ISSN (ON-LINE): 1806-3756

Jornal Brasileiro de **Pneumologia**

PUBLICAÇÃO OFICIAL DA SOCIEDADE BRASILEIRA DE PNEUMOLOGIA E TISIOLOGIA

Volume 50
July
2024

Research Methods
SBPT

Volume 50, Research Methods July 2024

SCS - Quadra 1 - Bloco K - Sala 203 - Ed. Denasa Brasília - DF - 70398-900

sbpt@sbpt.org.br | 08000 61 6218 / (061) 3245-1030 ou 3245-6218

Acompanhe nossos conteúdos:
@pneumosbpt

ISSN (ON-LINE): 1806-3756

Continuous and Bimonthly Publication, J Bras Pneumol. v. 50, Research Methods, July 2024

Associação Brasileira de Editores Científicos
ABEC

Publicação Indexada em:
Latindex, LILACS, Scielo Brazil, Scopus, Index Copernicus, ISI Web of Knowledge, MEDLINE e PubMed Central (PMC)

Disponível eletronicamente nas versões português e inglês:
www.jornaldepneumologia.com.br e www.scielo.br/jbpneu

NATIONAL LIBRARY OF MEDICINE

PMC

ISI Web of Knowledge℠

SCOPUS

SciELO Brazil

INDEX COPERNICUS INTERNATIONAL

latindex

EDITOR-IN-CHIEF
Marcia Margaret Menezes Pizzichini - Universidade Federal de Santa Catarina, Florianópolis - SC

DEPUTY EDITOR
Bruno Guedes Baldi - Universidade de São Paulo, São Paulo - SP

ASSOCIATE EDITORS
André Prato Schimidt - Universidade Federal do Rio Grande do Sul, Porto Alegre, RS | **Area:** Critical Care and Mechanical Ventilation
Bruno do Valle Pinheiro - Universidade Federal de Juiz de Fora, Juiz de Fora - MG | **Area:** Terapia intensiva/ Ventilação mecânica
Carlos Gustavo Verrastro - Universidade Federal de São Paulo, São Paulo - SP | **Area:** Imagem
Danilo Cortozi Berton - Universidade Federal do Rio Grande do Sul, Porto Alegre - RS | **Area:** Respiratory Physiological
Denise Rossato Silva - Universidade Federal do Rio Grande do Sul, Porto Alegre - RS | **Area:** Tuberculosis and Respiratory Infections
Edson Marchiori - Universidade Federal Fluminense, Niterói - RJ | **Area:** Image
Fernanda Carvalho de Queiroz Mello - Universidade Federal do Rio de Janeiro - Rio de Janeiro - RJ | **Area:** Tuberculosis and Respiratory Infections
Gilberto Castro Junior - Instituto Brasileiro de Controle do Câncer - São Paulo - SP | **Area:** Oncology
Giovanni Battista Migliori - Director WHO Collaborating Centre for TB and Lung Diseases, Fondazione S. Maugeri, Care and Research Institute, Tradate, Italy | **Area:** Tuberculosis and Respiratory Infections
Ian Pavord - Respiratory Medicine - University of Oxford | **Area:** Asthma
Jaqueline Sonoe Ota Arakaki - Universidade Federal de São Paulo, São Paulo - SP | **Area:** Pulmonary Circulation/ Pulmonary Hypertension
Klaus Irion - School of Biological Sciences, The University of Manchester, United Kingdom | **Area:** Image
Leonardo Araújo Pinto - Pontifícia Universidade Católica do Grande do Sul, Porto Alegre - RS | **Area:** Pneumopediatrics
Paul Jones - Respiratory Medicine at St George's, University of London | **Area:** COPD
Paulo Manuel Pêgo Fernandes - Universidade de São Paulo, São Paulo - SP | **Area:** Thoracic surgery
Pedro Rodrigues Genta - Universidade de São Paulo, São Paulo - SP | **Area:** Sleep
Regina Maria de Carvalho-Pinto - Universidade de São Paulo, São Paulo, SP | **Area:** Asthma/Other Chronic Respiratory Diseases
Rodrigo Silva Cavallazzi- Respiratory Medicine at St George's, University of London University of Louisville - Kentucky - USA | **Area:** UTI e Infecções Respiratórias
Rosemeri Maurici da Silva - Universidade Federal de Santa Catarina, Florianópolis - SC | **Area:** Infections and bronchiectasis
Simone Dal Corso - Universidade Nove de Julho, São Paulo (SP), Brasil. | **Area:** Respiratory physiotherapy/Exercise
Suzana Erico Tanni - Universidade Estadual Paulista "Julio de Mesquita Filho" - Botucatu - SP | **Area:** COPD and Epidemiology
Ubiratan de Paula Santos - Universidade de São Paulo - São Paulo - SP | **Area:** Smoking/Environmental and occupational respiratory diseases
Zafeiris Louvaris - University Hospitals Leuven, Leuven, Belgium | **Area:** Respiratory physiology

EDITORIAL COUNCIL
Alberto Cukier - Universidade de São Paulo, São Paulo - SP
Álvaro A. Cruz - Universidade Federal da Bahia, Salvador - BA
Ana C. Krieger - Weill Cornell Medical College - New York - USA
Ana Luiza Godoy Fernandes - Universidade Federal de São Paulo, São Paulo - SP
Antonio Segorbe Luis - Universidade de Coimbra, Coimbra - Portugal
Ascedio Jose Rodrigues - Universidade de São Paulo - São Paulo - SP
Brent Winston - University of Calgary, Calgary - Canada
Carlos Alberto de Assis Viegas - Universidade de Brasília, Brasília - DF
Carlos Alberto de Castro Pereira - Universidade Federal de São Paulo, São Paulo - SP
Carlos M. Luna - Hospital de Clínicas, Universidad de Buenos Aires, Buenos Aires - Argentina
Carmen Silvia Valente Barbas - Universidade de São Paulo, São Paulo - SP
Celso Ricardo Fernandes de Carvalho - Universidade de São Paulo, São Paulo - SP
Dany Jasinowodolinski - Universidade de São Paulo, São Paulo - SP
Denis Martinez - Universidade Federal do Rio Grande do Sul, Porto Alegre - RS
Douglas Bradley - University of Toronto, Toronto, ON - Canadá
Emílio Pizzichini - Universidade Federal de Santa Catarina, Florianópolis - SC
Fábio Biscegli Jatene - Universidade de São Paulo, São Paulo - SP
Frank McCormack - University of Cincinnati School of Medicine, Cincinnati, OH - USA
Geraldo Lorenzi Filho - Universidade de São Paulo, São Paulo - SP
Gilberto de Castro Junior - Universidade de São Paulo, São Paulo - SP
Gustavo Javier Rodrigo - Hospital Central de las Fuerzas Armadas, Montevidéu - Uruguay
Ilma Aparecida Paschoal - Universidade de Campinas, Campinas - SP
C. Isabela Silva Müller - Vancouver General Hospital, Vancouver, BC - Canadá
J. Randall Curtis - University of Washington, Seattle, Wa - USA
John J. Godleski - Harvard Medical School, Boston, MA - USA
José Alberto Neder - Queen's University - Ontario, Canada
José Antonio Baddini Martinez - Universidade de São Paulo, Ribeirão Preto - SP
José Dirceu Ribeiro - Universidade de Campinas, Campinas - SP
José Miguel Chatkin - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre - RS
José Roberto de Brito Jardim - Universidade Federal de São Paulo, São Paulo - SP
José Roberto Lapa e Silva - Universidade Federal do Rio de Janeiro, Rio de Janeiro - RJ
Kevin Leslie - Mayo Clinic College of Medicine, Rochester, MN - USA
Luiz Eduardo Nery - Universidade Federal de São Paulo, São Paulo - SP
Marc Miravitlles - University Hospital Vall d'Hebron - Barcelona, Catalonia - Spain
Marisa Dolhnikoff - Universidade de São Paulo, São Paulo - SP
Marli Maria Knorst - Universidade Federal do Rio Grande do Sul, Porto Alegre - RS
Mauro Musa Zamboni - Instituto Nacional do Câncer, Rio de Janeiro - RJ
Nestor Muller - Vancouver General Hospital, Vancouver, BC - Canadá
Noé Zamel - University of Toronto, Toronto, ON - Canadá
Oliver Augusto Nascimento - Universidade Federal de São Paulo - São Paulo - SP
Paul Noble - Duke University, Durham, NC - USA
Paulo Francisco Guerreiro Cardoso - Universidade de São Paulo, São Paulo - SP
Paulo Manuel Pêgo Fernandes - Universidade de São Paulo, São Paulo - SP
Peter J. Barnes - National Heart and Lung Institute, Imperial College, London - UK
Renato Sotto Mayor - Hospital Santa Maria, Lisboa - Portugal
Richard W. Light - Vanderbili University, Nashville, TN - USA
Rik Gosselink - University Hospitals Leuven - Bélgica
Robert Skomro - University of Saskatoon, Saskatoon - Canadá
Rubin Tuder - University of Colorado, Denver, CO - USA
Sérgio Saldanha Menna Barreto - Universidade Federal do Rio Grande do Sul, Porto Alegre - RS
Sonia Buist - Oregon Health & Science University, Portland, OR - USA
Talmadge King Jr. - University of California, San Francisco, CA - USA
Thais Helena Abrahão Thomaz Queluz - Universidade Estadual Paulista, Botucatu - SP
Vera Luiza Capelozzi - Universidade de São Paulo, São Paulo - SP

ISSN (ON-LINE): 1806-3756

JBP
Jornal Brasileiro de Pneumologia

Expediente

BRAZILIAN THORACIC SOCIETY

Office: SCS Quadra 01, Bloco K, Asa Sul, salas 203/204. Edifício Denasa, CEP 70398-900, Brasília, DF, Brazil. Tel. +55 61 3245-1030/+55 08000 616218. Website: www.sbpt.org.br.
E-mail: sbpt@sbpt.org.br

The Brazilian Journal of Pulmonology (ISSN 1806-3756) is published once every two months by the Brazilian Thoracic Society (BTS). The statements and opinions contained in the editorials and articles in this Journal are solely those of the authors thereof and not of the Journal's Editor-in-Chief, peer reviewers, the BTS, its officers, regents, members, or employees. Permission is granted to reproduce any figure, table, or other material published in the Journal provided that the source for any of these is credited.

BTS Board of Directors (2023-2024 biennium):
President: Margareth Maria Pretti Dalcolmo - RJ
President Elect (2025/2026 biennium): Ricardo Amorim Corrêa - MG
Secretary-General: Ricardo Luiz de Melo - DF
Director, Defense and Professional Practice: Octávio Messeder - BA
CFO: Maria Enedina Claudino Aquino Scuarcialupi - PB
Scientific Director: Valeria Maria Augusto - MG
Education Director: Clystenes Odyr Soares Silva - SP
Director, Communications: Waldo Luis Leite Dias de Mattos - RS
Editor-in-Chief of the Brazilian Journal of Pulmonology: Marcia Margaret Menezes Pizzichin - SC

AUDIT COMMITTEE (2023-2024 biennium):
Active Members: Elie FISS - SP, Eduardo Felipe Barbosa Silva - DF, Flávio Mendonça Andrade da Silva - MG

Alternates: Marcelo Tadday Rodrigues - RS, Carlos Alberto de Assis Viegas - DF, Fabio José Fabricio de Souza - SC

COORDINATORS, BTS DEPARTMENTS:
Thoracic Surgery: Artur Gomes Neto - AL
Sleep-disordered Breathing: Ricardo Luiz de Menezes Duarte - RJ
Respiratory Endoscopy: Luis Renato Alves - SP
Pulmonary Function: André Luís Pereira de Albuquerque - SP
Imaging: Danny Warszawiak - PR
Lung Diseases: Alexandre Todorovic Fabro - SP
Pediatric Pulmonology: Luiz Vicente Ribeiro Ferreira da Silva Filho - SP

COORDINATORS, BTS SCIENTIFIC COMMITTEES:
Asthma: Lilian Serrasqueiro Ballini Caetano - SP
Lung Cancer: Gustavo Faischew Prado - SP
Pulmonary Circulation: Veronica Moreira Amado - DF
Advanced Lung Disease: Paulo Henrique Ramos Feitosa - DF
Interstitial Diseases: Karin Mueller Storrer - PR
Environmental and Occupational Respiratory Diseases: Eduardo Algranti - SP
COPD: Luiz Fernando Ferreira Pereira - MG
Epidemiology: Suzana Erico Tanni Minamotos - SP
Cystic Fibrosis: Samia Zahi Rached - SP
Respiratory Infections and Mycoses: José Tadeu Colares Monteiro - PA
Pleura: Philppe de Figueiredo Braga Colares - SP
Smoking: Paulo Cesar Rodrigues Pinto Correa - MG
Intensive Care: Arthur Oswaldo de Abreu - RJ
Tuberculosis: Denise Rossato Silva - RS

ADMINISTRATIVE SECRETARIAT OF THE BRAZILIAN JOURNAL OF PULMONOLOGY
Address: SCS Quadra 01, Bloco K, Asa Sul, salas 203/204. Edifício Denasa, CEP 70398-900, Brasília, DF, Brazil. Tel. +55 61 3245-1030/+55 08000 616218.
Editorial Manager: Luana Maria Bernardes Campos.
E-mail: jbp@jbp.org.br | jbp@sbpt.org.br
Distribution: Free to members of the BTS and libraries

SUPPORT:
CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico
CAPES
Ministério da Educação
Ministério da Ciência, Tecnologia e Inovação
GOVERNO FEDERAL BRASIL PAÍS RICO É PAÍS SEM POBREZA

ISSN (ON-LINE): 1806-3756

Jornal Brasileiro de Pneumologia

Continuous and Bimonthly Publication, J Bras Pneumol. v. 50, Research Methods, July 2024

Contents

ISSN (ON-LINE): 1806-3756

Jornal Brasileiro de Pneumologia

Continuous and Bimonthly Publication, J Bras Pneumol. v. 50, Research Methods, July 2024

Contents

ISSN (ON-LINE): 1806-3756

Jornal Brasileiro de Pneumologia

Continuous and Bimonthly Publication, J Bras Pneumol. v. 50, Research Methods, July 2024

Contents

Preface

Juliana Carvalho Ferreira[1,2], Cecilia Maria Patino[1,3], Marcia Pizichinni[4]

We are happy to present this Clinical Research Methodology and Career Development e-book in Respiratory, Critical Care, and Sleep Medicine, a compilation of short, stand-along articles that have been published in the *Jornal Brasileiro de Pneumologia* (JBP) since 2015. The release of the e-book coincides with the 30-year celebration of the American Thoracic Society (ATS) Methods in Epidemiologic, Clinical, and Operations Research (MECOR) Program.

The ATS MECOR Program is a research method training course aiming at strengthening the capacity among clinicians, investigators, academicians, and public health professionals who primarily work in the fields of pulmonary, sleep, and critical care medicine. The MECOR program develops competency in the design, conduct, publication, and evaluation of research that are relevant to local settings. The program was created by the ATS in 1994 and first piloted in Latin America; and now it is offered in six different countries/regions worldwide. Since 2012, MECOR Latin America has been run by the *Asociación Latinoamericana de Tórax* (ALAT) in partnership with the ATS and local respiratory societies that host the course each year. In 2024, we celebrate 30 years of the program, and the course will be offered in Belo Horizonte, Brazil, with support from SBPT.

The program has trained over 1,800 students across the globe, the majority of whom from Latin America, where it has been offered every year since its creation 30 years ago. In Latin America, the majority of MECOR Faculty are MECOR Alumni, and the previous three Latin American course directors were alumni, attesting to the success of the program in building research capacity in our region, which we are very proud of. In Brazil, as in other Latin American countries, many of the pulmonologists in current and past leadership positions are MECOR alumni. Not by coincidence, the last four JBP Editors-in-Chiefs were MECOR alumni.

It was one of these alumni, Dr. Rogério de Souza, who, in 2015, as the JBP editor-in-chief, suggested that Cecy Patino and Juliana Ferreira, who were respectively Program Director and co-director of the MECOR program in Latin America, started a continued education series based on MECOR curriculum as a short article collection. The idea was very innovative, and we did not know at the time how it would be received by the readership. The aim of the series was to provide information on research methods to a broad readership of the JBP, many with little prior knowledge on the subject matter. One of the challenges was to write about complex topics, such as odds ratios, noninferiority trials, and regression analysis, in a language that was accessible to readers, but also technically accurate, as a short article. Cecy and Juliana were enthusiastic about the idea and wrote a list of potential topics on epidemiology, statistics, and methods in research. The first topic in the series was entitled "What does the p value really mean?" It may seem like a simple topic, but many people get it wrong, and explaining the statistical basis for truly understanding it in 500 words required a lot of work. When it was published, we were unsure if it would be of interest to the readers. But the article was very well received, with growing number of views and downloads, close to 300 per month on the website in 2023. It has also received more than 100 citations in Google Scholar. Therefore, everyone was enthusiastic about continuing the series.

When Dr. Bruno Baldi became JBP Editor-in-Chief in 2019, he invited Cecy and Juliana to continue with the series. At that point, in addition to Cecy and Juliana, other MECOR faculty were invited to co-author articles in the series, with the goal of bringing new ideas and perspectives. More recently, they have expanded the contributing authors to the MECOR alumni participating as teaching assistants in the MECOR Program to co-author articles in the series, as a means to continue to build research capacity.

In 2023, Dr. Marcia Pizzichini became JBP Editor-in-Chief, and not only did she suggest that the series should continue, but she also was highly enthusiastic about a long-time dream to make a compendium with all of the articles. The 30-year celebration of the program seemed like the perfect moment to do it. Having served as MECOR faculty in Latin America for many years, Marcia suggested that an e-book be created, and that printed copies were to be offered as a gift to students and faculty during the MECOR course in Belo Horizonte, from 19 to 24th of August 2024. The e-book has 51 articles, organized in four sections: Designing Research Protocols, Study Designs, Statistical Analysis: Tips for Clinical Researchers, and Advice For Young Researchers. All articles are one-to-two pages long, written in accessible language, and targeting clinicians who are interested in understanding and conducting research.

As we evaluate this compendium, we conclude that the main objective has been achieved. Similarly to what we see in the MECOR Program, it is difficult to

1. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.
4. Programa de Pós Graduação em Ciências Médicas Universidade Federal de Santa Catarina (UFSC), Florianópolis,SC

CC BY © 2024 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3756

measure if it was successful using objective metrics. The fact that the articles have had many views and downloads and that may have several citations is one way to measure success. But how do we measure the impact of an article on the research productivity in a country or a region? How do we measure the impact of an educational program such as MECOR in the research career of its alumni? We have some evidence: the MECOR program has been offered in Latin America every year for 30 years, thanks to the support from the ATS, ALAT, SBPT, and other respiratory societies in Latin America; thanks to the committed and knowledgeable faculty who have been donating their time to travel to a different country and teaching young investigators; thanks to the committed and curious students who have invested their time and resources to learn about research methodology. Over the years, alumni have become faculty, then course directors; they have risen to leadership roles in their Institutions, become respected researchers in their fields, professors in universities all over Latin America, and Editors-in-Chief of impactful research journals, such as JBP. Lastly, the MECOR program has built a strong community that has enabled and prioritized teaching, collaborative learning, and state-of-the-art research with the overall goal to improve the health in our communities in Latin America. This in turn has transformed our own lives and the lives of others. If that is not success, we don't know what is.

J Bras Pneumol. 2016;42(6):403-403
http://dx.doi.org/10.1590/S1806-37562016000000354

Developing research questions that make a difference

Cecilia Maria Patino[1,2], Juliana Carvalho Ferreira[2,3]

BACKGROUND

A clinical research question is defined as an ***uncertainty*** about a health problem that points to the need for meaningful understanding and deliberate investigation. [1] For clinicians interested in conducting high-quality clinical research, it is essential to recognize the fact that the research process starts with developing a question about a specific health-related area of interest. This is important because once the research question is defined, it has an impact on every remaining component of the research process, including generating the hypothesis and defining the appropriate study design, as well as the study population, study variables, and statistical approach. However, conceiving a sound research question is not an easy task; it requires having a particular set of personal skills and utilizing structured approaches.

DEVELOPING AND WRITING A RESEARCH QUESTION

Developing a research question starts by identifying a clinical problem that is important to patients, being related to managing and ultimately improving their health. The process requires clinician scientists to be curious about and attentive to day-to-day practice outcomes, as well as to be avid readers of the scientific literature, to participate in scientific activities (e.g., journal clubs), and to have access to a scientific mentor or collaborators interested in clinical research.

The research question itself should meet certain criteria, as summarized by the acronym FINGER, which stands for Feasible, Interesting, Novel, Good (for your career), Ethical, and Relevant (Chart 1).[1] We recommend going through the FINGER criteria systematically and discussing all issues with a mentor or colleague before writing the study protocol and conducting a study that will answer the proposed research question.

Once the research question has been defined, it should be written out in such a way that the answer can be expressed as either a number, typical of descriptive research questions (e.g., a prevalence related to disease burden, such as "What is the prevalence of asthma among favela residents in Brazil?"), or as a yes or no, typical of studies about associations between exposures and outcomes (e.g., "Is living in a *favela* in Brazil associated with increased mortality among adults with asthma?"). In addition, if the researcher has a hypothesis about the answer to the research question,[1] it is important that it be written out using a comprehensive approach, as summarized by the acronym PICOT, which stands for **P**opulation (the population to be included in the study), **I**ntervention (treatment applied to participants in the treatment arm), **C**omparison (treatment applied to the control group), **O**utcome (the primary outcome variable), and **T**ime (follow-up time to measure the outcome).[2]

INVESTING THE TIME AND EFFORT TO COME UP WITH A HIGH-QUALITY, WELL-WRITTEN RESEARCH QUESTION IS WORTH IT!

As clinician scientists who train clinicians to become successful researchers, we cannot emphasize enough the importance of investing one's time wisely to develop a high-quality research question. Researchers who conceive and clearly state a research question about an important health-related problem are at an advantage because they are more likely to convince key individuals to provide them with the necessary resources and support to carry out the study, as well as to increase the reporting quality of the paper to be published.[3]

Chart 1. Expanded descriptions of the recommended criteria for developing a good research question.

FINGER Criteria	
Feasible	Access to an adequate number of participants Research team has technical expertise to conduct the study Affordable: costs are reasonable and funding is available Can be completed in a reasonable time period
Interesting	Results of the study will be of interest to the research community
Novel	Provides new findings, extends or refutes previous findings
Good	For your career: fits into your career development plan
Ethical	Risk to participants is low/acceptable, considered ethical by peers and the Institutional Review Board
Relevant	To improve scientific knowledge, inform clinicians and health policy, and to impact future research

REFERENCES

1. Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. Designing clinical research. 4th ed. Philadelphia (PA): Lippincott Williams and Wilkins; 2013.
2. Haynes R. Forming research questions In: Haynes R, Sackett D, Guyatt GH, Tugwell P, editors. Clinical Epidemiology: How to do Clinical Practice Research. Philadelphia, PA: Lippincott Williams & Wilkins; 2006. p. 3-14.
3. Rios LP, Ye C, Thabane L. Association between framing of the research question using the PICOT format and reporting quality of randomized controlled trials. BMC Med Res Methodol. 2010;10:11. http://dx.doi.org/10.1186/1471-2288-10-11

1. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
2. Methods in Epidemiologic, Clinical and Operations Research (MECOR) Program, American Thoracic Society, New York, NY, USA, and Asociación Latinoamericana de Tórax, Montevideo, Uruguay.
3. Divisão de Pneumologia, Instituto do Coração – InCor – Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brasil.

BY-NC © 2016 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

Crafting a research protocol: a stepwise comprehensive approach

Ricardo Gassmann Figueiredo[1,2], Cecilia María Patino[2,3], Juliana Carvalho Ferreira [2,4]

PRACTICAL SCENARIO

A group of researchers plan to conduct a cross-sectional study to estimate the prevalence of frailty in elderly patients with moderate to severe asthma and to report a measure of association between asthma control and frailty.[1] The research protocol outlines the complex interactions of asthma control in frail patients and motivation to address this research question. Study design, objectives, methods, ethical issues, risks, and impact were also detailed in the protocol.

WHAT IS A RESEARCH PROTOCOL?

A well-structured research protocol guides researchers through the intricate process of conducting rigorous research. A research protocol is designed to be concise and self-contained, and to summarize the core aspects of the study. Self-discipline is vital in this process, as it requires the investigator to structure the central concepts of the study and reveal particular issues that demand attention.[2] The research protocol often serves as the foundation for the development of manual of operating procedures, which includes comprehensive information on the organization and policies of the study, as well as an operational approach to the procedures outlined in the study protocol; therefore, both documents complement each other.

ELEMENTS OF A RESEARCH PROTOCOL

The research protocol framework (outlined in Chart 1) usually includes a title, rationale, background information, objectives, methodology, data management, statistical plan, quality control, ethics, budget, developing plan, timeline, references, and appendices, although the sections included vary depending on institutional templates.

The title should be concise, descriptive, and engage readers, effectively reflecting the core of the research.[3] The background section outlines the driving factors and motivation for conducting the research. It should provide a broad context, elucidate the problem, address specific knowledge gaps, and establish the rationale for the study. In our practical example, the authors provided background information about how the multidimensional aspects of

Chart 1. Research protocol stepwise approach.

Step	Description
Title	Concise, reflecting study main ideas, and attracting reader's attention
Background and rationale	What is the problem? Why is it important? What is known about it?
Objectives	Specific, measurable, and established prior to carrying out the study
Relevance and study design	Contributions of the study to the field, aligned with rationale and objectives
Methods	Participants, exposures/intervention, outcomes, study setting, eligibility criteria, participant timeline, sample size, recruitment, and blinding Detailed script: How will the study be conducted? Why was the described design chosen?
Data collection, access, and management	Methods for data storage, security, privacy, and treatment of missing data
Statistical plan	Descriptive statistics, hypothesis testing, sample size, and power calculation
Quality control	Credibility of the research: instruments, data collection, data acquisition
Ethics	Ethical dilemmas, application to ethics research committees, Informed consent form
Roles and responsibilities	Affiliations, roles, and responsibilities of protocol contributors[3]
Budget	Detailed expenses: personnel, equipment, consumables, logistics
Funding	Sources of financial support
Dissemination plan	Effective communication of research findings
Timeline	Be realistic about project management throughout the research
References	Check publishers' guidelines, consider using reference manager software
Appendices	Extensive descriptions of procedures, questionnaires, and informed consent forms
Protocol version	Indicator of version and date of the protocol

1. Programa de Pós-Graduação em Saúde Coletiva, Universidade Estadual de Feira de Santana – PPGSC-UEFS – Feira de Santana (BA) Brasil.
2 Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
3 Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.
4. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2023 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3756

frailty are imbricated into proper asthma management in patients with advanced age. This section should align with the objectives, highlighting the potential impact of the study. Research objectives should be clear, measurable, precise, and set before conducting the study.[2] After the statement of the primary objective, secondary aims might be appropriate. The objectives will guide the study design and methodology, directing attention toward the intended research outcomes.

The methods section is a detailed blueprint of the research project and the basis for the manual of operating procedures. It should detail the study design, participant selection (eligibility, sampling, and recruitment), variables, data acquisition, data management (storage, security, privacy, and treatment of missing data), statistical plan, and sample size calculation. The scientific robustness of the study relies on its methodology, ensuring validity and replicability. The statistical plan should clearly outline the analysis methods, software used, and criteria for determining statistical significance. Quality control mechanisms uphold the internal validity of the study. This segment should describe measures to minimize bias and ensure data quality.[2] Steps might include regular data verification, calibration and certification of instruments, as well as research personnel training.

Ethical considerations are paramount in research. This section should document the issues that are likely to raise ethical concerns, including informed consent forms, confidentiality, data protection, and potential ethical dilemmas.[3] Moreover, it should also mention approvals obtained from institutional review boards. The budget section details the financial requirements of the research. It includes costs with personnel, equipment, materials, logistics, consumables, and contingencies. A realistic and well-planned timeline is crucial for successful project management.

Deficiencies in effectively disseminating and transferring research-based knowledge into clinical practice can impair the potential benefits of the research project. Therefore, most health research funding agencies expect commitment from investigators to disseminate the study findings actively. Integrating a dissemination plan in the research protocol will facilitate effective communication of research outcomes to the scientific community and those who can apply the knowledge in real-world situations.

KEY MESSAGES

1. A comprehensive research protocol not only provides a roadmap for the implementation of the study but also ensures that the research question is addressed according to high-quality research standards.
2. Quality control is essential to improve internal validity of the study.
3. A structured approach to conducting research reduces the likelihood of misleading conclusions and biases, ensuring validity and reproducibility of the study.

REFERENCES

1. Figueiredo RG, Pinheiro GP, Arata V, Leal MFM, Santana CVN, Tiraboschi TL, et al. Impact of frailty in elderly patients with moderate to severe asthma. PLoS One. 2022;17(7):e0270921. https://doi.org/10.1371/journal.pone.0270921
2. Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. Designing Clinical Research. Lippincott Williams & Wilkins; 2011.
3. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200-207. https://doi.org/10.7326/0003-4819-158-3-201302050-00583

Twelve tips to manage a research project—advice for the young investigator

Flavia Fonseca Fernandes[1,2,3], Juliana Carvalho Ferreira[1,4], Cecilia Maria Patino[1,5]

ILLUSTRATIVE SCENARIO

During the 2023 ATS-ALAT MECOR course recently held in Panama, a group of early career physicians from Latin America shared their experience on conducting research in their home countries. They expressed excitement about getting involved in research studies but were feeling anxious at the prospect of successfully carrying out both clinical and research activities. In search of guidance, they created a group to exchange experience and, inspired by the book by Hulley et al.,[1] they developed the following guide:

BEGINNER'S GUIDE TO RESEARCH PROJECT MANAGEMENT: 12 ESSENTIAL TIPS

1. Choose a familiar topic: start by selecting a topic that you enjoy and are knowledgeable about. This will make your research project more pleasurable, and likely to achieve high quality. It will also allow you to identify knowledge gaps more easily. Make sure you carry out a comprehensive review of the literature available, identifying high-quality articles which could potentially be references for your project and finalized manuscript.
2. Find a mentor, not just an advisor: look for a mentor who has experience in mentoring, shares your interest in the topic, has availability to meet regularly, and can guide you effectively.

Figure 1. Tips to research project management: choose a knowledgeable topic, find a mentor and a team, understand your institution and the ethics committee, plan your research, make a reasonable schedule, consult a statistician, and disseminate the results.

