

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# Prospective Cohort Studies in Medical Research

---

Samer Hammoudeh, Wessam Gadelhaq and  
Ibrahim Janahi

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.76514>

---

## Abstract

Cohort studies are the analytical design of observational studies that are epidemiologically used to identify and quantify the relationship between exposure and outcome. Due to the longitudinal design, cohort studies have several advantages over other types of observational studies. The purpose of this chapter is to cover the various characteristics of prospective cohort studies. This chapter is divided into three main sections. In the first we introduce the concept and ranking of cohort studies, as well as the advantages and disadvantages. In the second we focus on the design of cohort studies, mainly its prospective aspect, and the distinguishing features from the retrospective type. The section also covers the essential characteristics of a cohort study design and its varied applications in medical research. In the third we go over examples of prospective studies in the medical field. For each, an overview of the study design is given, along with a random selection of study findings/impact, strengths and weaknesses.

**Keywords:** observational study, cohort study, prospective cohort study, longitudinal study, study design, epidemiology, medical research

---

## 1. Cohort studies

### 1.1. Introduction

The term “cohort” originates from Latin “cohors” [1]. A term that was used in the military back in Roman times, which referred to a unit that is comprised of 300–600 men, of which each 10 cohorts were named a legion [2]. In the field of epidemiology, Frost was the first to introduce the term “cohort study” back in 1935 [3]. Cohort refers to a group of individuals that share a common factor or a defining characteristic [4, 5], or in other words, cohort is a

certain component of a specific population that can be measured and followed throughout time [6]. Cohort studies are classified under the non-experimental type of studies [4], which are observational by default [7].

A cohort study follows people as groups, two or more, from exposure to outcome [2, 8]. The two groups would be categorized based on their exposure status to “exposed” and “unexposed” [4, 9, 10]. If there were multiple groups then these would be categorized either by the type or level of exposure [4]. The main characteristic of a cohort study is that it follows participants in a forward manner, from the presence of the exposure to the presence of the outcome [2, 9–11]. Or as De Rango describes it: using a longitudinal pattern, a cohort study, follows a group or groups of individuals over time in order to ascertain the incidence of a predetermined outcome after being exposed to a certain factor, whether being a risk factor, medication, or intervention [12]. Cohort studies can either be prospective (concurrent) or retrospective (non-concurrent) [9].

## **1.2. Ranking of cohort studies**

Researchers agree that cohort studies, as related to the hierarchy of evidence, rank below meta-analysis, systematic review and randomized controlled trial, but rank higher than case-control studies, cross sectional studies, case series/reports [13–16]. As newer models or classifications of the hierarchy of evidence have emerged, where meta-analysis and systematic reviews have been removed from the hierarchy and repositioned as a magnifying glass or a lens through which evidence from other types of studies can be viewed or scrutinized; cohort studies remain below randomized controlled trials and higher than the other types [17]. Cohort studies provide information on the relationship between exposure and outcome when a randomized controlled trial is not possible to conduct for whatever reason [6, 15].

## **1.3. Advantages of cohort studies**

Cohort studies are the design of choice when randomization is not practical or ethical [6, 18]. They are also useful in the study of infections [9] and for hypothesis generation [19]. Due to the design of cohort studies, and since temporal sequence is present, both incidence rate and cumulative incidence can be calculated [2, 8, 20–22]. They also allow for the measurement of relative risk (RR) [2, 8, 23], hazard ratio [8], and attributable risk [8, 23]. Furthermore, they allow for the study of multiple outcomes that can be associated with a single type of exposure [2, 20] or multiple exposures [18]. Additionally, they allow for the study of rare exposures [2, 18, 20]. Finally, cohort studies have lower risk of encountering survivor bias [2], and recall bias [9, 21]. Survivor bias occurs when focusing only on those who survived or made it through a certain criteria or point, and ignoring those that didn’t, such as studying rapidly fatal diseases [2].

## **1.4. Disadvantages of cohort studies**

Among the disadvantages of cohort studies is selection bias, which may occur when the participants are not representative of the population or of the patient grouping that they fall under. This in turn will influence how well or not the results can be generalized to the rest of

the population, in what is known as external validity [2, 12, 18, 24, 25]. This will be covered later in section three of this chapter under aspects of cohort studies. Another disadvantage is that causation cannot be established from cohort studies [18, 20], as it would require an experimental design in order to determine any causal effect [20]. However, due to the longitudinal design of cohort studies, they may aid in studying a certain causal hypothesis [20]. A third disadvantage is that they require a large sample size, which might pose an issue when dealing with outcomes that take a long time to develop [10]. Finally, cohort studies cannot be used to study rare outcomes [23].

## 2. Prospective cohort studies

### 2.1. Types of cohort studies

Cohort studies are either prospective or retrospective [1, 2, 18]. In the former, the researcher would assess exposure at baseline and then follow the person over time in order to determine the outcome such as the development of a disease [9, 18, 20, 21, 26]. In the latter, the order is reversed, as a cohort is established after the follow up has been conducted, or the outcome has developed, and exposure is then assessed in a retrospective manner [9, 18, 20, 21, 27]. Merrill indicates that the outcome status at the start of the study is what determines the overall study type. If the outcome has not yet developed then it is a prospective study, and if the outcome has already developed then it is a retrospective study [23]. Cohort studies can also be classified based on whether or not participants are replaced once they are lost. If those that drop out or are lost to follow up are replaced with new participants, then this would be classified as a dynamic or an open cohort. In the case that those lost do not get replaced, then it would be classified as a fixed or closed cohort [4, 20].

