

---

# Vaccine Safety Study as an Interesting Case of "Over-Matching"

---

M. Catherine DeSoto and Robert T. Hitlan

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/53876>

---

## 1. Introduction

Increasing levels of diagnosed cases of autism have alarmed parents and health officials, but the cause has not been established. It has been hypothesized that vaccination itself, or some component in vaccines, may be somehow related to the onset of autism in some cases (DeLong, 2011; Gallagher & Goodman, 2010). Researchers have sought to alleviate such concerns. Although most studies report null effects, work continues to be published that suggests some reason for concern (Hewiston et al., 2010). Some skepticism of the safety of vaccines still exists, documented by scholars on either side of the issue (Austin, Schandley & Palombo, 2010, Destafano, 2007). As it is, the topic of vaccine safety and triggering of unintended outcomes is one of the most controversial topics in environmental health and toxicology.

After initial safety studies, case-control designs are often employed to continue to investigate both side effects and efficacy of inoculation. Matching is a technique used to improve signal to noise in research case-control designs. Matching cannot – or should not – be done in a way that artificially increases the chance that within strata exposure is the same. This happens when a matching variable is a strong predictor of exposure and is called over-matching. Here, we report a textbook case of overmatching within a widely-cited article. Focusing on the overmatching as a statistical concept, suggestions are made to standardize when overmatching may have occurred. It is important for statisticians to note when a study that fails to find an effect related to public health outcome has employed a design that would be expected a priori to result in a lack of effect.

It has been noted that some children received exposure to mercury significantly in excess of safety standards during the 1990's, before the level of thimerosal in vaccines was lowered (Geier & Geier, 2006), this has been suggested to increase odds of various developmental

disorders (Geier & Geier, 2006). The research by Price et al. (2010) spans the birth cohort years that saw a decline in thimerosal exposure and reports that thimerosal exposure was not associated with risk outcome of autism. Indeed, many studies have been published that find no negative effect of vaccination on developmental outcomes whatsoever (Parker, Schwartz, Todd, Pickering, 2004; see Destafano, 2007 for a review), indicating a lack of cause and effect between vaccination and autism. Here, we suggest that a recent widely cited study was flawed, and urge statisticians to carefully and critically review outcomes research on high stakes topics. It should be noted and understood that a flaw in such a study does not mean that vaccines cause autism, nor does it follow that one would properly assume that the flaw leads to the conclusion that vaccination is not safe. Rather the weight of scientific research as a whole should be deferred to.

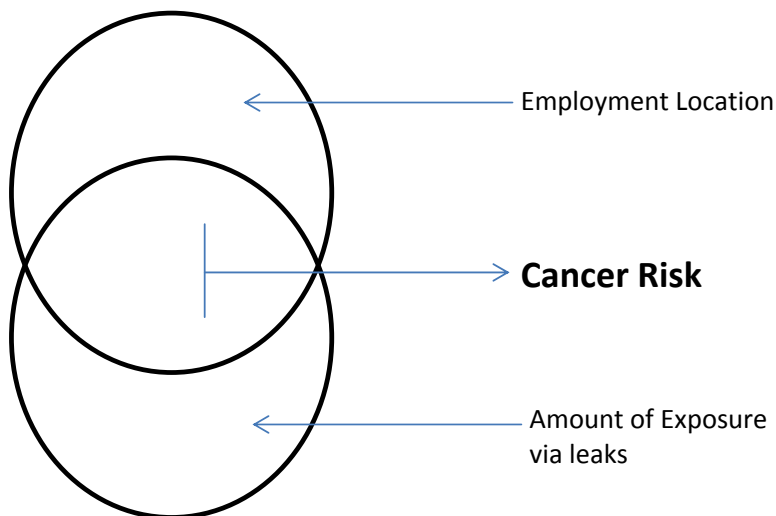
Conditional logistic regression (CLR) is a statistical technique used when the researchers have matched cases with controls on various parameters (e.g., age, gender). CLR is the often-used and appropriate way to analyze matched data sets (Rahman, Sakamoto & Fukui, 2003). To be clear, matching means that (as an example) for every 'case' that is male and aged 12, there is a control selected from a pool of possible controls that is also male and aged 12. If this were done, the researchers "matched on age and gender." A variant is to have two or three times the number of controls within each condition, or stratum. (Meaning for every male case who is age 12, there are three controls who are male and age 12.) The matched unit is called a stratum. When analyzing the data, CLR analyses are done within strata. When matching is done, only conditions (strata) that have cases and control pairs that vary on the risk factor contribute to the estimate of the effect of the risk factor (Miettinen, 1968). In other words, if exposure level within strata is the same, CLR cannot estimate the effect. As such, matching is a key design feature.

