# Cohort Studies: Prospective versus Retrospective

Anne M. Euser[a]    Carmine Zoccali[c]    Kitty J. Jager[b]    Friedo W. Dekker[a, b]

[a]Department of Clinical Epidemiology, Leiden University Medical Centre, Leiden, [b]ERA-EDTA Registry, Department of Medical Informatics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; [c]CNR-IBIM Clinical Epidemiology and Pathophysiology of Renal Diseases and Hypertension, Renal and Transplantation Unit, Ospedali Riuniti, Reggio Calabria, Italy

**Abstract**

Cohort studies form a suitable study design to assess associations between multiple exposures on the one hand and multiple outcomes on the other hand. They are especially appropriate to study rare exposures or exposures for which randomization is not possible for practical or ethical reasons. Prospective and retrospective cohort studies have higher accuracy and higher efficiency as their respective main advantages. In addition to possible confounding by indication, cohort studies may suffer from selection bias. Confounding and bias should be prevented whenever possible, but still can exert unknown effects in unknown directions. If one is aware of this, cohort studies can form a potent study design in nephrology producing, in general, highly generalizable results.

Copyright © 2009 S. Karger AG, Basel

## Introduction

Despite its high ranking in the hierarchy of clinical intervention studies [1], in nephrology practice the randomized clinical trial (RCT), as discussed in the previous article in this series, is frequently not possible due to practical and ethical constraints [2]. In these situations, the cohort study – either prospective or retrospective – often forms a suitable observational study design to reliably answer various research questions [3]. The aim of this article is to describe several issues related to the design of cohort studies, including their strengths and weaknesses, with special focus on the nomenclature of prospective versus retrospective study designs.

## Design

A key characteristic of a cohort study is that at the starting point of the study the subjects are identified and their exposure to a risk factor is assessed. Subsequently, the frequency of the outcome, usually the incidence of disease or death over a certain time span, is measured and related to exposure status. In this way, the effect of exposure on outcome can be expressed as a relative risk. This risk factor may be binary (risk factor present = index

Friedo W. Dekker, PhD
Department of Clinical Epidemiology/Directorate of Education
Leiden University Medical Centre, Room C7-G 91, PO Box 9600
NL–2100 RC Leiden (The Netherlands)
Tel. +13 73 126 2381, ext. 1210, Fax +13 73 126 6994, E-Mail f.w.dekker@lumc.nl

**Fig. 1.** Graphical representation of the timeline in a retrospective and prospective cohort study design.
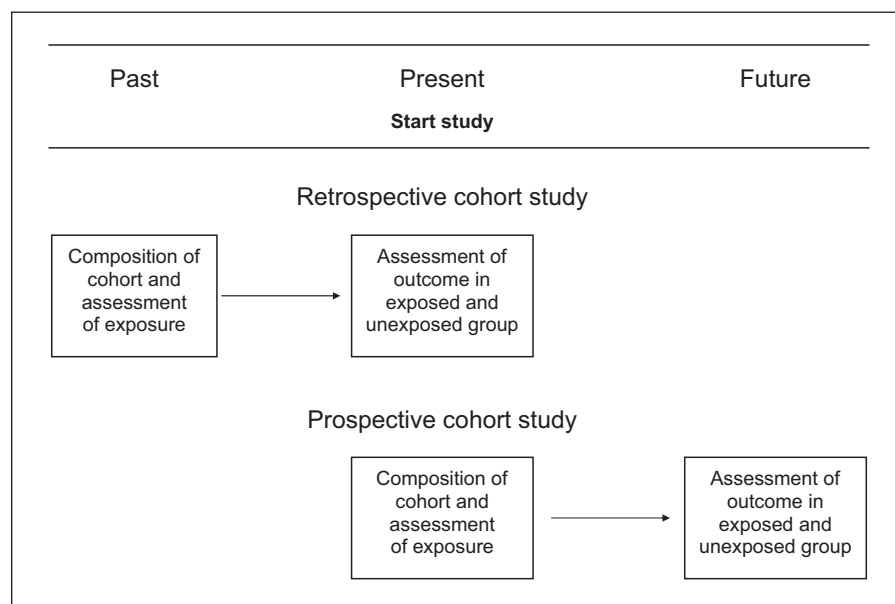
group versus not present = reference group), for instance diabetes yes/no. In contrast with an RCT, in a cohort study this exposure is not randomly assigned. Instead, exposure status is acquired by chance (e.g. genetic polymorphisms) or by choice (e.g. smoking). Alternatively, the risk factor can be continuous, for instance serum phosphate level or body mass index (BMI).

