

# Retrospective studies – utility and caveats

Keerthi Talari<sup>1</sup>, Mohit Goyal<sup>2</sup>



A thorough understanding of the pros and cons of the various study designs is critical to correct interpretation of their results. Retrospective studies are an important tool to study rare diseases, manifestations and outcomes. Findings of these studies can form the basis on which prospective studies are planned. Retrospective studies however have several limitations owing to their design. Since they depend on review of charts that were originally

not designed to collect data for research, some information is bound to be missing. Selection and recall biases also affect the results and reasons for differences in treatment between patients and lost follow ups can often not be ascertained and may lead to bias. Readers need to critically evaluate the methods and carefully interpret the results of retrospective studies before they put them to practice. Researchers should avoid over generalisation of results and be cautious in claiming cause-effect relationship in retrospective studies.

**Keywords:** Retrospective studies, chart review, limitations, bias, cause-effect relationship, interpretation

**Financial and Competing Interests:** No conflict of interests declared

**Correspondence to:**

Dr Mohit Goyal  
CARE Pain & Arthritis  
Centre  
Udaipur 313002  
India

**Email:**

dr.mohitgoyal@gmail.com

## Introduction

Evidence based practice is the basis of modern medicine. Accurate interpretation of any evidence demands a thorough understanding of research studies and their methodology. Appreciating the study design employed to answer a particular research question is critical as each design has its advantages and disadvantages.

Broadly, research designs are either observational or interventional. In observational studies, the researcher documents a naturally occurring relationship between the exposure and the outcome with no active intervention. Observational studies are either descriptive or analytical.<sup>1</sup> Descriptive studies merely describe data of a group of individuals and cannot establish relationships, while analytical studies attempt to establish an association or causal relationship between variables.<sup>2</sup> In interventional studies, investigators perform an intervention on a group of individuals to study particular outcomes.<sup>3</sup> Based on the directionality of the data inquiry, observational studies could be either prospective or retrospective.<sup>1</sup> In a prospective study, the outcome of interest has not occurred at the time of initiation of the study, and participants are followed up over a period to the outcome being studied. On the other hand, in retrospective studies, the outcome is already available and the investigator delves back into time to get data either from a participant or their medical records.<sup>4</sup> The important feature of note about retrospective studies is that the data

are never collected for research purposes and are only a part of the clinical database.<sup>5</sup> In this perspective, we would like to discuss retrospective studies, their limitations and the caution needed when interpreting their results.

## Retrospective studies – types and advantages

As discussed earlier, in retrospective studies, the outcome of interest has already occurred. Information on the variables being studied is usually obtained from medical records or depends on the participants' recall. Retrospective studies could either be descriptive or analytical. Descriptive retrospective studies are case series and cross sectional studies, while analytical retrospective studies are cross sectional, case control and cohort studies. A case series is a description of multiple, similar instructive cases; it can be used to study diseases that are rare and unusual in the population. Case descriptions are important as they can potentially help generate hypotheses which can be put to test through other study designs.<sup>6</sup> In a cross sectional study, the investigator makes all the measurements (both risk factor and the outcome of interest) in the same time frame. It can be either descriptive or may aim to establish a relationship between the risk factor and the outcome. While in a case control study, cases and controls, with and without the outcome of interest are identified and their exposure to a particular variable (risk factor) in the past is collected and analysed to establish a relation. In cohort studies, a group

<sup>1</sup>Consultant Rheumatologist, Yashoda Hospitals, Hyderabad, India; <sup>2</sup>Consultant Rheumatologist, CARE Pain & Arthritis Centre, Udaipur, India

of individuals with exposure to a risk factor are prospectively followed to determine the occurrence of the outcome.<sup>7</sup> While these are mostly prospective studies, a cohort study can be retrospective too. This is primarily done to reduce cost and duration of follow up. An association between the variable and the outcome may be derived from cohort studies.

Retrospective studies have a place in research and many of them have helped shape the clinical practices. An example of the utility of retrospective studies is the landmark paper that described the association between smoking and lung cancer.<sup>8</sup> The study revealed that smokers were at a significantly higher risk of developing carcinoma of the lung compared to non-smokers. Such a hypothesis could have never been put through the test of a randomised trial.

Another landmark retrospective chart review in the 1990s found that spinal anaesthesia was faster, easier to administer and more comfortable and safe for the patient for caesarean section, as compared to epidural anaesthesia.<sup>9</sup> Until then, epidural was the preferred mode of anaesthesia administration in caesarean section, but we have since seen a paradigm shift towards spinal anaesthesia; by 2009, 85% of obstetricians in the United States were using spinal anaesthesia for caesarean section and another 11% were combining it with epidural.<sup>10</sup> Some scenarios where retrospective studies can be of use are given in Box 1.

