Critical Thinking in Clinical Research

Applied Theory and Practice Using Case Studies

Edited by

FELIPE FREGNI & BEN M.W. ILLIGENS

Critical Thinking in Clinical Research

Critical Thinking in Clinical Research

Applied Theory and Practice Using Case Studies

Edited by

Felipe Fregni, MD, PhD, MMSc, MPH

Associate Professor of Epidemiology, Harvard T.H. Chan School of Public Health
Associate Professor of Physical Medicine and Rehabilitation
Harvard Medical School
Director, Collaborative Learning in Clinical Research Program
Principles and Practice of Clinical Research (PPCR) Program
Harvard T.H. Chan School of Public Health
Boston, MA

Ben M. W. Illigens, MD

Department of Neurology
Beth Israel Deaconess Medical Center
Harvard School of Medicine
Boston, MA

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and certain other countries.

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America.

Library of Congress Cataloging-in-Publication Data
Data Names: Fregni, Felipe, editor. | Illigens, Ben M. W., editor.
Title: Critical thinking in clinical research : applied theory and practice using case studies / edited by Felipe Fregni, Ben M.W. Illigens.
Description: New York, NY : Oxford University Press, [2018] |
Includes bibliographical references and index.
Identifiers: LCCN 2017028658 | ISBN 9780199324491 (alk. paper)
Subjects: | MESH: Clinical Studies as Topic | Statistics as Topic | Biomedical Research | Research Design | Ethics, Research
Classification: LCC R853.C55 | NLM W 20.55.C5 | DDC 610.72/4—dc23
LC record available at https://lccn.loc.gov/2017028658

1 3 5 7 9 8 6 4 2
Printed by WebCom, Inc., Canada

To Lucca Fregni, the light of our lives.

Felipe Fregni

To my wife Seung-Kyung Cecilia Lee, who will always
be in my heart, and our daughter Clara Eun-Bee Illigens.
You are my love and inspiration
and made me the person I am.

Ben M. W. Illigens

CONTENTS

CONTRIBUTORS

Priscila Caldeira Andrade, PharmD, MPH
Health Economics & Market Access Manager-Northern Region (Latin America)
Johnson & Johnson Medical
Sao Paulo, Brazil

Valeria Angelim, PhD
Alphacor Cardiologia Clinica e Diagnostica
Universidade de Sao Paulo (USP)
Sao Paulo, Brazil

Nazem Atassi, MD
Department of Neurology
Harvard Medical School
Massachusetts General Hospital
Boston, MA

Munir Boodhwani, BSc, MD
University of Ottawa Heart Institute
Ottowa, Canada

André Brunoni, MD, PhD
Universidade de Sao Paulo (USP)
Medical School
Sao Paulo, Brazil

Sandra Carvalho, MSc, PhD
Neuromodulation Center
Spaulding Rehabilitation Hospital
Harvard Medical School
Boston, MA

Laura Castillo-Saavedra, MD
University of Chicago Medicine
The Pritzker School of Medicine
Chicago, IL

Lívia Caroline Mariano Compte, MD
Universidade de Sao Paulo (USP)
Medical School
Sao Paulo, Brazil

Minou Djannatian, MD, MSc
Institute of Neuronal Cell Biology
Technische Universität München
Munich, Germany

Scott Evans, PhD
Director of the Statistical and Data Management Center (SDMC)
Antibacterial Resistance Leadership Group (ARLG)
Department of Biostatistics
Harvard T.H. Chan School of Public Health
Boston, MA

John Ferguson, MD
Sanofi Pharmaceutical
Boston, MA

Juliana C. Ferreira, MD, PhD
Respiratory Intensive Care Unit
Hospital das Clinicas
Universidade de Sao Paulo (USP)
Medical School
Sao Paulo, Brazil

Steven D. Freedman, MD, PhD
Chief, Division of Translational Research
Department of Medicine
Beth Israel Deaconess Medical Center
Boston, MA

Felipe Fregni, MD, PhD, MMSc, MPH
Associate Professor of Epidemiology, Harvard T.H. Chan School of Public Health
Associate Professor of Physical Medicine and Rehabilitation, Harvard Medical School
Director, Collaborative Learning in Clinical Research Program, Principles and Practice of Clinical Research (PPCR) Program
Harvard T.H. Chan School of Public Health
Boston, MA

Alexandra Gomez-Arteaga, MD
Universidad de Los Andes
Bogota, Colombia

Inês C. R. Henriques, MD
Faculty of Medicine of the University of Coimbra
Coimbra, Portugal

Ben M. W. Illigens, MD
Department of Neurology
Beth Israel Deaconess Medical Center
Harvard School of Medicine
Boston, MA

Rui Imamura, MD
Universidade de Sao Paulo (USP) Medical School
Sao Paulo, Brazil

Tamara Jorquiera, MD, MSc
Department of Medicine
University of San Martín de Porres
Lima, Peru

Faiza Khawaja, MD
Department of Clinical Pharmacology and Internal Medicine
The Michener Institute for Applied Health Sciences
Toronto, Canada

Catherine Lee, PhD
Department of Biostatistics
Harvard T.H. Chan School of Public Health
Boston, MA

Hang Lee, PhD
Department of Medicine
Massachusetts General Hospital
Harvard Medical School
Boston, MA

Jorge Leite, PhD
Laboratory of Neuromodulation
Spaulding Rehabilitation Hospital
Harvard Medical School
Boston, MA

Chin Lin, MD, PhD
University of Sao Paulo
Hospital das Clinicas (ICESP)
Sao Paulo, Brazil

Fernanda Lopes, PhD
Neuromodulation Center
Spaulding Rehabilitation Hospital
Harvard Medical School
Boston, MA

Ana R. Martins, RN
PPCR Program, ECPE
Harvard T.H. Chan School of Public Health
Boston, MA
Universidade de Coimbra
Coimbra, Portugal

Joseph Massaro, PhD
Professor of Biostatistics
Boston University School of Public Health
Boston, MA

Suely Reiko Matsubayashi, MD
PPCR Program, ECPE
Harvard T.H. Chan School of Public Health
Boston, MA

Lotfi B. Merabet, OD, PhD, MPH
Associate Professor
Department of Ophthalmology
Harvard Medical School
Boston, MA

Raquel Ajub Moyses, MD
Hospital das Clínicas
Universidade de Sao Paulo (USP)
Medical School
Sao Paulo, Brazil

Rui Nakamura, MD
Principles and Practice of Clinical
Research (PPCR) Program, ECPE
Harvard T.H. Chan School of
Public Health
Boston, MA

Vera Novak, PhD, MD
Department of Neurology
Beth Israel Deaconess Medical Center
Harvard School of Medicine
Boston, MA

Jessica Paulus, ScD
Investigator, Institute for Clinical
Research and Health Studies
Assistant Professor of Medicine
Tufts University School of Medicine
Medford, MA

Ana Isabel Penzlin, MD
University Hospital Carl Gustav Carus
Dresden, Germany

Mohammad Azfar Qureshi, MBBS
Department of Nephrology
King Abdulaziz Medical City
Saudi Arabia

Gustavo Rivara, MD, MSc
Pediatrician and Neonatologist
Universidad San Martín de Porres
Facultad de Medicina Humana Clinica
Delgado Lima, Perú

Timo Siepmann, MD
Departments of Neurology, Internal
Medicine, Clinical Pharmacology
Universitätsklinikum Dresden
Dresden, Germany

Claudia Kimie Suemoto, MD, PhD
Division of Geriatrics
University of Sao Paulo Medical School
Sao Paulo, Brazil

Alma Tamara Sanchez Jimenez, MD
PPCR Program, ECPE
Harvard T.H. Chan School of
Public Health
Boston, MA

Rita Tomás, MD
Adjunct Professor
Escola Superior de Tecnologia da Saúde
de Lisboa
Instituto Politécnico de Lisboa
Clínica CUF Alvalade
Portugal Football School,
Portuguese Football Federation
Lisbon School of Health Technology
Polytechnic Institute of Lisbon
Lisbon, Portugal

Keiko Ueda, MD
Principles and Practice of Clinical
Research (PPCR) Program
Harvard T.H. Chan School of
Public Health
Boston, MA

Clarissa Valim, MD, MSc, ScD
Department of Immunology and
Infectious Diseases
Harvard T.H. Chan School of
Public Health
Boston, MA

UNIT I

Basics of Clinical Research

1

BASICS OF CLINICAL RESEARCH
INTRODUCTION TO CLINICAL RESEARCH

Authors: *Rui Nakamura, Faiza Khawaja, Laura Castillo-Saavedra,*[*] *Felipe Fregni, and Steven D. Freedman*

Case study authors: *Felipe Fregni, Fabio Pinna, and André Brunoni*

The whole history of science has been the gradual realization that events do not happen in an arbitrary manner, but that they reflect a certain underlying order, which may or may not be divinely inspired.
—Stephen W. Hawking

INTRODUCTION

The search for knowledge about ourselves and the world around us is a fundamental human endeavor. Research is a natural extension of this desire to understand and to improve the world in which we live.

This chapter focuses on the process of clinical trials, ethical issues involved in the history of clinical research, and other issues that may be unique to clinical trials. As clinical trials are, perhaps, the most regulated type of research—subject to provincial, national, and international regulatory bodies—reference will be made to these regulations where appropriate.

The scope of research is vast. On the purely physical side, it ranges from seeking to understand the origins of the universe down to the fundamental nature of matter. At the analytic level, it covers mathematics, logic, and metaphysics. Research involving humans ranges widely, including attempts to understand the broad sweep of history, the workings of the human body and the body politic, the nature of human interactions, and the impact of nature on humans—the list is as boundless as the human imagination.

CLINICAL RESEARCH

Clinical research is a branch of medical science that determines the safety and effectiveness of medications, devices, diagnostic products, nutrition or behavioral changes,

[*] The first three authors contributed equally to the work.

and treatment regimens intended for human use. Clinical research is a structured process of investigating facts and theories and exploring connections. It proceeds in a systematic way to examine clinical conditions and outcomes, to establish relationships among clinical phenomena, to generate evidence for decision-making, and to provide the impetus for improving methods of practice.

Clinical trials are a set of procedures in medical research and drug development that are conducted to allow safety and efficacy data to be collected for health interventions. Given that clinical trials are experiments conducted on humans, there is a series of required procedures and steps for conducting a clinical trial. There are several goals tested in clinical trials, including testing whether the drug, therapy, or procedure is safe and effective for people to use. The overall purpose of a clinical trial is acquisition of new knowledge, not the treatment of patients per se.

A clinical trial, also known as patient-oriented research, is any investigation involving participants that evaluates the effects of one or more health-related interventions on health outcomes. Interventions include, but are not restricted to, drugs, radiopharmaceuticals, cells and other biological products, surgical procedures, radiologic procedures, devices, genetic therapies, natural health products, process-of-care changes, preventive care, manual therapies, and psychotherapies. Clinical trials may also include questions that are not directly related to therapeutic goals—for example, drug metabolism—in addition to those that directly evaluate the treatment of participants.

Clinical trials are most frequently undertaken in biomedical research, although research that evaluates interventions, usually by comparing two or more approaches, is also conducted in related disciplines, such as psychology. The researcher leading a clinical trial is often (but not always) a clinician, that is, a health-care provider (e.g., physician, dentist, naturopath, physiotherapist, etc.). Although various types and forms of clinical trials have methodological differences, the ethical principles and procedures are the same and are applicable to all.

History of Experimentation in Clinical Research

Clinical research has a long and rich history, dating back to as early as 2737 BCE; the first clinical trial is documented in the Old Testament. Since this first trial, the field has changed and progressed immensely, with the refinement of research methodology and practice. The most progress in the methodology and use of clinical trials has occurred in the past 50 years, changing significantly the landscape of clinical research.

2737 BCE

Shen Nung, legendary emperor of China, is considered the father of Chinese medicine. In addition to being credited for the technique of acupuncture, he purportedly experimented with and classified hundreds of poisonous and medical herbs, which he tested in a series of studies on himself.