Matching cannot – or should not – be done in a way that artificially increases the chance that within strata exposure is the same; this happens when a matching variable is a significant predictor of exposure and is called overmatching.

Proper design can have important implications and researchers are appropriately cognizant of the possible perils of failing to take enough care in considering the matching design. If matching is used, researchers are wise to give explicit consideration to ensure that the problem of overmatching is avoided when attempting to accurately estimate risk of an exposure of interest (Sasieni and Castanon, 2009; Al-Taiar et al., 2009; Vidal et al., 2008; Agudo & Gonzalez, 1999; Cullison et al., 2007). And this problem has long been known (see for example, West, Schuman, Lyon, Robison & Allred, 1984). In their consensus paper on outcomes research, the American Thoracic Society noted that, "Overmatching, matching for a variable that is associated with the exposure but not the outcome, will reduce the statistical power of the study," (p. 364). Improper matching cannot later be undone via analysis and the effect of the matched variables cannot be checked, once matching has been done (Rubinfeld et al., 1999). How could this happen? Usually, this arises when a researcher fails to realize he or she is essentially matching on the exposure variable, and inadvertently the researcher matches the effect out.

To illustrate overmatching, a fictitious example will be briefly discussed, followed by an actual example from the literature. Assume the question is whether radiation exposure in nuclear plant workers contributes to cancer. A hundred cancer cases are found, and a control group of 700 is identified. Then, each case is matched with one from the control group on gender, smoking, job location, and age. The researchers match on these variables to increase efficiency (because they think these variables might independently account for disease risk). We will keep this as one to one matching for simplicity, but a 1:3 matching would essentially work the same.

In this example, overmatching would happen if the researchers are looking for effects of radiation but fail to consider that while which power plant the worker is employed might have some independent influence on disease risk (which is why it is matched), location could also be a major determinant of radiation exposure. For example, imagine Plant L often had radiation leaks, while Plant S had better safety. If one then matches on where one works, all of the variance unique to a particular plant is matched out. In such a case, an effect for radiation – even if huge could be missed. It will be clear if one considers that this would be like testing if radiation was related to cancer in Japanese nuclear power plant workers after controlling for location with one of the locations being Fukushima (Figure 1). If participants who developed cancer were matched on where they worked – the researchers may not detect any true health effects of the radiation exposure from the nuclear meltdown at Fukushima compared to working at other plants that did not have a meltdown. The researchers would have matched out any effects associated with where they worked.



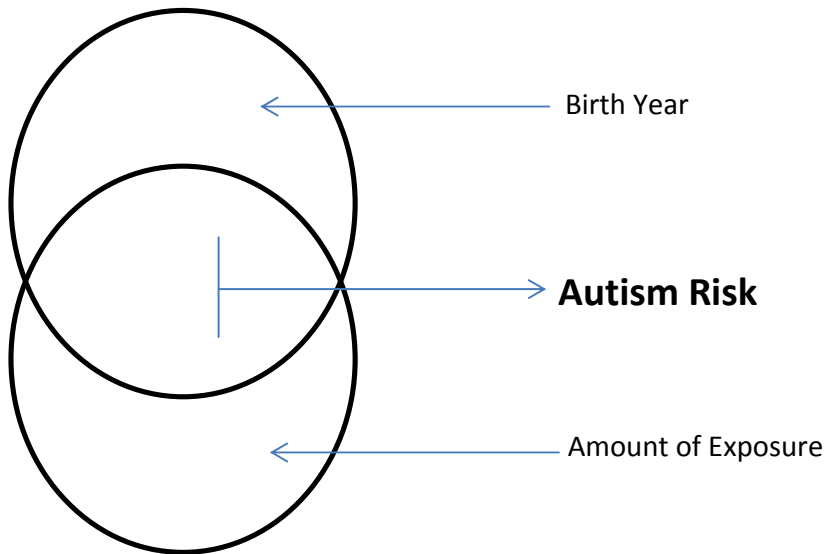
**Figure 1.** Overlapping variance: Illustration of Overmatching on Radiation Exposure; In this fictitious example, matching on the nuclear power plant of employment in the design of the study would be overmatching because it would remove the largely overlapping variance associated with radiation due to the Fukushima leak, obscuring the effect