3. Understand your institution: identify what resources are available in your department or explore alternatives such as public databases and collaborations with other research groups.
4. Plan your research: develop a feasible research protocol that describes the research question and hypothesis, study design, study population, the intervention or exposure, and expected outcomes. Write out the research plan following your institution's guidelines. Be reasonable!
5. Develop a comprehensive statistical analysis plan: make sure you understand your study design, variables, and the appropriate statistical tests and processes. Whenever possible, consult and work with a statistician prior to data collection and throughout the study.
6. Make ethics a priority: ensure that your project addresses ethical aspects, including risks, benefits, and compliance with research guidelines.
7. Consult the Research Ethics Committee: engage with your institution's Ethics Committee or Scientific Council to understand their requirements and to receive valuable guidance. Some offer free consultations that can save you a lot of wasted time going back and forth with the project.
8. Create a manual of procedures: develop a manual of procedures that describes, in detail, how the research will be conducted, data collection procedures, and data storage processes. Test data collection forms to ensure they are appropriate for the study.
9. Set a realistic schedule: make sure to create a schedule that accommodates unexpected delays and that the duration of each step is feasible in your environment. Research often encounters unforeseen obstacles.
10. Ensure data collection uniformity: train your data collection team using a standardized procedure plan to ensure consistency.
11. Expect the unexpected: keep your focus when unforeseen situations or delays arise during planning and execution. These are common in research. Stay calm and carry on.
12. Disseminate research results: the work is only half done after data are collected. The final step is documenting your findings for an oral presentation, a poster at conferences, or an original manuscript for publication.

REFERENCE

1. Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. Designing clinical research. 4th ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2013.

1. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Serviço de Pneumologia, Instituto de Gestão Estratégica em Saúde do Distrito Federal, Brasília, Brasil.
3. Unidade de Pneumologia, Hospital Regional da Asa Norte, Secretaria de Saúde do Distrito Federal, Brasília, Brasil.
4. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
5. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.

BY-NC © 2023 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3756

Inclusion and exclusion criteria in research studies: definitions and why they matter

Cecilia Maria Patino[1,2,a], Juliana Carvalho Ferreira[1,3,b]

PRACTICAL SCENARIO

A cross-sectional multicenter study evaluated self-reported adherence to inhaled therapies among patients with COPD in Latin America.[1] Inclusion and exclusion criteria for the study are shown in Chart 1. The authors found that self-reported adherence was low in 20% of the patients, intermediate in 29%, and high in 51%; and that poor adherence was associated with more exacerbations in the past year, a lower smoking history, and a lower level of education. The authors concluded that suboptimal adherence to inhaled therapies among COPD patients was common and that interventions to improve adherence are warranted.

BACKGROUND

Establishing inclusion and exclusion criteria for study participants is a standard, required practice when designing high-quality research protocols. Inclusion criteria are defined as the key features of the target population that the investigators will use to answer their research question.[2] Typical inclusion criteria include demographic, clinical, and geographic characteristics. In contrast, exclusion criteria are defined as features of the potential study participants who meet the inclusion criteria but present with additional characteristics that could interfere with the success of the study or increase their risk for an unfavorable outcome. Common exclusion criteria include characteristics of eligible individuals that make them highly likely to be lost to follow-up, miss scheduled appointments to collect data, provide inaccurate data, have comorbidities that could bias the results of the study, or increase their risk for adverse events (most relevant in studies testing interventions).

It is very important that investigators not only define the appropriate inclusion and exclusion criteria when designing a study but also evaluate how those decisions will impact the external validity of the results of the study. Common errors regarding inclusion and exclusion criteria include the following: using the same variable to define both inclusion and exclusion criteria (for example, in a study including only men, listing being a female as an exclusion criterion); selecting variables as inclusion criteria that are not related to answering the research question; and not describing key variables in the inclusion criteria that are needed to make a statement about the external validity of the study results.

IMPACT OF THE INCLUSION AND EXCLUSION CRITERIA ON THE EXTERNAL VALIDITY OF THE STUDY

In our example, the investigators described the inclusion criteria related to demographic characteristics (age ≥ 40 years of age and male or female gender), clinical characteristics (diagnosis of COPD, stable disease, outpatient, and current or former smoker); and exclusion criteria related to comorbidities that could bias the results (sleep apnea, other chronic respiratory diseases, and acute or chronic conditions that could limit the ability of the patient to participate in the study). On the basis of these inclusion and exclusion criteria, we can make a judgment regarding their impact on the external validity of the results. Making those judgments requires in-depth knowledge of the area of research, as well as of in what direction each criterion could affect the external validity of the study. As an example, the authors excluded patients with comorbidities, and it is therefore possible that the levels of nonadherence reported would not be generalizable to COPD patients with comorbidities, who most likely would show higher levels of nonadherence due to their more complex medication regimens.

Chart 1. Inclusion and exclusion criteria for a cross-sectional multicenter study of patients with COPD in Latin America.[1]

Inclusion criteria	Exclusion criteria
• Adults ≥40 years of age • Diagnosis of COPD at least for 1 year • At least one spirometry in the last year with a post-bronchodilator $FEV_1/FVC < 0.70$ • Current or former smokers (> 10 pack-years) • Stable disease (no recent exacerbation)	• Diagnosis of sleep apnea or any other chronic respiratory disease • Any acute or chronic condition that would limit the ability of the patient to participate in the study • Refusal to give informed consent

REFERENCES

1. Montes de Oca M, Menezes A, Wehrmeister FC, Lopez Varela MV, Casas A, Ugalde L, et al. Adherence to inhaled therapies of COPD patients from seven Latin American countries: The LASSYC study. PLoS One. 2017;12(11):e0186777. https://doi.org/10.1371/journal.pone.0186777
2. Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. Designing Clinical Research. 3rd ed, Philadelphia, PA: Lippincott Williams & Wilkins; 2007.

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
a. http://orcid.org/0000-0001-5742-2157; b. http://orcid.org/0000-0001-6548-1384

BY-NC © 2018 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

Types of outcomes in clinical research

Juliana Carvalho Ferreira[1,2], Cecilia Maria Patino[1,3]

PRACTICAL SCENARIO

In a randomized trial evaluating the efficacy of a new drug for pulmonary arterial hypertension (PAH), patients were randomly assigned to receive the new drug or a placebo. The primary composite outcome was the time to the first PAH-related event (worsening of symptoms, initiation of treatment with prostanoids, lung transplantation, or atrial septostomy) or to death. Secondary outcomes included changes in the 6-minute walk distance (6MWD) and adverse events.

DEFINITIONS

Outcomes (also called events or endpoints) are variables that are monitored during a study to document the impact that a given intervention or exposure has on the health of a given population. Typical examples of outcomes are cure, clinical worsening, and mortality. The primary outcome is the variable that is the most relevant to answer the research question. Ideally, it should be patient-centered (i.e., an outcome that matters to patients, such as quality of life and survival).

Secondary outcomes are additional outcomes monitored to help interpret the results of the primary outcome: in our example, an increase in the 6MWD is inversely associated with the need for lung transplantation. They can also provide preliminary data for a larger study. For example, a preliminary trial that uses 6MWD as the primary outcome may include mortality as a secondary outcome if the power of the study to detect a difference in mortality is low. Although investigators may be tempted to monitor several outcomes, the effort and cost to monitor various outcomes may be prohibitive. Therefore, it is essential to decide which outcome(s) to monitor (Table 1).

Surrogate outcomes are biomarkers intended to substitute for a clinical outcome, for example, 6MWD as a marker of disease severity in PAH. Surrogate outcomes are typically continuous variables and occur earlier than does the clinical outcome, reducing costs, study duration, and size. Surrogates are commonly used as the primary outcome in phase I and II clinical trials. However, they may lead to false interpretations of the efficacy of the intervention if the surrogate is not a very good predictor of the clinical outcome.

Composite outcomes are made up of multiple variables. In our practical scenario, the primary outcome was composed of several clinical outcomes related to disease progression. Using composite outcomes has the advantage of increasing the power of the study when each of the events is rare and when events are competitive (patients who die cannot have a lung transplant). However, the interpretation of results can be misleading: if the intervention reduces the occurrence of the composite outcome, it does not necessarily mean that it reduces the occurrence of all of its components.

IMPORTANT CONSIDERATIONS

- The study outcomes should be stated *a priori* (before the researcher looks at the results) in order to avoid the risk of drawing false conclusions by testing every possible variable until one is statistically significant.
- The sample size calculation should be carried out to detect a clinically relevant effect of the intervention on the primary outcome, although calculations can also be made for secondary outcome variables, which may increase the sample size but also increase trial validity.
- More importantly, the choice of the most suitable outcome should be based on the research question and the corresponding hypothesis.

Table 1. Types of outcomes.

Outcome	Patient-centered	Composite	Surrogate
Asthma	Asthma control (questionnaire)	Hospitalization or a > 20% decline in asthma control	FEV_1, peak flow, eosinophils
PAH	2-year survival	Lung transplantation or death	6MWD, PASP
ARDS	Hospital survival	Time to extubation or tracheotomy	PaO_2/FiO_2 ratio, ventilator-free days

PAH: pulmonary arterial hypertension; 6MWD: six-minute walk distance; and PASP: pulmonary artery systolic pressure.

RECOMMENDED READING

1. Pulido T, Adzerikho I, Channick RN, Delcroix M, Galiè N, Ghofrani HA, et al. Macitentan and morbidity and mortality in pulmonary arterial hypertension. N Engl J Med. 2013;369(9):809-18. https://doi.org/10.1056/NEJMoa1213917
2. Haynes B, Sackett DL, Guyatt GH, Tugwell P. The tactics of performing therapeutic trials. In: Haynes B, Sackett DL, Guyatt GH, Tugwell P. Clinical Epidemiology: How to Do Clinical Practice Research. 3rd ed. Philadelphia: Lippincott, Williams and Wilkins, 2005.
3. Patino CM, Ferreira JC. Developing research questions that make a difference. J Bras Pneumol. 2016;42(6):403. http://dx.doi.org/10.1590/s1806-37562016000000354

1. Methods in Epidemiologic, Clinical and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração – InCor – Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

BY-NC © 2017 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

J Bras Pneumol. 2016;42(2):162-162
http://dx.doi.org/10.1590/S1806-37562016000000114

What is the importance of calculating sample size?

Cecilia Maria Patino[1,2], Juliana Carvalho Ferreira[1,3]

PRACTICAL SCENARIO

In a controlled, randomized clinical trial on the management of asthma in pregnant women, researchers evaluated the effect of implementing a program in which a portable device was used for asthma control, as assessed by the Asthma Control Questionnaire (ACQ). In clinical research, our goal is to make an inference regarding something about a population by studying a sample of that population. This sample has to be representative of the target population, and the number of participants must be appropriate. It should be large enough that the probability of finding differences between groups by mere chance is low and that of detecting true, clinically significant differences is high. However, the number of participants should not be so large that resources are wasted or participants are exposed to unnecessary risk. Therefore, in the study design phase, it is essential to perform sample size calculation. To perform this calculation, one must define the key characteristics of the study, such as the study design, the primary endpoint, the expected variability, the degree of certainty desired, and the predicted number of participants who will drop out of the study. To define these parameters and calculate the ideal sample size, we need to obtain deep knowledge of the field of research in question by reviewing the literature and biostatistics.

In our example, the researchers tested the effect that using a new device had on asthma control (the primary outcome) compared with the usual treatment. They estimated that the difference between the groups would be 0.55 points on the ACQ, with a standard deviation of 0.66 points, a power of 80%, and a level of significance of 5%. In addition, they estimated that 25% of the participants could be lost to follow-up. Using those data, the authors calculated that they needed to include 72 participants. At the end of the study, the researchers analyzed the results of 69 participants and showed that the new intervention improved asthma control in pregnant women.

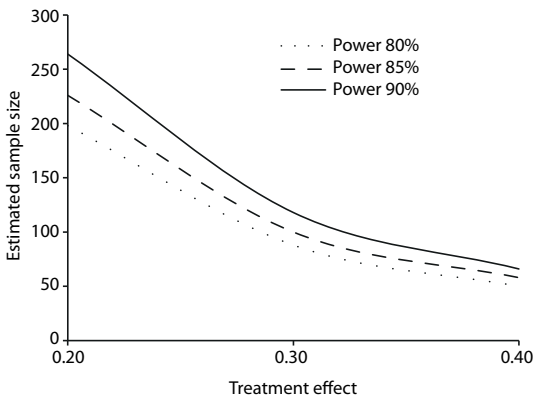

Figure 1. Relationship between the size of the treatment effect and the estimated sample size. On the x axis, we show hypothetical values for the size of the treatment effect, expressed as scores on the asthma symptoms questionnaire. We considered a fixed degree of variability (standard deviation) of 0.5 points and a level of significance of 5%. As the size of the treatment effect increases, the estimated size of the sample decreases. It is also clear that for the same effect size, choosing a higher power to detect the effect of the treatment causes an increase in the sample size.

BASIC CONCEPTS

Power

In biostatistics, power is defined as the probability of obtaining a statistically significant result when there is a real difference between treatments. In general, a power of at least 80% is needed in order to ensure a high probability of observing the effect, if any, of the intervention. To increase the power to detect differences, it is necessary to increase the sample size (Figure 1).

Critical level of significance

The critical level of significance is usually ≤ 5%. If we want greater certainty that a difference observed in the study population is not coincidental, we need to increase the sample size.

Effect size and variability

The greater the effect of the new intervention on the outcome is, the smaller is the sample size needed in order to prove it. Conversely, to show smaller effects, it is necessary to increase the sample size. If there is great variability of the effect in the population, we will also need a larger sample size (Figure 1).

It should be borne in mind that the sample size calculation is based on estimates and assumptions that can be inaccurate and is therefore subject to error. It is also important to be realistic when choosing the estimates employed in calculating the sample size. Highly optimistic choices about the effect size increase the risk of calculating an insufficient number of participants for the sample, whereas highly pessimistic choices can make the study unviable by resulting in a sample size that is too large to be practical.

RECOMMENDED READING

1. Zairina E, Abramson MJ, McDonald CF, Li J, Dharmasiri T, Stewart K, et al. Telehealth to improve asthma control in pregnancy: a randomized controlled trial. Respirology. 2016 Mar 31. doi: 10.1111/resp.12773 [Epub ahead of print]
2. Hulley SB, Cummings SR; Browner WS; Grady DG. Delineando a Pesquisa Clínica. 3rd ed. Porto Alegre: Artmed; 2008.
3. Ferreira JC, Patino CM. What does the p value really mean? J Bras Pneumol. 2015;41(5):485. http://dx.doi.org/10.1590/S1806-37132015000000215

1. Divisão de Pneumologia, Instituto do Coração – InCor – Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brasil.
2. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
3. Methods in Epidemiologic, Clinical and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax.

BY-NC © 2016 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

J Bras Pneumol. 2018;44(3):183-183
http://dx.doi.org/10.1590/S1806-37562018000000164

Internal and external validity: can you apply research study results to your patients?

Cecilia Maria Patino[1,2,a], Juliana Carvalho Ferreira[1,3,b]

CLINICAL SCENARIO

In a multicenter study in France, investigators conducted a randomized controlled trial to test the effect of prone vs. supine positioning ventilation on mortality among patients with early, severe ARDS. They showed that prolonged prone-positioning ventilation decreased 28-day mortality [hazard ratio (HR) = 0.39; 95% CI: 0.25-0.63].[1]

STUDY VALIDITY

The validity of a research study refers to how well the results among the study participants represent true findings among similar individuals outside the study. This concept of validity applies to all types of clinical studies, including those about prevalence, associations, interventions, and diagnosis. The validity of a research study includes two domains: internal and external validity.

Internal validity is defined as the extent to which the observed results represent the truth in the population we are studying and, thus, are not due to methodological errors. In our example, if the authors can support that the study has internal validity, they can conclude that prone positioning reduces mortality among patients with severe ARDS. The internal validity of a study can be threatened by many factors, including errors in measurement or in the selection of participants in the study, and researchers should think about and avoid these errors.

Once the internal validity of the study is established, the researcher can proceed to make a judgment regarding its external validity by asking whether the study results apply to similar patients in a different setting or not (Figure 1). In the example, we would want to evaluate if the results of the clinical trial apply to ARDS patients in other ICUs. If the patients have early, severe ARDS, probably yes, but the study results may not apply to patients with mild ARDS**.** External validity refers to the extent to which the results of a study are generalizable to patients in our daily practice, especially for the population that the sample is thought to represent.

Lack of internal validity implies that the results of the study deviate from the truth, and, therefore, we cannot draw any conclusions; hence, if the results of a trial are not internally valid, external validity is irrelevant.[2] Lack of external validity implies that the results of the trial may not apply to patients who differ from the study population and, consequently, could lead to low adoption of the treatment tested in the trial by other clinicians.

INCREASING VALIDITY OF RESEARCH STUDIES

To increase internal validity, investigators should ensure careful study planning and adequate quality control and implementation strategies—including adequate recruitment strategies, data collection, data analysis, and sample size. External validity can be increased by using broad inclusion criteria that result in a study population that more closely resembles real-life patients, and, in the case of clinical trials, by choosing interventions that are feasible to apply.[2]

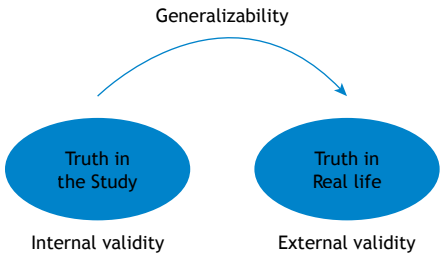

Figure 1. Internal and external validity.

REFERENCES

1. Guérin C, Reignier J, Richard JC, Beuret P, Gacouin A, Boulain T, et al. Prone positioning in severe acute respiratory distress syndrome. N Engl J Med. 2013;368(23):2159-68. https://doi.org/10.1056/NEJMoa1214103
2. Akobeng AK. Assessing the validity of clinical trials. J Pediatr Gastroenterol Nutr. 2008;47(3):277-82. https://doi.org/10.1097/MPG.0b013e31816c749f

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
a. http://orcid.org/0000-0001-5742-2157; b. http://orcid.org/0000-0001-6548-1384

BY-NC © 2018 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

Intercurrent events in clinical research: the norm, not the exception

L. Paloma Rojas-Saunero[1,2], Cecilia María Patino[1,3], Juliana Carvalho Ferreira[1,4]

PRACTICAL SCENARIO

A pulmonary and critical care team designed a randomized controlled clinical trial (RCT) to evaluate the efficacy and safety of a new drug vs. standard of care on reducing the number of hospitalizations due to COPD exacerbations within one year among adult patients. Both interventions (new drug and standard of care) required daily inhaler use throughout the follow-up period; therefore, daily adherence was monitored. Importantly, researchers anticipated a nonpreventable problem related to longitudinal studies: participants could die due to common comorbidities, which constitutes an intercurrent event.

INTERCURRENT EVENTS

In RCTs, intercurrent events are defined as events that occur after treatment initiation and either affect the ability to measure the intervention of interest or prevent the occurrence of an outcome over follow-up (Figure 1). Adverse events that lead to study arm crossover or discontinuation of the assigned treatment are considered intercurrent events, because they prevent the continuation of the assigned intervention. Since these events impact the interpretation of study results, we must consider them when defining the research question, study design, and analysis.

COMPETING EVENTS—WHEN THE OUTCOME CANNOT OCCUR

Competing events are a particular type of intercurrent events; they prevent the study outcome from occurring.[1] In our example, participants who died due to other reasons and prior to a hospitalization due to a COPD exacerbation could no longer experience the study outcome. Consequently, death determines that the risk of a COPD exacerbation-related hospitalization for these participants is zero. Therefore, when calculating the measure of effect, such as risk difference or risk ratio between the two study arms, we must be careful when interpreting the results, because part of the effect of the new intervention when compared with the existing one may be explained by how and/or if the intervention affected the competing event.

CENSORING EVENTS—WHEN THE OUTCOME CANNOT BE MEASURED

Competing events can also be defined as a censoring event in some settings. Censoring events are those that may prevent investigators from measuring the outcome, as opposed to preventing it from happening. For example, when participants move to another geographical area and follow-up is interrupted, investigators cannot determine whether those participants were hospitalized, died for other reasons, or remained outcome-free. Censoring

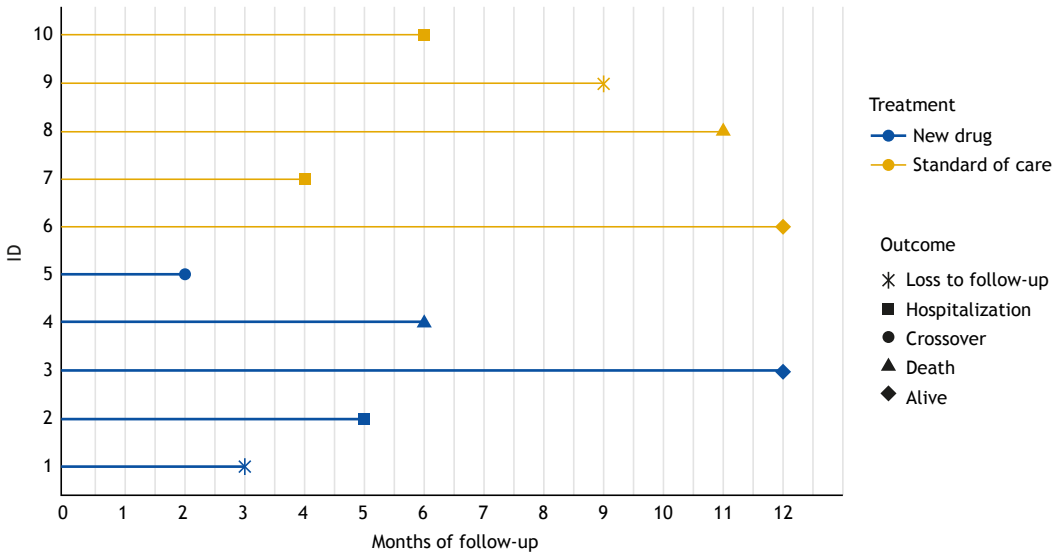

Figure 1. Illustration of different intercurrent events that may happen over follow-up in a controlled trial (or in an observational study). Each identification (ID) represents the path followed by a different participant.

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Department of Epidemiology, University of California Los Angeles, Los Angeles (CA) USA.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.
4. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2022 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3756

relies on the assumption that participants who were lost to follow up (censored) share similar measured and unmeasured demographic and clinical characteristics with those who remained in the study. This assumption is called the "independent censoring assumption", and it requires extensive expertise on the topic of interest to evoke it.[1]

In our example, defining death due to other causes as a censoring event would require that we conceptualized this event as preventable by study design (just as loss to follow-up) and assumed that those who died are demographically and clinically comparable to those who remained in the study. Similarly, having a lung transplant could be considered a competing event, because lung transplant patients no longer have COPD and thus cannot be hospitalized for COPD exacerbations. However, treating lung transplant as a censoring event could be reasonable in some settings.[2]

Careful consideration of the impact of intercurrent events on study results is necessary when designing analytical approaches in order to provide accurate and reliable results to inform patient care and health policies.

KEY POINTS

- Identify all potential intercurrent events that can occur over follow-up when designing RCTs or longitudinal observational studies
- Clear definitions of intercurrent events help
  - refine research questions
  - identify data that need to be collected longitudinally to account for intercurrent events
  - facilitate result interpretation and implications
- Report frequency (absolute and relative numbers) of intercurrent events across the variable
- Consult with an expert when conducting longitudinal studies with competing events

REFERENCES

1. Geskus RB. Data Analysis with Competing Risks and Intermediate States. 1st ed. Boca Raton (FL): Chapman & Hall/CRC Biostatistics Series; 2016.
2. van Geloven N, le Cessie S, Dekker FW, Putter H. Transplant as a competing risk in the analysis of dialysis patients. Nephrol Dial Transplant. 2017;32(suppl_2):ii53-ii59. https://doi.org/10.1093/ndt/gfx012

Loss to follow-up and missing data: important issues that can affect your study results

Juliana Carvalho Ferreira[1,2,a], Cecilia Maria Patino[1,3,b]

Investigators used data from the São Paulo State Tuberculosis Control Program database to evaluate the association between the place of diagnosis and treatment outcomes. They reported that 25% of new tuberculosis cases were diagnosed in emergency facilities and that, in comparison with patients diagnosed in outpatient settings, they were more likely to have unsuccessful treatment outcomes, including loss to follow-up (in 12%) and death (in 10%).[1]

LOSS TO FOLLOW-UP AND TYPES OF MISSING DATA

To study associations of an exposure (in cohort studies) or of an intervention (in randomized controlled trials) with clinical outcomes prospectively, investigators collect exposure/intervention information at study entry and then follow study participants over time and collect outcome data. If follow-up is incomplete or interrupted, leading to missing data at the end of the study, this could impact the internal validity of the study. Participants with missing data, compared with those with complete data, may differ systematically, for example when loss to follow-up is related to the death of participants.

Missing data occur for multiple reasons: 1. the variable of interest is not measured by the research team (e.g., forgetting to measure weight at baseline); 2. the study participant misses a scheduled study visit or test; or 3. the variable is measured but the team fails to register the variable value on the data collection form. However, loss to follow-up, where patient data is not available until the end of the follow-up period, is the most critical mechanism of missing data, because it might include missing outcome data, which is crucial to answer the research question.

Missing data are classified as missing completely at random (MCAR), when missingness is not related to the exposure, covariates, or the outcome; missing at random (MAR), when missingness is related to the exposure or confounders, but not the outcome; and missing not at random (MNAR), when missingness might be related to the outcome (Chart 1).[2] Loss to follow-up and missing data can threaten the internal validity of a study even if the mechanism is MCAR, in which we consider that the remaining participants are a random sample of the initial study population, because the study power will be decreased. If the mechanism is MAR, adjustments and imputation methods can be used, but that might introduce biases to the study. If the mechanism is MNAR, there is a serious risk of biased results.[3]

HOW TO DEAL WITH LOSS TO FOLLOW-UP AND MISSING DATA

The best strategy to avert missing data is to prevent loss to follow-up. Designing the study carefully, training staff, implementing data quality procedures, and developing mechanisms to retain and contact participants are key. Additionally, there are statistical methods available to deal with missing data, but these procedures should be planned a priori and with consultation of a biostatistician.[3] However, there are situations when investigators might not overcome problems related to missing data because the mechanism of missingness is MNAR. In that case, losses to follow-up of 20%, for example, can result in serious biases and, therefore, should not be considered "acceptable". Remember, missing data are common and best practices include thinking about it early on when defining the research question and writing the protocol.

Chart 1. Types of missing data and strategies to minimize them.

Type of missing data	Example	Strategies to minimize missing data
MCAR	Participant moves to another state and abandons the study; a test result is lost at the lab	Develop standardized collection forms; monitor data quality; keep participant contact information up to date
MAR	In a cohort of COPD patients, participants with mild disease are more likely to abandon the study because they are asymptomatic	Offer benefits and incentives to retain participants; regularly contact participants; conduct a pilot study to identify risk factors for loss to follow-up; and develop strategies to overcome them
MNAR	Loss to follow-up is higher among tuberculosis patients who have serious adverse events due to tuberculosis drugs than among patients who tolerate treatment, and treatment nonadherence is related to death	Offer adequate support for study participants; develop strategies to retain participants with a high risk of loss to follow-up; and develop alternative methods to measure the outcome even for participants lost to follow up

MCAR: missing completely at random; MAR: missing at random; and MNAR: missing not at random.

REFERENCES

1. Ranzani OT, Rodrigues LC, Waldman EA, Prina E, Carvalho CRR. Who are the patients with tuberculosis who are diagnosed in emergency facilities? An analysis of treatment outcomes in the state of São Paulo, Brazil. J Bras Pneumol. 2018;44(2):125-133. https://doi.org/10.1590/s1806-37562017000000384
2. Pedersen AB, Mikkelsen EM, Cronin-Fenton D, Kristensen NR, Pham TM, Pedersen L, et al. Missing data and multiple imputation in clinical epidemiological research. Clin Epidemiol. 2017;9:157-166. https://doi.org/10.2147/CLEP.S129785
3. Kristman V, Manno M, Côté P. Loss to follow-up in cohort studies: how much is too much? Eur J Epidemiol. 2004;19(8):751-60.

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
a. http://orcid.org/0000-0001-5742-2157; b. http://orcid.org/0000-0001-6548-1384

BY-NC © 2019 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

J Bras Pneumol. 2016;42(5):310-310
http://dx.doi.org/10.1590/S1806-37562016000000296

Randomization: beyond tossing a coin

Juliana Carvalho Ferreira[1,2], Cecilia Maria Patino[1,3]

INTRODUCTION

Randomization is a research strategy used in order to increase the validity of clinical trials evaluating the effect of interventions (e.g., drugs or exercise). It involves the random allocation of participants to either intervention or control groups and requires that participants have an equal chance of being allocated to either group. When properly implemented, randomization prevents selection bias and produces comparable study groups in terms of known and unknown baseline risk factors. For randomization to work, investigators and participants must be unable to predict to which group each of the participants will be allocated—this is called allocation concealment; in addition, investigators must be unable to change the allocation of any participant after randomization.

COMMONLY USED RANDOMIZATION STRATEGIES

Simple randomization is equivalent to tossing a coin: a new participant has an equal chance of being assigned to intervention or control groups, independently of previous assignments. Instead of tossing a coin, however, a randomization list is generated by a computer and used to prepare sequentially numbered, sealed envelopes, or, preferably, that list is administered by a central telephone service or website. The advantages of simple randomization are that it is inexpensive and easy to implement. The disadvantages include the risk of producing imbalances in the number of participants in the groups, as well as in the distribution of baseline risk factors, in studies with small sample sizes (N < 100; Figure 1).

In block randomization, the randomization list is a random sequence of blocks of participants instead of individual participants. The blocks have a pre-determined size; for example, four participants in one block, with six possible intervention and control sequences. This strategy ensures that intervention and control groups are balanced in terms of the number of participants (Figure 1). To ensure allocation concealment using this method, random variation of block sizes should be used (four to eight participants per block).

Stratified randomization is an alternative when balance for key baseline risk factors is desired. Each new participant is first classified into strata according to baseline characteristics (e.g., age or disease severity), and each stratum has a separate randomization list. Thereafter, once the participants are categorized into their stratum, they are randomized to either the intervention or the control groups. Stratification should be carried out using few relevant strata in order to work well. Stratified and block randomization strategies can be combined so that patients are first categorized into a stratum and then randomized in blocks.

