Scott Evans¹ / Daniel B. Rubin² / John H. Powers^{3,4} / Dean Follmann⁵

Analysis Populations in Anti-Infective Clinical Trials: Whom to Analyze?

¹ Department of Epidemiology and Biostatistics, Milken Institute School of Public Health of The George Washington University, Innovations in Design, Education, and Analysis (IDEA) Committee, The Biostatistics Center, 6610 Executive Blvd., Suite 750 Rockville, United States of America, E-mail: SCID@degruyter.com

² Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

³ Clinical Research Directorate/Clinical Monitoring Research Program, Leidos Biomedical Research, Inc., NCI Campus at Frederick, Frederick, MD 21702, USA

⁴ George Washington University School of Medicine, Washington DC, USA

⁵ National Institute of Allergy and Infectious Diseases of the National Institutes of Health, Rockville, United States of America

Abstract:

Investigators can choose to analyze different patient populations in clinical trials. The different analysis populations answer different types of research questions, estimate different quantities, and evaluate the robustness of the trial results. Various analysis populations have different strengths and weaknesses depending on the type of question being addressed and the potential for bias from the selection of various groups of trial participants. We discuss analysis populations in the context of anti-infective clinical trials.

Keywords: analysis populations, anti-infectives, clinical trials, intent-to-treat

DOI: 10.1515/scid-2017-0002

Received: July 5, 2017; **Revised:** February 3, 2018; **Accepted:** June 20, 2018

1 Introduction

A specified number of participants are enrolled in randomized controlled trials (RCTs) but investigators also specify which patients will be included in various analysis populations once the trial is concluded. While the intention-to-treat (ITT) analysis population is usually the most clinically relevant and represents results based on how patients present in real-world settings, investigators can choose to analyze different patient populations in clinical trials. The different analysis populations answer different types of research questions, estimate different quantities (estimands), and evaluate the robustness of the trial results. Various analysis populations have different strengths and weaknesses depending on the type of question being addressed and the potential for bias from the selection of various groups of trial participants.

Commonly in clinical trials evaluating interventions to treat bacterial infections, a provisional syndromic diagnosis (e. g. pneumonia, urinary tract infection) is made based upon clinical presentation (e. g. signs and symptoms) followed by confirmation from other data (radiology, antibody tests, cultures and other laboratory tests) possibly including isolation of the suspected offending pathogen(s). Optimally, pathogen(s) can be identified prior to randomization. However, often syndromic diagnoses are inaccurate and rapid accurate diagnostics are unavailable (e. g. as in diagnosis and confirmation of ventilator associated pneumonia), and administration of interventions may be needed prior to confirmation of disease, pathogen identification or drug susceptibility information, due to the presumed impact of early administration of effective interventions on morbidity and mortality in serious and life threatening diseases. Microbiological specimens are obtained before randomization and treatment often is initiated empirically based on presumptive diagnosis and presence of pathogen(s) with microbiological confirmation (or not being obtained after randomization).

Researchers designing, analyzing, and interpreting clinical trials that evaluate interventions to treat these diseases must decide which trial participants to analyze in this setting of delayed knowledge of offending pathogen(s). Frequently several analysis populations are defined prior to study initiation in these clinical trials.

Scott Evans is the corresponding author.

© 2018 Walter de Gruyter GmbH, Berlin/Boston.

This content is free.

2 Definitions

Figure 1 shows a flow diagram defining analysis populations for a typical clinical trial (Gillings and Koch 1991). Potential trial participants first are screened for eligibility. Screening failures are excluded from further follow-up prior to randomization. Exclusion of these patients does not affect the internal validity of the trial but may affect the generalizability or applicability of the results to specific non-studied populations. Participants with clinical disease suspected to be caused by targeted pathogen(s) and meeting other inclusion criteria are randomized. All randomized participants are *assigned* to a study intervention forming an analysis population called the *intent-to-treat or intention-to-treat population (ITT)*. This analysis population takes its name from the fact that patients are analyzed according to their randomized assignment group regardless of whether or not they actually receive the assigned intervention or whether they follow the protocol as specified. Therefore the “intention” to treat, whether or not the participants received the assigned intervention, is the basis for inclusion in this group. Since all randomized participants will be analyzed, the ITT principle states that all randomized study participants should be followed to study completion regardless of treatment status or adherence. Requiring analyses of study participants that did not take their assigned therapy is counterintuitive in many settings to some and thus a difficult concept to understand. Analyses of the ITT population can be considered an evaluation of the treatment “strategy” rather than an evaluation of treatment under ideal conditions (e. g. when participants adhere to therapy) and may be more reflective of real-world conditions.

