Check for updates

RESEARCH ARTICLE

REVISED Prospective versus retrospective: a simple matter of timing?

[version 2; peer review: 1 approved, 1 approved with reservations]

Mohammad Kamran Ikram[1,2], Jeremy A Labrecque[1], Mohammad Arfan Ikram[1]

[1]Department of Epidemiology, Erasmus University Medical Centre, Rotterdam, The Netherlands
[2]Department of Neurology, Erasmus University Medical Centre, Rotterdam, The Netherlands

v2 **First published:** 18 Oct 2024, **4**:225
https://doi.org/10.12688/openreseurope.18640.1
Latest published: 10 Jul 2025, **4**:225
https://doi.org/10.12688/openreseurope.18640.2

Abstract

Background

The terms “prospective” and “retrospective” are commonly used to describe the timing of a study. However, there is much confusion surrounding the use of these terms. In this paper we describe two common definitions used for these terms and explore whether the application of these different definitions may partly explain the confusion surrounding the use of the terms “prospective” and “retrospective”.

Methods

The most used definitions explored in this paper are a) time of recording of exposure status with respect to the occurrence of an outcome, and b) timing of the accumulation of person-time with respect to the study’s conduct. Subsequently, we apply these definitions to various real-world scenarios.

Results

Using these definitions, we show that there are specific scenarios in which these definitions may lead to concordant conclusions. However, more worryingly, we also present scenarios in which the two definitions lead to discordant conclusions, thereby partly explaining the confusion caused by these terms. Finally, we propose to use the terms “prospective” and “retrospective” to describe an aspect of a study not covered by other epidemiological concepts (e.g. bias, study design), namely the chronological relationship between the implementation of the eligibility criteria (which may include obtaining

Open Peer Review

Approval Status ✓ ?

	1	2
version 2 (revision) 10 Jul 2025	✓ view	? view
version 1 18 Oct 2024	? view	

1. **Sophie Bots**, Utrecht University, Utrecht, The Netherlands

2. **Annie Y Ng**, Kheiron Medical Technologies, London, UK

Any reports and responses or comments on the article can be found at the end of the article.

informed consent) and the accumulation of person-time.

Conclusion

In this paper, we show that different definitions used may partly explain the confusion surrounding the term prospective/retrospective. A consistent application of a clear definition will in the future reduce this confusion in the medical literature.

Plain language summary
In this paper, we examine the terms “prospective” and “retrospective”, commonly used to describe the timing of a study. However, there is confusion surrounding the use of these terms. Here, we explore two common definitions: a) time of recording of exposure status with respect to the occurrence of an outcome, and b) timing of the accumulation of person-time with respect to the study’s conduct. Using these definitions, we show that in specific scenarios these definitions may lead to different conclusions, thereby partly explaining the confusion caused by these terms. We conclude this paper by suggesting the use of the terms "prospective" and "retrospective" to specifically describe an element of a study not addressed by other epidemiological concepts (such as bias or study design). This element pertains to the temporal relationship between the application of eligibility criteria (including obtaining informed consent) and the subsequent accumulation of person-time. By adhering to a consistent definition of these terms, we aim to minimize the existing confusion within the medical literature.

Keywords
Study Design, Prospective, Retrospective, Longitudinal Studies, Eligibility Criteria



This article is included in the European Research Council (ERC) gateway.

Corresponding author: Mohammad Kamran Ikram (m.ikram@erasmusmc.nl)

Author roles: Ikram MK: Conceptualization, Methodology, Writing – Original Draft Preparation; **Labrecque JA**: Conceptualization, Methodology, Writing – Review & Editing; **Ikram MA**: Conceptualization, Methodology, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (project: ORACLE, Ggrant agreement No: 678543) (.project: ORACLE)
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2025 Ikram MK *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Ikram MK, Labrecque JA and Ikram MA. **Prospective versus retrospective: a simple matter of timing? [version 2; peer review: 1 approved, 1 approved with reservations]** Open Research Europe 2025, **4**:225 https://doi.org/10.12688/openreseurope.18640.2

First published: 18 Oct 2024, **4**:225 https://doi.org/10.12688/openreseurope.18640.1

REVISED **Amendments from Version 1**

The additional references, as suggested by the reviewer, were added to the 'Methods' section. In our original paper we re-iterated the advice given by the STROBE guidelines either to make researchers aware of this advice or remove from the guidelines. However, the paper by Dekkers et al. advices to discourage the use of the terms 'prospective' and 'retrospective'. This advice has now been added to our manuscript. Furthermore, the paper by Schneeweiss et al. provides an alternative to use graphical depictions to clarify longitudinal study designs. We have also added this advice to our manuscript. However, despite these recent additional advices the confusion surrounding these terms still remain prevalent in the literature. Hence, the addition of these references has further strengthened the rationale for our current paper.

