Elliott A, Zetlmeisl M, McMahon R, Thaker G (2002). Use of telephone screens improves efficiency of healthy subject recruitment. *Psychiatric Research* **113**, 295–301.

Kendler KS (1990). The super-normal control group in psychiatric genetics: possible artifactual evidence for coaggregation. *Psychiatric Genetics* **1**, 45–53.

Lee WJ, Bindman J, Ford T, Glozier N, Moran P, Stewart R, Hotopf M (2007). Bias in psychiatric case-control studies: literature survey. *British Journal of Psychiatry* **190**, 204–209.

Rothman KJ, Greenland S, Lash TL (2008). *Modern Epidemiology*, 3rd edn. Lippincott-Raven: Philadelphia.

Schechter D, Levobitch R (2005). Normal controls are expensive to find: methods to improve cost-effectiveness of the screening evaluation. *Psychiatric Research* **136**, 69–78.

Schlesselman JJ (1982). *Case-Control Studies: Design, Conduct, Analysis*. Oxford University Press: New York.

Schwartz S, Link BG (1989). The 'well control' artefact in case/control studies of specific psychiatric disorders. *Psychological Medicine* **19**, 737–742.

Susser E, Schwartz S, Morabia A, Bromet EJ (2006). *Psychiatric Epidemiology: Searching for the Causes of Mental Disorders*. Oxford University Press: New York.

Talati A, Fyer SJ, Weissman MM (2008). A comparison between screened NIMH and clinically interviewed control samples on neuroticism and extraversion. *Molecular Psychiatry* **13**, 122–130.