An Analysis of the Feasibility of N-of-1 Clinical Trials for the Market Approval of Pharmaceuticals

An Analysis of the Feasibility of N-of-1 Clinical Trials for the Market Approval of Pharmaceuticals

Meghan G. Hart

A Thesis in the Field of Biotechnology

for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

May 2018

Abstract

Randomized controlled trials are the cornerstone necessary for the testing and analysis of pharmaceuticals prior to gaining market approval. Implementation of these trials since the establishment of the Food and Drug Administration has resulted in the approval of thousands of drugs. Importantly, they have also alerted the Agency to drugs that are not safe or well-tolerated by the general population or are not efficacious, as thousands of people are enrolled in multiple trials over the course of clinical development. Despite the benefits of conducting randomized controlled trials and their ability to assess market readiness for the patient population for which the drug is intended, results from clinical trials are not always indicative of how the drug will perform in an individual. Many approved drugs do not have the same efficacy levels in individual patients as demonstrated in clinical trials, leading to patient frustration with the pharmaceutical industry.

More advanced treatment options are emerging in the form of personalized medicine which considers disease management of a single patient. The increase in this trend elicits the need for assessing the role of randomized controlled trials in pharmaceutical development. Personalized medicines should be evaluated with modified studies that assess each patient on an individual basis. Randomized controlled trials do not have the ability to do this, as they measure effects in a group of patients as a whole. The use of N-of-1 clinical trials is a necessity as the field trends towards personalized

treatment, as they measure the effect of a drug on the individual and can additionally draw conclusions of safety and efficacy for a group of patients.

An extensive literature search was performed to evaluate the use of N-of-1 clinical trials with respect to the pharmaceutical industry. Major aspects of clinical development, such as trial design, statistical analysis, data reporting, and regulatory guidelines were analyzed to determine the comprehensiveness of these trials. This case study reveals that N-of-1 clinical trials have established design criteria, methods of statistical analysis, and reporting guidelines that would allow these trials to be used in pharmaceutical development. They are satisfactory to use alone for rare diseases and small patient populations and may be used as supplementary trials for diseases with larger populations or for post-marketing studies for assessment of drug use for additional indications.

Acknowledgments

First and foremost, I would like to express sincere gratitude to Dr. Sujata K. Bhatia for serving as my thesis director. I am extremely appreciative and thankful for Dr. Bhatia's interest in my area of focus and for providing support to me over the course of this project. I am additionally grateful for her contributions and guidance relative to the development of my thesis, without which I would not have been as successful. The idea for this thesis would not have been conceptualized if it were not for a thought-provoking discussion and brainstorming session with Ankita Deshpande, my colleague at Genocea Biosciences. Finally, I would like to thank my family, friends, and colleagues for all the support and encouragement during the extent of my studies, but especially throughout the past year while I was completing this endeavor.

Table of Contents

List of Tables

List of Figures

Chapter I.

Introduction

From its inception in 1938 through 2013, the Food and Drug Administration (FDA) has approved less than 1500 New Molecular Entities (NMEs) (Kinch et al, 2014). While this number doesn't include biologics, it is representative of the fact that only 10 percent of drugs entering clinical trials are approved (Hay et al, 2014). This results in a large majority of inaccessible drugs that could be advantageous to patients seeking treatment. Further, it has been shown that market-leading drugs are not effective for everyone taking them, leading to off-label prescriptions and usage (Schork, 2015). While randomized controlled trials (RCTs) are advantageous in predicting how a drug will benefit the population, the data are not always indicative of how helpful the drug will be to an individual. Within the industry, there is an increasing acceptance of the fact that interventions need to be determined on a personalized basis (Lillie et al, 2011).

The concept of N-of-1 clinical trials has become more relevant as the concept of personalized medicine has evolved. N-of-1 trials, also called single subject clinical trials, are based upon an individual patient as the population (Lillie et al, 2011). Currently, the standard design of these trials follows an individual in response to both placebo and active treatment to determine if treatment is efficacious (Kravitz et al, 2014). Treatments may be administered one after another for ease of comparison or a washout period may occur between treatments (Kravitz et al, 2014). Safety is also observed upon administering interventions; however, this type of trial differs from RCTs as it determines the optimal intervention for one individual as opposed to a population (Lillie et al, 2011).

These clinical trials are used routinely in educational and learning environments, yet have not generally been pursued in a biomedical setting (Lillie et al, 2011), as RCTs are the gold standard accepted by the FDA. However, it has been suggested that N-of-1 trials "provide the strongest evidence for the decisions of the individual patient" (Chen et al, 2014).

