

patients have CFTR genotypes that do not respond to current CFTR modulators. Additional follow-up data, collected over a longer period of time, are clearly necessary to fully establish the effects of elixacaftor–tezacaftor–ivacaftor on lung transplantation in eligible patients with advanced disease. Nevertheless, our data suggest that clinically significant improvements in lung function, body weight, and gas exchange as well as symptoms and quality of life will allow healthcare teams to postpone lung transplantation in many patients.

The data provided in our study therefore support granting access to elixacaftor–tezacaftor–ivacaftor to all eligible patients throughout the world and seem paramount in the care of patients with CF, albeit with a careful monitoring of long-term effectiveness and potential adverse outcomes. ■

Author disclosures are available with the text of this letter at www.atsjournals.org.

Pierre-Régis Burgel, M.D., Ph.D.*
Institut Cochin, Université de Paris
Paris, France

Cochin Hospital
Paris, France
and

ERN-Lung CF Network
Frankfurt, Germany

Jennifer Da Silva, B.S.
Cochin Hospital
Paris, France
and

ERN-Lung CF Network
Frankfurt, Germany

Jean-Louis Paillasseur, Ph.D.
EffiStat
Paris, France

Clémence Martin, M.D., Ph.D.
Institut Cochin, Université de Paris
Paris, France

Cochin Hospital
Paris, France
and

ERN-Lung CF Network
Frankfurt, Germany

On behalf of all the authors

ORCID ID: 0000-0003-0903-9828 (P.-R.B.).

*Corresponding author (e-mail: pierre-regis.burgel@aphp.fr).

References

- Burgel P-R, Durieu I, Chiron R, Ramel S, Danner-Boucher I, Prevotat A, et al.; French Cystic Fibrosis Reference Network Study Group. Rapid improvement after starting elixacaftor–tezacaftor–ivacaftor in patients with cystic fibrosis and advanced pulmonary disease. *Am J Respir Crit Care Med* 2021;204:64–73.
- Burgel PR, Durieu I, Chiron R, Mely L, Prevotat A, Murriss-Espin M, et al.; French Cystic Fibrosis Reference Network study group. Clinical response to lumacaftor–ivacaftor in patients with cystic fibrosis according to baseline lung function. *J Cyst Fibros* 2021;20:220–227.
- McColley SA, Konstan MW, Ramsey BW, Stuart Elborn J, Boyle MP, Wainwright CE, et al. Lumacaftor/ivacaftor reduces pulmonary exacerbations in patients irrespective of initial changes in FEV₁. *J Cyst Fibros* 2019;18:94–101.
- Middleton PG, Mall MA, Dřevínek P, Lands LC, McKone EF, Polineni D, et al.; VX17-445-102 Study Group. Elexacaftor-tezacaftor-ivacaftor for cystic fibrosis with a single Phe508del allele. *N Engl J Med* 2019;381:1809–1819.
- Burgel PR, Munck A, Durieu I, Chiron R, Mely L, Prevotat A, et al.; French Cystic Fibrosis Reference Network Study Group. Real-life safety and effectiveness of lumacaftor-ivacaftor in patients with cystic fibrosis. *Am J Respir Crit Care Med* 2020;201:188–197.

Copyright © 2021 by the American Thoracic Society



Informing Healthcare Decisions with Observational Research Assessing Causal Effect: An American Thoracic Society Statement Not Ready for Implementation

To the Editor:

A recently published American Thoracic Society Statement concluded that observational studies (OS) should be included in guideline development and used in clinical decision-making in absence of high-quality randomized controlled trials (RCTs) and can “contribute compelling evidence for causal inference” (1). The authors contend that OS have better generalizability and/or external validity, less publication bias, imprecision, and inconsistency, and lower cost; enroll larger sample sizes; have fewer limitations resulting from lack of equipoise; and can be used to assess cause and effect. A more evenly balanced consideration of these contentions is needed.

