

Feature Article

Why Epidemiological and Clinical Intervention Studies Often Give Different or Diverging Results?

Henrik E. Poulsen^{1,2}, Jon Trærup Andersen¹, Niels Keiding^{2,3}, Tina Ken Schramm⁴, Rikke Sørensen⁴, Gunnar Gislason⁴, Emil Loldrup Fosbøl³ and Christian Torp-Pedersen^{2,4}

¹Department of Clinical Pharmacology Q7642, Rigshospitalet, University Hospital Copenhagen, Copenhagen, Denmark

²University of Copenhagen, Faculty of Health Sciences, Copenhagen, Denmark

³Department of Biostatistics, Faculty of Health Science, Institute of Public Health, University of Copenhagen, Copenhagen, Denmark

⁴Department of Cardiology, Gentofte University Hospital, Hellerup, Denmark

In this response, we will interpret “epidemiological” as “observational studies”, and “clinical intervention” studies as “randomised studies”. A prototype epidemiological study would be a cohort study, and a prototype clinical intervention study would be a randomised, placebo-controlled double blinded study.

The question is of interest not only because there sometimes is a discrepancy between epidemiological and clinical intervention studies but also because when the subject is of interest to the public, news media coverage may provide a biased view on the magnitude of the problem, so the “often” in the question is not correct. Furthermore, detailed review of analyses/re-analysis of the discrepant trials can provide close agreement, and also provide focus on better ways of analysing epidemiological studies. We would like to draw attention to subjects that illustrate this, one concerns the benefit/harm of the use of hormone replacement therapy (HRT) and the other is the benefit/harm of the use of antioxidant vitamins.

For HRT, the interested reader should read the paper and subsequent discussion by Prentice et al. (1), and the latest conclusion (2). In this particular case, the discrepancy between epidemiology and clinical intervention studies seems to be reduced or eliminated when the important variable “time since start of exposure” is included in the analysis.

In other cases, that is, the antioxidant intervention trials (see below) other issues are important, for example, it might not be the antioxidants in the diet that are the health-promoting substances or the dose levels chosen in the clinical trials that is too high as indicated by very high plasma levels in the treated

group (3). In other cases the problems may arise due to other issues.

The basic idea behind observational studies is to examine groups of people or patients for exposures and associate the results with a disease outcome. Exposure in the epidemiological sense is very broad and can include drugs, diet, talking over telephone, etc. One example of an observational study: people who eat large amounts of fruits and vegetables have lower occurrence of certain cancers (4). Based on this, it has been recommended to the European populations that they should eat more fruits and vegetables because of a cancer preventive effect. The basic assumption behind the jump from association/exposure to clinical advice is that the association demonstrated represents a causal relation. The other possibility is that fruit intake is also related to other health-promoting factors. For an epidemiological study to be accepted as demonstrating a causal relationship, it is necessary that you can convincingly adjust the result for confounders. The problem is that there may be factors of importance and which are either unknown or poorly understood. The possibility of hidden bias may cause the analysis to provide a false result.

The basic idea behind randomised trials is different: take a defined group or population of people or patients, divide them at random into two groups, and treat the two groups (preferably blinded) with two different interventions for a defined period and measure if there is a difference in disease outcome. The strength of this approach is that the randomisation process removes bias. This approach is considered superior to the observational study, quite simply because it tests whether the change in “treatment/intervention” leads to a change in disease, whereas the observational study assumes that if such a treatment/intervention were done, it would lead to a change in disease. However, the clinical intervention trial is almost always performed in tightly defined groups, which narrows the possibil-

Received 1 September 2008; accepted 3 September 2008

Address correspondence to: Prof. Henrik E. Poulsen, Department of Clinical Pharmacology, University Hospital Copenhagen, 20 Tagensvej, Copenhagen DK-2200, Denmark.

E-mail: henrik.enghusen.poulsen@rh.regionh.dk

ity of generalising the results. A clinical intervention trial cannot always answer the clinician's questions because of ethics or other hindrances.

Because of the actual experimental intervention, the clinical intervention trial has been considered an unbeatable standard in clinical research and established the hierarchy of study designs for intended effects of therapy. However, it should also be acknowledged that observational studies allow studies of associations between a wide variety of exposures or characteristics and clinical outcomes of interest. On the other hand, clinical intervention trials study the effect of a given treatment or intervention at high cost and severe demand on logistics, and in practice cannot provide information on long-term effects. The following heavily advocated hierarchy [of study design for intended effects of intervention (5)] is thus valid only under these restrictions, and is particularly relevant in cases of new drug treatments:

1. Evidence obtained from at least one properly randomised controlled trial.
2. (a) evidence obtained from well-designed controlled trials without randomisation, (b) evidence obtained from well-designed cohort or case-control analytical studies, preferably from more than one centre or research group, (c) evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments (such as, the results of the introduction of penicillin treatment in the 1940s) could also be regarded as this type of evidence.
3. Opinions of respected authorities based on clinical experience, descriptive studies and case reports, or reports of expert committees.