In a cohort study, usually one group of people is followed over time, e.g. new dialysis patients, and many exposures can be measured at baseline. Alternatively, one can select 2 specific groups at baseline on having or not-having the exposure, e.g. gadolinium-contrast exposure versus a sample of comparable people without gadolinium-contrast exposure, and subsequently study the occurrence of renal disease.

*Example 1*

An example of a cohort study is provided by de Mutsert et al. [4] who studied the association between nutritional status as measured by subjective global assessment (SGA) and mortality in chronic dialysis patients. To that end, a cohort was formed of Dutch incident dialysis patients who started their first renal replacement therapy (the Netherlands Cooperative Study on the Adequacy of Dialysis NECOSAD-II study cohort). Baseline was defined as 3 months after the start of dialysis and at this time, the SGA was performed in all 1,601 participants. Subsequently, subjects were followed-up until a maxi-

mum of 7 years after the start of the study, with a mean follow-up time of 2.7 years. SGA levels were divided into 3 categories. It was shown that, adjusted for age, sex, treatment modality, primary kidney disease and co-morbidity, severe protein wasting at baseline was associated with a 2-fold increase in mortality.

**Prospective versus Retrospective**

Usually, 2 types of cohort studies are distinguished: those that are prospective and those that are retrospective. The study described above is a typical example of a prospective cohort study in which exposure is assessed at baseline and the researcher follows the subjects in time to study the development of disease or mortality. In a retrospective design, the researcher starts the study at the time follow-up has already been completed. Retrospectively, eligible subjects are identified, a cohort is composed and exposures are assessed at baseline. Thereafter, the subsequent disease occurrence or death is studied during the historical observation period (fig. 1). Typically, a prospective design has been ranked higher in the hierarchy of evidence than a retrospective design [1]. However, as argued by Vandenbroucke [5], some epidemiologists consider any follow-up study synonymous with 'prospective', as follow-up always goes forward in time. Besides, studies are often arbitrarily labeled 'prospective' in an attempt to

increase their impact. Altogether, this may lead to a lot of confusion and semantics. Both forms have their own merits and possible weaknesses – as will be discussed below – but should not be classified automatically into superior and inferior. Instead of using the labels prospective or retrospective too easily, researchers should better give an explanation of what exactly has been done in the abstract and methods section [5].

*Example 2*

A retrospective cohort study was published by Voormolen et al. [6] who studied the association between plasma phosphate and decline in renal function in pre-dialysis patients. The study has a retrospective design, because in the year 2003 incident pre-dialysis patients (chronic kidney disease stage IV–V) were included who were referred to pre-dialysis care in the years 1999–2001. These patients were identified in the administration of the participating hospitals, and the laboratory measurements at baseline were noted from the patient files. Subsequently, the medical course of these patients, especially the decline in renal function, was followed through the medical charts until start of dialysis, death, or 1 January 2003. From these data it could be calculated that renal function declined faster with higher phosphate levels at baseline. Also, a relative risk of death (of 1.25) could be calculated for every mg/dl increase in phosphate. Thus high plasma phosphate showed to be an independent risk factor for a more rapid decline in renal function and a higher mortality during the pre-dialysis phase.

## Strengths and Weaknesses

As mentioned above, apart from their close relationship, prospective and retrospective cohort studies do have different strengths and weaknesses. The major strength of a prospective cohort study is the accuracy of data collection with regard to exposures, confounders, and endpoints, but this is realized at the cost of an inevitable loss of efficiency, for this design is both expensive and time-consuming because of a usually long follow-up period. Vice versa, the retrospective design is a very time-efficient and elegant way of answering new questions with existing data, but one has no choice other than to work with what has been measured in the past, often for another purpose (e.g. patient care) than the one under investigation.

A major advantage of cohort studies in general is the possibility to study multiple exposures and multiple outcomes in one cohort. Even rare exposures can be studied, for the index group can be selected on this exposure. Besides, the combined effect of multiple exposures on disease risk can be determined, e.g. the effect of low birth weight and prematurity on adult renal function [7]. Hypothesis generation is another advantage that has been associated with cohort studies. However, no study generates hypotheses – only researchers do, using study data. Therefore, instead the term 'hypothesis screening' has been proposed by Rothman et al. [8], as, despite possible biases, a cohort study is considered to be a relatively easy way to pick up associations between many exposures and outcomes. For example, in a population-based cohort study it can be studied quite simply that a BMI is associated with an increased risk of chronic kidney disease [9]. Yet, this doesn't clarify how this association can be explained, for BMI is only the result of and a proxy for many other variables, which are partly unknown. Several underlying etiological hypotheses can be generated now, which can subsequently be tested in other so-called confirmatory studies, often with a more experimental design. Finally, because a cohort study has usually broader inclusion criteria and less exclusion criteria compared to an RCT, its results may be more generalizable to clinical practice.