### Limitations of retrospective studies

While retrospective studies save on funds and time and are useful in studying rare diseases and rare outcomes, they are marred by their fallacies. Retrospective studies depend on data that were entered into a clinical database and not collected for research. Since the data was not collected in a predesigned proforma as per the specific requirements of the study, in most of the cases some data would inevitably be missing. Also, certain variables that have the potential to impact the outcome may not have been recorded at all.<sup>11</sup>

The United States Renal Data System (USRDS) is a robust database with details of patients with end stage renal disease (ERDS) across the country. They make the data available freely to the researchers for analysis.<sup>12</sup> Retrospective analyses of the same data by different researchers have shown conflicting conclusions regarding reuse of dialysers and patient mortality. While some studies found increased mortality related to reuse, another found that there were other confounders that could have influenced the mortality in the reuse group.<sup>13-15</sup> Also, different analyses of the USRDS data by various researchers have shown similar, better and even worse outcomes with haemodialysis when compared to peritoneal dialysis in patients with ERDS.<sup>16,17,18</sup> Older charts are likely to have missing information and unavailability of information on confounders leads to bias. Sometimes where the disease is very severe at presentation, other supposedly minor abnormalities may not be recorded and remain unaccounted for during retrospective analysis. This is an example of how unknown and unnoticed biases

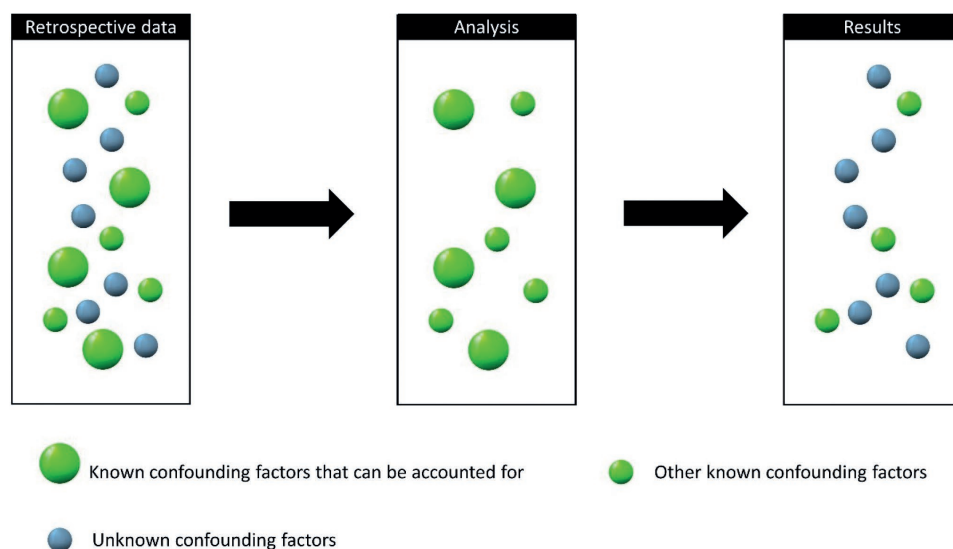
#### Box 1 Conditions where retrospective studies are useful

- Rare diseases where the population needed to study is too large to identify the few who develop the outcome of interest
- Rare exposures where the number of people exposed may be too low and hence may take too long for enough numbers to develop the outcome of interest
- Where the duration between the exposure and the outcome is too long, thus reducing the feasibility of a prospective study
- Where there are funding constraints towards planning a prospective study

creep into retrospective studies and all of these cannot be accounted for.

Many times, the investigator fills in missing data by looking at records at different time points (previous or next visit) which is also fallacious. It is also common practice to ask the patients to recall certain details to collect data for retrospective studies. This introduces a systematic error called 'recall bias'.<sup>19</sup> Patients may not be able to recall or describe the details accurately, resultantly certain details may be omitted or altered. The accuracy and the volume of memory of a certain event are usually dependent on the impact of that event.<sup>20</sup> Thus, there is always the chance that an event of lesser magnitude and the finer details will not be accurately recalled. Human memory is imperfect and study results based on them cannot be relied upon. In many manuscripts, while the authors may not explicitly state that certain information was retrieved through recall, it may be evident to the reader on careful reading. A note should, however, be made here that it is incumbent on the authors' part to declare that certain information was retrieved through recall by participants and it should also be listed as a limitation of the study.

In most of the retrospective studies, it is assumed that except for the variable under study which differs between the cases and the controls, they are otherwise similar in all other respects. Researchers would argue that this is what confounding is all about, and adequate adjustment (matching and regression analysis) for such confounding factors was made. However, besides the recognised ones, there are always some unknown confounding factors that remain unrecognised in retrospective studies.<sup>21</sup> It is often difficult to identify appropriate study and control groups in retrospective studies. Since in these studies, researchers have no control over the exposure of cases versus controls, these unrecognised confounders may influence the results (Figure 1). This way a false association may be derived between the variable of interest and the outcome even when no true association exists. On the other hand, in a prospective study, certain unknown risk factors are identified and new variables that can influence the outcome may also be recognised. In retrospective studies, it may not always be possible to ascertain the reason for the lost follow ups.