A now classic paper by Marsh, Hutton and Binks (2002) refers to a real research example and is entitled, "Removal of radiation dose response effects: an example of over-matching." It details how a true effect can be missed if the researchers overmatch. According to the authors, "If the exposure itself leads to the confounder or has equal status with it, then stratifying by the confounder will also stratify by the exposure, and the relation of the exposure to the disease will be obscured. This is called over-matching and leads to biased estimates of risk," (p. 1235). After previous work had suggested that radiation did predict leukemia, the more recent case-control study failed to indicate any relation between radiation and leukemia. The matched factors in the new study that showed no increased for leukemia as a result of radiation included: date of birth, gender, and "date of entry". "Date of entry" was a measure of what years the workers worked in the industry. The data was properly analyzed given the matched design by conditional logistic regression, yet failed to find a known effect.

This prompted the study of the statistics used, with a focus on the matching process. It was noted that some things are appropriate to match on, for example, gender. "Because of the underlying difference of the risks of leukemia between the sexes," being male versus female affects the outcome, and it is important not to accidentally have more males in the case group as this would be a confound. On the other hand, Marsh et al. clearly showed that radiation exposure varied by year, that is some years were higher than others and this was indeed a major source of radiation variation (see figure 3, Marsh et al., 2001). "The general decline in median dose shows that dose and time are associated. The situation seems to be one where dose is partially 'explained' by date of entry, both being related to time;" in sum, "this seems to have had the effect that workers in the same matched set have broadly similar recorded doses. The apparent over-matching on date of entry has distorted the parameter estimate of the risk of leukemia on cumulative dose by introducing matching (at least partially) on dose," (Marsh et al., 2002).

What is the take home message of this classic report on the problem of overmatching? When researchers match on a variable closely associated with the risk factor exposure, then actual effects will not be-- and cannot be-- detected. This danger is written about by various other authors as well. Richard Monson in his text, "Occupational Epidemiology" notes "over matching is a problem in case control studies." Monson emphasizes that "there should be no possibility that the factor is part of the causal pathway linking exposure and disease under study." (p. 41). If this is even remotely possible, Monson advises matching should not be done on that variable. Monson discussed an example where overmatching resulted in underestimating the effect of estrogen use on endometrial cancer. Here the matching was on a correlate of intrauterine bleeding, which in effect controlled for a symptom of the cancer itself.

Price et al. do not mention overmatching as a potential concern. The risk factor of interest is thimerosal exposure via its inclusion in vaccine ingredients. There are two things that have a systematic and predictable effect on how much thimerosal exposure a child would receive: 1) the vaccine schedule a child is born into/national recommendations, and 2) which manufacturer a given provider is using for the vaccines (e.g. for the same years, Smith, Kline and Beecham were using thimerosal in their HepB vaccine, while Merck did not).



**Figure 2.** Controlling for Birth Year is overmatching due to the overlap with Amount of Exposure; similar to the radiation risk for leukemia written about by Marsh, controlling for time is (at least partly) controlling for exposure, which varies with birth year. The matching on birth year is matching on the exposure. This seems to have had the effect that children in the same matched set have similar recorded exposures to thimerosal, removing much of the variance

Price et al. matched out both of these variations in exposure. This has the effect of ensuring that the control group is nearly identical with the case group on the risk factor, which prevents its effect from being accurately measured. Considering cumulative exposure for the first 7 months of life, the overall mean for the full data set is 102.88 micrograms/Hg and a standard deviation of 42.2. The means for the cases and matched controls is 100.0 and 103.2 micrograms of Hg; this similarity (less than one tenth of the standard deviation) is forced by the matching on the variables that define exposure. Birth year dictates which vaccine schedule a child is born under as well as which batch brands and formulations are available on the market at a given time. Doctors within a practice will be using the same manufacturer across children (vaccines are ordered in large batches from a given manufacturer; the Vaccine Data Set used by Price et al. documents that the same providers use the same manufacturer. Thus, this is a text book case of overmatching: variables were matched on that essentially define exposure. It is well known that matching on a variable that is associated only with exposure, not with disease, reduces statistical efficiency (Zondervan et al, 2002; Rubinfeld et al., 1999; Day, Byar, & Green, 1980) and that care needs to be taken to avoid this in a case-control research design.