With a firm belief in the hierarchy, it is very easy to answer the question "Why do epidemiological and clinical intervention studies often give different or diverging results?" by a very simple answer: Clinical intervention trials are superior to observational studies, so in case of non-agreement, the observational study was biased and therefore gives a false estimate of effect.

It has been advocated: "If you find a study was not randomised (*i.e.* a clinical intervention study in the present context), we'd suggest that you stop reading it and go on to the next article" (6). The answer, unfortunately, is not as simple as that.

We have several objections to this simplistic attitude towards epidemiological studies. It is well appreciated that also clinical intervention trials have limitations and problems. For example, the patient populations investigated often differ from those receiving drugs in everyday practice. Trials on treatments that give effects after long-time periods, years or decades, can suffer from baseline drift and changes in life-style, etc. that can obscure intervention effects. Sometimes it is impossible to fulfil the very important blinding, and difficult to sustain the intervention, for example, calorie-restriction or other dietary intervention. Also, effects may be dependent on unknown factors, for

example, genetic factors or variables like single nucleotide polymorphisms, where one non-classifiable group will have opposite effects to the remaining part of the population, and bring about an overall negative and deleterious effect, even though a large subgroup may benefit from the intervention. Finally, clinical intervention trials are extremely resource demanding in time, money, intellectual resources, available subjects/patients, etc. In summary, randomised clinical trials are only performed for highly selected questions, their complexity results in short follow-up and populations, which questions the general validity.

Therefore, to only conduct and accept clinical intervention trials will greatly limit knowledge and it is simply necessary to include epidemiological studies while accepting their limitations. Fortunately, techniques have been developed that increase the validity of the epidemiological approach. A multivariable approach using techniques such as regression models, for example Cox regression, has improved the analysis of epidemiological studies, but also has limitations. The use of propensity scores and matching has in many cases provided very close relationships between randomised studies and observational studies. Also, case-crossover and related techniques, where each patient serves as his own control at a different time period have provided important results in situations where a randomised study is impossible.

Regarding adverse drug or treatments effects, it is very costly and difficult to conduct clinical trials. Adverse effects must be assumed to be much less frequent than the intended effects, so to get sufficient power in the trials on adverse effects, the demand on observation number and so forth increases dramatically, and in many cases makes a randomised study impossible.

A first example is the divergent results between epidemiology and clinical intervention studies: large scale observational studies showed a 35–45% reduction of myocardial infarction with hormone replacement therapy (HRT) with oestrogen and progesterone (7), which was in line with estrogens protecting women from development of arteriosclerosis. A controlled trial challenged this (8), including evidence from three further intervention trials, and a meta-analysis concluded that "HRT users had a significant increased incidence of breast cancer, stroke, and pulmonary embolism, and significantly reduced incidence of colorectal cancer and fractured neck of the femur"; there were no significant change in coronary heart disease (9). The impact on HRT use was profound, with a reduction to 50% (10). As mentioned in the introductory paragraphs above, careful reanalyses have been performed, which make these apparent discrepancies much less dramatic or even non-existent.

A second example is that of beta-carotene, where numerous observational studies showed reduced cancer incidence in groups with high intake of vegetables containing beta-carotene [see references in ref. (4)]. A large intervention trial (3) showed an opposite effect with increased incidence of lung cancer in

subjects given tablets with beta-carotene. The same study showed that in the placebo group; pre-trial levels of beta-carotene in plasma were associated with a reduced lung cancer incidence.

Agreement between epidemiology and clinical intervention studies has been summarised in New England Journal of Medicine:

In 21 observational studies reported between 1985 and 1998 and matching clinical intervention trials there was little evidence of larger estimates of effects in observational studies compared with intervention trials (11). Among the 21 studies, 19 different intervention studies provided similar estimates as in controlled trials, and only slightly higher estimates in two epidemiological studies compared with intervention trials. This study indicates that the discrepancy between clinical and epidemiological studies is not as profound as suggested by the stories in news media.

