

# Observational Studies

Dean R Hess

- Introduction
- Retrospective Studies, Prospective Studies, and Bias
- Multi-Center Versus Single-Center
- Prospective Observational Study Designs
  - Cross Sectional
  - Prospective Longitudinal Cohort
  - Interrupted Time Series
  - Quasi Experimental
  - Pre-Post
- Retrospective Study Designs
  - Case Study
  - Case Series
  - Single Cohort
  - Retrospective Comparison Cohort
  - Case Control
  - Matched Case Control
  - Prospective Study With Historical Controls
  - Pre-Post
- Data Registries
- Administrative Databases
- Secondary Analysis
- Writing the Paper
- Summary

Studies can be observational or experimental. With an observational study, the investigator does not determine the assignment of subjects, and there might not be a control group. If there is a control group, assignment of the independent variable (exposure or intervention) is not under the control of the investigator. Observational studies can be rigorously conducted, but the lack of random assignment of the exposure/intervention introduces confounding and bias. Thus, the quality of evidence resulting from observational studies is lower than that of experimental randomized controlled trials (RCTs). An observational study might be performed if an RCT is unethical, impractical, or outside the control of the investigator. There are many types of prospective and retrospective observational study designs. However, an observational study design should be avoided if an experimental study is possible. Sophisticated statistical approaches can be used, but this does not elevate an observational study to the level of an RCT. Regardless of quality, an observational study cannot establish causality. *Key words: case control; case series; cohort study; cross-sectional study; interrupted time series; matched case control; observational study; prospective study; quasi experimental; retrospective study.* [Respir Care 2023;68(11):1585–1597. © 2023 Daedalus Enterprises]

## Introduction

Clinical study designs can be experimental or observational. For experimental studies, subjects are assigned to control and interventional (exposure) groups by random assignment—a randomized controlled trial (RCT)—and outcomes are compared between the 2 groups. Probability sampling, or random selection of subjects from the population of interest, is used in experimental designs. Non-probability sampling is used in observational studies where study subjects are not chosen at random, and outcomes are available for retrospective or prospective analysis. For observational studies, the investigator does not determine the assignment of subjects, and there might not be a control group. If there is a control group, assignment of the independent variable (exposure or intervention) is not under the control of the investigator. Observational studies can be rigorously conducted, but the lack of random assignment of the exposure/intervention introduces confounding and bias. Thus, the quality of evidence from observational studies is lower than that of RCTs. Many observational studies are pragmatic, meaning that they are designed to show real-world effectiveness with no attempt to understand the underlying mechanisms.<sup>1</sup>

An observational study might be performed rather than an experimental study because an RCT is unethical, impractical, or outside the control of the investigator. For example, it is unethical to conduct an RCT in which subjects are assigned to vaping or not to determine whether vaping leads to COPD. This is also impractical due to the long lag time between exposure and disease. And it might be outside the control of the investigator if local laws ban the sale of electronic cigarettes and their use. Other ways that an RCT might be impractical include lack of resources (eg, funding, time), expertise, and rare occurrences of a disease or therapy.

Table 1 lists the steps in an observational study design, and Table 2 lists questions to ask during the conception and design of the study. The questions in Table 2 might also be considered by peer reviewers and readers of the published study. There may be as much time and planning for an observational as for an RCT. Just as for an RCT, the study

Table 1. Observational Study Design

Clinical observations.
Literature search; consult with a librarian.
Develop the study idea/question/hypothesis.
Consult an expert/mentor.
Design the study.
Consider statistical issues such as sample size, confounding, minimizing bias; consult with a statistician.
Write the protocol: where data will be found, what data will be needed, how data will be collected.
Obtain permissions, including Institutional Review Board.
Collect data.
Analyze data; consult with a statistician.
Publish the findings.

design must be carefully articulated, and approval from an institutional review board (IRB) or ethics committee is needed. The help of a statistician is strongly advised.

An observational study is considered inferior to an RCT. However, a Cochrane report found that there was little evidence for significant effect estimate differences between observational studies and RCTs.<sup>2</sup> That said, an observational study can never determine causation; the determination of cause and effect requires an experimental study design. There are several observational study designs that I will describe in this paper, including examples from papers published in *RESPIRATORY CARE*. Whereas it is important for investigators to understand the nuances of observational study designs, it is equally important for readers to understand observational study designs so that they can critically assess the published literature.

## Retrospective Studies, Prospective Studies, and Bias

A retrospective study uses existing data that have been recorded for reasons other than research.<sup>3</sup> In medical research, these are colloquially called chart reviews because the data source is often the medical record, recognizing that in most cases today this is the electronic medical record. Figure 1 contrasts retrospective and prospective studies. A retrospective design is discouraged when a prospective study is feasible. One useful application of a retrospective study is as a pilot study that is completed in anticipation of a prospective study. A retrospective study can focus the study question, clarify the hypothesis, determine baseline measurements of the primary outcome to ensure an appropriate sample size, and evaluate the feasibility of a prospective study. Advantages and disadvantages of retrospective studies are shown in Table 3.

An observational study should have internal and external validity.<sup>4</sup> Internal validity of a study is its ability to measure what it sets out to measure. Bias and confounding threaten the internal validity of a study. External validity relates to

---

Dr Hess is affiliated with Massachusetts General Hospital, Boston, Massachusetts; and Northeastern University, Boston, Massachusetts.

Dr Hess discloses relationships with Daedalus Enterprises, American Association for Respiratory Care, American Respiratory Care Foundation, University of Pittsburgh, Lungpacer, Jones & Bartlett, McGraw Hill, and UpToDate. Dr Hess is managing editor of *RESPIRATORY CARE*.

Dr Hess presented a version of this paper at the AARC Summer Forum 2022, held July 26–28, 2022, in Palm Springs, California; and at the AARC Congress 2022, held November 9–11, 2022, in New Orleans, Louisiana.

Correspondence: Dean R Hess PhD RRT FAARC. E-mail: dhess@aacrc.org.

DOI: 10.4187/respcare.11170

Table 2. Considerations When Planning an Observational Study

What database(s) are available? Electronic medical records, registry, administrative, other?
Will the database provide the data to support the question/hypothesis?
What is the study population, the intervention/exposure, and the outcome(s)?
Do you have confidence in the veracity of the database to minimize bias?
Are data available to control for confounding?
Are data readily accessible, electronic transfer?
Can the anonymity of subjects be protected?
If several databases are needed, can you control for temporal bias?
If data are missing from one database, is it possible to find that data elsewhere?
Is a statistician available who is skilled in advanced statistical methods to control for bias and confounding?
Will you be satisfied that an observational study cannot establish causality?

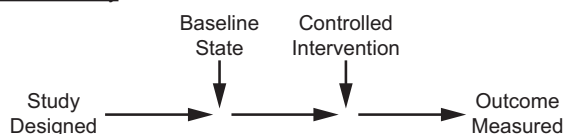
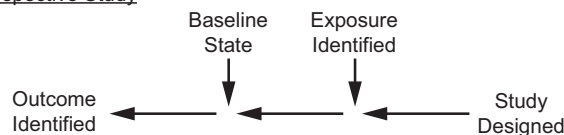
Prospective StudyRetrospective Study

Fig. 1. Prospective versus retrospective study.

generalizability, can the results of a study be applied outside of the constraints of the study? The value of the study is seriously compromised if its results cannot be applied to external local practice.

