# Evidence-Based Respiratory Care

Dean R Hess

**Evidence-based medicine is the integration of individual clinical expertise with the best available research and the patient's values and expectations. The efficient approach to finding the best evidence is to identify systematic reviews or evidence-based clinical practice guidelines. Respiratory therapies that are supported by evidence include lung-protective ventilation and noninvasive ventilation for individuals with COPD. Evidence does not support postoperative incentive spirometry or intermittent mandatory ventilation. The principles of evidence-based**

**medicine are a valuable approach to respiratory care practice.** *Key words: clinical practice guideline; evidence-based medicine; meta-analysis; systematic review.* [Respir Care 0;0(0):1–•. © 0 Daedalus Enterprises]

## Introduction

The practice of respiratory care should be evidence-based. But this was not always the case, and evidence-based medicine (EBM) as we recognize it today has its roots in the late 20th century. The role of evidence-based respiratory care was recognized by the editorial board of RESPIRATORY CARE, who published a Journal Conference on the topic in 2000.[1] In 1988, Pierson[2] challenged us to use the scientific approach, the backbone of evidence-based practice, in respiratory care practice. Early in my career, during the 1970s, respiratory care was largely anecdotal. Many of us were shocked at the time when the Proceedings of the Conference on the Scientific Basis of Respiratory Therapy, commonly called the Sugarloaf Conference, were published in 1974.[3] The conference participants boldly stated that many accepted respiratory therapy practices were not supported by science.

The emergence of EBM can be traced to 1992,[4] with publication of the *Users' Guides to the Medical Literature*.[5] Today EBM has permeated all aspects of health care practice. Respiratory care practice demands evidence for the efficacy and safety of the treatments we use. Too often we rely on our experience, which is the lowest level of evidence. As the novelist Michael Crichton wrote in 1971, "*In my clinical experience* is a phrase that usually introduces a statement of rank, prejudice, or bias. The information that follows it cannot be checked, nor has it been subjected to any analysis other than some vague tally of the speaker's memory . . . the biases of eminent men are still biases."[6] In this paper, I cover the important implications of EBM for respiratory care practice. This is an extension of papers I have previously published on this subject.[7,8]

Dr Hess is Managing Editor of RESPIRATORY CARE and is affiliated with the Massachusetts General Hospital and Northeastern University, Boston, Massachusetts.

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

## What Is Evidence-Based Medicine?

EBM is the integration of individual clinical expertise with the best available evidence from systematic research and the patient's values and expectations.[9] The best evidence is extrapolated to the patient's unique pathophysiology. EBM does not devalue clinical skills and clinical judgment. Rather, EBM demands a skilled clinician to inform its judicious application. The practice of EBM requires us to apply the evidence to the right patient, at the right time, in the right place, at the right dose, and using the right resources.

Clinical evidence comes from real clinical research among intact patients. Bench studies, simulations, animal studies, and other types of physiologic studies can support human studies. However, we should recognize that these are low-levels of evidence, and we need to be cautious about extrapolating these data to patient care. The best clinical evidence is not static, but changes when new and better evidence becomes available.

The practice of EBM does not discount patient values and expectations. For example, a compelling body of high-level clinical evidence supports the use of noninvasive ventilation (NIV) in patients with COPD exacerbation, yet some patients prefer not to receive NIV. As another example, clinician bias might suggest the use of a pressurized metered-dose inhaler (pMDI) rather than a nebulizer for delivery of an inhaled bronchodilator; however, evidence suggests that that the 2 approaches yield similar outcomes. If a patient prefers to use a nebulizer rather than a pMDI, the patient's choice should be respected.

## Hierarchy of Evidence

All evidence is not created equal, which has led to the concept of a hierarchy of evidence (Table 1).[5] The highest level of available evidence is used when making clinical

Table 1. Hierarchy of Evidence

| |
|---|
| *N* of 1 randomized trial |
| Systematic reviews of randomized trials |
| Single randomized trial |
| Systematic reviews of observational studies addressing patient-important outcomes |
| Single observational study addressing a patient-important outcome |
| Physiologic studies |
| Unsystematic clinical observations |

Table 2.    Characteristics of High-Level Studies

| Characteristic | Description |
|---|---|
| Prospective | A prospective investigation is a study in which subjects are enrolled to evaluate the outcome in question. |
| Randomization | Allocation of individuals to groups by chance. This differs from systematic allocation (eg, even and odd days of the month) or allocation at the convenience or discretion of the investigator. |
| Blind (or masked) | The participant of interest is unaware of whether subjects have been assigned to the experimental or control group. Subjects, clinicians, those monitoring outcomes, assessors of outcomes, data analysts, and those writing the paper can all be blinded. |
| Placebo | Interventions without biologically active attributes. |
| Concealment | Randomization is concealed if the person who is making the decision about enrolling a subject is unaware of whether the next subject enrolled will be entered in the treatment or control group. |
| Parallel design | With a parallel study design, subjects are randomly assigned to the treatment or control group, an intervention is applied, and the outcome is identified for each subject. This is different from a crossover study, in which subjects receive both the treatment and control intervention. |

decisions. Note that evidence always exists, thought it may be of a low level. The best available evidence may be an unsystematic clinical observation or a generalization from a physiologic study. Note that experience is low-level evidence, but opinion is not evidence. Accordingly, we should not base clinical decisions on opinion, even if that opinion comes from a respected source. Thus, it is important to separate experience from opinion.

Randomization is an important attribute of higher levels of evidence. The highest level of evidence is an *N*-of-1 randomized controlled trial (RCT). In the *N*-of-1 RCT, a patient undergoes pairs of treatment periods in which the patient receives a target treatment in one period and a placebo or alternative in the other.[10-14] The order of the target treatment and control is randomized, and quantitative ratings are made for each treatment. The *N*-of-1 RCT continues until both the patient and clinician conclude that there is, or is not, benefit from the target treatment.

There are some therapies for which there has not been a randomized trial, and for which one might argue that a randomized trial is either unethical or unnecessary. For example, it is unnecessary that a randomized trial be conducted to study the survival benefit of mechanical ventilation in patients with apnea. Randomized trials of the best approach to mechanical ventilation are necessary, but a control group of patients with apnea who do not receive mechanical ventilation is clearly unethical. In a clever paper arguing this point, Smith et al[15] show that an RCT is not necessary to determine whether parachutes are effective in preventing major trauma related to gravitational challenge.

In respiratory care, the evidence to support many therapies is weak. Just because a therapy is unproven does not mean that it is wrong. But it also does not mean that it is right. And just because a therapy is new does not necessarily mean that it is better. Too often new therapies are quickly embraced, only to learn later that they are ineffective. Intermittent mandatory ventilation[16] and high-frequency oscillatory ventilation are examples.[17,18]

High-level studies are prospective, randomized, blinded, placebo-controlled, concealed allocation, parallel design, and assess patient-important outcomes (Table 2). A patient-important outcome like mortality is favored over a physiologic outcome like an improvement in arterial blood gases. There are a number of examples in which an improvement in physiologic outcomes does not lead to patient-important outcomes. High tidal volumes in patients with ARDS improve arterial blood gases, but mortality is lower for smaller tidal volumes.[19] For patients with ARDS, inhaled nitric oxide improves $P_{aO_2}$ but not mortality.[20] High-frequency oscillatory ventilation improves $P_{aO_2}$ but does not improve survival in patients with ARDS.[17,18] Aggressive recruitment maneuvers might improve gas exchange but contribute to a higher likelihood of mortality.[21-23] As clinicians, we should be wary of targeting therapies solely to improve physiologic parameters such as $P_{aO_2}$. The injury attributed to a short-term improvement in a physiologic measure like arterial blood gases could contribute to harm in the long term. Treatments like aggressive lung recruitment strategies are appealing on the surface but might not be supported by best evidence.