The authors propose that OS produce higher levels of generalizability/external validity because efficacy RCTs are frequently conducted in academic centers and use numerous inclusion and exclusion criteria. These are not problems with RCTs *per se*, however, as investigators can specify sites where studies should be conducted and can define inclusion criteria as narrowly or broadly as they wish (2). The authors correctly note that pragmatic RCTs address many of these concerns and these preserve the critical element of randomization. Accordingly, the benefits of randomization need not be killed on the altar of generalizability/external validity.

The authors state that publication bias, imprecision, inconsistency, and lack of equipoise adversely affect RCTs, but these concerns apply to

Ⓢ This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). For commercial usage and reprints, please contact Diane Gern (dgern@thoracic.org).

Supported by Department of Defense RPMRP Clinical Trial Award: Sigh Ventilation to Reduce the Incidence and/or the Severity of the Acute Respiratory Distress Syndrome (Principal Investigator: R.K.A.) and by PCS-1504-30430: Roflumilast or Azithromycin to Prevent COPD Exacerbations (PECORI; Principal Investigator: J. Krishnan).

Originally Published in Press as DOI: 10.1164/rccm.202102-0492LE on June 3, 2021

both OS and RCTs (2). Importantly, RCTs can quantify the extent of lack of equipoise because patient flow diagrams should specify the number of patients excluded by physician preference (2).

The authors indicate that examining treatment effects in large numbers of patients provides more power than is achieved in RCTs. More power allows investigating smaller effect sizes, but the smaller the effect size, the more likely any perceived causality will be due to confounding (see below). Given a large enough sample size, analyses will always show significant results unless the true effect is zero.

The authors propose that larger sample sizes allow looking for treatment heterogeneities. The main criticism of OS is that, without randomization, unmeasured confounders and other biases cannot be excluded. This concern is multiplied when subgroups are selected, as the process of identifying these subgroups may result in additional unmeasured confounders (3, 4).

The authors note that OS are less expensive to perform. But if interventions based on OS are subsequently found to be ineffective (see below), or potentially even harmful, including them in guidelines that direct decision-making may actually *increase* the cost of care and, more importantly, may harm patients.

The authors indicate that the Newcastle-Ottawa Scale and the ROBINS-I tool can be used to assess the quality of OS, but they fail to note that the accuracy of both of these instruments has been questioned (5–7).

The authors contend that OS can be used to assess cause and effect using the Hill criteria (8). At least five of the nine Hill criteria merit comment with regard to their use in medical research.

Criterion 1 (strength of association) depends, in part, on the prevalence of the condition in question as well as on the statistical methods used in the analysis (9).

Criterion 2 (consistency) can only be invoked after all the relevant details of a causal mechanism are understood, and this is never the case in medicine.

Criterion 3 (specificity) indicates that associations are more likely to be causal when they lead to a single effect, a most unlikely scenario in medicine (although this may be less of an issue when applied to genetic associations).

Criterion 5 (biological gradient) refers to an association being causal if a dose–response relationship is observed. Hill noted that complex dose–response relationships may exist, and others have indicated that monotonic dose–response relationships are overly simplistic for most causal associations (8).

Criterion 6 (plausibility) is frequently not based on knowledge or data but rather on prior beliefs; hence the appearance of Bayesian reanalyses of recently published RCTs and development of the concepts defining behavioral economics for which Thaler, Kahneman, and Tversky shared the Nobel Prize.

Difficulties linking cause with effect were identified at least as early as the mid-1700s by David Hume in his *Treatise on Human Nature*: “Were ideas entirely loose and unconnected, chance alone would join them” (10). As Hill himself noted, “None of my nine [criteria] can bring indisputable evidence for or against the cause and effect hypothesis” (8).

