Epidemiology Publish Ahead of Print

DOI: 10.1097/EDE.0000000000001088

Test-negative designs: differences and commonalities with other case–control studies with 'other patient' controls

Jan P Vandenbroucke (1), Neil Pearce (2)

(1) Jan P Vandenbroucke, [Corresponding author] , Leiden University Medical Center, Dept. Clinical Epidemiology, PO Box 9600, 2300 RC Leiden, The Netherlands, E-mail: J.P.Vandenbroucke@lumc.nl, Tel: +31-71-5264037; Fax: +31-71-5266994; And: Professor of Clinical Epidemiology, Aarhus University Denmark; Honorary Professor, London School of Hygiene and Tropical Medicine, UK

(2) Neil Pearce, Department of Medical Statistics, London School of Hygiene and Tropical Medicine, UK

Word count: about 5041 (without abstract, references, or appendix)

Running Title: Test-negative designs, case–control, patient controls

Disclosure of Conflict of Interest: the authors declare to have no financial conflicts of interest.

Funding: This work was supported by the European Research Council under the European Union's Seventh Framework Programme [FP7/2007-2013 / ERC grant agreement number 668954]. The Centre for Global NCDs is supported by the Wellcome Trust Institutional Strategic Support Fund (097834/Z/11/B).

Acknowledgments: We acknowledge our debt to comments on a previous version of this paper by Ivo Foppa, Mike Jackson, Marc Lipsitch , Sheena Sullivan, Lorenzo Richiardi, Ken Rothman and Kate Tilling. All opinions as well as remaining inadequacies are our own.

Abstract

Test-negative studies recruit cases who attend a healthcare facility and test positive for a particular disease; controls are patients undergoing the same tests for the same reasons at the same healthcare facility and who test negative. The design is often used for vaccine efficacy studies, but not exclusively, and has been posited as a separate type of study design, different from case–control studies because the controls are not sampled from a wider source population. However, the design is a special case of a broader class of case–control designs that identify cases and sample 'other patient' controls from the same health care facilities. Therefore, we consider that new insights into the test-negative design can be obtained by viewing them as case–control studies with 'other patient' controls; in this context, we explore differences and commonalities, to better define the advantages and disadvantages of the test-negative design in various circumstances. The design has the advantage of similar participation rates, information quality and completeness, referral/catchment areas, initial presentation, diagnostic suspicion tendencies, and preferences by doctors. Under certain assumptions, valid population odds ratios can be estimated with the test-negative design, just as with case–control studies with 'other patient' controls. Interestingly, directed acyclic graphs are not completely helpful in explaining why the design works. The use of test-negative designs may not completely resolve all potential biases, but they are a valid study design option, and will in some circumstances lead to less bias, as well as often being the most practical option.

Keywords: Methods, test-negative design, case–control studies, patient controls, vaccine efficacy

Introduction

Test-negative studies are frequently used in epidemiology, particularly in the context of vaccine effectiveness or antibiotic resistance,[1-8] but also in studies of risk factors for various conditions such as venous thrombosis[9] or Reye syndrome.[10] Such studies involve recruiting cases who attend a healthcare facility and test positive for a particular disease; controls are then patients undergoing the same tests for the same reasons at the same healthcare facility and who test negative. Although this approach is widely used, there is some debate and confusion about what type of study this is, and what the potential advantages, disadvantages, and biases are. We will argue that the test-negative design can be seen as belonging to a wider class of case–control studies that use 'other patient controls' (i.e. other attendees of the same or similar healthcare facilities). This perspective clarifies the status of test-negative studies as a design option, clarifies discussions about potential biases, and offers additional insights into its strengths and weaknesses.

History and terminology

To the best of our knowledge, Miettinen was among the first to describe in full the basic idea of test-negative case–control studies in his 1985 "Theoretical Epidemiology" [See[11] page 79]: He argued that "An idealized reference series, as a sample of that base would consist of a phenocopy of the illness under study, occurring independently of the determinant". He explained that the persons or events making up this reference series would come under attention in the same way as the cases, but that a last-minute differential diagnosis would separate them from the cases. That reference series would share all the selectivity of the cases, and therefore would validly estimate the distribution of the exposure in the source population for the cases.

Thus, Miettinen essentially describes these as case–control studies, albeit with a non-standard control group.