Finally, throughout the manuscript we have removed some language mistakes to further improve the readability of our paper. These edits were meant to correct previous errors/typos. Content-wise there were no changes to the text.

Any further responses from the reviewers can be found at the end of the article

Introduction

In any research, after formulating a clear research question based on an underlying hypothesis, a crucial part is the design of the study. The cornerstone of observational studies trying to answer etiological questions are the cohort and case-control designs. When describing the design of a study, besides the terms "cohort" and "case-control", another set of closely related terms is used: namely prospective versus retrospective. Although initially (erroneously) used as synonyms for the terms "cohort" and "case-control", respectively, these terms are increasingly used to reflect aspects of the timing of data collection within a study. Though these terms are meant to clarify a design aspect of a study, in the current medical literature these terms lead to much confusion regarding the design and more worryingly tend to be implicitly used to judge the quality of the findings from such a study. In particular, the term "retrospective" has obtained a negative connotation in terms of the quality of a study. It is also important to realize that these terms were originally developed in the mid-twentieth century when initial observational studies dealt with the association of one clearly defined risk factor and one outcome. In recent decades the nature of data collection and analysis has changed considerably: e.g. availability of administrative data (data originally collected for other purposes such as reimbursement claims, medical files) for research purposes and more recently, the advent of large biobanks that collect data without clearly-specified aims. Another aspect is that data processing nowadays typically does not take place at one time point, but the interval between collection to processing may be spread over a longer time period. Furthermore, the distinction between which data is being collected for study purposes, and which data is being collected for routine clinical care is becoming blurred. These developments have further added to the confusion regarding the use of these terms. In this paper, we review the various definitions commonly used, discuss their merits and caveats with practical examples and propose a revised definition for the use of the terms "prospective" and "retrospective". Being aware of the differences between these definitions in the medical literature will enable both the epidemiologist and the clinician to properly appreciate this aspect of the design of a study.

Methods

Current practice

The terms "prospective" and "retrospective" have become firmly rooted in the health scientist's vocabulary: in 2016 Caram-Deelder *et al.*[1] reported that in 1975, only 1% of all papers and 15% of the epidemiological papers used these terms, whereas in this century these percentages increased to 6% and 31%, respectively. However, despite this increased usage of these terms, instead of clarifying a study's design, they have caused much confusion in the medical literature. Therefore, in 2007, the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines addressed this issue and recommended either to refrain from using these terms or to clearly define what is meant by them[2–4]. However, despite these recommendations, these terms are still frequently used without clearly describing what they mean. In particular, Caram-Deelder *et al.*[1] showed that in 150 observational studies in leading journals over half of the papers (86/150) still used these terms with only four papers clearly defining what they meant by these terms. The conclusion by the authors was either there should be an effort to make people aware of this advice or remove this from the STROBE guidelines. Others have discouraged the use of these terms[5]. Despite these efforts, the situation in the medical literature seems unaltered. As an alternative the use of graphical depictions have been recommended to clarify longitudinal study design[6]. Overall, it seems nearly impossible to get rid of these terms. At the same time, the confusion surrounding the terms "prospective" and "retrospective" is being aggravated by the ever-changing nature of data collection in observational studies. Given this confusion, it is important to note that whereas in the training of an epidemiologist/health scientist, ample time is spent on the cohort and case-control designs, the terms prospective/retrospective usually only receive a passing mention.

Common definitions

There are different definitions used to describe the terms "prospective" and "retrospective", as was already highlighted by Rothman *et al.*[7] in their seminal book "Modern Epidemiology" on p.95–96 and re-iterated in their latest edition:[8]

1. The first definition (Definition A) commonly used to distinguish between these terms is the timing and order in which the recording of exposure status and the occurrence of disease took place. An important aspect in this definition is the idea that the exposure measurement could not be influenced by the disease status. If this were true, then the study would be classified as prospective, and retrospective otherwise.

2. The second definition (Definition B) uses the timing of the accumulated person-time with respect to the study's conduct. In this case, when the person-time accumulates after the study begins, it is classified as

prospective. In contrast, when the person-time accumulated before the study is conducted, it is termed retrospective, even if the exposure status was recorded before the disease occurred.

3. Finally, some (early) papers used a third way to define the terms "prospective" and "retrospective", namely they are interchangeably used with the terms "cohort" and "case-control" studies, respectively. However, this definition does not convey any additional information beyond the terms used for the study design and hence from the outset we recommend to discard this definition.