Approximately 600 BCE

The first experiment that could be considered a trial can be found in the Old Testament. The book of Daniel describes how under King Nebuchadnezzar II,

children of royal blood and certain children from the conquered Israel were recruited to be trained as the king's advisors over a period of three years during which they would be granted to eat from the king's meat and wine. Daniel, however, requested from the officer in charge of the diet that he and three other Hebrew children would be allowed to have only legumes and water. When the officer expressed concerns about the "inferior" diet, Daniel suggested a 10-day trial period, after which the officer would assess both groups of children. At the end of this "pilot study," Daniel's group was noticeably healthier than the group of children who were relegated to the diet of wine and meat. Therefore Daniel and the other three children were permitted to continue with their diet for the entire training period, after which they displayed superior wisdom and understanding compared to all other advisors of the king (Old Testament, Daniel 1:5–20).

1537

Ambroise Paré, a French surgeon during the Renaissance, accidentally carried out a clinical study when he ran out of elderberry oil, which after being boiled was used as the standard treatment for gun wounds at that time. He then used a mixture of egg yolk, turpentine (a pine tree–derived oil), and rose oil instead, and he soon noticed that patients treated with this mixture had less pain and better wound healing than those patients who had received the standard treatment [1].

1747

The first reported systematic experiment of the modern era was conducted by James Lind, a Scottish physician, when he was sailing on the Salisbury. After many of the seamen developed signs of scurvy, Lind selected 12 similar sick sailors and split them into six groups of two. All groups were given the same diet, but each group was treated differently for scurvy: group 1 received cider, group 2 vitriol (sulfuric acid), group 3 vinegar, group 4 seawater, group 5 oranges and lemon, and group 6 nutmeg and barley water (British herbal tea). The group that had the fruits recovered from scurvy within just six days. Of the other treatments, vinegar "showed the best effect" [Dr. James Lind. "Treatise on Scurvy." Published in 1753, in Edinburgh]. (Lind's experiment had little short-term impact; he was reluctant to believe in fruits being a sole remedy for scurvy, and citrus fruits were difficult to preserve and also expensive. It wasn't until 1790 that the Royal Navy made fresh lemons a standard supplement. In 1932 the link between vitamin C and scurvy was finally proven.)

In 1747, Dr. James Lind tested several scurvy treatments on crew members of the British naval ship Salisbury and discovered that lemons and oranges were most effective in treating the dreaded affliction.

1863

Austin Flint (US physician, graduate of Harvard Medical School, class of 1833) is revered for having conducted the first study with a placebo. A placebo is considered a substance or procedure with no therapeutic effect. In 1863 Flint tested a placebo on prisoners with rheumatic fever and compared their response

to the response of patients who had received an active treatment (although not in the same trial). (Austin Flint murmur is a murmur associated with aortic regurgitation. This trial is somewhat problematic ethically [research conducted on a vulnerable population] and methodically [placebo and active treatment were not tested at the same time, and active treatment for rheumatic fever was questionable to be active]).

1906

The 1906 Pure Food and Drug Act imposed purity standards on products and drugs and mandated accurate labeling with content and dose.

1923

The idea of randomization was introduced to clinical trials in 1923. Randomization involves participants randomly receiving one of the treatments, one being a placebo and one being the new drug.

1938: The Federal Food, Drug, and Cosmetic (FDC) Act

This act required that new drugs be shown to be safe before marketing, thus starting a new system of drug regulation. It also provided that safe tolerances be set for unavoidable poisonous substances.

1943

Blind clinical trials—in which neither group knows which treatment they are receiving—also emerged in the twentieth century. The first double-blind controlled trial—Patulin for the common cold—was conducted, and the first widely published randomized clinical trial was conducted on Streptomycin as a treatment for pulmonary tuberculosis [2].

1944

Multicenter clinical trials were introduced, in which multiple studies were conducted at various sites, all using the same protocol to provide wider testing, generalization, and better statistical data.

1947

The Nuremberg Code was developed, which outlines 10 basic statements for the protection of human participants in clinical trials.

1964

The Declaration of Helsinki was developed, which outlines ethical codes for physicians and for the protection of participants in clinical trials worldwide.

1988

The US Food and Drug Administration (FDA) was given more authority and accountability over the approval of new drugs and treatments.

1990

The International Conference on Harmonization (ICH) was assembled to help eliminate differences in drug-development requirements for three global pharmaceutical markets: the European Union, Japan, and the United States. The ICH initiatives promote increased efficiency in the development of new drugs, improving their availability to patients and the public.

2000

A Common Technical Document (CTD) was developed. The CTD acts as a standard dossier used in Europe, Japan, and the United States for proposing data gathered in clinical trials to respective governing authorities.

THE HISTORY OF ETHICS IN CLINICAL RESEARCH

Disasters in Clinical Research

A significant part of the ethical regulations in clinical research was catalyzed by disastrous events that took place at the beginning of the new era of clinical research, when experimentation and the development of novel treatments began to take place in a more systematic way. We review these events in the context of US and global changes in ethical regulations. Although changes in ethical regulations had a similar pathway worldwide, there have been some differences between US and global changes in ethical regulations for clinical research.

US Disasters and Responses

Sulfanilamide Cold Syrup

The months of September and October of 1937 served as the tragic background for one irreparable clinical disaster. The medication responsible for this disaster was known as "Elixir sulfanilamide," and was commonly and safely used for the treatment of streptococcal infections for several years, mainly in the form of powder and tablets. On June 1937, the S. E. Massengill Company, located in Bristol, Tennessee, created and distributed 633 doses of the same compound but in liquid form, by dissolving the existing powder into ethylene glycol. The distributing laboratory controlled the new substance by testing its flavor, appearance, and fragrance and decided that it was satisfactory. Nonetheless, as safety testing was not a legal requirement by the FDA at the time, the new preparation was not tested for safety. The company then failed to understand that ethylene glycol was a toxic compound, used frequently as antifreeze.

Approximately 1 month after the distribution of the new liquid preparation of Elixir sulfanilamide, the first report of associated deaths was made public. The

American Medical Association (AMA) was the first to publicly announce the toxicity of the new compound and to warn physicians and patients against its lethal effects. The S. E. Massengill Company was also notified; the company then sent telegrams to distributors, pharmacists, and physicians, asking them to return the product, but failed to explain the reason for the request, thus undermining the urgency of the situation and the lethal effects of the product. At the request of the FDA, the company was forced to send out a second announcement, which was clear about the toxicity of the product and the importance of the situation.

The next step taken by the FDA was to make sure all of the products were returned safely; to do so, they had to locate all of the stores, dispensers, pharmacists, physicians, and buyers. This proved to be a difficult task: many of the company's salesmen were not willing to help by providing the information required to locate the recipients of the Elixir; pharmacies had no clear record of buyers; and many physicians didn't keep documentation of the patients to whom the compound was prescribed, nor their addresses. Some physicians decided to abstain from helping authorities and lied about their prescription trail, afraid that they could be held liable for prescribing the medication. In spite of these circumstances, the relentless efforts of the FDA and local authorities, as well as the help of the AMA and the media, allowed for the recovery of 234 of the 240 gallons of the drug that had been distributed. In several cases, legal action through federal seizure was required. The FDA had to refer to the compound's branding name "Elixir" to file federal charges that would allow them to complete their task. The misbranding charge was brought to the company for distributing a compound as an elixir, meaning it was dissolved in alcohol, when it was actually dissolved in ethylene glycol.

The victims were many, including young children—most sick with throat infections—young workers, older patients, mothers, and fathers. Most of the victims would suffer the lethal effects of the substance for 10–21 days before succumbing to death. The symptoms were mainly associated with severe renal failure and included oliguria, edema, nausea, vomiting, abdominal pain, and seizures.

The Response: The Federal Food, Drug, and Cosmetic Act of 1938

Given these unfortunate events and other disasters associated with medications that had not been properly tested, the FDA emphasized the need for stricter control of the production and distribution of new drugs that could ensure the welfare of patients and consumers. The enactment of the Federal Food, Drug, and Cosmetic Act of 1938 provided a new system for drug control and safety, which not only solidified the conditions required for the release of new medications, but also stimulated medical research.

Thalidomide

Thalidomide is considered a derivative of glutamic acid, first synthesized unsuccessfully in Europe by Swiss Pharmaceuticals in 1953. A German company, Chemie Grunenthal, then remarketed it in 1957 as an anticonvulsant. Given its sedative effects, it was also commercialized as a sleeping aid. It became a very popular medication, considered effective and safe, and was highly sought and prescribed due to its lack of

hangover effect, as supported by small clinical studies conducted in Europe and in which the subjects were unaware of the medication being tested [3,4]. By the end of the decade, more than 10 pharmaceutical companies in all five continents were selling the drug. It was even said that the medication became nearly as famous as aspirin [5]. Its use became widespread and unregulated; it was prescribed for treatment of almost every condition, from simple colds to anxiety and asthma. Some even went as far as to attribute to the medication some beneficial effects on diabetes, cancer, and autoimmune diseases [6].

Soon after, some doctors started to recommend the off-label use of thalidomide for the relief of morning sickness in pregnant women. This use also became widespread in the late 1950s and early 1960s. After this unprecedented success in Europe, the company William S. Merrel made a formal application to the FDA for marketing thalidomide as an over-the-counter medication, which could be used for the treatment of many ailments, from anorexia and poor school performance, to asthma and tuberculosis.

The petition was assigned for review to a new FDA reviewer, Dr. Frances Kelsey. After her initial review of the available information, she became concerned with some of the initial reports of adverse effects associated with the use of thalidomide and the lack of well-designed safety studies [7]. The initial reports released regarding safety concerns were associated with irreversible peripheral neuropathy in chronic users of thalidomide [8]. At the time, previous research had shown that medications associated with peripheral nerve irritation could lead to growth retardation and birth defects, shown mainly in rabbit fetuses. These concerns were dismissed by the company's CEO, who claimed that the peripheral neuropathy was associated with isolated cases and was reversible upon discontinuation of the drug [6,9].

Given this information, Dr. Kelsey put a hold on the approval of the request by Merrel's company and asked for further data regarding safety studies and data on the effects of the biologically active compound in pregnant patients, as one of the indications for the new drug would be morning sickness. The data were not available, which led to a significant delay in the approval process. At the same time, tragedy struck as new reports of adverse events were made public [7].

It was a bittersweet surprise for both the American public and lawmakers when they found out that the only thing that kept the ill-proved drug from being commercialized in the United States was a cautious FDA agent. She had reported the safety proofs provided by the company as testimonial and not as the result of well-designed and executed studies, rendering the application incomplete and withdrawing it from approval. After the studies made clear that there was a causal association between thalidomide and severe birth defects, there was a clear urgency in recapitulating in the face of such tragedy.

The thalidomide disaster brought to light the conditions of a still ill-equipped regulatory process on drug synthesis and marketing, accounting for the weakness of the FDA in regulating efficacy, safety, test conduction, and accountability. The way in which the agency operated at the time rendered the process ineffective and left almost all of the responsibility and control in the hands of drug manufacturers. The FDA had a 60-day period to prove that a new drug was not safe and contradicted the studies conducted by the pharmaceutical company; otherwise the product would be approved. The companies controlled the conduction of clinical tests, so they could

complete new studies without FDA approval or even subjects' consent, and even worse, after approval, all new data regarding the drug were considered private. All of these situations clearly showed that pharmaceutical companies had an upper hand in the game.

Response: The Kefauver-Harris Drug Amendments of 1962

At around the same time as the thalidomide disaster started to spark, Democratic Senator Estes Kefauver proposed a bill to the Senate that included price control on the products of pharmaceutical companies and tougher safety controls. The bill was widely discussed and debated, but was never passed. Well after the disaster became clear, interest in the bill revived, and Senator Kefauver modified it to include only the changes to safety control, including some of the weakness that were brought to light by Dr. Kelsey's story. The new bill made important modifications to the existing system, including the abolishment of the approval time period, which then shifted the weight of accounting for drug safety and efficacy from the FDA to the pharmaceutical companies. It also mandated that such proofs be based on well-designed and conducted clinical studies using up-to-date techniques and practices. All clinical trials had to be approved by the FDA, and all participants had to be properly informed and consented. Also, all information regarding efficacy and safety was to be shared with the FDA during all stages of the marketing process, and the agency had the freedom to make new sets of rules and limitations regarding the approval process, including immediately removing drugs from the market [10–12].