Across the different years, the average cumulative exposure varies from 42.3 micrograms to 125.46 micrograms; while within the birth year stratas, the mean exposures do not vary by more than 15 micrograms. Birth year is a variable that defines exposure due to changes in recommendations regarding the vaccine schedule and changes in vaccine formulas that oc-

curred at different times. The above panels suggest that variance within the matched variable (year) is small compared to the variance between birth years: birth year is accounting for much variance in thimerosal exposure.

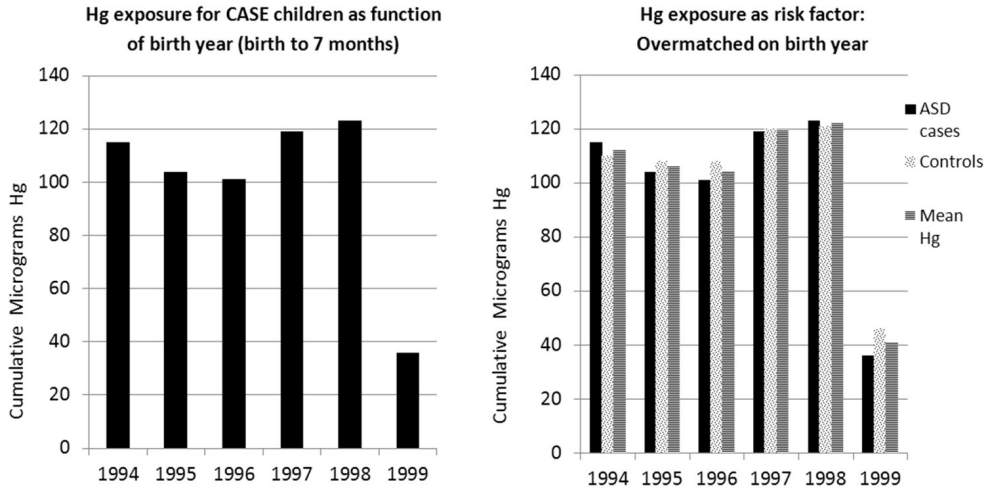
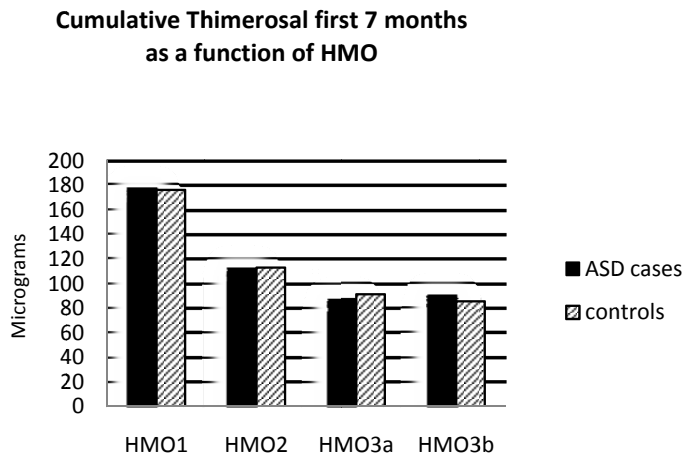


Figure 3. The difference across birth years on the risk factor of interest

During the past decades, there have been three main exposure sources of thimerosal: DPT/DTaP, then Hepatitis B and Hib vaccines, while flu shots are currently the primary source in the USA today. The Hib/Hep B introductions came in during the late 1980s and early 1990s. The recognition that the cumulative mercury burden may have been too high came in 1999, and mercury levels dropped for most vaccines given to children in the USA. Some people have raised concerns that the increase in autism is associated with the changes in thimerosal exposure; that is, the increase in autism is thought to be a function of the increases in the number and amount of mercury containing vaccines. Whether or not one finds this model persuading, matching on birth year is questionable if the goal is to test the model that differences in thimerosal exposure via vaccine schedule increase ASD risk since -- as most people are aware -- birth year essentially dictates which vaccine guidelines a child is born into. It could be that the authors intended to control for hypothesized changes in diagnostic criteria trends across the six birth years. The problem is that diagnostic effects on risk is not measured while birth year effects on exposure are clear.