**Figure 1** All confounders cannot be accounted for in retrospective studies

Non responders and those developing adverse effects or complications have a higher chance of being lost to follow up, leading to bias. The various sources of error or bias in retrospective studies and measures to minimise them are listed in Table 1.

For example, an imaginary study looked at the records of all 50,000 patients with migraine that attended the outpatient services of an institution in the last five years. All of these patients had undergone history taking and

clinical examination. For a mean follow up duration of six months, it was found that women had a significantly higher nonsteroidal anti-inflammatory drug (NSAID) requirement per month (five pills of NSAID at optimum dose in women versus three in men,  $p < 0.01$ ). When other factors were adjusted for, women still had a higher NSAID requirement. It was concluded that women with migraine had a lower threshold for NSAID use. Since all the patients were included and the data were retrieved from records, selection and recall bias were eliminated. But since at the time of their creation, the

**Table 1** Confounders and sources of error or bias in retrospective studies and how to minimise or account for them


Source of bias	How it creeps in	How to minimise
Baseline characteristics	Differences in baseline characteristics of the groups, that have the potential to impact the outcome	Choose appropriate comparator group
Selection of subjects	The study subjects may not be representative of the population and reasons for non-selection may not be ascertainable	Stringent and validated (where applicable) selection criteria
Chart selections	Data was not collected for research, resultantly some charts are excluded due to missing of certain crucial information	Document, mention in manuscript and list the same as a limitation
Missing data in charts	Since the data was not collected for research, even the included charts are bound to have some missing information	Document, mention in manuscript and list the same as a limitation
Reliance on recall	Accuracy of missing data added by asking patients to recall may be limited by inability to describe or inaccurate memory	If recall is a major part, subjects must have comparable educational status
Assumptions	Investigators may try to complete missing data by assuming and approximating from data available at different time points	Authors must avoid such practices
Lack of homogeneity	Different people are involved at different times in patient care and data entry, especially when studies look at charts over many years	Plan studies such that these errors are minimised
Prescription bias	Prescriptions may have varied according to patients' risk profiles and the exact reasons may not have been recorded	Maximum details must be retrieved from records and accounted for
Loss of follow up	Reasons for lost follow-ups often cannot be ascertained in retrospective studies and can potentially bias the results	Manuscript should mention these exclusions as limitations
Generalisation of results by authors	Due to selection bias, results of retrospective studies are often not generalisable to the whole population	List the limitations and do not over-generalise
Claiming cause-effect	Retrospective studies generally only establish association and not cause-effect between risk factor and outcome	Such claims must be avoided

records were not aimed at studying the association between gender and NSAID requirement, other variables that can affect the use of NSAIDs may not have been recorded. NSAID requirement and prophylactic treatment at baseline, frequency of episodes of headache and severity are factors that can confound the results of such a retrospective study. Migraine has been found to be more prevalent in women, and it is possible that they have more frequent and severe episodes. Thus, it may not be a prudent conclusion to make.

In prospective studies, as we move forward, multiple outcomes can be assessed and analysed at different time frames.<sup>21</sup> The results obtained by a retrospective study are limited; multiple outcomes cannot be studied at a time. This is because we identify the outcome first and the risk variable is looked at in the past. Concerning the recent retrospective analysis of the use of hydroxychloroquine for COVID-19 which was later retracted (for different reasons), certain findings were reported which demanded extreme caution on the part of the reader.<sup>22</sup> The registry in this study comprised data of hospitalised patients from different countries, which might have had different guidelines for hospitalisation and varying protocols for the administration of antimalarials and macrolides in patients with COVID-19. For example, the national guidelines in India at different times have advocated the use of these agents only in patients with severe disease and requiring intensive care, or for prophylaxis in health care workers.<sup>23</sup> There were significant differences in comorbidities between the survivors and non-survivors. Some of these comorbidities had been previously reported to be associated with a severe disease course in COVID-19 patients.<sup>24</sup> Despite

having a higher burden of cardiac comorbidities, significantly fewer non survivors were on ACE inhibitors and statins. Also, it should have been acknowledged that cardiac complications unrelated to medications had already been documented in patients with COVID-19, and were more likely to occur in those with severe disease.<sup>25,26</sup> Then, there are hitherto unexplained factors which have led to widely varied mortality from COVID-19 in different countries. Authors used exploratory multivariate analysis, which has its limitations and may not be the best tool to establish a cause-effect relationship. While retrospective studies with sound methods and data collection may provide an association between the variable and the outcome, they are generally not suited to determine causation.<sup>27</sup>

## Conclusion

While they are valuable tools to study diseases, exposures and outcomes that are rare, retrospective studies are rife with inherent limitations. The reader must understand and account for such limitations while analysing the results and before applying them in the clinic. Retrospective studies are the right first step to formulate a hypothesis that may otherwise need exorbitant funding on a prospective design. When a disease is common enough, the results of a retrospective study need to be confirmed in a prospective study, so that unknown factors that could have influenced the study results are identified and accounted for. A true causal relationship can only be established by properly conducted prospective studies where there is an option of taking care of biases of different kinds. 