Tuskegee Study

This study, also known as the "The Tuskegee Study of Untreated Syphilis in the Negro Male," was conducted in Alabama by the United States Public Health Service (USPHS) and the Tuskegee Institute between 1932 and 1972 [13]. During this period of time, hundreds of African-American males did not receive proper and standard care for syphilis, with the intention to document the natural course of syphilis infection if it was left untreated. During the 40 years that the study took place, many of the enrolled subjects, who came from a poor, rural area in Alabama, died of syphilis, and many of their descendants were born with congenital syphilis. Directors, researchers, and collaborators of the study observed the tragic effects of the disease, completely indifferent to the suffering of their subjects, and even decided to continue their study after penicillin was proven to effectively treat and cure the infection [13,14].

In 1928, scientists from Oslo published a study conducted in white males with untreated syphilis that refuted the long-lived belief that the effects of syphilis depended on the race of the affected. It was thought that this infection had more severe neurologic effects on people of Caucasian descent, while it produced more severe cardiovascular effects in population of African-American descent, but the study from Oslo showed that most of the infected white males had severe affectation of their cardiovascular system, but very few ever developed neurosyphilis [15]. This finding amazed physicians and researchers in the United States, which led them to plan and execute a similar study, which would be carried out in a population with high prevalence of the infection. American scientists chose the city of Tuskegee because 35%–45% of

the population in the area was seropositive for syphilis. The initial design proposed observing the untreated subjects for a period of 6–8 months, after which the subjects would be treated with the standard care: salvarsan and bismuth, both fairly effective but toxic. The initial purpose of the study was to benefit the health of the poor population enrolled, as well as to understand and learn more about the disease, its prevention, and its cure. This led to the support of many of the local hospitals and physicians, including African-American doctors and organizations [14,16].

Researchers initially enrolled 600 men, 201 as healthy controls and 399 seropositive for syphilis but not yet aware of their diagnosis. The subjects proceeded from Macon County in Alabama, where the city of Tuskegee was located; they were mostly illiterate men, lured into the study by the promise of free medical care, free daily meals, and US$50 for burial expenses for their participation. Throughout their participation in the study, the participants were not informed of their infection status, nor did they receive treatment. In many instances, researchers used deception to assure cooperation and avoid dropouts, including making the burial policy contingent to their previous authorization for an autopsy. After the initial allotted time for the study was completed, many of the participating researchers decided it was necessary to continue the study and obtain more clinical information. As the study continued, the great economic crisis of 1929 was gestating, which led to withdrawal of the main funding source. Researchers thought this would mean the end of the experiment, as it would be impossible to afford treatment for all participants, but soon they proposed continuing the study without offering standard care to patients, leading to a complete deviation from the initial proposal and to the resignation of one of the initial creators of the study, Dr. Taliaferro Clark [13,17].

It is important to note that during the 40 years of the Tuskegee experiment, the study was never kept secret, and many articles were published in medical journals describing initial discoveries and important information obtained from the research [18–21]. Despite its controversial and irregular techniques, many argued that the contribution of this study to science far outweighed its detrimental effects on the health of the studied population. One of the main contributions of the experiment was the development of the Venereal Disease Research Laboratory (VDRL), a non-treponemal diagnostic test now widely used. This sign of research and medical progress was instrumental for establishing a renowned position for the United States in the international research scenario, which served as an impetus for the ambition of many of the participating researchers [13,17].

By 1947, penicillin had long been established as the most effective treatment for syphilis and it was widely used for such purpose, leading to a significant decrease in the prevalence of the disease. Its efficacy was so clear that many even argued that syphilis would be completely eradicated in the near future. Nonetheless, researchers of the Tuskegee study continued to deny proper treatment to their subjects, and they were specifically warned against the use of penicillin and carefully shielded from receiving any information regarding its benefits. By the end of the study period, only 74 of the original 399 men were alive; 128 of them had died of syphilis and related complications, 40 of their wives had contracted the disease, and 19 children were diagnosed with congenital syphilis [16,22]. The relentless ambition of the Tuskegee researchers continued in spite of the establishment of the Nuremberg Code in 1947, the declaration of Helsinki in 1964, and the position of the Catholic Church urging

physicians and scientists to always respect patients, superseding all scientific or research objectives [16,23].

In 1966, Peter Buxtun was the first researcher to speak publicly about his ethical concerns for the Tuskegee study. He warned the Division of Venereal Diseases about the practices and techniques being used in the study, but the Centers for Disease Control (CDC), now in charge of the experiment, argued for the importance of completing the study and obtained the support of local and national medical associations. In 1972, Buxtun took the story to the national press, which led to widespread outrage [16]. Congressional hearings were held, many researchers and physicians testified, and the deplorable objectives and practices of the study were exposed. The CDC appointed a review committee for the study, which finally determined that it was not justifiable from an ethical and medical standpoint, leading to its termination.

The Response: Belmont Report—Ethical Principles and Guidelines for the Protection of Human Subjects of Research, Report of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research

The aftershocks of these revelations highlighted some of the main issues that afflicted the medical and scientific community of the time. As a response, the United States Congress passed the National Research Act in 1974. Its main concern was to guarantee stricter regulation and control of clinical trials. Now studies would always be required to pass through an Institutional Review Board, obtain inform consent from all study participants, and always inform and counsel participants regarding their diagnosis and study practices and techniques [24].

The National Research Act also created a Commission within the Department of Health and Human Services, destined to shape bioethics within the United States. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research issued and published the Belmont Report in 1979. This report established three core principles that should always guide the design and review of clinical trials: respect for persons, beneficence, and justice. The first refers to one of the critical issues of the Tuskegee study—autonomy—and includes the mandatory need to obtain an informed consent from every study participant. This consent would be based on the truth, and not deception or misinformation. Beneficence refers to the basic principle of "do no harm," based on the conception of minimizing risk and maximizing benefits for the research participants. And finally, justice assures safe procedures that do not intend to exploit the participants, while arguing for a fair distribution of costs and benefits [25].

Global Disasters and Responses

Nazi Human Experimentation

The twentieth century was the background for spectacular international advances in the medical and scientific field, but some of them were associated with troubling events taking place against the international backdrop of a deadly war. The atrocities

and horrendous experiments that took place during this time led to the death of most of the participants, but the few survivors were able to narrate their suffering and leave a record for history.

This period of science should be discussed and described in the light of the historical and sociopolitical events that took place at the time. Mainly, we need to point out two determining factors in order to understand the beginning of this period of Nazi experimentation and the nature of such experiments: (1) the political structure of Nazi Germany, based on a totalitarian system, and (2) the racial hygiene paradigm that arose from both political and social movements at the time. The origin of the latter preceded by roughly two decades that of the Nazi government; nonetheless, it was the totalitarianism of the time that allowed for such an ideology to flourish and to give rise to the scientific questions that were later addressed by researchers and physicians of the Nazi regime. All of them took place in a setting in which no legal or ethical boundaries existed, leading to the ideal conditions for such experimentation to take place [26].

Based on these ideas, many German scientists and physicians, mainly geneticists, found in the newly formed Nazi government the opportunity to put into practice their theories and discoveries, while at the same time, the government found in the researchers an opportunity to legitimize its political and social beliefs of racial superiority. The research scenario was further darkened by the complete violation of the civil rights of the Jewish population, which was then rendered as freely available "guinea pigs" for any research agenda. Resources were redistributed to any scientific quest that would improve the health of the superior race, and the focus of most research programs became heredity and fitness [27]. The most striking aspect of all of the experimentation that took place in the Nazi Germany is that there is no direct proof that any of the researchers involved were forced to participate, or that any of the research techniques and practices were merely imposed by the government [28,29].

The idea of biological inferiority led to unimaginable cruelty and disrespect for the unconsenting subjects. It is impossible to name and include all of the examples of such cruelty for the purpose of this review, but it is important to mention that most of the experiments conducted in the concentration camps followed the strict guidelines of clinical research of the time, some of them pursuing questions in accordance with the scientific progress of the time, though some of them used obsolete or outdated practices [26]. Some of the methods and results could even be considered innovative and helpful, as certain experiments were later continued by the conquering armies, aided by the same German physicians, but following the newly established laws of ethical research. It is fair to say, however, that regardless of the results or objectives of each study, their methods were always brutal, and researchers had a complete disregard for human life and suffering [28]. The main justification for their actions was based on the ideal of preserving the health and well-being of the population, at the same time that new critical knowledge was gained from such endeavors.

Response: The Nuremberg Code

After the end of World War II, all of the participants and collaborators of the Nazi government were brought to trial. The judgment of Nazi physicians in Nuremberg, known as the Nuremberg trial, is considered the precipitating event for the start of modern

research ethics. From this trial, the founding principles of ethical research were established under the Nuremberg Code. It outlined 10 critical aspects for the conduction of experimentation with humans [30]:

- Obtaining voluntary consent from all participants, which should be based on complete and sufficient knowledge of the purpose, methods, and duration of the experiment.
- The main objective was to render fruitful results for the health and well-being of society, which could not be accomplished by any other method.
- Prior knowledge and animal research should be used as the base for any new research development.
- Researchers should avoid all unnecessary mental or physical suffering to participants.
- Experiments should not incur in any permanent injury or death, unless physicians were also considered participants in the study.
- Risk:benefit ratio should be evaluated, and risk should never be greater than possible benefits.
- All necessary tools and facilities should be available and provided to prevent injury, suffering, or death of subjects.
- Only highly trained and skilled personnel should conduct the research.
- Subjects are free to withdraw their participation if they reached their maximum point of physical and/or mental tolerance.
- Researchers should be willing and able to end any experiment if they believed that there was a risk to subjects' integrity.

Thalidomide in the International Context

As previously mentioned, the thalidomide disaster was of international magnitude. The initial reports of adverse events in both adults and newborns originated in Europe and Australia. Most of the affected patients were outside the United States. The first reports of the possible teratogenic effects of thalidomide came from Dr. William McBride, an Australian obstetrician, and Dr. Widuking Lenz, a German pediatrician, who proposed an association between the increase in birth defects and the use of thalidomide in pregnant women [31].

By 1958 the drug was sold as an over-the-counter medication in West Germany, and by 1961 46 countries were doing the same. One of the biggest concerns of its early distribution is that the medication was marketed as "completely safe," and it took several years before this claim would start crumbling. The pharmaceutical company initially dismissed the reports of irreversible peripheral neuropathy caused by long-term use of the drug. These reports were only considered serious on 1961, when West Germany restricted the market of thalidomide to only by prescription and forced the company to remove the "non-toxic" argument from the label. The second hit to the reputation of thalidomide started when doctors prescribing the medication to pregnant women noted a congruent increase in birth deformities, reaching an unprecedented incidence, with physicians delivering several cases of phocomelia per month. The reports suggested that the use of thalidomide during the first trimester of the pregnancy, probably even before the mother knew she was pregnant, could

be associated with the problem. Again, the company denied the reports and qualified them as a cheap intent to murder a perfectly safe drug. Due to public pressure, the distribution of the medication was halted in Germany, but continued in other countries. Only when the news arrived of birth deformities did each country establish restrictions on the medication. Canada was the last country to stop the sale of thalidomide in 1962.

It is estimated that from the late 1950s to the early 1960s, over 13.000 babies were born with several deformities, including phocomelia, secondary to use of thalidomide during pregnancy. Many of them died soon after birth, but many lived long lives, with many surviving until 2010.

Response: Declaration of Helsinki in 1964

This document, prepared by the World Medical Association, consisted of an authoritative attestation of the importance of conducting previous concise and serious review of any research protocol with human experimentation, encouraging every researcher to apply the principles promulgated by the declaration. The initial document established clear boundaries for the conduct of clinical research, but left certain freedom for the investigator to decide if there were special circumstances for the experiment that allowed subjects to participate without previous informed consent.