Moreover, HMO is not known to be a significant predictor of the outcome of autism diagnosis, so potential reasons to match on this variable are less clear. As Hansson and Khamis (2008) write in their paper on matched-sample logistic regression, "Generally, matching will increase the efficiency of the study when the matching variable is a strong outcome determinant, but will actually reduce it when the matching variable is strongly related to the exposure variable (over-matching)," (p.595-596). Meittinem (1969) states that, "matching reflects

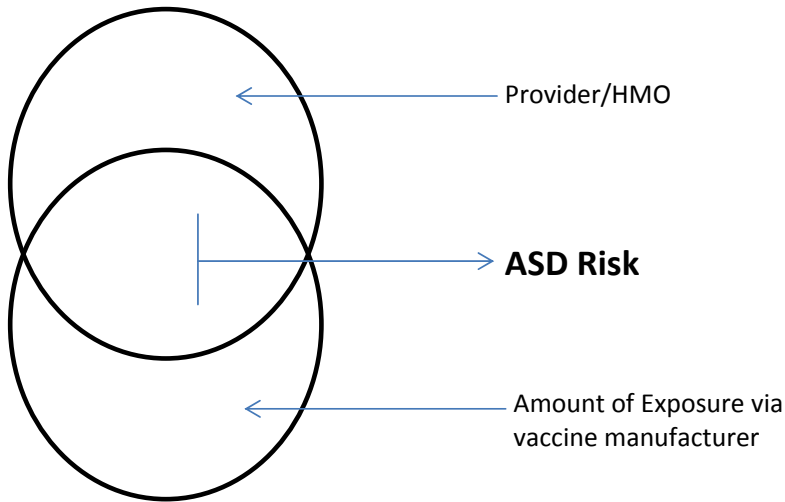
the notion that the probability  $P$  of response is related to  $M$ ," (p. 340) meaning that when one matches, one infers that the matching variable effects the probability of risk (here for autism). HMO / health care provider was a major determinant of thimerosal exposure, but we are not aware of papers that identify HMO is an independent risk for autism. Thus, it should not have been matched. What was needed was a design that compared persons with different exposures. "Studies with uniform developmental assessments of children with a range of cumulative thimerosal exposures are needed," (Vertraeten et al., 2003). Here Price et al., began with such a data set, but then matched on birth year and HMO, matching out exposure differences and negating comparisons of different exposures (see Miettinen, 1969 for a mathematical discussion).



**Figure 4.** The apparent over-matching on HMO distorts the estimate of the risk of autism on thimerosal by introducing matching on exposure. If one matches on provider, one is matching on the vaccine manufacturer. There are different manufacturers available, but a given provider will be using one or the other. This seems to have had the effect that children in the same matched set have similar recorded exposures to thimerosal. Again, this removes this variance and obscures the effect

The model Price et al. were trying to test was whether thimerosal exposure via the US vaccination schedule was associated with any increased risk of autism. To do this, they needed to compare persons with and without high levels of exposure. They did not do this because due to the conditional logistic regression matched on both birth year and HMO they have inadvertently made sure that cases were only compared to controls with the same exposure. Because Price et al. did not mention the possibility of overmatching, we assume this did not occur to the research team. We assume this was accidental, but it does underscore the need to have a balanced research team that does not start with assumptions that might flaw the design. For example, assuming that the increase in autism is only due to diagnostic changes would lead to controlling for birth year, which might have been flagged by someone who does not share this assumption. It is harder to understand why HMO would be matched. Overall, this is unfortunate because the question

of vaccine safety is high stakes. There are concerns that a proper test of the full vaccine schedule has not been properly tested, and that the safety tests that exist have been designed by the vaccine industry itself. Such concerns about conflicts of interest may be preventing otherwise willing parents to adhere to the full vaccine schedule. Vaccines have been and will continue to be a huge benefit to humanity. But this paper is flawed. Unfortunately, there is not an analytic fix for overmatching: it is design flaw.



**Figure 5.** Which manufacturer a given provider used for the vaccines varied by HMO. Manufacturers differed in their thimerosal use. For example, in 2002, Smith, Kline and Beecham were using thimerosal in their HepB vaccine, while Merck did not. While the data set is careful to note manufacturer and Hg in the associated batch and manufacturer, but CLR matching on HMO results in comparing cases to controls who had the same levels of exposure

The Price et al. research is an interesting case of overmatching that we think is of general interest in the field of epidemiology. To avoid misunderstanding, we wish to state that this research does not support the argument that vaccines or thimerosal in vaccines cause autism. It is however, uninformative to the question.