Figure 1: Flow diagram of analysis populations from a typical clinical trial.

A modified ITT (mITT) population is often defined excluding those trial participants in the ITT population that *did not receive* the intended study interventions despite being assigned to an intervention. Usually the difference in numbers of participants included between ITT and mITT is negligible. If the trial is blinded then the difference between ITT and mITT is likely trivial. But if the study is open-label (not blinded or unblinded) then there is the possibility that participants will opt not to participate if they are assigned an intervention different from which they had hoped, potentially causing bias.

Further exclusions of trial participants may occur if participants did not meet some entry criteria based on factors measured *before* randomization. For instance, in a pneumonia trial, chest x-rays measured before

randomization may come back after randomization as devoid of infiltrates, meaning that those participants may not have the clinical disease under study (e. g. have bronchitis instead of pneumonia). After treatment initiation, microbiological identification of the offending pathogen(s) measured before randomization becomes available. Participants with the targeted pathogen(s) become part of the microbiological mITT (micro-mITT) population. Exclusion based on pre-randomization factors maintains the benefits of randomization with respect to the expectation of balance of confounding factors. However, if the number of participants excluded is large, then imbalances may occur due to a resulting sample size for the mITT and micro-mITT populations that are too small for the benefits of randomization to be realized. Differences in the numbers of participants included in the mITT and micro-mITT analyses can differ considerably from the ITT analysis population.

An evaluable or per protocol (PP) population may be defined by further refining of the mITT and micro-mITT populations potentially based on factors that occur after randomization e. g. “major protocol violations”, including requiring a specified amount of drug and protocol adherence, completion of follow-up, lack of concomitant therapy, concentrations or exposure to study drugs in pharmacokinetic/pharmacodynamics analyses, and availability of important endpoint data. The definition of the PP population is not well-defined or unique, potentially varying between trials. The numbers of participants in the PP and ITT analysis populations can be considerable.

Prior research has shown a lack of consistency and clarity regarding how these populations are defined from trial to trial. Therefore, a first step is to define analysis populations clearly prior to the trial initiation, while matching these populations to the specific research questions under consideration (Gravel, Opatrny, and Shapiro 2007; Hollis and Campbell 1999). Regulatory guidances also have not been consistent regarding recommendations on primary analysis populations, e. g. ITT suggested for pneumonia, and micro-mITT for urinary tract infections (US Food and Drug Administration 2014, 2015). Greater clarity is needed on the choice of analysis populations and the reasoning for those choices in relation to the research questions posed and benefit:risk analyses.

3 Different Research Questions

Which populations should be the basis for the primary analysis for confirmatory conclusions and on which the sample size of the study is based prior to study initiation? Antimicrobials often are biologically active in vitro against specific pathogens. This leads to a desire by some researchers to analyze only trial participants with confirmation of targeted pathogens (i. e. micro-mITT). Other investigators believe that results should be analyzed only in those participants who follow the protocol exactly since that is how the intervention might “work” optimally. Thus, they conduct analyses only on participants with the targeted pathogens who adhered to the intervention and protocol (i. e. PP).

The selection of whom to analyze has consequences regarding the clinical utility/applicability/pragmatism for (i) empiric therapy choices, and (ii) confirmatory therapy choices, and (iii) the scientific and statistical integrity and of the results. A first step is recognition of the different definitions and important distinctions between specific scientific research questions and associated estimands addressed by analyses of the different populations. The research question of interest and how the results might be applied in the future determines the appropriate analysis population.

Consider the following research questions of potential interest when designing, conducting, analyzing, reporting and applying the results of clinical trials evaluating anti-infective therapies.

1. The clinical applicability for empiric therapy: How do the interventions compare with regard to the *strategy to treat clinical disease* as the patients present and are treated in clinical practice (regardless of the offending pathogens(s))?
2. The clinical applicability for confirmatory therapy: How do the interventions compare with regard to the *strategy to treat clinical disease caused by a specific pathogen(s) of interest*?
3. The questions related to biological activity of an intervention rather than efficacy in treating human disease, evaluating the potential for use if the intervention can be tolerated and adhered to: How do the interventions compare with respect to *biological activity vs. a specific pathogen(s)*?