In 1975, a revision was made that included the need for the review of research protocols by independent review boards. This declaration recognized that medical progress was based on research that will eventually include experimentation with human subjects, but clarified that the goal of gaining new knowledge should never surpass the need to protect the rights and health of the patients, including those participating in medical research. (For more information, see: the Declaration of Helsinki, World Medical Association; http://www.wma.net/en/30publications/10policies/b3/.)

THE PROCESS OF TESTING NEW INTERVENTIONS: STUDY PHASES

Based on the lessons learned from the major disasters in creation and commercialization of novel therapeutic drugs, the process for the development of a new drug or device has been systematized, and safety has become a major issue. The importance of safety evaluation is such that it is currently the first issue to be assessed during the development process, and it is also measured throughout the other phases of development.

Let us quickly summarize the drug development phases:

Preclinical study: Consists in completing a rigorous animal testing previous to application before the FDA for an investigational new drug (IND). Most drugs that undergo animal testing do not make it to human experimentation.

Phase I: Its main goal is to determine the pharmacokinetic and pharmacodynamics parameters of the drug and its safety in human subjects. It is mostly conducted in a group of 20–80 healthy volunteers. The main parameter to determine progression to Phase I is proof of safety (mainly no severe toxic effects).

Phase II: Its main goal is to obtain preliminary data on the efficacy of the medication on a given population. Usually, the study will be conducted in a group of diseased patients, which can range from a dozen to 300. The study should be controlled, meaning that the diseased population receiving the new drug being studied has to be compared to a control diseased population receiving either placebo or any standard medication available. This phase continues to evaluate for drug safety and short-term side effects.

Phase III: If evidence of effectiveness is shown in phase II studies, then the process can continue to phase III. The main goal of this phase is to assess effectiveness and safety. For this purpose, larger study populations should be evaluated and "real-life" conditions emulated, in order to assess the behavior of the drug when given at different doses, in heterogeneous populations or compared against the standard of care. The number of patients can range from several hundred to 3,000–10,000.

Phase IV: This phase is also known as post-marketing survey. It takes place after the drug has been approved by the FDA and has been put in the market. Post-marketing surveillance and commitment studies allow the FDA to collect further information on safety, efficacy, and tolerability profile of any given drug.

CASE STUDY: THERE IS NO FREE LUNCH—THE COST AND BENEFIT OF SEEKING THE CURE OF HIV INFECTION

Fabio Pinna, André Brunoni, and Felipe Fregni

Introduction

John Geegs is a PhD researcher form Pennsylvania University. He is one of the most prominent researchers in gene therapy, having published several influential articles in prestigious journals. Despite having a brilliant career with a solid position in academia, he has big dreams. As he describes himself, he is a "dreamer" who pursues nothing more than the final cure for HIV infection.

Since earning his PhD in Molecular, Cell, and Developmental Biology in one of the most competitive programs at UCLA, he has been dedicating his career to translational research. As a scientist, he has a passion for basic laboratory research, but also has an urgent need to apply these findings to clinical situations. In fact, he fits well with the NIH (National Institutes of Health) profile of a clinical translational scientist. As defined by this agency, "Translational research . . . is the process of applying discoveries generated during research in the laboratory, and in preclinical studies, to the development of trials and studies in humans [32]." It is also commonly known as taking knowledge from "bench to bedside." As a matter of fact, like most basic science research, Prof. Geegs's scientific breakthroughs have not yet been translated into major changes in medical therapy for humans. Truly, he deeply believes that gene therapy will be used in the near future in the treatment of severe conditions, such as cancer or HIV.

At his laboratory at UPenn, he mentors five brilliant young doctors—one of them, Dr. Ryan Stevenson, has the same dreams and passion for gene therapy as Prof. Geegs. Dr. Stevenson was born in Sweden but moved at a young age to the United States. His mother had a terrible condition—Li-Fraumeni syndrome (a rare autosomal dominant disorder associated with the *CHEK2* and the *TP53* genes that greatly increases the risk of developing several types of cancer, particularly in children and young adults, including breast cancer, brain tumors, sarcomas, leukemia, adrenocortical carcinoma, and others)—and she died at a very young age of a brain tumor. Luckily, Dr. Stevenson did not have the mutation of the p53 suppressor gene, but one of his sisters did, and she died at age 18 of adrenal carcinoma. He then decided to devote his life to medicine, and soon after he graduated with a degree in medicine in Chicago, he moved to UPenn for a PhD in molecular biology. His initial plans were to do a master's program and thereafter to start his residency in Oncology, but the outstanding reputation of Prof. Geegs and the belief that he could do more for patients as a researcher made him decide to follow this pathway. He felt that all the suffering that he had been through with his family's disease could bring something positive in the future.

Dr. Stevenson knows that although gene therapy is still in its infancy, it has been showing satisfactory results in the treatment of several hereditary diseases. As he likes to explain to the graduate students of Prof. Geegs (Dr. Stevenson is his teaching assistant in his course, Applied Molecular Biology) "gene therapy is based on the idea of gene insertions into one's cells and tissues in order to replace a mutant allele.

The replacement of non-functional genes can be done by insertion of multiple vectors being the use of different types of viruses—probably the most effective method!"

A Cold Winter in Pennsylvania and the Jesse Gelsinger Case

It was a late afternoon of a cold winter in Pennsylvania. Although Prof. Geegs does not enjoy cold weather, as he lived most of his life in Southern California, for him, winter is a particularly productive season, as students are often busy with final exams and grant deadlines are usually in the spring. On one of these cold days, in a staff meeting, Prof. Geegs and his team were having a pleasant chat while having coffee and talking about designing a new HIV phase I study using gene therapy. Although the conversation was collegial and pleasant, suddenly there was a moment of awkwardness in the room, and Prof. Geegs with a sad look said to his staff, "OK—let us discuss the "elephant in the room"—we shall never forget the Gelsinger case!" and he started talking about it:

"As not all of you know, let me narrate again this important ethical case. Jesse Gelsinger was an 18-year-old patient who suffered from onithine transcarbamilase defciency, an X-linked genetic disease of the liver that is characterized by the liver's inability to metabolize ammonia, a byproduct of protein breakdown. Indeed, the disease is usually fatal at birth. However, Jesse Gelsinger had the non-severe form of the disease, as some of his cells were normal, which enabled him to survive on a restricted diet and special medications. Gelsinger was really excited to join a UPenn phase I trial that aimed at developing a new drug for children born with the severe form of the disease. He was the last subject of his group. During the phase I study, Gelsinger was injected with adenoviruses carrying a corrected gene to test the safety of the procedure. Four days later, he died by immune response triggered by the use of the vector used to transport the gene into his cells. Consequently, this led to multiple organ failure and brain death. Afterward, the FDA (Food and Drug Administration) ran an investigation on this case and concluded that this phase I trial broke several rules of conduct, such as:

1. Inclusion of Gelsinger as a substitute for another volunteer who dropped out, despite having high ammonia levels that should have led to his exclusion from the trial;
2. Failure by the university to report that two patients had experienced serious side effects from the gene therapy; and
3. Failure to mention the deaths of monkeys given a similar treatment in the informed consent documentation.

After this terrible incident, which shot down human research at UPenn for several months and led to a detailed investigation, UPenn paid the parents an amount of money in settlement. Both the university and the principal investigator (PI) had serious financial stakes."

Finishing his thoughts, Prof. Geegs concluded, "The Gelsinger case was an important setback for phase I studies in gene therapy."

Dr. Stevenson, thinking about his family with Li-Fraumeni syndrome, said, "The thought of benefits at any cost has brought up terrible lessons for humankind such

as the Tuskegee Study in 1932 [a syphilis experiment conducted between 1932 and 1972 in Tuskegee, Alabama, by the US Public Health Service in which impoverished African-Americans with syphilis were recruited in order to study the natural progression of the untreated disease] or the thalidomide case in 1959 in Germany [a drug that was used to inhibit morning sickness during pregnancy and resulted in thousands of babies being born with abnormalities such as phocomelia]. We should not disregard the issue of ethics and regulatory requirements in any phase of a drug trial, especially phase I!"

Prof. Geegs looked at his watch and realized he was late to meet a group of researchers from Japan who had come to visit his laboratory. He then wrapped up the discussion, "Guys, let us continue this discussion tomorrow; and I also want you to do a bit of research on the phases of a trial, so we can continue our discussion."

Phases of a Trial

In the investigation of a new drug, sequences of clinical trials must be carried out. Each phase of a trial seeks to provide different types of information about the treatment in relation to dosage, safety, and efficacy of the investigational new drug (IND).

Preclinical research: Before using an IND in humans, tests should be taken in the laboratory usually using animal models. If an IND shows good results in this phase, then researchers are able to request permission to start studies in humans.

Phase I trial: The aim of this phase is to show that the IND is safe. Data are collected on side effects, timing, and dosage. Usually dosage is increased until a maximum dosage (predetermined) or development of adverse effects are found. It usually requires a small sample size of subjects and it helps researchers to understand the mechanism of action of a drug. Much of the pharmakocinetics and pharmacodynamics of INDs are researched in this phase. Also during this phase, the drug is usually tested in healthy subjects, except for some drugs such as oncologic and HIV drugs.

Phase II trial: Once an IND is found to be safe in humans, phase II trials focus on demonstrating that it is effective. This is also done in relatively small sample sizes, in studies often referred to as "proof-of-principle" studies. The response rate should be at least the same as standard treatment to encourage further studies. These small trials are usually placebo-controlled.

Phase III trial: Also referred to as pivotal studies, they represent large studies with large samples and are usually (but not always) designed as a randomized, double-blinded trial comparing the IND to the standard treatment and/or placebo. Successful outcomes in two phase III trials would make a new drug likely to be approved by the FDA.

Phase IV trial: Also referred to as post-marketing studies, in phase IV trials, approved drugs are tested in other diseases and populations and usually in an open-label fashion.

Early Morning Meeting and Gene Therapy in HIV

The next morning, Prof. Geegs arrived in the lab's small conference room. Prof. Geegs started the meeting, saying, "Because Dr. Wang is our new post-doctoral fellow from Beijing, China, I want someone to explain our preliminary HIV study to her." Dr. Stevenson, the senior post-doc, quickly volunteered: "It will be a pleasure to do so.

Our team at the University of Pennsylvania School of Medicine has recently reported the first clinical test of a new gene therapy based on a disabled AIDS virus carrying genetic material that inhibits HIV replication. In this first trial, we studied five subjects with chronic HIV infection who had failed to respond to at least two antiretroviral regimens, giving them a single infusion of their own immune cells that had been genetically modified for HIV resistance. In the study, viral loads of the patients remained stable or decreased during the study, and one subject showed a sustained decrease in viral load. T-cell counts remained steady or increased in four patients during the nine-month trial. Additionally, in four patients, immune function specific to HIV improved." Prof. Geegs, who was extremely excited about these findings (and the approval for the paper's publication in Proceedings of the National Academy of Sciences (PNAS)), could not resist interrupting and added, "Overall, our results are significant, because it is the first demonstration of safety in humans for a lentiviral vector (of which HIV is an example) for any disease." Although Dr. Wang was still jet-lagged from her long trip to the United States, she added, "Thank you so much, Dr. Stevenson. In fact, we appreciate the work of Prof. Geegs in Beijing and it is a wonderful opportunity to be here in the lab. What is the next step now?" Prof. Geegs responded, "Our results are good, but they are preliminary—meaning that we shall replicate it in a larger population. We have much more work to do. In the study we are planning, each patient will now be followed for 15 years."

Stevenson completed with the details of this new study, "The new vector is a lab modified HIV that has been disabled to allow it to function as a 'Trojan horse,' carrying a gene that prevents new infectious HIV from being produced." He continued, "Essentially, the vector puts a wrench in the HIV replication process. Instead of chemical- or protein-based HIV replication blockers, this approach is genetic and uses a disabled AIDS virus which carries an anti-HIV genetic payload. This approach enables patients' own T-cells, which are targets for HIV, to inhibit HIV replication—via the HIV vector and its anti-viral cargo."