On the other hand, a major disadvantage of cohort studies is that it is not possible to establish causal effects. The exposure has not been allocated randomly and there is always a possibility that the association found may be explained by other variables that differ between exposed and non-exposed subjects and that also have an association with the outcome studied, so-called confounders. If these other variables were measured they can be adjusted for in the analysis, but frequently these factors are unmeasured, measured imprecisely, or even unknown. For instance, when comparing survival in hemodialysis versus peritoneal dialysis patients in a cohort study, statistical adjustment for known confounders may not suffice to arrive at 2 groups that have truly the same prognosis at baseline. In that case, the comparison suffers from confounding by indication. Moreover, cohort studies are susceptible to selection bias. For example, when the risk of decreased renal function is tested in obese versus non-obese subjects in a secondary care setting of all patients newly referred to an internist, selection bias will occur if general practitioners (GPs) have especially referred obese patients with albuminuria. This because GPs already suspected decreased renal function in them because of the obesity, while in slim people the GP may have refrained from proteinuria testing.

This so-called referral bias is a form of selection bias, because the index group has been formed not only by exposure, but was also more likely to have the disease of interest at the beginning of the study. Next, selection bias can be introduced in a cohort study by a low response if this non-response is selective, i.e. different in those people that have both the exposure and an increased risk of developing the disease. This could happen when prevalent instead of incident patients are used to form a cohort of dialysis patients; some of the patients already died due to the risk factor studied before they could have been included in the cohort. A comparable bias can be introduced not at the time of inclusion, but with a selective loss-to-follow-up. Loss-to-follow-up is almost never completely random. Often, disease status cannot be measured because subjects do not show up as they have no complaints and/or are too busy, or on the other hand, they are too ill to go anywhere. This will bias the results when this selective follow-up rate differs between index and reference group. When for example, quality of life is studied in incident hemodialysis versus peritoneal dialysis patients, those with a low quality of life and a depression will not fill in the questionnaire. If this percentage of depression with concomitant non-response is higher in one of the treatment modalities, the outcome of the remaining subjects in this group will form an inflated, biased score. Obviously, this form of selection bias can only be prevented by assuring a high percentage of participation and follow-up.

## Conclusion

Cohort studies form a suitable study design to assess associations between multiple exposures on the one hand and multiple outcomes on the other hand. They are especially appropriate to study rare exposures or exposures for which randomization is not possible for practical or ethical reasons. Prospective and retrospective cohort studies have higher accuracy and higher efficiency as their respective main advantages. In addition to possible confounding by indication, cohort studies may suffer from selection bias. Confounding and bias should be prevented whenever possible, but still can exert unknown effects in unknown directions. If one remains aware of this, cohort studies can form a potent study design in nephrology, producing most of the time highly generalizable results.

## References

1 Vandenbroucke JP: Observational research, randomised trials, and two views of medical science. PLoS Med 2008;5:e67.
2 Noordzij M, Dekker FW, Zoccali C, Jager KJ: Study designs in clinical research. Nephron Clin Pract 2009;113:c218–c221.
3 Jager KJ, Stel VS, Wanner C, Zoccali C, Dekker FW: The valuable contribution of observational studies to nephrology. Kidney Int 2007;72:671–675.
4 de Mutsert R, Grootendorst DC, Boeschoten EW, Brandts H, van Manen JG, Krediet RT, et al: Subjective global assessment of nutritional status is strongly associated with mortality in chronic dialysis patients. Am J Clin Nutr 2009;89:787–793.
5 Vandenbroucke JP: Prospective or retrospective: what's in a name? BMJ 1991;302:249–250.
6 Voormolen N, Noordzij M, Grootendorst DC, Beetz I, Sijpkens YW, van Manen JG, et al: High plasma phosphate as a risk factor for decline in renal function and mortality in pre-dialysis patients. Nephrol Dial Transplant 2007;22:2909–2916.
7 Hallan S, Euser AM, Irgens LM, Finken MJ, Holmen J, Dekker FW: Effect of intrauterine growth restriction on kidney function at young adult age: the Nord Trøndelag Health (HUNT 2) Study. Am J Kidney Dis 2008;51:10–20.
8 Rothman KJ, Greenland S, Lash TL: Types of epidemiological studies; in Modern Epidemiology. Philadelphia, Lippincott Williams and Wilkins, 2008, pp 87–99.
9 Hallan S, de Mutsert R, Carlsen S, Dekker FW, Aasarod K, Holmen J: Obesity, smoking, and physical inactivity as risk factors for CKD: are men more vulnerable? Am J Kidney Dis 2006;47:396–405.