#### 2.1.1. *Prospective cohort studies*

Prospective cohort studies, as the name indicates, observes a group of people after being exposed to a certain factor in order to investigate the outcome, following the natural sequence of time, starting with the present and looking forward in time [12, 18, 20], which in turn provides true risk (absolute) estimates for the groups under investigation [26]. It is considered the gold standard among observational studies [8]. Under this type of study, the researcher would have control over data collection methodology, as well as the overall cohort study set up, which gives prospective cohort studies an advantage over retrospective cohort studies [9]. Further advantages and disadvantages of prospective cohort studies are discussed below.

##### 2.1.1.1. *Advantages of prospective cohort studies*

Euser et al. highlight the major advantage of prospective cohort studies as being accurate in regards to the information collected about exposures, endpoints, and confounders [18]. Others list the following as advantages of prospective cohort studies; first: the exposure has already been measured before the outcome has occurred, which allows for the assessment of

temporal sequence [28]. This allows for the calculation of incidence and the determination of the disease process [2, 12, 20, 23]. Second: elimination of recall bias, as there is no need for any recollection of information since the data is being collected in a prospective manner [7]. However, Kip et al. reported that recall bias can pose an issue in prospective cohort studies if the exposure is self-reported, brief, and requires multiple measurements, such as stress episodes [29]. Third: It allows for the study of exposures where randomization is not practical or ethical [12]. Fourth: it allows for the study of rare exposures [20]. Fifth: it allows for the study of multiple outcomes [20, 26].

#### *2.1.1.2. Disadvantages of prospective cohort studies*

Among the disadvantages of prospective cohort studies is the loss to follow up, which is common among cohort studies. This can ultimately lead to differential loss to follow up among those exposed and unexposed, which in turn can complicate the interpretation of the results [2, 7, 12, 18, 24]. Another disadvantage is that they are time consuming if follow up periods are far apart. This would be resource consuming as well, which would make prospective cohort studies not suitable for the study of outcomes that take long time to develop [18, 20, 24, 26]. A third disadvantage is that they are expensive to conduct [18, 20, 30]. The third section of this chapter is dedicated to providing examples of prospective cohort studies.

#### *2.1.2. Retrospective cohort studies*

As previously described, retrospective cohort studies, also known as historic [28] or historical [24] cohorts, use data that has already been collected, such as databases of healthcare records, in order to investigate the association between the exposure and the outcome [22, 24, 26, 28]. Although the outcome has already occurred, the design of retrospective cohort studies is similar to those of prospective cohort studies [22]. They also have similar advantages and disadvantages [26, 28]. Hess indicates that retrospective studies in general are useful as pilot studies for future prospective studies [31].

Retrospective cohort studies have advantages and disadvantages. They are time efficient and cheap since the data has been collected previously and is available for scrutiny [18, 20, 26]. Additionally, since the exposure has already been measured before the outcome has occurred, this allows for the assessment of temporal sequence [28]. However, retrospective cohort studies use information that has been collected in the past for another objective other than the current study [18], and in some cases, collected for a purpose that is not related to medical research [9]. Due to this factor, the investigator lacks control over the collection of data [24, 26, 27]. Additionally, the measurement of exposure and outcome might be inconsistent or inaccurate, which can become a source of bias [24, 27, 28, 31, 32].

Examples of retrospective cohort studies:

High plasma phosphate as a risk factor for decline in renal function and mortality in pre-dialysis patients [18, 33]. In this study, Voormolen et al. followed the clinical course among incident pre-dialysis patients, using medical charts, to study the decline in kidney function and its association with plasma phosphate levels [18, 33].

Assessment of female sex as a risk factor in atrial fibrillation in Sweden: nationwide retrospective cohort study [28, 34]. In this study, Friberg et al. investigated gender differences in the incidence of stroke among those with atrial fibrillation using the Swedish hospital discharge registry [28, 34].

Outcomes of care by hospitalists, general Internists, and family physicians [35]. In this study, Lindenauer et al. collected data from various hospitals in the USA, and compared the outcome of patients treated by the three types of care provider [35].

### 2.1.3. Aspects of cohort studies

#### 2.1.3.1. Validity

Validity is the epidemiological assessment to the lack of systematic error [4, 11]. There are two types of validity: internal validity and external validity [4, 11, 25]. *Internal validity* refers to the inferences made from the study that are related to the same source population [4, 5, 11, 25, 36], as to whether or not the study has measured what it had originally planned on measuring [25, 36]. For an example, if the exposure caused the observed change in the outcome, then the study would be considered to have high internal validity [11]. On the other hand, if the observed change in the outcome was caused by a systematic error (bias), then the study would be considered to have low internal validity [11]. Threats or violations to internal validity will be discussed later in this section under bias.

*External validity* refers to the degree to which the study results can be generalized to other populations [4, 5, 11, 25, 36]. For example, if the study participants were not representative of the general population, then the study results cannot be generalizable to others [12]. The highest level of external validity occurs when the results can be generalized to three other domains: other populations, other environments, and other times [36]. External validity can be improved by using random selection [37].

It is essential to have internal validity in order to establish external validity; that is the study must have internal validity in the first place in order to have external validity [4, 11]. For an example, if the exposure caused the observed change in the outcome, then the results can be generalizable to others. If the observed change was caused by any other factor, then the results cannot be generalized to others [4, 11]. Based on the validity hierarchy, cohort studies are considered to have low internal validity, while the external validity is high [11, 16].