4. Besides the terms "prospective" and "retrospective", a third term "ambispective" is used as a hybrid term for studies that incorporate both prospective and retrospective aspects.

Results

Classic scenarios

Although these definitions appear in various textbooks on epidemiology[7,8], the use of the terms "prospective" and "retrospective" is nearly always explained using some simple textbook examples. However, an extensive description and application of these terms to various complex situations is lacking. Here, we describe different scenarios in which we apply these two definitions and examine whether both definitions give the same results, or whether there are instances in which a study is labelled as "prospective" according to one definition and "retrospective" using another. Here it is important to note that the research question in all examples is whether ***exposure at t_0 causes an outcome at t_1***.

As a first step we will start with a cross-sectional study (Scenario 1, Figure 1). In this case, the study starts at t_0. At that same time point, the researchers also ascertain the exposure and outcome status of their study population. In view of this concomitant measurement of exposure and outcome, the underlying actual research question is about the co-occurrence of two states labelled as "exposure" and "outcome", instead of our original question described above whether exposure causes the outcome. However, even in this cross-sectional case, it could be argued that the researchers still want to answer an underlying causal question and the exposure measured at t_0 reflects (or maybe misrepresents) the exposure status of an earlier time point. However, this second interpretation is like scenario 2b discussed below.

Moving to our next four scenarios, we want to point out that each of these can be viewed from two perspectives: either from a cohort perspective, that is defining a cohort and exposure status, then moving on to the outcome; or a case-control perspective defining cases and identifying the underlying cohort, then moving on to the selection of a control sample and determining exposure status. Also, it is crucial to realize that the design of a study (cohort/case-control) is independent of the timing of data collection (prospective/retrospective). Finally, we want to point out that in these four scenarios the exposure status is obtained at ***one time point***, that incorporates all stages from initial raw data collection to the final data on exposure status used in the analysis phase. This is crucial as later on we will describe situations in which the determination of exposure status is spread over time (e.g. raw data collection at one timepoint and subsequent processing taking place at a later time point) leading to further confusion when applying the terms prospective and retrospective, because this "spreading out" aspect of data collection was not incorporated in the original definitions of these terms. We will start by describing two scenarios in which both definitions lead to the same conclusion (Figure 2).

Concordant scenarios

In Scenario 2a, we have a study that starts at time point t_0. At that time point the exposure status is also determined (including raw data collection and processing), and subsequently the outcome is determined at t_1. As already stated, this could

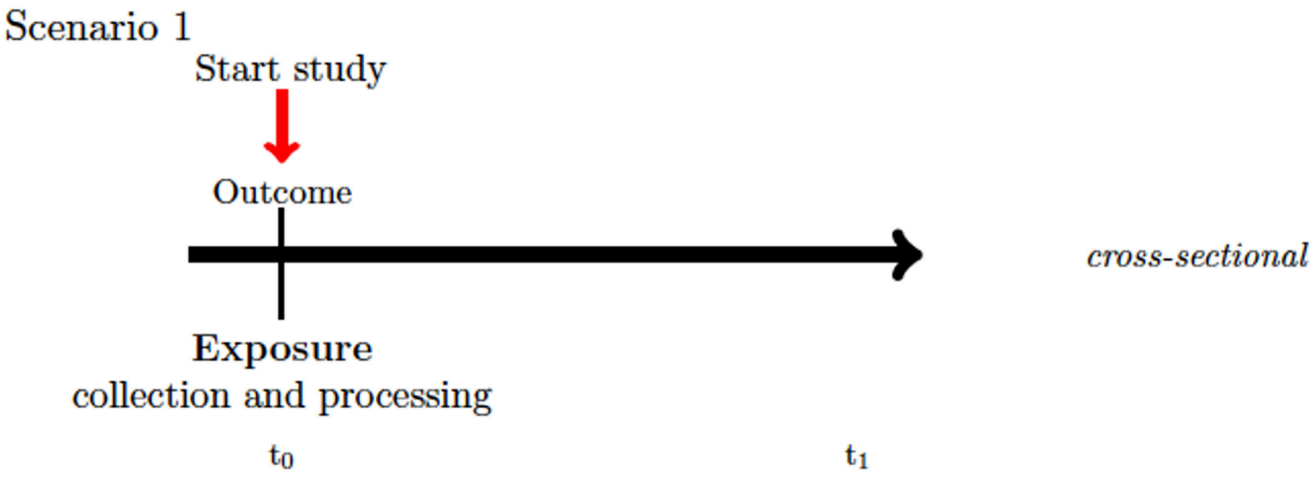


Figure 1. Cross-sectional study.