The major reason OS should not be included in guidelines, direct decision-making, or infer causality is the concern for discrepant results. Although a number of articles and meta-analyses report that, on average, the effect sizes of RCTs and OS are similar, the more important question is how often the results of RCTs provide conclusions that are *opposite* to those of OS. Gershon and colleagues (11) suggest that these discrepancies occur because of problems in methodology or design

rather than being due to the intrinsic limitations of OS. Ioannidis (12), however, reviewed 45 articles (6 OS, 39 RCTs) that likely used very high-quality methodology as they were published in highly peer-reviewed journals and had more than 1,000 citations each. He found that subsequently published studies contradicted five of the six observational studies (83%) but only 9 of the 39 RCTs (23%; $P = 0.008$). I previously cited 19 studies of interventions relevant to Pulmonary and Critical Care (2). Fifteen of these were published in highly peer-reviewed journals. All of the OS reported clinically important benefits whereas all of the RCTs found no effects. It seems unlikely that all of these discrepant findings can be attributed to methodologic problems.

Increasingly complex statistical approaches have been developed in an attempt to circumvent the problem of unmeasured confounders in OS, but the authors of the Statement should acknowledge that many statisticians have concluded that no strategy adequately adjusts for confounding by indication (13–18).

Three of the Statement’s coauthors recently proposed that the universally accepted hierarchy of evidence ranking RCTs above OS should be altered such that OS and RCTs provide equal strengths of evidence (11). The Statement’s conclusions seem to be another attempt to change this hierarchy. As David Hume also noted: “Nothing is more usual and more natural for those, who pretend to discover any thing new to the world in philosophy and the sciences, than to insinuate the praises of their own systems” (8).

OS have important roles in medical research. These include suggesting hypotheses that merit testing by RCTs, tracking rare events (e.g., medication side effects), describing aspects of diseases, and contributing to prior beliefs. But if one accepts the dictum *primum non nocere*, OS should not be incorporated into guidelines. If OS are used in clinical decision-making, physicians should recognize that they are acting on lower strengths of evidence and are subjecting patients to interventions that subsequent RCTs will frequently find to be ineffective. ■

Author disclosures are available with the text of this letter at www.atsjournals.org.

Richard K. Albert, M.D.*
University of Colorado
Denver, Colorado

*Corresponding author (e-mail: richard.albert@cuanschutz.edu).

References

- Gershon AS, Lindenauer PK, Wilson KC, Rose L, Walkey AJ, Sadatsafavi M, *et al*. Informing healthcare decisions with observational research assessing causal effect: an official American Thoracic Society Research Statement. *Am J Respir Crit Care Med* 2021;203:14–23.
- Albert RK. “Lies, damned lies ...” and observational studies in comparative effectiveness research. *Am J Respir Crit Care Med* 2013; 187:1173–1177.
- Lantz B. The large sample size fallacy. *Scand J Caring Sci* 2013;27:487–492.
- Overhage JM, Overhage LM. Sensible use of observational clinical data. *Stat Methods Med Res* 2013;22:7–13.
- Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol* 2010;25:603–605.
- Hartling L, Milne A, Hamm MP, Vandermeer B, Ansari M, Tsertsvadze A, *et al*. Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. *J Clin Epidemiol* 2013;66:982–993.

7. Bero L, Chartres N, Diong J, Fabbri A, Ghersi D, Lam J, *et al.*. The risk of bias in observational studies of exposures (ROBINS-E) tool: concerns arising from application to observational studies of exposures. *Syst Rev* 2018;7:242.
8. Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;58:295–300.
9. Fedak KM, Bernal A, Capshaw ZA, Gross S. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol* 2015;12:14.
10. Hume D. A treatise on human nature. London: Clarendon Press; 1739.
11. Gershon AS, Jafarzadeh SR, Wilson KC, Walkey AJ. Clinical knowledge from observational studies everything you wanted to know but were afraid to ask. *Am J Respir Crit Care Med* 2018;198:859–867.
12. Ioannidis JPA. Contradicted and initially stronger effects in highly cited clinical research. *JAMA* 2005;294:218–228.
13. MacMahon S, Collins R. Reliable assessment of the effects of treatment on mortality and major morbidity, II: observational studies. *Lancet* 2001;357:455–462.
14. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovich C, Song F, *et al.*; International Stroke Trial Collaborative Group; European Carotid Surgery Trial Collaborative Group. Evaluating non-randomised intervention studies. *Health Technol Assess* 2003;7:iii–x, 1–173.
15. Bosco JLF, Silliman RA, Thwin SS, Geiger AM, Buist DSM, Prout MN, *et al.* A most stubborn bias: no adjustment method fully resolves confounding by indication in observational studies. *J Clin Epidemiol* 2010;63:64–74.
16. Vandembroucke JP. When are observational studies as credible as randomised trials? *Lancet* 2004;363:1728–1731.
17. Vandembroucke JP. Why do the results of randomised and observational studies differ? *BMJ* 2011;343:d7020.
18. Sox HC, Greenfield S. Comparative effectiveness research: a report from the Institute of Medicine. *Ann Intern Med* 2009;151:203–205.