With personalized medicine on the rise and the lengthy time required to bring a drug to market, new options are needed for clinicians to make informed decisions about treatment for their patients. The introduction of genetic screening has improved treatment options; for example, it has been shown that tumors with a certain genetic makeup will not respond to certain drugs (Lillie et al, 2011). While screening is important, its application is not possible for every disease requiring treatment, and more personalized studies are needed to connect the right treatment with the right patient. Due to cost and feasibility, the concept of randomized controlled trials no longer seems fit for situations in which a small number of patients could benefit from an off-label treatment. In this instance, the use of N-of-1 clinical trials is necessary for determining the best course of therapeutic action. Further, N-of-1 studies could be extremely applicable in the rare disease space, as the patient population is limited and may not be large enough for a RCT. However, as this is a new concept in the field, there are many questions surrounding the implementation of these trials. The number of times treatment must be repeated to prove safety and efficacy along with the regulatory process as a whole must be considered, as these studies may not follow the typical processes of approved drugs. This case study will serve as an analysis of the concept of N-of-1 clinical trials with

respect to pharmaceuticals and will assess the feasibility of the use of these trials for product licensure and treatment of individual patients.

History of the Pharmaceutical Industry and the Drug Development Process

The Food and Drug Administration (FDA) originated in 1848 due to the need for chemical analyses in the agriculture industry, making it the oldest agency of the United States government (About FDA: History, 2015). The first regulations were enacted in 1906 with the passing of the Pure Food Drugs Act, which ruled against commerce between states in regards to misbranded and contaminated food and drugs. Although given its official name in the 1930s, the mission of FDA has not evolved since its inception. Rather than creating and then later removing laws that no longer fit the time period, the FDA's regulations and guidelines represent an evolution of the changes in the industry (About FDA: History, 2015).

Clinical trials have been occurring informally for centuries, beginning with the use of medical observations made by ancient civilizations (Junod, n.d.). However, it wasn't until the early twentieth century that well-controlled trials occurred. The 1938 Food, Drug, and Cosmetics (FD&C) Act was the first official requirement for pre-market safety evaluations, requiring the FDA to review pre-clinical and clinical data for drugs seeking market approval. It was this enactment that allowed the FDA to require more data if necessary to approve a drug candidate (Junod, n.d.). While rules and regulations for approving drugs were becoming stricter, the Thalidomide drug crisis, identified in 1961, prompted the Drug Amendments of 1962. Thalidomide was patented in Germany in 1954 and gained licensure in 1958 in the UK (Mandel, 2015). The drug was produced to help pregnant women with morning sickness. However, after children of mothers

taking Thalidomide were born with severe birth defects (Mandel, 2015), it was discovered that the drug's molecules could pass through the placental wall, thereby causing harm to the fetus. While this drug was never approved in the United States, this crisis prompted the FDA to more strictly enforce its drug approval process. The 1962 Drug Amendments stated that drug approvals would rely on scientific testing to prove not only safety but also "substantial testing" of efficacy gathered by adequate, well-controlled clinical trials (About FDA: History, 2015).

FDA Drug Approval Process

In order for a drug to make it to market and be accessible to patients, clinical trials must be conducted as part of a formalized process mandated by the FDA. For most interventional pharmaceuticals, randomized controlled trials are conducted many times to prove safety and efficacy of the product. Prior to market approval, there are three clinical trial phases that must be completed and obtain statistically significant data. Once data collection and analysis has occurred, a New Drug Application (NDA) or Biological License Agreement (BLA) is submitted, and, after review, the product is either accepted or denied.

To initially begin clinical trials after pre-clinical work is completed, an Investigational New Drug Application (IND) must be filed with the FDA (Brody, 2016). The IND contains information and data regarding the drug substance based on *in vitro* and *in vivo* studies. Once approved, this IND allows for testing of the product in human subjects in Phase 1 studies. Phase 1, or "first in human studies", assess safety, typically in healthy volunteers (Brody, 2016). However, in instances where diseases are life-threatening, the actual patient population may be used, for example with cancer (Lipsky

et al, 2011). Phase 1 studies typically enroll 20-100 people and monitor subjects for adverse events (AEs). Dosage level is also examined through dose escalation studies (Lipsky et al, 2011) in order to find the proper dose for subsequent trials (Brody, 2016). In many cases, a small group of study participants receives a small dose of the study drug. As the drug is tolerated, more cohorts of patients receive increased doses of the drug until dose-limiting toxicities (DLTs) are reached. The dose just below the DLT, the maximum tolerated dose (MTD), is typically what is used in later phase clinical trials (Brody, 2016). Phase 1 studies are generally completed in a minimum of several months (The Drug Development Process, 2005).