Adaptive randomization uses computer algorithms that take into consideration baseline risk factors and the allocation of previous participants to allocate the next participant. The advantage of this method is that it accommodates more baseline risk factors than stratification and produces optimized group balance at the same time. However, it is more complex and requires a web-based randomization center available 24 h a day.

HOW TO CHOOSE

Simple randomization is easy to implement, is inexpensive, and can be a good option for large trials (N > 200). Block randomization is a good option when balance in the number of participants in each group is desired. Stratification is a good option to provide balance for important covariates. Adaptive randomization methods may be a good option when the trial structure includes statisticians and information technology support. For all methods, adequate implementation is paramount to ensure allocation concealment and to prevent manipulation and selection bias.

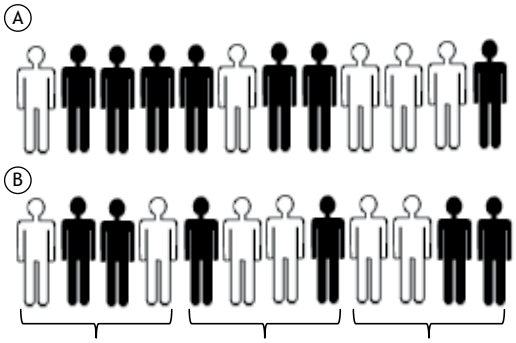

Figure 1. A) Simple randomization of 12 participants (black for intervention, white for control). This random sequence resulted in 7 subjects assigned to intervention and 5 to the control group; B) Block randomization of 12 participants with blocks of 4, resulting in 6 participants in each group.

REFERENCES

1. Kang M, Ragan BG, Park JH. Issues in outcomes research: an overview of randomization techniques for clinical trials. J Athl Train. 2008;43(2):215-21. http://dx.doi.org/10.4085/1062-6050-43.2.215
2. Vickers AJ. How to randomize. J Soc Integr Oncol. 2006;4(4):194-8. http://dx.doi.org/10.2310/7200.2006.023

1. Methods in Epidemiologic, Clinical and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax. Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração – InCor – Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

BY-NC © 2016 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

Dealing with confounding in observational studies

Cristiane Fumo-dos-Santos[1], Juliana Carvalho Ferreira[1,2]

PRACTICAL SCENARIO

Investigators in a large academic center in São Paulo, Brazil, examined the association between the use of protective ventilation, defined as a tidal volume < 8 mL/kg of predicted body weight and plateau pressure < 30 cmH_2O, and survival in patients with severe COVID-19. They also collected data about severity of disease at ICU admission, need for renal replacement therapy, and several ventilatory parameters. They found that the use of protective ventilation was associated with improved survival, with an adjusted hazard ratio of 0.73 (95% CI, 0.57-0.94; p = 0.013).

CAUSAL INFERENCE IN OBSERVATIONAL STUDIES

In epidemiological studies, investigators do not assign interventions, but rather classify individuals as exposed or non-exposed to risk factors for developing an outcome. When a statistically significant association is found, several possible explanations need to be considered:

1. The association is real, and the predictor (protective ventilation, in our example) is truly a cause of the outcome (survival, in our example).
2. The association is real, but it is an effect-cause relationship: the outcome (survival) causes the predictor (protective ventilation). In this example it would not be plausible to consider this possibility, but there are many cases that this makes sense.
3. The association is due to chance—random error. Because we usually consider a p value < 0.05 as significant, and the p value in our example was 0.013, there is 1.3% probability that chance is the explanation for this association.
4. The association is not real, it is the result of a systematic error (bias), resulting from methodological aspects of the study, such as systematically underestimating the predicted body weight of patients.
5. The association is real, but it is confounded by the effect of other(s) variable(s) associated with both the outcome and the predictor.

WHAT IS CONFOUNDING?

Confounding derives from the Latin *confundere*, to mix. The classical definition of a confounder is any third variable that is associated with the exposure of interest, that is a cause of the outcome of interest, and that does not reside in the causal pathway between exposure and outcome (Figure 1A). For example, in our practical scenario, the investigators considered that lung compliance, among other variables, was a potential confounder, because low lung compliance is a cause of death (therefore, reducing survival), and it is also associated with the predictor—when compliance is very low, it may be more challenging to apply protective ventilation. Severity of disease at admission, on the other hand, was not treated as a confounder by the investigators, because although it is highly associated with the outcome (death), it does not have a causal relationship with the predictor of interest (protective ventilation).[1] Even though we can have confounders in experimental research, it is a more important issue to be considered in observational studies.[2]

WHY SHOULD WE CARE ABOUT IDENTIFYING CONFOUNDERS?

Confounders can lead to overestimation or underestimation of the effect of the main predictor on the outcome of interest, making the effect not reliable and interfering with our ability to draw causal inferences in observational studies.[2] Therefore, statistical strategies are recommended to control for or to adjust the analysis for confounders in order to observe the true, isolated effect of the predictor of interest on the outcome.

We should not identify a confounder based on statistical testing but on prior clinical knowledge or on the pathophysiology of the process that we are studying.[1] One of the most accepted strategies to identify a confounder is using prior knowledge about the outcome of interest to build causal models, especially graphical criteria.[2] This approach is important because the traditional way to identify the confounder, as described earlier, is often inadequate in more complex structures.[3]

HOW CAN WE DEAL WITH CONFOUNDERS?

The best way to deal with confounders is to plan in advance. A randomized controlled trial randomly assigns individuals to the intervention and control arms of the study, dispersing the known and unknown confounders into each arm. However, this design is not suitable to answer many important research questions.[1]

Selecting individuals with the same characteristic is also a strategy: to study reduced lung function in asthma, researchers may exclude people with obesity. The problem with this strategy is that the results do not apply to all individuals with asthma, but only for non-obese asthma patients. Another way to deal with confounding

1. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2023 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3756

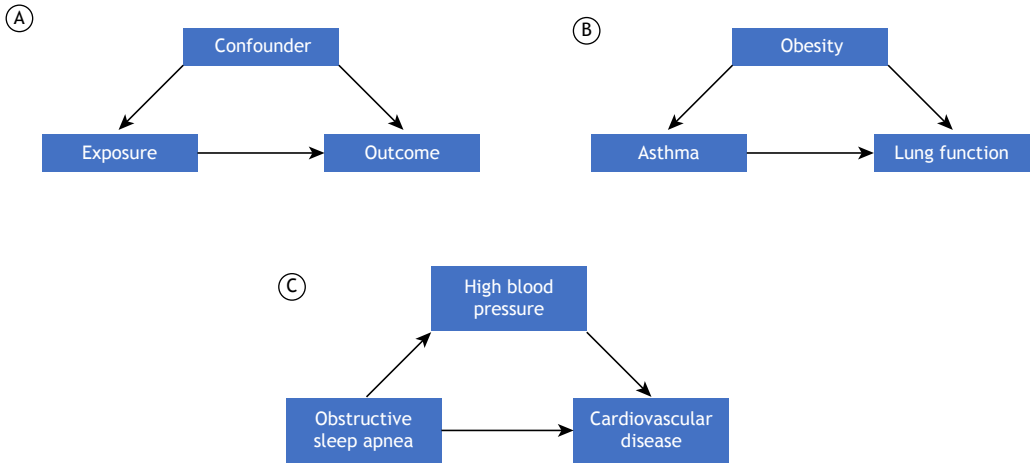

Figure 1. In A, representation of the confounding pattern: the variable is related to exposure, it is a cause of the outcome, and it is not in the causal pathway between the main exposure and the outcome of interest. In B, obesity is a confounder in the relation between asthma and lung function since obesity may worsen asthma and may cause a reduction in lung function. Adjusting for obesity is advised in this scenario. In C, the model represents a mediation effect —obstructive sleep apnea may lead to cardiovascular disease (direct effect), but obstructive sleep apnea may also lead to high blood pressure, which causes cardiovascular disease (indirect effect). In this case, adjusting for high blood pressure is not appropriate.

is matching individuals: the researcher selects the same number of participants with and without obesity both in the exposed and in the non-exposed group.[1] Again, however, the manipulation results in reduced generalizability of the results.

The most commonly used strategy to deal with confounders is controlling (or adjusting) for confounders during the statistical analysis since regression models can address several predictors at the same time.[3] In this case, it is really important to build a causal model and adjust only for confounders, instead of adjusting for all variables based on p values, for example.

The main message is that confounders can interfere with causal inference in observational studies, and we need to plan ahead to identify, measure, minimize, and adjust for confounders in order to use the results of observational studies to guide future research and clinical decision making.

REFERENCES

1. van Stralen KJ, Dekker FW, Zoccali C, Jager KJ. Confounding. Nephron Clin Pract. 2010;116(2):c143-c147. https://doi.org/10.1159/000315883
2. Santos S, Zugna D, Pizzi C, Richiardi L. Sources of confounding in life course epidemiology. J Dev Orig Health Dis. 2019;10(3):299-305. https://doi.org/10.1017/S2040174418000582
3. Lederer DJ, Bell SC, Branson RD, Chalmers JD, Marshall R, Maslove DM, et al. Control of Confounding and Reporting of Results in Causal Inference Studies. Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals [published correction appears in Ann Am Thorac Soc. 2019 Feb;16(2):283]. Ann Am Thorac Soc. 2019;16(1):22-28. https://doi.org/10.1513/AnnalsATS.201808-564PS

What is a manual of procedures and why do we need one?

Juan C Calderon[1,2,3], Juliana C Ferreira[1,4], Cecilia M Patino[1,5]

PRACTICAL SCENARIO

At the beginning of this millennium, the PLATINO multi-country cross-sectional study was planned by a group of researchers to measure the burden of COPD among adults from major cities in Latin America (São Paulo, Mexico City, Montevideo, Caracas, and Santiago).[1] The investigators measured the prevalence of COPD and its risk factors using a standardized paper questionnaire and standardized pulmonary function methodology. The authors reported that the prevalence of COPD was higher than expected and ranged from 7.8% (95% CI, 5.9-9.7) in Mexico City to 19.7% (95% CI, 17.2-22.2) in Montevideo, Uruguay.[2] To yield accurate results, investigators collected data across study sites using the same methodological standards by writing and implementing a manual of procedures (MOP).[1]

WHAT IS A MOP?

Although the research process begins with an idea derived from an observation, the next steps include writing a research question, a protocol, and a detailed manual of procedures to guarantee accurate data collection process, as well as a comprehensive data analysis plan to provide reliable answers to important research questions that impact population health. At each step of the research protocol, inconsistencies in data collection procedures can lead to excess variation and/or error, jeopardizing the validity and reliability of the results. Therefore, the MOP is necessary to document and standardize all research procedures. The research process evolves from the initial research protocol into a fully detailed operations manual, describing every procedure stated in the protocol, and encompasses study organization, policies, participant recruitment and enrollment, randomization, blinding procedures, variable measurements, quality control, data management practices, and the statistical plan (Table 1).

All types of research study designs need a MOP, but MOPs are especially relevant for randomized clinical trials and cohort studies, because the types of procedures

Table 1. Outline of a manual of procedures (MOP).

Procedure	Example
Introduction	Overview of study goals and rationale
Study protocol	In many cases, the study protocol is the first document included in the MOP
Organization	Description of the team of investigators and research study units in multicentric studies
Recruitment and enrollment procedures	Procedures to identify and recruit potentially eligible patients and check for inclusion and exclusion criteria; informed consent
Randomization & blinding	Procedure used to randomize patients using a website
Ethical considerations	Protection of participant confidentiality and privacy Compliance with ethical guidelines and regulations
Clinic visits	Which variables are to be measured at baseline and at specific time-points during follow-up and how they will be measured
Intervention/exposure/predictors & comparison group	Detailed description of the intervention and the control groups (for unblinded studies)
Study variables	Detailed instructions on how the primary outcome variable and other important variables will be measured, including adverse events that will be measured
Quality control	Responsibilities, training of data collection team, and equipment calibration and maintenance
Data management	How data will be collected and stored; confidentiality; plan for backups
Data analysis	A detailed data analysis plan
Emergency procedures	Protocols for managing medical emergencies during study visits
Communication plan	Procedures for communication among study staff, investigators, and participants Contact information for key study personnel
Appendices	Questionnaires, forms

1. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Universidad Espíritu Santo, Samborondón, Ecuador.
3. Respiralab Research Group, Guayaquil, Ecuador.
4. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
5. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.

CC BY © 2024 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3756

used in these studies are complex and need to be well standardized, but descriptive studies are not exceptions.[1] In the case of the PLATINO study, despite cultural, social, and economic disparities among Latin-American cities, the use of a MOP, which included standardized instructions on how to perform a spirometry test, enabled the execution of high-quality research, addressing crucial research questions and establishing the prevalence and severity of COPD in major cities of Latin America.

WHAT GOES IN A MOP?

All study procedures should be included in the MOP. In the PLATINO study, for example, there was a section dedicated to training and certification of study staff to perform spirometry. In addition, there was detailed description on how to perform an acceptable spirometry test and the accepted variability for selected maneuvers. Even simpler procedures, such as the technique to measure anthropometric variables were detailed, as well as the sampling procedure. As individuals were recruited from Portuguese and Spanish speaking cities, the questionnaires were validated for both languages. The MOP meticulously describes procedures, anticipates data variability, addresses recruitment and follow-up errors, and delineates strategies to minimize bias and maximize quality control, assuring validity and generalizability to the target population.

TIPS TO WRITE A MOP

- Describe in depth all the procedures used in the study, like a recipe.
- Be systematic and anticipate all possible misunderstandings; be accurate.
- Ask for feedback from the study team to check for any errors.
- Implement strategies to minimize errors and plan for possible solutions.
- Remember, all investigators and study staff will consult the MOP and are expected to follow the instructions.

REFERENCES

1. Menezes AM, Victora CG, Perez-Padilla R; PLATINO Team. The Platino project: methodology of a multicenter prevalence survey of chronic obstructive pulmonary disease in major Latin American cities. BMC Med Res Methodol. 2004;4:15. https://doi.org/10.1186/1471-2288-4-15
2. Menezes AM, Perez-Padilla R, Jardim JR, Muiño A, Lopez MV, Valdivia G, et al. Chronic obstructive pulmonary disease in five Latin American cities (the PLATINO study): a prevalence study. Lancet. 2005;366(9500):1875-1881. https://doi.org/10.1016/S0140-6736(05)67632-5

Choosing wisely between randomized controlled trials and observational designs in studies about interventions

Juliana Carvalho Ferreira[1,2], Cecilia Maria Patino[2,3]

Randomized controlled trials (RCTs) are the gold standard for evaluating the efficacy of interventions, because they avoid key sources of bias by randomly allocating participants to the treatment or control. That feature of the study design makes RCTs the highest ranked type of study within the Evidence-Based Medicine framework grading system. However, not all questions about health interventions can be answered with an RCT. Observational studies may be more appropriate to study certain aspects about interventions and thus complement RCTs.

In some situations, it is unfeasible or unethical to randomize patients to a treatment, such as a surgical intervention, if surgeons are uncomfortable performing an unfamiliar procedure. In addition, observational studies are better suited to evaluate the incidence of adverse events of interventions because they have less strict inclusion and exclusion criteria, which allows a broader spectrum of the target population to be included. While RCTs are usually the best option to test efficacy (the effect of the intervention under ideal conditions), observational studies are a valuable option to evaluate effectiveness (the effect of an intervention in real life).

Some advantages of observational studies include the following: they are usually less expensive than RCTs, they have no ethical roadblocks in assigning participants to treatment or control groups, and placebos are rarely used (Table 1).

CHOOSING WISELY

The choice between an observational study and an RCT should be based on the specific research question. Observational designs are appropriate when it is reasonable to assume that characteristics that influence clinicians to choose a given intervention are not related to the study outcome. For example, in a comparison between the impact of radiosurgery and that of surgical lung resection on the survival of lung cancer patients, an observational study would not be appropriate, because the choice between radiosurgery and lung resection is influenced by tumor size and patient performance status, which also influence survival independently of the treatment option. In contrast, observational studies are often used to study the effectiveness of vaccination to protect against infectious diseases, because the characteristics that influence the decision to get vaccinated are not major determinants of the risk of being infected.

MINIMIZING BIAS

When conducting observational studies to test interventions, the investigator needs to design strategies to minimize bias resulting from imbalances in competing risk factors (confounders) across the intervention and control groups. In the design phase, a typical strategy involves measuring known confounders at baseline and later adjusting for those confounders during the analysis phase by using multivariable models. Another strategy includes combining confounding variables associated with the intervention and creating a new variable, called a propensity score, that can be used, for example, to match participants at baseline or adjust for confounders during analysis. However, the efficiency of such methods is limited to known and adequately measured confounders.

BEYOND STUDY DESIGN

When evaluating the medical literature, clinicians should consider not only the design (RCT or observational) but also the quality of a given study. RCTs and observational trials both contribute to advancing knowledge in health care, which can guide clinical decision-making and public health policy.

Table 1. Comparison between randomized controlled trials and observational studies.

Aspect	RCTs	Observational studies
Randomization	Yes	No
Risk of selection bias	Low	Can be high
Risk of imbalances in baseline risk factors	Low	High
Cost	++++	++
Complexity	++++	++
Duration	++	++++
Appropriate for evaluating efficacy	++++	++ to +++
Appropriate for evaluating effectiveness	+	++++
Appropriate for identifying adverse events	++ to +++	++++

RCTs: randomized controlled trials.

RECOMMENDED READING

1. Concato, J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med. 2000:342(25):1887-92. http://dx.doi.org/10.1056/NEJM200006223422507
2. Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ. 1996;312(7040):1215-8. http://dx.doi.org/10.1136/bmj.312.7040.1215

1. Divisão de Pneumologia, Instituto do Coração – InCor – Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (SP) Brasil.
2. Methods in Epidemiologic, Clinical and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.

BY-NC © 2016 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

Randomized controlled trials: advantages and pitfalls when studying causality

Diego Caruso[1,2], Juliana C Ferreira[2,3]

PRACTICAL SCENARIO

A pharmaceutical company has developed a new drug to improve asthma control and they are asking a respected team of investigators to design a study to compare "betteraline" (new drug) with "normalraline" (usual care). The investigators believe that the best design should be a randomized controlled trial (RCT) comparing both drugs and measuring the improvement in FEV_1 after three months of treatment as the main outcome. However, they are worried about costs, time commitment, and the need for an organized team to minimize follow-up losses, as well as about the logistics to measure the primary outcome. They wonder what pros and cons of performing an RCT are in this case.

In clinical and epidemiological research, analytical studies aim to assess the potential cause-effect association between an intervention and an outcome to ensure that causation is the best possible explanation among all available options.

To establish causation, the research question we would like to answer is: what would the outcome be if patients received an experimental intervention (factual scenario) compared with what would have happened if the same patients had received a control treatment, at the same moment of their lives, under identical conditions (counterfactual scenario)? Because we cannot test that in real life, the best substitute is to randomly select "similar" patients to receive either the intervention or the control and to compare outcomes. The outcome of the control group is the counterfactual scenario.[1] Although not perfect, this model served as the central concept inspiring the inception of randomized experiments and their statistical inference by Ronald Fisher circa 1920.

ADVANTAGES OF RCTS

RCT is a robust design because participants are randomly assigned to receive the intervention or control, which ensures that both known and unknown potential confounders are balanced at baseline in the two (or more) study groups. This process is achieved in two steps. First, the generation of a random list; second, allocation concealment, which is a procedure to prevent investigators from knowing to which group the next patient will be assigned. There are a few ways to do this, such as using sealed opaque envelopes or using digital automated response systems accessed by phone or over the Internet.

Any attempt to manipulate the process disrupts the balance that we are trying to achieve. Another advantage of RCTs is that measurement of variables during the study is prospective and ensures that all participants have measurements taken in the same manner throughout the study, avoiding information bias, minimizing missing data, and increasing internal validity.

Masking, when possible, is another advantage of RCTs. The participants, the researchers who follow the patients during the study, the researchers who are responsible for defining whether or not the participants experienced the outcome, and/or the statistician who analyzes the data may be prevented from knowing the assignment of each participant in order to minimize bias.

Performing an RCT requires a lot of preparation, with a carefully designed study protocol, a manual of procedures (for example, specific instructions to perform spirometry), a team, and an experienced leader. That takes time and money; therefore, a realistic schedule and budget are essential.

IMPORTANT CONSIDERATIONS AND PITFALLS

Participants in an RCT are not selected at random from the population of interest. They are usually referred by their doctors or self-referred by seeing advertisements or receiving recommendations from other patients, which might affect generalizability. In addition, the wonders of randomization are at the heart of RCTs, but like any vital organ, it can be affected by certain conditions:

- Crossover: patients who are assigned to one of the study arms but, due to unexpected reasons, receive the treatment of the other study arm. For instance, participants assigned to the intervention group obtain inhalers containing "normalraline" at a pharmacy.
- Nonadherence: some participants may not adhere to the assigned treatment. In our example, a patient may decide to stop using his/her asthma inhalers. If this proportion is high, or if it occurs more frequently in one arm than in another, it becomes a potential bias.
- Loss to follow-up: if a participant drops out of the study and cannot be contacted, it cannot be determined whether they experienced the study outcome or not, affecting the interpretation of the results.[2]

1. Hospital Dr. César Milstein, Buenos Aires, Argentina.
2. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2024 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3756

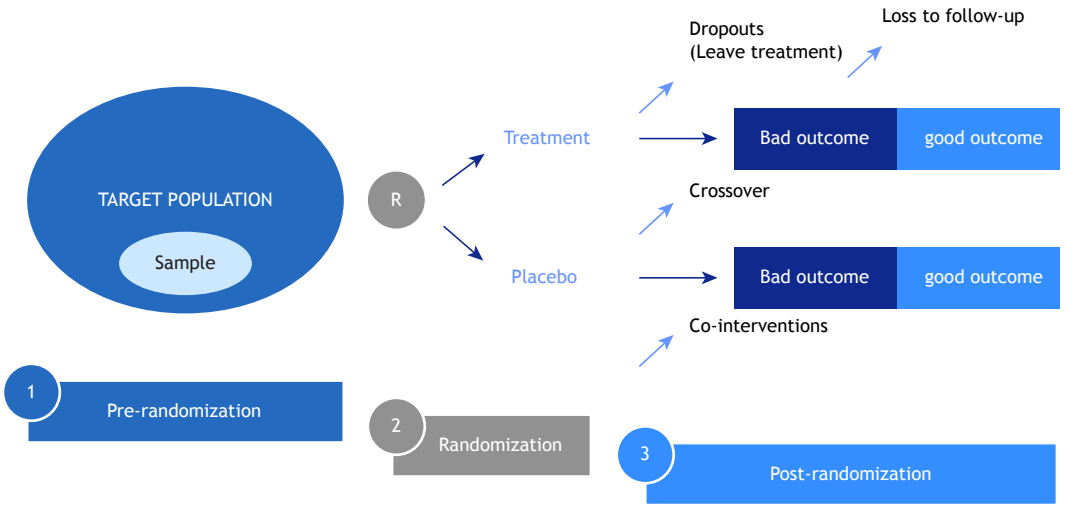

Figure 1. Framework and potential pitfalls in randomized clinical trials. Loss to follow up, dropout, co-interventions and crossover can happen in either of the study arms. R: randomization.

- Co-interventions: when participants receive interventions other than the main intervention, it may be difficult to know whether to attribute the benefit to the study intervention or to the co-intervention. In our example, the addition of corticosteroids to achieve asthma control is a co-intervention.

The investigators have decided to perform an RCT to test if "betteraline" is superior to usual care to treat asthma, because RCT is the most robust design to determine causality if all premises are met. To obtain valid results, the study will need careful planning, time, resources, and a dedicated team.

REFERENCES

1. Höfler M. Causal inference based on counterfactuals. BMC Med Res Methodol. 2005;5:28. https://doi.org/10.1186/1471-2288-5-28
2. Bell ML, Kenward MG, Fairclough DL, Horton NJ. Differential dropout and bias in randomised controlled trials: when it matters and when it may not. BMJ. 2013;346:e8668. https://doi.org/10.1136/bmj.e8668

Why is conducting pragmatic clinical trials so important?

Ignacio Esteban[1,2], Juliana Carvalho Ferreira[1,3], Cecilia Maria Patino[1,4]

PRACTICAL SCENARIO

Researchers in Finland have designed a randomized clinical trial (RCT) in which adults older than 65 years of age will be randomized (1:1) to receive either high-dose or standard-dose of a quadrivalent influenza vaccine. The main outcome is cardiorespiratory hospitalizations up to 6 months post-vaccination. They propose to use a pragmatic design and implement follow-up for up to 11 months post-vaccination using the Finnish national health registries. Here, we analyze the design of this pragmatic clinical trial (PCT) to discuss its importance in evidence-based decision-making.[1]

PCTS: ADVANTAGES AND DIFFERENCES FROM EXPLANATORY CLINICAL TRIALS

RCTs are the gold standard study design to determine the safety and efficacy of new interventions. However, the "ideal scenario" in which clinical trials are conducted may be far removed from the true needs and the decision-making process of health personnel and the population. RCTs can be classified as explanatory trials (also called phase III of drug development), in which the main objective is to confirm a clinical or physiological hypothesis in a very controlled environment, or as pragmatic trials, as part of the post-marketing phase or the so-called phase IV, with the objective of testing the new intervention in real-world scenarios, thus helping understand the true impact of the introduction of the drug or technology under study.[2]

Choosing a PCT over an explanatory clinical trial (ECT) depends on the stage of development of the intervention and the level of pragmatism desired to increase the generalizability of the results. This decision implies modifications to typical aspects of RCTs to improve the feasibility of the study.

ECTs are usually carried out in research centers with trained professionals, while PCTs can be carried out in multiple types of health care centers (hospitals, clinics, and private practices) and by different health professionals, several of whom without prior research training; in our example, the study takes place in over 40 health care stations.[1] This increases the generalizability of the results.

Participants in PCTs, as those in our example, tend to be a heterogeneous population with minimal criteria for

Table 1. Differences between explanatory clinical trials and pragmatic clinical trials.

	Explanatory clinical trials	Pragmatic clinical trials
General		
Objective	Efficacy and safety of a new intervention.	Effectiveness and long-term safety. Optimization of generalizability of trial results.
Recruitment	Active recruitment is needed.	Less strict. May utilize disease registries.
Participants	Highly selected (many exclusion criteria)	Similar to patients who would receive the intervention if it became standard of care.
Study design		
Delivery of the intervention	Requires frequent study visits for intervention administration and safety evaluations.	Trial procedures and data-collection requirements are minimized. Intervention is administered as in normal practice.
Safety endpoints	Precise collection and description of adverse events.	Long-term safety data in some cases, often less complex and similar to standard of care.
Randomization	Present. Removes significant differences between groups.	Can be more complex. Could be performed on a patient, cluster, or clinician level.
Risk of bias	Minimal.	Higher due to less control over other variables. Can be addressed.
Other		
Ethical considerations	Participants' informed consent, and Institutional Review Board and regulatory entities' approval are required.	Less complicated. Requirements may be waived in some cases
Funding	Usually by industry.	Variable. Co-financed by industry and government.

1. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. INFANT Foundation, Buenos Aires, Argentina.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
4. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.

BY-NC © 2022 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário

their selection and with a wider age range, whereas participants in ECTs are frequently more homogeneous and highly selected or share a common pathology.[2]

Both types of clinical trials use a control group; however, PCTs usually utilize another active arm group, many times a standard-of-care group instead of a placebo group. In our example, the high-dose group is being compared against the standard-dose group. Endpoints in PCTs are usually patient-centered such as deaths, hospitalizations, symptoms, disability, and quality of life, which facilitates data collection with a more flexible surveillance system.[2] The follow-up of participants in ECTs typically requires multiple visits to the study site, while the follow-up in PCTs is less strict; in our example, the researchers periodically collect registry data provided by electronic health records of the public health care system. Other major characteristics of both studies are summarized in Table 1.

Due to concerns about adherence and less stringent follow-up in PCTs, high-quality data collection, robust statistical design, and blinding when defining and adjudicating the study endpoint are key to obtaining reliable results.[2] The Pragmatic Explanatory Continuum Indicator Summary (PRECIS-2)[3] is a useful tool on how to conduct and increase the robustness of PCTs.

KEY POINTS

1. PCTs are increasingly in use in clinical research because they offer evidence about interventions under real-life circumstances.
2. PCTs provide information of paramount importance for new interventions, development processes, and public health decision-making, informing clinical practice.
3. The correct implementation of PCTs with a robust statistical design, high-quality data collection, and follow-up are essential to increase their validity.

REFERENCES

1. Hollingsworth R, Palmu A, Pepin S, Dupuy M, Shrestha A, Jokinen J, et al. Effectiveness of the quadrivalent high-dose influenza vaccine for prevention of cardiovascular and respiratory events in people aged 65 years and above: Rationale and design of a real-world pragmatic randomized clinical trial. Am Heart J. 2021;237:54-61. https://doi.org/10.1016/j.ahj.2021.03.007
2. Ford I, Norrie J. Pragmatic Trials. N Engl J Med. 2016;375(5):454-463. https://doi.org/10.1056/NEJMra1510059
3. PRECIS-2 [homepage on the Internet]. Aberdeen: University of Aberdeen; c2016 [cited 2022 Oct 1]. Available from: http://www.precis-2.org/

Why should noninferiority clinical trials be performed?

Patricio Maskin[1,2], Juliana Carvalho Ferreira[1,3], Cecilia María Patino[1,4]

PRACTICAL SCENARIO

A double-blind, randomized, placebo-controlled, noninferiority clinical trial assessed the efficacy of adding five extra days of β-lactam treatment (amoxicillin plus clavulanate) versus placebo after three days of that therapy on clinical cure among clinically stable, moderately severe, community-acquired pneumonia adult patients admitted to 16 hospitals in France.[1] Results showed that 77% of the participants in the placebo group and 68% of the participants in the β-lactam group were considered clinically cured—the between-group difference was 9.4% (95% CI: −0.38 to 20.04). The authors concluded that treatment for three days was noninferior to treatment for eight days, and that these results could lead to important reductions in antibiotic consumption and decrease hospital costs.

NONINFERIORITY TRIALS

Among the types of randomized controlled trials, superiority trials are the most common. However, sometimes it is important to evaluate whether a new intervention is noninferior (equal or not worse) than an existing treatment in terms of efficacy, but exhibits other additional benefits, such as lower costs, fewer side/adverse effects, easier administration, or improved adherence.[2]

In our example, the authors chose a noninferiority clinical trial because their goal was to assess whether a shorter antibiotic regimen was not worse than the standard therapy, within a predefined noninferiority margin (NIM). The NIM is defined as an acceptable clinically difference in efficacy that is a trade-off for other advantages of the new treatment, such as shorter duration in our example. As long as the new treatment is not worse than the standard of care by this margin, the new treatment is considered noninferior.