There are many potential sources of bias in retrospective studies. Bias is a systematic deviation from the truth. This affects the veracity of the study findings, potentially making the study invalid. Confounding is a distortion of the apparent effect of an independent variable on the dependent variable because of the influence of one or more other variables on the outcome. Confounding is a mixing of the effects of the independent variable with other influences on the outcome. It is important that investigators recognize the potential for confounding and bias and use strategies to minimize these in the study design.

To the extent possible, confounding and bias should be recognized during study design. Statistical methods might be necessary to control for their effect, eg, stratification, multivariable methods, inverse probability weighting, and propensity score matching.<sup>5,6</sup> Propensity score matching is a non-experimental causal inference statistical technique that attempts to balance groups based on confounding

Table 3. Advantages and Disadvantages of Retrospective Study Designs

Advantages

- Inexpensive.
- Uses existing records.
- Allows study of rare occurrences.
- Easier to assess conditions where there is a long latency between exposure and disease.
- Can generate hypothesis that is then tested prospectively.

Disadvantages

- Relies on accuracy of the record or recall of individuals (recall bias); garbage in → garbage out.
- Important data may not be available; nothing in → nothing out.
- Difficult to control bias and confounders: no randomization, no blinding.
- May not gain access to important information; restricted by statute or institutional regulations.
- Difficult to establish cause and effect.
- Results are, at best, hypothesis generating.

From reference 3.

factors to make them comparable. The input of a statistician is needed to consider sophisticated statistical methods during the study design. A common issue is that there might be important confounders unrecognized by the investigator, or known confounders may be unavailable in a retrospective study.

There are some sources of bias that are particularly important to retrospective studies. Selection bias occurs when subjects are included who are not representative of the population being studied. It can also occur if subjects are excluded who should be included. This can adversely affect the findings. In retrospective studies, it can be difficult to accurately categorize the subjects enrolled. In RCTs, randomization and concealed allocation are used to minimize the risk of selection bias.

Accuracy is a major problem in retrospective studies because the investigator cannot be certain that the data were correctly recorded or documented at the time of the event. Missing data are a major problem in retrospective studies. For example, bias is introduced into a retrospective study of lung-protective ventilation if the clinicians did not record plateau pressures or if they did not record them correctly. How data from missing cases are handled can further contribute to bias. There are several ways to adjust for missing data including removal of cases with missing data, substituting data from the last measured value, and imputation.<sup>7</sup> Each of these approaches has advantages and disadvantages, and help from a statistician is usually needed to apply these correctly. However, even the most sophisticated statistical adjustments will not completely remove the risk of confounding and bias.

Recall bias occurs if the study participants cannot correctly recall a past event in a retrospective study. Imagine a

study in which the investigators are interested to learn whether smoking as a teenager results in COPD in individuals age 60–65 y. Subjects are asked to recall their smoking habits as a teenager. They might recall correctly whether they smoked. But likely they will not recall how much they smoked, thus making it difficult to accurately quantify the dose of the exposure.

Performance bias occurs when subjects do not perform the intervention correctly. Imagine that you are doing a retrospective study evaluating outcomes of subjects with asthma who are using a pressurized metered-dose inhaler (pMDI) versus nebulizer for inhaled drug delivery. In a retrospective study, it is impossible to know if the subjects were using the device correctly (or at all). If subjects using the pMDI did not have good technique or if they were not adherent, that would bias the outcome in favor of the nebulizer. It would also clarify the importance of correct technique and adherence to therapy.

Temporal bias occurs when a study does not consider the effects of time on the outcome. For example, the outcome of 2 cohorts might be studied, one cohort for the first 6 months of the year and the other cohort for the second 6 months of the year. However, there might be confounders related to time of year that invalidate the results. For example, students typically graduate mid-year and begin their jobs soon after graduation. An evaluation of staff performance might be confounded by temporal bias due to the greater inexperience of staff in the second 6 months of the year. A better design is to compare performance of the first months of the year to the previous year. However, that might be biased due to the mix of staff from one year to the next.

Another form of temporal bias is associating of data that were not collected at the same time. Blood gas data might be associated with chest radiography results. However, if the chest radiograph was done at 6 AM and the blood gas was done at 11 AM, the results might not reflect the same patient condition, as there might be an improvement or worsening of the patient's condition over 5 h. As another example, imagine that you want to do a real-world study on pulse oximetry accuracy. You design a study of mechanically ventilated subjects and retrospectively examine the medical records of those subjects to collect pairs of arterial oxygen saturation and  $S_{pO_2}$ . Knowing that  $S_{pO_2}$  might not be recorded at the exact time that arterial blood was collected, you decide to record the  $S_{pO_2}$  within 1 h of the arterial blood collection. The instability that is common in mechanically ventilated patients suggests that it is quite likely that the  $S_{pO_2}$  at the time it was recorded might be different than it was at the time of the arterial blood collection. Additional bias might be introduced by using the time that the arterial blood sample was analyzed rather than the time that the sample was collected. A retrospective study design would also be unable to account for the poor quality of  $S_{pO_2}$  signal as occurs with poor perfusion, shock, and vasopressors.

Transcription errors can also be an issue with retrospective studies. Imagine that the person entering  $P_{aO_2}$  data into a registry transposes digits so that a  $P_{aO_2}$  of 96 mm Hg is recorded as 69 mm Hg. A transcription error like this will result in an error when summary data like mean and SD are calculated. Another example is recording a plateau pressure as 52 cm  $H_2O$  rather than 25 cm  $H_2O$ . In this case, it might be tempting for the investigator to assume that the correct plateau pressure must be 25 cm  $H_2O$ . But if 52 cm  $H_2O$  was indeed correct, this would bias the findings.

The Risk Of Bias In Non-Randomized Studies - of Interventions (ROBINS-I) can be used to assess bias in an observational study.<sup>8-10</sup> This tool assesses bias across 7 domains: (1) bias due to confounding, (2) bias in selection of participants, (3) bias in classification of interventions, (4) bias due to deviations from intended interventions, (5) bias due to missing data, (6) bias in measurement of outcomes, and (7) bias in selection of the reported result. Considering these domains, a judgment can be made for low risk of bias, moderate risk of bias, serious risk of bias, or critical risk of bias. ROBINS-I is most relevant when evaluating bias in a published study.

In an attempt to minimize the perception of bias, a study might be called a retrospective analysis of prospectively collected data. This designation is confusing. How can a study be both retrospective and prospective at the same time? This is important when considering bias, as prospective versus retrospective is defined based on the exposure and outcome in relation to when data are collected. A retrospective analysis of prospectively collected data is in fact a retrospective study. The term “prospectively collected” does little to reduce the risk of bias related to exposure to the intervention but could increase the quality of the collected data. But the investigator often has no control over the data recorded or the veracity of the data. Data routinely entered into the medical record are prospectively recorded, but that does not mean that the data were correctly collected and recorded. Arguably, the data might be more robust for a registry under the control of the investigator, particularly if there is robust auditing and quality control in place.

### Multi-Center Versus Single-Center

Observational studies can occur in a single setting (ie, single-center) or in many settings (ie, multi-center). A multi-center study increases external validity and is more generalizable. More centers contributing means that the results might be more applicable to another practice. But the external validity might not be as strong as first appears. When examining the study details, it might be noted that a few centers contributed most of the study results, thus limiting the external validity. Moreover, the greater external validity comes at the expense of internal validity. In a large multi-center study, it is difficult to assure that all participants apply

the intervention correctly. Outcomes for complex interventions such as high-frequency ventilation, extracorporeal membrane oxygenation (ECMO), or recruitment maneuvers are heavily influenced by the experience and skills of the teams at each center, which can lead to performance bias, and that might negatively affect the study's findings. The heterogeneity of the multi-center study can be both a positive (increased external validity) and a negative (decreased internal validity).