In the literature about test-negative designs for vaccine effectiveness, the 1980 study on "Pneumococcal Disease after Pneumococcal Vaccination" by Broome et al.[12] is often credited for being the first test-negative vaccine effectiveness study. As quoted from the abstract of that paper, the design aimed to compare the frequency of different serotypes in "...35 isolates of *Streptococcus pneumonia* isolated from blood or cerebrospinal fluid one month or longer after the patient had received commercially available pneumococcal vaccine with serotypes of 392 isolated from unvaccinated persons surveyed in a study of the nationwide distribution of pneumococcal serotypes." Both series consisted of body material samples of persons with pneumococcal disease, and thus represented incident cases of disease. The analysis (See the equations in the Methods section of Broome et al.[12]) is a comparison between the incidence of "pneumococcal disease with serotypes present in the vaccine" between two cohorts of *unknown* size: one vaccinated and one not. However, the relative size of the unknown cohorts is approximated by the relative numbers of persons with pneumococcal disease with "serotypes not present in the vaccine", emanating from these two cohorts. This approach is commonly known as a Proportional Mortality (or Morbidity) Ratio study in occupational epidemiology, (see[13] page 72) and can be rewritten as a Morbidity Odds Ratio analysis (the latter is essentially a case–control study sampling from a 2x2 table comparing exposed and unexposed cohorts[14]). This proportional estimation was later referred to as an 'indirect cohort analysis' in a study on pneumococcal vaccine efficacy by Shapiro et al.[15]

In recent reviews and publications about vaccine effectiveness of influenza vaccination, the 'test-negative design' is posited as a separate entity, possibly different from case–control studies.[1, 4, 16] A focused discussion on a particular design in a particular application (such as vaccine effectiveness) has benefits, as it is rooted in and enriched by subject matter knowledge. On the other hand, we consider that the test-negative design has commonalities with a broader class of designs that identify cases and sample 'other patient' controls from the attendees of the same or similar health care facilities. In this general class of case–control studies, controls are intended to be sampled from the catchment population for a healthcare facility; this is achieved by sampling controls from persons who present at the same or similar facilities but have a different disease. Common examples of this approach include hospital-based case–control studies (like the first case–control studies on smoking and lung-cancer[17-19]), or case–control studies in which cases and controls are chosen from the same disease register (e.g. cancer case–control studies with other cancer controls).[20]

There are various reasons for using other patient controls.[20, 21] Sometimes this is the most practical option, if cases can only be identified through a particular healthcare facility, and the catchment area for the facility cannot be enumerated; other patient controls therefore represent a practical option. Other reasons include the need to address several difficult-to-define and difficult-to-control selection biases, such as diagnostic referral bias; i.e., if not all relevant cases can be identified, and there are selection processes which influence case identification, then other patient controls may be used if it is considered that they will have gone through the same selection processes. Finally, a further reason can be to reduce recall bias, such as in studies of

cancer or congenital malformations where other cancers or other malformations are used as controls, sometimes in a sensitivity analysis.[22]

Test-negative designs are a particular type of such case–control studies, the distinctive feature being that they sample controls from persons who present themselves with *similar signs and symptoms* (to those of the cases) to the *same health care facilities*, but in which the case-disease of interest is ruled out by the test. In this situation, the (test-negative) controls will have other reasons for their signs and symptoms (e.g., in vaccine effectiveness studies of influenza, controls may have been infected with other viruses that cause flu-like symptoms), but may be assumed to have gone through a similar selection and referral process to that of the (test-positive) cases. Test-negative sampling of controls has been applied to studies of widely diverse topics, such as oral contraceptives and venous thrombosis,[9] Reye syndrome and aspirin use,[10] and risk factors for antibiotic resistance.[2]