These questions have different levels of clinical utility for different situations and are addressed utilizing different analysis populations (Table 1). Analyses that address these questions also have varying levels of scientific integrity due to the potential for bias.

Table 1: Recommended analysis populations for various research foci in clinical trials.

Research Focus	Primarily Useful to	Whom To Analyze	Clinical Utility/Pragmatism for Empiric Therapy	Clinical Utility/Pragmatism for Confirmatory Therapy	Preserves Integrity of Randomization
Evaluate strategy to treat clinical disease, i. e. empiric therapy in real-world setting	Today's clinicians and patients	ITT (or mITT if blinded and uniformly assessed and small difference in number of participants with ITT)** mmITT**	High	High	Yes***
Evaluate strategy to treat clinical disease caused by specific pathogen(s), i. e. confirmatory therapy*	Future clinicians and patients (i. e. with development of rapid point of care diagnostics)		Today: Variable (indirect) Future: possibly high with development of rapid diagnostics assuming affects do not change over time	High	Yes***
Understanding biological mechanisms of action, or evaluating potential for use if therapy can be tolerated/adhered to	Biologists, chemists	PP	Variable	Variable	No. Subject to the biases of observational studies

*Equivalent to "Evaluate strategy to treat clinical disease" when pathogen(s) can be identified prior to randomization and subsequent initiation of treatment".

**If a noninferiority trial then supplement ITT/mITT/micro-mITT analyses with PP analyses to preserve assay sensitivity.

***Integrity can be lost with missing data.(1) 1. Fleming TR. Addressing missing data in clinical trials. Ann Intern Med. 2011;154(2):113–7.

4 Different Strengths and Weaknesses

The most important evaluation for patients and clinicians from a practical perspective is the comparison of the interventions based on the strategy of treating clinical disease which reflects how patients present in practice (Gail 1985). This "clinical question" is addressed using the ITT analysis population. Since this question is evaluated using ITT, the protection that randomization provides from confounding and selection bias is preserved. This means the intervention groups have the expectation of balance with respect to all factors (measured or unmeasured, known or unknown) except for intervention assignment. It is this expectation of balance with respect to all potentially confounding variables that provides the foundation for valid statistical inference. If the prevalence of missing data is minimal, then causality related to the interventions can be confidently assessed without other assumptions such as testing to confirm the intervention's mechanism of action (e. g. negative cultures) (Altman and Bland 1999).

The relationship between randomization and causal inference is important, since the desire to assure that observed results are "due to" the studied intervention and that the intervention is achieving its stated mechanism of action may be one reason for use of endpoints that combine symptoms and microbiological results, such as clinical trials in urinary tract infections. Urine culture results are biomarkers used as surrogate endpoints, which should reflect how patients feel, function or survive (Fleming and Powers 2012). However, in acute diseases symptoms, function, and survival can be measured directly, and recent evidence shows a poor correlation between symptoms and urine culture results (Monsen et al. 2014). A primary analysis that includes a microbiological (or other biomarker) outcome, makes evaluation of ITT analysis populations challenging, since

only patients with positive cultures at baseline and/or with follow-up culture results at the time of endpoint analysis can be evaluated in an micro-mITT or PP analysis. However, as discussed below such a micro-mITT analysis may limit the clinical applicability or validity of study findings, if culture results are not available at the time of patient presentation, or if culture results do not reflect direct patient centered outcomes.

Comparison of the interventions with regard to confirmatory therapy decisions, (i. e. the strategy to treat clinical disease caused by a specific pathogen(s) of interest) is an important question for patients and clinicians when the offending pathogen is known at the time of clinical presentation. This makes the “clinical question” and the “pathogen-specific question” identical. The intervention comparison is made using the micro-mITT population (targeting trial participants with the pathogens of interest) since the micro-mITT population would be identical to the ITT population when rapid accurate diagnostics are available.