Phase 2 clinical trials are larger, enrolling 100-300 people with the disease of interest (Lipsky et al, 2011). They are sometimes divided into Phase 2a and 2b studies and are conducted to determine efficacy while also continuously measuring safety (Brody, 2016). Dose frequency and delivery method are additionally examined during these studies (Lipsky et al, 2011). Phase 2 studies have strict inclusion and exclusion criteria meant to focus on the population of interest to determine how efficacious the drug will be for the subset of people affected by the disease or ailment. Often, drugs that are neither safe enough nor effective enough are discontinued in this phase (Lipsky et al, 2011). The duration of Phase 2 studies is typically around two years (The Drug Development Process, 2005).

Phase 3 trials represent the last stage of clinical development (Brody, 2016). Clinical trials in this phase are the largest, with subject enrollment reaching 300-3000 individuals (The Drug Development Process, 2005). A large number of subjects is needed in this phase in order to ensure drug-related toxicities are recognized and to confirm

efficacy in a broader population (Brody, 2016). This phase typically consists of two pivotal studies measuring safety and efficacy in a more generalized population and seeks to confirm previously generated data. Data obtained in these studies are important for the drug label and package insert (Brody, 2016). The inclusion and exclusion criteria for these trials is not as strict, as they aim to prove efficacy in a wide variety of individuals suffering from the disease or condition across many clinical sites. Phase 3 studies generally last multiple years (Lipsky et al, 2011).

Due to the regulations of clinical trials and the amount of time it takes to complete them adequately, in addition to the drug discovery and preclinical study process, the length of time required for FDA approval of a pharmaceutical product is at least 12 years. Further, the cost to develop a drug is exorbitant, costing 2.6 billion dollars (Tufts, 2016). This, along with the added cost of drugs that fail in preclinical or clinical development, accounts for the high cost of drugs, thereby affecting payers and ultimately the patient through increased drug prices, often for treatments that are life-saving. As RCTs are the standard for the drug approval process, there has not been much research conducted on alternative solutions to these issues.

Randomized Controlled Trials

Randomized controlled trials are used in intervention studies and typically include at least two groups, one being the intervention and the other a control (or placebo) group. After enrollment, study subjects are randomly assigned to one group (Akobeng, 2005), referred to as randomization. Viewed as the gold standard for evaluating interventions, RCTs are “the most scientifically rigorous method of hypothesis testing available.” Although the two groups are administered a different intervention, they are monitored in

the same manner for the same period of time, thus allowing the investigator to draw conclusions based on similarities or differences between the two groups. In many cases, these trials are double-blind, where both the investigator and the subject are unaware of the intervention that is assigned to the study subject. This is an important piece of trial design, as it allows for the exclusion of selection bias.

Bias in clinical trials is undesirable, as it may influence the outcome by bringing into question whether the intervention was effective or not. By randomly allocating subjects to treatment groups, participants are fairly distributed in terms of elements such as subject age, sex, or disease status (Akobeng, 2005). In addition to selection bias, which occurs before the trial begins, there are other types of biases that can be present during and after a trial. Performance bias can occur if there are any differences in treatment between trial groups in terms of care or exposure to factors other than the intervention (Higgins et al, 2011). Conversely, detection bias can occur if there are differences in determination of outcomes between trial groups due to patient characteristics. Blinding of participants can assist with minimizing both of these biases, as the investigator or sponsor is unaware of which groups participants are assigned to. Finally, attrition bias can occur when there are unequal amounts of withdrawals in the study groups due to experiencing adverse events or lack of treatment efficacy (Higgins et al, 2011).

The introduction of biases in randomized controlled trials is detrimental, as it can affect internal validity, or the reliability or accuracy of the study (Pannucci et al, 2010). By controlling for high internal validity, investigators can be confident that the efficacy reported is accurate for the highly-specified group of individuals. However, high internal validity compromises a study's external validity, and it is difficult to maintain a balance

of the two (Pannucci et al, 2010). If more personalized studies were acceptable for use in drug testing, many of these biases would no longer be an issue that sponsors or investigators would have to focus on during trial design and implementation, as they would be designing a trial for one individual and not a group of people.

