Errors in epidemiological exposure assessment: Implications for study results

Martin Röösli[#1]

[#]*Institute of Social and Preventive Medicine, University of Bern*
Finkenhubelweg 11
3012 Bern, Switzerland
[1]Roeoesli@ispm.unibe.ch

***Abstract*— The main aim of epidemiological exposure assessment is to measure the exposure of interest as accurately as possible. Unfortunately, errors in exposure assessment are unavoidable to a certain extent. However, exposure misclassification does not automatically lead to severe bias in the risk estimates, because the bias depends on the magnitude and the nature of exposure misclassification. There are scenarios where a large misclassification error introduces negligible bias only, while in other situations a presumably small misclassification error already hampers the interpretation of the risk estimates. Therefore, the impact of bias due to misclassification has to be discussed for every specific study setting.**

The aim of this paper is to present different types of exposure misclassification and to clarify their impact on the study result using case-control studies about mobile phone use and brain tumour risk as examples.

I. BACKGROUND

In practice, for most types of environmental exposure it is almost impossible to obtain accurate exposure information for a period covering many decades, including the details of how the exposure varied over time during this period. For this reason the primary goal of epidemiological exposure assessment is to find a good proxy representative for the exposure of interest which allows dividing the study collective accurately into an exposed and a non-exposed group (or into groups which are exposed to a varying degree) [1].

In the case of head exposure to radio and microwave frequency electromagnetic fields such an exposure proxy may be the lifetime cumulative number of phone calls, the cumulative duration of phone calls or the time since first subscription to a mobile phone operator. Depending on the information source, e.g. questionnaire or operator data, the obtained information is subject to error, which may result in exposure misclassification.

II. AIMS

The aim of this paper is to introduce different types of exposure misclassification. Their impact on the study results will be exemplified through case-control studies on mobile phone use and brain tumour risk. The concept of a case-control study is to compare the exposure of cases with the exposure of a random sample, which is representative for the whole population the cases were derived from. Up to now several case-control studies on brain tumour risk and use of mobile phones have been published. Exposure assessment in all of these case-control studies is based on self-reported cellular phone use. Within the INTERPHONE collaboration, several validation studies have been performed where self reported information was compared with objective information recorded by operators or software modified phones [2-5]. These studies found moderate to high correlation between recalled and actual phone use for the last six months. The correlations ranged from 0.5 to 0.8 (weighted kappa: 0.2-0.6) across eleven different countries and were of the same order for number and duration of calls. On average, individuals underestimated the number of calls by a factor of 0.92 and overestimated the duration of calls by a factor of 1.42. In Denmark, a comparison of self-reported mobile phone use with operator data from 1982 to 1995 yielded a fair agreement (kappa value: 0.3) [4]. It has to be noted that not only self-reported information is limited but also operator data. The subscriber is not necessarily the person who is using the phone.

III. EXPOSURE ASSESSMENT ERRORS

A. Sensitivity and specificity

The reliability of an exposure assessment is measured as sensitivity and specificity. To simplify matters let us assume a binary exposure status: either being exposed or not exposed (e.g. regularly using a mobile phone vs. not using mobile phones). With respect to true exposure status, exposure assessment can result in the division of individuals into four groups (see Table 1):

(a) those classified as exposed who are really exposed;
(b) those classified as exposed who are in fact unexposed;
(c) those classified as unexposed who are in fact exposed;
(d) those classified as unexposed who are truly unexposed.

Sensitivity refers to the proportion of people being exposed and being (correctly) classified as exposed (=a/(a+c)). Specificity refers to the proportion of people being unexposed and being (correctly) classified as unexposed (=d/(b+d)).

TABLE 1
CROSS TABULATION OF THE FOUR POSSIBLE COMBINATIONS OF EXPOSURE CLASSIFICATION WITH TRUE EXPOSURE STATUS.

		True exposure status	
		exposed	*not exposed*
Exposure classification	*exposed*	a	b
	not exposed	c	d

An exposure assessment with a sensitivity of 90 percent and a specificity of 80 percent means that 90 percent of the exposed people are correctly classified as exposed and 80 percent of the unexposed individuals are correctly classified as unexposed. The remaining study participants are erroneously assigned to the wrong exposure category.