Dr. Cameron, an extremely educated research fellow from Australia, then made a comment, "I believe that it is wonderful to go in this direction instead of drugs only as they have significant toxicity, but in the first trial, patients were still taking the drug. Do you think patients would be able to stay off drugs with this gene therapy, Prof. Geegs?"

Prof. Geegs liked to stimulate his fellows to think, and he asked Stevenson to respond—which he quickly did, with a subtle smile, "That is an excellent point, which is why, in this second trial using the new vector with HIV patients, we will select a group of patients who are generally healthier and use six infusions rather than one—we therefore want to evaluate the safety of multiple infusions and test the effect of infusions on the patients' ability to control HIV after removal of their anti-retroviral drugs. The hope is that this treatment approach may ultimately allow patients to stay off antiretroviral drugs for an extended period. This would be a great breakthrough for this laboratory."

Prof. Geegs quickly concluded, "But we should never forget the Gelsinger case as, you know, *fool me once, shame on you; fool me twice, shame on me . . .*. Our group should then reflect on the ethical implications in this case. I want you guys thinking about this subject tonight and send an email to the group with your conclusions. Looking forward to hearing back from you!"

Dr. Cameron—The First Email: Email Subject, "Too Risky for Subjects"

He starts the message with his usual politeness:

Greetings my dearest colleagues,
We obviously cannot predict or control all the possible side effects that can occur—or will probably occur, considering the risks of our investigational therapy. Although the pilot trial was OK, it had only a few patients and we do not know very well what the long-term effects of gene therapy may be. In addition, we cannot even submit our study to a grant; reviewers would kill our proposal very quickly. I would respectfully suggest to go back to our lab and think again about the next steps.

Warmest regards,
Cameron

Dr. Stevenson—The Quick Emotional Reply: Email Subject, "Risks Are Justified Based on Potential Individual Benefits"

Stevenson, thinking about his family, goes directly to the point:

Thanks Cameron—Remember our first study! We saw a significant decrease in viral load in two patients, and in one patient, a very dramatic decrease. There is hope here . . .. Imagine that you have HIV, you would like to enroll in a trial that could make you stop taking medications and perhaps be cured . . .. *No pain, no gain!* . . . we should go on with our trial!

Dr. Wang—The Late Response: Email Subject, "Risks Are Justified Based on the Knowledge Being Produced—Benefits for Society"

Dr. Wang had not been sleeping well due to problems adjusting. She then replied at 3 a.m. to the group:

Dear All,
Thank you for sharing so much knowledge. One point that I believe we should consider is the potential benefit for society. In fact I have reservations of having a trial that might benefit individuals. This is called "therapeutic misconception"—when subjects interpret a clinical trial as therapy rather than producing knowledge. But in this case as this study might benefit future patients, I think there is a reasonable justification for this trial.

Best wishes,
Sleepless Wang :)

The Next Morning

After reading all the emails, Prof. Geegs called all the fellows into his office, "OK! I enjoyed the discussion. Now I want everyone to rest and perhaps enjoy the last winter

weekend and next Monday we will discuss this ethics issue again. By now, I just ask you to reflect on combining ethics, benefits, and minimizing risks."

CASE DISCUSSION

This case illustrates how ethical dilemmas can influence the design of any given study. Particularly, the readers should pay special attention to the study phases of a given study and how to design a study of a novel intervention while keeping the safety of subjects as a main concern. The readers need also to identify that a clinical goal should not be applicable to a design of a given study; the clinician-scientist needs to use a different "hat" when designing and conducting a clinical study.

CASE QUESTIONS FOR REFLECTION

The following questions can be used to reflect on this case:

1. What challenges does Prof. Geegs face in choosing the next steps for his HIV study?
2. What are Prof. Geegs's main concerns?
3. What should he consider in making this decision?

FURTHER READING

Articles

Azeka E, Fregni F, Auler Junior JO. The past, present and future of clinical research. *Clinics*. 2011; 66(6): 931–2. [PMCID: PMC3129946] (Outstanding article in order to grasp the outline of clinical research)

Bhatt A. Evolution of clinical research: a history before and beyond James Lind. *Perspect Clin Res*. 2010 Jan–Mar; 1(1): 6–10. [PMCID: PMC3149409]

Blixen CE, Papp KK, Hull AL, et al. Developing a mentorship program for clinical researchers. *J Contin Educ Health Prof*. 2007 Spring; 27(2): 86–93. [PMID: 17576629]

Drucker CB. Ambroise Paré and the Birth of the Gentle Art of Surgery. Yale J Biol Med. 2008 December; 81(4): 199–202. [PMCID: PMC2605308]

Glickman SW, McHutchison JG, Peterson ED, et al. Ethical and scientific implications of the globalization of clinical research. *N Engl J Med*. 2009 Feb 19; 360(8): 816–823. [PMID: 19228627]

Goffee R, Jones G. Managing authenticity: the paradox of great leadership. *Harv Bus Rev*. 2005 Dec; 83(12): 86–94, 153. [PMID: 16334584]

Herzlinger RE. Why innovation in health care is so hard. *Harv Bus Rev*. 2006 May; 84(5): 58–66, 156. [PMID: 16649698]

Kottow MH. Should research ethics triumph over clinical ethics? *J Eval Clin Pract*. 2007 Aug; 13(4): 695–698. [PMID: 17683317]

Murgo AJ, Kummar S, Rubinstein L, Anthony JM, et al. Designing phase 0 cancer trials. *Clin Cancer Res*. 2008 Jun 15; 14(12): 3675–3682. [PMCID: PMC2435428]

Umscheid CA, Margolis DJ, Grossman CE. Key concepts of clinical trials: a narrative review. *Postgrad Med*. 2011 Sep; 123(5): 194–204.

Wilson JM. Lessons learned from the gene therapy trial for ornithine transcarbamylase deficiency. *Mol Genet Metab*. 2009 Apr; 96(4): 151–157. [PMID: 19211285].

Books

Beauchamp TL, Childress JF. *Principles of biomedical ethics*, 7th ed. New York: Oxford University Press; 2012. (7th edition of the first major American bioethics textbook, written by co-author of the Belmont Report)

Gallin JI, Ognibene FP. *Principles and practice of clinical research*, 3rd ed. Burlington, VT: Academic Press; 2012.

Groopman J. *How doctors think*. Boston, MA: Mariner Books; 2008.

Robertson D, Gordon HW. *Clinical and translational science: principles of human research*. Burlington, VT: Academic Press; 2008.

Online

FDA CFRs for regulations in details. www.fda.com

http://cme.nci.nih.gov/

http://firstclinical.com/journal/2008/0806_IIR.pdf

http://ohsr.od.nih.gov/

http://www.availclinical.com/clinical-study/clinical-trials-history/

http://www.cioms.ch

http://www.fda.gov

http://www.james.com/beaumont/dr_life.htm

http://www.niehs.nih.gov/research/resources/bioethics/whatis/

http://www.wma.net

https://web.archive.org/web/20131021022424/http://www.innovation.org/index.cfm/InsideDrugDiscovery/Inside_Drug_Discovery

https://web.archive.org/web/20120526134445/http://www.nihtraining.com/cc/ippcr/current/downloads/HisClinRes.pdf

https://web.archive.org/web/20160401233046/http://www.cancer.org/treatment/treatmentsandsideeffects/clinicaltrials/whatyouneedtoknowaboutclinicaltrials/clinical-trials-what-you-need-to-know-phase0

REFERENCES

1. Bull JP, A study of the history and principles of clinical therapeutic trials. MD Thesis, University of Cambridge, 1951. Available at http://jameslindlibrary.org/wp-data/uploads/2014/05/bull-19511.pdf
2. Streptomycin Treatment of Pulmonary Tuberculosis. *BMJ*. 1948; 2(4582): 769–782.
3. Burley DM, Dennison TC, Harrison W. Clinical experience with a new sedative drug. *Practitioner*. 1959; 183: 57–61.
4. Lasagna L. Thalidomide: a new nonbarbiturate sleep-inducing drug. *J Chronic Dis*. 1960; 11: 627–631.
5. Ghoreishni K. Thalidomine. In: *Encyclopedia of toxicology*. Burlington, VT: Academic Press; 2014: 523–526.
6. RiceE.Dr.FrancesKelsey:TurningthethalidomidetragedyintoFoodandDrugAdministration Reform. 2007. Available at http://www.section216.com/history/Kelsey.pdf
7. Kelsey FO. Inside story of a medical tragedy. *U.S. News & World Report*. 1962; 13: 54–55.

8. Fullerton PM, Kremer M. Neuropathy after intake of thalidomide (distaval). *Br Med J*. 1961; 2(5256): 855–858.
9. Mintz M. *By prescription only*, 2nd ed. Boston: Houghton Mifflin; 1967.
10. Tighter rein on drug testing. *Business Week*. January 12, 1963.
11. The Kefauver Hearings: the drug industry finally has its day and does quite well. *Science*. 1961; 134(3494): 1968–1970.
12. The drug hearings: to no one's surprise, Kefauver and the AMA do not agree on what should be done. *Science*. 1961; 134(3472): 89–92.
13. Cuerda-Galindo E, Sierra-Valentí X, González-López E, et al. Syphilis and human experimentation from World War II to the present: a historical perspective and reflections on ethics. *Actas Dermosifiliogr*. 2014; 105(9): 847–853.
14. Jones, J. *Bad blood: The Tuskegee syphilis experiment—a tragedy of race and medicine*. New York: The Free Press; 1981.
15. Gjestland T. The Oslo study of untreated syphilis: an epidemiologic investigation of the natural course of the syphilitic infection based upon a re-study of the Boeck-Bruusgaard material. *Acta Derm Venereol Suppl (Stockh)*. 1955; 35(Suppl 34): 3–368; Annex I–LVI.
16. Heller J. Syphilis victims in U.S. study went untreated for 40 years: syphilis victims got no therapy. *New York Times*. July 26, 1972. Associated Press.
17. Carmack HJ, Bates BR, Harter LM. Narrative constructions of health care issues and policies: the case of President Clinton's apology-by-proxy for the Tuskegee syphilis experiment. *J Med Humanit*. 2008; 29(2): 89–109.
18. Edmundson WF, Ackerman JH, Gutierrez-Salinas E, et al. Study of the TPI test in clinical syphilis. 1. Untreated early symptomatic syphilis. *AMA Arch Derm Syphilol*. 1954; 70(3): 298–301.
19. Olansky S, Simpson L, Schuman SH. Environmental factors in the Tuskegee study of untreated syphilis. *Public Health Rep*. 1954; 69(7): 691–698.
20. Peters JJ, Peers JH, Olansky S, et al. Untreated syphilis in the male Negro; pathologic findings in syphilitic and nonsyphilitic patients. *J Chronic Dis*. 1955; 1(2): 127–48.
21. Schuman SH, Olansky S, Rivers E, et al. Untreated syphilis in the male negro; background and current status of patients in the Tuskegee study. *J Chronic Dis*. 1955; 2(5): 543–558.
22. Katz RV, Kegeles SS, Kressin NR, et al. The Tuskegee Legacy Project: willingness of minorities to participate in biomedical research. *J Health Care Poor Underserved*. 2006; 17(4): 698–715.
23. Yoon C. Families emerge as silent victims of Tuskegee syphilis experiment. *New York Times*, May 12, 1997.
24. Stuart H. *History of medical ethics*. Oxford: Elsevier; 1998: 3: 165–175.
25. Van Valey T. *Ethical guidelines and codes of conduct in social and behavioral research*. Oxford: Elsevier; 2015: 37–43.
26. Weindling P. *Nazi medicine and the Nuremberg trials*. New York: Houndmills; 2004.
27. Lopez-Munoz F, Alamo C, García-García P, et al. The role of psychopharmacology in the medical abuses of the Third Reich: from euthanasia programmes to human experimentation. *Brain Res Bull*. 2008; 77(6): 388–403.