Design of randomized controlled trials is important to consider in order to create both internally and externally valid trials. While internally valid trials keep the possibility of bias to a minimum, externally valid trials obtain results that can be generalized to a defined group of patients (Rothwell, 2005). Despite being so highly regarded, RCTs are not perfect. One of the major flaws with RCTs is their lack of external validity; their results cannot be generalized to individualized patients (Clay, 2010). Due to the large sample size needed to attain statistical significance, results are often indicative of central tendencies of a large sample size which does not translate to the patient as an individual (Clay, 2010). Further, the real-world population may not be represented by the sample of individuals participating in a RCT (Clay, 2010). Although Phase 3 studies have a large sample size and less restricted eligibility criteria in order to closely mimic the actual population who would be receiving the drug, a true heterogeneous population is hard to achieve. Patients are excluded from trials due to eligibility criteria, and this exclusion makes the population more homogeneous. Patient characteristics, especially those that make a patient drastically different from the majority of the patient population, can also influence external validity of a trial through efficacy and safety. Further, enrichment strategies such as placebo run-in periods or companion diagnostic assays that aim to exclude non-responders from studies decrease the external validity of trials (Rothwell, 2005).

Other factors influencing a lack of external validity are the length of treatment and follow-up along with patient-centered outcomes (Rothwell, 2005). Length of treatment or follow-up periods are important in order to measure the effect of the drug on improving an ailment. Many trials are designed to follow acute response to treatment but do not follow efficacy long-term. However, these initial responses may not be a good predictor of long-term efficacy. The lack of long-term follow up can create issues with external validity, as beneficial effects of the treatment may not be experienced in clinical practice (Rothwell, 2005).

Similarly, clinical outcomes serve to increase external validity if they corroborate with patient priorities rather than clinician priorities (Rothwell, 2005). Patient-centered outcomes generally are focused on improvement of quality of life, such as mental, emotional, or general health, as opposed to improving physical effects, which are typically measured and are the focus of clinicians. Clinical outcomes need to be relevant to the patient, despite whether this causes a decrease in statistical power (Rothwell, 2005). The issue of decreasing external validity due to intervention length, follow-up, and clinical outcomes could be addressed by initiating clinical trials focusing on the individual. The length and follow-up of treatment would become more personalized based on the needs and opinions of the individual receiving the intervention, rather than simply following a schedule created for a group of patients. Individual focus on the patient can also prioritize clinical outcomes and can more easily involve patient opinion in the outcomes that are important to them.

Another flaw with RCTs is their compliance with guidelines for Good Clinical Practice (GCP), set forth by the International Conference on Harmonization (ICH). GCP

regulations state: “Before a trial is initiated, foreseeable risks and inconveniences should be weighed against the anticipated benefit for the individual trial subject and society. A trial should be initiated and continued only if the anticipated benefits justify the risks” (The Principles of ICH GCP, n.d.). Thus, a participant of a clinical trial should benefit from the intervention rather than simply be exposed to risks. This is problematic in RCTs, as many Phase 1 studies use healthy volunteers to determine safety; therefore, there is no benefit. Personalized clinical studies would alleviate this ethical issue, as the only participants in the study would be those who have a disease or disorder and are in need of treatment.

Importantly, it has been shown that the top 10 most profitable drugs in the United States only help treat between 4 and 25 percent of the people who are taking them (Schork, 2015). This is due to the fact that results from randomized controlled trials are not representative of an entire population. Drug efficacy should be held to a higher standard, and this is only possible with personalized treatment options stemming from personalized clinical research. The use of N-of-1 clinical trials is a necessary measure for accurately predicting the response of an individual to drugs, as they examine specifically how the drug will affect the patient as opposed to generalizing a drug’s efficacy to an individual. While there are data supporting N-of-1 trials in non-pharmaceutical fields, the use of these trials needs to be assessed to determine feasibility of use for drug approval in the United States.

Randomized Controlled Trials Are Not Ideal for Market Re-approval

While RCTs are necessary for drug approval, they may not always be the best option, as they can leave gaps in treatment if the drug does not get approved due to