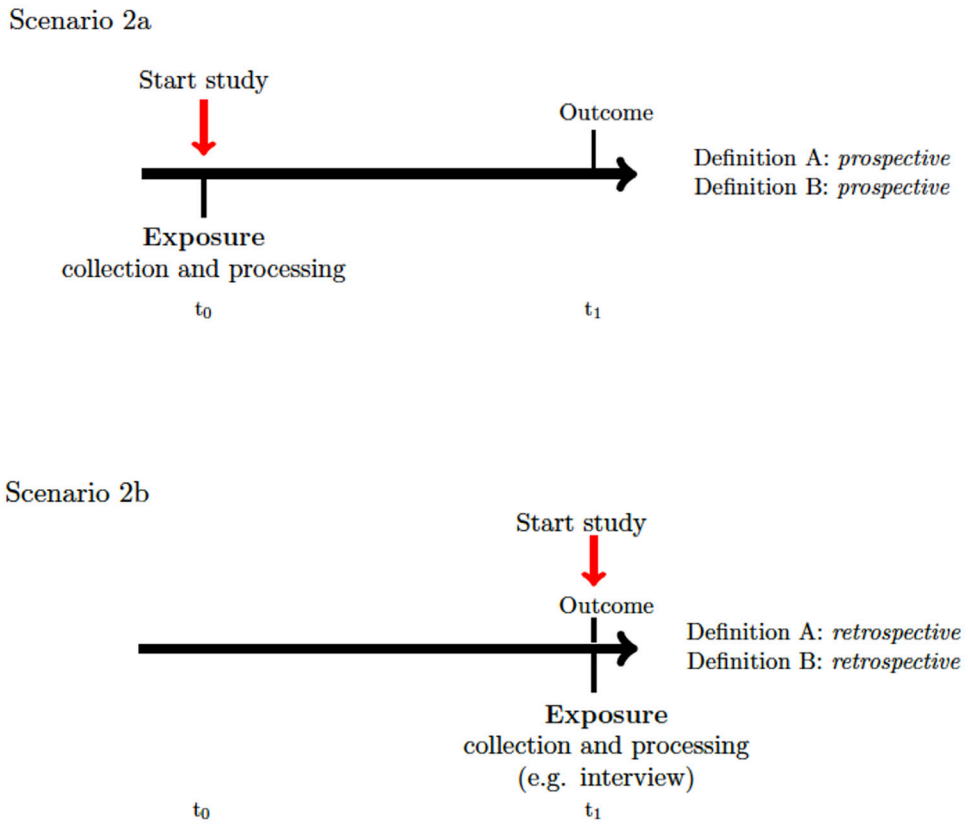


Figure 2. Two scenarios in which the two definitions give concordant use of the terms prospective and retrospective.

be a cohort study or a (nested) case-control study design. In this case, according to both definitions this situation will be labelled as "prospective", albeit on slightly different grounds: using definition A, the reasoning would be that the timing and order in which the recording of exposure status took place precedes the outcome. Using Definition B, the argument would be that the person-time accumulated after the study had started. Despite this difference in underlying arguments, both definitions would agree to label this study as "prospective".

Another scenario in which both definitions would lead to a concordant conclusion is Scenario 2b. In this case, a researcher at t_1 decides to start a study by identifying either a cohort or selecting cases/controls and then asks the participants to provide information about an exposure that occurred earlier (time point t_0). The outcome status is known at t_1, so there is no accumulation of person-time after t_1. Using definition A: as the data are collected through e.g. interview at timepoint t_1 (to ascertain the actual exposure status at t_0), the timing and order of exposure collection does not precede the outcome and there is thus a possibility that the exposure measurement could be influenced by the outcome status. So, this study according to definition A will be labelled as "retrospective". Using definition B, the accumulation of person-time occurred before the start of the study, hence again this study would be labelled as "retrospective".

Discordant scenarios

After considering these two concordant scenarios, we will discuss two discordant scenarios, that is the two definitions lead to different conclusions (Figure 3). In Scenario 3a, the researcher decides to conduct a study at timepoint t_1, by either identifying an existing cohort (e.g. registry, GP files) or selecting cases/controls from a clinic setting. In both settings, the outcomes are determined at t_1, whereas in this case data on the exposure at timepoint t_0 happened to be present as part of routine clinical practice (interview data or blood measurements) or entered into a registry by an insurance company (administrative data). Now, using definition A, the researcher will reason that the timing and order of collection of exposure status preceded the outcome, even though the study started at t_1, so labelling this study as "prospective". In addition, this researcher might say that the exposure status could not be influenced by the disease status as the exposure data were already present at t_0, further strengthening the conclusion that this is "prospective". Finally, even though the exposure data were not collected as part of a research protocol, but as routine daily practice, according to definition A this will not