#### 2.1.3.2. Bias

Bias is a study systematic error in the design, conduct, or analysis that can be categorized into three main categories: selection bias, information bias, and confounding [4, 25, 38]. *Selection bias* occurs when the sample chosen for the study is not obtained randomly, so that the sample chosen is no longer representative of the overall population [4, 25, 38, 39]. This type of bias includes three types: attrition bias, non-respondent bias, and the healthy entrant effect [38]. Attrition bias, or loss to follow up bias, occurs due to dropouts or death, which can be encountered in studies with long follow up durations (prospective) [23]. Non-respondent bias occurs



when those that respond are different than those that don't respond. For example, nonsmokers are more likely to return questionnaires about smoking than smokers are [25]. The healthy entrant effect or the healthy worker effect occurs when there are differences between those that are exposed and those that are not exposed. For an example, when comparing working individuals to the general population, as workers are more likely to be healthier than the general population. In order to avoid this type of bias, it is recommended to use two similar groups, such as using two groups of working individuals [23].

*Information bias* (measurement bias) [25], occurs when the data obtained is being recorded inaccurately [4, 25, 38–40]. This type of bias can be differential (nonrandom) or nondifferential (random) as related to the outcome [4, 9, 23, 25]. The former is dependent on other variables and leads to overestimation or underestimation of any possible association, while the latter is independent from other variables and leads to underestimation of any possible association [4, 9, 23], and if the exposure was dichotomous, this type leads to bias towards the null [9]. Non differential is more commonly encountered in cohort studies [9]. Information bias can be reduced by using standardized assessment tools that have been validated [9]. Information bias is also known as classification bias, observation bias [25], or misclassification bias [23].

*Confounding*: confounding is a distortion of the effect [4, 25] that may lead to overestimation or underestimation of an effect, or even reversing the direction of an effect [4]. A confounding factor is a risk factor that is associated with the exposure and influences the outcome, however, is not related to the causation sequence [4, 25, 39]. Unlike selection and information bias, confounding can be controlled for prior to study initiation, or after study completion [25]. Controlling for confounding factors can be accomplished through: restriction, matching, stratification, and using multivariate techniques [23, 25, 27].

Restriction would involve excluding those with the confounding factor [23, 25]. If the confounding factor is categorical, then participants that fall within that category would be excluded [4], such as if smoking was considered to be a confounding factor, then those that smoke would be excluded [25]. If the confounding factor was continuous, such as age, then a range of that variable would be used to restrict the confounding [4]. Matching would involve choosing two groups that are similar to each other as much as possible [23, 25, 41], such as matching by gender or age [39]. Matching can be either individual matching or frequency matching. The former involves matching on an individual participant level, while the latter refers to matching on a group level [4]. Overmatching may occur when matching is being used, which may reflect on the statistical efficiency, validity, or cost efficiency of the study [4]. After the completion of the study, and during the analysis stage, stratification can be used to control for confounding by dividing the groups into several subgroups that are based on the confounding factor [23, 25, 39, 41]. Multivariate techniques are also used during the analysis stage and allow for the control of multiple factors [25, 39, 41].

#### 2.1.3.3. *Exposure and risk*

Exposure must be determined using a clear and accurate definition [2, 22], which in some cases may involve levels of exposure [2]. This helps in eliminating possible selection bias [2]. The challenge becomes greater when there are multiple exposure assessments over an

extended period of time [30]. The validity and the cost are two important aspects that must be taken into consideration when selecting an exposure measurement tool [30].

Both groups, those exposed and those that are not exposed should be at risk of eventually developing the outcome at some stage [2]. The exclusion criteria should exclude those that are not at risk of developing the outcome [24]. For an example, a study investigating the role of antipsychotics in the development of diabetes, should exclude those with diabetes to start with, since they are not at risk [10]. This helps in eliminating possible selection bias [2].

#### *2.1.3.4. Outcomes*

Outcomes should have a clear and specific definition from the beginning of the study [2, 22], which must be measurable as well [2, 22]. Outcomes should also be measured in a similar manner across all participants [2, 22]. This helps in eliminating possible information bias [2]. It is recommended to use measurement tools that have been previously validated when dealing with secondary data, and to blind those who are assessing the outcome when dealing with primary data [10].

#### *2.1.3.5. Controls*

The comparison group or controls (unexposed group) should be similar to the exposed group in all possible aspects, but differ in regards to the exposure itself [2]. Three types of controls can be used, with the first being the most preferable: internal comparisons, other external cohorts, and the general population [2].

#### *2.1.3.6. Follow up*

To avoid loss to follow up and its consequent effects on the validity of the study results; measures should be taken in order to minimize the attrition rate [2, 22, 24, 27, 42]. Some of these actions include excluding those that are at high risk of not committing to the study, providing incentives for participation, collecting personal information that would allow or facilitate future contact, and maintaining ongoing contact on regular basis during the conduction period of the study [2, 23, 24, 27]. The maximum acceptable limit for loss to follow up is 20% [23, 24, 42].

#### *2.1.3.7. Precision*

Precision is based on the absence of random error or chance [4, 11]. This random variation can be due to the sample itself, or how it was selected, or how it was measured [4, 11]. Standard deviations and confidence intervals are useful in determining the precision of a study, as a large standard deviation or a wide confidence interval would indicate low precision [11]. Random error or variation can be reduced by increasing the sample size [4, 27, 43], improving how you sample and how you measure, in addition to using the appropriate statistical methods [43].