generalization to the entire population. Additionally, if a drug is approved and then later removed from the market, RCTs need to be performed again to seek re-approval, which is a lengthy process. In 1999, the drug rofecoxib, known by its trade name of Vioxx, was launched by Merck as a safer option than common non-steroidal anti-inflammatory drugs (NSAIDs) for the treatment of pain associated with osteoarthritis (Krumholz et al, 2007). The mechanism of action of NSAIDs is to block types 1 and 2 cyclooxygenase (COX) enzymes, which, when produced, are pain-signaling. Both COX enzymes produce prostaglandins which, while promoting inflammation and pain, have other functions (Zarghi et al, 2011). Importantly, COX-1 prostaglandins contribute to blood clotting (Zarghi et al, 2011) while COX-2 prostaglandins are vasodilators, thus inhibitors of platelet aggregation (Ricciotti et al, 2011). NSAIDs bind non-selectively to both COX enzymes, and while this inhibition causes anti-inflammatory responses, their binding also causes gastrointestinal and toxic renal side effects due to the blocking of COX-2 (Zarghi et al, 2011). To attenuate these side effects, rofecoxib and other COX-2 inhibitors were developed with the intention of selectively inhibiting COX-2, thereby eliminating the sometimes-extreme side effects seen with non-selective inhibitors.

Despite these selective COX-2 inhibitors being efficacious, people taking rofecoxib experienced extreme cardiovascular side effects, including heart attack and death (Zarghi et al, 2011). An estimated 60,000 people died as a result of taking Vioxx (Herper, 2005). Both rofecoxib and another COX-2 inhibitor, valdecoxib, were removed from the market, and others, including NSAIDs were asked to update their labels to include gastrointestinal and cardiovascular risks (Center for Drug Evaluation and Research, 2005). While one selective COX-2 inhibitor remained on the market, it was

given a black box warning to alert consumers of potential cardiovascular side effects (Center for Drug Evaluation and Research, 2005). As the mechanism of cardiovascular disease was unknown, the FDA needed to take action to protect against potential side effects of the drug, even if there was some underlying mechanism contributing to the issues.

Further research on the mechanism of action was conducted and in 2008, a study was published describing "acceleration of cardiovascular disease by a dysfunctional prostacyclin receptor mutation" (Arehart et al, 2008). The prostacyclin receptor mutation, R212C, was associated with cardiovascular disease in high risk groups (those with underlying stress or injury) (Arehart et al, 2008). While this research was not available at the time the FDA suggested removing the drugs from the market, it sheds light on the fact that genetic testing could have prevented this situation by exposing a predisposition to cardiovascular risk factors. Similar genetic tests are available prior to prescription of another COX-2 inhibitor, Celecoxib (Dean, 2016). Further, while RCTs with companion diagnostic assays could be used to reintroduce the drug to the market, this is both a lengthy and costly process. Smaller, more personalized studies would be more beneficial, as they would take less time to complete and would be less expensive. It would be advantageous to use a different type of trial design to reintroduce this class of drug that is still helpful to a large majority of people, despite the extreme side effects that occurred in a small subset. The use of personalized studies for the purpose of drug introduction or reintroduction to the market should be evaluated to determine whether these studies are a viable option for approval.

Randomized Controlled Trials Are Not Satisfactory for Off-Label Drug Use

While drugs that are approved generally work for much of the population they are intended for, there are instances where people taking drugs do not benefit. This leads to off-label drug prescription. Off-label treatment occurs when a drug is prescribed for an indication not included on the product label. This commonly occurs with the prescription of tricyclic antidepressants used to treat pain, or Ativan, an anti-anxiety drug, used to treat nausea as a result of cancer treatment (The American Cancer Society, 2015). While off-label treatment can be successful, there is a lack of information regarding the usage of these drugs for indications other than what they were approved for. Therefore, there is a higher risk of side effects or reactions due to the medication (The American Cancer Society, 2015).

Currently, to obtain an additional indication for a drug, sponsors must conduct randomized controlled trials following the standard process previously described in order to prove the drug is safe and efficacious against the desired indication. This is time-consuming and costly for the sponsor, yet may be beneficial in the end to both the patients and the pharmaceutical manufacturer if the new indication has a large enough population. However, it would be more beneficial to the patient if there was a formalized process for collecting information regarding off-label drug use. This information could be collected through clinical trials if they were to occur. Rather than conducting RCTs, an alternative trial design would assist in the collection of data regarding patients who are prescribed drugs off-label, and this information could be analyzed collectively to determine whether a certain drug is suitable for other indications not initially examined.

The Rise of Personalized Medicine and the Need for More Personalized Studies

With the increase in popularity of personalized medicine, pharmacogenomics is becoming more and more important. Pharmaceuticals that gain approval status do so because they are proven to be effective for a statistically significant number of individuals. However, prescription drugs are not effective for everyone. Because of this, genetic screening is becoming commonplace to determine whether a drug will or will not be effective (Lillie et al, 2011). This is very common in cancer, where screening tests and companion diagnostics are being used more frequently. Upon diagnosis of breast cancer, the level of human epidural growth factor 2 (HER2) on the surface of cancer cells is measured in order to determine treatment options (HER2 Testing for Breast Cancer, 2015). Depending on the levels of HER2 present on the tumor, HER2 targeted therapy or another option will be discussed (HER2 Testing for Breast Cancer, 2015).