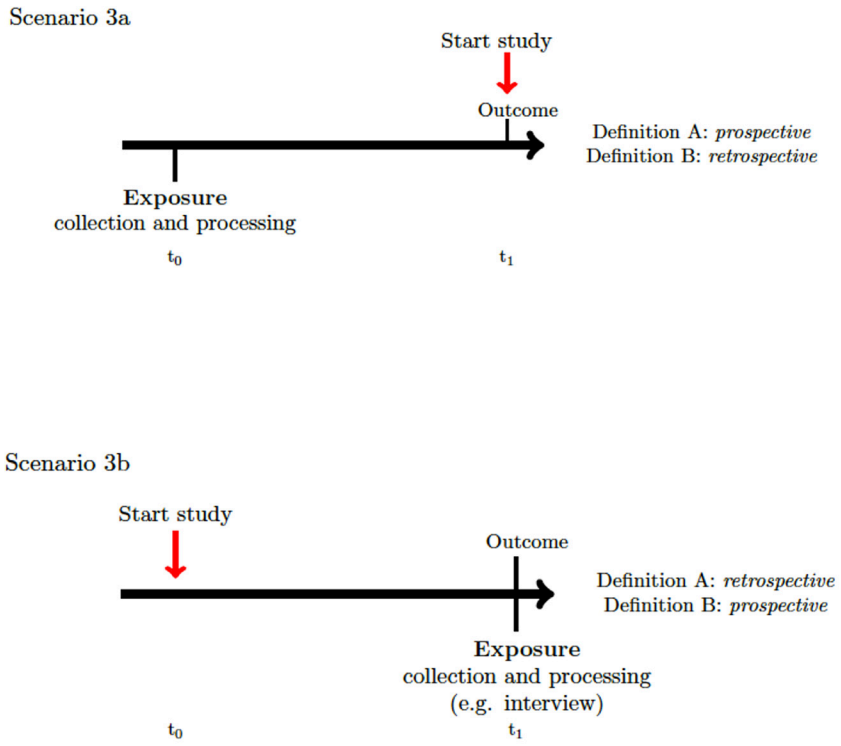


Figure 3. Two scenarios in which the two definitions give discordant use of the terms prospective and retrospective.

have any influence on our conclusion. Taking this argument to its logical end, identifying a "historical" cohort in which exposure collection preceded the outcome would always be labelled a "prospective cohort study".

Now, applying definition B to scenario 3a will lead to a different conclusion, namely: the study is setup at timepoint t_1, whereas the accumulated person-time has already occurred before this timepoint, so this study will be labelled as "retrospective". This example nicely highlights the fact that although the two definitions may seem similar, but in fact they are based on different aspects of the study and thereby can lead to different conclusions. The first main difference between the two definitions is that definition A focusses on the ***timing of measurement of exposure with respect to the outcome***, whereas definition B focusses on the ***accumulation of person-time***. This discordant scenario may explain some of the confusion in the medical literature.

Moving to the next discordant example we have Scenario 3b, in which there is a study that started at t_0 with collection of some initial exposure data and outcome determination at t_1. At a later stage, the researcher becomes interested in a ***new*** exposure variable. The researcher wants to know the value of this new exposure at timepoint t_0, but decides to collect these data through an interview at t_1. Now, using definition A, the researcher concludes that the timing and order of this new data collection does not precede the outcome, terming this as "retrospective". In contrast, using definition B, the researcher will argue that the study had already started at t_0 and the person-time was already being accumulated after the start of the study, so this is a "prospective" study. The fact that the new variable is collected at timepoint t_1 through e.g. an interview may lead to (recall) bias, but in definition B, this is irrelevant with respect to the use of the terms "retrospective" and "prospective". Furthermore, this scenario highlights a second difference between the two definitions, namely definition A focusses ***more on the specific research question level (exposure-outcome relation)***, whereas definition B tends to focus ***on timing of the overall conduct of the study***. Of course, if a study only focusses on one research question (as was the case in most of the original epidemiological studies conducted in the 1950s), then these two levels are the same. However, studies nowadays comprise many research questions, many of them explicitly formulated at later time points, thereby leading to this difference in application of these two definitions at two different levels.

When considering the abovementioned four scenarios, a ***third difference between the two definitions is that in definition A there is a tendency to take the possibility of bias into account, whereas this aspect is irrelevant in definition B***. This aspect is encapsuled in the idea that the exposure measurement could not be influenced by the disease. However, we want to point out that incorporating the possibility of bias into the definitions of "prospective" and "retrospective" unnecessarily complicates the situation further and should not be part of the definition. The issue of disease status influencing exposure should

be dealt with in the discussion of biases (e.g. information bias) and should not be part of the timing of data collection.