### Prospective Observational Study Designs

#### Cross Sectional

In a cross-sectional study, the investigator measures the outcome and the exposures in the study participants at the same time.<sup>11,12</sup> It is an observational study in which there is no prospective or retrospective follow-up. Once the subjects are selected, the investigators collect data and assess associations between outcomes and exposures. Cross-sectional studies can be classified as descriptive or analytical, depending on whether the outcome variable is assessed for potential associations with exposures or risk factors.<sup>12</sup> Descriptive cross-sectional studies characterize the prevalence of one or more outcomes in a population. Analytical cross-sectional studies collect data for both exposures and outcomes at a specific point in time to compare outcome differences between exposed and unexposed subjects. Because exposures and outcomes are measured at the same time, it is difficult to determine whether the exposures preceded or followed the outcomes. In a repeated, or serial, cross-sectional study, data are collected on the same target population at different time points. At each time point, investigators sample different subjects of the target population. Repeated cross-sectional studies are used for analyzing population changes over time but cannot be used to look at individual changes.

Jesus et al<sup>13</sup> conducted a cross-sectional study to investigate the association between caregiver burden and quality of life when caring for patients receiving long-term oxygen therapy (LTOT). Caregivers and patients were evaluated at a single point in time. The outcome was caregiver burden and quality of life, and the exposure was caring for a patient receiving LTOT. Validated tools were used to assess caregiver burden and quality of life and physical function of the patients. Increased dyspnea and dependence on activities of daily living in patients on LTOT were associated with a higher burden to their caregivers. The increased caregiver burden was associated with worse quality of life.

Bellani et al<sup>14</sup> evaluated the epidemiology, patterns of care, and mortality for subjects with ARDS. This was a cross-sectional study conducted for 4 consecutive weeks in the winter of 2014 with a convenience sample of 459 ICUs from 50 countries across 5 continents. The exposure was ARDS. The primary outcome was ICU incidence of ARDS, and secondary outcomes were clinician recognition of ARDS, application of lung-protective ventilation, use of

adjunctive interventions, and clinical outcomes. They reported an ARDS incidence of about 10% of ICU admissions. ARDS was underrecognized, undertreated, and associated with a high mortality rate. Of greatest concern, 35% of subjects received a tidal volume of > 8 mL/kg predicted body weight, and plateau pressure was measured in only 40%.

#### Prospective Longitudinal Cohort

A cohort is a group of individuals who share a baseline state (eg, COPD, ARDS), exposure/intervention (eg, noninvasive ventilation [NIV], standard care), or outcome (intubation, mortality). Prospective longitudinal observational cohort studies follow individuals over time.<sup>15</sup> These studies are useful to evaluate risk factors and the development of disease or the outcomes of treatments over time. Recall bias is excluded by collecting data prospectively. However, bias can be introduced if there are dropouts over time. It is also necessary to consider the role of confounders. Cost and time are important considerations when embarking on a prospective longitudinal observational study.

The Framingham Heart Study is a longitudinal cohort study that began in 1948 and continues to this day.<sup>16</sup> The town of Framingham, located about 20 miles west of Boston, Massachusetts, was selected as the first large epidemiological study of cardiovascular disease. In 1948, Framingham represented a typical middle-class American community with a stable population. The study now spans 3 generations of white subjects and 2 cohorts comprised of racial and ethnic minority groups. These cohorts have had extensive longitudinal follow-up with well-studied phenotypes. They have provided important information on cardiovascular and non-cardiovascular physiology over the life span and have identified major risk factors for cardiovascular disease. The Framingham Heart Study may be the most important and best-known longitudinal study ever conducted. This is an observational study because the investigators identify exposures and outcomes, but they do not control which subjects receive which exposures.

Sebbane et al<sup>17</sup> conducted a longitudinal study to evaluate the effect of weight loss on functional residual capacity (FRC) in obese subjects. The exposure was weight loss, and the outcome was changes in pulmonary function testing. Consecutive morbidly obese adult subjects enrolled in a bariatric surgery program were included. They found that reduction in FRC can be recovered following gastropasty-induced weight loss, despite residual mild to moderate obesity.

Although this study design often enrolls human subjects, it can also be applied to devices in a bench study. Awad et al<sup>18</sup> conducted a longitudinal evaluation of compressor/nebulizer performance. Compressor/nebulizers were operated for 1 h twice daily 5 d/week for 24 weeks. Compressor flow/pressure characteristics were measured every 6 weeks. They found

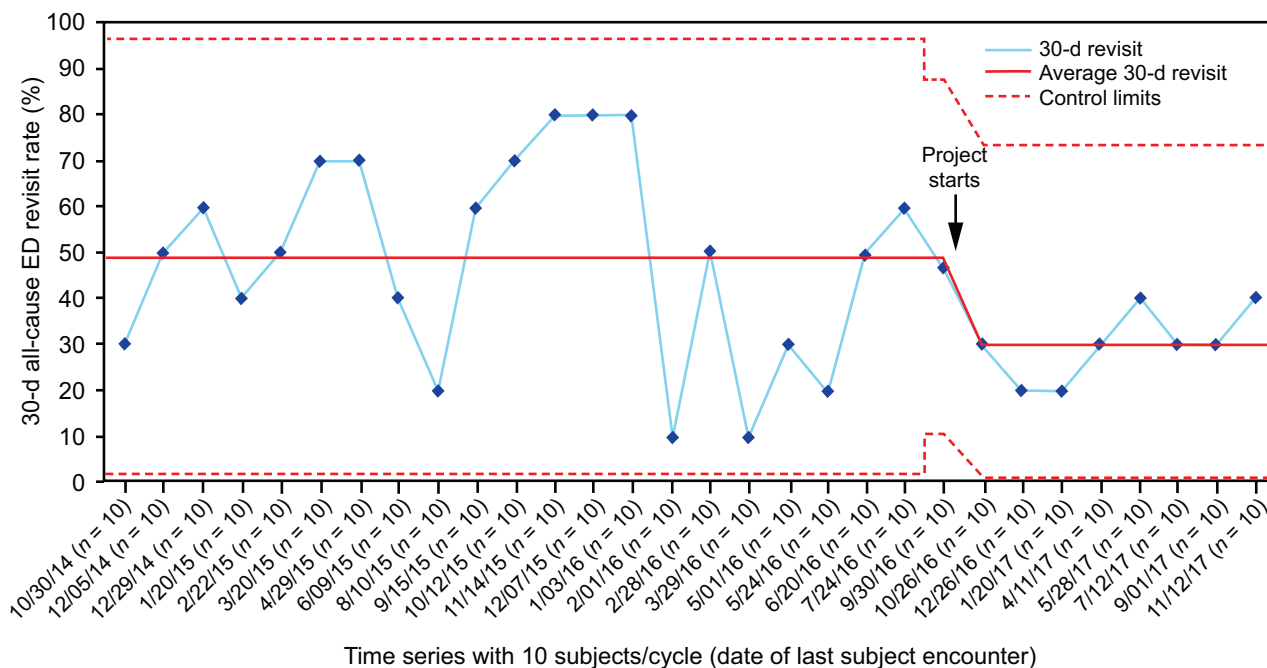


Fig. 2. Example of an interrupted time series. Thirty-day all-cause emergency department (ED) revisit rate among subjects with COPD exacerbation who were discharged from ED observation unit (ED-Obs) (outcome measure) is displayed on a statistical process control p-chart. Due to month-to-month variation in the number of subjects discharged from ED-Obs, the time series on the x axis is based on every 10 consecutive subjects. The dates on the x axis signify the date of last encounter with the subject. The 30-d all-cause ED revisit rate (%) is on the y axis. The baseline period extends from August 2014–September 2016, showing a stable system with common-cause (random) variation only. The last subgroup in the baseline period has 15 subjects, after which the bundle testing began in August 2016. During the post-bundle period, the data show a system shift with 8 consecutive observations below the center line (49%). New control limits were hence calculated with a new average of 30%. From reference 20. ED = emergency department.

that long-term use of compressor/nebulizers in a regimen like that of subjects with cystic fibrosis (CF) affected their performance. From a practical standpoint, they also found that maximal flow without and with the nebulizer could help identify compressors that are likely to fail.