Further, more recent tests have been developed to measure levels of PD-L1 on tumor cells from patients diagnosed with non-small cell lung cancer (NSCLC) or melanoma. These tests, approved as either companion or complementary diagnostic tests, direct the patient's physician as to whether to administer monoclonal antibody (mAb) therapy or use another type of treatment. While these tests are important aspects of determining the best treatment for patients, they more so serve to stratify groups of patients into responders and non-responders based on the tumor genetic makeup (Lillie et al, 2011).

Other strategies aim for a more targeted approach, such as neoantigen vaccines. These therapies predict important tumor neoantigens that, upon administration, will mount an immune response to fight the tumor. Because each patient's tumor may have different neoantigens, each vaccine created could potentially be different. In this case,

while the therapy is fundamentally the same, it is in fact different for each person. Thus, if personalized vaccines are given and responses are based on the individual, do trial design and analyses geared towards randomized controlled trials still apply? Because of the shift to personalized medicine, there is a large need to begin to design clinical trials based around the patient as an individual, not as a large comprehensive group.

Different Trial Design is Necessary for Rare Disease Indications

One stipulation of randomized controlled trials is a large population of interest which is needed to ensure enough enrollment to sufficiently power the study will be achieved. In the rare disease space, this is sometimes difficult to accomplish. According to the FDA's Orphan Drug Act, rare disease status is defined as "any disease or condition which affects less than 200,000 persons in the United States" (Commissioner, 2013). The FDA provides financial incentives to companies who develop drugs for orphan diseases so that these diseases are acknowledged. Incentives include tax credits, funding, and market exclusivity for a seven-year period (Information for Consumers, 2016).

There are several challenges with developing drugs for rare diseases, but the most evident is that the patient population can be small, therefore it is difficult to conduct studies that are powered adequately (Augustine et al, 2013). Patients may be located globally, which challenges clinical trial enrollment along with collection of data for natural history studies (Pariser, 2014). For diseases that do have drugs in development, clinical trials can become a question of morality for parents as to whether they should enroll their child into a clinical trial where they may be administered placebo (Augustine et al, 2013). RCTs are lengthy and costly to conduct, and some patients simply do not have the time to wait for the entire process to be completed for drug approval. While

accelerated approvals are offered by the FDA, there is a great need to have alternate clinical trials designed to fit the needs of this small patient population while maintaining reputability in proving efficacy.

N-of-1 Clinical Trials

An alternative to randomized controlled trials, used often in education or learning settings for behavioral or psychological assessments, is N-of-1 clinical trials (Lillie et al, 2011). N-of-1 trials focus on intervention results for the patient as an individual rather than in a group setting. Although not typically pursued in a medical setting, with the rise of precision medicine, this type of trial design could prove to be fruitful and would address many of the issues the industry is faced with regarding RCTs. Because N-of-1 trials are focused on the individual, they provide more accurate treatment options for the patient, as the patient is not simply receiving a drug that has been generalized to an entire group of individuals. The patient's own data will help to decide the treatment plan that is best suited for that individual (Lillie et al, 2011), consequently generating stronger evidence-based data that the best treatment is being administered. Further, this type of trial design would be less expensive in terms of cost per patient, as costs average $36,500 per patient for RCTs (Biopharmaceutical Industry-Sponsored Clinical Trials: Impact on State Economies, 2015) as compared to an estimated cost per patient of less than $2,000 for N-of-1 participants (Scuffham et al, 2008). This is due to the increase of the use of Contract Research Organizations (CROs) for RCTs for the management and facilitation of clinical trials along with the shift to principal investigators who are non-academic physicians in the private sector (Bothwell et al, 2006). N-of-1 trials are also faster to complete than RCTs, as timelines depend on the individual only, not a large group of

people. Thus, the time a patient needs to wait before he can be prescribed treatment could decrease from years to months, depending on the length of the intervention.