## 2. Suggestions for avoiding the problem of publishing overmatched results

One way to conceptualize the problem of overmatching in conditional logistic regression is the preemptive removal of variance that should stay available for the hypothesized predictor variable to attempt to account for. The total variance in a data set can be defined using the average squared distance from the mean score for each participant:  $s^2 = SS/df$ . A question related to overmatching concerns how much of the total variance is taken out beforehand (matched out)? How much is too much? 10%? 90%? 50%? We would propose that the per-



cent removed before testing should normally be small compared to the total. Further, the removal of this variance should only occur when there is authentic need: when the potential matching variable is likely related to the outcome of interest via a path that is distinct from the risk variable of interest in a case-control design.

As elaborated above, matching is appropriate only if the matching variable is a strong predictor of the outcome of interest, but it is not appropriate when the matching variable is strongly related to the exposure risk variable. We offer three suggestions to help objectively identify, and thus avoid, the problem of overmatching.

**Empirical Support.** Before matching, first and foremost, researchers should locate studies that suggest the potential match is likely correlated to the probability of the outcome occurring. These should be cited to support the need to match on that variable. If there is no reason to think the matching variable relates to the outcome, there is no reason to match it.

**Remaining Variance.** Next, once the participants have been selected as a matched data set, researchers can check to get an idea how much variance in the exposure variable is actually accounted for by the matching variable  $M$ . If only a small amount of the variance is left after the various matching, matching on the variable(s) cannot be justified and an unmatched or lesser matched set of participants is called for. Specifically, a check to see if too much of the total variance in the outcome of interest is matched out could be done by requesting Partial Eta Squared. Partial Eta Squared represents the proportion of the total variance that is explained by the between factor when an ANOVA is performed. Specifically, one can take the extra step of analyzing the variance in the risk factor of interest (e.g., thimerosal exposure) as a function of the matched variable (e.g., HMO or BirthYear). In this example, using thimerosal exposure as the dependent variable, the total SS is 23507522. The SS associated with the Birth Year is 1485471. This gives Partial Eta Squared = .456, meaning that about 46% of the total variance in thimerosal exposure is fully explainable based on Year of Birth. When one matches on this, only about half (54%) of the variance is left.

HMO, the other variable matched on, removed about 30% of the variance.

The percent that should be left would depend on the research question and causal assumptions, but we suggest that if a matched variable is removing more than a fourth (25%) of the variance (corresponding to a large effect size, Cohen, 1977), matching is unlikely to be warranted for this reason alone and welcome commentary on this benchmark proposal.

**Relative relations.** Finally, there are times when it could be proper to match on a variable that accounts for variance in the risk factor being tested. A recent case coincidentally also related to vaccines helps to illustrate this more. It had been pointed out that the enormous benefits of the flu vaccine among the elderly appeared to far surpass even the effect that a total eradicating of flu from the vaccinated population could account for (Jefferson, 2006). After additional investigation, much of the original effect appears to be due to the tendency for seriously ill and/or less healthy elderly persons not to have the flu shot. To be clear, most of the flu vaccine effect on mortality was found to be due to health of the participants independent of the flu shot (Jackson et al., 2006). In this case, if this had been a case control design, the risk factor would be flu vaccine and the probability outcome of interest would

hospitalization or death. In such a case control study, it would be proper to match on preexisting health, even though one would find that health accounts for some of the variability in getting or not getting the risk factor (flu vaccine). BUT: health would also relate to the mortality outcome, and even more strongly. It is this strong relationship that is key. If the variable is more strongly related to the outcome – this serves to justify matching.

To objectively quantify this, one needs to know how strongly M is related to the Risk Factor R; and then how strongly M is related to the probability of response P. A problem is that different types of data can make precise comparisons of effect size hard to judge.

Assume that M would be HMO, R would be mcg Thimerosal exposure, and P would be ASD diagnosis. It would be desirable to compare the size of this relationship M to R with the relationship of M to P. It would be ideal if one could simply compute correlations for M and R and for M and P. However, in most cases this would not work: the scales are not all continual, and even if one were to employ a Spearman correlation, it would not be apparent how to code something like HMO to insure a linear relationship. What if HMO 2 was associated with an increase in thimerosal, and HMO 1 and 3 both had low levels? This would result in a low correlation due to the curvilinear relationship, even IF much of the variance were in fact associated with HMO. On the other hand, the relationship between HMO and thimerosal (M and R) can be checked via ANOVA easily enough since R is continual and M is categorical. ANOVA would not work for testing association between M and P because both M and P are both categorical in this case. Chi – Square would be appropriate. However, regardless of the correct hypothesis test, all hypothesis tests are in fact unified by the p value.