#### *2.1.3.8. Analysis of data*

The main statistical term or product of cohort studies is the relative risk or risk ratio [6, 21], which represents the risk of developing the outcome among those that are exposed in relation



to those that are not exposed [20]. An RR that is equivalent to 1 indicates an absence of any type of association. An RR that is greater than 1 would indicate that there is a positive correlation between the exposure and risk of developing a disease. An RR that is smaller than 1 would indicate the presence of a protective effect between the exposure and the outcome [12]. Other outcome measures include: hazard ratios, survival curves, and life-table rates [2]. Some of the common statistical analysis involving cohort studies include: analysis of variance (ANOVA), multivariate analysis of variance (MANOVA), mixed effect regression model, and generalized estimating equation models [7].

#### 2.1.3.9. Reporting

The reporting of prospective cohort studies should follow the STROBE guidelines [12], which also apply to other observational studies [41, 44]. This acronym stands for: Strengthening the Reporting of Observational Studies in Epidemiology. These guidelines were designed by a group of international scholars including journal editors, epidemiologists, statisticians and researchers in order to set universal standards when reporting observational studies. It is comprised of a 22 item checklist that precisely dictates what should be reported under each section of an article [44–47]. Sessler and Imrey indicate that the most crucial ones are related to the study: objectives, methodology, definitions, source of data, statistical analysis, participants, and results [41]. Further information can be found at <http://www.strobe-statement.org/>.

Bookwala et al. outlined three main factors that aid in evaluating prospective cohort studies in their article titled “the three-minute appraisal of a prospective cohort study”. These are related to (1) comparison groups selection; (2) the impact of confounding variables; (3) type of analytical strategy used [48]. Finally, the equator network (which is supported by the University of Oxford, UK, and aims to improve the quality and transparency of health research) provides guidelines and instructions for the reporting of various kinds of studies. These can be found at [www.equator-network.org](http://www.equator-network.org). Additional information regarding what to look for in a cohort study, as well as evaluation checklists can be found elsewhere [2, 8, 11, 25, 39, 48, 49]. The next section of this chapter will cover examples of famous prospective cohort studies from the medical field.

### 3. Examples of prospective cohort studies

#### 3.1. The Framingham Heart Study

##### 3.1.1. Overview

The Framingham heart study, initiated in 1948 by The National Heart Institute (currently the National Heart, Lung, and Blood Institute) [50], is considered to be the longest, ongoing, prospective cohort study in the history of the USA [51]. Others view it as a live model that illustrates the cohort design [52]. The study was based on the hypothesis that arteriosclerosis and hypertensive cardiovascular disease are the result of several causation factors combined rather than an individual factor [53]. Based on this, the aim of the study was to investigate

the factors that contribute to the development of cardiovascular disease (CVD) by following a large cohort of individuals over a long period of time [50]. Back then in 1951, when the first article about the study was published, little was known about arteriosclerosis and hypertensive cardiovascular disease [53].

The original cohort included 5209 participants, ages 30–62 years, that were recruited at the beginning of the study in the town of Framingham, Massachusetts, USA [50]. The same cohort has been followed since initiation every two years for physical, laboratory, and lifestyle examinations [50]. The second generation, the offspring cohort, was recruited in 1971 and included 5124 participants. While 1994 witnessed the enrollment of the first Omni cohort ( $n = 506$ ), in order to diversify the study population. More recently in 2002, the third generation cohort ( $n = 4095$ ) was enrolled, while in 2003 the new offspring cohort ( $n = 103$ ), and the second Omni group ( $n = 410$ ) was enrolled [50]. The study continues to follow these cohorts every 2–6 years [54]. This multi generation, multi ethnicity, enrollment design aided significantly in the study of genetics in relation to a wide range of factors and illnesses [51, 54].

Based on the Framingham study data, since initiation and through November 2017, a total of 3561 articles have been published so far [55]. The accumulation of knowledge that has risen from this study has shed the light on cardiovascular disease risk factors [50, 51, 56], by further expanding on our understanding of chronic illnesses such as diabetes, obesity, metabolic syndrome and nonalcoholic fatty liver disease [51, 57]. Such risk factors include high blood pressure, high cholesterol levels, smoking, obesity, diabetes, and physical inactivity [50, 57].

The study was the basis of which the Framingham risk score was built on [56]. Initially published by Wilson et al. in 1998 [58], it allows for the calculation of a 10 year risk estimate of developing coronary heart disease (CHD) based on the levels of different variables [56, 58]. This would allow for the undertaking of preventive measures [56]. Later on in 2002, the Adult Treatment Panel of the National Cholesterol Education Program used the risk score as a foundation for its risk calculator [56].

### *3.1.2. Study findings*

The study website (<https://www.framinghamheartstudy.org/about-fhs/research-milestones.php>) covers a long list of findings, among those; cigarette smoking was discovered to increase ones risk of developing heart disease back in 1960. In 1970, high blood pressure was discovered to increase ones risk of stroke. In 1988, the beneficial effects of HDL cholesterol were discovered. In 2002, the study found that obesity is considered a risk factor leading to heart failure. More recently in 2010 sleep apnea was linked to a higher risk of stroke [59]. More information and a full list of research milestones can be found elsewhere [59].

### *3.1.3. Strengths and weaknesses*

In addition to what had been previously discussed regarding the benefits of the prospective design of the study, a high retention rate is among the strengths of the Framingham Heart Study as participants continue to return for their follow up visits despite the years [54].