N-of-1 trial design is derived from standard design strategies used in randomized controlled trials, ensuring design integrity is not lost (Lillie et al, 2011). However, what is unique about N-of-1 trials is that the design can be agreed upon by physician and patient. Despite this, trials are designed to ensure safety, which is measured throughout the trial, similar to RCTs. (Lillie et al, 2011). Dosing periods are likewise determined, along with whether washout periods will or will not occur (Kravitz et al, 2014). Randomization, an important piece of design intended to avoid bias, can still be used in terms of order of administration of the intervention, whether it be placebo or drug product. This is important in N-of-1 trials, as attaining balance in the assignment of intervention will help to prevent bias from confounding factors that may affect the trial outcome. This could be achieved by alternating the administration of treatment A and treatment B, for example: ABBABAAB (Kravitz et al, 2014). In this case, the patient is exposed to one treatment followed directly by the other, which allows for a direct comparison to be made between treatments. These treatments are then repeated and data is collected again to protect against random error. Whereas sample size is important for RCTs to achieve significance, repetition is important in N-of-1 trials (Kravitz et al, 2014). Blinding is suggested for N-of-1 studies and, for studies involving drugs, is possible by compounding the different interventions in the same manner (Kravitz et al, 2014); however, this is not always a feasible option.

While N-of-1 trials have been used in other fields, their presence in the pharmaceutical industry is scarce. However, they are a compelling alternative to

randomized controlled trials for diseases that have small populations or in the case of personalized medicine. Their use could facilitate the treatment of patients by finding the right treatment in a shorter period of time and eliminating the chance that the patient is only administered placebo (Lillie et al, 2011). Further, the cost to sponsors could decrease, as trial expenses may not be limited to one institution and could be distributed amongst many depending on patient location. Additionally, the design of these trials can be more heterogenous than in RCTs, providing that objective evidence which promotes one intervention over another is found (Lillie et al, 2011). Despite these benefits, research is needed to determine whether these types of trials are feasible for the pharmaceutical industry and, if so, how precedent will be set for the incorporation into the FDA approval process for a drug, biologic, or medical device.

Chapter II.

Case Study Methods

This case study will analyze the components of N-of-1 clinical trials and assess how well established and suitable they are for use of pharmaceutical approval by the FDA. Initially, the foundation for the application of N-of-1 trials will be established by evaluating key factors such as trial design, statistical analysis, and regulatory considerations. Examples of N-of-1 trials will be presented and analyzed in order to fully understand how they can be applied. N-of-1 studies that have been used for exploratory research or feasibility studies will be assessed, as they represent instances where N-of-1 trials were used in pharmaceutical settings.

Next, the need for N-of-1 trials as opposed to the use of randomized controlled trials will be investigated to further strengthen the argument. N-of-1 clinical trial design will be analyzed to determine how it can benefit products that are currently in development, such as drugs for rare diseases or in the cancer field, or how they could have served as beneficial to products that have already completed clinical trial studies. One or two drugs will be examined and N-of-1 trial design will be applied to these products. The benefits and disadvantages of using this alternative trial design for product approval will be assessed based upon justifications found in the literature promoting N-of-1 trials over RCTs. While these justifications represent trial design in general, in this analysis, they will be applied to specific cases and expanded upon to prove the value of N-of-1 trials.

Finally, a high-level financial overview of N-of-1 clinical trials will be examined to assess the readiness of determining the economic impact of an intervention on a per patient basis as opposed to a group of patients. The ability to determine the value of N-of-1 clinical trials from a cost perspective will be evaluated by comparing the evaluation method to the established method for randomized controlled trials. As a result of this case study, recommendations will be made regarding the development, implementation, and utilization of N-of-1 clinical trials in the licensure of pharmaceutical products.

Chapter III.

Results

N-of-1 clinical trials are a relatively new type of trial design that is beginning to be explored in the pharmaceutical industry. With the increase in treatments centered around the patient as an individual, there is a need for the exploration of a clinical trial design that reflects personalized treatment. Traditional N-of-1 trials are most effective and applicable for chronic diseases with treatments that do not have long-lasting effects, such that treatments can be alternated and results generated in a period that is not time consuming for the patient (Duan et al, 2013). These are prospective studies that can be placebo-controlled.

To maintain clinical trial integrity, randomization and blinding should be included in the trial design. In the case of N-of-1 trials, the trial participant receives the treatments in question in a repeated fashion over the course of a specified period. For example, subjects receiving two different treatments, A and B, would be randomized to these treatments prior to administration. While there are many options for treatment assignment, balance in randomization is important in the design of N-of-1 trials to avoid bias in outcomes of treatment that could be generated in a time-dependent manner due to clinical or environmental effects (Kravitz et al, 2014). A strong randomization scheme is ABBABAAB in order to avoid potential confounders (Kravitz et al, 2014).