The p value is a function of the size of the effect and the sample size. Different types of statistical tests have different probability distributions, but the total area under the curve has a constant meaning across tests. The percent of area covered means the same thing in any test, regardless of the precise shape of the curve associated with a particular statistical test (correlation, ANOVA, Chi-square). A small p value could be due to a large effect, or it could be due to a very small effect and a very large sample. It should be stated that when sample sizes are similar, it will not be unduly affected by sample size differences. Since the sample will be the same for testing M to P or testing M to R, we propose the p values are the most readily available means to index the comparison.

Compute a measure for the relationship of M and P and the associated p value. (e.g., HMO and ASD:  $\chi^2(2) = 1.59, p = .45$ )

Compute a measure for the relationship of M and R and the associated p value. (e.g., HMO and Thimerosal exposure:  $F(2,1090) = 237, p < .0001$ ).

The p value in all cases should be smaller for the M to R relationship, compared to the M to P relationship test. This will serve to demonstrate that even if the Matching variable does bear some relationship to the risk factor for the outcome probability, there is clearly a stronger relationship to the outcome itself, thus objectively justifying the matching. (e.g., the p value of .45 indicates no relationship exists between the matched variable and the outcome of interest, while the p value <.0001 indicates that matched variable is related to the expo-

sure variable being tested. It is well known that matching on a variable that is associated only with exposure, not with disease, reduces statistical efficiency in a case – control design (Zondervan et al, 2002; Rubenfeld et al., 1999; Day, Byar, & Green, 1980), and this in essence, defines the problem of overmatching).

To sum, variables such as birth year, HMO, age, gender, address should first and foremost be matched if and only if there is a truly justifiable rationale to expect they have an independent causal pathway to the outcome; "matching will increase the efficiency of the study when the matching variable is a strong outcome determinant, but will actually reduce it when the matching variable is strongly related to the exposure variable (over-matching)," (Hanson & Khamis, 2008, p.595-596). Second, if the majority of the variance in the risk factor being tested is removed by matching, before the hypothesis is tested, extreme caution in reporting a lack of effect is warranted. Finally, recalling that sample size will be held constant, testing the relationships of M to R and P and comparing the p values can be used to justify matching in the context of the matching variable removing variance relating to the risk factor. We would propose that overmatching has and will continue to be a problem in matched case control designs, but suggest that employing the three checks above will serve to lessen deleterious effects associated with publishing overmatched results.

We welcome comments on these proposals.

## Acknowledgement

This work was partially funded by a small grant awarded to the second author, Robert T. Hitlan from Safeminds. We thank Safeminds for their support.

## Author details

M. Catherine DeSoto and Robert T. Hitlan

University of Northern Iowa, Cedar Falls, USA

## References

- [1] Austin D. W., Shandley K. A., Palombo E. A. 2010. Mercury in vaccines from the Australian childhood immunization program schedule. *J. Toxicol. Environ Health A*, 73: 637-40.
- [2] Agudo A. and González C. A. 1999. Secondary matching: A method for selecting controls in case-control studies on environmental risk factors. *Int. J. Epidemiol.*, 1999: 1130-1133.