B. Differential exposure misclassification (systematic)

Exposure misclassification can be either systematic (differential) or random (non-differential). The former means that the misclassification depends on the disease status. This can happen for instance in a case-control study on mobile phone use if cases reflect more intensely about past exposure situations than controls and thus are more likely to report use of mobile phones than controls (recall bias). This is a serious problem as it creates a systematic bias in the study resulting in either an overestimation or an underestimation of the true effect estimate, depending on the type of systematic exposure misclassification.

Example 1: Let us assume that 60% of the population is regularly using mobile phones (i.e. at least once a week during the last six month) [6]. Further, I assume that exposure assessment is perfect with the exception of unexposed cases who tend to overestimate their mobile phone use. As a consequence, 20% of the unexposed cases are erroneously classified as exposed. If there is no exposure-disease association in reality the true Odds ratio (OR) is 1.0. However, due to systematic misclassification of the unexposed cases the observed OR would be 1.4. If the true risk was 2.0, the observed risk would be 2.7. In either case, the observed risk is overestimated, because the exposure status of unexposed cases is systematically overestimated.

C. Non-Differential exposure misclassification (random)

In many situations exposure classification is non-differential; i.e. the error is random and does not differ between cases and controls. In order to explain the impact of a non-differential exposure misclassification, I will present some simple examples.

Example 2: Again I assume that 60% of the population is regularly using a mobile phone. Further I assume that there is no risk present (OR=1) and that non-differential exposure misclassification occurred in a case-control study. It can be easily modelled that under such circumstances the observed OR will be unchanged, i.e. OR=1, regardless how large the exposure misclassification is. The only problem is that the random data fluctuation increases according to the increasing imprecision of the exposure assessment. Thus, the observed OR's scatter randomly around 1. Therefore, in some studies an over- or underestimated OR may occur due to chance.

Example 3: The situation is different if there is a true association between the exposure and the disease. Thus, in example 3 I assume a doubling of the brain tumour risk for people regularly using a mobile phone. The hypothetical corresponding distribution of mobile phone use (=exposure) of 3000 control persons and 1500 cases is shown in table 2.

TABLE 2
DISTRIBUTION OF THE *TRUE* EXPOSURE STATUS OF CASES AND CONTROLS IF THERE IS AN OR OF 2 FOR REGULAR MOBILE PHONE USE.

	Cases	*Controls*
exposed	a_1=1125	b_1=1800
not exposed	a_0=375	b_0=1200c

The OR is obtained by multiplying a_1 with b_0 divided by a_0*b_1. Now, I assume that 10% of the truly exposed cases and controls are erroneously classified as unexposed (sensitivity=0.9) and 20% of the truly unexposed cases and controls are erroneously classified as exposed (specificity=0.8). Unfortunately, assuming exposure misclassification means that we cannot observe the 'true' situation as described in table 2; instead our data collection yields an erroneous table as shown in table 3. For instance, the observed 1088 exposed cases consist of 90% (sensitivity) of the real exposed cases as well as 20% (1-specificity) of the unexposed cases.

TABLE 3
DISTRIBUTION OF THE *OBSERVED* EXPOSURE STATUS OF CASES IF THE SENSITIVITY OF THE EXPOSURE ASSESSMENT IS 0.9 AND THE SPECIFICTIY IS 0.8. (TRUE OR IS ASSUMED TO BE 2.).

	Cases	*Controls*
exposed	**1088** (=0.9*1125 +0.2*375)	**1860** (=0.9*1800 +0.2*1200)
not exposed	**413** (=0.1*1125 +0.8*375)	**1140** =(0.1*1800 +0.8*1200)

The obtained OR of 1.6 (=(1087.5*1140)/(1860*412.5)) is obviously a substantial underestimation of the true exposure-disease association. Except for the impact of possible random data fluctuation, non-differential exposure misclassification always leads to an underestimation of the true exposure response association. The extent of underestimation depends on the sensitivity and the specificity of the exposure assessment.

In Fig. 1 the impact of exposure misclassification for varying assumptions about the sensitivity and specificity is shown for a true OR of 2 and an exposure prevalence of 60%. If a binary exposure classification is not correlated to the true exposure status, the sensitivity and specificity are 0.5. In such a case the observed OR is 1.0 (see Fig. 1). A negative correlation between exposure assessment and the true exposure status results in a sensitivity and/or a specificity below 0.5 and an observed OR below unity (1.0).