Examples of studies, where randomised studies are non-feasible:

A classical example where randomisation is impossible is the relationship between the use of mobile phone and driving and crashes. With the use of a case-crossover design, it could be established that the use of a mobile phone 10 min before a crash was associated with a four-fold increased likelihood of crashing (12). Another example is malformation due to drug use in relationship with pregnancy. Using an epidemiological approach and the nationwide Danish records on drug dispensing from pharmacies and the nationwide national registers of hospital discharge diagnoses, we have been able to find increased rates of malformation with the use of certain antibiotics. Again, a randomised study could never be carried out, as it would be un-ethical, logistically and financially impossible to randomise almost a million potentially pregnant women to treat with an active drug or placebo. Another example is the use of non-steroidal anti-inflammatory drugs in the general and healthy population. We showed that the risk of death is increased with high doses of such drugs, with multivariate Cox analysis, with propensity score analysis, and with case-crossover analysis of more than 1 million people (13). Such a design minimises the risk of uncontrolled bias. A randomised study of that size is theoretically possible, but will never be done because of logistic and fiscal limitations.

Although the hierarchy of study design, given above, clearly places the randomised controlled trial at the top, one should realize for discovery, explanation and idea generation; the progress still depends on a process in the reverse order, starting with anecdotal cases and series, and observations of epidemiological nature.

A critical and intelligent scrutiny of clinical data, which being from clinical intervention trials or from epidemiological studies, is always needed to progress our understanding and intervention strategy in prevention and disease.

REFERENCES

1. Prentice, R. L., Pettinger, M., and Anderson, G. L. (2005) Statistical issues arising in the Women's Health Initiative. *Biometrics* **61**, 899–911.
2. Prentice, R. L., Chlebowski, R. T., Stefanick, M. L., Manson, J. E., Langer, R. D., Pettinger, M., Hendrix, S. L., Hubbell, F. A., Kooperberg, C., Kuller, L. H., Lane, D. S., McTiernan, A., O'Sullivan, M. J., Rossouw, J. E., and Anderson, G. L. (2008) Conjugated equine estrogens and breast cancer risk in the Women's Health Initiative clinical trial and observational study. *Am. J. Epidemiol.* **167**, 1407–1415.
3. Albanes, D., Heinonen, O. P., Huttunen, J. K., Taylor, P. R., Virtamo, J., Edwards, B. K., Haapakoski, J., Rautalahti, M., Hartman, A. M., and Palmgren, J. (1995) Effects of alpha-tocopherol and beta-carotene supplements on cancer incidence in the alpha-tocopherol beta-carotene cancer prevention study. *Am. J. Clin. Nutr.* **62**, 1427S–1430S.
4. Vainio, H. and Bianchini, F., ed. (2003) *IARC Handbooks of cancer prevention: Fruit and Vegetables*. IARC Press: Lyon France.
5. Concato, J., Shah, N., and Horwitz, R. I. (2000) Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N. Engl. J. Med.* **342**, 1887–1892.
6. Sackett, D. L., Richardson, W. S., Rosenberg, W., and Haynes, R. B. (1997) *Evidence based medicine: how to practice and teach EBM*. Churchill Livingstone: New York.
7. Stampfer, M. J. and Colditz, G. A. (1991) Estrogen replacement therapy and coronary heart disease: a quantitative assessment of the epidemiologic evidence. *Prev. Med.* **20**, 47–63.
8. Rossouw, J. E., Anderson, G. L., Prentice, R. L., LaCroix, A. Z., Kooperberg, C., Stefanick, M. L., Jackson, R. D., Beresford, S. A., Howard, B. V., Johnson, K. C., Kotchen, J. M., and Ockene, J. (2002) Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the women's health initiative randomized controlled trial. *JAMA* **288**, 321–333.
9. Beral, V., Banks, E., and Reeves, G. (2002) Evidence from randomised trials on the long-term effects of hormone replacement therapy. *Lancet* **360**, 942–944.
10. Lokkegaard, E., Lidegaard, O., Moller, L. N., Agger, C., Andreasen, A. H., and Jorgensen, T. (2007) Hormone replacement therapy in Denmark, 1995–2004. *Acta. Obstet. Gynecol. Scand.* **86**, 1342–1351.
11. Benson, K. and Hartz, A. J. (2000) A comparison of observational studies and randomized, controlled trials. *N. Engl. J. Med.* **342**, 1878–1886.
12. McEvoy, S. P., Stevenson, M. R., McCart, A. T., Woodward, M., Haworth, C., Palamara, P., and Cercarelli, R. (2005) Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study. *BMJ* **331**, 428.
13. Fosbol, E. F., Gislason, G. H., Jacobsen, S., Folke, F., Hansen, M. L., Schramm, T. K., Sørensen, R., Rasmussen, J. N., Andersen, S. S., Abildstrom, S. Z., Andersen, J. T., Poulsen, H. E., Rasmussen, S., Kober, L., Torp-Pedersen, C. (unpublished) Risk of myocardial infarction and death associated with the use of non-steroidal antiinflammatory drugs (NSAID) among healthy individuals: a nationwide cohort study.