When offending pathogen(s) are unknown at the time of clinical presentation and prescribing of interventions, the clinical utility of analyses restricting to the pathogen-specific micro-mITT population is lessened. Exclusion of participants who do not have the pathogen of interest ignores results in those participants even though they received the intervention as they would in practice. These results are important both in terms of evaluating “real-world” effectiveness as well as the evaluation of harms. Evaluating results in these participants could influence risk-benefit decision-making if these participants accrue no benefit yet still experience harms of the intervention. The treatment effect in the ITT analysis population represents the actual effect in patients as they present in clinical practice. This effect size may be smaller than in the micro-ITT and therefore may have implications for sample size selection prior to trial initiation.

If there are several micro-mITT populations e. g. defined by patients with several types of pathogens evaluated separately, then the chance of false-positive results increases due to multiple comparisons. The more comparisons investigators make, the more likely they are to find a false positive result due to chance. Statistical methods can be employed to correct for multiple comparisons and should be specified in advance of starting the trial.

Comparison of the interventions’ biological activity vs. a specific pathogen(s) may be of interest to chemists and biologists studying pharmacokinetics, mechanisms of action, and causal pathways; or researchers creating new (or modifying old) interventions, or generating hypotheses for choosing interventions to carry forward into future randomized trials. These analyses are often conducted utilizing the PP population. This analysis population excludes participants based upon factors that occur after randomization such as adherence or drug exposure. Clearly data from a patient that does not take a drug will not be informative regarding pharmacokinetics of the drug. However, exclusion of these patients may lose important information about drug tolerability. PP analyses do not preserve the protection that randomization provides from confounding and selection bias, i. e. exclusions of participants from the PP population may not be independent of treatment assignment and thus the expectation of balance of all other factors cannot be guaranteed. PP analyses are thus subject to the biases of observational studies (Yusuf et al. 1991). These biases may not be able to be addressed by multivariable analyses by known factors such as age and gender, and severity of illness measures. Since these analyses are not based on how patients present but exclude participants post-randomization that initiate the intervention at the time of randomization, the clinical applicability of the analyses may be limited. The generalizability of the results may be limited since trial participants that are excluded may have different baseline factors for the outcomes of interest compared to those participants that adhere. For instance, it is known that participants who adhere to interventions differ substantially in their baseline risk of outcomes compared to participants who do not adhere (Horwitz and Horwitz 1993). For example, one large randomized trial showed that adherers on placebo had improved survival compared to non-adherers on placebo, showing this effects was not due to the intervention but to differences in patients characteristics and their risk of outcomes independent of treatment (Coronary Drug Research Group 1980). PP analyses may be viewed as an attempt to compare treatments under ideal conditions, e. g. among patients that can tolerate and adhere to therapy. These conditions may not exist in real-world practice and indeed adherence in clinical trials may be higher than in the real-world. These analyses may be of interest if measures to improve adherence or reduce toxicity could be advanced.

5 Considerations for Noninferiority Clinical Trials

It has been posed (International Conference on Harmonization 1998) that while ITT analysis populations may be most appropriate for superiority trials which evaluate whether test interventions are superior to a control, the PP analysis population should be given greater consideration in noninferiority trials, which evaluate whether test interventions are not worse than a control arm by more than an unacceptable amount. The suggested reasoning is that the PP analysis may be more conservative in a noninferiority trial because e. g. poor adherence can reduce assay sensitivity attenuating treatment effects, and thus poor trial conduct could drive an ineffective

intervention toward a spurious finding of noninferiority. US Food and Drug Administration (FDA) guidance regarding noninferiority trials exists (US Food and Drug Administration 2010, 2016).

Likewise, in noninferiority trials that evaluate antimicrobials for complicated urinary tract infections and certain other diseases, it has been argued [3] that the ITT analyses should be de-emphasized in favor of the micro-mITT analysis in patients with microbiologically identified target pathogens. The rationale is that because patient outcomes would be expected to be very similar between randomized antimicrobial groups in patients who do not truly have bacterial disease, including these subjects in the analysis could bias toward a noninferiority conclusion.

Although this reasoning is sound in the setting of non-inferiority trials, it also points out limitations inherent in the non-inferiority trial design. There are important tradeoffs to be recognized between the ability of a trial to reliably answer an important clinical question and the sensitivity to detect differences between interventions. If the goal is to answer the pragmatic question of whether the strategy of assigning the test intervention is noninferior to the strategy of assigning the control intervention, then the ITT analysis provides the most direct information.