stand by this framework for decision-making but acknowledge the need to update the framework as new evidence emerges. ■

Author disclosures are available with the text of this letter at www.atsjournals.org.

Andrea S. Gershon, M.D., M.Sc.*
Sunnybrook Health Sciences Centre
Toronto, Ontario, Canada
and

University of Toronto
Toronto, Ontario, Canada

Peter K. Lindenauer, M.D., M.Sc., M.H.M.
University of Massachusetts Medical School–Baystate
Springfield, Massachusetts

Jerry A. Krishnan, M.D., Ph.D.†
University of Illinois at Chicago
Chicago, Illinois

On behalf of all the authors

ORCID IDs: 0000-0002-0246-594X (A.S.G.); 0000-0002-8414-0507 (P.K.L.); 0000-0001-5525-4778 (J.A.K.).

*Corresponding author (e-mail: andrea.gershon@ices.on.ca).

†J.A.K. is Associate Editor of *AJRCCM*. His participation complies with American Thoracic Society requirements for recusal from review and decisions for authored works.

Copyright © 2021 by the American Thoracic Society



Reply to Albert

From the Authors:

We thank Dr. Albert for the comments about our ATS Research Statement (1). We agree that well-performed randomized clinical trials (RCTs) can produce high-quality evidence for making inferences about the causal effects of an intervention on outcomes. However, for most cases, evidence from high-quality RCTs for outcomes that are critical to decision-making does not exist or is insufficient for informing a course of action with confidence. For example, of 19 guideline recommendations in the recently published 2020 asthma guideline update from the National Asthma Education and Prevention Program, only 3 were based on high-quality evidence (2). The ATS Research Statement explains the framework proposed by the Grading of Recommendations Assessment, Development and Evaluation working group in cases in which there is insufficient evidence from RCTs. We

Ⓐ This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). For commercial usage and reprints, please contact Diane Gern (dgern@thoracic.org).

The authors are the co-chairs of the official American Thoracic Society Document Informing Health Care Decisions with Observational Research Assessing Causal Effect: An Official ATS Research Statement.

Originally Published in Press as DOI: 10.1164/rccm.202104-0851LE on June 3, 2021

References

1. Gershon AS, Lindenauer PK, Wilson KC, Rose L, Walkey AJ, Sadatsafavi M, *et al.* Informing healthcare decisions with observational research assessing causal effect. An official American Thoracic Society Research Statement. *Am J Respir Crit Care Med* 2021;203:14–23.
2. Cloutier MM, Baptist AP, Blake KV, Brooks EG, Bryant-Stephens T, DiMango E, *et al.*; Expert Panel Working Group of the National Heart, Lung, and Blood Institute (NHLBI) administered and coordinated National Asthma Education and Prevention Program Coordinating Committee (NAEPPCC). 2020 focused updates to the Asthma Management Guidelines: a report from the National Asthma Education and Prevention Program Coordinating Committee Expert Panel Working Group. *J Allergy Clin Immunol* 2020;146:1217–1270.

Copyright © 2021 by the American Thoracic Society



Multiple Manifestations of Systemic Sclerosis Affect Walk Distance

To the Editor:

We welcome the novel report on the efficacy of B-cell depletion in the treatment of pulmonary arterial hypertension (PAH) associated with

Ⓐ This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). For commercial usage and reprints, please contact Diane Gern (dgern@thoracic.org).

Originally Published in Press as DOI: 10.1164/rccm.202104-0938LE on June 9, 2021