In this paper, we will discuss test-negative designs in this broader framework of case–control studies with other patient controls, to define more clearly their benefits and their drawbacks in several circumstances. We will argue that test-negative studies have their own characteristics and potential advantages, disadvantages and biases, but they also have much in common with other case–control studies which use other patient controls. The consideration of test-negative studies can therefore be enhanced, and the potential biases clarified, by considering the broader literature on such studies. We will focus on biases that may occur in test-negative studies in comparison with the gold standard of an ideal population-based case–control study with complete case-ascertainment, random sampling of controls (possibly conditional on certain characteristics), correct classification of exposure, and 100% response rates. Such a case–control study would be seen as nested in a cohort study with ‘missing data at

random and by design'(see[23] and[24] page 112). Implicit in this design is that it involves a primary definition of the source population,[11] i.e., a population that is defined beforehand and from which cases emanate and controls are selected over calendar time (in the past or in the future). In focusing on the potential differences with this design, we will not consider potential problems that the test-negative design may have in common with all case–control studies. Thus, our aim is to discuss how a test-negative study can yield the same findings as would have been obtained with the gold-standard approach; we acknowledge that even a gold-standard study may yield biased findings (e.g. due to residual confounding) in comparison with an idealized randomized controlled trial, but we will not consider these more general issues here. Also, we will refrain from discussing issues specific for test-negative designs for influenza vaccine effectiveness studies because these issues have been discussed extensively elsewhere[1, 4-8], and because any exposure–disease association studied by test-negative designs will have its own particularities. Rather we focus on the general case–control aspects of test-negative designs.

General principles of case–control studies with other patient controls

We will briefly consider the general issues of case–control studies with 'other patient' controls before considering the specific issues of test negative studies. We will do so in the context of the traditional classification of bias into selection bias, information bias, and confounding.[24]

Selection bias

Selection bias can occur in case–control studies due to biased selection of cases and/or controls from the source population, either by choices of the investigators, or in the pathways that make patients present themselves in particular health care settings, or the pathways to diagnosis that arise in those health care settings.

If all of the cases generated by the source population over the risk period (i.e., from an idealized cohort) are included in a case–control study then there can be no selection bias with regards to selection of cases. However, selection bias can occur if only some cases are identified, and this is related to the exposure under study.

Selection bias can also occur due to biased selection of controls. According to standard theory, controls should be sampled (conditionally) at random from the source population and risk period, and should be free of and at risk of the study disease at the time of sampling – thereby the controls should represent the exposure distribution of the source population for cases. However, if the cases are identified through a healthcare facility or group of facilities (e.g. a particular hospital, or groups of hospitals, or in a series of different outpatient clinics or GP practices), then the source population (the catchment or referral area of the facilities) may be difficult to enumerate and may have socio-economic, lifestyle and/or health characteristics that differ from the general population. Choosing controls from the general population may thus lead to bias with regard to the estimation of the prevalence of exposure(s) in the true source population. Therefore controls are often chosen from patients referred to the same facility for symptoms or diseases with similar referral patterns (i.e. patients who would have come to that hospital for the case-disease). This approach was used, for example, in several of the first case–control studies of lung cancer and smoking that included other patients in the same wards or other cancer patients.[17-19] This approach has continued to be used in cancer epidemiology, exploiting the existence of cancer registries.[20] It has also been advised and commonly used in case–control studies in pharmacoepidemiology.[25]

A problem also arises if a subgroup of cases with specific characteristics is targeted for investigation of those characteristics, for example, in studies of risk factors for urinary tract infections with bacterial strains that are resistant to particular antibiotics. If one were to sample controls from the general population, the case–control analysis might mix up general risk factors for urinary tract infections with specific risk factors for infection with a resistant strain. Therefore, such studies are often done by using controls who also have urinary tract infections, but with sensitive strains: cases and controls have the same disease by the same general agent (e.g., urinary tract infection by *E. coli*) but with different antibiotic sensitivity characteristics.[26] As antibiotic sensitivity does not influence the virulence of the bacteria (their propensity to be infective and/or invasive), both groups will have similar risk factors for urinary tract infections. However, over and above these general risk factors, the cases will have additional risk factors for being differentially infected with resistant strains.

A further potential problem with biased selection of controls involves self-selection mechanisms, leading to differences in participation rates. While persons who acquired a particular disease and become cases may quite readily participate in a study about past exposures, this may not be true for general population controls. In contrast, other patient controls may more readily participate, given that they might have similar concerns to the cases about the causes of their disease or their signs and symptoms.