Among the weaknesses is that the study was conducted in one population residing in one locality [7], which in turn reflects on the ability to generalize findings to other populations [58]. Another weakness is that the study cohort was not randomly selected, as investigators had to use volunteers in order to obtain the necessary sample. The final cohort ended up being more healthy when compared to the general population [7, 60].

### 3.2. The Nurses' Health Study (NHS)

#### 3.2.1. Overview

This National Institutes of Health (NIH) funded study started in 1976 [61], and as of today includes more than 275,000 participants and counting, as the Nurses' Health Study 3 is still recruiting subjects [62]. The study looks into the risk factors that have been implicated in major chronic diseases among women [62]. Initially, the study focused on heart disease, cancer, smoking, and contraceptive methods [61]. As the study evolved, it investigated many other lifestyle factors, characteristics, and diseases [61, 63].

The original cohort of the study has been followed up on by mail every two years, with a minimum response rate of 90% [61]. The second cohort, under NHS 2, was enrolled in 1989 and included 116,430 women. These also were followed up on using mail every two years. A food frequency questionnaire was added in 1991 and was mailed out every four years, with a response rate of 85–90%. Later on blood and urine samples were collected from participants [61]. The third cohort, under NHS 3, was enrolled in 2010 and is still enrolling, with a goal of diversifying the study population to include other ethnic backgrounds [61].

#### 3.2.2. Study findings

The study website (<http://www.nurseshealthstudy.org/about-nhs/key-contributions-scientific-knowledge>) covers numerous study findings, such as reporting lower risk of colon cancer and polyps with higher levels of vitamin D [64]. Also among the findings, Giovannucci et al. reported lower risk rates of colon cancer with longer duration of aspirin usage [65]. Baer et al. reported on mortality related risk factors among the NHS cohort [66]. Other findings related to breast cancer, CHD, stroke, colon cancer, hip fracture, cognitive function, and eye disease, in relation to cigarette smoking, oral contraceptives, post-menopausal hormone therapy obesity, alcohol, and diet can be found elsewhere [64, 67–79]. More recently Colditz et al. summarized the findings and impact of the three NHS studies in an article published in the American Journal of Public Health [80].

#### 3.2.3. Strengths and weaknesses

With focus on women, it is considered to be the longest and largest running prospective cohort study that investigates the role of lifestyle on health [63]. Among the strengths of this study is that it included multiple assessments of the various lifestyle characteristics and exposure factors [63, 80], in turn, it also contributed to the methodology of lifestyle assessment in general,

which has been used in other studies [63, 80]. Additionally, it allowed for the calculation of mortality rates [63]. As for the weaknesses, white women dominated the original cohort, which reflects on the generalizability of the study results [4, 63].

### **3.3. The Caerphilly Prospective Study (CAPS)**

#### *3.3.1. Overview*

Also known as the Caerphilly Heart Disease Study, this study was conducted in Caerphilly, South Wales, UK, and focused on ischemic heart disease (IHD) in relation to hormones, hemostatic factors, and lipids [81]. As the study evolved, other investigations were included which looked into cognitive function, stroke and hearing problems [81].

The study included four phases. In the first phase, 2512 males, ages 45–59 years, were recruited in 1979. The procedures included blood tests, electrocardiogram (ECG), clinical history, lifestyle and IHD related questionnaires [81]. The second phase ran from 1984 to 1988 and included 447 males. An audiometry test was added to the list of investigations that were included in the first phase [81]. Phase 3 took place from 1989 to 1993 and added a cognitive function test and a bleeding time test [81]. Phase 4 was conducted from 1993 to 1997, which included the audiometry and cognitive function tests originally included in the second and third phases, respectively [81]. Follow up was conducted at a later stage through mail. The study has accumulated in a total of 150 studies and counting [81].

#### *3.3.2. Study findings*

Among the findings of the Caerphilly Prospective study; Elwood et al. showed that adopting a healthy lifestyle was associated with lower rates of chronic disease, as well as less cognitive impairment and dementia [82]. In other findings, Mertens et al. reported an inverse association between CVD and adopting a healthy diet [83], while Bolton et al. reported an inverse association between mid-life lung function and arterial stiffness among men [84]. Additional findings can be found elsewhere [85–91].

### **3.4. Conclusion**

The three sections of this chapter covered the two types of cohort studies. Observational studies in general and cohort studies in specific are a good source of information when an experiment is not feasible. Prospective cohort studies provide valuable information when studying the relationship between exposure and outcome. As with any type of study, prospective cohort studies come with advantages and disadvantages that need to be taken into consideration when interpreting the results of these studies.

### **Conflict of interest**

The author(s) declare no conflict of interest.

## Abbreviations

ANOVA	analysis of variance
CAPS	Caerphilly Prospective Study
CVD	cardiovascular disease
CHD	coronary heart disease
ECG	electrocardiogram
IHD	ischemic heart disease
RR	relative risk
MANOVA	multivariate analysis of variance
NHS	Nurses' Health Study
NIH	National Institutes of Health

## Author details

Samer Hammoudeh, Wessam Gadelhaq and Ibrahim Janahi\*

\*Address all correspondence to: [ijanahi@hamad.qa](mailto:ijanahi@hamad.qa)