Interestingly, the effect of sensitivity and specificity depends on the exposure prevalence [7]. Fig. 2 shows the observed OR if the exposure prevalence is only 5% (e.g. regularly using a mobile phone for at least 10 years). It is striking that there is only a minor underestimation of the OR if the specificity is 1, even if sensitivity is very low (e.g. 0.1). In contrast, even a slightly reduced specificity of 0.9 results in a substantial underestimation of the true exposure-disease association, even if sensitivity is perfect. This means that in such a situation it is very important that those who are

considered as unexposed are indeed unexposed; whereas we do not have to worry much about the accuracy of the exposure assessment for exposed individuals. Even missing many of the real exposed individuals and consider them as unexposed does not introduce much bias in the risk estimate. The opposite is true if the exposure is highly prevalent (e.g. 95%). In this case sensitivity is very important whereas specificity is unimportant.

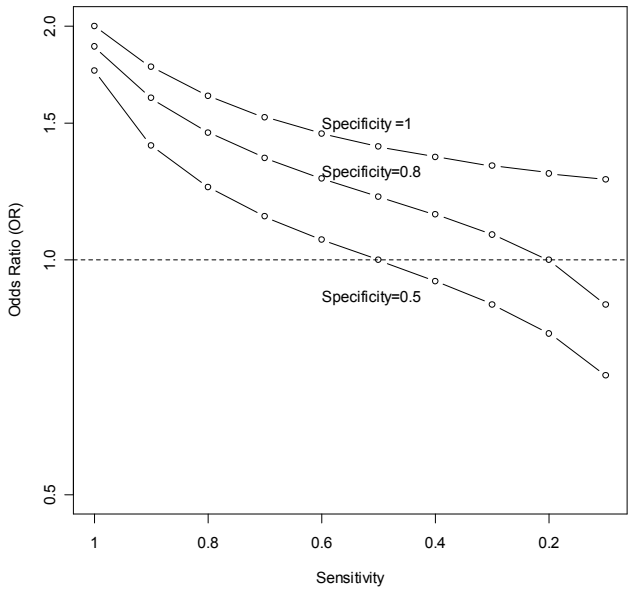

Fig. 1 Observed risk (OR) for different assumptions for the sensitivity and specificity of the exposure assessment if the true risk is 2 and the exposure prevalence in the population is 60%.

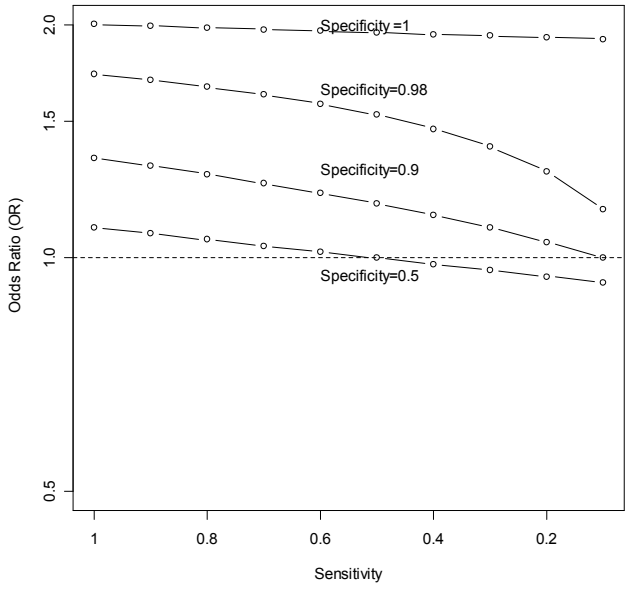

Fig. 2 Observed risk (OR) for different assumptions for the sensitivity and specificity of the exposure assessment if the true risk is 2 and the exposure prevalence in the population is 5%.

To simplify, standard errors are not calculated in the examples. However, with increasing exposure misclassification the standard errors are increasing as well. Therefore, the likelihood to miss a true exposure-disease association is additionally increasing.