Information bias

Even if all cases (or a random samples of cases) are identified from the source population and risk period, and controls are sampled in an unbiased manner, bias can still occur because of differences in the validity of exposure information in cases and controls. A particular concern is that persons with disease may have a much better recall of potentially ‘bad’ exposures than healthy persons who are asked about similar

exposures ‘out of the blue’. This is a common experience in studies of events during pregnancy and the occurrence of congenital malformations: mothers who deliver a child with a congenital malformation may recall and report many possible causes that occurred during the pregnancy, while mothers who delivered a seemingly healthy baby may have less recall of anything untoward during their pregnancy. That is the reason that investigators have used mothers with babies with a different congenital malformation as ‘other patient’ controls.[22]

Confounding

Confounding can occur in all epidemiologic studies, and is not specific to studies involving other patient controls. However, studies with other patient controls can be used in innovative ways to control for types of confounding that are difficult to specify or measure. An example arises from studies of specific asthma medications as a cause of asthma mortality. A series of case–control studies in New Zealand sampled asthma deaths as cases, while controls were patients who were admitted to hospital with a non-fatal, but apparently severe enough asthma attack to warrant hospitalisation. The reason for using ‘hospital controls’ was to achieve ‘indirect matching’ on asthma severity, since there was prior evidence that asthmatics who died and asthmatics who were admitted to hospital with a non-fatal attack were similar with regards to chronic asthma severity, which is a major determinant of prescribing.[27] On the contrary, it was known that asthmatics who were never admitted to hospital generally had a lower chronic severity. The contrast was therefore between asthmatics who had a severe attack and died (often without making it to hospital), and asthmatic who had a severe attack and survived long enough to make it to hospital and did not die.

The meaning of the odds ratio with 'other patient controls'

As the theory of 'modern' case–control studies has evolved, it has become generally accepted that the controls should reflect the exposure prevalence (past or recent or cumulative, etc) of the source population that generated the cases over the relevant risk period.[28] This permits a straightforward interpretation of case–control odds ratios as rate ratios, risk ratios, or prevalence odds ratios, depending on how the controls were selected.[29]

Patients with other diseases are mentioned as potential sources of controls in several standard publications on case–control studies (see[21] and[24] page 119). Patients with other diseases can be seen as sampled from a 'secondary base', i.e., a source population that is defined only *after* cases are sampled, and by the way the cases are sampled, e.g., the catchment population of a medical facility or a registry (See[11] page 54-55). Patients with other conditions will reflect that source population. The term 'secondary' distinguishes from the 'primary' base, mentioned above when describing the 'gold standard'.

A caveat that was repeatedly emphasized in the literature[11, 20, 21, 25] was that the control condition(s) should not be associated positively or negatively with the exposure under study. This is obvious nowadays: one would not, for example, consider using patients with other lung conditions as controls in a study of smoking and lung cancer as many such conditions are also associated with smoking. In the early days it was not that obvious. For example, Wynder et al.[17] in their epochal 1950 study on smoking and lung cancer used controls from a chest disease ward for a sensitivity analysis about possible interviewer bias (interviewers were not aware of the diagnosis and all patients had chest disease); they were perhaps lucky to find a strong association with smoking nevertheless. Careful control selection always was an

important consideration in pharmacoepidemiologic case–control studies wherein it was proposed that control diseases should neither be positively nor negatively associated with the drug exposure under investigation.[25] And if there is an association, the important consideration then becomes whether the design would underestimate or overestimate exposure in the controls with respect to that in the source population.

Specific characteristics of test-negative studies: opportunities and drawbacks

In test-negative studies, controls are enrolled from those who present with similar signs and symptoms, preferably sampled from a time frame before the final diagnosis; and usually from the same health care facilities. Given that the controls have similar signs and symptoms but not the disease of interest, there will be other reasons for their signs and symptoms, such as other diseases. In a broader sense, the 'test' in the 'test-negative design' is a procedure to achieve a diagnosis, and the controls can also be viewed as 'negative to the diagnosis' of the case disease of interest; thus the design might also be described as the 'same symptom – different diagnosis' design.

We will consider several examples of test-negative studies in the light of the general principles of using other patient controls. We will consider the potential biases that test-negative study designs try to remedy, as well as potential biases that they may create. In each instance, some potential biases relate to the selection and/or classification of the cases, or the controls, or to both.