### Interrupted Time Series

With an interrupted time series design, data are collected at multiple and equally spaced time points before and after an intervention.<sup>19</sup> It is important to know the exact time when an intervention occurs. The data pattern observed post intervention is evaluated to determine if it is different from that pre intervention. Zafar et al<sup>20</sup> evaluated the effect of a COPD care bundle in an emergency department (ED) observation unit. The intervention was the care bundle, and the outcome was ED revisits. An interrupted time-series analysis was used. The outcome shows a shift in the interrupted time series (Fig. 2).

### Quasi Experimental

Quasi experiments are studies that evaluate interventions but do not use true randomization. These are non-randomized or pseudo-randomized interventional studies. These

designs are used when it is not feasible or not ethical to conduct an RCT. Because quasi-experimental studies are not randomized, there is an increased risk of bias. For convenience, a study might allocate subjects by birth date (eg, odd years for intervention and even years for control), day of admission (eg, every other day or every other month to intervention or control), geography (eg, zip code), center (some implementing the intervention and some not), room number (eg, odd or even), or equipment availability (eg, available or not). There is a high risk of selection bias because the investigator can predict the allocation and thus affect assignment. Although these are comparison studies, they should not be confused with RCTs using subject-level randomization.

### Pre-Post

One example of a quasi-experimental design is a before-and-after (pre-post) study. A pre-post study measures outcomes in participants before an intervention and then again afterward. Lena et al<sup>21</sup> evaluated the effects of tracheostomy on patient-ventilator asynchronies and respiratory system mechanics. The study population consisted of 20 adult subjects receiving mechanical ventilation; the intervention was a tracheostomy, and the outcome was

asynchrony and respiratory mechanics. It is unethical to randomize subjects to tracheostomy or not to study the effect on asynchrony and mechanics. Subjects were evaluated during the 24 h before tracheostomy and the 24 h after tracheostomy. The short time between the 2 study periods minimized the risk of temporal bias. They found that tracheostomy did not affect patient-ventilator asynchronies or respiratory mechanics within 24 h before and after the procedure.

Given the known health effect of cigarette smoking, it is unethical to randomize individuals to smoking or not to assess health outcomes. But the implementation of a smoking ban creates the opportunity to conduct a natural experiment with a quasi-experimental design to evaluate the health-related effects of smoking. Helena, Montana, is a geographically isolated community that imposed a ban on smoking in workplaces and public places on December 3, 2002. This allowed a natural quasi-experimental study to examine the association of the smoking ordinance (intervention) with admissions for myocardial infarction (MI) from within Helena (intervention) and from outside Helena, where the ordinance did not apply (control).<sup>22</sup> The number of MI admissions fell from an average of 40 in the years before the ordinance to 24 during the 6 months after the law was in effect. From outside Helena during the same period, MI admissions increased from 12 in the years before to 18 after the law was in effect. To control for the effect of temporal bias, the same calendar months were compared before and after the implementation of the smoking ordinance. Identification of MI was extracted from the medical record, which might introduce bias if the diagnosis was incorrect. The authors concluded that smoke-free laws may be associated with a rapid effect on morbidity from heart disease. Because this was an observational study, a statement of cause and effect cannot be made, but an RCT with this question will never be done for ethical reasons.

In a pre-post design, Acho et al<sup>23</sup> evaluated the impact of a mechanical ventilation curriculum on respiratory therapist recognition of patient-ventilator asynchrony. Respiratory therapists from 2 academic medical centers enrolled in a one-day mechanical ventilation course. Prior to and following the course (intervention), they were asked to identify abnormalities on a 5-question, multiple-choice ventilator waveform exam (outcome). They reported a modest improvement in test scores after the intervention.

### Retrospective Study Designs

#### Case Study

A case report is a description of a single unusual and/or instructive case, such as symptoms not previously observed with a given medical condition or an unexpected or new combination of medical conditions in one case. A case

report is evidence but at the lowest level in the hierarchy of evidence. Due to the low level of evidence and high risk of bias, case reports are not published at all in some journals, including *RESPIRATORY CARE*, and are published selectively in other journals. A case report does not have much potential impact on the science of respiratory care. However, it does have a high value for teaching others.<sup>24</sup> Case reports are usually written by clinicians actively involved in the case described. As such, it can be difficult to distinguish between a good outcome due to the care provided or despite the care provided. Case reports of bad outcomes are usually not published, as the poor outcome does not reflect well on the author and brings attention to bad practice.

#### Case Series

A case series is a report of multiple similar unusual or instructive cases.<sup>3</sup> A retrospective case series can be used to study a disease that occurs infrequently or to generate a hypothesis that can be tested more rigorously in a prospective study. An important case series is the paper first describing ARDS.<sup>25</sup> This report was a description of 12 subjects with ARDS and was also the first report on the use of PEEP in individuals with ARDS. The first description of intermittent mandatory ventilation (IMV) was a case series of 6 subjects.<sup>26</sup> Based on this case series, IMV was shortly thereafter incorporated into ventilators as a mode and was adopted as an approach to weaning patients from mechanical ventilation. Although this small case series changed practice at the time, RCTs later reported that IMV was not the best approach to liberating patients from mechanical ventilation.<sup>27,28</sup>

#### Single Cohort

Single-cohort studies can be prospective or retrospective. Advantages of a single-cohort retrospective study are that they allow pooling of experience with a new or unusual disease or treatment. They can also generate a hypothesis to be tested more rigorously. Disadvantages include that investigators self-select (selection bias) and that there is no control group. A single-cohort retrospective study can be considered a more expansive case series. Descriptive statistics and logistic regression analysis are commonly used with a single-cohort retrospective study.

Moy et al<sup>29</sup> conducted a multi-center, nationwide retrospective cohort study on consecutive adults mechanically ventilated during air medical transport in the prehospital environment. The exposure was mechanical ventilation, and the outcome was low tidal volume ventilation. They used logistic regression analysis to adjust for potentially confounding variables. Female sex was an independent predictor of non-protective ventilation (adjusted odds ratio [OR] 6.79,  $P < .001$ ).