The NIM is the largest reduction in efficacy of the new treatment compared with the standard of care that is acceptable to the expert community, and can be challenging to establish. If the width of the NIM is too narrow, a clinically acceptable alternative intervention might be considered inferior, and if it is too wide, an inferior intervention might be considered noninferior. Choosing the NIM requires both statistical and clinical consideration, and it is defined a priori and reported in the study protocol. Guidelines recommend defining the NIM based on a comprehensive review of the historical evidence of the efficacy of the current standard of care, which must also be the comparator.

Once the NIM is set (< 10% in our example),[1] the study hypothesis is described. The null hypothesis (H0) states that the between-group difference is larger than

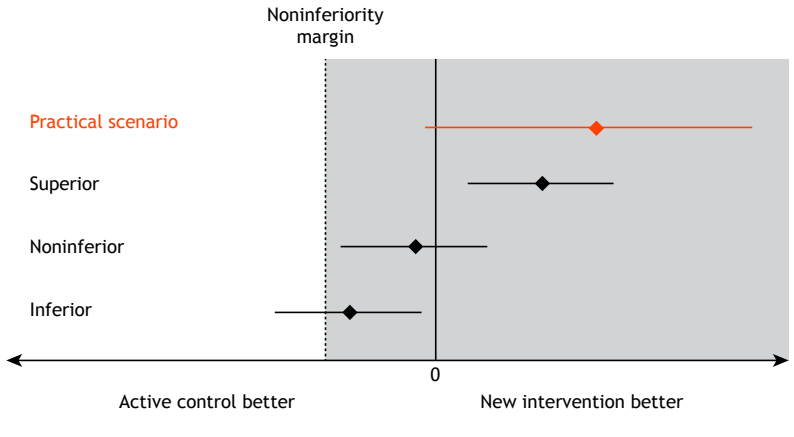

Figure 1. The figure shows possible result scenarios of a noninferiority clinical trial. The mean between-group differences (black circles) and respective 95% CIs (black error bars) for three potential scenarios comparing the new intervention with the active control are shown. The null hypothesis is that the difference between the new intervention and the active control is beyond the noninferiority margin, shown in the shadowed area. In our example (in red)[1], the lower bound of the 95% CI is within the noninferiority margin; therefore, the null hypothesis is rejected and noninferiority can be claimed.

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Unidad de Terapia Intensiva, Sección Neumonologia, Centro de Educación Médica e Investigaciones Clínicas – CEMIC – Buenos Aires, Argentina
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
4. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.

BY-NC © 2022 Sociedade Brasileira de Pneumologia e Tisiologia

the noninferiority margin (i.e., the new intervention is inferior) and the alternative hypothesis (H1) is that the between-group difference is smaller than the noninferiority margin (i.e., the new intervention is at least not worse). The statistical approach involves calculating the 95% CI of the mean difference (in our example)[1] in efficacy between the groups and evaluating if the lower bound of the 95% CI is greater than the noninferiority margin (Figure 1). The null hypothesis is rejected, and noninferiority can be claimed, when the lower bound of the 95% CI is smaller than this margin.

KEY POINTS

1) A noninferiority trial is the appropriate design to answer a research question when a new intervention is not expected to be superior to the standard of care in terms of efficacy, but it is not unacceptably inferior either, and offers additional advantages.

2) The noninferiority margin is the largest reduction in efficacy of the new treatment compared with the standard of care that is acceptable, so that the new treatment is not "unacceptably worse." This predefined margin should be based on clinical and statistical considerations.

REFERENCE

1. Dinh A, Ropers J, Duran C, Davido B, Deconinck L, Matt M, et al. Discontinuing β-lactam treatment after 3 days for patients with community-acquired pneumonia in non-critical care wards (PTC): a double-blind, randomised, placebo-controlled, non-inferiority trial [published correction appears in Lancet. 2021 Jun 5;397(10290):2150]. Lancet. 2021;397(10280):1195-1203. https://doi.org/10.1016/S0140-6736(21)00313-5
2. Kaji AH, Lewis RJ. Noninferiority Trials: Is a New Treatment Almost as Effective as Another?. JAMA. 2015;313(23):2371-2372. https://doi.org/10.1001/jama.2015.6645

Case reports: narratives highlighting clinical experiences that inform practice and future research studies

Cecilia Maria Patino[1,2,a], Juliana Carvalho Ferreira[1,3,b]

PRACTICAL SCENARIO

The authors of a case report[1] describe the success and safety of treating a 65-year-old patient admitted to the hospital with severe COPD and huge emphysematous bullae in the right middle lobe and bronchoscopic lung volume reduction (BLVR). They placed an endobronchial valve in the right middle bronchus after confirming that there was no collateral ventilation.[1] The authors describe the complete treatment regimen given to the patient and report that the huge bullae remained small one week later and disappeared at two months, with the improvement of pulmonary function, symptoms, and quality of life with no signs of obstruction, pneumonia, or pneumothorax. The authors concluded that BLVR may serve as an alternative treatment among selected patients with giant emphysematous bullae.

BACKGROUND

A case report is a comprehensive narrative that provides a clear and detailed description of unique medical experiences with patients that can impact both clinical and research practices. It is very important to publish those experiences, such as in our example, as case reports. Such narratives serve to increase existing knowledge on important clinical topics and to provide insights into new or rare diseases and to nonconventional patient care that can later be more formally evaluated using more sophisticated study designs, such as randomized controlled trials.

In the case report described above,[1] the authors report their experience treating a patient with a specific clinical pattern of COPD with BLVR, because patients with both COPD and giant emphysematous bullae have been excluded from previous treatment studies. The reason for excluding patients with this clinical presentation was that this type of emphysema is a predictor of operative mortality.

We highly recommend that all clinicians, and especially clinician-scientists, take the time to report interesting and unique cases of patients they treat in their home setting that could eventually affect the health of similar patients worldwide. Publishing case reports is an important first step in contributing to answer new questions and guiding informed patient-centered clinical practices. Additionally, for early-stage clinicians, a case report is sometimes the first opportunity to become a published author, since there is no requirement for design or implementation of a clinical research study. Although case reports are at the base of the evidence-based pyramid and are often mistakenly perceived as unimportant in medical science, we highlight that the evidence-based pyramid serves as a guide for clinicians' decision-making processes as a reminder that decisions and recommendations for patients should be based on research data resulting from robust study designs, such as clinical trials, but it does not imply that case reports are not valid.

USING ESTABLISHED GUIDELINES TO PUBLISH CASE REPORTS

To write and publish high-quality case reports, we highly recommend the use of the CARE (CAse REport) Statement and Checklist[2] for the accurate report of information that should be provided in each section of the case report (Table 1).

Table 1. Examples of some of the items on the CARE (CAse REport) Statement and Checklist.[2]

Item	Topic	Checklist Item description
1	Title	Include in the title the name of the study design: "case report".
3a	Abstract Introduction	What is unique about this case? What does it add to the medical literature?
5a	Patient Information	De-identify demographic information and other specific patient information.
9a	Therapeutic Intervention	Describe types of intervention (such as pharmacological, surgical, preventive, and self-care).
10b	Follow-up and outcomes	Provide important follow-up diagnostic/nondiagnostic test results.
11a	Discussion	Discuss the strengths and limitations in your approach to this case.

REFERENCES

1. Lee EG, Rhee CK. Bronchoscopic lung volume reduction using an endobronchial valve to treat a huge emphysematous bullae: a case report. BMC Pulm Med. 2019;19(1):92. https://doi.org/10.1186/s12890-019-0849-z
2. Riley DS, Barber MS. Kienle GS, Aronson JK, von Schoen-Angerer T, Tugwell P, et. al. CARE guidelines for case reports: explanation and elaboration document. J Clin Epidemiol. 2017;89:218-235. https://doi.org/10.1016/j.jclinepi.2017.04.026

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
a. http://orcid.org/0000-0001-5742-2157; b. http://orcid.org/0000-0001-6548-1384

BY-NC © 2019 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

Case series: an essential study design to build knowledge and pose hypotheses for rare and new diseases

Carlos Arturo Torres-Duque[1,2,3], Cecilia Maria Patino[1,4], Juliana Carvalho Ferreira[1,5]

PRACTICAL SCENARIO

At the end of December of 2019, a pneumonia outbreak of unknown origin appeared in China. Soon afterwards, the causative virus was identified—SARS coronavirus 2 (SARS-CoV-2), and the disease was named coronavirus disease 2019 (COVID-19). In January of 2020, Chinese investigators published a detailed case series describing the characteristics and outcomes of 41 adults with confirmed COVID-19.[1] The study showed that 15% of those patients died during the study period. That case series[1] was extremely important because it was the first published description of the impact of the new disease, helping clinicians around the world to face a new pandemic.

CONCEPTS AND APPLICATION

A case series includes a description of the characteristics and outcomes among a group of individuals with either a disease or an exposure (which can be an intervention) over a period of time and without a control group. Data are collected retrospectively or prospectively, and there is no randomization. The objective is to describe the population and outcomes, rather than compare risks across groups. Therefore, a case series differs from cohort studies because the latter compares the risk between two groups (exposed and unexposed) and allows for the estimation of an absolute risk for the occurrence of a given outcome in the exposed group and of a relative risk in comparison with the unexposed group.

The case series design is not considered the strongest source of evidence due to the absence of a control group and the risk of bias, in particular selection bias, since typical or severe cases of the disease are more easily identified, and rare presentations or mild cases may not be included. In the Chinese report,[1] for example, patients with less severe COVID-19 were not hospitalized and therefore were not included in the case series. However, case series are particularly important when a new disease or treatment emerges, because it provides descriptive information and contributes to building knowledge and generating hypotheses. Case series is also an appropriate study design to describe new treatments, previously unknown medication adverse events, and rare diseases.[2]

METHODOLOGY AND QUALITY OF CASE SERIES STUDIES

- Inclusion criteria - A precise operational definition of a "case" is crucial for the reliability of the study.
- Sampling - Two strategies are possible: 1) based on disease or exposure; 2) based on a specific outcome.
- Selection of variables of interest - A detailed selection and a clear definition of predictive variables of interest are necessary, as well as test results, interventions, complications, adverse events, and outcomes.
- Systematic collection of data and robust analysis - They assure the quality of a case series study.

Table 1 presents a tool for evaluating the methodological quality of case series.[2]

Table 1. A tool for evaluating the methodological quality of case series.

Domains	Leading explanatory questions
Selection	1. Were all the potentially eligible patients included or is the selection method unclear to the extent that other patients with similar presentations may not have been reported?
Definition of exposure and outcomes	2. Was the exposure adequately and clearly defined? 3. Was the outcome adequately and clearly defined?
Causality	4. Were other alternative causes that may explain the observation ruled out? 5. Was there a challenge/rechallenge phenomenon? 6. Was there a dose-response effect? 7. Was follow-up long enough for outcomes to occur?
Reporting	8. Are the cases described with sufficient details to allow other investigators to replicate the research or to allow practitioners to make inferences related to their own practice?

Adapted from Murad et al.[2] Questions 4, 5 and 6 are more relevant for adverse drug events.

REFERENCES

1. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China [published correction appears in Lancet. 2020 Jan 30]. Lancet. 2020;395(10223):497-506. https://doi.org/10.1016/S0140-6736(20)30183-5
2. Murad MH, Sultan S, Haffar S, Bazerbachi F. Methodological quality and synthesis of case series and case reports. BMJ Evid Based Med. 2018;23(2):60-63. https://doi.org/10.1136/bmjebm-2017-110853

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. CINEUMO Research Center, Fundación Neumológica Colombiana, Bogota, Colombia.
3. Universidad de la Sabana, Bogota, Colombia.
4. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
5. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2020 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

J Bras Pneumol. 2017;43(5):330-331
http://dx.doi.org/10.1590/S1806-37562017000000330

Understanding diagnostic tests. Part 1.

Juliana Carvalho Ferreira[1,2], Cecilia Maria Patino[1,3]

PRACTICAL SCENARIO

Investigators studied the diagnostic accuracy of serum procalcitonin levels to diagnose parapneumonic pleural effusions (PPE) and differentiate it from other causes of pleural effusions. They found that procalcitonin (with a cut-off value of 0.195 ng/mL) had a sensitivity of 83% and a specificity of 80% to diagnose PPE and accurately diagnosed individuals with PPE.[1]

USING DIAGNOSTIC TESTS IN CLINICAL PRACTICE

Clinicians are frequently faced with the challenge of diagnosing a disease based on diagnostic test results. Most diagnostic tests used in clinical practice, however, are not perfect and produce false positive results (the test is positive, but the patient does not have the disease) and false negative results (the test is negative, but the patient has the disease). Therefore, learning to interpret the properties of diagnostic tests is a critical competency for clinicians and researchers. In this article, we discuss sensitivity and specificity. In the forthcoming parts, we will discuss positive and negative predictive values, and receiver operating characteristic (ROC) curves.

Sensitivity and specificity are important measures of a diagnostic test because they give us an idea of how well a new diagnostic test performs when compared with an existing gold standard test. Sensitivity is defined as the proportion of subjects with the disease who have a positive test. In the example in Table 1, true positives (n = 39) divided by the total number of subjects with disease (n = 47) results in 83%. Specificity is defined as the proportion of subjects without the disease who have a negative test. In the example, true negatives (n = 81) divided by total number of subjects without the disease (n = 101) results in 80%.

When a new diagnostic test is evaluated, the investigator sets a cut-off point which defines whether the test is positive or negative, and there is always a trade-off between sensitivity and specificity. In our example, if the cut-off point for a positive procalcitonin test was decreased from 0.195 ng/mL to 0.095 ng/mL, it might detect more cases of PPE, decreasing the false negative rate and increasing sensitivity, but the test would also be positive in more subjects without PPE, increasing the false positive rate and decreasing specificity. This trade-off between sensitivity and specificity for several possible cut-off points can be used to plot a ROC curve and describe the overall test performance in discriminating between presence and absence of the disease; we can also use sensitivity and specificity to calculate likelihood ratios, as we will see later in this series.

Sensitivity and specificity are useful measures to evaluate the performance of a diagnostic test but are not very helpful for personalized clinical decision making.[2] When a clinician is facing a patient with a positive test result, the most important question is: what is the probability that, given that the test is positive, the patient has the disease? The sensitivity of the test does not tell us that; it tells us the probability of a positive test, given that the patient has the disease. We will address more relevant clinical measures of diagnosis in part 2 of this series.

Table 1. Diagnostic performance of serum procalcitonin testing for identifying parapneumonic pleural effusion.

Result	PPE		Total
	+	–	
PCT+	a = 39	b = 20	59
PCT–	c = 8	d = 81	89
Total	47	101	148

Data obtained from He et al.[1] PCT: procalcitonin; PPE: parapneumonic pleural effusion. Sensitivity (light grey column) = a/(a + c). Specificity (dark grey column) = d/(b + d).

REFERENCES

1. He C, Wang B, Li D, Xu H, Shen Y. Performance of procalcitonin in diagnosing parapneumonic pleural effusions: A clinical study and meta-analysis. Medicine (Baltimore). 2017;96(33):e7829. https://doi.org/10.1097/MD.0000000000007829
2. Altman DG, Bland M. Diagnostic tests. 1: Sensitivity and specificity. BMJ. 1994;308(6943):1552. https://doi.org/10.1136/bmj.308.6943.1552

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

BY-NC © 2017 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

ERRATUM

Manuscript: Understanding diagnostic tests. Part 1.

Publication: J Bras Pneumol. 2022;43(5):330

DOI: http://dx.doi.org/10.1590/S1806-37562017000000330

On page 330 of the original publication, in Table 1, last line of the text, where is written
Specificity (dark grey column) = b/(b + d)

Should be read
Specificity (dark grey column) = d/(b + d)

Understanding diagnostic tests. Part 2.

Cecilia Maria Patino[1,2], Juliana Carvalho Ferreira[1,3]

PRACTICAL SCENARIO

Investigators studied the diagnostic accuracy of serum procalcitonin levels in diagnosing parapneumonic pleural effusions (PPE) and differentiating it from other causes of pleural effusions. They found that procalcitonin had a positive predictive value (PPV) of 66% and a negative predictive value (NPV) of 91%.[1]

PPV AND NPV OF DIAGNOSTIC TESTS

In the previous article[2] we discussed two common features of diagnostics tests, sensitivity and specificity, which are important characteristics that describe the accuracy of a test. In this article, we focus on important features of a diagnostic test that help us understand how well a new test diagnoses a disease based on the results of the gold standard: PPV and NPV.

The PPV of a diagnostic test is the proportion of individuals who test positive to the new test and have the disease according to the gold standard (the proportion of true positives). When a diagnostic test has a high PPV, there is a high probability that a patient has the disease being investigated when the patient has a positive test.

The NPV of a diagnostic test is the proportion of individuals who test negative to the new test and do not have the disease according to the gold standard (the proportion of true negatives). When a test has a high NPV, there is a high probability that a patient does not have the disease being investigated when the patient has a negative test. In our example, the PPV was 66% (39/59), and the NPV was 91% (81/89), according to the results of the new test among 148 individuals (Table 1).

The PPV and the NPV of a new test depend on the prevalence of the disease in the population; thus, their results will change across populations with higher or lower prevalence of the disease when compared with the population where the test is first reported. If the prevalence of the disease is high in a given population, PPV increases and NPV decreases. Thus, the results of predictive values are not fixed characteristics of the test and cannot be generalized across populations with different prevalences of the disease.[3] There is an easy way to calculate PPV and NPV, based on Bayes' theorem, using previously reported results and taking into account the local disease prevalence.[2]

Table 1. Diagnostic performance of serum procalcitonin testing for identifying parapneumonic pleural effusion.

	PPE		Total
	+	-	
PCT+	a = 39	b = 20	59
PCT-	c = 8	d = 81	89
Total	47	101	148

Data obtained from He et al.[1] PCT: procalcitonin; PPE: parapneumonic pleural effusion. Sensitivity = a/(a + c); specificity = b/(b + d); Positive predictive value (light gray row) = a/(a+b); and negative predictive value (dark gray row) = d/(d+c).

PPV and NPV are also important indicators when screening the general population. A screening test with high sensitivity and specificity may still have low PPV if the prevalence of the disease is low in that population. For example, when screening for cancer in asymptomatic adults, if the NPV of the test is high, negative results are helpful to rule out the presence of the disease; however, if the PPV is low, a positive result has a higher probability of being a false positive.

PPV and NPV are more useful than sensitivity and specificity for clinicians because they estimate the probability of the disease (or its absence), given the test result. In the next, final part of this series about diagnostic tests, we will discuss likelihood ratios and ROC curves.

REFERENCES

1. He C, Wang B, Li D, Xu H, Shen Y. Performance of procalcitonin in diagnosing parapneumonic pleural effusions: A clinical study and meta-analysis. Medicine (Baltimore). 2017;96(33):e7829. https://doi.org/10.1097/MD.0000000000007829
2. Ferreira JC, Patino CM. Understanding diagnostic tests. Part 1. J Bras Pneumol. 2017;43(5):330. https://doi.org/10.1590/S1806-37562017000000330
3. Altman DG, Bland M. Diagnostic tests. 2: predictive values. BMJ. 1994;309(6947):102. https://doi.org/10.1136/bmj.309.6947.102

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2017 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

Understanding diagnostic tests. Part 3.

Juliana Carvalho Ferreira[1,2,a], Cecilia Maria Patino[1,3,b]

In the previous articles from this series[1,2] we discussed important characteristics used in order to evaluate diagnostic tests: sensitivity, specificity, positive predictive value, and negative predictive value. In this final part, we discuss positive likelihood ratio (LR+), negative likelihood ratio (LR−), and ROC curves.

LIKELIHOOD RATIOS

LRs combine sensitivity and specificity to quantify how helpful a new diagnostic test is in changing (increasing or decreasing) the probability of having a disease compared with the prevalence of that disease (pretest probability) in the population studied. The LR+ of a test is the probability of a positive result in patients with the disease divided by the probability of a positive result in patients without the disease, whereas LR− is the probability of a negative result in patients with the disease divided by the probability of a negative result in patients without the disease. LR+ ranges from 1 to infinity, and an LR+ of 1 indicates that the probability of a positive test result is the same for patients with and without the disease; therefore, the test is useless. An LR+ greater than 1 supports the presence of the disease, and the greater LR+ is, the more a positive test result increases the probability of the disease when compared with the pretest probability. LR− ranges from 1 to 0, and the closer the LR is to 0, the lower the probability of the disease is if the test result is negative.

ROC CURVES

We use ROC curves to make a global assessment of the value of a diagnostic test by calculating the area under the curve (AUC). The values of the AUC can vary from 0 to 1.0, and values over 0.8 indicate that the diagnostic test has very good accuracy. The ROC curve plots sensitivity (true positives) against "1 – specificity" (false negatives) for all the possible cut-off values of the new test (Figure 1). As we have previously discussed, there is always a trade-off between sensitivity and specificity when we define a cut-off value for quantitative test results. If a new test were perfect, there would be a complete separation of values between patients with and without the disease, the cut-off value would be the lowest value among patients with disease, and the AUC would be 1. However, since there are no perfect tests, there will always be some false positive or some false negative results. The more accurate a test is, the greater the AUC is, which is the probability that a random person with the disease has a higher value of the measurement than a random person without the disease.[3]

MAKING SENSE OF DIAGNOSTIC TEST PERFORMANCE CHARACTERISTICS

If you are wondering which of the parameters described is more useful to evaluate a diagnostic test—sensitivity, specificity, LRs, or ROC curve—the answer is: it depends! Each parameter describes a specific characteristic of the test, and depending on how you will use the test, one or another may be more useful. Now that you understand these concepts, interpreting a test result will be much more than just looking at the result.

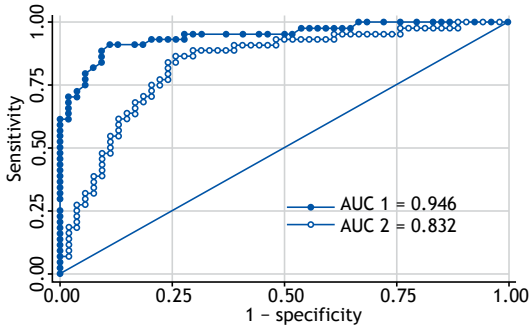

Figure 1. ROC curve plotting sensitivity vs. "1 – specificity" for two different tests. Both tests have good accuracy; however, test 1 (closed circles) has an area under the curve (AUC) of 0.946 and test 2 has an AUC of 0.832 (open circles), meaning that test 1 has overall better accuracy to discriminate between patients with and without the disease. This figure was created with fictitious data.

REFERENCES

1. Ferreira JC, Patino CM. Understanding diagnostic tests. Part 1. J Bras Pneumol. 2017;43(5):330. https://doi.org/10.1590/s1806-37562017000000330
2. Patino CM, Ferreira JC. Understanding diagnostic tests. Part 2. J Bras Pneumol. 2017;43(6):408. https://doi.org/10.1590/s1806-37562017000000424
3. Altman DG, Bland JM. Diagnostic tests 3: receiver operating characteristic plots. BMJ. 1994■309(6948):188. https://doi.org/10.1136/bmj.309.6948.188

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.
a. http://orcid.org/0000-0001-6548-1384; b. http://orcid.org/0000-0001-5742-2157

BY-NC © 2018 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

J Bras Pneumol. 2017;43(4):252-252
http://dx.doi.org/10.1590/S1806-37562017000000241

Prognostic studies for health care decision making

Cecilia Maria Patino[1,2] Juliana Carvalho Ferreira[1,3]

PRACTICAL SCENARIO

A multicenter retrospective cohort study was conducted to develop and validate a prognostic model to predict 1-year mortality among adult patients receiving at least 14 uninterrupted days of mechanical ventilation. Likely prognostic variables were chosen, *a priori*, based on published literature and clinical judgment (10 variables). During the development phase of the study, the prognostic variables were included in a logistic regression model to evaluate how well each variable predicted 1-year mortality by calculating discrimination (ability to correctly classify patients into those who did and did not die) using ROC curves and area under the curve (AUC). The authors found that 5 of the 10 variables maximized the prognostic capability of the model for 1-year mortality (age, platelet count, vasopressor use, hemodialysis, no trauma diagnosis) showing very good discrimination (AUC = 0.80; 95% CI: 0.76-0.83). For the validation phase, the authors used the β-coefficient values estimated for each variable in the development cohort logistic regression model to predict 1-year mortality in a new cohort of patients and showed that discrimination was also very good (AUC = 0.78; 95% CI: 0.72-0.83), thus showing that the 5-variable model was valid. Then the authors created a clinical prediction rule, a point system used to easily calculate the probability of 1-year mortality for each patient, based on the strength of association of each variable (β-coefficient) with mortality in the development model. All β-coefficients were assigned 1 point except for the category "age ≥ 65 years," which was assigned 2 points. Lastly, the authors validated this point system by showing that, as the number of points increased, the probability of 1-year mortality increased.[1]

WHY PROGNOSTIC STUDIES ARE USEFUL

The overall goal of prognostic research for clinical settings is to help clinicians, patients, and families make informed health-care decisions based on information available on each patient in the present to predict outcomes in the future. In our example, identifying patients at high risk of dying within 1 year justifies clinicians' recommendation for closer outpatient-monitoring after discharge. Additionally, it helps patients and family think about appropriate end-of-life decisions for those at very high risk of dying, as well as to identify individualized interventions to prevent future hospitalizations due to respiratory failure.

HOW TO DEVELOP A CLINICAL PREDICTION RULE

The process involves designing a retrospective or prospective cohort study that measures prognostic variables among participants at study baseline (entry), that follows them during a pre-specified time and that assesses whether they develop the outcome or not. Using data from a subset of the participants, called the development cohort, a logistic regression model with the outcome (in our example, 1-year mortality) as the dependent variable and plausible predictive variables as independent variables is built and the AUC is calculated (Figure 1). In a second step, the mathematical equation (β-coefficients) from the development model is tested in another subgroup of similar patients, called the validation cohort. The clinical prediction rule is built by assigning points to each predictive variable based on their strength of association with the outcome.[2]

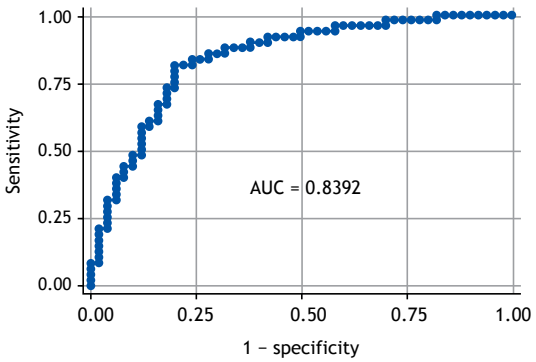

Figure 1. The ROC curve is used to quantify model discrimination by plotting the true positive rate (sensitivity) against the false positive rate (1 – specificity) for different possible cut-off values of a prognostic model. The greater the area under the curve (AUC), the better the model discriminates the subjects with the outcome from those without it. This figure was created with fictitious data.

REFERENCES

1. Hough CL, Caldwell ES, Cox CE, Douglas IS, Kahn JM, White DB, et al. Development and validation of a mortality prediction model for patients receiving 14 days of mechanical ventilation. Crit Care Med. 2015;43(11):2339-45. https://doi.org/10.1097/CCM.0000000000001205
2. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10(2):e1001381. https://doi.org/10.1371/journal.pmed.1001381

1. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.
2. Methods in Epidemiologic, Clinical and Operations Research (MECOR) program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
3. Divisão de Pneumologia do Instituto do Coração (InCor) do Hospital das Clínicas, Faculdade de Medicina da Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2017 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

Systematic reviews: a brief overview

Natalia Causada Calo[1,2], Juliana Carvalho Ferreira[1,3], Cecilia Maria Patino[1,4]

SCENARIO

A randomized controlled trial (RCT) was conducted to determine if drug B improves survival when compared with drug A in patients with condition Y. A systematic review (SR) can also answer this same question; however, it is important to differentiate between these study designs.

THE PROCESS OF A SYSTEMATIC REVIEW

SRs summarize the body of research from primary studies that address a well-defined research question. It also evaluates the quality of the studies and their conclusions using a systematic and reproducible approach.[1] SRs commonly answer questions related to therapy, diagnosis, or prognosis. They are particularly useful when similar studies show conflicting results, when various studies with a small number of participants show inconclusive results, or when practice guidelines are being developed.

The conduct of an SR follows a strict methodological process, which includes the definition of inclusion and exclusion criteria and evaluation of the risk of errors (bias).[1] The process (or "system") for an SR is summarized in Table 1.

First, the PICOT format can be used to define the different components of the research question: the population (P), the intervention or exposure (I), the comparison group (C), the outcome (O), and the type of study design (T). These components will depend on the nature of the study question (intervention, diagnosis, or prognosis). In our hypothetical example, we are interested in comparing the effects of two drugs (interventions) on survival, and the most appropriate study design is an RCT.

Once the question and the detailed study eligibility criteria have been defined, a comprehensive literature search is conducted. This step is elaborate and often requires a librarian who provides the "language" for the search. In contrast to a search that we often conduct as clinicians, in order to conduct an SR, the search has to use clear terms, be comprehensive and reproducible, and be performed across all important medical databases, including the gray literature. Once the search is completed, researchers screen the list of references for eligibility. Typically, two researchers complete this step and the data extraction that follows. A key component of an SR is the evaluation of the quality of the studies included. Different tools are available according to the nature of the question.[2] For RCTs, for example, questions about randomization and allocation concealment are asked. For prognostic studies, it is essential to understand if patient selection is representative.

Once all steps are completed, data are summarized and often analyzed to provide quantitative estimates with their corresponding confidence intervals. This last part corresponds to the meta-analysis, which will be discussed in a forthcoming article. In some cases, an SR does not include a meta-analysis; when this occurs, a transparent report of the methodology should be provided.

KEY CONCEPTS

- An SR is a summary of the evidence that addresses a well-defined research question in a systematic and reproducible manner.
- One study of interventions, diagnosis, or prognosis alone is unlikely to represent the entirety of the evidence. SRs are useful because they summarize the body of evidence after a comprehensive and reproducible medical literature search and assessment of the risk of bias.

Table 1. The process of a systematic review.