The *potential selection biases* that test-negative studies intend to address include diagnostic suspicion and/or referral mechanisms that affect case-ascertainment. In the study of oral contraceptives as a cause of venous thrombosis, the argument was used that persons with milder signs and symptoms would only be diagnosed or referred for diagnostic evaluation if they used oral contraceptives, thus biasing the estimate of case exposure in comparison with that of all cases in the source population. One early

solution was to stratify for severity of illness, on the assumption that case-ascertainment would be almost complete in the group with very severe signs and symptoms, and that oral contraceptive use therefore would not affect referral in this subgroup. From such stratification/restriction it was found that the odds ratio was even higher in persons with severe symptoms (which was the inverse of what would be expected if the hypothesized selection bias existed).[30] A more recent solution was to sample cases and controls from the diagnostic facility to which they were referred in order to rule out venous thrombosis by imaging.[9, 31] The cases were persons in which the suspicion was confirmed by imaging; the controls were those in which the imaging did not confirm the suspicion (in those patients there would be other reasons for their signs and symptoms). Thus, all cases and controls would be referred with similar reasons to suspect venous thrombosis (i.e., similar signs and symptoms). In these more recent studies that used a 'test-negative' design, the elevated odds ratio of venous thrombosis with oral contraceptives was confirmed. The earliest approach to remedy this potential selection bias involved 'general population' controls, with the cases being stratified or restricted on severity; the later approach involved using 'other patient' controls. Both approaches yielded similar results, which were of the same magnitude as that of the original studies in which the bias might allegedly have occurred – and in which no precaution against this bias was taken.

A second and related mechanism of biased case-ascertainment relates to 'health-seeking behavior'. This possibility was raised in studies of influenza vaccine effectiveness. It is possible that persons who present themselves with milder forms of flu-like symptoms may be persons who are more prone to seeking health care, and that these persons would also be more prone to have been vaccinated for influenza virus. Test-negative designs have been used in which persons presenting themselves

with potential influenza to a health care facility were virologically tested as to whether they were infected with an influenza virus or not. The former become (test-positive) cases, the latter become (test-negative) controls.[1] This makes it clear that the distribution of health seeking behavior amongst these cases and controls should be similar: some would present with very severe symptoms (and the problem of health seeking behavior is moot), while others would present with mild symptoms; in the latter, there may be health seeking behavior related to influenza vaccinations.[1, 4, 5] Sullivan et al.[1] also consider the gradient of health seeking behavior, which in their opinion opens the possibility that the test-negative design would not completely block the effect of health seeking behavior. However, this would only apply if the average level (and range) of health seeking behavior was different in those who are test positive and those who are test negative. This is not different from the situation when health seeking behavior is a dichotomous variable. Indeed, one can imagine some 'hypochondriac' persons who will rush to a doctor with very mild symptoms, while others, somewhat more robust, will visit the doctor only when they have severe symptoms. Amongst the 'robust', some persons with mild influenza symptoms will not be identified because they did not present themselves, but the same would happen if they had a mild non-influenza viral infection and therefore would be potential controls. Thus, in a test-negative design the controls will certainly meet a kind of counterfactual definition of 'controls [that] would have been identified as cases had they developed the health outcome of interest during the period of observation of the study base.'[32] Self-selection in presenting oneself to a health care facility is in fact a general issue in case–control studies wherein a visit to a health care facility is needed in order to be included in a study.

This reasoning also makes it clear that it does not help to have better diagnostic criteria for the case-condition to circumvent this selectivity in self-presentation. Of course, good case-definitions, can be important for other purposes, such as reducing misclassification. Stratification according to severity of symptoms is an analytic strategy described above for *standard* case control studies, such as the studies of oral contraceptives and venous thrombosis, as a safeguard against diagnostic suspicion bias.[30] As explained above, test-negative design on the same association made the potential bias equally strong amongst cases and controls.[9, 31]

Selection bias and confounding were of particular concern in the study of aspirin as a potential cause of Reye syndrome in children.[10] A noteworthy aspect of this example is that exposure to a medication can be a consequence of early symptoms. In this situation, choosing controls from other attendees at the same healthcare facility with similar signs and symptoms can have methodologic advantages. Reye syndrome is a rare but potentially lethal complication of some viral infections in children (in particular chickenpox) characterized by failure of several organs (liver, brain). Based on case series and initial case–control studies, it was postulated that the risk of the syndrome was highly increased if children with such viral diseases received aspirin to alleviate the signs and symptoms of that viral disease. This hypothesis was met with strong skepticism because it was held that giving aspirin to children with symptoms of viral origin would be a standard medical therapy, i.e. that there would be confounding by indication. Because of that concern, a case–control comparison was set up with controls being chosen from children with alternate conditions, i.e., children who also had been seen with similar initial symptoms in the same health care facilities. In those children the diagnosis of Reye syndrome might have been considered, but their disease was judged not to have evolved into Reye syndrome. That study very strongly