As another example, Mosher et al<sup>30</sup> performed a retrospective cohort study of hospitalized subjects with COPD

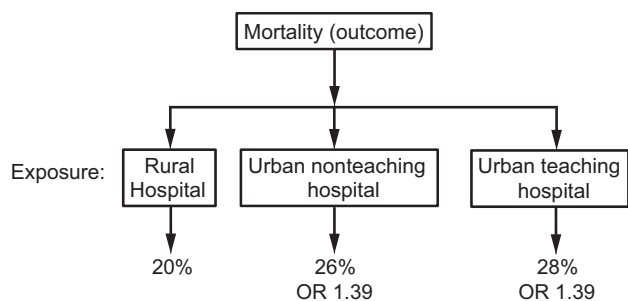


Fig. 3. An example of a retrospective comparison cohort. The outcome (mortality) is identified, and then exposures (rural hospital, urban non-teaching hospital, and urban teaching hospital) are compared. From reference 31. OR = odds ratio.

exacerbation. The intervention was NIV, and the outcome was NIV failure. Sophisticated statistical analysis was used including unadjusted multinomial logistic regression models to examine the association between NIV treatment outcomes and 17 recipient characteristics. They found that the first 8 h following initiation of NIV was a critical time when subjects were at risk for life-threatening decompensation. Their findings also suggest that careful consideration should be given to increasing age, body mass index (BMI), bicarbonate level, and creatinine level as these factors associated with NIV failure or persistent treatment.

### Retrospective Comparison Cohort

With a retrospective comparison cohort, the outcome is identified and then exposures are compared among groups. The results can be analyzed statistically using descriptive statistics, inferential statistics, OR, and relative risk. This can be illustrated in the study by Burton et al.<sup>31</sup> They used the 2014 National Inpatient Sample database to evaluate the association between hospital urbanicity and mortality. The outcome was mortality, and the exposure was hospital urbanicity. The authors performed a mixed-effects logistic regression analysis adjusting for sociodemographic variables and medical comorbidities. The exposure variable (hospital urbanicity) was rural hospitals, non-teaching urban hospitals, and teaching urban hospitals. The odds of in-patient mortality were significantly higher among urban teaching (OR 1.39,  $P < .001$ ) and urban non-teaching hospitals (OR 1.39,  $P < .001$ ) compared to rural hospitals (Fig. 3).

Eidman et al<sup>32</sup> conducted a retrospective review of subjects at least 3 months of age with acute respiratory failure requiring NIV who were admitted to a pediatric ICU (PICU). Subjects were stratified to those receiving continuous dexmedetomidine versus those not receiving sedation. A statistical method called augmented inverse probability weighting was used to create an equivalent baseline between the dexmedetomidine and no-sedation groups, thus adjusting for the confounding effects of lower age,

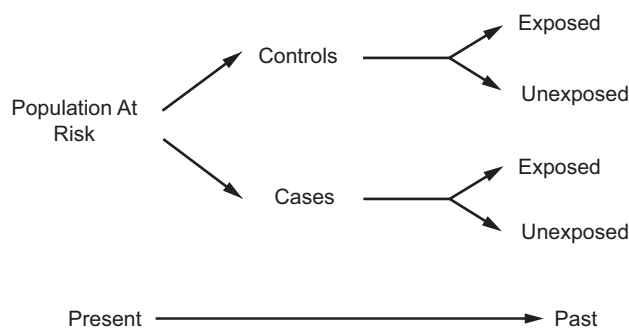


Fig. 4. Case-control study scheme.

developmental delay, and intellectual disability. Binomial logistic regression was then performed to determine the odds of intubation for dexmedetomidine subjects versus no-sedation subjects. A time-to-intubation analysis of subjects requiring intubation was performed using a Kaplan-Meier analysis. This illustrates the rigorous statistical methodology that can be used to adjust for the effects of confounding. Of 500 patients admitted to the PICU and receiving NIV during the study period, only 108 were included in the study (60 receiving dexmedetomidine and 48 no sedation), raising the potential for selection bias. Sedation scoring was not available for many subjects, illustrating the bias of missing data in retrospective studies. The authors concluded that dexmedetomidine may allow tolerance of NIV in acute respiratory failure without increasing risk for intubation, especially in preschool age patients and those with developmental delay or intellectual disability. They correctly acknowledge that a larger prospective multi-center study, ideally an RCT, is needed to support their conclusions.

### Case Control

A case-control study compares outcomes with and without an exposure (Fig. 4). It assumes that cases differ from controls only in having the exposure. Greater exposure in cases reflects increased risk or benefit. The degree of exposure to a risk factor is compared between the 2 groups. The exposure and outcome are determined retrospectively. Ideally, the data collectors should be blinded to whether a subject is a case or a control, and they should be blind to the study hypothesis. The cases and the controls must be assessed for exposure in the same way. Descriptive statistics, inferential statistics, and OR can be used to analyze the results.

In a classic case-control study, Doll and Hill<sup>33</sup> studied the relationship between lung cancer (outcome) and cigarette smoking (exposure). Subjects with lung cancer (cases) were identified in 20 London hospitals. An equal number of controls was identified among hospitalized subjects of the same age with diagnoses other than lung cancer. For males, the



OR was 14, meaning that male smokers had a 14 times greater chance of lung cancer than male non-smokers.

Certain types of bias are unique to case-control studies. For hospital-based case-control studies, the study population is the collection of clinical records of the participating hospital. However, the cases and the controls may have had different hospital admission rates, as seen with Berkson bias. Berkson's original example was a retrospective study examining a risk factor for a disease in a statistical sample from an in-patient population.<sup>34,35</sup> Because samples were taken from an in-patient population rather than from the general public, this resulted in a spurious negative association between the disease and the risk factor. For example, many patients with COPD exacerbation are not admitted to the hospital. A case-control study of COPD exacerbation using hospital records would select only the most severe cases.

For population-based case-control studies, the study population is the collection of subjects who become cases if they develop disease. This can result in cases that are not representative of the intended population as seen with Neyman bias (prevalence-incidence bias or survival bias).<sup>36</sup> It is a bias that occurs when prevalence cases are sampled and exposure affects disease and disease-associated mortality.<sup>37,38</sup> Neyman<sup>36</sup> identified this as a potential bias in the interpretation of case-control studies.

Neyman<sup>36</sup> used a fictitious example of lung cancer in smokers and non-smokers. Imagine that 10% of smokers get lung cancer and 20% of non-smokers get lung cancer. A prospective study would report a relative risk of lung cancer of 0.5 for smokers (10%/20%). However, in this fictitious example, 90% of smokers who get lung cancer survive, but only 5% of non-smokers who get lung cancer survive. Now imagine a case control study with an initial population of 10,000 smokers and 10,000 non-smokers. At the time of a case-control study one year later, the population has about 9,900 smokers and 8,100 non-smokers; this difference is due to the higher mortality among non-smokers. This leaves 900 with cancer among the smokers and 100 with cancer among the non-smokers, resulting in a relative risk of lung cancer is 7.37 (900/9,900)/(100/8,100). This difference in relative risk between the prospective study and the case control study is the result of excluding individuals who died from lung cancer. Use incidence cases to avoid Neyman bias. Incidence is limited to new cases only, whereas prevalence includes all cases, both new and preexisting.

Excluding patients who died usually makes the disease appear less severe. On the other hand, excluding patients who have recovered makes the disease seem more severe. More time between exposure and investigation makes it more likely for individuals to die or recover, therefore excluding them from the analysis. Thus, this bias is more likely to impact long-lasting diseases. Case-control studies are most susceptible to Neyman bias, but it can also occur in cross-sectional, experimental, and cohort studies.