1. Definition of the question: PICOT format	Systematic review	Systematic review + Meta-analysis
2. Inclusion and exclusion criteria		
3. Literature search for studies		
4. Screening of studies for eligibility		
5. Data collection from studies		
6. Assessment of risk of bias of the studies included		
7. Analysis of results (synthesis)	Meta-analysis	
8. Interpretation of the results		
9. Conclusions on the estimates		

PICOT: P: population; I: intervention/exposure; C: control group or comparator; O: outcome; and T: type of study design.

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Department of Medicine, Division of Gastroenterology, University of Ottawa, Ottawa, ON, Canada
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
4. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

BY-NC © 2020 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

- A quantitative analysis (meta-analysis) often accompanies the summary of the evidence, yielding a higher precision in the results than individual studies.

REFERENCES

1. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions Version 6.0 (Updated July 2019). Chichester, West Sussex ; Hoboken NJ :Wiley & Sons; 2019.
2. Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ. 2020;368:l6890. https://doi.org/10.1136/bmj.l6890

Meta-analyses: a primer for clinicians

Bruno L Ferreyro[1,2], Cecilia M Patino[1,3], Juliana Carvalho Ferreira[1,4]

PRACTICAL SCENARIO

Investigators conducted a systematic review (SR) study that included eight randomized controlled trials (RCTs) comparing drug A vs. drug B for the treatment of condition Y. The outcome of interest was 30-day all-cause mortality. Each study reported an effect estimate (OR) to compare the two drugs. Investigators then generated a pooled estimate to summarize the overall effect across the studies. How is that achieved in a meta-analysis?

WHAT IS A META-ANALYSIS?

A meta-analysis is a statistical approach that combines results from individual studies identified in an SR and calculates a pooled estimate of the magnitude and direction of treatment effects.[1] Consequently, the overall sample size and the precision of the estimate increase, and the width of confidence intervals decreases. The combined treatment effect is estimated by calculating a weighted average across individual study estimates. The weight assigned to each study result is related to the precision of each estimate, which in turn is related to the sample size of the study. Therefore, larger studies have a greater influence on the final pooled estimate.

Frequently, a meta-analysis follows an SR of individual RCTs or observational studies. Depending on the nature of the research question, a meta-analysis can be used to answer questions about intervention effectiveness, diagnostic/prognostic test accuracy, and disease burden (prevalence and incidence).

SRs often include studies with distinct features that lead to clinical, methodological, and statistical heterogeneity. Clinical heterogeneity arises from differences in study participants, interventions, or outcome definitions. Methodological heterogeneity arises, for example, when some of the RCTs included are blinded, and others are not. In a meta-analysis, statistical heterogeneity is formally assessed by calculating the I^2 statistic, which ranges from 0% to 100%. An $I^2 > 50\%$ indicates high heterogeneity, which should raise the question of whether it is reasonable to perform a meta-analysis or not and to prompt the search of potential underlying reasons for heterogeneity.

FOREST PLOTS: A VISUAL SUMMARY OF META-ANALYSIS RESULTS

A forest plot[2] is the key graphical representation of the major findings of an SR and meta-analysis. In our example (Figure 1), each row represents one of the 8 RCTs included in the SR with their respective effect estimates (OR and 95% CI). The bottom row represents the pooled estimate of the effect, that is, the result of the meta-analysis. Each individual study has a different relative weight; for example, study 7 has the largest weight, which is likely associated with a high precision of the estimate (smaller CI). Notably, specific estimates of most individual studies are not statistically significant (95% CI includes the value of 1), whereas the pooled estimate shows a statistically significant beneficial effect of drug A vs. drug B. The heterogeneity of the study was 45%, estimated by the I^2 statistic.

SRs combined with meta-analyses are often considered as one of the highest levels of analysis in evidence-based medicine because they combine the results of various RCTs/observational studies and offer a more precise estimate of the effect size of a given intervention. They can be very useful for clinical decision making, although their results are only as good as the studies included in the analysis.

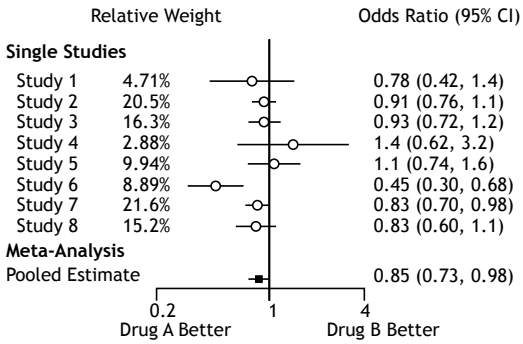

Figure 1. An example of a forest plot.

REFERENCES

1. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews Interventions version 6.0 (updated July 2019). Chichester (UK): Wiley & Sons; 2019.
2. University of Oxford. Centre for Evidence-Based Intervention [homepage on the Internet]. Oxford: University of Oxford [cited 2016 Jun 1]. How to analyze the forest plot: assess the heterogeneity amongst the studies. Available from: https://www.spi.ox.ac.uk/how-to-interpret-the-sample-forest-plot

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Department of Medicine, Sinai Health System, University Health Network, Toronto, ON, Canada.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
4. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2020 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

J Bras Pneumol. 2018;44(6):241-241
http://dx.doi.org/10.1590/S1806-37562018000000364

Critical appraisal of the literature. Why do we care?

Juliana Carvalho Ferreira[1,2,a], Cecilia Maria Patino[1,3,b]

PRACTICAL SCENARIO

Investigators conducted a noninferiority, double-blind clinical trial involving 4,215 patients with mild asthma, randomly assigned to receive twice-daily placebo plus budesonide-formoterol used as needed vs. maintenance therapy with twice-daily budesonide plus terbutaline as needed. They found that budesonide-formoterol used as needed was noninferior to twice-daily budesonide concerning the rate of severe asthma exacerbations but was inferior in controlling symptoms.[1]

HOW TO CRITICALLY APPRAISE THE MEDICAL LITERATURE

As clinicians, when we read a paper reporting the benefit of a given intervention, we make a judgment regarding whether we should use those results to inform how we care for our patients. In our example, after reading the paper, we ask ourselves: should a clinician working in a public hospital in Brazil start prescribing budesonide-formoterol as needed rather than maintenance budesonide for her patients with mild asthma? What criteria should guide her decision to adopt a new intervention? One may think that if a study is published in a high-impact, peer-reviewed journal, it is of high quality and should therefore be used to guide clinical decision making. However, if the population included in the study or the context is different from her population, that may not be the case. Therefore, examining the external validity of a study is critical to informing local practice.

Other commonly used criteria are related to evaluating the quality of the evidence by evaluating the type of study design used. The pyramid of evidence puts meta-analyses at the top (as providing the highest quality of evidence), followed by systematic reviews and randomized controlled trials; then come observational studies (cohort, case-control, and cross-sectional studies); whereas case reports and case series are categorized as offering the lowest quality of evidence. Although those criteria may be helpful, making a detailed appraisal of a paper, taking into account aspects other than the study design, is a skill that researchers and clinicians can learn and apply when reading the literature.

Critical appraisal is the **systematic** evaluation of clinical research papers that helps us establish if the results are valid and if they could be used to inform medical decision in a given local population and context. There are several published guidelines for critically appraising the scientific literature, most of which are structured as checklists and address specific study designs.[2] Although different appraisal tools may vary, the general structure is shown in Table 1.

The items in Table 1 are a guide to appraising the content of a research article. There are also guidelines for appraising the quality of reporting of health research which focus on the reporting accuracy and completeness of research studies.[3] These two types of appraisal (content and reporting) are complementary and should both be used, because it is possible that a research paper has high reporting quality but is not relevant to the context in question.

KEY MESSAGE

Critical appraisal of the literature is an essential skill for researchers and clinicians, and there are easy-to-use guidelines. Clinicians have the responsibility to help patients make health-related decisions, which should be based on high-quality, valid research that is applicable in their context.

Table 1. How to appraise medical literature.

QUESTION	WHAT TO LOOK FOR
Does this study address a clearly focused, important question?	The research question should be clearly stated, and the scope of the study should be focused
Was the study design appropriate for the research question?	The chosen design should be suited to answering the research question
Did the study use valid methods to address this question?	Adequate participant allocation, intervention administration, and outcome assessments
Was systematic bias avoided or minimized?	The groups being compared should be as similar as possible except for the intervention/exposure being studied
Was the primary outcome adequately evaluated?	Assessments should be blinded when possible, measured objectively, and performed for all (or most) participants
Are these valid, significant results applicable to my patient or population?	The study intervention should be available, affordable, and acceptable in your clinical context

REFERENCES

1. Bateman ED, Reddel HK, O'Byrne PM, Barnes PJ, Zhong N, Keen C, et al. As-Needed Budesonide-Formoterol versus Maintenance Budesonide in Mild Asthma. N Engl J Med. 2018;378(20):1877-1887. https://doi.org/10.1056/NEJMoa1715275
2. Critical Appraisal Skills Programme-CASP [homepage on the Internet]. Oxford (UK): CASP; c2018 [cited 2018 Nov 1]. CASP Checklists. Available from: https://casp-uk.net/casp-tools-checklists/
3. Equator network [homepage on the Internet]. Oxford (UK): Centre for Statistics in Medicine, University of Oxford [cited 2018 Nov 1]. Available from: https://www.equator-network.org/

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
a. http://orcid.org/0000-0001-5742-2157; b. http://orcid.org/0000-0001-6548-1384

BY-NC © 2018 Sociedade Brasileira de Pneumologia e Tisiologia

◄ Voltar ao sumário

Clinical practice guidelines: how do they help clinicians and patients make important decisions about health?

Juliana Carvalho Ferreira[1,2], Cecilia Maria Patino[1,3]

PRACTICAL SCENARIO

In 2017, a clinical practice guideline (CPG) about the use of mechanical ventilation in adult patients with acute respiratory distress syndrome (ARDS), sponsored by three medical societies, recommended the use of lower tidal volumes (4-8 mL/kg of predicted body weight) and lower inspiratory pressures (plateau pressure < 30 cmH_2O). The CPG classified this recommendation as "strong" and with "moderate confidence in effect estimates".[1]

INTRODUCTION

When clinicians and patients make health-related decisions, they should consider the potential benefits and harms of diagnostic procedures and interventions, as well as patient values and preferences. When the benefits outweigh the harms, the diagnostic procedure or intervention should be recommended, or otherwise, avoided. However, in times of information abundance, how can we facilitate this decision-making process for both clinicians and patients? CPGs offer recommendations about specific clinical questions and provide a summary of the evidence—and its quality—to help the decision making of clinicians and patients.

HOW ARE RECOMMENDATIONS MADE?

In the past, recommendations were commonly based on expert opinion, but this process was often based on low quality evidence and thus may not have represented the best choice for the patient. Since then, formal systems have been created, such as the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) system, which uses rigorous methodological processes.[2] As an example, the Brazilian Thoracic Association recently adopted GRADE as a formal approach to develop Brazilian CPGs, which will be published in the JBP.

GRADE offers a systematic approach to develop CPGs, including the formulation of clinical questions aligned with patient-centered outcomes, systematic literature review, and a structured appraisal process to evaluate the quality of the evidence, which ultimately informs the recommendations. Randomized controlled trials usually provide the highest quality of evidence, but five limitations can impact on study quality: study limitations (biases), imprecision, inconsistency across studies, indirectness of evidence, and publication bias.

The process of writing CPG recommendations is rigorous. A CPG should be clearly written to avoid ambiguity and use standard approaches. The strength of a recommendation reflects the extent to which one can be confident that the desirable effects of an intervention outweigh undesirable effects. Chart 1 shows what a strong or conditional recommendation means for clinicians, patients, and policy makers. Four key factors determine the strength of a recommendation: balance between the desirable and undesirable consequences; quality of the evidence; variability in values and preferences; and costs.

In our example, the CPG makes a strong recommendation for using low tidal volumes and inspiratory pressures for patients with ARDS, because the evidence suggests that the benefits outweigh the harms. The recommendation includes a statement about the quality of the evidence, considered moderate, implying that, although the panel recommends the intervention, they acknowledge the fact that the quality of evidence is not high and that further research is likely to have an impact on our confidence in the estimate of the effect of the intervention.

Finally, it is important to remember that recommendations from CPGs are only a guide for decision making and should always be put into context, considering patient preferences, values, and perspectives, as well as local available resources.

Chart 1. Examples of recommendations that inform patients, clinicians, and policy makers for the decision making.

	Strong recommendation	Conditional recommendation
Patients	Most informed patients would choose the recommended management, and only a minority would not accept it	Most informed patients would choose the recommended management, but many would not
Clinicians	Most patients should receive the recommended course of action	Clinicians must ensure that patients' care is in keeping with their values and preferences
Policy makers	The recommendation can be adopted as a policy in most situations	There is a need for substantial debate and stakeholder involvement

REFERENCES

1. 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. https://doi.org/10.1164/rccm.201703-0548ST
2. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-6. https://doi.org/10.1136/bmj.39489.470347.AD

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
a. http://orcid.org/0000-0001-6548-1384; b. http://orcid.org/0000-0001-5742-2157

BY-NC © 2019 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

Measures of frequency: calculating prevalence and incidence in the era of COVID-19

Marcia Margaret Menezes Pizzichini[1,2], Cecilia Maria Patino[1,3], Juliana Carvalho Ferreira[1,4]

PRACTICAL SCENARIO

Coronavirus disease 2019 (COVID-19), caused by the coronavirus designated SARS-CoV-2, has become a pandemic despite global efforts to prevent its spread. The first confirmed case of COVID-19 in Brazil was reported on February 26 of 2020. Until May 11 of 2020, a total of 168,331 Brazilians had a confirmed diagnosis of COVID-19, of whom 89,429 (53.1%) were still infected, 67,384 (40%) had been cured, and 11,519 (6.8%) had died. The number of new cases on May 11th was 5,632, and the reported incidence was 80.1/100.000 population.[1]

MEASURES OF MORBIDITY AND MORTALITY OF COVID-19

Counting mild-to-severe and asymptomatic cases of COVID-19 is essential to describe and interpret local epidemic responses. In this scenario, repeated estimates of prevalence and incidence inform trajectory trends of the disease and guide the decision-making process related to control measures and resource allocation.[2]

Prevalence is defined as the proportion of a population who has the disease at one time point (Table 1). Cross-sectional studies are commonly used in order to conduct prevalence studies because they examine the disease at one particular time point. The prevalence of confirmed COVID-19 cases on May 11th was 0.08%, estimated as the number of cases of COVID-19 on that day divided by the population at risk (the Brazilian population).[1] Because measures of prevalence include both new and existing cases, they do not provide a complete picture of the natural history of the disease. In addition, the calculation of COVID-19 prevalence in Brazil on May 11th might not be accurate, because the data reported by the Brazilian Ministry of Health did not include extensive testing for SARS-CoV-2 across the full spectrum of the disease severity; therefore, the number of reported cases were likely to represent the more severe ones (because most tests were performed in symptomatic individuals and not in the general population) and, as a consequence, underestimating the actual disease prevalence.

Incidence is a measure of the occurrence of new cases during a specified period in a population at risk for the disease. Prevalence focuses on new and existing cases of the disease, whereas incidence focuses only on new cases (Table 1). To estimate incidence, all individuals in the denominator (population at risk) must have the potential to be in the numerator (those who develop the disease). Estimates of incidence require longitudinal follow-up (e.g., hours, days, or years). The study design of choice is cohort studies involving individuals at risk but without the disease at baseline who are followed through time and are evaluated if they develop the disease. Finally, the incidence also depends on the frequency of the disease, the definition of cases, and the population at risk. In the Brazilian scenario, the incidence of confirmed cases of COVID-19 on May 11th was 2.7/100,000 population at risk (Table 1).

KEY POINTS TO INTERPRET PREVALENCE AND INCIDENCE ESTIMATES

1. Accurate definitions of cases and noncases are critical to defining prevalence and incidence.
2. Both prevalence and incidence estimates can be misleading if the number of cases is underestimated due to barriers in accessing information regarding diagnosis and health care practices or if only patients with severe disease are submitted to diagnostic tests.
3. The timing of the estimates of prevalence and incidence must be taken into account when interpreting these measures. For example, the estimates might be lower in the beginning of an outbreak when compared with the epidemic later.

Table 1. Incidence and prevalence of COVID-19 on May 11 of 2020 in Brazil.[1]

Measure	Definition	How to calculate	Equation	Result
Prevalence	Existing cases of a disease at a point in time divided by the population at risk of having the disease	Cases of COVID-19 on May 11th ÷ Population at risk	168,331 ÷ 210 mi	0.08%
Incidence	New cases of a disease in a defined population over a period of time (a day, for example) divided by the population at risk	New cases of COVID-19 within a day ÷ Population at risk on May 11th[a]	5,632 ÷ 209,837,301	2.7/100,000

mi: million (Brazilian population). [a]Brazilian population minus the total number of confirmed cases on May 11th.

REFERENCES

1. Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde [homepage on the Internet]. Brasília: Ministério da Saúde; c2020 [cited 2020 May 11]. COVID-19 Painel Coronavírus. Available from: https://covid.saude.gov.br
2. Lipsitch M, Hayden FG, Cowling BJ, Leung GM. How to maintain surveillance for novel influenza A H1N1 when there are too many cases to count. Lancet. 2009;374(9696):1209-1211. https://doi.org/10.1016/S0140-6736(09)61377-5

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Programa de Pós-Graduação em Ciências Médicas, Universidade Federal de Santa Catarina – UFSC – Florianópolis (SC) Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
4. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2020 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

J Bras Pneumol. 2015;41(5):485-485
http://dx.doi.org/10.1590/S1806-37132015000000215

What does the p value really mean?

Juliana Carvalho Ferreira[1,3], Cecilia Maria Patino[2,3]

WHY CALCULATE A P VALUE?

Consider an experiment in which 10 subjects receive a placebo, and another 10 receive an experimental diuretic. After 8 h, the average urine output in the placebo group is 769 mL, versus 814 mL in the diuretic group—a difference of 45 mL (Figure 1). How do we know if that difference means the drug works and is not just a result of chance?

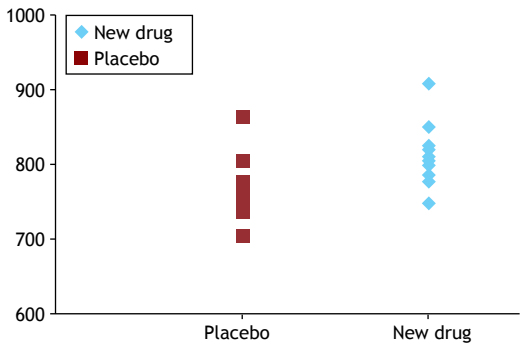

Figure 1. Urine output (mL) for each subject in the placebo (squares) and new drug groups (diamonds).

The most common way to approach this problem is to use statistical hypothesis testing. First, we state the null hypothesis of no statistical difference between the groups and the alternative hypothesis of a statistical difference. Then we select a statistical test to compute a test statistic, which is a standardized numerical measure of the between-group difference. Under the null hypothesis, we expect the test statistic value to be small, but there is a small probability that it is large, just by chance. Once we calculate the test statistic, we use it to calculate the p-value.

The p value is defined as the probability of observing the given value of the test statistic, or greater, under the null hypothesis. Traditionally, the cut-off value to reject the null hypothesis is 0.05, which means that when no difference exists, such an extreme value for the test statistic is expected less than 5% of the time.

Now let us go back to our case: we are comparing means and assuming that the data is normally distributed, so we use a t-test and compute a t-statistic of 2.34, with a p value of 0.031. Because we use a 0.05 cutoff for the p value, we reject the null hypothesis and conclude that there is a statistically significant difference between groups. So what does "p = 0.031" mean? It means that there is only a 3% probability of observing a difference of 45 mL in the average urine output between groups under the null hypothesis. Because this is a very small probability, we reject the null hypothesis. It does **not** mean that the drug is a diuretic, nor that there is 97% chance of the drug being a diuretic.

MISCONCEPTIONS ABOUT THE P VALUE

Clinical versus statistical significance of the effect size

There is a misconception that a very small p value means the difference between groups is highly relevant. Looking at the p value alone deviates our attention from the effect size. In our example, the p value is significant but a drug that increases urine output by 45 mL has no clinical relevance.

Nonsignificant p values

Another misconception is that if the p value is greater than 5%, the new treatment has no effect. The p value indicates the probability of observing a difference **as large or larger** than what was observed, under the null hypothesis. But if the new treatment has an effect of smaller size, a study with a small sample may be underpowered to detect it.

Overinterpreting a nonsignificant p value that is close to 5%

Yet another misconception is that if the p value is close to 5%, there is a trend towards a group difference. It is inappropriate to interpret a p value of, say, 0.06, as **a trend towards a difference**. A p value of 0.06 means that there is a probability of 6% of obtaining that result by chance when the treatment has no real effect. Because we set the significance level at 5%, the null hypothesis should not be rejected.

Effect sizes versus p values

Many researchers believe that the p value is the most important number to report. However, we should focus on the effect size. Avoid reporting the p value alone and preferably report the mean values for each group, the difference, and the 95% confidence interval—then the p value.

RECOMMENDED LITERATURE

1. Glantz SA. Primer in Biostatistics, 5th ed. New York: McGraw-Hill; 2002.

1. Divisão de Pneumologia, Instituto do Coração – InCor – Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brasil.
2. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
3. Methods in Epidemiologic, Clinical and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax.

BY-NC © 2015 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

J Bras Pneumol. 2015;41(6):565-566
http://dx.doi.org/10.1590/S1806-37562015000000314

Confidence intervals: a useful statistical tool to estimate effect sizes in the real world

Cecilia Maria Patino,[1,2] Juliana Carvalho Ferreira[2,3]

PRACTICAL SCENARIO

A prospective cohort study evaluated the association between the presence of asthma and the risk of developing obstructive sleep apnea (OSA) in adults. Adults were randomly recruited from a population-based list of state employees and were followed for four years. Participants with asthma, when compared with those without, had a higher risk of developing OSA in four years (relative risk [RR] = 1.39; 95% CI: 1.06-1.82; p = 0.03)

BACKGROUND

When conducting clinical research, we usually recruit a subgroup of the population of interest in order to increase study efficiency (fewer costs and less time). This subgroup of individuals, the study population, are those individuals who meet the inclusion criteria and agree to participate in the study (Figure 1). We then complete the study and calculate an effect size (e.g., a mean difference or a relative risk) to answer our research question. This process (inference) involves using data collected from the study population to estimate the true effect size in the population of interest, i.e., the source population. In our example, investigators recruited a random sample of state employees (source population) who were eligible and agreed to participate in the study (study population) and reported that asthma increases the risk of developing OSA in the study population (RR = 1.39). To take into account a sampling error due to recruiting only a subgroup of the population of interest, they also calculated a 95% confidence interval (around the estimate) of 1.06-1.82, indicating a 95% probability that the true RR in the source population would be between 1.06 and 1.82.

DEFINITION

A confidence interval is a measure of imprecision of the true effect size in the population of interest (e.g., difference between two means or a relative risk) estimated in the study population. That imprecision is due to the sampling error caused by taking subsamples of the population of interest. However, the estimate calculated in the study population is always the best estimate of the effect size in the source population.

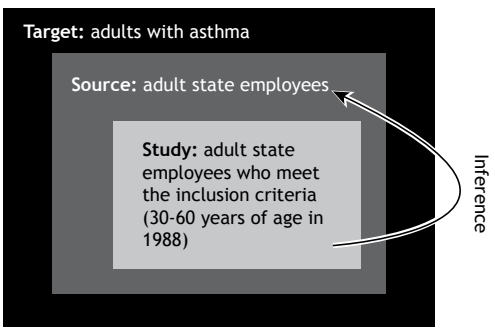

Figure 1. Research populations.

WHY DO WE NEED CONFIDENCE INTERVALS?

We need confidence intervals to indicate the amount of uncertainty or imprecision around the effect size calculated, using the study sample to estimate the true effect size in the source population. Calculating the confidence interval is a strategy that takes into account sampling error: the study effect size and its´ confidence interval represent plausible values for the source population, and the narrower the confidence interval is, the more certain we are that the estimate from the study population represents the true effect size in the source population.

CONFIDENCE INTERVALS: INTERESTING FACTS

The most common width of confidence intervals reported in the literature is the 95% confidence interval. However, if we are interested in more or less confidence, 90% or 99% confidence intervals can be used.

The confidence interval represents the uncertainty of the effect size in the source population, not in the study population. When calculating a confidence interval, the width of the interval is determined by the sample size (i.e., the individuals who agreed to be studied), the amount of measurement error of the study, and the degree of confidence required.

There is a unique relationship between the 95% confidence interval and a two-sided 5% level of significance. When the 95% confidence interval for differences in effect does not include 0 for absolute measures of association (e.g., mean differences) or 1 for relative measures of association (e.g., odds ratios), it can be inferred that the association is statistically significant ($p < 0.05$). The advantage of the 95% confidence interval over the p value is that it provides information about the size of the effect, the uncertainty of the population estimate, and the direction of the effect.

Confidence intervals should always be used in order to describe the major findings of a research study. The relevant confidence intervals should be shown not only in the text of the paper but also in the abstract.

1. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
2. Methods in Epidemiologic, Clinical and Operations Research–MECOR–Program, American Thoracic Society/Asociación Latinoamericana de Tórax.
3. Divisão de Pneumologia, Instituto do Coração – InCor – Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brasil.

BY-NC © 2015 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

RECOMMENDED READING

1. Teodorescu M, Barnet JH, Hagen EW, Palta M, Young TB, Peppard PE. Association between asthma and risk of developing obstructive sleep apnea. JAMA. 2015;313(2):156-64. http://dx.doi.org/10.1001/jama.2014.17822
2. Sedgwick P. Understanding confidence intervals. BMJ. 2014;349:g6051. http://dx.doi.org/10.1136/bmj.g6051
3. Sedgwick P. Confidence intervals: predicting uncertainty. BMJ. 2012;344:e3147. http://dx.doi.org/10.1136/bmj.e3147
4. Ferreira JC, Patino CM. What does the p value really mean? J Bras Pneumol. 2015;4195):485. doi: 10.1590/S1806-37132015000000215.
5. Gardner MJ, Altman DG. Statistics with confidence: confidence intervals and statistical guidelines. British Medical Journal. Belfast: The Universities Press Ltd; 1989.

J Bras Pneumol. 2018;44(5):353-353
http://dx.doi.org/10.1590/S1806-37562018000000303

Meeting the assumptions of statistical tests: an important and often forgotten step to reporting valid results

Cecilia Maria Patino[1,2,a] Juliana Carvalho Ferreira[1,3,b]

PRACTICAL SCENARIO

A secondary analysis of a randomized controlled trial was conducted to evaluate whether atopic status was associated with asthma severity and asthma control among inner-city adolescents in the USA.[1] To answer this question, the authors evaluated the differences between atopic and non-atopic patients, in terms of asthma control and asthma severity scores, using a t-test, and reported the results as means ± SDs. The results showed that the atopic patients, when compared with non-atopic patients, had similar asthma control scores (18.1 ± 4.2 vs. 18.2 ± 3.7; p = 0.95) but worse asthma severity scores (5.5 ± 2.9 vs. 4.7 ± 2.8; p = 0.04).

BACKGROUND

As part of the process of answering research questions using quantitative methods, investigators select a statistical analytical approach based on various characteristics of the study, such as the nature of the variables collected (e.g., continuous, categorical, time-to-event variables), as well as the study design. Once the analysis is completed, it is expected that investigators take an additional step in the analysis process to make sure that the a priori assumptions of the statistical test selected are met in the dataset assembled for the study.

All statistical tests have underlying assumptions that need to be met so that the test provides results that are valid (*without unacceptable error*) regarding the parameter the test is calculating (e.g., mean, proportion, odds ratio, etc.). In our example, the authors used a t-test to calculate the mean and the standard deviation of asthma control and severity asthma scores in atopic and non-atopic patients using data collected from the study population as a means to represent the truth in similar patients from the source population (adolescents with asthma in the USA). This process, called inference, is only valid if the assumptions of the statistical test are met (Table 1).

It is good practice, as investigators, to acknowledge that the assumptions of the statistical tests used to answer their research question have been evaluated and whether they were met or not. If the assumptions of the tests are met, which should be reported in the results section of the study, this assures the scientific community that the results of the study have met one of the important criteria related to their validity. However, it has been suggested that the assumptions of statistical techniques are often not checked[2] or reported. Reasons for not assessing assumptions include: 1. researchers being unaware of the assumptions of the statistical tests used in the study, such as a t-test, ANOVA, or regression analysis; 2. researchers being unaware of standard approaches used to check assumptions of statistical tests and evaluate if they are violated or not; 3. researchers being unaware of how to remedy violations of the assumptions of a statistical model or how to select a new test when violations cannot be remedied; and 4. researchers being confident in the robustness of the statistical test used and choosing not to check its assumptions.

As educators and investigators, we all need to contribute to the overall goal of reporting high quality research conducted among the populations we serve. Testing the assumptions of statistical tests or models used to answer our research questions is a good start!

Table 1. Example of assumptions of a statistical test.

Statistical test	Assumption	How to corroborate
t-Test	Sampling: The participants in the study are randomly sampled from the source population.	Check the protocol.
	Sample size: The sample size calculated for the study is achieved.	Check sample size calculation in the protocol and check if sample size was reached by the number of participants included in the study.
	Normal distribution: The scale of measurement of the outcome variable is continuous and is normally distributed (or at least symmetric).	Conduct descriptive statistics on the outcome variable and create a graph showing the distribution which should follow a bell curve.
	Homogeneity of variances: The variance (standard deviation) of the data collected on the continuous variable across the two comparison groups is similar.	Use valid statistical methods to test for homogeneity.

REFERENCES

1. Arroyave WD, Rabito FA, Carlson JC, Sever ML, Lefante J. Asthma severity, not asthma control, is worse in atopic compared with nonatopic adolescents with asthma. Ann Allergy Asthma Immunol 2016;116(1):18-25. https://doi.org/10.1016/j.anai.2015.10.015
2. Hoekstra R, Kiers HA, Johnson A. Are assumptions of well-known statistical techniques checked, and why (not)? Front Psychol. 2012;3:137. https://doi.org/10.3389/fpsyg.2012.00137

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
a. http://orcid.org/0000-0001-5742-2157; b. http://orcid.org/0000-0001-6548-1384

BY-NC © 2018 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

Estimating risk in clinical studies: odds ratio and risk ratio

Joaquin Maritano Furcada[1,2], Cecilia Maria Patino[1,3], Juliana Carvalho Ferreira[1,4]

PRACTICAL SCENARIO

A retrospective cohort study evaluated the association between type of ventilatory support and mortality among adult patients with interstitial lung disease and acute respiratory failure.[1] In comparison with noninvasive ventilation (NIV), invasive mechanical ventilation (IMV) increased mortality, with an OR of 26.0 (95% CI: 5.9-116.6) and a risk ratio (RR) of 2.2 (95% CI: 1.7-2.9), as detailed in table 1.