Data collection spread over time

Thus far, we discussed several scenarios, of which two cases resulted in discordant conclusions. As was already noted earlier, over the last few decades the nature of data collection and storage has changed considerably. It is important to realize that determining exposure status is not simply a matter of one time point, but consists of different aspects, including raw data collection (e.g. blood samples, imaging, -omics) and subsequent processing to obtain the final exposure status (e.g. genetic variant, imaging marker). All these aspects may be accomplished at one time point, but may – as is increasingly the case – be spread out over multiple timepoints, in many cases over years. In the previous scenarios, exposure status was either determined at timepoints t_0 or t_1. However, in the following two scenarios various aspects of the data collection process are spread out over the period t_0 and t_1, which further complicates the designation "prospective" and "retrospective". It is important to realize that the original definitions of these terms did not accommodate the possibility of data collection spread out over time. Hence, the application of these terms is dependent on how the time of data collection is interpreted, as we will discuss in the following examples.

Scenarios 4a/b (Figure 4) are more akin to Scenarios 2a/3a, but with the difference that the original raw data collection took place at t_0, but the final determination happened at t_1. In scenario 4a, the study started at t_0, which could be a cohort or a nested-case control study, the initial raw data (e.g. blood samples) were already collected at t_0, but the actual measurement of a biomarker was done at t_1. In this case, although the "spread out" aspect of data collection is not incorporated in definition A, the researcher using this definition could argue that it is the actual raw data at t_0 that needs to be considered and not the final determinations at t_1, so labelling this study as prospective. In the case of definition B based on the accumulated person-time, this would also be a prospective study as the person-time accumulation started at t_0. Again, in the case of definition B, this does not accommodate the "spread out" aspect of data collection.

Finally, if the study started at t_1 (Scenario 4b), we would again have a discordant conclusion. This situation could arise in a registry setting or a clinical setting in which for example clinical data or blood samples were collected as routine practice, but only at a later stage t_1 does a researcher decide to use these for research purposes in a study. The difference with scenario 3a is that in the current case the processing of the exposure occurs at t_1 instead of t_0. According to definition A,

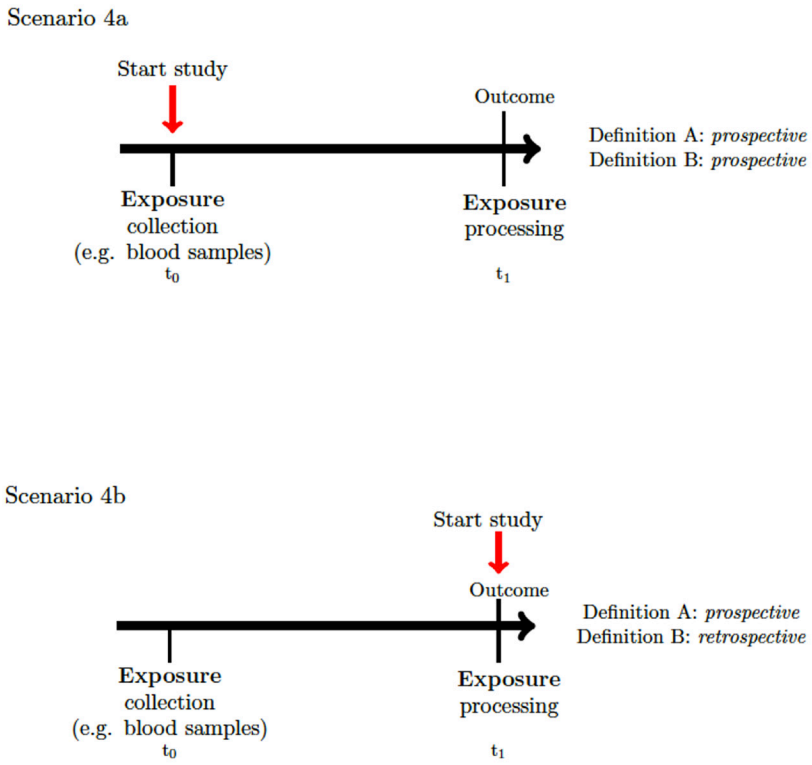


Figure 4. Scenarios in which the data collection process is spread over time.

as the original samples were collected at t_0 and thus preceded the outcome, this would be a prospective study. However, according to definition B, the study starts at t_1, and the accumulated person-time has already occurred, hence this would be called a retrospective study. Here again, the subtle difference between definitions A and B with respect to the timing of the exposure-outcome relation and timing of person-time leads to a discordant conclusion.