D. Systematic and non-systematic misclassification combined

It is relatively simple to model the effect of differential and non-differential exposure misclassification for a binary exposure variable in a case-control study. With respect to a specific study, however, it is more difficult to determine the extent of both types of misclassification. Table 4 shows how the observed OR changes if both, differential and non-differential exposure misclassification are present. In these examples it is assumed that the exposure of cases is overestimated compared to controls. This can either be due to higher sensitivity for cases compared to controls, lower specificity or both effects combined.

TABLE 4
OBSERVED RISK (OR) FOR DIFFERENT ASSUMPTIONS ABOUT THE TRUE RISK AND ABOUT THE EXTENT OF SYSTEMATIC AND RANDOM EXPOSURE MISCLASSIFICATION.

Sensitivity		**Specificity**		**OR**	
cases	controls	cases	controls	true	observed
0.9	0.8	0.6	0.8	2	2.7
0.8	0.8	0.6	0.8	2	1.8
0.9	0.8	0.5	0.8	2	3.1
0.8	0.8	0.5	0.8	2	2.1
0.9	0.7	0.6	0.7	2	2.9
0.8	0.7	0.6	0.7	2	2.0
0.9	0.7	0.5	0.7	2	3.4
0.8	0.7	0.5	0.7	2	2.2
0.9	0.8	0.6	0.8	1	1.8
0.8	0.8	0.6	0.8	1	1.4
0.9	0.8	0.5	0.8	1	2.2
0.8	0.8	0.5	0.8	1	1.7
0.9	0.7	0.6	0.7	1	2.0
0.8	0.7	0.6	0.7	1	1.5
0.9	0.7	0.5	0.7	1	2.4
0.8	0.7	0.5	0.7	1	1.8

In most of the chosen examples, the (erroneously) observed OR is larger than the true exposure-disease association. That means that overestimation of the OR from differential exposure misclassification dominates the dilution effect from non-differential exposure misclassification. However, in one example the true effect is underestimated (2nd row) and in another example both effects are compensating each other and the true OR of 2 is observed (6th row). If there is no real risk (true OR=1), the observed OR is always overestimated.

IV. DISCUSSION

In epidemiological studies exposure misclassification is to a certain extent unavoidable. The effect of exposure misclassification can be modelled for specific assumptions and study settings. To simplify matters I chose an example with the binary outcome exposed vs. not exposed. The described principles can also be applied to exposure measures

on a continuous scale (e.g. cumulative duration of mobile phone use). The main conclusion was that differential exposure misclassification biases the risk estimates away from unity and non-differential exposure misclassification tends to shift the risk estimates towards unity (OR=1). This is a general pattern, which occurs as long as one deals with individual-level exposure variables. Counter-intuitively the effect of non-differential exposure misclassification is changing if exposure is assessed on a group-level. A group-level exposure assessment may be used in a cross-sectional or a cohort study when the group membership is defined by determinants such as occupation or residential area. When group-level exposure data are used, the underlying error model is the Berkson model [8]. It is less well appreciated that if the Berkson model holds, then the estimate of exposure effect obtained by ordinary linear regression is in fact unbiased and robust to random exposure misclassification [9]. However, the standard error is increased, resulting in less power or precision.

V. Conclusion

In epidemiology systematic exposure misclassification is a serious problem, because it shifts risk estimates away from unity yielding falsely positive study results. An appropriate exposure proxy is chosen in a way that misclassification will be only non-differential (i.e. the same for cases and non-cases), and as small as achievable. Such random errors generally shift risk estimates towards unity. When epidemiological studies are evaluated one has to consider potential systematic and random exposure misclassification. Thus, the following rule of thumb can be applied for interpretation of study results:

- Non-differential exposure misclassification is of particular concern in studies that show no association between exposure and disease. If so, the differentiation between 'no true association' and substantial under-estimation of the true exposure response association due to random exposure misclassification is crucial.
- Differential exposure misclassification is of particular concern in studies that showed an association between exposure and disease. In this case the observed risk can either be 'a true association' or 'a biased association' due to systematic errors in the exposure assessment.

In practice, however, it is not easy to determine whether and to what extent systematic and/or random exposure misclassification has actually occurred. Moreover, random data fluctuation, confounding or other types of bias can superpose the effects from exposure misclassification. Vrijheid and colleagues evaluated the effect from recall errors and selection bias in epidemiological studies of mobile phone use and cancer risk [10]. They showed that random recall errors of plausible levels can lead to a large underestimation of the risk of brain cancer whereas differential errors in recall had very little additional impact in the presence of large random errors. In Denmark a comparison between self-reported and operator data yielded little evidence of systematic exposure misclassification. Sensitivity for cases was 29% and for controls 31%. Specificity was 93% for cases and 95% for controls [4].