## Finding the Evidence

### Asking the Question

Finding the best evidence begins with a clearly articulated question. Often this is in the form of PICO: patient or problem, intervention, comparison, and outcomes.[5,24]

### Source Literature

Primary studies supply the source evidence, but the information they contain requires critical assessment before application to clinical problems. Methods for conducting an online search to identify source literature are described in detail elsewhere.[25-27] A commonly used search engine is Google (http://www.google.com). Conducting a search

Table 3.    Comparison of Narrative and Systematic Reviews

| Characteristic | Narrative Review | Systematic Review |
|---|---|---|
| Topic or question | Broad overview without a specific question | Narrow scope, with a specific question to be answered |
| Literature sources | Limited sources, typically not specified | Wide variety of specifically named databases |
| Literature selection | Unspecified, potentially biased | Predetermined specific criteria for selection of papers |
| Literature appraisal | Unspecified, variable, and potentially biased | Critical review using specific criteria |
| Literature synthesis | Qualitative summary | Quantitative summary |
| Recommendations | Opinion, potentially biased, not evidence-based | Evidence-based |

From Reference 28.

using Google can be overwhelming and usually involves a considerable amount of time filtering through the search results. Much of the information that is found may be irrelevant, and the validity of the information may be outdated or incorrect. Much of what is found in a simple Google search has not been subjected to peer review. The Google Scholar search engine (http://scholar.google.com) searches scholarly information such as articles, dissertations, books, abstracts, and full text from publishers. Google Scholar ranks material and links to other documents that cite an important item you have identified.

PubMed (www.pubmed.gov) is easy to use but can be overwhelming because of the large number of articles (ie, > 30 million). Publicly available online since 1996, PubMed is maintained by the National Center for Biotechnology Information of the National Library of Medicine (NLM), located at the National Institutes of Health. PubMed searches across the NLM resources: MEDLINE, PubMed Central, and Bookshelf. MEDLINE is the largest component of PubMed and consists of citations from journals selected for inclusion. PubMed Central is a full-text archive that includes articles collected for archiving in compliance with funding policies. Bookshelf is a full-text archive of books, reports, and databases. The PubMed User Guide is a useful primer on the use of this database (https://pubmed.ncbi.nlm.nih.gov/help/).

To search PubMed, enter search terms in the query box. The Advanced search feature allows content to be limited, such as to a specific journal, author, or publication dates. Search filters can be selected from the left of the search page, and the Related Articles feature in PubMed uses a word-weighted algorithm to compare words from the Title and Abstract of each citation. A comprehensive PubMed search for purposes of identifying the best evidence is overwhelming. Few individuals have the time to read all of the papers identified in a PubMed search, assess the validity of the evidence, and develop strategies to incorporate such evidence into everyday practice.

CINAHL (Cumulative Index to Nursing and Allied Health Literature; www.cinahl.com) is a nursing and allied health database. EMBASE (www.embase.com) is a large European database similar to MEDLINE in scope and content, with

strengths in drugs and allied health disciplines. Up to 70% of citations in EMBASE are not included in MEDLINE. Ovid (www.ovid.com) provides medical information services to individuals in medical schools, hospitals, and academic institutions. Ovid provides access to a large selection of databases, including MEDLINE, CINAHL, and the other bibliographic databases. Most journals, including RESPIRATORY CARE, provide searching of journal contents by keyword or author directly from their home pages. The Web of Science (https://clarivate.com/webofsciencegroup/solutions/web-of-science/) is a global citation database; search results on this database highlight the number of times papers have been cited.

A useful strategy is to scan journal contents monthly for relevant articles. This is easily accomplished by subscribing to the email alerts provided free of charge by journals. It is also useful to use a citation manager to organize full-text articles. Examples include Papers, EndNote, Reference Manager, RefWorks, Mendeley, and others. Zotero (https://www.zotero.org) is a free, easy-to-use tool to collect, organize, cite, and share references. Most citation management software allows full text to be imported and allows one to conduct the online search from within the program and download articles directly to the citation manager.

All clinicians should be able to perform a basic literature search to inform their practice. However, most clinicians need help to conduct a comprehensive search. As print has given way to digital media, the expertise of medical librarians has shifted to online searches. The assistance of a medical librarian is essential when conducting a literature review to support a systematic review, meta-analysis, or clinical practice guideline (CPG).

## Systematic Reviews

Narrative and systematic reviews are compared in Table 3. Narrative reviews are often limited by incomplete searches of the literature, intentional or unintentional bias by the authors, and failure to account for the quality of individual publications. Narrative reviews typically fail to deal effectively with studies having conflicting results, as they are written by experts who rely on their own experience and
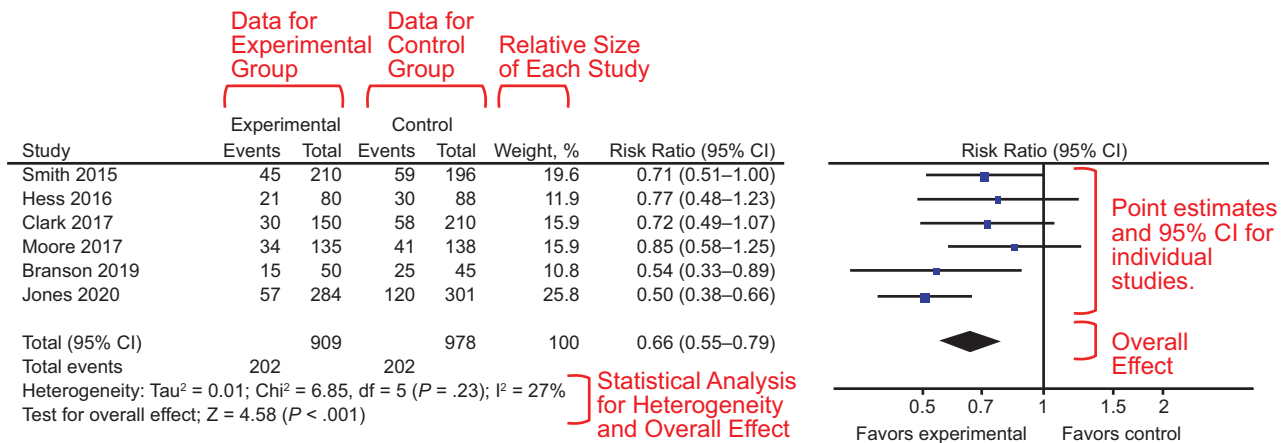
Fig. 1. Anatomy of a forest plot. The outcome is placed on the *x* axis and the vertical line at 1.0 represents the no-effect line. Confidence intervals (CIs) that cross this vertical line indicate that the groups have equal risk of the outcome of interest. The square boxes are the point estimates for each study and are often presented as boxes of different sizes, which correspond to the weight given to the study. A larger box reflects a higher weight assigned to that study. The cumulative treatment effect is represented by the diamond symbol at the bottom of the diagram. The center of the diamond represents the point estimate of the combined result, and the width of the diamond represents the 95% CI of the point estimate.

expertise rather than on a critical assessment of all available evidence.[28]

A systematic review is a summary of the literature that uses explicit methods, is based on a thorough literature search, performs a critical appraisal of individual studies, and uses statistical techniques to combine study results (ie, meta-analysis).[28-31] In a systematic review, the primary evidence is identified and appraised. Unlike the narrative review, a systematic review uses explicit methods. A systematic review details the methods by which papers were identified in the literature, it uses predetermined criteria for selection of papers to be included in the review, and it critically assesses the evidence and bases the review on the strength of that evidence.

Several sources can be searched for systematic reviews. PubMed can be searched using "systematic reviews" as article type. OVID can be searched using the specific databases "EBM Reviews." The Cochrane Database (www.cochrane.org) is a rich source of systematic reviews, including many related to respiratory care. Several syntheses of Cochrane Reviews have been published in RESPIRATORY CARE.[32-34] Systematic reviews and CPGs may become outdated and should be supplemented by recent RCTs published after the publication date of the review or CPG.

**Meta-Analysis**

Meta-analysis is a statistical method that combines the results of several independent studies.[28,29,31,35] Since it is based on a literature review, the meta-analysis is observational rather than experimental; ie, it is not original research. The person conducting the meta-analysis has limited control over the availability of studies or the information reported in the individual studies. The studies included in the meta-analysis should be comparable, but the degree of comparability is subjective and is determined by the person conducting the meta-analysis. Included studies should be identified from a comprehensive review of the literature, and unpublished data should be included to reduce the risk of publication bias. A meta-analysis uses statistical methods to combine the results of several studies into a single pooled metric. A meta-analysis of RCTs is a higher level of evidence than a single RCT.

In a meta-analysis, if the odds ratio, risk ratio, or relative risk exceeds 1, the likelihood of the outcome is greater in the treatment group. On the other hand, if it is below 1, the outcome is less likely in the treatment group. If the value is close to 1, the outcomes in the treatment and control groups are similar. If the CI overlaps 1, the results are not significantly different from one another, whereas a wider CI indicates a less precise treatment effect, which is often due to smaller sample sizes.