Discussion

In this paper, we described the different definitions used for the terms "prospective" and "retrospective" and applied these to different contexts, revealing instances where these interpretations could yield conflicting conclusions, thereby showing that in certain situations these definitions may lead to discordant conclusions. These instances of discordance offer valuable insights into the on-going confusion surrounding the usage of these terms in medical literature. Furthermore, considering the guidelines laid out in the 2007 STROBE guidelines – either to avoid using these terms or to provide explicit definitions - we can reasonably deduce from the current medical literature that the medical research community has not effectively adhered to the initial part of the STROBE recommendation. Consequently, it becomes imperative, as we move forward, to establish clear and unambiguous definitions for these terms.

It is important to underscore that the terms "prospective" and "retrospective" have most value if these address aspects of a study that are not already covered by other existing terminology. Firstly, elements relating to potential bias, as outlined in Definition A, are therefore preferably not included in the definition of these terms. Secondly, the temporal relationship between exposure and outcome, as explicated in Definition A, is already captured by another term, namely "longitudinal". Thirdly, the aspect of accruing person-time, as defined in Definition B, might not always be evident, particularly in situations where large volumes of data are being collected (simultaneously as part of routine clinical practice and biobanking) without a clear specification when the study actually started. Finally, we want to point out that in the definitions discussed in this paper it is not always clear what is meant by the "start of a study" and the importance of the application of the eligibility criteria is not explicitly mentioned. Whereas eligibility criteria are an essential part of any clinical trial, in observational studies these are often not mentioned explicitly. Moreover, even if they are mentioned, they are mostly used to explain how the final study population was acquired. However, the ***timing of the implementation of the eligibility criteria*** is often missing and is not covered by any other epidemiological term. We propose that the terms "prospective" and "retrospective" should specifically be used to describe the relationship between the implementation of the eligibility criteria (which may include obtaining informed consent) and the accumulation of person-time. Hence, we define the moment a researcher decides to apply the eligibility criteria either to recruit subjects into a study or apply the eligibility criteria to an existing dataset as the "start of the study". If in a study the accumulation of person-time occurs after the implementation of the eligibility criteria, for example, a cohort where the eligibility criteria are applied at baseline and the participants are followed up for the outcome, then this will be a prospective study. In contrast, when the eligibility criteria are implemented and person-time has already been accrued, then this will be labelled as a retrospective study. Examples of the eligibility criteria being applied at the end are in a case-control study where cases and controls are asked about their past exposure or a study where consent is required to have access to data that have already been collected for other (e.g. clinical) purposes. Here, we would like to point out that obtaining informed consent is an integral part of the eligibility criteria. However, there are cases in which informed consent cannot be obtained (historical data, medical files). Even in these cases, this updated definition will not lead to any confusion, because a researcher still must select a timepoint when the eligibility criteria will be implemented and hence can decide how the accumulation of person-time is related to this moment. Furthermore, our proposed definition does not signal the existence of a bias, though it may be suggestive of the potential for some biases. Hence, this definition will not have any in-built negative connotation about the quality of the study. Our proposed definition is closer to Definition B. However, the main difference is that in this revised definition the moment of implementation of the eligibility criteria is highlighted explicitly. Finally, specifying this moment will not cause any confusion related to the spread-out nature of data collection, as explained in scenarios 4a/b.

In conclusion, we propose to use the terms "prospective" and "retrospective" to describe an aspect of a study not covered by other epidemiological concepts (e.g. bias, study design), specifically the relationship between the implementation of the eligibility criteria (which includes obtaining informed consent) and the accumulation of person-time. We believe that a consistent application of a clear definition will in the future reduce the confusion surrounding these terms in the medical literature and will help researchers better appreciate the overall design of a study.

Ethics approval

Not applicable.

Data availability

No data are associated with this article.

Statements & declarations

Author contributions

Conceptualization: all authors

Methods: all authors

Writing – Original Draft Preparation: MK Ikram

Writing – Review & Editing: JA Labrecque & MA Ikram

References

1. Caram-Deelder C, Kreuger AL, Rosendaal FR, *et al.*: **Continuing use of the terms prospective and retrospective and quality of reporting of observational studies: time to update the STROBE guideline?** *Int J Epidemiol.* 2016; **45**(2): 587–9.
PubMed Abstract | **Publisher Full Text**

2. Von Elm E, Altman DG, Egger M, *et al.*: **The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.** *Lancet.* 2007; **370**(9596): 1453–1457.
PubMed Abstract | **Publisher Full Text**

3. Vandenbroucke JP, von Elm E, Altman DG, *et al.*: **Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration.** *Epidemiology.* 2007; **18**(6): 805–835.
PubMed Abstract | **Publisher Full Text**