Incidence-prevalence bias results from the inclusion of prevalence cases in a study. Prevalence cases overrepresent those who live the longest.<sup>39</sup> It is well known that continuing to smoke after COPD diagnosis increases mortality. In an incidence cohort, observation starts at diagnosis of severe COPD. Imagine that the mortality in those who continue smoking is 10 deaths in 14 patient-years and in those who quit smoking is 10 deaths in 26.5 patient-years. The relative risk for mortality in the incidence cohort is  $(10/14)/(10/26.5) = 1.89$ . In a prevalence cohort, observation in the same subjects was started 1 y after diagnosis. Only 14 subjects are left from the original incidence cohort because many subjects who continued smoking had died. The mortality in the prevalence cohort was 6 deaths in 5 patient-years in those who continued smokers and 8 deaths in 17 patient-years in those who quit smoking, leading to a biased relative risk for mortality of  $(6/5)/(8/17) = 2.55$ , which is substantially greater than that derived from the incidence cohort. The only way to prevent this type of bias is through limiting inclusion to incidence cases.

Strengths of a case-control study include fewer constraints by the frequency of the disease and a shorter waiting time than a prospective cohort study. They can be used when ethical considerations do not allow an RCT, such as the study by Doll and Hill.<sup>33</sup> Case-control studies can be feasible when RCTs are not, and case-control studies cost less and have fewer practical restrictions. The weaknesses include a less well-defined target population and risk of selection bias, Berkson bias, Neyman bias, and other biases. A case-control study cannot definitively establish cause and effect.

### Matched Case Control

With a matched case-control study design, the control subjects are selected so they match (resemble) the cases in certain characteristics (eg, age, sex, comorbidity, severity of disease). The goal is to compare case and control subjects who have similar characteristics and thereby to adjust for potential confounders and increase the precision of the comparison. Girou et al<sup>40</sup> conducted a matched case-control study to determine whether NIV was associated with a lower risk of nosocomial pneumonia. A sample of 50 subjects who received NIV (cases) was matched to 50 subjects who received invasive ventilation (controls). The controls were matched to cases based on the same diagnosis at admission, age  $\pm 5$  y, Simplified Acute Physiology Score II  $\pm 6$  points, Logistic Organ Dysfunction score  $\pm 3$  points, and no contraindications to NIV. The OR was 0.31, meaning that the odds of nosocomial pneumonia were lower among the subjects who received NIV.

### Prospective Study With Historical Controls

The outcomes associated with the intervention can be compared retrospectively to a control group. The intervention

cohort can be studied either prospectively or retrospectively. The interventional group and control group are not concurrent. This design is particularly prone to concerns for bias and confounding. For example, the control group might be dissimilar to the intervention group. Or the care team for the control group might differ in skill and experience compared to that for the intervention group.

Abdallah et al<sup>41</sup> performed an observational retrospective study of subjects with severe COVID-19 requiring mechanical ventilation > 48 h and compared these subjects to historical controls without COVID-19 who received mechanical ventilation > 48 h between 2016–2019. Data were collected from the electronic medical records using a standardized data collection form. The study population was mechanically ventilated subjects; the exposure was COVID-19, and the outcome was postextubation stridor. A propensity score was used to adjust for potential risk factors of postextubation stridor in both groups, which included variables with  $P < .05$  by multivariate logistic regression, specifically COVID-19 status, female sex, and tube mobilization or re-intubation or prone positioning. Each of the 65 subjects with COVID-19 was matched on the propensity score to a control. They found that postextubation stridor affected nearly one quarter of subjects with COVID-19, which was significantly higher than that seen in controls. Independent risk factors for postextubation stridor were COVID-19, female sex, and tube mobilization or re-intubation or prone positioning.

### Pre-Post

A pre-post study can be prospective or retrospective. Truumees et al<sup>42</sup> performed a retrospective pre-post interventional study comparing hospital readmissions for subjects with COPD exacerbation that received standard of care in the home versus a respiratory therapist–led home COPD disease management program. Subjects discharged home after COPD exacerbation were enrolled in the pre-intervention group. Subsequently, an evidence-based home COPD disease management program (the intervention) was implemented by a respiratory therapist in the home. The home COPD Disease Management Program was implemented from April 2017–September 2019, and this served as the post-intervention group. The primary outcome was readmission rates at 30 d. They reported the COPD Disease Management Program was significantly associated with decreased readmission adjusting for demographics and smoking status.

### Data Registries

Data registries are an organized system that use observational study methods to collect data to evaluate outcomes for a disease, condition, or exposure. The registry serves a scientific, clinical, or policy purpose. Advantages of registry studies are they allow for large sample sizes to test the associations

between interventions and outcomes. But they have significant disadvantages as they often have large amounts of missing data; they rely on administrative data; they may lack granular data, and they are prone to selection bias when evaluating rare interventions. The lack of granular data is a major limitation as it prevents investigators from accounting for important confounders in their analyses. In most cases, it requires significant time and expertise to clean, sort, and analyze the data. If data from a registry study are used for research, oversight from the local IRB is needed.

Trauma registries document acute care delivered to hospitalized patients with trauma, designed to provide information that can be used to improve the efficiency and quality of trauma care. Indeed, the combination of trauma registry data at regional or national levels can produce very large databases that allow opportunities for the evaluation of patient outcomes and inter-hospital comparisons.<sup>43</sup> There are many other registries including tumor (cancer) registries, ICU registries, surgical registries, and many others. Registries are often related to rare disorders but not necessarily. Local registries can be linked to national registries.

The CF Foundation Patient Registry collects information on the health status of people with CF who receive care in CF Foundation–accredited care centers and agree to participate in the registry (<https://www.cff.org/medical-professionals/patient-registry>. Accessed April 10, 2023). This information is used to create CF care guidelines, assist care teams providing care to individuals with CF, and guide quality improvement initiatives at care centers. Investigators use the patient registry to study CF treatments and outcomes and to design CF clinical trials. Using de-identified CF Foundation Patient Registry data, Sears et al<sup>44</sup> reported that gaps in CF care were associated with reduced lung function, even after controlling for other known factors associated with pulmonary compromise.

The Extracorporeal Life Support Organization (ELSO) is an international nonprofit consortium of health care institutions, researchers, and industry partners. ELSO provides support to those delivering extracorporeal life support through continuing education, guidelines, original research, publications, and a comprehensive registry of ECMO patient data (<https://www.elseo.org>. Accessed April 11, 2023). Peetermans et al<sup>45</sup> evaluated the impact of BMI on outcomes in respiratory failure necessitating ECMO using a retrospective analysis of the ELSO Registry from January 1, 2010–December 31, 2020. They found that subjects with BMI  $\geq 35$  kg/m<sup>2</sup> treated with ECMO for respiratory failure have lower mortality risk and shorter stays, despite increased cardiovascular, device-related, and renal complications. No upper limit of BMI indicating futility of ECMO treatment could be identified. They conclude that BMI as a single parameter should not be a contraindication for respiratory ECMO. However, there is a potential for selection bias due to the criteria used to determine whether ECMO was offered to patients.

The National Emergency Airway Registry (NEAR) is a multi-center observational intubation registry coordinated through the Department of Emergency Medicine at Brigham and Women's Hospital in Boston, Massachusetts. The NEAR for Children (NEAR4KIDS) is a multi-center registry for advanced airway management in PICUs. It is based at Children's Hospital of Philadelphia in the Department of Anesthesiology and Critical Care Medicine, and the database is maintained with the support of the Department of Emergency Medicine, Brigham and Women's Hospital, in Boston, Massachusetts. Brown et al<sup>46</sup> performed an analysis of ED intubations from the NEAR registry to evaluate videolaryngoscopy compared to augmented direct laryngoscopy in adult ED tracheal intubations. Videolaryngoscopy used without any augmenting maneuver, device, or technique resulted in higher first-attempt success than direct laryngoscopy that was augmented by use of a bougie, external laryngeal manipulation, ramping, or combinations thereof. They also observed fewer esophageal intubations in the videolaryngoscopy cohort.