The results of a meta-analysis are displayed as a forest plot (Fig. 1). Following are some common metrics that may be reported in a meta-analysis:

- Event Rate: proportion of subjects in a group in whom an event is observed. Control event rate and experimental event rate are determined for control and experimental groups of subjects, respectively.

- Relative Risk (Risk Ratio): ratio of the risk of an event among an experimental group to the risk among the control group. A relative risk < 1 means benefit, a relative risk > 1 means harm, and a relative risk of 1 means the intervention has no effect.

- Relative Risk Reduction: estimate of the proportion of baseline risk that is removed by the therapy.

| Study | HFNC Events | HFNC Total | Control Events | Control Total | Weight, % | Risk Ratio M-H, Random, 95% CI | Risk Ratio M-H, Random, 95% CI |
|---|---|---|---|---|---|---|---|
| Azoulay 2017 | 77 | 187 | 263 | 649 | 26.2 | 1.02 (0.84–1.23) | |
| Azoulay 2018 | 150 | 388 | 170 | 388 | 29.6 | 0.88 (0.75–1.04) | |
| Coudroy 2016 | 21 | 60 | 30 | 55 | 9.9 | 0.64 (0.42–0.98) | |
| Frat 2016 | 8 | 26 | 13 | 30 | 4.0 | 0.71 (0.35–1.44) | |
| Lemiale 2015 | 4 | 52 | 2 | 48 | 0.8 | 1.85 (0.35–9.63) | |
| Lemiale 2016 | 40 | 90 | 48 | 90 | 16.1 | 0.83 (0.62–1.13) | |
| Roca 2015 | 13 | 22 | 16 | 18 | 11.4 | 0.66 (0.45–0.98) | |
| Tu 2017 | 4 | 20 | 9 | 18 | 2.1 | 0.40 (0.15–1.08) | |
| Total (95% CI) | | 845 | | 1296 | 100 | 0.83 (0.72--0.97) | |
| Total events | 317 | | 551 | | | | |

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 9.95$, $df = 7$ ($P = .19$); $I^2 = 30\%$

Test for overall effect; $Z = 2.39$ ($P = .02$)
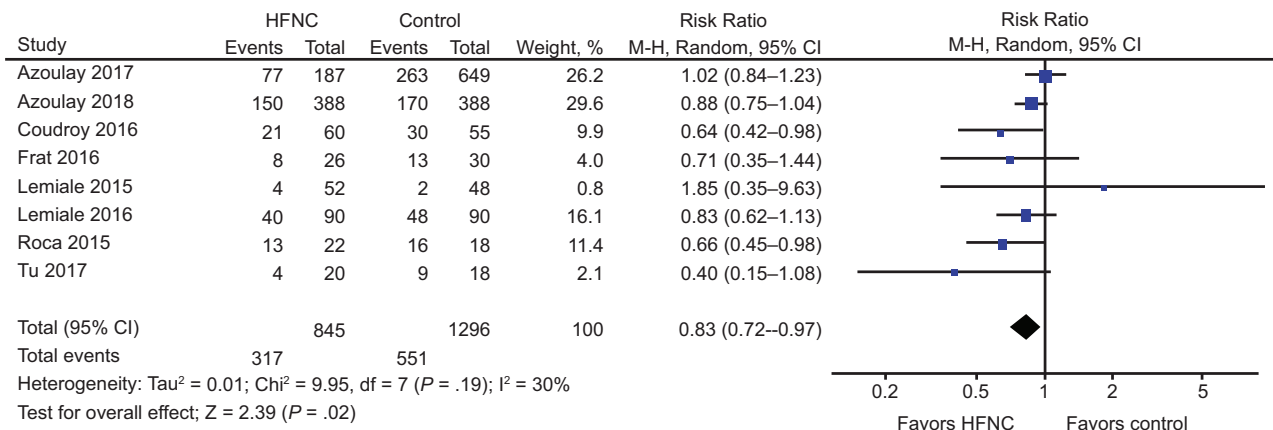
Favors HFNC    Favors control

Fig. 2. Forest plot for intubation rate associated with high-flow nasal cannula (HFNC) in immunocompromised subjects with acute respiratory failure. From Reference 37.

- Absolute Risk Reduction: difference in the experimental event rate and the control event rate.
- Number Needed to Treat (NNT): number of patients who need to be treated to avoid a bad outcome; the inverse of the absolute risk reduction.
- Odds Ratio: ratio of the odds of an event in an exposed group to the odds of the same event in a group that is not exposed.
- Heterogeneity ($I^2$): The percentage of variation across studies that is due to heterogeneity rather than chance. $I^2 < 30\%$ is low heterogeneity, 30–60% is medium heterogeneity, and > 60% is high heterogeneity.[36]

Kang et al[37] conducted a meta-analysis of the effect of high-flow nasal cannula (HFNC) in immunocompromised subjects with acute respiratory failure. Figure 2 is a forest plot for intubation rate associated with HFNC in immunocompromised subjects with acute respiratory failure. The relative risk (risk ratio) is 0.83, meaning that those who received HFNC had 0.83 times the risk of intubation as those who did not receive HFNC. The CI does not cross 1 (the vertical line of no difference), so the difference is significant ($P = .02$). The heterogeneity among studies is acceptable ($I^2 = 30\%$). For the 8 studies included in the meta-analysis, the intubation rate in subjects received HFNC was 317 of 845 (0.375), compared to an intubation rate of 551 of 1,296 (0.425) in those in the control group. The absolute risk reduction is 0.05 (ie, 0.425 − 0.375), and the NNT (1/0.05) is 20. This is interpreted to mean that, for every 20 patients with acute respiratory failure secondary to immunocompromise treated with HFNC, 1 intubation is avoided.

A few caveats are important when reading a meta-analysis. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines should be followed (http://www.prisma-statement.org). The quality of the meta-analysis is only as good as the quality of the included studies. Combining studies with a high risk of bias does not result in a meta-analysis with a low risk of bias.[35] A meta-analysis requires a mature evidence base. Immature evidence does not merit a meta-analysis. A meta-analysis including only a few poorly done studies does not advance knowledge of a subject.

**Clinical Practice Guidelines**

Evidence-based CPGs ask relevant questions, systematically search the literature using explicit methodology, grade the level of the evidence, make recommendations, and grade the recommendations on the basis of the strength of the evidence. The recommendations of the evidence-based guidelines are supported by evidence, and the level of evidence is unambiguous and defensible. CPGs are often accompanied by systematic reviews and meta-analysis. They are often supported by professional organizations, including the American Association for Respiratory Care.[38]

The GRADE (Grades of Recommendation, Assessment, Development, and Evaluation) approach is commonly used to evaluate the quality of supporting evidence and the strength of recommendations in health care.[39-41] The GRADE system provides a detailed stepwise process that defines the quality of the available evidence in the development of recommendations. The value of GRADE is not that it eliminates judgments or disagreements about evidence and recommendations, but that it makes them transparent.

CPGs are trustworthy only if the recommendations accurately reflect the underlying evidence about benefits and harms to individual patients.[42] This requires a rigorous process for assembling, evaluating, and summarizing the evidence. Often this is the result of a systematic review. Another requirement is the process used to decide the recommended strategies that best offer a favorable balance of harms and benefits. The Institute of Medicine developed
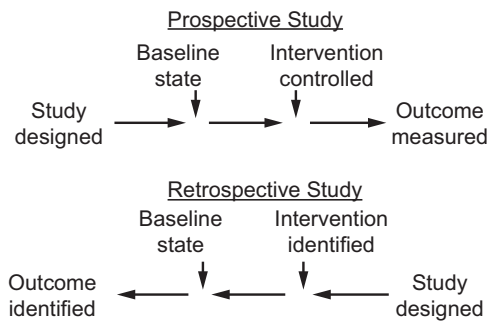
Fig. 3. Schema for prospective and retrospective studies. With a prospective study, the study is designed before data collection begins. A retrospective study uses available data that was recorded for reasons other than the study.

the Appraisal of Guidelines, Research and Evaluation (AGREE) system to set standards for rating the quality of the process of guideline development.