confirmed the association of aspirin and Reye syndrome. A nice analytic feature was the restriction to children who had either received aspirin or paracetamol for their viral symptoms; the odds ratio was 36 (95% CI 4.2 – 288), which was of a similar order of magnitude as found in other studies, among which one with several different control groups.[33] The study tried thus to mimic a randomized experiment in children with viral diseases: either give paracetamol or aspirin.

Another example of potential confounding by health seeking behavior, as well as confounding by indication, occurred in a study that assessed protein pump inhibitors as risk factors for acquiring a urinary infection with an *E. coli* strain with multiple resistance to antibiotics, instead of with a strain that is generally sensitive to antibiotics.[26] A potential problem is that antibiotic cultures are not carried out in many patients with signs and symptoms of urinary tract infection: when it is a first or an occasional urinary tract infection (in particular in women), often standard antibiotic treatment is given without cultures. On the other hand, patients with persistent or recurrent urinary tract infection and/or pre-existing medical diseases are more likely to have urine cultures taken and tested for antibiotic resistance; the same patients may also use protein pump inhibitors much more often. Because all associations are positive, including only patients with urine cultures in the study may have induced a selection bias that diminishes the association with protein pump inhibitors, which can be controlled (or reduced) by conditioning on (aspects of) medical history.[26]

Discussion

It is interesting to note that in the examples, several potential selection mechanisms are addressed at once by the test-negative design. The design has the advantage of similar participation rates, similar information quality and completeness, similar

referral/catchment areas, similar initial presentation, similar diagnostic suspicion tendencies, and finally similar preferences or choices by doctors.

Westreich et al., in a commentary entitled: "Beware of the test-negative design",[16] echo the view that test-negative designs are not really case–control studies as controls are not sampled at random from the source population. However, as noted above, there is a broad class of case–control studies in which controls are sampled from other health care attendees within the source population, and the test-negative design falls within this class. Also, it has been argued that the test-negative design is a cohort with some incomplete follow-up.[1, 34] However, there is no follow-up happening in a test-negative design: case or control status is determined at a single point in time: the time of the test.

As mentioned, the study of vaccine effectiveness that is considered the first used the reasoning of a Proportional Morbidity Ratio analysis (which one might recast as a case–control analysis) sampled from two cohorts.[12] As described above, under certain assumptions, valid population odds ratios can be estimated with controls that are not sampled from the source population, e.g. with other patient controls. The views echoed by Westreich et al. are at odds with standard case–control theory (see [21] and [24] page 119) and with several reviews that have upheld the validity of the test-negative design,[4, 5, 35] amongst others by simulation studies and probability modelling.

According to Jackson et al.,[5] "With the assumptions that (a) the distribution of non-influenza causes of acute respiratory infections does not vary by influenza vaccination status, and (b) vaccine effectiveness does not vary by health care-seeking behavior, the vaccine effectiveness estimate from the sample can generalized to the full source population that gave rise to the study sample." This statement is logical if one sees it in the light of a case–control design with 'patient controls'. Shi et al.[7] also add the

qualification that "the confounding effects resulting from non-random vaccination are similar for influenza and non-influenza acute respiratory infections" – we agree with this additional criterion with the proviso that it is the residual confounding effects (after adjustment for known confounders) that are relevant here.

The general caveat that the control condition should not be associated with the exposure, either positively or negatively is not often invoked in discussions of test-negative designs, in contrast with the more general literature on case–control studies with other patients as controls. In pharmacoepidemiology there is often careful consideration of which patients can or cannot be included as 'other patient controls'. In contrast, in test-negative designs there is usually little effort to know what other diseases might have caused the same signs and symptoms in the controls that tested negative for the case disease. This seems to be an under-evaluated aspect of test-negative designs for vaccine effectiveness. As an example of the subtleties involved, in the instance of studies of venous thrombosis and the risk of travel a test-negative design found no association with long-haul flights, in contrast to a host of other well-done case–control and cohort studies. [36] The reason might be that long-haul flights may also cause leg symptoms such as edema, without an underlying venous thrombosis; however, leg edema is also possible warning sign of venous thrombosis. Thus a referral for a patient with leg edema might be influenced by the history of a recent long flight and will yield controls with a similar flying history as the cases. We learn from this example that it may be a particular challenge to identify a phenocopy with the same symptoms, but which is independent of the determinant; however, this is not a qualitatively different issue than faced by any other 'secondary base' case–control study.