### Administrative Databases

Administrative databases are typically derived from discharge abstracts of patient records by trained abstraction personnel.<sup>47</sup> They are intended primarily for billing. Unlike data derived from research databases, administrative data are captured for all patients, usually at many hospitals, making them appealing for observational research. When using large administrative databases for observational research, the investigator must appreciate the limitations of the data. The most important is that the principal objective of recording the data was not for research purposes.

From the Centers for Medicare and Medicaid Services, Spitzer et al<sup>48</sup> obtained beneficiary denominator and standard analytic files for every individual hospitalized in an acute care hospital in 2012 with a principal diagnosis of COPD or a principal diagnosis of acute respiratory failure combined with a secondary diagnosis of COPD exacerbation. To identify patients who received pulmonary rehabilitation, and providers of pulmonary rehabilitation, they used health care common procedure coding system codes from the Medicare out-patient file, which contains claims data from institutional out-patient providers (ie, hospital out-patient-based facilities), and carrier files, which contain claims from non-institutional providers (ie, physicians' offices). Using this administrative data, they concluded that participation rates for pulmonary rehabilitation after hospitalization were extremely low.

### Secondary Analysis

Secondary, or post hoc, data analysis research may be limited to descriptive, exploratory, and correlational designs and

Table 4. Considerations When Writing the Paper From an Observational Study

Introduction
Background.
Question/aims/hypothesis.
Establish that an observational study is appropriate.
Methods
Define the participants, exposures/interventions, outcomes.
Describe the data source(s).
Describe the data types.
Explain how bias and confounding are addressed.
Describe the sample size calculation.
Explain the statistical methods used.
Results
Provide tables and figures to show summary data.
Use parametric and non-parametric analysis correctly.
Provide the statistical results used to control for confounding and bias.
Discussion
Start with the important findings.
Interpret the findings.
Relate findings to similar studies in the literature.
Discuss the generalizability of the findings (or lack thereof).
Explain mechanisms that might support the study findings.
Acknowledge the limitations.
Do not overstate the importance of the results.

Consult the Standards of Reporting of Observational Studies in Epidemiology checklist (<https://www.strobe-statement.org/checklists>).

nonparametric statistical tests.<sup>49</sup> By their nature, secondary data analysis is observational and retrospective, and the investigator cannot examine causal relationships. Secondary data can be defined as data gathered for one reason being repurposed to answer a new question, whereas primary data are data collected specifically for the purposes of answering a new question.<sup>50</sup> Pena-Lopez et al<sup>51</sup> evaluated the effect of short-acting sedative-analgesic drugs to protect against the development of ventilator-associated events in children. This was a secondary analysis of the EUVAE study. They found that the use of continuous short-acting drugs, such as remifentanyl or propofol, were a strong protective factor against the development of pediatric ventilator-associated events. Burr et al<sup>52</sup> performed a post hoc analysis evaluating factors associated with a positive view of respiratory therapy leadership using data from a prior study of respiratory therapist burnout.<sup>53</sup> Most respiratory therapists had a positive view of their leadership. A negative leadership score was associated with higher burnout and missing work.

### Writing the Paper

Table 4 lists considerations when writing the paper following completion of an observational study. The Standards of Reporting of Observational Studies in Epidemiology

(STROBE) initiative developed recommendations on what should be included in an accurate and complete report of an observational study (<https://www.strobe-statement.org/checklists>. Accessed April 18, 2023).<sup>54</sup> The STROBE Statement is a checklist of items that should be addressed in articles reporting on cohort, case-control, and cross-sectional studies. The checklist includes 22 items that are considered essential for good reporting of observational studies. The objective of STROBE is to standardize reporting. It is not a tool for assessing the quality of published observational research.

### Summary

There are many types of prospective and retrospective observational study designs. Due to a lack of randomization, causal relationships between exposures and outcomes cannot be established with an observational study design. Nonetheless, observational study designs can be used when an RCT is either unethical or impractical. However, an observational study design should be avoided if an RCT is possible. Bias and confounding are major issues with observational studies, and statistical methods are available to control for these influences on outcomes. Sophisticated statistical approaches can be used, but this does not elevate an observational study to the level of an RCT. In the end, an observational study might suggest, but cannot establish, causality.

### REFERENCES

- Ford I, Norrie J. Pragmatic Trials. *N Engl J Med* 2016;375(5):454-463.
- Anglemyer A, Horvath HT, Bero L. Health care outcomes assessed with observational study designs compared with those assessed in randomized trials. *Cochrane Database Syst Rev* 2014;2014(4):MR000034.
- Hess DR. Retrospective studies and chart reviews. *Respir Care* 2004;49(10):1171-1174.
- Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet* 2002;359(9302):248-252.
- Andersen LW, Kurth T. Propensity scores - A brief introduction for resuscitation researchers. *Resuscitation* 2018;125:66-69.
- Fu EL, Groenwold RHH, Zoccali C, Jager KJ, van Diepen M, Dekker FW. Merits and caveats of propensity scores to adjust for confounding. *Nephrol Dial Transplant* 2019;34(10):1629-1635.
- Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
- Igelstrom E, Campbell M, Craig P, Katikireddi SV. Cochrane's risk of bias tool for non-randomized studies (ROBINS-I) is frequently misapplied: a methodological systematic review. *J Clin Epidemiol* 2021;140:22-32.
- Schunemann HJ, Cuello C, Akl EA, Mustafa RA, Meerpohl JJ, Thayer K, et al; GRADE Working Group. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. *J Clin Epidemiol* 2019;111:105-114.
- Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions. *BMJ* 2016;355:i4919.
- Setia MS. Methodology series module 3: cross-sectional studies. *Indian J Dermatol* 2016;61(3):261-264.
- Wang X, Cheng Z. Cross-sectional studies: strengths, weaknesses, and recommendations. *Chest* 2020;158(1S):S65-S71.
- Jesus LA, Malaguti C, Evangelista DG, Azevedo FM, Franca BP, Santos LT, et al. Caregiver burden is associated with the physical function of individuals on long-term oxygen therapy. *Respir Care* 2022;67(11):1413-1419.
- Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A, et al; ESICM Trials Group. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. *JAMA* 2016;315(8):788-800.
- Caruana EJ, Roman M, Hernandez-Sanchez J, Solli P. Longitudinal studies. *J Thorac Dis* 2015;7(11):E537-540.
- Andersson C, Naylor M, Tsao CW, Levy D, Vasan RS. Framingham Heart Study: JACC Focus Seminar, 1/8. *J Am Coll Cardiol* 2021;77(21):2680-2692.
- Sebbane M, El Kamel M, Millot A, Jung B, Lefebvre S, Rubenovitch J, et al. Effect of weight loss on postural changes in pulmonary function in obese subjects: a longitudinal study. *Respir Care* 2015;60(7):992-999.
- Awad S, Williams DK, Berlinski A. Longitudinal evaluation of compressor/nebulizer performance. *Respir Care* 2014;59(7):1053-1061.
- Hudson J, Fielding S, Ramsay CR. Methodology and reporting characteristics of studies using interrupted time series design in health care. *BMC Med Res Methodol* 2019;19(1):137.
- Zafar MA, Loftus TM, Palmer JP, Phillips M, Ko J, Ward SR, et al. COPD care bundle in emergency department observation unit reduces emergency department revisits. *Respir Care* 2020;65(1):1-10.
- Lena E, Aquino-Esperanza J, Lopez-Aguilar J, Magrans R, de Haro C, Sarlabous L, et al; Asynchronies in the Intensive Care Unit (ASYNICU) Group. Longitudinal changes in patient-ventilator asynchronies and respiratory system mechanics before and after tracheostomy. *Respir Care* 2021;66(9):1389-1397.
- Sargent RP, Shepard RM, Glantz SA. Reduced incidence of admissions for myocardial infarction associated with public smoking ban: before and after study. *BMJ* 2004;328(7446):977-980.
- Acho M, Kriner E, Sartain NN, Chatterjee S, Sun J, Lee BW, et al. Impact of a mechanical ventilation curriculum on respiratory therapist recognition of patient-ventilator asynchrony. *Respir Care* 2022;67(12):1597-1602.
- Pierson DJ. Case reports in respiratory care. *Respir Care* 2004;49(10):1186-1194.
- Ashbaugh DG, Bigelow DB, Petty TL, Levine BE. Acute respiratory distress in adults. *Lancet* 1967;2(7511):319-323.
- Downs JB, Klein EF Jr, Desautels D, Modell JH, Kirby RR. Intermittent mandatory ventilation: a new approach to weaning patients from mechanical ventilators. *Chest* 1973;64(3):331-335.
- Brochard L, Rauss A, Benito S, Conti G, Mancebo J, Rekkik N, et al. Comparison of three methods of gradual withdrawal from ventilatory support during weaning from mechanical ventilation. *Am J Respir Crit Care Med* 1994;150(4):896-903.
- Esteban A, Frutos F, Tobin MJ, Alia I, Solsona JF, Valverdu I, et al. A comparison of four methods of weaning patients from mechanical ventilation. Spanish Lung Failure Collaborative Group. *N Engl J Med* 1995;332(6):345-350.
- Moy HP, Nayman BD, Olvera D, Monnin K, Pappal RD, Hayes JM, et al. Mechanical ventilation practices and low tidal volume ventilation in air medical transport patients: The AIR-VENT study. *Respir Care* 2022;67(6):647-656.
- Mosher CL, Weber JM, Adagarla BS, Neely ML, Palmer SM, MacIntyre NR. Timing of treatment outcomes and risk factors for failure of BPAP in patients hospitalized for COPD exacerbation. *Respir Care* 2022;67(12):1517-1526.
- Burton BN, Trivedi S, Beletsky A, Mitchell A, Nasser E, Cazares U, et al. The influence of hospital urbanicity on mortality in patients with