In our example, we calculated the OR and RR to answer the study question. It is important to understand the difference between these two statistical methods for calculating risk, which one is more applicable to answer this research question, and how they are interpreted.

The OR is defined as the ***ratio of odds*** of an event occurring, estimated by calculating the ratio of the number of times that the event of interest occurs to the number of times that it does not (ratio of events to non-events) between exposed and unexposed groups.[2] The RR is defined as the ***ratio of probabilities*** of an event occurring (ratio of events to subjects) between exposed and unexposed groups (Table 1). A positive association (increased risk) between exposure and outcome implies that the OR or RR is > 1.0, and a negative association (decreased risk) implies that the OR or RR is < 1.0.

HOW THE RR AND OR SHOULD BE INTERPRETED

The RR expresses that the "risk of the event (e.g., mortality) is X times larger/smaller in the exposed group than in the unexposed group". That statement is easily interpreted because it deals with probabilities (which range from 0 to 1). However, the OR is expressed as the ratio of the odds of event X in the exposed group to the odds of the same event in the unexposed group. Although we often use ORs to estimate RRs, they are different and ORs are not as intuitive to grasp and are therefore often misinterpreted.

The RR is commonly (and most correctly) used in order to estimate the risk of an event in randomized controlled trials, cohort studies, and cross-sectional studies, because all of those study designs calculate absolute risk and the RR can therefore be estimated. The OR is utilized to estimate risk in case-control studies, where the prevalence/incidence of the outcome cannot be estimated, given that the numbers of individuals with and without the outcome (cases and controls, respectively) are fixed by the investigators. The OR is also commonly used in order to calculate risk in cohort studies and randomized controlled trials when a logistic regression statistical model is employed to adjust for confounders or test for effect modification. However, we should bear in mind the fact that ORs and RRs are not equivalent.

In comparison with an RR, an OR tends to overestimate the strength of the association between exposure and outcome. However, the degree of overestimation is negligible in studies in which the outcome of interest occurs rarely (typically in < 10% of the participants). In our example, the ratio of the odds of death in the IMV group to that in the NIV group is 26 to 1 (OR = 26.0). In contrast, the risk of death in patients on IMV is 95%, compared with 43% for those on NIV (RR = 2.2). That RR is interpreted as the risk of death being 2.2 times higher in the IMV group than in the NIV group. This large difference between OR and RR is explained by the high proportion of participants who died in our example (41%). The OR estimates the RR more accurately when the study outcome is rare.

Table 1. Calculation of odds ratios and risk ratios for our example of noninvasive ventilation versus invasive mechanical ventilation in interstitial lung disease.

Group	Death	Survival	Total	Odds	Risk
IMV	39 (a)	2 (b)	41	Odds of death in IMV $\frac{a}{b} = \frac{39}{2} = 19.5$	Risk of death in IMV $\frac{a}{a+b} = \frac{39}{41} = 0.95$
NIV	32 (c)	43 (d)	75	Odds of death in NIV $\frac{c}{d} = \frac{32}{43} = 0.74$	Risk of death in NIV $\frac{c}{c+d} = \frac{32}{75} = 0.43$
Total	71	45	116	OR for death between groups $\frac{a}{b}/\frac{c}{d} = \frac{a*c}{b*d} = \frac{19.5}{0.74} = 26.4$	RR for death between groups $\frac{a}{a+b}/\frac{c}{c+d} = \frac{0.95}{0.43} = 2.2$

IMV: invasive mechanical ventilation; NIV: noninvasive ventilation; and RR: risk ratio. Adapted from Güngör et al.[1]

REFERENCES

1. Güngör G, Tatar D, Saltürk C, Çimen P, Karakurt Z, Kirakli C, et al. Why do patients with interstitial lung diseases fail in the ICU? a 2-center cohort study. Respir Care. 2013;58(3):525–531. https://doi.org/10.4187/respcare.01734
2. Sistrom CL, Garvan CW. Proportions, odds, and risk. Radiology. 2004;230(1):12–19. https://doi.org/10.1148/radiol.2301031028

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Pulmonary Medicine Department, Hospital Italiano de Buenos Aires, Buenos Aires, Argentina.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
4. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2020 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

Number needed to treat: a useful statistic to evaluate the impact of an intervention

Rogelio Perez-Padilla[1,2,a], Cecilia Maria Patino[1,3,b], Juliana Carvalho Ferreira[1,4,c]

PRACTICAL SCENARIO

A meta-analysis examined the effect of the use of daily medium-dose inhaled corticosteroids (ICS) on preventing exacerbations among preschoolers with recurrent wheeze. It summarized the results of 15 randomized clinical trials (RCTs) involving 3,278 individuals that showed that the use of daily ICS, compared with that of placebo, prevented exacerbations by 30% [risk ratio (RR) = 0.70; 95% CI: 0.61-0.79; number needed to treat (NNT) = 9].[1]

COMPARING RISKS

The impact of interventions can be estimated by comparing the incidence of the outcome (e.g., exacerbations) in the experimental group vs. a control group (e.g., placebo) by calculating an outcome ratio or difference across the intervention groups. The typical ratio calculated is the risk in the intervention group over the risk in the control group, designated the risk ratio (RR). In RCTs, the difference in risk between groups is called the absolute risk reduction (ARR), and it represents the proportion of outcomes reduced by the new intervention to the comparison group. Similar estimates can be calculated in observational studies replacing an intervention with the exposure of interest; for example, tobacco smokers compared with nonsmokers when reporting the risk of tobacco-related disease.

A statistic related to the ARR is the NNT, which is important because it provides an estimate of the number of patients that are required to be treated to avoid one additional patient from developing the outcome of interest (Table 1).[2]

The popularity of NNT has increased considerably, although this statistic is not necessarily easier to grasp than the ARR, either by patients or physicians. It is useful to remember that the lower the NNT, the higher the effectiveness of the intervention. In our example, an NNT of 9 is interpreted as follows: 9 children, on average, need to be treated with ICS to prevent 1 additional child from having an exacerbation.

More recently, RCTs also evaluate the impact of adverse events of an intervention by reporting the number needed to harm (NNH) in addition to the NNT. NNH is defined as the average number of individuals that would need to be exposed to a new intervention to produce one additional adverse outcome.

LIMITATIONS AND KEY POINTS

The reporting of NNT and NNH should always include confidence intervals and not only a point estimate.

The importance of taking into account the baseline risk to properly assess an intervention in an RCT cannot be overemphasized. Table 1 shows that as the incidence of the outcome in the control group increases, an identical RR results in greater ARRs and, consequently, lower NNT. Therefore, the effect can be exaggerated by simply reporting an RR of 0.75. In both examples, the risk is reduced by 25%, but NNT informs how many individuals must be treated in order to decrease that risk or absolute difference. It is recommended reporting both absolute and relative effect sizes.

The size and clinical impact of the effect of the intervention are important. Similar RCTs may have the same NNT, but their clinical relevance is different if the NNT refers to preventing one death or one COPD exacerbation when compared with preventing a small decrease in FEV_1 or another surrogate outcome.

Table 1. Comparing risks and interpreting results across different clinical scenarios. RR: risk ratio; ARR: absolute risk reduction; and NNT: number needed to treat.[2]

	Result	Interpretation
Low baseline risk (20% risk of death in the control group)		
Control group: n = 500; 100 (20%) deaths; Intervention group: n = 500; 75 (15%) deaths		
RR	15%/20% = 0.75	The intervention reduces risk **by** 25%
ARR	20% − 15% = 5%	The intervention reduces risk **in** 5%
NNT	1/5% =20	20 patients need to receive the intervention to prevent 1 death
High baseline risk (50% risk of death in the control group)		
Control group: n = 500; 250 (50%) deaths; Intervention group: n = 500; 188 (38%) deaths		
RR	38%/50% = 0.75	The intervention reduces risk **by** 25%
ARR	50% − 38% =12%	The intervention reduces risk **in** 12%
NNT	1/12% = 8	8 patients need to receive the intervention to prevent 1 death

REFERENCES

1. Kaiser SV, Huynh T, Bacharier LB, Rosenthal JL, Bakel LA, Parkin PC, et al. Preventing Exacerbations in Preschoolers with Recurrent Wheeze: A Meta-analysis. Pediatrics. 2016;137(6). pii: e20154496. https://doi.org/10.1542/peds.2015-4496
2. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med. 1988;318(26):1728-33. https://doi.org/10.1056/NEJM198806303182605

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Instituto Nacional de Enfermedades Respiratorias, Ciudad de Mexico, Mexico.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
4. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
a. http://orcid.org/0000-0002-1132-5308; b. http://orcid.org/0000-0001-5742-2157; c. http://orcid.org/0000-0001-6548-1384

BY-NC © 2019 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

What is the minimal clinically important difference, and why does it matter?

Maria do Socorro Simões[1,2], Cecilia M. Patino[1,3], Juliana Carvalho Ferreira[1,4]

PRACTICAL SCENARIO

In a hypothetical randomized controlled clinical trial, researchers compared the effect of bronchodilator A vs. bronchodilator B on FEV_1 in patients with COPD. Although the results showed that A was superior to B in improving FEV_1 and that this difference was statistically significant, did this change in FEV_1 result in fewer symptoms, or did it increase the participants' self-perceived ability to perform activities of daily living?

To answer these and other clinical outcome-related questions, it is crucial to understand the concept of the minimal clinically important difference (MCID).

DEFINING THE MCID

One of the many challenges of translating scientific evidence into clinical practice is the interpretation of data in light of clinical meaningfulness. We commonly find reports of statistical results, such as p-values, confidence intervals, and effect sizes. The MCID conveys results that are meaningful to patients. Depending on what outcome we are measuring, this change may be self-reported or objectively measured.

The MCID refers to the smallest change in an outcome that represents a meaningful change for the patient.[1,2] There are different methods to determine the MCID, but the major points are that the change has to be greater than the measurement error of the instrument that we are using to assess the outcome and it has to be large enough for patients to perceive the clinical change.

MCID IN RESEARCH AND CLINICAL SETTINGS

When designing studies that compare the effects of interventions, researchers should consider including thresholds for the MCID together with statistical significance.[2]

The MCID for a given test can be determined using expert consensus, using patient assessments anchoring the change to a subjective perception of change, or using statistical methods, which generally need validation. Interestingly, the same instrument may have different MCID thresholds according to specific study populations. For example, the six-minute walk test has different MCDI for patients with COPD, patients with heart failure, and apparently healthy adults.

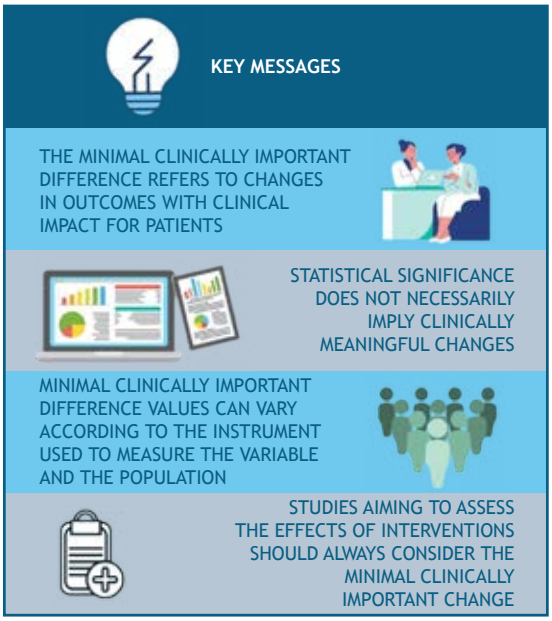

Figure 1. Key messages.

In the hypothetical trial of our practical scenario, the investigators found that the variation in FEV_1 was 241 ± 38 mL in the bronchodilator A group and 91 ± 14 mL in the bronchodilator B group. Considering that the MCID for FEV_1 in patients with COPD is 100 mL, we can conclude that bronchodilator A is statistically superior to bronchodilator B and that the change in FEV_1 is clinically meaningful.

CONCLUSION

Using patient-centered outcomes and aligning clinically relevant effects with statistical significance are important steps in the process of translating scientific clinical knowledge into evidence-based practice. Understanding the concept of the MCID is crucial to analyze and interpret the results of clinical interventions. In both research and clinical settings, we should consider MCIDs when analyzing and interpreting clinical outcome results (Figure 1).

REFERENCES

1. Maltenfort MG. The Minimally Important Clinical Difference. Clin Spine Surg. 2016;29(9):383. https://doi.org/10.1097/BSD.0000000000000446
2. Embry TW, Piccirillo JF. Minimal Clinically Important Difference Reporting in Randomized Clinical Trials. JAMA Otolaryngol Head Neck Surg. 2020;146(9):862-863. https://doi.org/10.1001/jamaoto.2020.1586

1. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Departamento de Ciências do Movimento Humano, Instituto de Saúde e Sociedade, Universidade Federal de São Paulo, São Paulo (SP) Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
4. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2021 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3756

Nonparametric statistical tests: friend or foe?

María Teresa Politi[1,2], Juliana Carvalho Ferreira[1,3], Cecilia María Patino[1,4]

PRACTICAL SCENARIO

The head of an ICU would like to assess if obese patients admitted for a COPD exacerbation have a longer hospital length of stay (LOS) than do non-obese patients. After recruiting 200 patients, she finds that the distribution of LOS is strongly skewed to the right (Figure 1A). If she were to perform a test of hypothesis, would it be appropriate to use a t-test to compare LOS between obese and non-obese patients with a COPD exacerbation?

PARAMETRIC VS. NONPARAMETRIC TESTS IN STATISTICS

Parametric tests assume that the distribution of data is normal or bell-shaped (Figure 1B) to test hypotheses. For example, the t-test is a parametric test that assumes that the outcome of interest has a normal distribution, that can be characterized by two parameters[1]: the mean and the standard deviation (Figure 1B).

Nonparametric tests do not require that the data fulfill this restrictive distribution assumption for the outcome variable. Therefore, they are more flexible and can be widely applied to various different distributions. Nonparametric techniques use ranks[1] instead of the actual values of the observations. For this reason, in addition to continuous data, they can be used to analyze ordinal data, for which parametric tests are usually inappropriate.[2]

What are the pitfalls? If the outcome variable is normally distributed and the assumptions for using parametric tests are met, nonparametric techniques have lower statistical power than do the comparable parametric tests. This means that nonparametric tests are less likely to detect a statistically significant result (i.e., less likely to find a p-value < 0.05 than a parametric test). Additionally, parametric tests provide parameter estimations—in the case of the t test, the mean and the standard deviation are the calculated parameters—and a confidence interval for these parameters. For example, in our practical

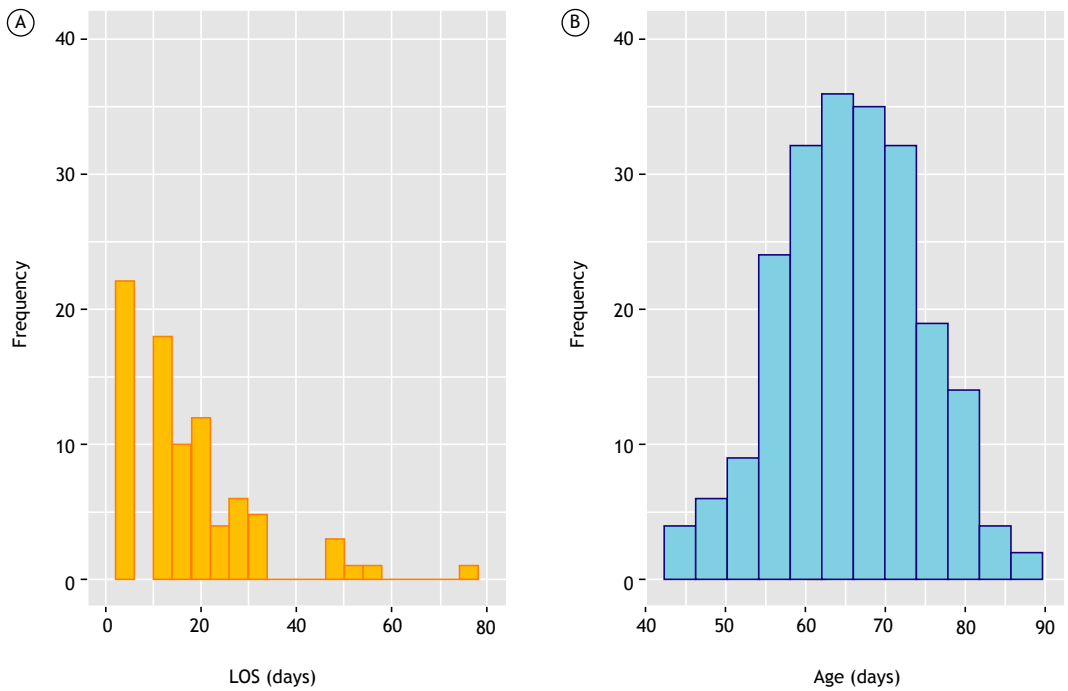

Figure 1. In A, hospital length of stay (LOS) of patients admitted for COPD exacerbations. The data clearly have a non-normal distribution and are skewed to the right. In B, age distribution of the same group of patients. The data are normally distributed (N = 200 patients).

1. Methods in Epidemiologic, Clinical, and Operations Research (MECOR) program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Laboratorio de Estadística Aplicada a las Ciencias de la Salud — LEACS — Departamento de Toxicología y Farmacología, Facultad de Ciencias Médicas, Universidad de Buenos Aires, Buenos Aires, Argentina.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
4. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

BY-NC © 2021 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3756

scenario, if the difference in LOS between the groups were analyzed with a t-test, it would report a sample mean difference in LOS between the groups and the standard deviation of that difference in LOS. Finally, the 95% confidence interval of the sample mean difference could be reported to express the range of values for the mean difference in the population. Conversely, nonparametric tests do not estimate parameters such as mean, standard deviation, or confidence intervals. They only calculate a p-value.[2]

HOW TO CHOOSE BETWEEN PARAMETRIC AND NONPARAMETRIC TESTS?

When sample sizes are large, that is, greater than 100, parametric tests can usually be applied regardless of the outcome variable distribution. This is due to the central limit theorem, which states that if the sample size is large enough, the distribution of a given variable is approximately normal. The farther the distribution departs from being normal, the larger the sample size will be necessary to approximate normality.

When sample sizes are small, and outcome variable distributions are extremely non-normal, nonparametric tests are more appropriate. For example, some variables are naturally skewed, such as hospital LOS or number of asthma exacerbations per year. In these cases, extremely skewed variables should always be analyzed with nonparametric tests, even with large sample sizes.[2]

In our practical scenario, because the distribution of LOS is strongly skewed to the right, the relationship between obesity and LOS among the patients hospitalized for COPD exacerbations should be analyzed with a nonparametric test (Wilcoxon rank sum test or Mann-Whitney test) instead of a t-test.

REFERENCES

1. Whitley E, Ball J. Statistics review 6: Nonparametric methods. Crit Care. 2002;6(6):509-513. https://doi.org/10.1186/cc1820
2. e Cessie S, Goeman JJ, Dekkers OM. Who is afraid of non-normal data? Choosing between parametric and non-parametric tests. Eur J Endocrinol. 2020;182(2):E1-E3. https://doi.org/10.1530/EJE-19-0922

J Bras Pneumol. 2017;43(3):162-162
http://dx.doi.org/10.1590/S1806-37562017000000170

Subgroup analysis and interaction tests: why they are important and how to avoid common mistakes

Juliana Carvalho Ferreira[1,2], Cecilia Maria Patino[1,3]

A randomized clinical trial was conducted to compare the effect of vitamin C vs. placebo on improving pulmonary function in newborns of pregnant smokers; and to test if this effect differed by maternal genotype.[1] Vitamin C improved pulmonary function in newborns compared to placebo (TPTEF:TE ratio 0.383 vs 0.345; p = 0.006); and this effect was stronger in newborns with mothers with a specific genotype (p-*interaction* < 0.001).[1]

BACKGROUND

When conducting clinical trials, investigators examine the effect of interventions on outcomes in the study population and often in subgroups of patients defined by baseline characteristics (e.g., demographics, prognostic factors). The goal is to understand if the magnitude of the effect of the intervention differs within categories of a subgroup; in our example, genotype subgroups. If the effect is different within subgroups we call this effect modification of the intervention on the outcome due to the additional presence of the subgroup variable. We commonly conduct a test for interaction, using multivariable models, to evaluate for statistically significant subgroup differences. If the p value is significant, we conclude that the effect of the intervention on the outcome differs within subgroups, in our example, maternal genotype.

Understanding treatment effects across patient subgroups is important because it helps identify patient groups that respond better or worse to the intervention. However, subgroup analyses should be done with caution to avoid common mistakes that either lead to false negative or positive findings, especially when they are not pre-specified in the analysis plan before starting the study. A common mistake is to compare the effect of treatment on the outcome separately within each subgroup. For example, comparing the effect of vitamin C vs. placebo on pulmonary function in newborns among mothers with one genotype and then separately among the mothers with another genotype. This approach is incorrect because it leads to multiple testing, which means that instead of using only one calculation to test for differences in effect across subgroups (p for interaction across genotype-groups in our example), we use two or more different calculations for each subgroup analysis. Every time we add a calculation, we no longer can use the standard significant level of p < 0.05. In this case, since there are two calculations we would need to divide the p value by 2 and use p < 0.025 as the significance level.[2] Thus, we would overestimate subgroup differences if we kept the significance level at 0.05. Another challenge with subgroup analysis is that results may suggest that there are subgroup differences but the p-value is not statistically significant because the sample size within each subgroup is too small (Figure 1).

SUBGROUP ANALYSIS TIPS

1. Identify a few subgroups that seem highly relevant to your research question *a priori* and justify your choices.
2. Do not compare the effects of treatment vs. control in each subgroup. There are specific statistical tests to determine if there is an interaction between the treatment effect and the variables that define subgroups, which are best performed with the aid of a statistician.
3. Before making changes in clinical practice, subgroup results should be replicated in other studies.

Figure 1. Consider a hypothetical randomized trial with 30 participants, 15 in the treatment group (9 men and 6 women) and 15 in the control group (7 men and 8 women). To test if the effect of treatment differs between men and women, the correct approach is to use a multivariate model including an interaction term (treatment vs. sex), but with 30 participants, such a model would probably be underpowered to detect clinically significant differences. Comparing the effect of treatment vs. control in women only (6 vs. 8 participants), then in men only (7 vs. 8 participants) would also be underpowered.

REFERENCES

1. McEvoy CT, Schilling D, Clay N, Jackson K, Go MD, Spitale P, et al. Vitamin C supplementation for pregnant smoking women and pulmonary function in their newborn infants: a randomized clinical trial. JAMA. 2014;311(20):2074-82. https://doi.org/10.1001/jama.2014.5217
2. Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine–reporting of subgroup analyses in clinical trials. N Engl J Med. 2007;357(21):2189-94. https://doi.org/10.1056/NEJMsr077003

1. Methods in Epidemiologic, Clinical, and Operations Research – MECOR – program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração – InCor – Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

BY-NC © 2017 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

Linear and logistic regression models: when to use and how to interpret them?

Horacio Matias Castro[1,2], Juliana Carvalho Ferreira[1,3]

PRACTICAL SCENARIO

A secondary analysis[1] of a study designated "Integrating Palliative and Critical Care," a cluster randomized trial, was conducted to explore differences in receipt of elements of palliative care among patients who died in the ICU with interstitial lung disease (ILD) or COPD in comparison with those who died of cancer. The authors used two methods of multiple regression analysis: linear regression to estimate the impact of COPD and ILD, in comparison with that of cancer, on the length of ICU stay, and logistic regression to evaluate the effects of COPD and ILD on the presence or absence of elements of palliative care. All regression models were adjusted for confounders (age, sex, minority status, education level, among others) of the association between the patient diagnosis and palliative care outcomes.

INTRODUCTION

Linear and logistic regressions are widely used statistical methods to assess the association between variables in medical research. These methods estimate if there is an association between the independent variable (also called predictor, exposure, or risk factor) and the dependent variable (outcome).[2]

The association between two variables is evaluated with simple regression analysis. However, in many clinical scenarios, more than one independent variable may be associated with the outcome, and there may be the need to control for confounder variables. When more than two independent variables are associated with the outcome, multiple regression analysis is used. Multiple regression analysis evaluates the independent effect of each variable on the outcome, adjusting for the effect of the other variables included in the same regression model.

WHEN TO USE LINEAR OR LOGISTIC REGRESSION?

The determinant of the type of regression analysis to be used is the nature of the outcome variable. Linear regression is used for continuous outcome variables (e.g., days of hospitalization or FEV1), and logistic regression is used for categorical outcome variables, such as death. Independent variables can be continuous, categorical, or a mix of both.

In our example, the authors wanted to know if there was a relationship between cancer, COPD, and ILD (baseline disease; the independent variables) with two different outcomes. One outcome was continuous (length of ICU stay) and the other one was categorical (presence or absence of elements of palliative care). Therefore, two models were built: a linear model to examine the association between baseline disease (chronic pulmonary disease or cancer) and length of ICU stay, and a logistic regression analysis to examine the association between the baseline disease and being in receipt of elements of palliative care.

HOW TO INTERPRET RESULTS OF REGRESSION ANALYSIS?

Regression models are performed within statistical packages, and the output results include several parameters, which can be complex to interpret. Clinicians who are learning the basics of regression models should focus on the key parameters presented in Chart 1.

In our example, the baseline disease—COPD, ILD, or cancer (the reference category)—is the independent variable, and length of ICU stay and receipt of palliative care elements are the outcomes of interest. In addition, the regression models also included other independent variables considered as potential confounders, such as age, sex, and minority status. In the linear regression model, the length of ICU stay for patients with ILD was longer than for those with cancer (β = 2.75; 95% CI, 0.52-4.98; p = 0.016), which means that, on average, having ILD increased the length of ICU stay in 2.75 days when compared with the length of ICU stay among cancer patients. In the logistic regression model, the authors found that patients with ILD, when compared with cancer patients, were less likely to have any documentation of their pain assessment in the last 24 h of life (OR = 0.43; 95% CI, 0.19-0.97; p = 0.042), which means that having ILD decreased the odds of documentation of pain assessment by more than half.

KEY POINTS

- Linear and logistic regressions are important statistical methods for testing relationships between variables and quantifying the direction and strenght of the association.
- Linear regression is used with continuous outcomes, and logistic regression is used with categorical outcomes.
- These procedures require expertise in regression model building and typically require the assistance of a biostatistician.

1. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Pulmonary Medicine Department, Hospital Italiano de Buenos Aires, Buenos Aires, Argentina.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2022 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3756

Chart 1. Most important parameters in regression analyses and their interpretations.

Parameter	Linear regression	Logistic regression
Direction and strength of the association between the independent variable and the dependent variable (outcome)	Beta coefficient: Describes the (expected) average change in the outcome variable for each one-unit change in the independent variable for continuous variables, or the average change in the outcome variable for one category of the independent variable compared with a reference category for categorical variables	OR: The OR for a continuous independent variable is interpreted as the change in the odds of the outcome occurring for every one-unit increase in the independent variable The OR for categorical independent variables is interpreted as the increase or decrease in odds between two categories (e.g., men vs women) OR = 1: no association; OR > 1: positive association or risk factor; and OR < 1: negative association or protective factor
Example (for a continuous independent variable)	The expected increase in FEV_1 for each centimeter increase in height	The expected increase in the odds of death for each increase of one year of age among patients with sepsis
Example (for a categorical independent variable)	The expected increase in FEV_1 for men compared with women with the same height and age	The expected increase in the odds of death for men compared with women among COVID-19 patients
Precision of the estimate	The 95% CI of the beta coefficient	The 95%CI of the OR
Statistical significance	The p value (significant when < 0.05)	The p value (significant when < 0.05)

REFERENCES

1. Brown CE, Engelberg RA, Nielsen EL, Curtis JR. Palliative Care for Patients Dying in the Intensive Care Unit with Chronic Lung Disease Compared with Metastatic Cancer. Ann Am Thorac Soc. 2016;13(5):684-689. https://doi.org/10.1513/AnnalsATS.201510-667OC
2. Bzovsky S, Phillips MR, Guymer RH, Wykoff CC, Thabane L, Bhandari M, et al. The clinician's guide to interpreting a regression analysis. Eye (Lond). 2022;36(9):1715-1717. https://doi.org/10.1038/s41433-022-01949-z

J Bras Pneumol. 2016;42(1):77-77
http://dx.doi.org/10.1590/S1806-37562016000000013

What is survival analysis, and when should I use it?

Juliana Carvalho Ferreira[1,2], Cecilia Maria Patino[2,3]

INTRODUCTION

Researchers used a national registry of lung cancer patients in the United States to identify the impact of tumor size and histological type on patient survival. They included 7,965 patients treated between 1988 and 2000. As can be seen in Figure 1, they found that survival times were shorter among the patients with larger tumors (> 5 cm) than among those with smaller tumors (< 1 cm).

The misconception that mortality and survival are interchangeable comes from the lay use of the terms. However, in biostatistics, survival is a concept derived from a specific analytical procedure, whereas mortality is a dichotomous outcome variable usually compared between or across two or more groups at a specific time point (for example, at five years). Survival, in turn, deals with a time-to-event variable: it measures the time between the beginning of observation until the occurrence of an event.

WHY USE SURVIVAL ANALYSIS?

Survival analysis is important when the time between exposure and event is of clinical interest. In our example, five-year survival among patients with tumors < 1 cm was 85%, compared with 52% among those with tumors > 5 cm. Of the patients in that latter group (the high-risk group), approximately half were dead in five years. However, knowing that survival after two years was 70% is also clinically relevant. For highly lethal diseases, like metastatic cancer, a subgroup submitted to a new treatment might have a survival advantage in the first three years but similar mortality after five years. Comparing mortality at the end of the period does not distinguish between longer and shorter survival times.

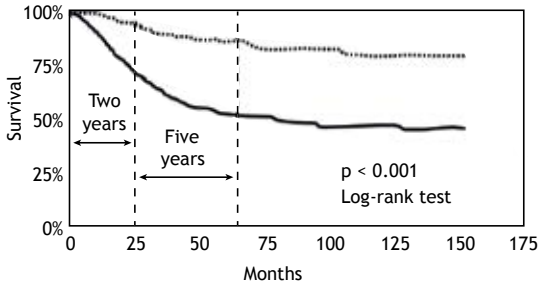

Figure 1. Survival for participants with larger tumors (> 5 cm, solid line) and for those with smaller tumors (< 1 cm, dotted line). Adapted from Ost et al. (Recommended reading).