One important criterion for the interpretability of a noninferiority trial is that the noninferiority margin should be less than the effect of the active control over a hypothetical inert placebo. Otherwise, demonstrating noninferiority would not necessarily imply that the test intervention is superior to a placebo. If this (active control minus placebo) effect is expected to be smaller for an ITT analysis than for other forms of analyses, one design option would be to use the ITT analysis but to specify a noninferiority margin that is smaller than that typically utilized. Justification would then need to be provided that the active control effect over placebo would exceed this noninferiority margin accounting for rates of non-adherence and non-microbiological diagnoses commensurate with an ITT analysis population. Development of rapid point of care diagnostics will help to streamline the design of both superiority and non-inferiority trials, and enhance benefit:risk assessments by excluding patients who would not benefit from the interventions but still might experience harm when the intervention is applied in clinical practice.

In the anti-infective setting it may also be possible to re-design some non-inferiority trials as superiority trials with analyses based on use of an ITT population. An inherent basis for non-inferiority hypotheses is a trade-off of the risk for some loss of effectiveness for superiority on other non-efficacy benefits. If a new intervention offers an advantage in terms of fewer harms or improved convenience, then the primary outcome could be based on a composite global utility endpoint in which a patient's score depends on both efficacy outcomes and "partial credit" from other factors (Evans and Follmann 2016; Evans et al. 2015).

Other aspects of clinical trial conduct such as missing data can affect any analysis population, and the choice of analysis population is not independent of other aspects of study design, analysis and conduct. A discussion of missing data is beyond the scope of this discussion, but counting all missing data as failures (as is common in anti-infective trials) may not be a conservative analysis, especially if there is an imbalance in missing data between the experimental and control groups (Altman 2009; Fleming 2011).

6 Safety and Benefit: Risk Assessment

In anti-infective clinical trials, it is common to define the Safety Population as comprised of all randomized subjects who receive any amount of study drug, with analyses conducted according to treatment received rather than randomized treatment. In most trials this analysis population tends to nearly completely overlap with the ITT population. Thus, one advantage of using the ITT population for efficacy analysis is that it may facilitate benefit:risk assessment because the results evaluating efficacy and harms will apply to the same group of patients. With the treatment strategy of empirically administering antimicrobial treatment before a microbiological diagnosis is known, *all* subjects in the ITT population could be exposed to risks of harms from an intervention. Consequently, if efficacy is assessed only in a subset with microbiologically confirmed disease, the benefits for this group would need to somehow be weighed against the risks in a different and larger patient population. The most pragmatic benefit:risk assessment would utilize an ITT population.

7 Summary

The selection of whom to analyze and include in various analysis populations is an important aspect to the design and analyses of clinical trials. The selection should be driven by the specific research question of interest. Most late stage confirmatory trials should have a more pragmatic "real-world" research focus to evaluate the

strategy of treating clinical disease using the ITT population as the results of these analyses provide the highest level of clinical utility and reflect clinical practice.

The selection of the analysis population has consequences regarding the scientific integrity (i. e. protection from confounding and bias) of the trial results, as populations other than ITT share many of the same issues as subgroup analyses since they are indeed subgroups of the ITT analysis population. The results of randomized clinical trials are least affected by bias when the analysis of study results includes all randomized patients (i. e. analysis population = ITT). Randomization protects against bias allowing for valid causal inferences. Exclusion of randomized participants from the analysis increases the likelihood that bias may affect study conclusions and decreases the clinical relevance of results.

Use of analysis populations other than ITT can be used to address other research questions and provide supportive information regarding the robustness of the trial results. Like subgroup analyses, other analyses that reach different conclusions both quantitatively and qualitatively from the ITT results should be viewed with caution given the increased chance of bias affecting these results.

The complexities of different analysis populations disappear if the ITT, mITT, micro-mITT, and PP populations are identical. While identical populations may be difficult to achieve, design choices can minimize the differences between the analysis populations. The development and implementation of rapid point of care diagnostics in clinical trials and clinical practice will provide a major step to harmonizing the various analysis populations.

Acknowledgements

Research reported in this publication was supported by the National Institute of Allergy And Infectious Diseases of the National Institutes of Health under Award Number UM1AI104681. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This article should not be construed to represent the Food and Drug Administration's views or policies. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health and should not be construed to represent the Food and Drug Administration's views or policies."