## References

- Ranganathan P, Aggarwal R. Study designs: Part 1 – An overview and classification. *Perspect Clin Res*. 2018; 9: 184–6.
- Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. *Lancet*. 2002; 359: 145–9.
- Thiese MS. Observational and interventional study design types; an overview. *Biochem Med (Zagreb)*. 2014; 24: 199–210.
- Manja V, Lakshminrusimha S. Epidemiology and Clinical Research Design, Part 1: Study Types. *Neoreviews*. 2014; 15: e558–e69.
- Goje SK. Longitudinal vs retrospective studies. *Am J Orthod Dentofacial Orthop*. 2017; 151: 10–1.
- Ortega-Loubon C, Culquichicon C, Correa R. The Importance of Writing and Publishing Case Reports During Medical Training. *Cureus*. 2017; 9: e1964.
- Mann CJ. Observational research methods. Research design II: cohort, cross sectional, and case-control studies. *Emerg Med J*. 2003; 20: 54–60.
- Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. *Br Med J*. 1950; 2: 739–48.
- Riley ET, Cohen SE, Macario A, et al. Spinal versus epidural anesthesia for cesarean section: a comparison of time efficiency, costs, charges, and complications. *Anesth Analg*. 1995; 80: 709–12.
- Aiono-Le Tagalao L, Butwick AJ, Carvalho B. A survey of perioperative and postoperative anesthetic practices for cesarean delivery. *Anesthesiol Res Pract*. 2009; 2009: 510642.
- Altman DG, Bland JM. Missing data. *BMJ*. 2007; 334: 424.
- Ward RA, Brier ME. Retrospective analyses of large medical databases: what do they tell us? *J Am Soc Nephrol*. 1999; 10: 429–32.
- Held PJ, Wolfe RA, Gaylin DS, et al. Analysis of the association of dialyzer reuse practices and patient outcomes. *Am J Kidney Dis*. 1994; 23: 692–708.
- Feldman HI, Kinosian M, Bilker WB, et al. Effect of dialyzer reuse on survival of patients treated with hemodialysis. *JAMA*. 1996; 276: 620–5.
- Collins AJ, Ma JZ, Constantini EG, et al. Dialysis unit and patient characteristics associated with reuse practices and mortality: 1989–1993. *J Am Soc Nephrol*. 1998; 9: 2108–17.
- Vonesh EF, Moran J. Mortality in end-stage renal disease: a reassessment of differences between patients treated with hemodialysis and peritoneal dialysis. *J Am Soc Nephrol*. 1999; 10: 354–65.
- Fenton SS, Schaubel DE, Desmeules M, et al. Hemodialysis versus peritoneal dialysis: a comparison of adjusted mortality rates. *Am J Kidney Dis*. 1997; 30: 334–42.
- Bloembergen WE, Port FK, Mauger EA, et al. A comparison of mortality between patients treated with hemodialysis and peritoneal dialysis. *J Am Soc Nephrol*. 1995; 6: 177–83.
- Coughlin SS. Recall bias in epidemiologic studies. *J Clin Epidemiol*. 1990; 43: 87–91.
- Smith MC, Bibi U, Sheard DE. Evidence for the differential impact of time and emotion on personal and event memories for September 11, 2001. *Applied Cognitive Psychology*. 2004; 17: 1047–55.

- 21 Euser AM, Zoccali C, Jager KJ, et al. Cohort studies: prospective versus retrospective. *Nephron Clin Pract.* 2009; 113: c214–7.
- 22 Mehra MR, Desai SS, Ruschitzka F, et al. RETRACTED: Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet.* 2020.
- 23 Government of India. Revised Guidelines on Clinical Management of COVID – 19. Available from: <https://www.mohfw.gov.in/pdf/RevisedNationalClinicalManagementGuidelineforCOVID1931032020.pdf>
- 24 Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med.* 2020; 382: 1708–20.
- 25 Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* 2020; 395: 497–506.
- 26 Shi S, Qin M, Shen B, et al. Association of Cardiac Injury With Mortality in Hospitalized Patients With COVID-19 in Wuhan, China. *JAMA Cardiol.* 2020; 5: 802–10.
- 27 Toftagen C. Threats to validity in retrospective studies. *J Adv Pract Oncol.* 2012; 3: 181–3.