28. Roelcke V. Nazi medicine and research on human beings. *Lancet*. 2004; 364(Suppl 1): s6–7.
29. Roelcke V, Maio G. *Twentieth century ethics of human subjects research*. Sttugart: Steiner; 2004.
30. Guraya SY, London NJM, Guraya SS. Ethics in medical research. *Journal of Microscop and Ultrestructure*. 2014; 2: 121–126.
31. Lenz W, Knapp K. Thalidomide embryopathy. *Arch Environ Health*. 1962; 5: 100–105.
32. National Institutes of Health. RFA-RM-07-007: Institutional Clinical and Translational Science Award (U54). Mar 2007, Available at http://grants.nih.gov/grants/guide/rfa-files/RFA-RM-07-007.html.

2

SELECTION OF THE RESEARCH QUESTION

Authors: *Keiko Ueda and Lotfi B. Merabet*

Case study authors: *André Brunoni and Felipe Fregni*

The difficulty in most scientific work lies in framing the questions rather than in finding the answers.
—Arthur E. Boycott (Pathologist, 1877–1938)

The grand aim of all science is to cover the greatest number of empirical facts by logical deduction from the smallest number of hypotheses or axioms.
—Albert Einstein (Physicist, 1879–1955)

INTRODUCTION

The previous chapter provided the reader with an overview of the history of clinical research, followed by an introduction to fundamental concepts of clinical research and clinical trials. It is important to be aware of and to learn lessons from the mistakes of past and current research in order to be prepared to conduct your own research. As you will soon learn, developing your research project is an evolutionary process, and research itself is a continuously changing and evolving field.

Careful conceptual design and planning are crucial for conducting a reproducible, compelling, and ethically responsible research study. In this chapter, we will discuss what should be the first step of any research project, that is, how to develop your own research question. The basic process is to select a topic of interest, identify a research problem within this area of interest, formulate a research question, and finally state the overall research objectives (i.e., the specific aims that define what you want to accomplish).

You will learn how to define your research question, starting from broad interests and then narrowing these down to your primary research question. We will address the key elements you will need to define for your research question: the study population, the intervention (x, independent variable[s]), and the outcome(s) (y, dependent variable[s]). Later chapters in this volume will discuss popular study designs and elements such as covariates, confounders, and effect modifiers (interaction) that will help you to further delineate your research question and your data analysis plan.

Although this chapter is not a grant-writing tutorial, most of what you will learn here has very important implications for writing a grant proposal. In fact, the most important part of a grant proposal is the "specific aims" page, where you state your research question, hypotheses, and objectives.

HOW TO SELECT A RESEARCH QUESTION

What Is a Research Question?

A research question is an inquiry about an unanswered scientific problem. The purpose of your research project is to find the answer to this particular research question. Defining a research question can be the most difficult task during the design of your study. Nevertheless, it is fundamental to start with the research question, as it is strongly associated with the study design and predetermines all the subsequent steps in the planning and analysis of the research study.

What Is the Importance of a Research Question?

Defining the research question is instrumental for the success of your study. It determines the study population, outcome, intervention, and statistical analysis of the research study, and therefore the scope of the entire project.

A novice researcher will often jump to methodology and design, eager to collect data and analyze them. It is always tempting to try out a new or "fancy" method (e.g., "Let's test this new proteomic biomarker in a pilot study!" or "With this Luminex assay we can test 20 cytokines simultaneously in our patient serum!"), but this mistake all too often makes the research project a "fishing expedition," with the unfortunate outcome that a researcher has invested hours of work and has obtained reams of data, only to find herself at an impasse, and never figuring out what to do with all the information collected. Although it is not wrong to plan an exploratory study (or a hypothesis-generating study), such study has a high risk of not yielding any useful information; thus all the effort to have the study performed will be lost. When planning an exploratory or pilot study (with no defined research question), the investigator must understand the goals and the risks (for additional discussion on pilot studies, see Lancaster et al. 2002).

It is important to first establish a concept for your research. You must have a preset idea or a working hypothesis in order to be able to understand the data you will generate. Otherwise, you will not be able to differentiate whether your data were obtained by chance, by mistake, or if they actually reflect a true finding. Also, have in mind ahead of time how you would like to present your study at a conference, in a manuscript, or in a grant proposal. You should be able to present your research to your audience in a well-designed manner that reflects a logical approach and appropriate reasoning.

A good research question leads to useful findings that may have a significant impact on clinical practice and health care, regardless of whether the results are positive or negative. It also gives rise to the next generation of research questions. Therefore, taking enough time to develop the research question is essential.

Where Do Research Questions Come From?

How do we find research questions? As a clinical research scientist, your motivation to conduct a study might be driven by a perceived knowledge gap, the urge to deepen your understanding in a certain phenomenon, or perhaps to clarify contradictory existing findings. Maybe your bench research implies that your findings warrant translation into a study involving patients in a clinical setting. Maybe your clinical work

experience gives you the impression that a new intervention would be more effective for your patients compared to standard treatment. For example, your results could lead you to ask, "Does this drug really prolong life in patients with breast cancer?" or, "Does this procedure really decrease pain in patients with chronic arthritis?"

Once you have identified a problem in the area you want to study, you can refine your idea into a research question by gaining a firm grasp of "what is known" and "what is unknown." To better understand the research problem, you should learn as much as you can about the background pertaining to the topic of interest and specify the gap between current understanding and unsolved problems. As an early step, you should consult the literature, using tools such as MEDLINE or EMBASE, to gauge the current level of knowledge relevant to your potential research question. This is essential in order to avoid spending unnecessary time and effort on questions that have already been solved by other investigators. Meta-analyses and systematic reviews are especially useful to understand the combined level of evidence from a large number of studies and to obtain an overview of clinical trials associated with your questions. You should also pay attention to unpublished results and the progress of important studies whose results are not yet published. It is important to realize that there likely are negative results produced but never published. You can inspect funnel plots obtained from meta-analyses or generated from your own research (see Chapter 14 in this volume for more details) to estimate if there has been publication bias toward positive studies. Also, be aware that clinical trials with aims similar to those of your study might still be ongoing. To find this information, you can check the public registration of trials using sites such as clinicaltrials.gov.

HOW TO DEVELOP THE RESEARCH QUESTION: NARROW DOWN THE QUESTION

Once you have selected your research topic, you need to develop it into a more specific question. The first step in refining a research question is to narrow down a *broad research topic* into a specific description (*narrow research question*) that covers the four points of *importance, feasibility, answerability*, and *ethicality.*

Importance: Interesting, Novel, and Relevant

Your research can be descriptive, exploratory, or experimental. The purpose of your research can be for diagnostic or treatment purposes, or to discover or elucidate a certain mechanism. The point you will always have to consider when making a plan for your study, however, is how to justify your research proposal. Does your research question have scientific relevance? Can you answer the "so what" question? You need to describe the importance of your research question with careful consideration of the following elements:

- *The disease (condition or problem)*: Novelty, unmet need, or urgency are important. What is the prevalence of this disease/condition? Is there a pressing need for further discoveries regarding this topic because of well-established negative prognoses (e.g., HIV, pancreatic cancer, or Alzheimer's disease)? Are existing treatment

options limited, too complex or costly, or otherwise not satisfactory (e.g., limb replacement, face transplantation)? Does the research topic reflect a major problem in terms of health policy, medical, social, and/or economic aspects (e.g., smoking, hypertension, or obesity)?

- *The intervention*: Is it a new drug, procedure, technology, or medical device (e.g., stem-cell derived pacemaker or artificial heart)? Does it concern an existing drug approved by the Food and Drug Administration (FDA) for a different indication (e.g., is Rituximab, a drug normally indicated for malignant lymphoma, effective for systemic lupus erythematosus or rheumatoid arthritis)? Is there new evidence for application of an existing intervention in a different population (e.g., is Palivizumab also effective in immunodeficiency infants, not only in premature infants to prevent respiratory syncytial virus)? Have recent findings supported the testing of a new intervention in a particular condition (e.g., is a β-blocker effective in preventing cardiovascular events in patients with chronic renal failure)? Even a research question regarding a standard of care intervention can be valuable if in the end it can improve the effectiveness of clinical practice.

Feasibility

In short, be realistic: novel research tends to jump right away into very ambitious projects. You should carefully prove the feasibility of your research idea to prevent wasting precious resources such as time and money:

- *Patients*: Can you recruit the required number of subjects? Do you think your recruitment goal is realistic? Rare diseases such as Pompe or Fabry's disease will pose a challenge in obtaining a sufficient sample size. Even common diseases, depending on your inclusion criteria and regimen of intervention, may be difficult to recruit. Does your hospital have enough patients? If not, you may have to consider a *multicenter* study. What about protocol *adherence* and *dropouts*? Do you expect significant deviations from the protocol? Do you need to adjust your sample size accordingly?
- *Technical expertise*: Are there any established measurements or diagnostic tools for your study? Can the outcome be measured? Is there any established diagnostic tool? Do we have any standard techniques for using the device (e.g., guidelines for echocardiographic diagnosis for congenital heart disease)? Is there a defined optimal dose? Can you operate the device, or can the skill be learned appropriately (e.g., training manual for transcutaneous atrial valve replacement)? *A pilot study* or *small preliminary study* can be helpful at this stage to help answer these preliminary questions.
- *Time*: Do you have the required time to recruit your patients? Is it possible to *follow up* with patients for the entire time of the proposed study period (e.g., can you follow preterm infant development at 3, 6, and 9 years of age)? When do you need to have your results in order to apply for your next grant?
- *Funding*: Does you budget allow for the scope of your study? Are there any research grants you can apply for? Do the funding groups' interests align with those of your study? How realistic are your chances of obtaining the required funding? If there are available funds, how do you apply for the grant?

- *Team*: How about your research environment? Do your *mentors* and *colleagues* share your interests? What kind of specialists do you need to invite for your research? Do you have the staff to support your project (technicians, nurses, administrators, etc.)?

Answerability

New knowledge can only originate from questions that are answerable. A broad research problem is still a theoretical idea, and even if it is important and feasible, it still needs to be further specified. You should carefully investigate your research idea and consider the following:

- Precisely define what is known or not known and identify what area your research will address. The research question should demonstrate an understanding of the *biology, physiology*, and *epidemiology* relevant to your research topic. For example, you may want to investigate the prevalence and incidence of stroke after catheterization and its prognosis before you begin research on the efficacy of a new anticoagulant for patients who received catheter procedures. Again, you may need to conduct a *literature review* in order to clarify what is already known. Conducting *surveys* (*interviews* or *questionnaires*) initially could also be useful to understand the current status of your issues (e.g., how many patients a year are diagnosed with stroke after catherization in your hospital? What kind of anticoagulant is already being used for the patients? How old are the patients? How about the duration of cauterization techniques? etc.).
- The standard treatment should be well known before testing a new treatment. Are there any established treatments in your research field? Could your new treatment potentially replace the standard treatment or be complementary to the current treatment of choice? *Guidelines* can be helpful for discussion (e.g., American College of Cardiology/American Heart Association guidelines for anticoagulant therapy). Without knowing the current practice, your new treatment may never find its clinical relevance.
- We also need information about clinical issues for *diagnostic tests and interventions*. Are you familiar with the diagnoses and treatment of this disease (e.g., computerized tomography or magnetic resonance imaging to rule out stroke after catherization)? Do you know the *current guidelines*?

Ethical Aspects

Ethical issues should be discussed before conducting research. Is the subject of your research a controversial topic? The possible ethical issues will often depend largely on whether the study population is considered vulnerable (e.g., children, pregnant women, etc.; see Chapter 1) [1]. You must always determine the possible *risks and benefits* of your study intervention [1].

Finally, you may want to ask for *expert opinions* about whether your research *question is answerable* and relevant (no matter how strong your personal feelings may be about the relevance). To this end, a presentation of your idea or preliminary results at a study meeting early on in the project development can help refine your question.