In outlining the arguments in this paper, we have not made use of directed acyclic graphs (DAGs), for reasons fully explained in the e*Appendix; http://links.lww.com/EDE/B579*, as we felt that the test-negative design involves potential biases that are not readily represented in DAGs, and that the previously published DAGs each had their merits but were wanting in other respects, in particular because the DAGs do not explain why the design works. Indeed, the test-negative design may reduce certain selection and confounding biases, but the biases may not completely disappear; the common problem with DAGs is that they do not address the quantitative strength of the associations that are potential biases and that are diminished in a test-negative design.

In this paper, we have described the test-negative design as belonging to a family of similar designs where cases and controls are selected from patients from the same or similar health care facilities. It represents a subtype of this approach, where controls are patients with similar clinical signs and symptoms who have tested negative for the 'case disease' in a further lab-based or imaging procedure.

These studies can give valid causal estimates of odds ratios in the source population, albeit only under certain assumptions. The use of such designs may not completely resolve all potential biases, but will in some circumstances lead to less bias, as well as being the most practical option.

REFERENCES

1. Sullivan SG, Tchetgen Tchetgen EJ, Cowling BJ. Theoretical Basis of the Test-Negative Study Design for Assessment of Influenza Vaccine Effectiveness. Am J Epidemiol. 2016;184(5):345-353.

2. Lipsitch M. Measuring and interpreting associations between antibiotic use and penicillin resistance in Streptococcus pneumoniae. Clin Infect Dis. 2001;32(7):1044-1054.

3. Schechner V, Temkin E, Harbarth S, Carmeli Y, Schwaber MJ. Epidemiological interpretation of studies examining the effect of antibiotic usage on resistance. Clin Microbiol Rev. 2013;26(2):289-307.

4. Foppa IM, Ferdinands JM, Chaves SS, Haber MJ, Reynolds SB, Flannery B, et al. The case test-negative design for studies of the effectiveness of influenza vaccine in inpatient settings. Int J Epidemiol. 2016;45(6):2052-2059.

5. Jackson ML, Nelson JC. The test-negative design for estimating influenza vaccine effectiveness. Vaccine. 2013;31(17):2165-2168.

6. Jackson ML, Phillips CH, Benoit J, Kiniry E, Madziwa L, Nelson JC, et al. The impact of selection bias on vaccine effectiveness estimates from test-negative studies. Vaccine. 2018;36(5):751-757.

7. Shi M, An Q, Ainslie KEC, Haber M, Orenstein WA. A comparison of the test-negative and the traditional case–control study designs for estimation of influenza vaccine effectiveness under nonrandom vaccination. BMC Infect Dis. 2017;17(1):757.

8. Lewnard JA, Tedijanto C, Cowling BJ, Lipsitch M. Quantifying biases in test-negative studies of vaccine effectiveness. bioRxiv. 2017.

9. Bloemenkamp KW, Rosendaal FR, Buller HR, Helmerhorst FM, Colly LP, Vandenbroucke JP. Risk of venous thrombosis with use of current low-dose oral contraceptives is not explained by diagnostic suspicion and referral bias. Arch Intern Med. 1999;159(1):65-70.

10. Forsyth BW, Horwitz RI, Acampora D, Shapiro ED, Viscoli CM, Feinstein AR, et al. New epidemiologic evidence confirming that bias does not explain the aspirin/Reye's syndrome association. JAMA. 1989;261(17):2517-2524.

11. Miettinen OS. Theoretical Epidemiology. Principles of Occurence Research in Medicine.: John Wiley and Sons; 1985.

12. Broome CV, Facklam RR, Fraser DW. Pneumococcal disease after pneumococcal vaccination: an alternative method to estimate the efficacy of pneumococcal vaccine. N Engl J Med. 1980;303(10):549-552.