CPGs using the GRADE methodology have been published related to mechanical ventilation for patients with ARDS,[43] NIV for acute respiratory failure,[44] and NIV for stable hypercapnic COPD.[45] However, strict adherence to the GRADE methodology has had unintended consequences, such as a long time from conception to completion and high costs. A modified Delphi process based on expert consensus might result in similar recommendations.[46] The AARC has adopted an approach to CPGs using a modification of the RAND/UCLA Appropriateness Method.[38] The first CPGs using this methodology were published early in 2021.[47,48]

### Social Media

The role of social media in finding the best evidence is receiving increasing attention. Social media should not be used as the primary source of evidence. Many of the claims posted on social media have not been subjected to the scientific method or peer review. Although patients and clinicians sometimes use social media to solicit input for clinical challenges, caution is urged for this approach because many of the responses are not based on high levels of evidence. Nevertheless, social media platforms can also facilitate communication, interactions, and connections among health care professionals. Many journals, including RESPIRATORY CARE, use social media to highlight published papers. Many investigators use social media to share their latest publications. Although this is largely celebratory, it also has the benefit of rapidly informing others of new research findings.[49,50] An interesting approach, a Twitter journal club, has been described to provide post-publication peer review of articles and helps Twitter users build a network of engaged people with similar clinical interests.[51]

Facebook is increasingly used as a platform to host virtual presentations, including journal clubs.

### Study Types

#### Prospective Versus Retrospective Studies

The basic schema for prospective and retrospective studies are shown in Figure 3. With a prospective design, the study is designed before data collection begins. Conversely, a retrospective study uses available data that was recorded for reasons other than the study. There are many sources of bias in a retrospective study. Most important, the investigator has no control over the veracity of the data. The study might ask subjects to recall events from the past, which introduces bias if subjects cannot recall correctly. Performance bias is an issue. Imagine that a study involves subjects' use of a pMDI; the investigator cannot know if the subjects used the device correctly. Missing data are a common issue with retrospective studies.

#### Randomized Controlled Trials

A trial is an experiment. Randomization is the process by which allocation of subjects to treatment groups is done by chance, without the ability to predict who is in each group. Control subjects are individuals who do not receive the experimental treatment but are otherwise similar to those in the experimental group. The RCT is the most important type of study for medical interventions in clinical care.[52] A clinical trial is a controlled experiment having a clinical event as the outcome measure. It is done in a clinical setting and enrolls subjects having a particular disease or condition. In a randomized clinical trial, subjects are randomly assigned to groups that compare different treatments. An RCT is interventional, prospective, randomized, blinded if possible, and has well-matched treatment and control groups.

Most RCTs are time-consuming and expensive. They typically have strict inclusion and exclusion criteria, which is important methodology but limits generalizability. Due to the length of the study, dropouts can occur and an intention-to-treat analysis is important. The RCT should evaluate a patient-important outcome. In critical care, this is often mortality.

#### Observational Studies

In an observational study, the investigator observes the effect of an intervention without changing who is or is not exposed to it. Cross-sectional studies sample a population at a given point in time; they are commonly used to determine a case rate (prevalence) or values for a test. A cohort study compares subjects exposed to a variable to those who
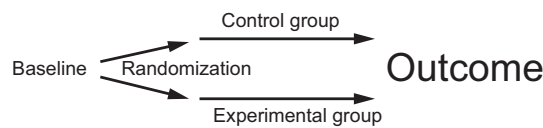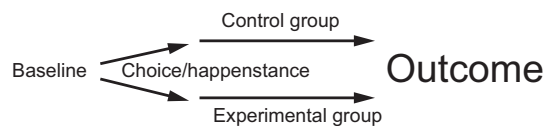
## Randomized Controlled Trial



## Observational Study



Fig. 4. Schema for randomized controlled trials and observational studies. With the randomized controlled trial, subjects are assigned to groups as the result of randomization. With an observational study, subjects are assigned due to chance or happenstance.

are not exposed. A case control study compares subjects with a problem (ie, the cases) and a similar group without the problem (ie, the controls). The schema for RCTs and observational studies is shown in Figure 4. Observational studies, particularly retrospective observational studies, can be used to generate hypotheses, but seldom will they alone change practice.

### Case Series

A case series describes characteristics of a population. The first description of ARDS in 1967 was a case series.[53] The Large Observational Study to Understand the Global Impact of Severe Acute Respiratory Failure (LUNG SAFE) study, a large case series, was published in 2016 and reported a 10.6% prevalence of ARDS for ICU admissions and reported that ARDS was underrecognized, undertreated, and associated with a high mortality rate.[54] A case report describes the care of a single patient. A case report is low-level evidence, and many journals, including RESPIRATORY CARE, no longer publish them.

### High-Quality Studies

A high-quality study uses methodology to reduce bias such as being prospective, randomized assignment to the control and experimental groups, blinding of subjects and investigators, placebo control, concealed allocation before group assignment (ie, the investigator does not know the assignment until after subject enrollment), and an appropriate sample size. A multi center study increases the generalizability of the findings. Using a patient-important outcome increases the value of the study. It is important to appreciate that there is no perfect study; almost all peer-reviewed studies have flaws. Despite these flaws, most also have value.

### Considerations When Reading a Published Paper

### The Underpowered Study Dilemma

Important statistical attributes of a study are its risk of a type-1 (alpha) error and a type-2 (beta) error. A type-1 error is the risk of stating that a difference is present when it is not (false positive). The $P$ value is the probability for a type-1 error. Traditionally, the $P$ value is set a priori at < .05, meaning the risk of a type-1 error is < 5% (ie, 1 in 20). Note that the $P$ value can never be 0; there is always some risk of a type-1 error, but it is unlikely if the $P$ value is very small. A type-2 error is the risk of stating that there is no difference between groups when there is (false negative). Traditionally, beta is set at 0.2. The power of a study (ie, 1 − beta) is the likelihood of detecting a difference between groups if one exists.

During the study design, investigators should work with a statistician to determine the sample size necessary to minimize the risk of a beta error; in other words, to ensure that the study is appropriately powered.[55] The sample size estimate is determined by the desired difference between groups and the expected variability. A smaller difference or greater variability will require a larger sample size. Unfortunately, sometimes the difference is smaller than anticipated, or the variability is greater, resulting in a difference that may be real but is not statistically significant. This is an underpowered study.

Consider the study by Combes et al[56] that evaluated the effect of extracorporeal membrane oxygenation in (ECMO) subjects with severe ARDS. The authors expected a mortality at 60 d of 60% in the group that did not receive ECMO and 40% in the group that received ECMO. They calculated that, for 80% power and an alpha of 5%, a sample of 331 subjects would be required. In addition, stopping rules were determined a priori, and the study was stopped after 249 subjects were enrolled. The 60-d mortality was 35% in the ECMO group and 46% in the control group. Many clinicians would consider this reduction in mortality to be clinically important. However, this did not reach statistical significance ($P = .09$) because the study was underpowered. This has resulted in confusion among clinicians. Those who favor the use of ECMO will argue that there was an 11% absolute risk reduction in mortality with the use of ECMO, but those opposed will point out that the result was not statistically significant.

The assumptions used for a power analysis (sample size determination) can be determined by prior knowledge, such as a review of data from published studies, or from a pilot study. A conundrum in original research of novel technology is lack of data upon which to calculate power. A pilot study is usually a small study designed to power a larger study. A pilot study is determined a priori and approved by an institutional review board as such. Sometimes a study is stopped early due to slow enrollment or for other reasons. This can

result in an underpowered study that investigators sometimes incorrectly name a pilot study. It is important to distinguish an underpowered study from a pilot study.[57]

Sometimes it is appropriate to stop a study early. The ARDS Network study of higher versus lower tidal volumes was powered for 1,000 subjects but was stopped early.[19] After enrollment of 861 subjects, there was a significantly lower mortality in the subjects randomized to a smaller tidal volume ($P = .007$). With a NNT of 11, continuing enrollment to the target would be unethical because it would expose subjects to a risk of death.