- acute respiratory failure: a national cohort retrospective analysis. *Respir Care* 2021;66(12):1789-1796.
32. Eidman DB, Clauss CL, Kelly SA, J MR, McCollum S, K GC. Dexmedetomidine for sedation during pediatric noninvasive ventilation. *Respir Care* 2022;67(3):301-307.
  33. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. *Br Med J* 1950;2(4682):739-748.
  34. Berkson J. Limitations of the application of four-fold table analysis to hospital data. *Biometrics* 1946;2(3):47-53.
  35. Berkson J. Limitations of the application of four-fold table analysis to hospital data. *Int J Epidemiol* 2014;43(2):511-515.
  36. Neyman J. Statistics; servant of all sciences. *Science* 1955;122(3166):401-406.
  37. Hill G, Connelly J, Hebert R, Lindsay J, Millar W. Neyman's bias revisited. *J Clin Epidemiol* 2003;56(4):293-296.
  38. Swanson DM, Anderson CD, Betensky RA. Hypothesis tests for Neyman's bias in case-control studies. *J Appl Stat* 2018;45(11):1956-1977.
  39. Jager KJ, Tripepi G, Chesnaye NC, Dekker FW, Zoccali C, Stel VS. Where to look for the most frequent biases? *Nephrology (Carlton)* 2020;25(6):435-441.
  40. Girou E, Schortgen F, Delclaux C, Brun-Buisson C, Blot F, Lefort Y, et al. Association of noninvasive ventilation with nosocomial infections and survival in critically ill patients. *JAMA* 2000;284(18):2361-2367.
  41. Abdallah GA, Ferre A, Gros A, Simon C, Bruneel F, Marque-Juillet S, et al. Postextubation stridor in severe COVID-19. *Respir Care* 2022;67(6):638-646.
  42. Truumees M, Kendra M, Tonzola D, Chiu S, Cerrone F, Zimmerman D, et al. The impact of a home respiratory therapist to reduce 30-day readmission rates for exacerbation of COPD. *Respir Care* 2022;67(6):631-637.
  43. Moore L, Clark DE. The value of trauma registries. *Injury* 2008;39(6):686-695.
  44. Sears EH Jr, Hinton AC, Lopez-Pintado S, Lary CW, Zuckerman JB. Gaps in cystic fibrosis care are associated with reduced lung function in the US Cystic Fibrosis Foundation Patient Registry. *Ann Am Thorac Soc* 2023.
  45. Peetermans M, Guler I, Meersseman P, Wilmer A, Wauters J, Meyns B, et al. Impact of BMI on outcomes in respiratory ECMO: an ELSO registry study. *Intensive Care Med* 2023;49(1):37-49.
  46. Brown CA 3rd, Kaji AH, Fantegrossi A, Carlson JN, Apr MD, Kilgo RW, et al; National Emergency Airway Registry (NEAR) Investigators. Videolaryngoscopy compared to augmented direct laryngoscopy in adult emergency department tracheal intubations: a National Emergency Airway Registry (NEAR) study. *Acad Emerg Med* 2020;27(2):100-108.
  47. Bonczyk CS, Jelly CA, Freundlich RE. The blessing and the curse of the administrative database. *Ann Am Thorac Soc* 2020;17(2):174-175.
  48. Spitzer KA, Stefan MS, Priya A, Pack QR, Pekow PS, Lagu T, et al. Participation in pulmonary rehabilitation after hospitalization for chronic obstructive pulmonary disease among Medicare beneficiaries. *Ann Am Thorac Soc* 2019;16(1):99-106.
  49. Wickham RJ. Secondary analysis research. *J Adv Pract Oncol* 2019;10(4):395-400.
  50. Cooke CR, Iwashyna TJ. Using existing data to address important clinical questions in critical care. *Crit Care Med* 2013;41(3):886-896.
  51. Pena-Lopez Y, Ramirez-Estrada S, Serrano-Megias M, Lagunes L, Rello J, Group ES; EUVAE Study Group. Short-acting sedative-analgesic drugs protect against development of ventilator-associated events in children: secondary analysis of the EUVAE study. *Respir Care* 2021;66(5):798-805.
  52. Burr KL, Hinkson CR, Smith BJ, Roberts KJ, Strickland SL, Hoerr CA, et al. Factors associated with a positive view of respiratory care leadership. *Respir Care* 2022;67(10):1236-1245.
  53. Miller AG, Roberts KJ, Smith BJ, Burr KL, Hinkson CR, Hoerr CA, et al. Prevalence of burnout among respiratory therapists amidst the COVID-19 pandemic. *Respir Care* 2021;66(11):1639-1648.
  54. von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP; STROBE initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007;370(9596):1453-1457.