Calculating survival is also useful for methodological reasons; for example, when study participants are lost to follow-up. When the study ends, investigators might not know if a given participant is dead or alive, but they know that he or she was alive at least until their last visit. In addition, some participants might be followed for less than five years because they enter the study at a later date. When the study ends, they might have not experienced the event because their follow-up was interrupted. In Figure 1, it can be seen that survival continues to decrease from two to five years. In survival analysis, data related to participants who did not experience the event by the end of the study or were lost to follow-up are censored: they contribute to the analysis up to the last point at which the investigators knew that the participants were still alive.

Survival analysis uses conditional probability; that is, the probability of surviving up to time *t*, given that a subject was alive at the beginning of a specified time interval. The Kaplan-Meier method is used in order to estimate survival probability at several time intervals and to graphically illustrate survival over time. The log-rank test is a nonparametric test used in comparing survival curves between two or more groups.

SOME INTERESTING FACTS ABOUT SURVIVAL

In survival analysis, censored data are not the same as missing data. Participants whose data are censored are not excluded and contribute time at risk to the analysis up to the last interval during which they were alive. Therefore, imputation methods are not needed.

Censoring due to loss to follow-up is only acceptable for a small percentage of cases and when the prognosis of participants lost to follow-up is assumed to be the same as those remaining in the study.

The outcome of survival analysis does not have to be time-to-death; it can be other time-to-event outcomes, such as time-to-pregnancy after fertility treatment and time-to-ventilator weaning.

RECOMMENDED READING

1. Ost D, Goldberg J, Rolnitzky L, Rom WN. Survival after surgery in stage IA and IB non-small cell lung cancer. Am J Respir Crit Care Med. 2008;177(5):516-23. http://dx.doi.org/10.1164/rccm.200706-815OC
2. Glantz SA. How to analyse survival data. In: Glantz SA. Primer in Biostatistics. 7th ed. New York: McGraw-Hill Medical; 2011. p.229-44.

1. Divisão de Pneumologia, Instituto do Coração – InCor – Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brasil.
2. Methods in Epidemiologic, Clinical and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

BY-NC © 2016 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

Receiver operating characteristic analysis: an ally in the pandemic

Jezreel Pantaleón García[1,2], Juliana Carvalho Ferreira[1,3], Cecilia Maria Patino[1,4]

PRACTICAL SCENARIO

From a global public health perspective, a diagnostic test that accurately discriminates between positive and negative COVID-19 cases is critical to allocate human and material resources to manage the pandemic.[1] The ongoing COVID-19 pandemic has led to the expeditious development of multiple diagnostic tests to detect the SARS-CoV-2 infection. Thus, clinicians, researchers, and policy makers need to understand how to interpret the performance level of such diagnostic tests[1] to support the multilevel decision-making process. Here, we provide an overview of a commonly used tool to evaluate the accuracy of diagnostic or prognostic tests: the ROC curve.

ROC ANALYSIS

We use ROC analysis to graphically display, compare, and evaluate the accuracy of current and novel diagnostic tests. In order to do so, ROC curves integrate three related measures of accuracy: sensitivity (true positives), specificity (true negatives), and AUC.[2] These measures are calculated for any diagnostic test by comparing the test result (positive or negative) against a well-known gold standard that determines the true disease status in each case.

UNDERSTANDING ROC CURVES

ROC curves are created by plotting sensitivity (true positives) on the *y* axis against 1 − specificity (true negatives) on the *x* axis for every value found in a sample of subjects with and without the disease. It is expected that higher values would be more common among the subjects with the disease, and lower values would be more common among the subjects without the disease. In a perfect test, an obvious cutoff threshold can be identified that differentiates subjects with the disease from those without the disease, sensitivity and specificity being both 100%. Such a perfect differentiation is rarely the case for tests in real life, so ROC curves plot the trade-off between sensitivity and specificity for all possible cutoffs and the overall test accuracy. To express the diagnostic accuracy of a test numerically, we calculate the AUC, which estimates the probability of a random subject with the disease to have a higher value on the test than a subject without the disease. The probability ranges from 0% (AUC = 0) to 100% (AUC = 1).

USING ROC CURVES

Relative shapes of ROC curves within the plot are a quick approach to estimate and compare the accuracy between diagnostic tests (Figure 1). A perfect diagnostic test (AUC = 1.0) correctly identifies all positive and all negative results as diseased and non-diseased, respectively, and would reach the far top left. In contrast, a test that is inaccurate, or similar to flipping a coin, would result in a 45-degree line (AUC = 0.5). These two extremes (perfect test and uninformative test) are often used as references: ROC curves closer to a perfect diagnostic test have a higher AUC and are more accurate than are those closer to the random error line (AUC ~0.5).[2] Therefore, comparing multiple ROC curves may be an intuitive strategy to help us decide which the most accurate test for our clinical practice is. However, since there is always a trade-off between sensitivity and specificity, tests should not be evaluated by the AUC alone. In some cases, a test is more useful when it has high sensitivity (and, therefore, lower specificity), as when you cannot afford to miss the diagnosis. An example is when you are using a test to diagnose COVID-19. In that case, a test with lower

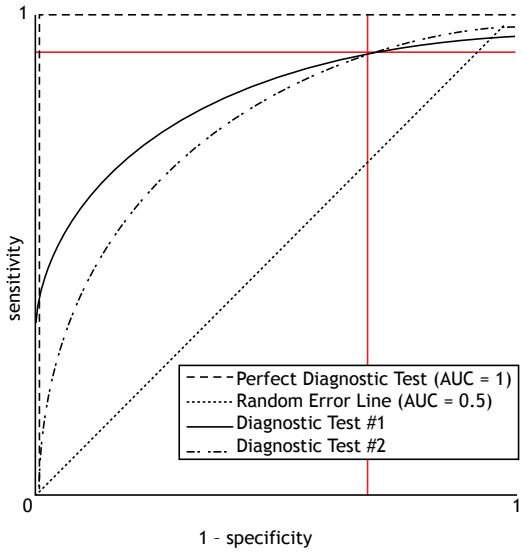

Figure 1. Comparative examples of ROC curves. ROC curve plots illustrating the accuracy performance of a perfect diagnostic test (AUC = 1), a random error line (AUC = 0.5) of an uninformative test, and two hypothetical diagnostic tests. Red lines depict a clinically relevant threshold of high sensitivity range in which the AUC of Diagnostic Test #2 outperforms Diagnostic Test #1.

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Department of Pulmonary Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
4. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

BY-NC © 2021 Sociedade Brasileira de Pneumologia e Tisiologia

AUC that has a high sensitivity may be more useful in certain clinical scenarios than a test with slightly higher AUC with lower sensitivity (and greater specificity).

REFERENCES

1. Butler-Laporte G, Lawandi A, Schiller I, Yao M, Dendukuri N, McDonald EG, et al. Comparison of Saliva and Nasopharyngeal Swab Nucleic Acid Amplification Testing for Detection of SARS-CoV-2: A Systematic Review and Meta-analysis [published correction appears in doi: 10.1001/jamainternmed.2021.0245]. JAMA Intern Med. 2021;181(3):353-360. https://doi.org/10.1001/jamainternmed.2020.8876
2. Ferreira JC, Patino CM. Understanding diagnostic tests. Part 3. J Bras Pneumol. 2018;44(1):4. https://doi.org/10.1590/s1806-37562018000000017

Propensity scores: a tool to help quantify treatment effects in observational studies

Cecilia Maria Patino[1,2], Juliana Carvalho Ferreira[1,3]

PRACTICAL SCENARIO

To evaluate the effect of early high-frequency oscillatory mechanical ventilation (MV) vs. conventional MV on duration of MV and in-hospital mortality among children with acute respiratory failure, a retrospective cohort study was conducted using data from a randomized controlled trial (RCT).[1] Multivariable models, adjusted for confounding factors using a propensity score (PS), showed that the children on high-frequency oscillatory MV, when compared with those on conventional MV, were less likely to discontinue MV (hazard ratio = 0.75; 95% CI: 0.64-0.89; p = 0.001) and not at increased risk of in-hospital mortality (odds ratio = 1.28; 95% CI: 0.92-1.79; p = 0.15).

BACKGROUND

To evaluate the effect of interventions on health-related outcomes, RCTs are considered the gold standard study design because randomization gives every study participant a pre-established probability of being assigned to either an intervention or a comparison group. The goal is to prevent selection bias and confounding[2] at baseline by yielding the two groups with a similar distribution of measured and unmeasured confounders so that study results reflect the effect of the intervention on the outcome.

When conducting an RCT is not a feasible or ethical option, observational studies about interventions using PS to mimic randomization effects may be an alternative. The PS is a new composite variable that is created by combining a set of confounding variables that increase the probability of an individual being assigned to a specific intervention (treatment A vs. treatment B) and then incorporated into the analysis. In our example, the goal was to evaluate the effect of two MV strategies (intervention) on duration of MV and in-hospital mortality (outcomes). To mimic the effects of randomization and make both groups similar regarding confounding variables, a PS was created, based on variables clinicians utilize to assign the specific MV strategy and included it in the multivariable analysis as a covariate to adjust for confounders.

PROPENSITY SCORE

Definition: a variable that results from calculating the likelihood (propensity) of each participant receiving a treatment conditional on values of variables thought to influence the decision to prescribe treatment A or B.

Variable selection: Researchers select variables for PS based on their effect as confounders or predictors of the exposure (the intervention). Typical variables included in PS are demographics (age, gender, and socioeconomic status), disease severity, and characteristics of the treatment environment (characteristics of physicians and their practice). The variables are included as exposure variables in a logistic regression model with the intervention as the outcome. This model calculates a score for each participant representing their estimated likelihood of receiving treatment A or B, conditional on a weighed score of the values of that participant on the set of exposure variables used to create the PS.

Analytical methods: Four[3] strategies are typically used in observational studies (Table 1), each having advantages and disadvantages. We recommend consulting with a biostatistician to guide all PS processes.

Table 1. Methods used in order to include propensity scores in observational studies.

Method	Description
Stratification	Strata are created with the participants that present with equal values in the propensity score. Weighted averages within strata are calculated before the multivariable analysis is conducted.
Matching	Each exposed participant (treatment A) is matched to an unexposed participant (treatment B) with same propensity score value before the multivariable analysis is conducted.
Inverse weighting	Two potential samples are created to represent samples that would have been observed if everyone had been exposed to the treatment or no one had been exposed to it.
Covariate adjustment	A regression model of the intervention on the outcome is fit to the both the intervention group (exposure) and the propensity score (covariate).

REFERENCES

1. Bateman ST, Borasino S, Asaro LA, Cheifetz IM, Diane S, Wypij D, et al. Early High-Frequency Oscillatory Ventilation in Pediatric Acute Respiratory Failure. A Propensity Score Analysis. Am J Respir Crit Care Med 2016:193(5):495-503. https://doi.org/10.1164/rccm.201507-1381OC
2. Ferreira JC, Patino CM. Randomization: beyond tossing a coin. J Bras Pneumol. 2016;42(5):310. https://doi.org/10.1590/S1806-37562016000000296
3. Williamson E, Morley R, Lucas A, Carpenter J. Propensity scores: from naïve enthusiasm to intuitive understanding. Stat Methods Med Res. 2012;21(3):273-93. https://doi.org/10.1177/0962280210394483

1. Methods in Epidemiologic, Clinical and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
3. Divisão de Pneumologia, Instituto do Coração – InCor – Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brasil.

BY-NC © 2017 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

J Bras Pneumol. 2016;42(4):240-240
http://dx.doi.org/10.1590/S1806-37562016000000225

Test for trend: evaluating dose-response effects in association studies

Cecilia Maria Patino[1,2], Juliana Carvalho Ferreira[2,3]

PRACTICAL SCENARIO

In a prospective cohort study, the association between maternal stress and asthma in children was evaluated. The authors were interested in determining the effect that increasing levels of maternal stress has on the prevalence of asthma in offspring by 6 years of age. Maternal stress was measured using a questionnaire and categorized as 0, 1-2, 3-4, or ≥ 5 negative pre- and post-natal life events; asthma was diagnosed by physicians. The authors reported that increasing levels of maternal stress, compared with zero negative events, were associated with increasing odds of offspring having asthma during childhood: for 1-2 events, OR = 1.30 (95% CI: 0.72-2.37); for 3-4 events, OR = 1.92 (95% CI: 1.03-3.57); and for ≥ 5 events, OR = 3.52 (95% CI: 1.79-6.93). The p for trend (ptrend) was < 0.01 (Figure 1).

BACKGROUND

When conducting studies that evaluate the association between a risk factor with more than two categories and with a natural ordering—in our example, maternal stress is an ordinal variable with four levels—we can choose to evaluate the association between each category of the risk factor and the outcome by comparing it with a reference level or determining whether increasing or decreasing levels of the risk factor are associated with increasing or decreasing levels of the outcome.[1,2] This analysis is called *testing for dose-response* or *testing for trend* of the effect of the risk factor on the outcome.[3]

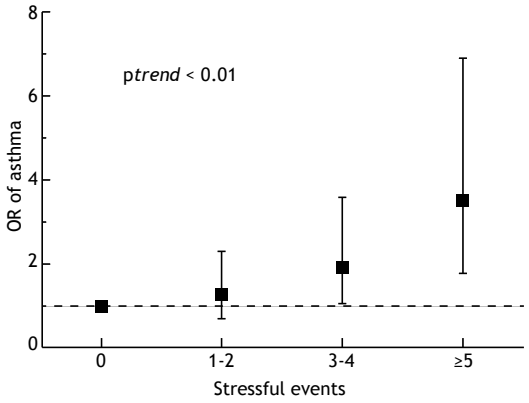

Figure 1. Association between the number of maternal pre- and post-natal stressful events and the odds of childhood asthma in offspring. Black squares represent ORs, and error bars are the 95% confidence intervals. A dose-response effect is confirmed with the test for trend (p*trend*)[2].

In our example, the authors were interested in determining whether an increasing number of maternal stressful life events was associated with increased odds of offspring being diagnosed with asthma. They first reported the effects that each level of maternal stress had on the diagnosis of asthma in offspring by reporting the OR and 95% CI compared with the lowest level of such stress; then, they reported the dose-response effect or trend of the effect of maternal stress on asthma in offspring by calculating and reporting the p*trend,* which was statistically significant. If we consider the OR for each level of stress, compared with zero negative events, we see that all levels of exposure to maternal stress increased the risk of asthma in offspring (all ORs > 1.0), although only 3-4 events and ≥ 5 events were statistically significant, indicated by the fact that the corresponding 95% CI did not include 1.0. However, the p*trend* indicates that increasing maternal stress across *all* levels increases the odds of physician-diagnosed asthma in offspring.

Regression methods are commonly used to test for trend.[3] When reporting a test for trend, we usually list each category of the risk factor and the strength of the effect (i.e., odds ratio) of each category on the outcome compared with the reference level, the p value at each level, and additionally the p*trend.* The p*trend* is the unique information we need in order to determine whether there is a dose-response effect.

WHY TEST FOR TREND?

As shown in our example, a test for trend can demonstrate a dose-response association between the risk factor and the outcome even if the association is not statistically significant for any particular level of exposure. Translated to clinical decision-making, knowledge of a dose-response association can help clinicians and patients understand that any increase in the level of exposure to a modifiable risk factor (e.g., maternal stress, cigarette smoking, and air pollution) increases the effect of that risk factor on a particular outcome.

RECOMMENDED READING

1. Lee A, Mathilda Chiu YH, Rosa MJ, Jara C, Wright RO, Coull BA, et al. Prenatal and postnatal stress and asthma in children: Temporal- and sex-specific associations. J Allergy Clin Immunol. Mar 4. pii: S0091-6749(16)00191-3. [Epub ahead of print] http://dx.doi.org/10.1016/j.jaci.2016.01.014
2. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health. 2005;95 Suppl 1:S144-50. http://dx.doi.org/10.2105/AJPH.2004.059204
3. Vitinghoff E, Glidden DV, Shiboski ST, McCulloch. Regression Methods in Biostatistics. Linear, Logistic, Survival and Repeated Measure Models. 2nd ed. New York, NY: Springer; 2012.

1. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
2. Methods in Epidemiologic, Clinical and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax.
3. Divisão de Pneumologia, Instituto do Coração – InCor – Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brasil.

BY-NC © 2016 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

Fragility index and fragility quotient in randomized clinical trials

Marcos Vinicius Fernandes Garcia[1], Juliana Carvalho Ferreira[2,3,4], Pedro Caruso[2,3]

CLINICAL SCENARIO

In a randomized clinical trial (RCT) by Meyer et al.,[1] tenecteplase plus heparin was compared with placebo plus heparin in patients with pulmonary embolism. The primary outcome (death or hemodynamic decompensation) occurred in 13 of 506 patients (2.6%) in the intervention group as compared to 28 of 499 patients (5.6%) in the control group (OR = 0.44; 95% CI: 0.23-0.87; p = 0.02).

RCT ROBUSTNESS

RCTs are expensive and time consuming, and they generally have limited sample sizes; therefore, results can be dependent on few events. In the abovementioned RCT,[1] despite the large sample size, if only 3 more patients in the intervention group had experienced the outcome, the p-value would be greater than 0.05, which means that if 16 patients, rather than 13 patients, in the experimental group had experienced the primary outcome, the study would not be significant. This number indicating how many additional events in one of the groups would be required to turn a statistically significant trial into a statistically non-significant trial is called the fragility index (FI).

FI

The FI is calculated by changing the status of 1 patient in the group with the fewest number of events (control or experimental) from "non-event" (not experiencing the primary outcome) to "event" and then recalculating a two-sided Fisher's exact test until p becomes ≥ 0.05.[2] Table 1 illustrates the calculation of the FI for the abovementioned RCT.[1] Therefore, the FI is a measure of robustness of clinical trial results; the smaller the FI is, the less robust the trial is considered to be. Although the FI has no formal cutoff, it serves as an additional indicator of how easily the statistical significance of an RCT depends on a small number of events. Also, as a rule of thumb, if the number of patients lost to follow-up is greater than the FI, the trial should be considered less robust.

The fragility quotient (FQ) is the FI divided by the sample size, and a low FQ indicates a less robust trial. The FQ for the abovementioned RCT[1] would be 3/1,005 = 0.003, which is small and also indicates that the trial is not robust. FQ provides a way to assess the vulnerability of studies with regard to sample size, especially when sample sizes vary widely between studies addressing the same intervention.

FI USE AND LIMITATIONS

A large FI does not necessarily indicate a conclusive result, and a small FI does not indicate that the RCT results are trivial. There is no clear consensus on defining what a "fragile" study is, but FI and FQ can help clinicians make health decisions considering the fragility of RCT results.

Table 1. Example of calculation of the fragility index.[a]

Study sample (N = 1,005)	Death or hemodynamic decompensation	Neither death nor hemodynamic decompensation	p
Study outcome			0.02
Intervention group	13	493	
Control group	28	471	
First step of FI calculation			0.027*
Intervention group	14	492	
Control group	28	471	
Second step of FI calculation			0.043*
Intervention group	15	491	
Control group	28	471	
Third step of FI calculation			0.065*
Intervention group	16	490	
Control group	28	471	

FI: fragility index. [a]Adapted from Meyer et al.[1] *Fragility index steps and p-values were calculated using the "R package: fragility index."

1. Cleveland Clinic Foundation, Cleveland (OH) USA.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Unidade de Terapia Intensiva, AC Camargo Cancer Center, São Paulo (SP) Brasil.
4. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.

BY-NC © 2023 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3756

The main limitation of the FI is that it applies only to RCTs with dichotomous outcomes. Another limitation is an FI equal to zero. While a trial uses chi-square analysis to calculate a p-value, FI is calculated using the Fisher's exact test, and therefore, an FI = 0 could occur in such trials when statistical significance was lost by simply changing the analysis from the chi-square test to the Fisher's exact test.

KEY MESSAGES

- The FI estimates the number of events needed to turn a statistically significant trial into non-significant. The smaller the FI is, the less robust the trial is.
- FI and FQ offer an alternative to the frequentist approach to RCT analysis and have been increasingly used in the critical appraisal of RCTs as an adjunctive tool for RCT interpretation.

REFERENCES

1. Meyer G, Vicaut E, Danays T, Agnelli G, Becattini C, Beyer-Westendorf J, et al. Fibrinolysis for patients with intermediate-risk pulmonary embolism. N Engl J Med. 2014;370(15):1402-1411. https://doi.org/10.1056/NEJMoa1302097
2. Tignanelli CJ, Napolitano LM. The Fragility Index in Randomized Clinical Trials as a Means of Optimizing Patient Care. JAMA Surg. 2019;154(1):74-79. https://doi.org/10.1001/jamasurg.2018.4318

Losing your fear of using R for statistical analysis

Iara Shimizu[1,2], Juliana Carvalho Ferreira[2,3]

WHAT IS R?

R is a programming language widely used in health research, as it provides a vast collection of software packages that encompass a wide range of data analysis techniques to conduct from simple to complex statistical analyses, to create graphics and figures, and to design websites and apps.

Because it is a programming language, it requires input through a command line, which may seem intimidating to nonprogrammers. However, even if you have never programmed before, there are several tips to get started and gradually learn how to use R, including preset software packages that extend the capabilities of basic R and allow users to perform specialized tasks without having to write all of the code from scratch.

ADVANTAGES OF R

Being an open-source program, R facilitates changes in analyses and ensures reproducibility of results in new datasets, allowing you to document and share your analyses in a systematic and organized way, in addition to the possibility of creating reproducible reports that combine code, text, and visualizations in a single document.

For researchers who do not want to use commercial statistical software, R is an option because it is free and adaptable to different operating systems.

One of the greatest advantages of R is its online support, since it has a very extensive and active community of users around the world, who continually develop new functionalities for the program and offer solutions to the various questions that may arise.

GETTING STARTED WITH R AND RSTUDIO

To get started with R programming for statistical analysis in health research, you can follow these steps, summarized in Figure 1.

1. Installing R and RStudio—Install R and RStudio on your computer. R is a programming language used for statistical computing and graphics, and RStudio is a program designed for working with R programming language, providing a user-friendly interface. Visit the official R website[1] and download the appropriate version for your operating system. There is also the possibility of using an online version that does not require installation of any software.[2]
2. Importing data into R—This involves extracting data from a file or database and importing them into an R data frame. You can import data into R from various sources, such as CSV files, Excel files, or databases, and manage the data by filtering, sorting, merging, or transforming datasets. If you do not have data of your own, R has a list of open data sets that can be used to gain hands-on experience and improve programming skills.
3. Learning R coding basics—Get familiar with the basic language of R. There are several free online tutorials, books, and resources available for learning R, such as the "Hands-On Programming with R"[3] for programming beginners. To use the R program, you enter the instructions, known as commands, which direct R to perform a specific task, such as calculate the mean of a variable or perform a t-test to compare two groups. If you type a command that R does not recognize, it will return an error message. If that happens, do not panic! Read the error message to understand the problem, review the command that you have typed for any mistakes or syntax errors, or search for solutions online using help pages in R or in communities and forums.
4. Statistical analysis—R has a system with a varied number of packages designed specifically for statistical analysis, such as basic R functions, stats package, survival analysis package, and more specialized packages; packages not included in

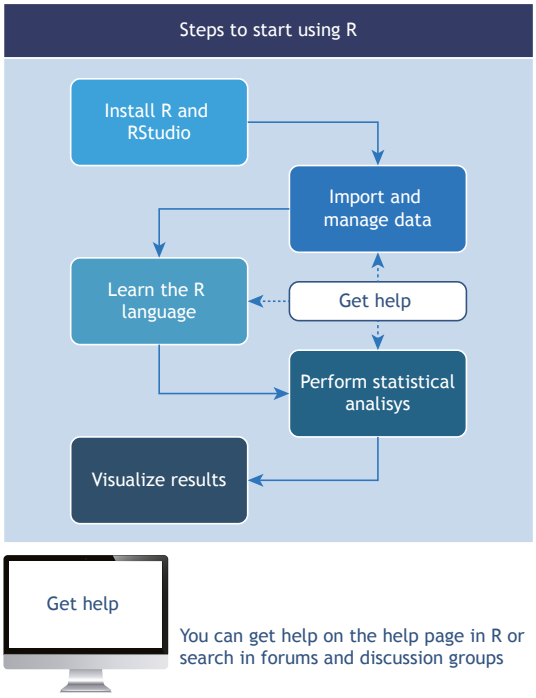

Figure 1. Steps to start using R.

1. Coordenação de Fisioterapia, Centro de Ciências da Saúde, Universidade Estadual do Piauí – UESPI – Teresina (PI) Brasil.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.

BY-NC © 2023 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3756

the basic software need to be installed. However, basic statistical analysis can be performed with a few simple and easy-to-learn commands.

5. Visualization of results—R provides a diverse range of packages that enable the generation of high-quality figures and charts. Researchers can create figures and charts that facilitate a deeper understanding of the data and aid in the communication of key results.
6. Getting help—Each R function comes with its own help page that you can access by typing the name of the function preceded by a question mark. R community forums and discussion groups allow you to submit a question or search through previously answered questions. Participating in the community will expose you to different perspectives, new techniques, and useful resources.

In conclusion, learning R programming for statistical analysis is like learning a new language: it may seem somewhat difficult in the beginning, but as you learn, it becomes easier. We recommend that, as a new user, you start with small projects to gradually build your skills and explore advanced techniques as you go.

ACKNOWLEDGEMENTS

Many thanks to Dr. Eduardo Leite Costa, who inspired and taught Dr. Ferreira to use R and continues to provide help when she most needs it.

REFERENCES

1. The R Foundation [homepage on the Internet]. Vienna, Austria: R Foundation for Statistical Computing; [cited 2023 Jun 17]. The R Project for Statistical Computing. Available from: https://www.r-project.org
2. posit Cloud [homepage on the Internet].Boston, MA: posit; c2023 [cited 2023 Jun 17]. Friction free data science. Available from: https://posit.cloud
3. Grolemund G. Hands-On Programming with R [monograph on the Internet]. Sebastopol, CA: O'Reilly Media; 2014. [cited 2023 Jun 17]. Available from: https://rstudio-education.github.io/hopr/

How to write a curriculum vitae – advice for young researchers

Juliana Carvalho Ferreira[1,2]

PRACTICAL SCENARIO

A pulmonary fellow in Brazil is planning to apply for a position as a student in a prestigious doctoral program in another state. She consulted the program's website and learned that, in order to apply, in addition to other documentation, she needed a complete, updated version of her curriculum vitae (CV) at *Plataforma Lattes*. She realizes that her CV is one of the most important documents in the application process, since its content may be compared with that of other candidates to the doctoral position. Therefore, she seeks help to make sure that her CV at *Plataforma Lattes* is not only up to date, but that it helps her stand out as a good candidate for the graduate program as well.

WHAT IS PLATAFORMA LATTES

Plataforma Lattes is a nationwide electronic platform for academic CV adopted by most funding agencies, academic institutions, and research institutes in Brazil. It was launched in 1999 by the Brazilian *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq, National Council for Scientific and Technological Development) and is integrated with databases from other institutions, such as SciELO, Scopus, and CrossRef, as well as university databases, giving the platform user the possibility of accessing information from the researcher's CV, including his/her connections with academic institutions and prior funding.[1]

Given that the *Plataforma Lattes* system is almost universally adopted by academic institutions in Brazil and provides standardization of the process to include items in academic CV, anyone accessing the platform can consult a researcher's CV and use it to evaluate and compare researchers, students, and institutions.

WHY IS THE CV SO IMPORTANT IN ACADEMIC MEDICINE?

A CV is a professional portfolio of a researcher's academic career. In many situations, it is the first impression that a future academic supervisor, employer, or funding agency will have of the researcher. It tells the researcher's professional story, and even at first glance, it should present information in a way that highlights strengths and expertise that show that the candidate is the perfect match for the role he/she is applying for.

One of the most important parts of a CV is the personal statement—described as the "Summary" in the platform—containing from one to three paragraphs, in which the researcher summarizes his/her academic history, current job position, accomplishments, and interests. In some instances, the personal statement is longer and used to apply to specific grants.[2] *Plataforma Lattes* shows the personal statement as the first and, therefore, most visible item in the CV. The platform generates a standardized text automatically, extracting information from other standardized items, such as education/training and current job positions. However, researchers can (and should) edit this section periodically, making sure that the text is well written and tells their professional story in a way that conveys what makes them fit for the position they are applying for.

Table 1. Elements to include in a curriculum vitae: the model at *Plataforma Lattes*.

Item	What to include
Brief personal details	Name, professional address
Personal statement[a]	One to three paragraphs summarizing career story, current job position, accomplishments and interests
Education (most recent first)	Undergraduate degree, specialization, residency, and fellowship for clinicians
Job positions (most recent first)	It may be a fellow position for young researchers
Research projects and funding	Current and completed projects, including funding details
Awards and prizes (most recent first)	It may include undergraduate prizes
Publications	Peer-reviewed manuscripts, book chapters, and conference abstracts
Teaching experience	It may include informal teaching such as teaching interns
Conference attendance	National and international conferences

[a]For some specific awards and grants, the personal statement may be longer and more focused on the position being sought.[2]

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2022 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário

WHAT ARE THE ELEMENTS TO BE INCLUDED IN THE CV?

Even young researchers, with little or no prior research experience, can and should have an impressive CV that is compatible with the position they are applying for. An experienced researcher, applying for a large grant, will highlight publications, teaching appointments, and prior successes with academic funding. Young researchers such as the fellow in the practical scenario should highlight their education, including the institutions where they were trained; their motivation to learn, such as attendance to academic and medical conferences; community service, such as volunteering in academic projects; and technical skills, such as the ability to communicate in foreign languages. Table 1 shows the basic structure of an academic CV at *Plataforma Lattes*. There are other models, of course, but the general structure is more or less the same.

TIPS TO WRITE A GREAT CV

1. Make sure that the most important information comes first (usually in the personal statement) and is well written, brief, and easy to read.
2. Make sure that you include all the important items in the order required by the institution.
3. If you are not using a standardized platform such as *Plataforma Lattes*, format your CV in a way that it is consistent in its use of fonts, line breaks, bullet points, and other details.
4. Proofread your CV to correct spelling mistakes and grammatical errors.
5. Be honest and consistent. Lying on your CV gets you a bad reputation.