### An RCT Allowing Crossovers Is Difficult to Interpret

The design of a RCT comparing treatments that cannot be blinded might allow subjects to crossover to the alternative therapy. This allows the alternative treatment if the subject meets failure criteria for the initially assigned treatment. This is desirable for ethical reasons due to equipoise between the treatment strategies. Related to respiratory care, this has been commonly applied to studies of HFNC and NIV. Note that a crossover design is only applicable when an outcome like intubation rate is used. A crossover design is not possible in a study like the ARDS Network studies, where mortality was the outcome. Intention-to-treat analysis evaluates the results of an RCT in which subjects are analyzed according to the group to which they were originally assigned, regardless of the treatment they received. The alternative approach, which is not recommended, is a per-protocol approach, where subjects are analyzed according to the therapy received.[58]

Crossovers and intention-to-treatment analysis can create confusion for the reader. Consider the paper by Doshi et al.[59] This was a multi-center RCT of adults presenting to the emergency department with respiratory failure requiring NIV. Patients were randomly assigned to receive NIV or HFNC. The intubation rate was higher for those assigned NIV, but the therapy failure rate was higher for those assigned HFNC. These results seem contradictory until the study design is examined. Crossover was allowed as a risk mitigation to support deferment of informed consent. In the HFNC group, the failure rate was 35%, but 85% of the failures were crossed over to NIV and 87% of those avoided intubations with the use of NIV. In the NIV group, only 20% failed, only 35% crossed over, and only 35% avoided intubations with the use of HFNC. An intention-to-treat analysis provided a favorable outcome for HFNC due to the beneficial effect of NIV in HFNC failures.

The study by Doshi et al[59] was designed a priori as a noninferiority study. The judgment of noninferiority should be based on 3 prerequisites: the new treatment (in this case HFNC) demonstrates therapeutic noninferiority to the standard treatment (in this case NIV), the new treatment (HFNC) exhibits therapeutic efficacy in a placebo-controlled trial if such a trial were performed, and the new treatment (HFNC)

offers benefits such as safety, tolerability, convenience, or cost.[60] Key to a noninferiority study is estimation of the non-inferiority margin, which is the maximum difference in outcomes between groups to allow a claim of noninferiority. Statistical reasoning and clinical judgment are used to choose this margin. When reading a published noninferiority study, it is important to make a judgment about the noninferiority margin selected by the investigators. In the case of Doshi et al,[59] the prespecified noninferiority margins are 15% and 20% for differences in intubation and failure rates, respectively. One might challenge whether this margin is too large, resulting in a type-2 error when stating that HFNC is noninferior to NIV.

Another example of a noninferiority study compared lower (0–6 cm $H_2O$) versus higher (8 cm $H_2O$) PEEP in subjects who did not have ARDS.[61] The noninferiority margin was set a priori at 1.6 ventilator-free days. The reader can judge whether that noninferiority margin is too high or too low. At day 28, subjects in the lower PEEP group had a median of 18 ventilator-free days and subjects in the higher PEEP group had a median of 17 ventilator-free days, which was within the predetermined noninferiority margin.

### A Not-Significant *P* Value Is Not a Trend

When the *P* value is not significant but near .05, some authors call this a trend. However, such a designation is statistically incorrect and without meaning. A *P* value near .05 but not significant does not reflect a trend but rather that the results are not significant and the study may be underpowered.

### Correlation Is Not Agreement

In respiratory care practice, we are often interested in how well one measurement compares to another. If the end-tidal $P_{CO_2}$ is not similar to the arterial $P_{CO_2}$, it is commonly said that the measures do not correlate. However, what we mean to say is that the measurements do not agree. Statistical assessment of agreement was first described by Bland and Altman in 1986.[62] The 3 important statistics when assessing agreement are:

- Bias: the mean difference between measurements
- Precision: the standard deviation (SD) of the differences
- Limits of agreement: bias ± 2 SD (or more precisely ± 1.96 SD)

It is now accepted practice that comparison between measurements should be reported using the method of Bland and Altman; reporting correlation is not appropriate.

The approach of Bland and Altman is illustrated in Figure 5. This is a hypothetical comparison of 50 simultaneous measurements of oxygen saturation from arterial blood

Mean diffrence: 0 (Bias)
SD of diffrence: 2.3% (Precision)
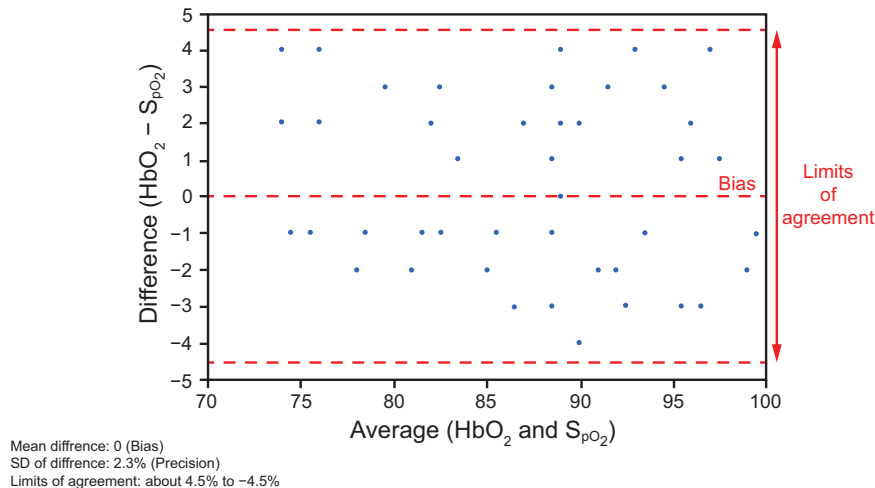Limits of agreement: about 4.5% to −4.5%

Fig. 5. Bland-Altman plot for a hypothetical evaluation of the accuracy of oxygen saturation measured with a pulse oximeter ($S_{pO_2}$) and from arterial blood ($HbO_2$). These are not real measurements; they are used only to illustrate a Bland-Altman plot.

($HbO_2$) and from a pulse oximeter ($S_{pO_2}$). Plotted on the $y$ axis are the differences between measurements and plotted on the $x$ axis are the average of the measurements; the average is used because it is not possible to know which measurement is the true value. The bias is zero and the precision is 2.3%. The limits of agreement are approximately +4.5% to –4.5%. This hypothetical example suggests that the $S_{pO_2}$ could be 4.5% greater than the $HbO_2$ or 4.5% less than the $HbO_2$.

### A Poorly Conducted Survey With a Low Response Rate Is Not Generalizable

Surveys can advance our understanding of practice, attitudes, and knowledge. However, like any study design, adherence to accepted methodology is important. The survey instrument must be carefully designed and validated. It is important to design a plan to maximize response rate. A survey with a poor response (eg, < 40%) is not generalizable.[63] Ideally, efforts should be undertaken to ensure that the responses of respondents are representative of those who do not respond.

### Use of Parametric Analysis When the Data Are Nonparametric

Without the input of a statistician, there is a tendency for investigators to summarize their data as means, SDs, and $t$ tests. However, this assumes that the data are continuous and normally distributed. Ordinal data, such as a dyspnea scores, are not parametric and should be analyzed with nonparametric analysis (eg, median, interquartile range, Mann-Whitney test).[64] Nonparametric analysis should also be used if continuous data are not normally distributed. A quick check on whether data are normally distributed is to compare the

mean and the SD. The mean $\pm$ 2 SDs should not overlap zero. For example, if the mean age of a group of subjects is 60 y with a standard deviation of 40 y, then the mean $\pm$ 2 SDs overlaps zero; these data should not be analyzed with parametric statistics. As another example, if the mean $S_{pO_2}$ for a group of subjects is 97% with a standard deviation of 5%, the mean $\pm$ 2 SDs exceeds an $S_{pO_2}$ of 100%, which is physiologically impossible.

### Use of Sophisticated Analysis Confuses Readers

Statistical approaches have advanced in recent years. Investigators correctly consult the help of biostatisticians to rigorously analyze their data and suggest the meaning of that analysis. Many readers of the literature do not have the necessary training and sophistication to independently vet this analysis. It is incumbent on the authors to explain their selection of statistical tests and how that informs the meaning to study findings.

### Statistical Significance Versus Clinical Importance

It is important to separate statistical significance and clinical importance. An underpowered study might report a clinically important finding that is not statistically significant. On the other hand, a large study might report a small difference that is statistically significant but is not clinically important. This again highlights the importance of determining an appropriate sample size when planning a study.

### Questions to Ask When Reading a Paper Reporting the Results of an RCT

There are several questions that should be asked when a convincing RCT is published:

- How large is the difference between groups?
- Are the results plausible?
- Are the results generalizable?
- Is there a proposed mechanism?
- Are there supporting RCTs or lower levels of evidence?