4. von Elm E, Altman DG, Egger M, *et al.*: **Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.** *BMJ.* 2007; **335**(7624): 806–8.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. Dekkers O, Groenwold R: **Study design: what's in a name?** *Eur J Endocrinol*. 2020; **183**(6): E11–E13.
PubMed Abstract | **Publisher Full Text**

6. Schneeweiss S, Rassen JA, Brown JS, *et al*.: **Graphical depiction of longitudinal study designs in health care databases.** *Ann Intern Med*. 2019; **170**(6): 398–406.
PubMed Abstract | **Publisher Full Text**

7. Rothman KJ: **Modern Epidemiology**. First Edition.

8. Lash T, VanderWeele TJ, Haneuse, *et al.*: **Modern Epidemiology**. Fourth Edition.

Open Peer Review

Current Peer Review Status: ✔ ?

Version 2

Reviewer Report 21 August 2025

https://doi.org/10.21956/openreseurope.20536.r56614

© **2025 Bots S.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔

Sophie Bots
Utrecht University, Utrecht, The Netherlands

The authors have incorporated my suggestions. I have no further comments.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: (pharmaco-)epidemiological methodology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 21 August 2025

https://doi.org/10.21956/openreseurope.20536.r46685

© **2025 Ng A.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Annie Y Ng
Kheiron Medical Technologies, London, UK

Summary:
The manuscript tackles an important and persistent source of confusion in epidemiological study design—the use of the terms “prospective” and “retrospective.”

The authors do a commendable job outlining the historical context, identifying competing definitions, and proposing a more precise and useful framing centered on the relationship between eligibility criteria and person-time accumulation.

Primary suggestion:

While the scenarios presented are helpful in illustrating the conceptual distinctions between definitions A and B, the manuscript would benefit from the inclusion of more concrete, real-world examples for each scenario. For instance, integrating published or hypothetical case studies across a range of research settings—such as registry-based studies, clinical trials, biobank studies, or EHR-based research—would greatly enhance the practical understanding and applicability of the concepts.

Additionally, readers may appreciate if a few of these case examples are explicitly mapped to both definitions (A and B) and the proposed new definition. This would underscore how the proposed terminology resolves ambiguity and could serve as a teaching tool for researchers and students.

Such additions would strengthen the clarity and impact of the manuscript and help it serve as a more comprehensive guide for those navigating the nuances of study timing terminology in the context of modern epidemiological research.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: clinical trials, trial design, retrospective and prospective studies

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Version 1

Reviewer Report 19 November 2024

https://doi.org/10.21956/openreseurope.20166.r45545

© **2024 Bots S.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Sophie Bots
[1] Utrecht University, Utrecht, The Netherlands
[2] Utrecht University, Utrecht, The Netherlands

The authors set out to clear up confusion surrounding the terms "prospective" and "retrospective" used in epidemiological research. I think this is an important topic, and laud the authors for addressing this.
Currently these terms are used interchangeably to indicate how data was collected (from the start of the study or before start of the study) and to differentiate between cohort designs (where exposure information is collected before the outcome occurs) and case-control designs (where exposure information is collected when/after the outcome occurs). This gives rise to confusion terminology like "retrospective cohort", where it is unclear whether the "retrospective" refers to data collection or study design choices. This issue is complicated by the fact that some inflate retrospective/prospective with risk of bias, resulting in the idea that a "retrospective cohort" is by definition more biased than a "prospective cohort", which is not true. To clear up this confusion, the authors suggest to reserve "prospective" and "retrospective" for the chronological relationship between the implementation of eligibility criteria (e.g., baseline) and the accumulation of follow-up time.

Major

My only major comment is to include some additional papers that have been published in recent years in their discussion to strengthen the piece. Please find my suggestions below:
Below, I suggest some other (more recent) papers that I think the authors could include to strenghten their piece.
Refer: 1,2

Minor

(*) There are some language mistakes (e.g. "typically comprise of" instead of "comprise").

References

1. Schneeweiss S, Rassen JA, Brown JS, Rothman KJ, et al.: Graphical Depiction of Longitudinal Study Designs in Health Care Databases.*Ann Intern Med*. 2019; **170** (6): 398-406 PubMed Abstract | Publisher Full Text
2. Dekkers O, Groenwold R: Study design: what's in a name?. *European Journal of Endocrinology*. 2020; **183** (6): E11-E13 Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and does the work have academic merit?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: (pharmaco-)epidemiological methodology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 19 Nov 2024

Mohammad Kamran Ikram

We thank the reviewer for the feedback and have included the additional references to the paper and revised the language mistakes.

Competing Interests: No competing interests were disclosed.