REFERENCES

1. Jericho BG, Ilgen JS, Gottlieb-Smith R, Simpson D, Sullivan GM. How to Write Your Curriculum Vitae. J Grad Med Educ. 2019;11(3):333-334. https://doi.org/10.4300/JGME-D-19-00221.1
2. Tay A. Sell yourself and your science in a compelling personal statement. Nature. 2021;593(7857):153-155. https://doi.org/10.1038/d41586-021-01101-z

J Bras Pneumol. 2018;44(4):260-260
http://dx.doi.org/10.1590/S1806-37562018000000217

Twelve tips to write an abstract for a conference: advice for young and experienced investigators

Juliana Carvalho Ferreira[1,2,a], Cecilia Maria Patino[1,3,b]

As we have returned from the successful XXXIX Brazilian Thoracic Society Conference in Goiânia, Brazil—where more than 600 abstracts have been presented—and prepare for the American Thoracic Society International Conference deadline for submitting abstracts by November this year, we would like to emphasize the importance of presenting high-quality scientific abstracts at such conferences.

Presenting clinical research results in the form of abstracts in national and international meetings is common and expected among clinical researchers in academic and nonacademic settings, giving researchers the opportunity to present their work in person, network with researchers working in the same field, receive feedback from peers, and publish their results as abstracts in conference proceedings.

Writing abstracts that are clear and informative, following both the conference and internationally endorsed reporting guidelines, is very important for various reasons: abstracts are used by conference program committees to select the best suited ones for oral presentations; abstracts are usually available online prior to the conference and attendees can select which presentations they will attend; abstracts are usually published and, therefore, may be cited by other authors on their peer-reviewed publications; and finally, health care professionals may base medical decisions on results of studies that have been published only as a conference abstract. Therefore, in order to guide investigators how to write high quality conference abstracts, we have developed 12 tips for young and experienced investigators:

1. **Identify and carefully follow specific guidelines** suggested by the conference. Usually an abstract contains the following: title, background/introduction, objectives, methods, results, and conclusion; however, this format varies across conferences. Pay close attention to information such as word limit and how the abstract should be structured.
2. **Follow internationally endorsed reporting guidelines** specifically developed for conference abstracts. The **E**nhancing the **QUA**lity and **T**ransparency **O**f health **R**esearch (EQUATOR) Network is an international initiative that seeks to improve the quality of published health research globally by developing reporting guidelines for several types of study designs.[1] Many reporting guidelines have extensions focusing specifically on abstracts.[2] Read them before starting to write your abstract.
3. **Think carefully about the title** because this is what readers look at first. Compose a clear, objective title and, whenever possible, include the study design. You can make it attractive, but avoid trying to be too clever (especially for beginners).
4. **Do not waste words on the introduction**. Be brief and straight to the point. Save space here, so you can provide more details in the methods and results sections, which are novel and particular to your study.
5. **Clearly state the objectives of the study**. The objective derives from your research question and should clearly align with results and conclusion.
6. **Make sure that the methods section is detailed enough**—but not too technical—and include the study design, setting, study participants, and eligibility criteria. You should also include a description of the important variables of the study, such as the exposure, intervention, predictors, and outcome, as well as the analytic approach used to answer the research question.
7. **Be precise and specific when writing the results**. Report the number of participants that were included the analysis, and, most importantly, always report the results that actually answer your research question (e.g., the difference between groups with a measure of precision such as an SD or 95% CI) and never just a p value.
8. **Be realistic in the conclusion**. Mention the impact of your study, but avoid speculating beyond what your results show; you can also mention future directions in the area of study, but avoid the overused "more studies are needed..."
9. **Perform a careful spell and language check**, especially if you are not writing in your native language.
10. **Avoid or minimize abbreviations**. Readers can feel frustrated when they have to go back to remember what an abbreviation stands for (e.g., EQUATOR in this paper).
11. **Get feedback** from your coauthors, mentor, and colleagues outside your team. The goal is to use their help to identify unclear sentences and missing or inaccurate information, as well as to make sure that the writing is high quality. They can also help you to make sure that the title, objectives, methods, results, and conclusion are all aligned with the research question.
12. **Do NOT wait until the last minute** to write and proofread the content. Writing and reviewing the abstract for quality always takes more time than you initially thought it would. Moreover, glitches in the submission process are always possible, so you want to give yourself time to contact the conference staff for help, if necessary.

REFERENCES

1. EQUATOR network [homepage on the internet]. Oxford: University of Oxford [cited 2018 Jun 1]. Available from: http://www.equator-network.org/
2. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, Schulz KF, the CONSORT Group. CONSORT for reporting randomised trials in journal and conference abstracts. Lancet. 2008;371(9609):281-3. https://doi.org/10.1016/S0140-6736(07)61835-2

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
a. http://orcid.org/0000-0001-5742-2157; b. http://orcid.org/0000-0001-6548-1384

BY-NC © 2018 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3713

How to prepare and present a poster at a conference and communicate your research findings effectively

Juliana Carvalho Ferreira[1,2,a], Cecilia Maria Patino[1,3,b]

The overall goal of presenting a poster at a conference is to communicate to and receive feedback from your peers who have similar research interests regarding your research findings, as well as to increase your scientific network. To make the most of this opportunity, in an environment where conference attendees are overwhelmed by new information, the presenters of a poster need to communicate their results effectively.

A POSTER IS NOT A MINI MANUSCRIPT

The most common mistake researchers make when designing and presenting a poster is to treat the poster as if it were a mini manuscript. Remember, the goal is to clearly and quickly communicate research content to attendees who are walking around a large area with many posters, with limited time and attention span to read extensive information.

Successfully communicating your study will rely on attracting attendees to your poster and making it easy for them to grasp key messages. Consider the 10-10 rule of thumb for poster viewing: attendees look at posters for 10 seconds and from 10 feet (3 meters) away.[1,2] During those 10 seconds, if they are attracted to the poster, they might want to read more information. Therefore, your key messages must be written using a large enough font size for people to read it from 3 meters away, and the information should be interesting and attractive enough so that attendees will want to come closer, read more, and ask questions.

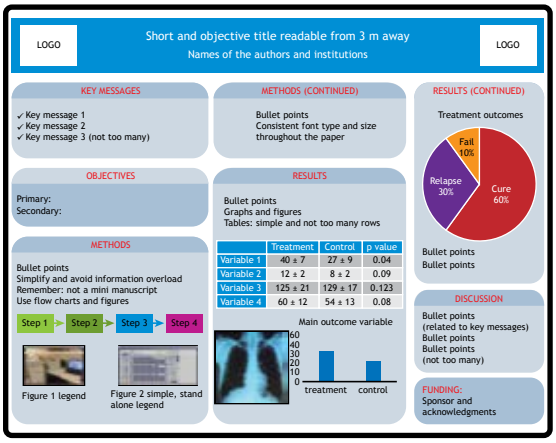

Figure 1. Poster model with examples of headers, figures, graphs, color, and font size.

HOW TO PREPARE YOUR POSTER

The most important advice is to avoid information overload. Remember: it is not a mini manuscript and attendees have limited time. We recommend that you use bullet points, minimize text, and simplify the language, with easy-to-read phrases. Use the active voice and avoid jargon and acronyms whenever possible. The title should be short and informative. The left upper corner is the first area attendees will look after reading the title and scan it during those 10 seconds, so avoid writing an introduction (remember, it is not a manuscript); instead, start with your key messages (Figure 1). Then state your objectives or research question, so that readers will know what your study is about. When describing your methods, avoid unnecessary details, use bullet points, simplify the text, and use flow charts or figures to illustrate the method process. Results are the most important information you will communicate. Use figures and graphs with legible font, clear axis, and, if possible, with the legend embedded in the graph/figure. Finish your poster with a discussion that aligns with your research question.

Check the conference guidelines for poster orientation and size. Use columns and headers to facilitate reading. Resist the temptation to fill all available space, leave some blank space to make the poster more attractive. It is important to use consistent wording, font, font size, and colors.

KEY MESSAGES

- A poster is not a mini manuscript; avoid communicating too much information
- Be mindful of using a small font size that is hard to read from a comfortable distance
- Substitute text for figures and graphs whenever possible
- Practice presenting your poster to your friends and colleagues at least five times
- Be prepared, look and act professionally, and make it worth the effort

REFERENCES

1. Hamilton CW. At a glance: a stepwise approach to successful poster presentations. Chest. 2008;134(2):457-459. https://doi.org/10.1378/chest.08-1078
2. Persky AM. Scientific Posters: A Plea from a Conference Attendee. Am J Pharm Educ. 2016;80(10):162. https://doi.org/10.5688/ajpe80455

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
a. http://orcid.org/0000-0001-6548-1384; b. http://orcid.org/0000-0001-5742-2157

BY-NC © 2019 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3713

Building clinical and translational research teams

Cecilia María Patino[1,2], Juliana Carvalho Ferreira[1,3]

PRACTICAL SCENARIO

A group of pulmonary and critical care investigators in Latin America are interested in developing a master's research program as a means of reducing pulmonary and critical care-related morbidity and mortality in urban settings. They conducted a scoping review of existing training models that they could use and came to agreement on evaluating the competency-based model in clinical and translational research[1] The overall goal of the Latin American program was to gather a cadre of investigators that conduct research proposed by the National Institutes of Health (NIH) in order to accelerate the development or adoption, as well as the dissemination, uptake, and implementation of new medical and health-related interventions to improve respiratory health in Latin America.

WHAT IS CLINICAL AND TRANSLATIONAL RESEARCH?

Translational science is the process of turning observations in the laboratory, clinic and community into interventions that improve the health of individuals and the public.[1] These interventions can include diagnostics, therapeutics, medical procedures, behavioral changes, access to health care, and health-related laws, in addition to how interventions are effectively disseminated, implemented, and evaluated in the community. This field focuses on understanding the scientific and operational principles underlying each step of the translational process, from developing new treatments to demonstrating their usefulness, as well as to disseminating and implementing the findings. The National Institutes of Health has divided this process into a spectrum of five different types of research areas: preclinical, clinical, dissemination, implementation, and public health,[1] but the spectrum is not necessarily linear, with each stage building upon and informing the other (Figure 1). Clinical and translational research prioritizes unmet needs that include preventing disease, overcoming disease, and decreasing the burden of disease in local communities. Translational teams produce crosscutting solutions for common and persistent challenges and emphasize creativity and innovation. They also leverage cross-disciplinary science teams, enhance efficiency and speed of research, use boundary-crossing partnerships, and use rigorous and reproducible research approaches.[2]

THE COMPETENCIES OF CLINICAL AND TRANSLATIONAL RESEARCH

Translational research teams should include professionals with diverse skills, and their core competencies go beyond each individual's specialization. The NIH training competency framework proposes both well-known as well as innovative competencies that the research students would need to develop and master to practice clinical and translational research successfully (Table 1).[1] In our practical scenario, once the group reaches a consensus on the competencies, the next step is to identify faculty with expertise and experience in clinical and translational research, as well as providing the necessary didactic and experiential training opportunities. It is important for the group that the students both "learn" and "conduct" clinical and translational research while in training. The group will evaluate the program during the following two years to report successes and adaptations of this program.

Table 1. Key competencies of a clinical and translational researcher.

Researcher	Competency
Rigorous researcher	Shows strong and state-of-the-art methodological and statistical skills that are rigorous and reproducible
Team player	Leverages and respects research expertise across team members
Boundary crosser	Broadly collaborates across disciplines to advance interventions
Process innovator	Innovates to overcome barriers to advancing intervention development and implementation
Domain expert	Has deep understanding and knowledge within one or more disciplines
Skilled communicator	Communicates well across a broad spectrum of audiences
Systems thinker	Evaluates external forces, interactions, and relationships across all stakeholders involved in developing and implementing successful interventions, including patients, family dynamics, medical professionals, and health care systems

Adapted from the National Institutes of Health.[1]

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2022 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3756

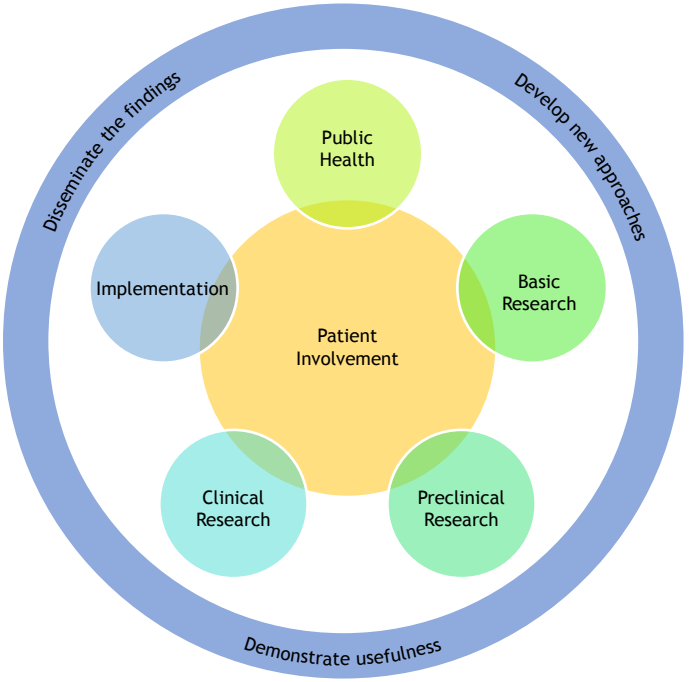

Figure 1. The translational science spectrum. Adapted from the National Institutes of Health.[3]

REFERENCES

1. U.S. Department of Health & Human Sciences. National Institutes of Health. National Center for Advancing Translational Sciences (NCATS) [homepage on the Internet]. Bethesda: NCATS; c2022 [updated 2022 Jun 6; cited 2022 Jun 10] Translational Science Principles [about 5 screens]. Available from: https://ncats.nih.gov/training-education/translational-science-principles
2. Gilliland CT, White J, Gee B, Kreeftmeijer-Vegter R, Bietrix F, Ussi AE, et al. The Fundamental Characteristics of a Translational Scientist. ACS Pharmacol Transl Sci. 2019;2(3):213-216. https://doi.org/10.1021/acsptsci.9b00022
3. U.S. Department of Health & Human Sciences. National Institutes of Health. National Center for Advancing Translational Sciences (NCATS) [homepage on the Internet]. Bethesda: NCATS; c2022 [updated 2022 Jun 6; cited 2022 Jun 10] Translational Science Spectrum [about 3 screens]. Available from: https://ncats.nih.gov/translation/spectrum

Implementation research and its role in public health and health policies

Victoria Sáenz[1,2], Cecilia María Patino[1,3], Juliana Carvalho Ferreira[1,4]

PRACTICAL SCENARIO

A ministry of health, worried about the increase in the prevalence of current smoking among their adult population, surveyed a sample of this population to identify barriers to smoking cessation. The results showed that current smoking cessation interventions, based on visiting a pulmonologist, were not feasible because of important barriers, including difficulties in scheduling an appointment and adhering to follow-up visits, leading to loss of motivation. These results were used to inform the development of a population-based smoking cessation intervention through a mobile app to evaluate its feasibility and effectiveness.

IMPLEMENTATION RESEARCH

Implementation research (IR) is a specific scientific approach that evaluates the effectiveness of incorporating evidence-based interventions and policies into the routine health care system. IR focuses on facilitators of and barriers to implementing evidence-based interventions in public and private health care systems and promotes the application, use, and sustainability of these interventions on a large scale (Figure 1).

IR evaluates different types of interventions, including newly developed medical devices and technologies, application of treatment protocols, service delivery programs, behavioral interventions, among others. Social science research methods, as well as methods for determining the cost of implementation strategies at different levels of the health care system[1] are used in IR.

HOW CAN IR IMPROVE PUBLIC HEALTH?

IR addresses health policy makers' priorities and the needs of real-world health decision makers. Although successful efforts have been made to close the research gap in changing health policies, health decision-making processes are highly complex and involve a great number of stakeholders. Conducting policy-driven research, such as IR, supports the use of research findings to inform health policy planning and its implementation by policy makers.[2]. Thus, the main role of IR is to improve the effectiveness of health care systems and health care delivery.

CONSIDERATIONS WHEN CONDUCTING IR

***Population*:** IR is ideally conducted within the population that will be affected by the health-related intervention. The selection of inclusion criteria should be broad and result in a study population that is truly representative of the target population, whereas exclusion criteria should be minimal. In our example, the population consisted of adult smokers from all regions of the country who have access to a smartphone.

***Intervention/Exposure*:** interventions that fall under IR are broad. They may be complex, and the research team should try to involve diverse stakeholders. In our example, the intervention was the use of an app to promote behavioral interventions for smoking cessation. Stakeholders included the ministry of health, the general population, health professionals working in smoke cessation programs, etc.

***Comparator*:** the analytical approach of IR differs from the approach used in clinical research. Usually, the intervention has already been shown to be effective within the controlled environment of a clinical trial. In IR, the objective is to test the application of an intervention in real-world settings and if it continues to be effective over time. Therefore, a comparison group may not be necessary, or historical controls may be used.

***Outcome*:** the outcomes are usually focused on feasibility, acceptance, adherence, and effectiveness in real-world scenarios where the intervention will be

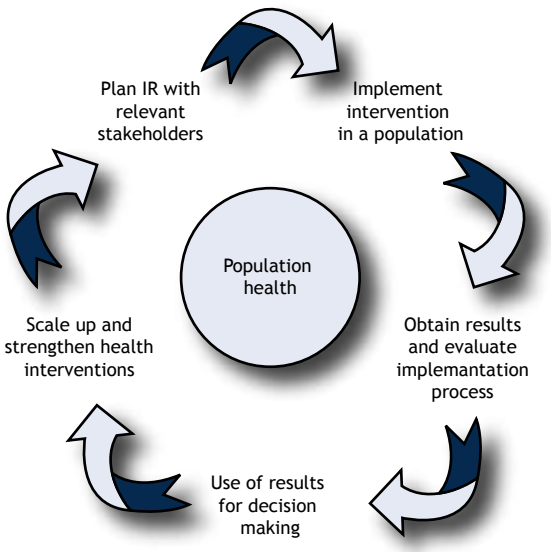

Figure 1. The process of implementation research.

1. Methods in Epidemiologic, Clinical, and Operations Research (MECOR) program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Roche Argentina, Buenos Aires, Argentina.
3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.
4. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.

BY-NC © 2021 Sociedade Brasileira de Pneumologia e Tisiologia

ISSN 1806-3756

implemented. In our example, outcomes include perceived usefulness and number of interactions with the app and, most importantly, smoking cessation rates among users. IR can evaluate various outcomes simultaneously, and the results should potentially be used for decision-making processes.

REFERENCES

1. Remme JH, Adam T, Becerra-Posada F, D'Arcangues C, Devlin M, Gardner C, et al. Defining research to improve health systems. PLoS Med. 2010;7(11):e1001000. https://doi.org/10.1371/journal.pmed.1001000
2. Langlois EV, Mancuso A, Elias V, Reveiz L. Embedding implementation research to enhance health policy and systems: a multi-country analysis from ten settings in Latin America and the Caribbean. Health Res Policy Syst. 2019;17(1):85. https://doi.org/10.1186/s12961-019-0484-4

Writing an effective response to reviewers: the goal is to improve the study and get it published!

Cecilia Maria Patino[1,2,a], Juliana Carvalho Ferreira[1,3,b]

PRACTICAL SCENARIO

We are very excited because a group of students from the Latin-American Methods in Epidemiologic, Clinical, and Operations Research (MECOR) program submitted a manuscript to an international journal and received a response stating that the manuscript is of interest but, based on the reviewers' comments, "major revisions are required". The students have contacted us because this is their first manuscript and they want to make sure that they respond to the reviewers effectively.

HOW TO STRATEGICALLY RESPOND TO REVIEWERS

As directors of the Latin-American MECOR program, we instill in our students to approach clinical research using sound methods from idea to publication. We provide many resources and guidelines related to research methodology, writing protocols, manuals of procedures, and original manuscripts. Here we provide a summary of recommendations, based on published literature[1,2] and our own experience, for responding to reviewer comments when a journal invites authors to resubmit.

First, we recommend that authors carefully read all of the comments made by the reviewers and distinguish between those that are positive and those in which the reviewers are criticizing or requesting revisions. It is very important to determine whether the comments can be adequately addressed and meet the expectations of the reviewers. Second, we strongly emphasize the need to quickly overcome feelings of frustration, sadness, and even a sense of unfairness. Remember that the article has not been rejected and the editor is giving you the opportunity to revise and resubmit; therefore, you need to get organized and respond to each comment carefully and politely, even when you do not agree with the reviewer (Table 1). Try to approach this effort with a positive attitude and use the comments made by the reviewers to your advantage. It is essential to prioritize the comments and make sure that the most important ones (those in which the reviewers request major changes) are addressed appropriately. For those who are new to this process, we highly recommend working with someone who has experience in responding to reviewers to help in this prioritization process. Third, when you resubmit your revised manuscript to the journal, the goal is to show the editor and reviewers that you have taken this process seriously by addressing every comment in detail and making all necessary changes. It is crucial that you communicate these revisions effectively through clear, simple, and straightforward language. If the authors are not native speakers of the language in which the journal is published (e.g., English), it is imperative that the final version of the response to reviewers be evaluated by an expert translator or editor.

In summary, an important goal of clinical researchers is to publish their work in peer-reviewed journals as a means of improving human health. That involves going through the peer review process and responding to the reviewer comments in an effective manner. We encourage all researchers who are starting to engage in publishing their work to develop a systematic approach when responding to reviewers. The process can be frustrating and tedious, but, in the end . . . the goal is to improve your manuscript and get it published!

Table 1. Process to successfully respond effectively to peer reviewer comments.

Task	Action 1	Action 2	Goals
Create a Word document into which you copy and paste each reviewer comment separately. Number the comments and label them according to the reviewer (e.g., Rev. 1 Comment 1).	Discuss each comment with your team and come to consensus on how to respond. You may agree or disagree with a given reviewer's comment, but you will need to politely respond to all comments.	Answer each comment separately, place them directly below the reviewer's comment (make sure there are no grammatical mistakes or misspellings), and make the corresponding changes in the revised manuscript.	This process shows the editor and the reviewers that you are dedicated and have taken the review seriously; and that you have made it easy for them to re-evaluate your manuscript.

Final products: 1. A letter to the editor in which the authors thank the editor for the opportunity to revise the paper and a list of all reviewer comments with the authors' responses; 2. A revised manuscript including all of the changes that have been made, which should be identified through the use of a different font, a different color, or *italicization*.

REFERENCES

1. Provenzale JM. Revising a manuscript: Ten principles to guide success for publication. AJR Am J Roentgenol. 2010;195(6):W382-7. https://doi.org/10.2214/AJR.10.5553
2. Happell B. Responding to reviewer's comments as part of writing for publication. Nurse Res. 2011;18(4):23-7. https://doi.org/10.7748/nr2011.07.18.4.23.c8632

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/*Asociación Latinoamericana del Tórax*, Montevideo, Uruguay.
2. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
3. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
a. http://orcid.org/0000-0001-5742-2157; b. http://orcid.org/0000-0001-6548-1384

BY-NC © 2019 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário

Building research capacity in Latin America and in Brazil: the MECOR program

Juliana Carvalho Ferreira[1,2], Irma de Godoy[3], Marcia Pizichinni[4], Ana Menezes[5], Cecilia María Patino[1,6]

WHAT IS MECOR?

The Methods in Epidemiologic, Clinical, and Operations Research (MECOR) is a training program created by the American Thoracic Society (ATS) to build clinical research capacity in low- and middle-income countries worldwide, with the purpose of improving respiratory health in these regions. The program started in Latin America in 1994 and now is offered in seven countries/regions around the globe. In Latin America, the program is a partnership among the ATS, the *Asociación Latinoamericana de Tórax* (ALAT, Latin American Thoracic Society), and four other respiratory societies: the *Sociedade Brasileira de Pneumologia e Tisiologia* (SBPT, Brazilian Thoracic Society), the Mexican Thoracic Society, the Colombian Thoracic Society, and the Argentinean Thoracic Society.

The objective of the MECOR Program is to train investigators, academicians, clinicians, and public health practitioners to design and conduct rigorous and reproducible scientific research that is relevant to the needs of the settings in which they work.[1] Ultimately, the program serves the purpose to improve global lung health through the development of local, national, and regional research capacity.[1]

MECOR IN BRAZIL

The program is held in a Latin American country during one week every year. The first MECOR course in Brazil was in São Paulo in 1997. Since then, the course has been held in Brazil seven times, including the 2021 MECOR course, which was online due to the COVID-19 pandemic. In these 27 years, 621 Brazilian students have attended MECOR, comprising approximately 35% of the students overall (Figure 1). However, when the course is held in Brazil, approximately 60% of the students are from different regions in Brazil.

IMPACT

Given that MECOR is a research capacity building program, the expected impact is that Latin American graduates thrive as researchers in respiratory medicine, critical care, and sleep medicine and produce science that has regional relevance and improves lung health in Latin America. The challenge has been measuring such impact, a task that the strategic planning committee within the program has prioritized. Although we do not have objective indicators of long-term impact yet, many graduates have become leaders in respiratory medicine, critical care, and sleep medicine, as well as professors in Latin American universities and internationally recognized researchers. We highlight that, since 2012, both program directors and course leaders have been MECOR graduates, all faculty is from Latin America, most of them are program graduates, and all teaching assistants are graduates. This shows that MECOR produces scientists/educators who can maintain the level of excellence of the program. We highlight another important product of the MECOR program in Latin America: the continuing education series published in this Journal. The co-directors of the MECOR program have been contributing to the JBP with papers on scientific methodology since 2015, producing more than 30 manuscripts that have garnered hundreds of citations.

MECOR AND SBPT PARTNERSHIP

The mission of the MECOR program—to build research capacity in respiratory medicine—is aligned with the SBPT mission, which is to promote continuous professional growth and excellence and stimulate partnerships and scientific research.[2] Over these 27 years of MECOR in Latin America, the partnership has resulted in the participation of many SBPT members in the MECOR program, many of whom have become SBPT leaders. It has also strengthened the partnership between SBPT with ALAT and ATS in a coordinated effort to contribute to the development of future leaders in respiratory medicine in Brazil and Latin America.

WHY APPLY TO THE MECOR PROGRAM?

During their time in the MECOR program, students learn how to design and implement a research project that is relevant to their setting, and how to analyze, interpret, and communicate their findings in the form of a scientific manuscript.[3] The process is completed over three years, and students learn many research skills, in addition to basic biostatistics, critical appraisal of the literature, and presentation skills. Importantly, they network with colleagues from other regions and countries and become part of a community which feels like a family. This new community is very often a career and life changing experience.

1. Methods in Epidemiologic, Clinical, and Operations Research–MECOR–program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Universidade Estadual Paulista, Faculdade de Medicina de Botucatu, Botucatu (SP) Brasil.
4. Programa de Pós-Graduação em Ciências Médicas, Universidade Federal de Santa Catarina, Florianópolis (SC) Brasil.
5. Programa de Pós-Graduação em Epidemiologia, Universidade Federal de Pelotas, Pelotas (RS) Brasil.
6. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.

BY-NC © 2021 Sociedade Brasileira de Pneumologia e Tisiologia
◄ Voltar ao sumário
ISSN 1806-3756

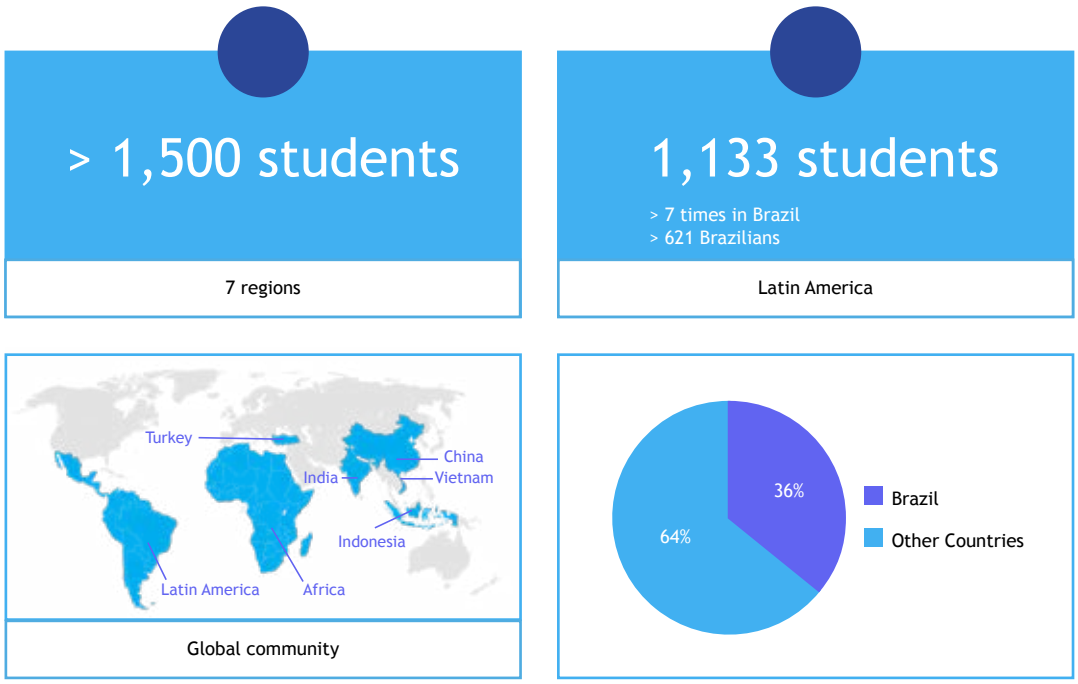

Figure 1. The MECOR program in Latin America and in Brazil.

REFERENCES

1. American Thoracic Society (ATS) [homepage on the Internet]. New York: ATS. MECOR program [cited 2021 Dec 1]. Available from: https://www.thoracic.org/about/global-public-health/mecor-program/index.php
2. Sociedade Brasileira de Tisiologia e Pneumologia (SBPT) [homepage on the Internet]. Brasília: SBPT [cited 2021 Dec 1]. Sobre a SBPT. Available from: https://sbpt.org.br/portal/quem-somos-sbpt/
3. Asociación Latinoamericana del Tórax (ALAT) [homepage on the Internet]. Montevideo: ALAT [cited 2021 Dec 1]. Métodos em Pesquisa Epidemiológica, Clínica e Operacional (MECOR). Available from: https://alatorax.org/pt/mecor

JBP
Jornal Brasileiro de **Pneumologia**
PUBLICAÇÃO OFICIAL DA SOCIEDADE BRASILEIRA DE PNEUMOLOGIA E TISIOLOGIA