The answers to these questions will inform whether the findings are incorporated into practice.

It is seldom that a single study will change practice. Only with subsequent validation should a study change practice. This is particularly true with a study performed at a single center. Sometimes a study is published and the results are rapidly translated into practice, but subsequent studies do not confirm the initially published study. An example is early goal-directed therapy. A single-center study published in the *New England Journal of Medicine* in 2001 reported that early goal-directed therapy provided significant benefits with respect to outcome in subjects with severe sepsis and septic shock.[65] The results were rapidly adopted worldwide and commonly referred to by the first author's name—the Rivers protocol. However, subsequent studies published in 2014 did not report a benefit for early goal-directed therapy.[66,67] Finally, a paper published in 2015 reported that, in subjects with septic shock who were identified early and received intravenous antibiotics and adequate fluid resuscitation, hemodynamic management according to a strict early goal-directed therapy protocol did not lead to an improvement in outcome.[68]

## Therapy That Evidence Supports

### Lung-Protective Ventilation

In the ARDS Network study,[19] 861 subjects with ARDS were randomly assigned to mechanical ventilation with a tidal volume of 12 mL/kg or 6 mL/kg on the basis of predicted body weight (not actual body weight). In addition, for the group that received the smaller tidal volume, plateau pressure was targeted at $\leq 30$ cm $H_2O$. Mortality for the control group (ie, with tidal volume 12 mL/kg) was 39.8%, and mortality for the experimental group (ie, with tidal volume 6 mL/kg) was 31%. The relative risk of mortality was lower for the 6 mL/kg group (0.79), with a relative risk reduction of 21% compared to the 12 mL/kg group. There was an absolute risk reduction for mortality of 8.8%, resulting in an NNT of 11 patients. A normal tidal volume is $\sim 6$ mL/kg predicted body weight.[69] Thus, the ARDS Network study might be described as a comparison of a normal tidal volume to a twice normal tidal volume.

The results of the ARDS Network are supported by lower levels of evidence.[70] The ARDS Network findings have also been confirmed in subsequent observational studies.[71,72] The tenets of lung-protective ventilation, volume and pressure limitation, are recommended in CPGs for mechanically ventilated patients with ARDS.[43] Unfortunately, lung-protective ventilation in patients with ARDS remains underutilized.[54]

In an RCT evaluating tidal volume selection during major surgery,[73] subjects were randomized to receive a tidal volume of 6 mL/kg or 10 mL/kg predicted body weight. Pulmonary complications occurred in 38% of the lower tidal volume group compared with 39% in the higher tidal volume group ($P = .64$). In an RCT assessing the effect of tidal volume in subjects without ARDS,[74] those assigned to a lower tidal volume started at a tidal volume of 6 mL/kg predicted body weight, whereas those assigned to an intermediate tidal volume group started at a tidal volume of 10 mL/kg. Both groups had a median of 21 ventilator-free days ($P = .71$).

The results of these studies might be interpreted to say that a tidal volume of 10 mL/kg might be used in patients without ARDS.[73,74] But the results could also be interpreted as there being no harm from a tidal volume of 6 mL/kg. Currently available evidence does not support a tidal volume > 10 mL/kg in mechanically ventilated patients without ARDS. It might be reasonable to target a tidal volume of 8 mL/kg in patients without ARDS.[75] It is also important to recognize that the ARDS Network protocol also allowed an increase in tidal volume to 8 mL/kg in certain situations. Even in patients receiving NIV, a tidal volume > 10 mL/kg leads to poorer outcomes.[76] A lung-protective strategy should be used with all mechanically ventilated patients.[77]

### NIV for COPD

High-level evidence supports the use of NIV in patients with COPD. In a meta-analysis of 12 RCTs, there was a significant reduction in mortality with the use of NIV ($P = .001$).[78] Mortality was 50 of 529 (9.5%) for those receiving NIV compared to 97 of 533 (18.2%) in the control group. This 8.7% absolute risk reduction translates to an NNT of 11. In other words, for every 11 patients with COPD who receive NIV, 1 life is saved. This is strong support for the use of NIV for COPD exacerbation.[79]

The use of NIV for COPD exacerbation is supported by CPGs.[44] There is a strong recommendation for patients with COPD exacerbation and respiratory acidosis (pH $\leq 7.35$). There is also a strong recommendation for a trial of NIV in patients considered to require endotracheal intubation and mechanical ventilation, unless the patient is immediately deteriorating. However, the evidence must be used wisely. NIV should not be used in patients with hypercapnia who are not acidotic in the setting of a COPD exacerbation.

A meta-analysis of 13 RCTs evaluated mortality for home NIV compared with no device in patients with COPD and stable hypercapnia.[80] NIV, compared with no device, was associated with lower risk of mortality ($P = .003$). Mortality was 22.3% for NIV and 28.6% without NIV. This 6.3% absolute risk reduction translates to an NNT of

Table 4. Common Reasons Clinicians Do Not Practice Evidence-Based Medicine

| |
| --- |
| They do not recognize the evidence (eg, laziness). |
|   They do not read the literature. |
|   They do not attend conferences. |
|   They do not talk to their colleagues. |
| They do not believe the evidence (eg, ignorance). |
|   They do not believe the right study was done (eg, wrong dose). |
|   The study is not consistent with lower levels of evidence (eg, animal studies). |
| They believe that incorporating the evidence into practice is someone else's job (eg, blaming). |
|   The system prevents incorporating the latest evidence. |
|   There is not enough time. |
|   The right resources are not available. |
|   There is no hospital policy supporting the practice change. |
| They do not think the evidence applies to their practice (eg, stubbornness). |
|   Their patients are sicker. |
|   Their patients are older. |
|   Their patients have always done fine with their expert treatment. |

From Reference 7.

16. CPGs suggest use of nocturnal NIV in addition to usual care for patients with chronic stable hypercapnic COPD.[45] Also suggested is targeted normalization of $P_{aCO_2}$ in patients with hypercapnic COPD on long-term NIV.[45]

## Therapy That Evidence Does Not Support

### Postoperative Incentive Spirometry

In the 1960s and 1970s, intermittent positive-pressure breathing was commonly prescribed in hopes of preventing postoperative pulmonary complications. In the mid-1970s, intermittent positive-pressure breathing was criticized as lacking sufficient evidence to support its use.[81,82] At the same time that intermittent positive-pressure breathing came under criticism as ineffective, the incentive spirometer was introduced by Bartlett et al.[83] Narrative reviews published in RESPIRATORY CARE suggest that the evidence supporting the use of incentive spirometry remains weak almost 50 y after the introduction of this therapy into practice.[84,85] CPGs do not support the routine use of incentive spirometry in the care of postoperative patients.[86,87] It might be argued that the replacement of intermittent positive-pressure breathing by incentive spirometry was the substitution of one unproven therapy for another.

### Intermittent Mandatory Ventilation

The first description of intermittent mandatory ventilation (IMV) was in 1973.[88] Of note, this report was a case series of 6 subjects, which led to widespread use of the mode. Included in the paper was a detailed description of how to configure the ventilator circuit to provide IMV. This could be easily implemented using equipment readily available in any respiratory care department. Because we *could* do this, few hesitated to question whether we *should* do this. Moreover, the use of IMV addressed a clinical problem, namely the poor triggers available on most ventilators at the time.

Two important studies published in the 1990s reported the poorest outcomes for ventilator weaning with gradual reductions in IMV, or synchronized IMV, compared to gradual reduction in pressure support or spontaneous breathing trials.[89,90] In a cohort of mechanically ventilated subjects, use of synchronized IMV compared with continuous mandatory ventilation did not offer any advantage in terms of clinical outcomes, despite a treatment-allocation bias that favored synchronized IMV.[16] An important lesson might be learned from the IMV experience. Back in 1973, many of us adopted use of IMV because it was new. But new is not necessarily better, which should be considered when other modes are introduced without high-level evidence.[91]

## Why Isn't the Best Evidence Implemented Into Practice?

There are a variety of reasons why practices persist that are not consistent with EBM (Table 4). Some clinicians do not accept a hierarchy of evidence, or grading of evidence, arguing instead that experiential evidence and evidence from physiologic trials is equally important as evidence from well-done RCTs. Clinical wisdom is required to determine how best evidence is applied to the care of an individual patient. High-level evidence does not exist for many respiratory care practices. However, high-level evidence should be implemented when it is available. EBM does not discredit the value of observational studies, physiologic studies, and clinical experience. Such lower levels of evidence are important to generate hypotheses, assess mechanisms, and establish proof of principle. However, lower levels of evidence should be set aside when a well-done RCT is available.