Hamad Medical Corporation, Doha, Qatar

## References

- [1] Morabia A. A History of Epidemiologic Methods and Concepts. Birkhäuser: Switzerland, Basel; 2004
- [2] Grimes DA, Schulz KF. Cohort studies: Marching towards outcomes. *The Lancet*. 2002; **359**(9303):341-345
- [3] Doll R. Cohort studies: History of the method. I. Prospective cohort studies. *Sozial- und Präventivmedizin*. 2001; **46**(2):75-86
- [4] Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, c2008; 2008
- [5] Bryman A. *Social research methods*. 4th ed. Oxford; New York: Oxford University Press; 2012
- [6] Mann CJ. Observational research methods. Research design II: Cohort, cross sectional, and case-control studies. *Emergency Medicine Journal*. 2003; **20**(1):54



- [7] Caruana EJ, Roman M, Hernández-Sánchez J, Solli P. Longitudinal studies. *Journal of Thoracic Disease*. 2015;**7**(11):E537-EE40
- [8] Thiese MS. Observational and interventional study design types; an overview. *Biochemia Medica*. 2014;**24**(2):199-210
- [9] Miller AB, Goff DC, BammannDr K, Wild P. Cohort Studies. In: Ahrens W, Pigeot I, editors. *Handbook of Epidemiology*. New York, NY: Springer; 2014. pp. 259-291
- [10] Gamble J-M. An introduction to the fundamentals of cohort and case-control studies. *The Canadian Journal of Hospital Pharmacy*. 2014;**67**(5):366-372
- [11] Carlson MDA, Morrison RS. Study design, precision, and validity in observational studies. *Journal of Palliative Medicine*. 2009;**12**(1):77-82
- [12] De Rango P. Prospective cohort studies. *European Journal of Vascular and Endovascular Surgery*. 2016;**51**(1):151
- [13] Costantino G, Montano N, Casazza G. When should we change our clinical practice based on the results of a clinical study? The hierarchy of evidence. *Internal and Emergency Medicine*. 2015;**10**(6):745-747
- [14] Burns PB, Rohrich RJ, Chung KC. The levels of evidence and their role in evidence-based medicine. *Plastic and Reconstructive Surgery*. 2011;**128**(1):305-310
- [15] Petrisor BA, Bhandari M. The hierarchy of evidence: Levels and grades of recommendation. *Indian Journal of Orthopaedics*. 2007;**41**(1):11-15
- [16] Perry-Parrish C, Dodge R. Research and statistics: Validity hierarchy for study design and study type. *Pediatrics in Review*. 2010;**31**(1):27-29
- [17] Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. *Evidence-Based Medicine*. 2016
- [18] Euser AM, Zoccali C, Jager KJ, Dekker FW. Cohort studies: Prospective versus retrospective. *Nephron Clinical practice*. 2009;**113**(3):c214-c217
- [19] Noordzij M, Dekker FW, Zoccali C, Jager KJ. Study designs in clinical research. *Nephron Clinical Practice*. 2009;**113**(3):c218-cc21
- [20] Levin KA. Study design IV: Cohort studies. *Evidence-Based Dentistry*. 2006;**7**:51
- [21] Munnangi S, Boktor SW. *Epidemiology, Study Design*. Treasure Island (FL): StatPearls Publishing LLC; 2017
- [22] Setia MS. Methodology Series Module 1: Cohort Studies. *Indian Journal of Dermatology*. 2016;**61**(1):21-25
- [23] Merrill RM. *Introduction to Epidemiology*. 6th ed. Burlington, MA: Jones & Bartlett Learning; 2012
- [24] Song JW, Chung KC. Observational studies: cohort and case-control studies. *Plastic and Reconstructive Surgery*. 2010;**126**(6):2234-2242

- [25] Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet*. 2002;**359**(9302):248-252
- [26] Jekel JF. *Epidemiology, Biostatistics, and preventive medicine*. 3rd ed. Philadelphia: Saunders/Elsevier, c2007; 2007
- [27] Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. *Designing Clinical Research*. Wolters Kluwer/Lippincott Williams & Wilkins; 2013
- [28] Sedgwick P. Retrospective cohort studies: advantages and disadvantages. *BMJ: British Medical Journal*. 2014;**348**
- [29] Kip KE, Cohen F, Cole SR, Wilhelmus KR, Patrick DL, Blair RC, et al. Recall bias in a prospective cohort study of acute time-varying exposures: example from the herpetic eye disease study. *Journal of Clinical Epidemiology*. 2001;**54**(5):482-487
- [30] White E, Hunt JR, Casso D. Exposure measurement in cohort studies: the challenges of prospective data collection. *Epidemiologic Reviews*. 1998;**20**(1):43-56
- [31] Hess DR. Retrospective studies and chart reviews. *Respiratory Care*. 2004;**49**(10):1171
- [32] Sedgwick P. Prospective cohort studies: Advantages and disadvantages. *BMJ: British Medical Journal*. 2013;**347**
- [33] 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. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association* 2007;**22**(10):2909-2916
- [34] Friberg L, Benson L, Rosenqvist M, Lip GYH. Assessment of female sex as a risk factor in atrial fibrillation in Sweden: Nationwide retrospective cohort study. *BMJ: British Medical Journal*. 2012;**344**
- [35] Lindenauer PK, Rothberg MB, Pekow PS, Kenwood C, Benjamin EM, Auerbach AD. Outcomes of care by hospitalists, general internists, and family physicians. *The New England Journal of Medicine*. 2007;**357**(25):2589-2600
- [36] Goodwin KA, Goodwin CJ. *Research in Psychology: Methods and Design*. 8th ed. Hoboken, NJ: Wiley; 2017
- [37] Trochim W, Donnelly JP. *The Research Methods Knowledge Base*. 3rd ed. Boston, MA: Cengage Learning; 2006
- [38] Sedgwick P. Bias in observational study designs: Prospective cohort studies. *BMJ: British Medical Journal*. 2014;**349**
- [39] Hammer GP, du Prel J-B, Blettner M. Avoiding bias in observational studies: Part 8 in a series of articles on evaluation of scientific publications. *Deutsches Ärzteblatt International*. 2009;**106**(41):664-668