- [3] Al-Taiara, A., Assabria, A., Al-Haboria, M., Azazya, A., Algabrib, A., Alganadib, M., Christopher J. M., Whittyc, and Jaffarc, S. 2009. Socioeconomic and environmental factors important for acquiring non-severe malaria in children in Yemen: A case-control study. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 1; 72-78.
- [4] Bonner, P. C., Schmidt, W. P., Belmain, S. R., Oshin, B., Baglole, D., and Borchert, M. 2007. Poor housing quality increases risk of rodent infestation and lassa fever in refugee camps of Sierra Leone. *Am. J. Trop. Med. Hyg.*, 77; 169-175.
- [5] Day, N. E., Byar, D. P., and Green, S. B. 1980. Overadjustment in case control studies. *Am. J. Epidemiol.*, 112: 696-706.
- [6] DeLong, G. A. 2011. Positive association found between autism prevalence and childhood vaccination uptake across the U.S. population. *J. Toxicol. Environ. Health A*, 74: 903-16.
- [7] DeStefano F. 2007. Vaccines and autism: Evidence does not support a causal association. *Clin. Pharmacol. Ther.*, 82: 756-759.
- [8] Gallagher, C. M. and Goodman, M. S. 2010. Hepatitis B vaccination of male neonates and autism diagnosis, NHIS 1997-2002. *J. Toxicol. Environ. Health A*, 73: 1665-1677.
- [9] Geier D. A., and Geier M. R. 2006. An evaluation of the effects of thimerosal on neurodevelopmental disorders reported following DTP and Hib vaccines in comparison to DTPH vaccine in the United States. *J. Toxicol. Environ. Health A*, 69: 1481-1495.
- [10] Hansson, L. and Khamis, H. J. 2008. Matched samples logistic regression in case-control studies with missing values: when to break the matches. *Statistical Methods in Medical Research*, 17: 595-607
- [11] Hewitson, L., Houser, L. A., Stott, C., Sackett, G., Tomko, J. L., Atwood, D., Blue, L., White, E. R. 2010. Delayed acquisition of neonatal reflexes in newborn primates receiving a thimerosal-containing hepatitis B vaccine: Influence of gestational age and birth weight. *J. Toxicol. Environ. Health A*, 73: 1298-1313
- [12] Jackson, L., Kackson M. et al. (2006). *Evidence of bias in estimates of influenza vaccine effectiveness in seniors*. *Int. J. Epidemiol.* 35 (2): 337-344.
- [13] Jefferson, T. 2006. Influenza vaccination: policy versus evidence. *British Medical Journal*, 333: 912-915.
- [14] Marsh, J. L., Hutton, J. L., and Binks, K. 2002. Removal of radiation dose response effects: an example of over-matching. *British Medical Journal*, 325: 327-330.
- [15] Miettinen, O.S. 1969. The matched pairs design in the case of all-or-none responses. *Biometrics*, 24: 339-52
- [16] Monson, R. R. 1990. *Occupational Epidemiology* (2<sup>nd</sup> Ed.). BocaRaton, FL: CRC Press.

- [17] Parker, S. K., Schwartz, B., Todd, J., and Pickering, L. K. 2004. Thimerosal-containing vaccines and autistic spectrum disorder: A critical review of published original data. *Pediatrics*, 114: 793-804.
- [18] Price, C. S., Thompson, W. W., Goodson, B., Weintraub, E. S., Croen, L., A., Hinrichsen, V. L., Marcy, M., Robertson, A., Eriksen, E., Lewis, E., Bernal, P., Shay, D., Davis, R. L., and DeStefano, F. 2010. Prenatal and infant exposure to thimerosal from vaccines and immunoglobulins and risk of autism. *Pediatrics*, 126: 655-664.
- [19] Rahman, M., Sakamoto, M. D., and Fukui, T. 2003. Conditional versus unconditional logistic regression in the medical literature *J. of Clin. Epidemiol.*, 56: 101-102.
- [20] Rubernfeld, G. D., Angus, D. C., Pinsky, M. R., Curtis, J. R., Coners, A. F., and Bernard, G. R. 1999. Outcomes research in critical care. Results of the american thoracic society critical care assembly workshop on outcomes research. *Am. J. Respir. Crit. Care Med.*, 160: 358-367
- [21] Sasieni, P., Castanon, A., and Cuzick, J. 2009. Effectiveness of cervical screening with age: Population based case-control study of prospectively recorded data. *British Medical Journal*, 339: b2968.
- [22] Vertraeten, T., Davis, R. L., DeStefano, F. 2003. Safety of thimerosal containing vaccines: A two-phased study of computerized health maintenance databases. *Pediatrics*, 112: 1039-1048.
- [23] Vidal, J. S., Vidailhet, M., Elbaz, A., Derkinderen, P., Tzourio, C., and Alperovitch, A. 2008. Risk factors of multiple system atrophy: A case-control study in French patients. *Movement Disorders*, 23: 797-803.
- [24] West, D. W., Schuman, K. L., Lyon, J. L., Robison, L. M., and Allred, R. 1984. Differences in risk estimation from a hospital and a population-based case-control study. *Int. J. Epidemiol.*, 13: 235-239.
- [25] Krina, T., Zondervan, L. R., Cardon, L. and Kennedy, S. H. 2002. What makes a good case-control study? *Human Reproduction*, 17: 1415-1423.