Politics has increasingly affected the uptake of science and evidence. Evidence should not be viewed as red or blue, but rather as black and white. The facts are the facts. Unfortunately, science and public health efforts clash with individual rights. For ourselves, our patients, and society as a whole, we should recognize certain truths that save lives: secondhand smoke exposure restrictions, vaccinations, wearing a seatbelt in a car, wearing a helmet when riding a motorcycle, and wearing a face covering during a pandemic. To embrace science and the scientific method is not elitist and should not be a political statement; it is responsible.

## Summary

EBM has permeated all parts of health care practice, including respiratory care. The principles of EBM provide the tools to incorporate the best evidence into everyday practice. The principles of EBM provide a valuable approach to improve respiratory care practice.

### REFERENCES

1. Mishoe SC, Hess DR. Evidence-based medicine in respiratory care. Respir Care 2000;46(11):1200.
2. Pierson DJ. Respiratory care as a science. Respir Care 1988;33(1):27-37.
3. Pierce AK, Saltzman HA. Conference on the scientific basis of respiratory therapy. Am Rev Respir Dis 1974;110(6 Pt 2):1-3.
4. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. JAMA 1992;268(17):2420-2425.
5. Guyatt G, Rennie D, Meade M, Cook D. Users' guides to the medical literature. Essentials of evidence-based practice, 2nd ed. New York: McGraw Hill; 2008.
6. Crichton JM. In my clinical experience. N Engl J Med 1971;285(26):1491.
7. Hess DR. What is evidence-based medicine and why should I care? Respir Care 2004;49(7):730-741.
8. Hess DR. Science and evidence: separating fact from fiction. Respir Care 2013;58(10):1649-1661.
9. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine. How to practice and teach EBM, 2nd ed. New York: Churchill Livingstone; 2000.
10. Sackett DL. Clinician-trialist rounds: 4. Why not do an N-of-1 RCT? Clin Trials 2011;8(3):350-352.
11. Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The N-of-1 clinical trial: the ultimate strategy for individualizing medicine? Per Med 2011;8(2):161-173.
12. Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature: a systematic review. Med Care 2011;49(8):761-768.
13. Hess DR. Airway clearance: physiology, pharmacology, techniques, and practice. Respir Care 2007;52(10):1392-1396.
14. Mirza RD, Punja S, Vohra S, Guyatt G. The history and development of N-of-1 trials. J R Soc Med 2017;110(8):330-340.
15. Smith GC, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ 2003;327(7429):1459-1461.
16. Ortiz G, Frutos-Vivar F, Ferguson ND, Esteban A, Raymondos K, Apezteguía C, et al. Outcomes of patients ventilated with synchronized intermittent mandatory ventilation with pressure support: a comparative propensity score study. Chest 2010;137(6):1265-1277.
17. Ferguson ND, Cook DJ, Guyatt GH, Mehta S, Hand L, Austin P, et al. High-frequency oscillation in early acute respiratory distress syndrome. N Engl J Med 2013;368(9):795-805.
18. Young D, Lamb SE, Shah S, MacKenzie I, Tunnicliffe W, Lall R, et al. High-frequency oscillation for acute respiratory distress syndrome. N Engl J Med 2013;368(9):806-813.
19. Acute Respiratory Distress Syndrome Network, Brower RG, Matthay MA, Morris A, Schoenfeld D, Thompson BT, Wheeler A. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. N Engl J Med 2000;342(18):1301-1308.
20. Dellinger RP, Zimmerman JL, Taylor RW, Straube RC, Hauser DL, Criner GJ, et al. Effects of inhaled nitric oxide in patients with acute respiratory distress syndrome: results of a randomized phase II trial. Crit Care Med 1998;26(1):15-23.
21. Writing Group for the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial Investigators, Cavalcanti AB, Suzumura EA, Laranjeira LN, Paisani DM, Damiani LP, et al. Effect of lung recruitment and titrated positive end-expiratory pressure (PEEP) vs low PEEP on mortality in patients with acute respiratory distress syndrome: a randomized clinical trial. JAMA 2017;318(14):1335-1345.
22. Hodgson CL, Cooper DJ, Arabi Y, King V, Bersten A, Bihari S, et al. Maximal recruitment open lung ventilation in acute respiratory distress syndrome (PHARLAP): a phase II, multicenter randomized controlled clinical trial. Am J Respir Crit Care Med 2019;200(11):1363-1372.
23. Mancebo J, Mercat A, Brochard L. Maximal lung recruitment in acute respiratory distress syndrome: a nail in the coffin. Am J Respir Crit Care Med 2019;200(11):1331-1333.
24. Wilton NK, Slim AM. Application of the principles of evidence-based medicine to patient care. South Med J 2012;105(3):136-143.
25. Hess DR. Information retrieval in respiratory care: tips to locate what you need to know. Respir Care 2004;49(4):389-399.
26. Chatburn RL. How to find the best evidence. Respir Care 2009;54(10):1360-1365.
27. Citrome L, Moss SV, Graf C. How to search and harvest the medical literature: let the citations come to you, and how to proceed when they do. Int J Clin Pract 2009;63(11):1565-1570.
28. Callcut RA, Branson RD. How to read a review paper. Respir Care 2009;54(10):1379-1385.
29. Uman LS. Systematic reviews and meta-analyses. J Can Acad Child Adolesc Psychiatry 2011;20(1):57-59.
30. Siddaway AP, Wood AM, Hedges LV. How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. Annu Rev Psychol 2019;70:747-770.
31. Delgado-Rodríguez M, Sillero-Arenas M. Systematic review and meta-analysis. Med Intensiva 2018;42(7):444-453.
32. Branson RD. Respiratory care and the cochrane collaboration. Respir Care 2020;65(4):581.
33. Hess DR. AARC Should noninvasive ventilation be used for treatment of acute cardiogenic pulmonary edema? a cochrane review summary with commentary. Respir Care 2020;65(4):573-574.
34. Willis LD. The role of noninvasive ventilation in cystic fibrosis: a cochrane review summary with commentary. Respir Care 2020;65(5):719-721.
35. Hernandez AV, Marti KM, Roman YM. Meta-analysis. Chest 2020;158(1S):S97-S102.
36. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-560.
37. Kang H, Zhao Z, Tong Z. Effect of high-flow nasal cannula oxygen therapy in immunocompromised subjects with acute respiratory failure. Respir Care 2020;65(3):369-376.
38. Hess DR. AARC Clinical practice guidelines: phase 4. Respir Care 2021;66(1):177-178.
39. Brozek JL, Akl EA, Compalati E, Kreis J, Terracciano L, Fiocchi A, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines part 3 of 3. The GRADE approach to developing recommendations. Allergy 2011;66(5):588-595.
40. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64(4):383-394.
41. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P, et al. GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. J Clin Epidemiol 2013;66(2):151-157.
42. Ransohoff DF, Pignone M, Sox HC. How to decide whether a clinical practice guideline is trustworthy. JAMA 2013;309(2):139-140.

43. Fan E, Del Sorbo L, Goligher EC, Hodgson CL, Munshi L, Walkey AJ, et al. An official American Thoracic Society/European Society of Intensive Care Medicine/Society of Critical Care Medicine Clinical Practice Guideline: mechanical ventilation in adult patients with acute respiratory distress syndrome. Am J Respir Crit Care Med 2017;195 (9):1253-1263.

44. Rochwerg B, Brochard L, Elliott MW, Hess D, Hill NS, Nava S, et al. Official ERS/ATS clinical practice guidelines: noninvasive ventilation for acute respiratory failure. Eur Respir J 2017;50(2):1602426.

45. Macrea M, Oczkowski S, Rochwerg B, Branson RD, Celli B, Coleman JM, et al. Long-term noninvasive ventilation in chronic stable hypercapnic chronic obstructive pulmonary disease. An official American Thoracic Society clinical practice guideline. Am J Respir Crit Care Med 2020;202(4):e74-e87.