HOW TO BUILD THE RESEARCH QUESTION

The next step of formulating a narrow research question is to focus on the primary interest (*primary question*): What is the most critical question for your research problem? You will define this primary question by addressing the key elements using the useful acronym PICOT (population, intervention, control, outcome, and time), while keeping in mind the importance, feasibility, answerability, and ethicality. Although PICOT is a useful framework, it does not cover all types of studies, especially some observational studies, for instance those investigating predictors of response (E [exposure] instead of I [intervention] is used for observational studies). But for an experimental study (e.g., a clinical trial), the PICOT framework is extremely useful to guide formulation of the research question.

Building the Research Question: PICOT

P (Population or Patient)

What is the *target population* of your research? The target population is the population of interest from which you want to draw conclusions and inferences. Do you want to study mice or rabbits? Adults or children? Nurses or doctors? What are the characteristics of the study subjects, and what are the given problems that should be considered? You may want to consider the pathophysiology (acute or chronic?) and the severity of the disease (severe end stage or early stage?), as well as factors such as geographical background and socioeconomic status.

Once you decide on the target population, you may select a sample as the *study population* for your study. The *study population* is a subset of the target population under investigation. However, it is important to remember that the *study population* is not always a perfect representation of the target population, even when sampled at random. Thus, defining the *study population* by the inclusion and exclusion criteria is a critical step (see Chapter 3).

Since only in rare cases will you be able to study every patient of interest, you will have to identify and select whom from the target population you want to study. This is referred to as the study sample. To do this requires choosing a method of selection or recruitment (see Chapter 7).

A specific study sample defined by restricted criteria will have a reduced number of covariates and will be more homogeneous, therefore increasing the chance of higher internal validity for your study. This also typically allows for the study to be smaller and potentially less expensive. In contrast, a restricted population might make it more difficult to recruit a sufficient number of subjects. On the other hand, recruitment can be easier if you define a broad population, which also increases the generalizability of your study results. However, a broad population can make the study larger and more expensive [2].

I (Intervention)

The *I* of the acronym usually refers to "intervention." However, a more general and therefore preferable term would be "independent variable." The independent variable

is the explanatory variable of primary interest, also declared as *x* in the statistical analysis. The independent variable can be an intervention (e.g., a drug or a specific drug dose), a prognostic factor, or a diagnostic test. *I* can also be the exposure in an observational study. In an experimental study, *I* is referred to as the fixed variable (controlled by the investigator), whereas in an observational study, *I* refers to an exposure that occurs outside of the experimenter's control.

The independent variable precedes the outcome in time or in its causal path, and thus it "drives" the outcome in a cause-effect relationship.

C (Control)

What comparison or control is being considered? This is an important component when comparing the efficacy of two interventions. The new treatment should be superior to the *placebo*, when there is no standard treatment available. Placebo is a simulated treatment that has no pharmaceutical effects and is used to mask the recipients to potential expectation biases associated with participating in clinical trials. On the other hand, *active controls* could be used when an established treatment exists and the efficacy of the new intervention should be examined at least within the context of non-inferiority to the standard treatment. Also the control could be baseline in a one-group study.

O (Outcomes)

O is the dependent variable, or the outcome variable of primary interest; in the statistical analysis, it is also referred to as *y*. The outcome of interest is a random variable and can be a clinical (e.g., death) or a surrogate endpoint (e.g., hormone level, bone density, antibody titer). Selection of the primary outcome depends on several considerations: What can you measure in a timely and efficient manner? Which measurement will be relevant to understand the effectiveness of the new intervention? What is routinely accepted and established within the clinical community? We will discuss the outcome variable in more detail later in the chapter.

T (Time)

Time is sometimes added as another criterion and often refers to the follow-up time necessary to assess the outcome or the time necessary to recruit the study sample. Rather than viewing time as a separate aspect, it is usually best to consider time in context with the other PICOT criteria.

What Is the Primary Interest in your Research?

Once you have selected your study population, as well as the dependent and independent variables, you are ready to formulate your primary research question, the major specific aim, and a hypothesis. Even if you have several different ideas regarding your research problem, you still need to clearly define what the most important question of your research is. This is called your *primary question*. A research project may also contain additional secondary questions.

The primary question is the most relevant question of your research that should be driven by the hypothesis. Usually only one primary question should be defined at the beginning of the study, and it must be stated explicitly upfront [3]. This question is relevant for your sample size calculation (and in turn, for the power of your study—see Chapter 11).

The specific aim is a statement of what you are proposing to do in your research project.

The primary hypothesis states your anticipated results by describing how the independent variable will affect the dependent variable. Your hypothesis cannot be just speculation, but rather it must be grounded on the research you have performed and must have a reasonable chance of being proven true.

We can define more than one question for a study, but aside from the primary question, all others associated with your research are treated as *secondary questions.* Secondary questions may help to clarify the primary question and may add some information to the research study. What potential problems do we encounter with secondary questions? Usually, they are not sufficiently powered to be answered because the sample size is determined based on th*e primary question.* Also, type I errors (i.e., false positives) may occur due to multiple comparisons if not adjusted for by the proper statistical analysis. Therefore, findings from secondary questions should be considered exploratory and hypothesis generating in nature, with new confirmatory studies needed to further support the results.

An *ancillary study* is a sub-study built into the main study design. Previous evidence may convince you of the need to test a hypothesis within a sub-group ancillary to the main population of interest (e.g., females, smokers). While this kind of study enables you to perform a detailed analysis of the subpopulation, there are limitations on the generalizability of an ancillary study since the population is usually more restricted (see Further Readings, Examples of Ancillary Studies).

Variables

It is important to understand thoroughly the study variables when formulating the study question. Here we will discuss some of the important concepts regarding the variables, which will be discussed in more detail in Chapter 8.

We have already learned that the *dependent variable* is the outcome, and *the independent* variable is the intervention. For study design purposes, it is important to also discuss how the outcome variables are measured. A good measurement requires reliability (precision), validity ("trueness"), and responsiveness to change. Reliability refers to how consistent the measurement is if it is repeated. Validity of a measurement refers to the degree to which it measures what it is supposed to measure. Responsiveness of a measurement means that it can detect differences that are proportional to the change of what is being measured with clinical meaningfulness and statistical significance.

Covariates are independent variables of secondary interest that may influence the relationship between the independent and dependent variables. Age, race, and gender are well-known examples. Since covariates can affect the study results, it is critical to control or adjust for them. Covariates can be controlled for by both planning (inclusion and exclusion criteria, placebo and blinding, sampling and randomization, etc.)

and analytical methods (e.g., covariate adjustment [see Chapter 13], and propensity scores [see Chapter 17]).

- *Continuous (ratio and interval scale), discrete, ordinal, nominal (categorical, binary) variables: Continuous data* represent all numbers (including fractions of numbers, floating point data) and are the common type of raw data. Discrete data are full numbers (i.e., integer data type; e.g., number of hospitalizations). Ordinal data are ordered categories (e.g., mild, moderate, severe). Nominal data can be either categorical (e.g., race) or dichotomous/binary (e.g., gender). Compared to other variables, continuous variables have more power, which is the ability of the study to detect an effect (e.g., differences between study groups) when it is truly present, but they don't always reflect clinical meaningfulness and therefore make interpretation more difficult. *Ordinal and nominal data* may better reflect the clinical significance (e.g., dead or alive, relapse or no relapse, stage 1 = localized carcinoma, etc.). However, ordinal and categorical data typically have less power, and important information may be lost (e.g., if an IQ less than 70 is categorized as developmental delay in infants, IQs of 50, 58, and 69 will all fall into the same category, while an IQ of 70 or more is considered to be normal development, although the difference is just 1 point). This approach is called categorization of continuous data, where a certain clinically meaningful threshold is set to make it easier to quickly assess study results. It is important to note that some authors differentiate between continuous and discrete variables by defining the former as having a quantitative characteristic and the latter as having a qualitative characteristic. This is a somewhat problematic classification, especially when it comes to ordinal data.
- *Single* and *multiple variables*: Having a *single variable* is simpler, as it is easier for clinical interpretation. Multiple valuables are efficient because we can evaluate many variables within a single trial, but these can be difficult to disentangle and interpret. *Composite endpoints* are combined multiple variables and are also sometimes used. Because each clinical outcome may separately require a long duration and a large sample size, combining many possible outcomes increases overall efficiency and enables one to reduce sample size requirements and to capture the overall impact of therapeutic interventions. Common examples include MACE (major adverse cardiac events) and TVF (target vessel failure: myocardial infarction in target vessel, target vessel reconstruction, cardiac death, etc.). Interpretation of the results has to proceed with caution, however (see section on case-specific questions) [9].
- *Surrogate variables (endpoints) and clinical variables (endpoints): Clinical variables* directly assess the effect of therapeutic interventions on patient function and survival, which is the ultimate goal of a clinical trial. Clinical variables may include mortality, events (e.g., myocardial infarction, stroke), and occurrence of disease (e.g., HIV). A clinical endpoint is the most definitive outcome to assess the efficacy of an intervention. Thus, clinical endpoints are preferably used in clinical research. However, it is not always feasible to use clinical outcomes in trials. The evaluation of clinical outcomes presents some methodological problems since they require long-term follow-up (with problems of adherence, dropouts, competing risks, requiring larger sample sizes) and can make a trial more costly. At the same time, the clinical endpoint may be difficult to observe. For this reason, clinical scientists often use alternative outcomes to substitute for the clinical outcomes. So-called *surrogate*

endpoints are a more practical measure to reflect the benefit of a new treatment. Surrogate endpoints (e.g., cholesterol levels, blood sugar, blood pressure, viral load) are defined based on the understanding of the mechanism of a disease that suggests a clear relationship between a marker and a clinical outcome [8]. Also, a biological rationale provided by epidemiological data, other clinical trials, or animal data should be previously demonstrated. A surrogate is frequently a continuous variable that can be measured early and repeatedly and therefore requires shorter follow-up time, smaller sample size, and reduced costs for conducting a trial. Surrogate endpoints are often used to accelerate the process of new drug development and early stages of development, such as in phase 2 [10]. As a word of caution, too much reliance on surrogate endpoints alone can be misleading if the results are not interpreted with regard to validation, measurability, and reproducibility (see Further Reading) [4].

HOW TO EXPRESS A RESEARCH QUESTION

Hypothesis

Once a narrow research question is defined, you should clearly specify a hypothesis in the study protocol. A *hypothesis* is a statement about the expected results that predicts the effect of the independent on the dependent variable. A research hypothesis is essential to frame the experimental and statistical plan (statistics will be discussed in Unit II of this volume) and is also important to support the aim of the study in a scientific manuscript.

Types of Research Questions

To refine the research question and form the research hypothesis, we will discuss three types of research questions that investigate group differences, correlations, or descriptive measures. This classification is particularly important in discussing which statistical analysis is appropriate for your research question [5].

- *Basic/complex difference (group comparison) questions*: Samples split into groups by levels associated with the independent variable are compared by considering whether there is a difference in the dependent variable. If you have only one independent variable, the question is classified as a *basic difference question* (e.g., drug A will reduce time to primary closure in a 5-mm punch biopsy vs. placebo) and you would rely on a t-test or one-way analysis of variance (ANOVA) for the analysis. If you have two or more independent variables (e.g., drugs A and B led to a 15-mg/dl reduction in LDL cholesterol versus placebo, but there was no reduction with only drug A), this then becomes a *complex difference question* and is analyzed by other statistical methods, such as a factorial ANOVA.
- *Basic/complex associational (relational/correlation) questions*: The independent variable is correlated with the dependent variable. If there is only one dependent variable and one independent variable (e.g., is there a relationship between weight and natriuretic peptide levels?), it is called a *basic associational question*, and in this

situation, a correlation analysis is used. If there is more than one independent variable associated with one dependent variable (e.g., smoking and drinking alcohol are associated with lung cancer), it is called a *complex associational question*, and multiple regression is used for statistical analysis.

- *Basic/complex descriptive question*: The data are described and summarized using measures of central tendency (means, median, and mode), variability, and percentage (prevalence, frequency). If there is only one variable, it is called a *basic descriptive question* (e.g., how much MRSA isolates occur after the 15th day of hospitalization?); for more than one variable, a classification of *basic/complex descriptive question* is used.