- [40] Boyko EJ. Observational research--Opportunities and limitations. *Journal of Diabetes and its Complications*. 2013;**27**(6):642-648
- [41] Sessler DI, Imrey PB. Clinical research methodology 2: Observational clinical research. *Anesthesia & Analgesia*. 2015;**121**(4):1043-1051
- [42] Dettori JR. Loss to follow-up. *Evidence-Based Spine-Care Journal*. 2011;**2**(1):7-10
- [43] Schoenbach V, Rosamond W. *Understanding the Fundamentals of Epidemiology: An Evolving Text*. University of North Carolina at Chapel Hill; 2000
- [44] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Preventive Medicine*. 2007;**45**(4):247-251
- [45] Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *PLoS Medicine*. 2007;**4**(10):e297
- [46] Vandenbroucke JP, Elm E, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (strobe): Explanation and elaboration. *Annals of Internal Medicine*. 2007;**147**(8):W-163-WW-94
- [47] Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *Epidemiology* 2007;**18**(6):805-835
- [48] Bookwala A, Hussain N, Bhandari M. The three-minute appraisal of a prospective cohort study. *Indian Journal of Orthopaedics*. 2011;**45**(4):291-293
- [49] Tooth L, Ware R, Bain C, Purdie DM, Dobson A. Quality of reporting of observational longitudinal research. *American Journal of Epidemiology*. 2005;**161**(3):280-288
- [50] FHS. History of the Framingham Heart Study 2017. Available from: <https://www.framinghamheartstudy.org/about-fhs/history.php>
- [51] Long MT, Fox CS. The Framingham Heart Study--67 years of discovery in metabolic disease. *Nature Reviews Endocrinology*. 2016;**12**(3):177-183
- [52] Oppenheimer GM. Framingham Heart Study: the first 20 years. *Progress in Cardiovascular Diseases*. 2010;**53**(1):55-61
- [53] Dawber TR, Meadors GF, Moore FE. Epidemiological approaches to heart disease: The Framingham Study. *American Journal of Public Health and the Nations Health*. 1951;**41**(3):279-286
- [54] Tsao CW, Vasan RS. Cohort Profile: The Framingham Heart Study (FHS): overview of milestones in cardiovascular epidemiology. *International Journal of Epidemiology*. 2015;**44**(6):1800-1813

- [55] FHS. Framingham Heart Study Bibliography 2017. Available from: <https://www.framinghamheartstudy.org/fhs-bibliography/index.php>
- [56] Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular diseases: A historical perspective. *Lancet*. 2014; **383**(9921):999-1008
- [57] Hajar R. Framingham contribution to cardiovascular disease. *Heart Views : The Official Journal of the Gulf Heart Association*. 2016;**17**(2):78-81
- [58] Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;**97**(18):1837-1847
- [59] FHS. Research Milestones 2017. Available from: <https://www.framinghamheartstudy.org/about-fhs/research-milestones.php>
- [60] FHS. Epidemiological Background and Design: The Framingham Study 2017. Available from: <https://www.framinghamheartstudy.org/about-fhs/background.php>
- [61] NHS. History 2016. Available from: <http://www.nurseshealthstudy.org/about-nhs/history>
- [62] NHS. About NHS 2016. Available from: <http://www.nurseshealthstudy.org/about-nhs>
- [63] Colditz GA, Manson JE, Hankinson SE. The Nurses' Health Study: 20-year contribution to the understanding of health among women. *Journal of Women's Health*. 1997;**6**(1):49-62
- [64] NHS. Key contributions to scientific knowledge 2016. Available from: <http://www.nurseshealthstudy.org/about-nhs/key-contributions-scientific-knowledge>
- [65] Giovannucci E, Egan KM, Hunter DJ, Stampfer MJ, Colditz GA, Willett WC, et al. Aspirin and the risk of colorectal cancer in women. *The New England Journal of Medicine*. 1995;**333**(10):609-614
- [66] Baer HJ, Glynn RJ, Hu FB, Hankinson SE, Willett WC, Colditz GA, et al. Risk factors for mortality in the nurses' health study: A competing risks analysis. *American Journal of Epidemiology*. 2011;**173**(3):319-329
- [67] Bernstein AM, Pan A, Rexrode KM, Stampfer M, Hu FB, Mozaffarian D, et al. Dietary protein sources and the risk of stroke in men and women. *Stroke*. 2012;**43**(3):637-644
- [68] Chiuve SE, Fung TT, Rexrode KM, Spiegelman D, Manson JE, Stampfer MJ, et al. Adherence to a low-risk, healthy lifestyle and risk of sudden cardiac death among women. *JAMA*. 2011;**306**(1):62-69
- [69] Liu S, Willett WC, Stampfer MJ, Hu FB, Franz M, Sampson L, et al. A prospective study of dietary glycemic load, carbohydrate intake, and risk of coronary heart disease in US women. *The American Journal of Clinical Nutrition*. 2000;**71**(6):1455-1461
- [70] Hu FB, Manson JE, Stampfer MJ, Colditz G, Liu S, Solomon CG, et al. Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. *The New England Journal of Medicine*. 2001;**345**(11):790-797