It is important for epidemiologists to deal actively with potential exposure misclassification. In the planning stage one should evaluate whether sensitivity or specificity of the exposure assessment is more crucial in a given setting. This allows spending the available funds in a most efficient way. For example, for rare exposures it is not efficient to focus on the exposed individuals, but it is important to determine accurately those who are not exposed.

Epidemiologists are also encouraged to collect data to estimate the extent of differential and non-differential exposure misclassification in their study. If such information is available with reasonable accuracy, it is possible to adjust the exposure-disease association for the presence of such errors using regression calibration or simulation extrapolation [11]. Such adjustment is extremely helpful for interpreting the results of epidemiologic studies.

References

[1] Rothman KJ, Greenland S. 1998. Modern Epidemiology, Second Edition Ed. Little, Brown and Company, Boston, MA.

[2] G. Berg, J. Schuz, F. Samkange-Zeeb, and M. Blettner, "Assessment of radiofrequency exposure from cellular telephone daily use in an epidemiological study: German Validation study of the international case-control study of cancers of the brain--INTERPHONE-Study," *J Expo Anal Environ Epidemiol*, vol. 15, pp. 217-24, 2005.

[3] F. Samkange-Zeeb, G. Berg, and M. Blettner, "Validation of self-reported cellular phone use," *J Expo Anal Environ Epidemiol*, vol. 14, pp. 245-8, 2004.

[4] J. Schuz and C. Johansen, "A comparison of self-reported cellular telephone use with subscriber data: agreement between the two methods and implications for risk estimation," *Bioelectromagnetics*, vol. 28, pp. 130-6, 2007.

[5] M. Vrijheid, E. Cardis, B. K. Armstrong, A. Auvinen, G. Berg, K. G. Blaasaas, J. Brown, M. Carroll, A. Chetrit, H. C. Christensen, I. Deltour, M. Feychting, G. G. Giles, S. J. Hepworth, M. Hours, I. Iavarone, C. Johansen, L. Klaeboe, P. Kurttio, S. Lagorio, S. Lonn, P. A. McKinney, L. Montestrucq, R. C. Parslow, L. Richardson, S. Sadetzki, T. Salminen, J. Schuz, T. Tynes, and A. Woodward, "Validation of short term recall of mobile phone use for the Interphone study," *Occup Environ Med*, vol. 63, pp. 237-43, 2006.

[6] A. Lahkola, A. Auvinen, J. Raitanen, M. J. Schoemaker, H. C. Christensen, M. Feychting, C. Johansen, L. Klaeboe, S. Lonn, A. J. Swerdlow, T. Tynes, and T. Salminen, "Mobile phone use and risk of glioma in 5 North European countries," *Int J Cancer*, vol. 120, pp. 1769-1775, 2007.

[7] G. Neubauer, M. Feychting, Y. Hamnerius, L. Kheifets, N. Kuster, I. Ruiz, J. Schuz, R. Uberbacher, J. Wiart, and M. Roosli, "Feasibility of future epidemiological studies on possible health effects of mobile phone base stations," *Bioelectromagnetics*, vol. 28, pp. 224-30, 2007.

[8] J. Berkson. "Are there two regressions?" *Journal of American State Association, vol*. 45, pp. 164-180, 1950.

[9] L. Sheppard, R. L. Prentice, M. A. Rossing. "Design considerations for estimation of exposure effects on disease risk, using aggregate data studies." *Stat Med,* vol. 15, pp. 1849-1858, 1996

[10] M. Vrijheid, I. Deltour, D. Krewski, M. Sanchez, and E. Cardis, "The effects of recall errors and of selection bias in epidemiologic studies of mobile phone use and cancer risk," *J Expo Sci Environ Epidemiol*, vol. 16, pp. 371-84, 2006.

[11] K. Fung, D. Krewski "Evaluation of quasi-likelihood, regression calibration, and SIMEX methods in logistic regression when one of the predictors is subject to additive measurement error. *J Epidemiol Biostat*, vol. 4, pp. 65–74, 1999.