13. Checkoway H, Pearce N, Kriebel D. Research methods in occupational epidemiology. 2nd ed. New York: Oxford University Press; 2004. xiv, 372 p. p.

14. Miettinen OS, Wang JD. An alternative to the proportionate mortality ratio. Am J Epidemiol. 1981;114(1):144-148.

15. Shapiro ED, Berg AT, Austrian R, Schroeder D, Parcells V, Margolis A, et al. The protective efficacy of polyvalent pneumococcal polysaccharide vaccine. N Engl J Med. 1991;325(21):1453-1460.

16. Westreich D, Hudgens MG. Invited Commentary: Beware the Test-Negative Design. Am J Epidemiol. 2016;184(5):354-356.

17. Wynder EL, Graham EA. Tobacco smoking as a possible etiologic factor in bronchiogenic carcinoma; a study of 684 proved cases. J Am Med Assoc. 1950;143(4):329-336.

18. Levin ML, Goldstein H, Gerhardt PR. Cancer and tobacco smoking; a preliminary report. J Am Med Assoc. 1950;143(4):336-338.

19. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. Br Med J. 1950;2(4682):739-748.

20. Smith AH, Pearce NE, Callas PW. Cancer case–control studies with other cancers as controls. Int J Epidemiol. 1988;17(2):298-306.

21. Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection of controls in case–control studies. II. Types of controls. Am J Epidemiol. 1992;135(9):1029-1041.

22. Rothman KJ. Epidemiology, an introduction New York: Oxford University Press; 2002, pp 98-99.

23. Wacholder S. The case–control study as data missing by design: estimating risk differences. Epidemiology. 1996;7(2):144-150.

24. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology, 3rd edition. Philadelphia: Wolters Kluwer, Lippincott, Williams & Wilkins; 2008.

25. Jick H, Vessey MP. Case–control studies in the evaluation of drug-induced illness. Am J Epidemiol. 1978;107(1):1-7.

26. Sogaard M, Heide-Jorgensen U, Vandenbroucke JP, Schonheyder HC, Vandenbroucke-Grauls C. Risk factors for extended-spectrum beta-lactamase-producing Escherichia coli urinary tract infection in the community in Denmark: a case–control study. Clin Microbiol Infect. 2017;23(12):952-960.

27. Pearce N, Hensley MJ. Epidemiologic studies of beta agonists and asthma deaths. Epidemiol Rev. 1998;20(2):173-186.

28. Vandenbroucke JP, Pearce N. Case–control studies: basic concepts. Int J Epidemiol. 2012;41(5):1480-1489.

29. Knol MJ, Vandenbroucke JP, Scott P, Egger M. What do case–control studies estimate? Survey of methods and assumptions in published case–control research. Am J Epidemiol. 2008;168(9):1073-1081.

30. Stolley PD, Tonascia JA, Tockman MS, Sartwell PE, Rutledge AH, Jacobs MP. Thrombosis with low-estrogen oral contraceptives. Am J Epidemiol. 1975;102(3):197-208.

31. Realini JP, Encarnacion CE, Chintapalli KN, Rees CR. Oral contraceptives and venous thromboembolism: a case–control study designed to minimize detection bias. J Am Board Fam Pract. 1997;10(5):315-321.

32. Checkoway H, Pearce N, Kriebel D. Selecting appropriate study designs to address specific research questions in occupational epidemiology. Occup Environ Med. 2007;64(9):633-638.

33. Hurwitz ES, Barrett MJ, Bregman D, Gunn WJ, Pinsky P, Schonberger LB, et al. Public Health Service study of Reye's syndrome and medications. Report of the main study. JAMA. 1987;257(14):1905-1911.

34. Lipsitch M, Jha A, Simonsen L. Observational studies and the difficult quest for causality: lessons from vaccine effectiveness and impact studies. Int J Epidemiol. 2016;45(6):2060-2074.

35. Haber M, An Q, Foppa IM, Shay DK, Ferdinands JM, Orenstein WA. A probability model for evaluating the bias and precision of influenza vaccine effectiveness estimates from case–control studies. Epidemiol Infect. 2015;143(7):1417-1426.

36. Kraaijenhagen RA, Haverkamp D, Koopman MM, Prandoni P, Piovella F, Buller HR. Travel and risk of venous thrombosis. Lancet. 2000;356(9240):1492-1493.