- [71] Hu FB, Bronner L, Willett WC, Stampfer MJ, Rexrode KM, Albert CM, et al. Fish and omega-3 fatty acid intake and risk of coronary heart disease in women. *JAMA*. 2002; **287**(14):1815-1821
- [72] Ayas NT, White DP, Manson JE, Stampfer MJ, Speizer FE, Malhotra A, et al. A prospective study of sleep duration and coronary heart disease in women. *Archives of Internal Medicine*. 2003; **163**(2):205-209
- [73] Curhan GC, Willett WC, Knight EL, Stampfer MJ. Dietary factors and the risk of incident kidney stones in younger women: Nurses' Health Study II. *Archives of Internal Medicine*. 2004; **164**(8):885-891
- [74] Holmes MD, Chen WY, Feskanich D, Kroenke CH, Colditz GA. Physical activity and survival after breast cancer diagnosis. *JAMA*. 2005; **293**(20):2479-2486
- [75] Eliassen AH, Missmer SA, Tworoger SS, Spiegelman D, Barbieri RL, Dowsett M, et al. Endogenous steroid hormone concentrations and risk of breast cancer among premenopausal women. *Journal of the National Cancer Institute*. 2006; **98**(19):1406-1415
- [76] Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, et al. A genome-wide association study identifies alleles in *FGFR2* associated with risk of sporadic postmenopausal breast cancer. *Nature Genetics*. 2007; **39**(7):870-874
- [77] Song H, Ramus SJ, Tyrer J, Bolton KL, Gentry-Maharaj A, Wozniak E, et al. A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. *Nature Genetics*. 2009; **41**(9):996-1000
- [78] Beral V, Doll R, Hermon C, Peto R, Reeves G. Ovarian cancer and oral contraceptives: collaborative reanalysis of data from 45 epidemiological studies including 23, 257 women with ovarian cancer and 87, 303 controls. *Lancet*. 2008; **371**(9609):303-314
- [79] Sarwar N, Gao P, Seshasai SR, Gobin R, Kaptoge S, Di Angelantonio E, et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: A collaborative meta-analysis of 102 prospective studies. *Lancet*. 2010; **375**(9733):2215-2222
- [80] Colditz GA, Philpott SE, Hankinson SE. The impact of the Nurses' Health Study on population health: Prevention, translation, and control. *American Journal of Public Health*. 2016; **106**(9):1540-1545
- [81] Bristol Uo. Caerphilly Prospective Study: About 2017. Available from: <https://www.bristol.ac.uk/population-health-sciences/projects/caerphilly/about/>
- [82] Elwood P, Galante J, Pickering J, Palmer S, Bayer A, Ben-Shlomo Y, et al. Healthy lifestyles reduce the incidence of chronic diseases and dementia: Evidence from the Caerphilly Cohort Study. *PLoS One*. 2013; **8**(12):e81877
- [83] Mertens E, Markey O, Geleijnse JM, Lovegrove JA, Givens DI. Adherence to a healthy diet in relation to cardiovascular incidence and risk markers: evidence from the Caerphilly Prospective Study. *European Journal of Nutrition*. 2017



- [84] Bolton CE, Cockcroft JR, Sabit R, Munnery M, McEniery CM, Wilkinson IB, et al. Lung function in mid-life compared with later life is a stronger predictor of arterial stiffness in men: the Caerphilly Prospective Study. *International Journal of Epidemiology*. 2009;**38**(3):867-876
- [85] Patterson CC, Smith AE, Yarnell JW, Rumley A, Ben-Shlomo Y, Lowe GD. The associations of interleukin-6 (IL-6) and downstream inflammatory markers with risk of cardiovascular disease: The Caerphilly Study. *Atherosclerosis*. 2010;**209**(2):551-557
- [86] Gallacher J, Bayer A, Lowe G, Fish M, Pickering J, Pedro S, et al. Is sticky blood bad for the brain?: Hemostatic and inflammatory systems and dementia in the Caerphilly Prospective Study. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 2010;**30**(3):599-604
- [87] Sarkar C, Gallacher J, Webster C. Built environment configuration and change in body mass index: The Caerphilly Prospective Study (CaPS). *Health & Place*. 2013;**19**:33-44
- [88] McEniery CM, Spratt M, Munnery M, Yarnell J, Lowe GD, Rumley A, et al. An analysis of prospective risk factors for aortic stiffness in men: 20-year follow-up from the Caerphilly prospective study. *Hypertension (Dallas, Tex: 1979)*. 2010;**56**(1):36-43
- [89] Livingstone KM, Lovegrove JA, Cockcroft JR, Elwood PC, Pickering JE, Givens DI. Does dairy food intake predict arterial stiffness and blood pressure in men?: Evidence from the Caerphilly Prospective Study. *Hypertension (Dallas, Tex: 1979)*. 2013;**61**(1):42-47
- [90] Eicher JD, Xue L, Ben-Shlomo Y, Beswick AD, Johnson AD. Replication and hematological characterization of human platelet reactivity genetic associations in men from the Caerphilly Prospective Study (CaPS). *Journal of Thrombosis and Thrombolysis*. 2016;**41**(2):343-350
- [91] Gallacher JE, Yarnell JW, Sweetnam PM, Elwood PC, Stansfeld SA. Anger and incident heart disease in the caerphilly study. *Psychosomatic Medicine*. 1999;**61**(4):446-453