46. Schoenberg NC, Barker AF, Bernardo J, Deterding RR, Ellner JJ, Hess DR, et al. A Comparative analysis of pulmonary and critical care medicine guideline development methodologies. Am J Respir Crit Care Med 2017;196(5):621-627.

47. Mussa CC, Gomaa D, Rowley DD, Schmidt U, Ginier E, Strickland SL. AARC Clinical Practice Guideline: management of adult patients with tracheostomy in the acute care setting. Respir Care 2021;66(1):156-169.

48. Volsko TA, Parker SW, Deakins K, Walsh BK, Fedor KL, Valika T, et al. AARC Clinical Practice Guideline: management of pediatric patients with tracheostomy in the acute care setting. Respir Care 2021;66(1):144-155.

49. Dol J, Tutelman PR, Chambers CT, Barwick M, Drake EK, Parker JA, et al. Health researchers' use of social media: scoping review. J Med Internet Res 2019;21(11):e13687.

50. Chan WS, Leung AY. Use of social network sites for communication among health professionals: systematic review. J Med Internet Res 2018;20(3):e117.

51. Topf JM, Hiremath S. Social media, medicine and the modern journal club. Int Rev Psychiatry 2015;27(2):147-154.

52. Zabor EC, Kaizer AM, Hobbs BP. Randomized controlled trials. Chest 2020;158(1S):S79-S87.

53. Ashbaugh DG, Bigelow DB, Petty TL, Levine BE. Acute respiratory distress in adults. Lancet 1967;2(7511):319-323.

54. Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A, et al. 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.

55. Wang X, Ji X. Sample size estimation in clinical research: from randomized controlled trials to observational studies. Chest 2020;158 (1S):S12-S20.

56. Combes A, Hajage D, Capellier G, Demoule A, Lavoué S, Guervilly C, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome. N Engl J Med 2018;378(21):1965-1975.

57. Abbott JH. The distinction between randomized clinical trials (RCTs) and preliminary feasibility and pilot studies: what they are and are not. J Orthop Sports Phys Ther 2014;44(8):555-558.

58. McCoy CE. Understanding the intention-to-treat principle in randomized controlled trials. West J Emerg Med 2017;18(6):1075-1078.

59. Doshi P, Whittle JS, Bublewicz M, Kearney J, Ashe T, Graham R, et al. High-velocity nasal insufflation in the treatment of respiratory failure: a randomized clinical trial. Ann Emerg Med 2018;72(1):73-83.

60. Kaul S, Diamond GA. Good enough: a primer on the analysis and interpretation of noninferiority trials. Ann Intern Med 2006;145(1):62-69.

61. Algera AG, Pisani L, Serpa Neto A, den Boer SS, Bosch FFH, Bruin K, et al. Effect of a lower vs higher positive end-expiratory pressure strategy on ventilator-free days in ICU patients without ARDS: a randomized clinical trial. JAMA 2020;324(24):2509-2520.

62. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1 (8476):307-310.

63. Story DA, Tait AR. Survey research. Anesthesiology 2019;130 (2):192-202.

64. Nayak BK, Hazra A. How to choose the right statistical test? Indian J Ophthalmol 2011;59(2):85-86.

65. Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. N Engl J Med 2001;345(19):1368-1377.

66. Pro CI, Yealy DM, Kellum JA, Huang DT, Barnato AE, Weissfeld LA, et al. A randomized trial of protocol-based care for early septic shock. N Engl J Med 2014;370(18):1683-1693.

67. Peake SL, Delaney A, Bailey M, Bellomo R, Cameron PA, Cooper DJ, et al. Goal-directed resuscitation for patients with early septic shock. N Engl J Med 2014;371(16):1496-1506.

68. Mouncey PR, Osborn TM, Power GS, Harrison DA, Sadique MZ, Grieve RD, et al. Trial of early, goal-directed resuscitation for septic shock. N Engl J Med 2015;372(14):1301-1311.

69. Villar J, Kacmarek RM, Hedenstierna G. From ventilator-induced lung injury to physician-induced lung injury: why the reluctance to use small tidal volumes? Acta Anaesthesiol Scand 2004;48(3):267-271.

70. Dreyfuss D, Saumon G. Ventilator-induced lung injury: lessons from experimental studies. Am J Respir Crit Care Med 1998;157(1):294-323.

71. Needham DM, Colantuoni E, Mendez-Tellez PA, Dinglas VD, Sevransky JE, Dennison Himmelfarb CR, et al. Lung protective mechanical ventilation and two year survival in patients with acute lung injury: prospective cohort study. BMJ 2012;344:e2124.

72. Needham DM, Yang T, Dinglas VD, Mendez-Tellez PA, Shanholtz C, Sevransky JE, et al. Timing of low tidal volume ventilation and intensive care unit mortality in acute respiratory distress syndrome. A prospective cohort study. Am J Respir Crit Care Med 2015;191(2):177-185.

73. Karalapillai D, Weinberg L, Peyton P, Ellard L, Hu R, Pearce B, et al. Effect of intraoperative low tidal volume vs conventional tidal volume on postoperative pulmonary complications in patients undergoing major surgery: a randomized clinical trial. JAMA 2020;324(9):848-858.

74. Simonis FD, Serpa Neto A, Binnekade JM, Braber A, Bruin KCM, Determann RM, et al. Effect of a Low vs intermediate tidal volume strategy on ventilator-free days in intensive care unit patients without ARDS: A randomized clinical trial. JAMA 2018;320(18):1872-1880.

75. Rubenfeld GD, Shankar-Hari M. Lessons from ARDS for non-ARDS research: remembrance of trials past. JAMA 2018;320(18):1863-1865.

76. Carteaux G, Millán-Guilarte T, De Prost N, Razazi K, Abid S, Thille AW, et al. Failure of noninvasive ventilation for de novo acute hypoxemic respiratory failure: role of tidal volume. Crit Care Med 2016;44 (2):282-290.

77. Rackley CR, MacIntyre NR. Low tidal volumes for everyone? Chest 2019;156(4):783-791.

78. Cabrini L, Landoni G, Oriani A, Plumari VP, Nobile L, Greco M, et al. Noninvasive ventilation and survival in acute care settings: a comprehensive systematic review and meta-analysis of randomized controlled trials. Crit Care Med 2015;43(4):880-888.

79. Hess DR. The evidence is in: noninvasive ventilation saves lives. Crit Care Med 2015;43(4):927-928.

80. Wilson ME, Dobler CC, Morrow AS, Beuschel B, Alsawas M, Benkhadra R, et al. Association of home noninvasive positive pressure ventilation with clinical outcomes in chronic obstructive pulmonary disease: a systematic review and meta-analysis. JAMA 2020;323(5):455-465.

81. Baker JP. Magnitude of usage of intermittent positive pressure breathing. Am Rev Respir Dis 1974;110(6 Pt 2):170-177.

82. Cheney FW Jr, Nelson EJ, Horton WG. The function of intermittent positive pressure breathing related to breathing patterns. Am Rev Respir Dis 1974;110(6 Pt 2):183-187.

83. Bartlett RH, Gazzaniga AB, Geraghty TR. Respiratory maneuvers to prevent postoperative pulmonary complications: a critical review. JAMA 1973;224(7):1017-1021.

84. Eltorai AEM, Szabo AL, Antoci V Jr, Ventetuolo CE, Elias JA, Daniels AH, et al. Clinical effectiveness of incentive spirometry for the prevention of postoperative pulmonary complications. Respir Care 2018;63(3):347-352.

85. Kotta PA, Ali JM. Incentive spirometry for prevention of postoperative pulmonary complications after thoracic surgery. Respir Care 2021;66(2):327-333.

86. Restrepo RD, Wettstein R, Wittnebel L, Tracy M. Incentive spirometry: 2011. Respir Care 2011;56(10):1600-1604.

87. Strickland SL, Rubin BK, Drescher GS, Haas CF, O'Malley CA, Volsko TA, et al. AARC clinical practice guideline: effectiveness of nonpharmacologic airway clearance therapies in hospitalized patients. Respir Care 2013;58(12):2187-2193.

88. 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.

89. Brochard L, Rauss A, Benito S, Conti G, Mancebo J, Rekik 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.

90. Esteban A, Frutos F, Tobin MJ, Alía I, Solsona JF, Valverdú 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.

91. Hess DR. Ventilator modes: where have we come from and where are we going? Chest 2010;137(6):1256-1258.