Where Should You State Your Research Question?

Finally, where should you state your hypothesis? You may be writing for a research grant, research protocol, or manuscript. Usually, research questions should be stated in the introduction, immediately following the justification ("so what") section. Research questions should be clearly stated in the form of a hypothesis, such as "We hypothesize that in this particular population (P), the new intervention (I) will improve the outcome (O) more than the standard of care (C)."

A Research Question Should Be Developed over Time

It is important that the investigator spend a good amount of time developing his or her study question. During this process, everything we discussed in this chapter needs to be reviewed and the research question then needs to be refined as this process takes place. A good planning, starting with the research question, is one of the key components for a study's success.

Related Topics for Choosing the Research Question

Selecting the Appropriate Control in Surgical Studies or Other Challenging Situations

Let's think about various situations. Can we use placebo (sham) or another procedure as a control in a surgical trial? What exactly can be considered a placebo in surgical studies? How do we control for a placebo effect in surgical procedures?

Placebos can be used for the control group in clinical studies in comparison to a new agent if no standard of care is available. In order to fully assess the placebo effect in the control arm, participants have to be blinded. The control group could either have no surgery at all or undergo a "sham" procedure, but both options might be unethical depending on the given patient population [6]. In surgical studies, the control group usually receives the "traditional" procedure. In all cases, blinding might be very challenging and even impossible on certain levels (e.g., the surgeon performing the procedure). What about acupuncture? What would you consider a good control? What about cosmetic procedures?

Using Adverse Events as the Primary Research Question

Important questions concerning adverse effects can be answered in a clinical trial. However, as the typical clinical trial is performed in a controlled setting, the information regarding adverse effects is not always generalizable to the real-world setting. Thus, the clinical translation of the results needs careful consideration when carrying out a safety-focused study. The adverse reports from phase 4 (post-marketing marketing surveillance) are considered more generalizable information in drug development, although minimum safety data from phase 1 are required to proceed to subsequent study phases.

Also, it might not be easy to formulate a specific research question regarding adverse effects, as they might not be fully known in the early stage of drug development. This will also make it difficult to power the study properly (e.g., how many patients do we need to examine to show the statistically meaningful difference?).

When the Research Question Leads to Other Research Questions

Medical history is filled with interesting stories about research questions. And sometimes, it is not the intended hypothesis to be proven that yields a big discovery. For example, Sildenafil (Viagra) was initially developed by Pfizer for the treatment of cardiovascular conditions. Although clinical trials showed Sildenafil to have only little effect on the primary outcomes, it was quickly realized that an unexpected but marked "side effect" occurred in men. Careful investigation of clinical and pharmacological data generated the new research question, "Can Sildenafil improve erectile dysfunction?" This question was then answered in clinical trials with nearly 5,000 patients, which led to Sildenafil's FDA approval in 1998 as the first oral treatment for male erectile dysfunction [7]. The investigator must be attentive to novel hypothesis that can be learned from a negative study.

CASE STUDY: FINDING THE RESEARCH QUESTION

André Brunoni and Felipe Fregni

Dr. L. Heart is a scientist working on cardiovascular diseases in a large, busy emergency room of a tertiary hospital specialized in acute coronarian syndromes. While searching PUBMED, she found an interesting article on a new drug—which animal studies have demonstrated to be a powerful anti-thrombotic agent—showing its safety in healthy volunteers. She then feels that it would be the right time to perform a phase II trial, testing this new drug in patients presenting myocardial infarction (MI). She sees this as her big career breakthrough. However, when Dr. Heart starts writing a study proposal for the internal review board (ethics committee), she asks herself, "What is my research question?"[1]

Introduction

Defining the research question is, perhaps, the most important part of the planning of a research study. That is because the wrong question will eventually lead to a poor study design and therefore all the results will be useless; on the other hand, choosing an elegant, simple question will probably lead to a good study that will be meaningful to the scientific community, even if the results are negative. In fact, the best research question is one that, regardless of the results (negative or positive), produces interesting findings. In addition, a study should be designed with only one main question in mind.

However, choosing the most appropriate question is not always easy, as such a question might not be feasible to be answered. For instance, when researching acute MI, the most important question would be whether or not a new drug decreases mortality. However, for economic and ethical reasons, such an approach can only be considered when previous studies have already *suggested* that the new drug is a potential candidate. Therefore, the investigator needs to deal with the important issue of feasibility versus clinical relevance. Dr. Heart soon realized that her task would not be an easy one, and also that this task may take some time; she kept thinking about one of the citations in an article she recently read: "One-third of a trial's time between the germ of your idea and its publication in the *New England Journal of Medicine* should be spent fighting about the research question."[2]

"So What?" Test for the Research Question

Dr. Heart knows that an important test for the research question is to ask, "So what?" In other words, does the research question address an important issue? She knows,

[1] Dr. André Brunoni and Professor Felipe Fregni prepared this case. Course cases are developed solely as the basis for class discussion. The situation in this case is fictional. Cases are not intended to serve as endorsements or sources of primary data.

[2] Riva JJ, Malik KM, Burnie SJ, Endicott AR, Busse JW. What is your research question? An introduction to the PICOT format for clinicians. *J Can Chiropr Assoc*. 2012 Sep; 56(3):167–71.

for example, that the main agency funding in the United States, the NIH (National Institutes of Health), considers significance and innovation as important factors to fund grant applications. Dr. Heart also remembers something that her mentor used to tell her at the beginning of her career: "A house built on a weak foundation will not stand." She knows that even if she has the most refined design and uses the optimal statistical tests, her research will be of very little interest or utility if it does not advance the field. But regarding this point, she is confident that her research will have a significant impact in the field.

Next Step for the Research Question: How to Measure the Efficacy of the Intervention

Dr. L. Heart is in a privileged position. She works in a busy hospital that receives a significant amount of acute cardiovascular patients. She also has received huge departmental support for her research, meaning that she can run a wide range of blood exams to measure specific biological markers related to death in myocardial infarction. Finally, she has a PhD student who is a psychologist working with quality of life post-MI. Therefore, she asks herself whether she should rely on biological markers, on the assessment of quality of life, or if she should go to a more robust outcome to prove the efficacy of the new drug. She knows that this is one of the most critical decisions she has to make. It was a Friday afternoon. She had just packed up her laptop and the articles she was reading, knowing that she will have to make a decision by the end of the weekend.

Dr. Heart is facing a common problem: What outcome should be used in a research study? This needs to be defined for the research question. She knows that there are several options. For instance, the outcome might be mortality, new MI, days admitted to the emergency room, quality of life, specific effect of disease such as angina, a laboratory measure (cholesterol levels), or the cost of the intervention. Also, she might use continuous or categorical outcomes. For instance, if she is measuring angina, she might measure the number of days with angina (continuous outcome) or dichotomize the number of angina days in two categories (less than 100 days with angina vs. more or equal to 100 days with angina). She then lays out her options:

- *Use of clinical outcomes (such as mortality or new myocardial infarction)*: She knows that by using this outcome, her results would be easily accepted by her colleagues; however, using these outcomes will increase the trial duration and costs.
- *Use of surrogates (for instance, laboratorial measurements)*: One attractive alternative for her is to use some biomarkers or radiological exams (such as a catheterism). She knows a colleague in the infectious disease field who only uses CD4 for HIV trials as the main outcome. This would increase the trial feasibility. However, she is concerned that her biomarkers might not really represent disease progression.
- *Use of quality of life scales*: This might be an intermediate solution for her. However, she is still concerned with the interpretation of the results if she decides to use quality of life scales.

More on the Response Variable: *Categorical or Continuous?*

Even before making the final decision, Dr. Heart needs to decide whether she will use a continuous or categorical variable. She wishes now that she knew the basic concepts of statistics. However, she calls a colleague, who explains to her the main issue of categorical versus continuous outcomes—in summary, the issue is the trade-off of power versus clinical significance.

A categorical outcome usually has two categories (e.g., a yes/no answer), while a continuous outcome can express any value. A categorical approach might be more robust than a continuous one, and it also has more clinical significance, but it also decreases the power of the study due to the use of less information.[3] She is now at the crossroad of feasibility versus clinical significance.

Choosing the Study Population

Now that Dr. Heart has gone through the difficult decision of finding the best outcome measure, she needs to define the target population—that is, in which patients is she going to test the new drug? Her first idea is to select only patients who have a high probability of dying—for instance, males who smoke, are older than 75 years, with insulin-dependent diabetes and hypercholesterolemia. "Then," she thinks, "it will be easier to prove that the new drug is useful regardless of the population I study. But does that really sound like a good idea?"

The next step is to define the target population. Dr. Heart is inclined to restrict the study population, as she knows that this drug might be effective to a particular population of patients and therefore this increases her chances of getting a good result. In addition, she does remember from her statistical courses that this would imply a smaller variability and therefore she would gain power (power is an important currency in research, as it makes the study more efficient, decreasing costs and time to complete the study). On the other hand, she is concerned that she might put all her efforts in one basket—this is a risky approach, as this specific population might not respond, and she knows that broadening the population also has some advantages, for instance, the results would be more generalizable and it would be easier to recruit patients. But this would also increase the costs of the study.

But How about Other Ideas?

After a weekend of reflection, Dr. Heart called the staff for a team meeting and proudly explained the scenario and stated her initial thoughts. The staff was very eager to start a new study, and they made several suggestions: "We should also use echocardiography to assess the outcome!"; "Why don't we perform a genotypic analysis on these patients?"; "We need to follow them until one year after discharge." She started to become anxious again. What should she do with these additional suggestions? They all seem to be good ideas.

[3] These concepts will be discussed in details in Unit II of this volume.

When designing a clinical trial, researchers expose a number of subjects to a new intervention. Therefore, they want to extract as much data as possible from studies. On the other hand, it might not be possible to ask all of the questions, since this will increase the study's duration, costs, and personnel. Also, researchers should be aware that all the other outcomes assessed will be *exploratory* (i.e., their usefulness remains in suggesting possible associations and future studies) because studies are designed to answer a primary question only—and, as a principle of statistics, there is a 5% probability of observing a positive result just by chance (if you perform 20 tests, for instance, one of them will be positive just by chance!). But Dr. Heart knows that she can test additional hypotheses as secondary questions. She knows that there is another issue to go through: the issue of primary versus secondary questions.

Defining Her Hypothesis

After going through this long process, Dr. Heart is getting close to her research question. But now she needs to define the study hypotheses. In other words, what is her educated guess regarding the study outcome?

An important step when formulating a research question is to define the hypothesis of the study. This is important in terms of designing the analysis plan, as well as estimating the study sample size. Usually, researchers come up with study hypotheses after reviewing the literature and preliminary data. Dr. Heart can choose between a simple and a complex hypothesis. In the first case, her hypothesis would only have one dependent variable (i.e., the response variable) and one independent variable (e.g., the intervention). Complex hypotheses have more than one independent and/or dependent variable and might not be easy to use in planning the data analysis.

By the end of the day, Dr. Heart was overwhelmed with the first steps to put this study together. Although she is confident that this study might be her breakthrough and she needs to get her tenure track position at the institution where she works, she also knows she has only one chance and must be very careful at this stage. After wrestling with her thoughts, she finished her espresso and walked back to her office, confident that she knew what to do.

CASE DISCUSSION

Dr. Heart is a busy and ambitious clinical scientist and wants to establish herself within the academic ranks of her hospital. She has some background in statistics but seems to be quite inexperienced in conducting clinical research. She is looking for an idea to write up a research proposal and rightly conducts a literature research in her field of expertise, cardiovascular diseases. She finds an interesting article about a compound that has been demonstrated to be effective in an animal model and safe in healthy volunteers (results of a phase I trial). She now plans to conduct a phase II trial, but struggles to come up with a study design. The most vexing problem for her is formulating the research question.

Dr. Heart then reviews and debates aspects that have to be considered when delineating a research question. The main points she ponders include the following: determining the outcome with regard to feasibility (mainly concerning the time of