Funding

This work were supported by National Cancer Institute, National Institutes of Health, Grant Number: HHSN261200800001E and National Institute of Allergy And Infectious Diseases of the National Institutes of Health, Grant Number: UM1AI104681.

References

- Altman, DG. 2009. "Missing Outcomes in Randomized Trials: Addressing the Dilemma." *Open Medicine: A Peer-reviewed, Independent, Open-Access Journal* 3: e51–3.
- Altman, DG, and JM Bland. 1999. "Statistics notes. Treatment Allocation in Controlled Trials: Why Randomise?" *British Medical Journal* 318: 1209.
- Coronary Drug Research Group. 1980. "Influence of Adherence to Treatment and Response of Cholesterol on Mortality in the Coronary Drug Project." *The New England Journal of Medicine* 303: 1038–41.
- Evans, SR, and D Follmann. 2016. "Using Outcomes to Analyze Patients Rather than Patients to Analyze Outcomes: A Step Toward Pragmatism in Benefit: RiskEvaluation." *Statistics in Biopharmaceutical Research* 8: 386–93.
- Evans SR, Rubin D, Follmann D, Pennello G, Huskins WC, Powers JH. et al 2015. "Desirability of Outcome Ranking (DOOR) and Response Adjusted for Duration of Antibiotic Risk (RADAR)." *Clinical Infectious Diseases* 61 (5): 800–806. DOI: 10.1093/cid/civ495.
- Fleming, TR. 2011. "Addressing Missing Data in Clinical Trials." *Annals of Internal Medicine* 154: 113–17.
- Fleming, TR, and JH Powers. 2012. "Biomarkers and Surrogate Endpoints in Clinical Trials." *Statistics in Medicine* 31: 2973–84.
- Gail, MH. 1985. "Eligibility Exclusions, Losses to Follow-Up, Removal of Randomized Patients, and Uncounted Events in Cancer Clinical Trials." *Cancer Treatment Reports* 69: 1107–13.
- Gillings, D, and G Koch. 1991. "The Application of the Principle of Intention-to-Treat to the Analysis of Clinical Trials." *Drug Information of Journal* 25: 411–24.
- Gravel, J, L Opatrny, and S Shapiro. 2007. "The Intention-to-Treat Approach in Randomized Controlled Trials: Are Authors Saying What They Do and Doing What They Say?" *Clinical Trials* 4: 350–6.

- Hollis, S, and F Campbell. 1999. "What is Meant by Intention to Treat Analysis? Survey of Published Randomised Controlled Trials." *British Medical Journal* 319: 670–74.
- Horwitz, RI, and SM Horwitz. 1993. "Adherence to Treatment and Health Outcomes." *Archives of Internal Medicine* 153: 1863–68.
- International Conference on Harmonization. 1998 "Statistical Principles for Clinical Trials E9." http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf.
- Monsen, TJ, SE Holm, BM Ferry, and SA Ferry. 2014. "Mecillinam Resistance and Outcome of Pivmecillinam Treatment in Uncomplicated Lower Urinary Tract Infection in Women." *APMIS: Acta Pathologica, Microbiologica, et Immunologica Scandinavica* 122: 317–23.
- US Food and Drug Administration. 2010. "Guidance for Industry: Antibacterial Drug Products: Use of Noninferiority Trials to Support Approval." <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm070951.pdf>.
- US Food and Drug Administration. 2014. "Guidance for Industry: Community-Acquired Bacterial Pneumonia: Developing Drugs for Treatment." <https://www.fda.gov/downloads/drugs/guidances/ucm123686.pdf>.
- US Food and Drug Administration. 2015. "Guidance for Industry: Complicated Urinary Tract Infections, Developing Drugs for Treatment." <https://www.fda.gov/downloads/Drugs/.../Guidances/ucm070981.pdf>.
- US Food and Drug Administration. 2016. "Non-Inferiority Clinical Trials to Establish Effectiveness: Guidance for Industry." <https://www.fda.gov/downloads/Drugs/Guidances/UCM202140.pdf>.
- Yusuf, S, J Wittes, J Probstfeld, and HA Tyroler. 1991. "Analysis and Interpretation of Treatment Effects in Subgroups of Patients in Randomized Clinical Trials." *JAMA: Journal of the American Medical Association* 266: 93–98.