While it is possible to treat the patient with each therapy just once in order to look for an effect, repetition is needed in order to design a more powerful study that is free of

errors. As this type of trial design includes just one subject, more data points are needed for increased power. Thus, increasing the amount of times a subject receives a treatment increases the confidence that the results generated are accurate. When possible, these studies should also be blinded to provide more integrity and limit the amount of bias present in the study.

Another consideration of N-of-1 trial design is the washout period. Washout periods are used in RCTs prior to drug administration in order to remove any potential effects from other drugs taken prior to enrolling in the trial. N-of-1 trials do not need to include washout periods if the drug has a short therapeutic half-life (Kravitz et al, 2014). If this is not the case, adequate washout periods must be included in order to ensure that the effects from one treatment segment do not affect the subsequent treatment. However, the benefit of not using a washout period is that the patient does not have to go without treatment for any period of time. An important consideration when assessing different drugs in an N-of-1 design is lasting effects that could be caused by one drug that may affect the second drug. If this is the case, washout periods become critical for clinical trial design (Kravitz et al, 2014).

Currently, N-of-1 trials are more commonly used in clinical practice, where a patient and doctor use the design to determine which treatments are most efficacious and beneficial to the patient. While this is a practical way to use these trials, they are beginning to be explored more in clinical trial applications for drug development and could be applied further than clinical practice for use in clinical research. Regulations for filing single patient INDs have been set forth by the FDA as part of the Code of Federal Regulations (CFR) (Code of Federal Regulations, 2016). Often, single patient INDs are

filed by physicians for cases of compassionate use. Part 312.310 of CFR21 outlines the requirements for filing these INDs. Additionally, part 312.315 discusses the filing of INDs for intermediate-sized groups of patients, which could be the preferred filing mode if there are many single patient INDs filed for the same investigational drug (Code of Federal Regulations, 2016). The existence of these regulatory processes were developed for the purposes of clinical practice; however, they have application for clinical development and should be explored further for these cases.

Applications of N-of-1 Clinical Trials for Pharmaceutical Evaluation

Despite being uncommon, there have been individual patient trials that have been performed to evaluate the feasibility of their use for clinical development. Many different conditions have been evaluated in single case designs, a few of which will be summarized in order to understand the different methods of their applications.

In March 2014, an N-of-1 trial was registered in the United Kingdom to determine the effect of ephedrine as an add-on therapy for patients suffering from myasthenia gravis (MG). Myasthenia gravis is a rare autoimmune disease affecting neuromuscular transmission due to the production of auto-antibodies against the neuromuscular junction (Vrinten et al, 2015). Clinical symptoms of this disease are fluctuating muscle weakness and fatigue, both of which cause interruptions in daily activities (Vrinten et al, 2015). Typically, acetylcholinesterase inhibitors (AChEIs) are used to treat MG by enhancing neuromuscular transmission by blocking the enzyme responsible for breaking down acetylcholine; however, not all patients respond. Those that do not are reliant upon high doses of corticosteroids or other immunosuppressive medication which can lead to serious side effects. Previous clinical observations by Vrinten et al, a group from the UK

and The Netherlands who initiated this trial, have described the use of ephedrine, a substituted amphetamine, as a combination therapy with AChEIs or low doses of prednisone as a preferred alternative to second line treatment. Mechanism of action and effects have been studied yet are not well understood. However, this combination therapy seems to be well-tolerated by patients. (Vrinten et al, 2015).

While there was a potential benefit with ephedrine, it had not been fully evaluated in a RCT and therefore could only be used off-label. Off-label usage is not ideal as it can cause issues with reimbursement and access. Further, as a rare disease, the patient population of those suffering from MG who could benefit from this treatment may not be large enough to conduct a well-powered RCT, which is where N-of-1 trials could be applicable. As part of a project aiming to investigate whether N-of-1 trials are an appropriate means of achieving market approval and reimbursement, Vrinten et al and their groups from University College London and various University Medical Centers in The Netherlands examined the use of ephedrine as an add-on treatment to AChEIs in an N-of-1 format.

The study was designed as follows: each patient participated in a randomized, placebo-controlled, double-blinded, multiple crossover N-of-1 trial which was followed by an optional open-label extension phase (Vrinten et al, 2015). Three cycles consisting of two one-week long treatments per cycle were aimed to be completed per patient, totaling six weeks of treatment per patient. The order of ephedrine and placebo treatments was randomized, and treatment was administered for five days with a two-day washout period. Following treatment, a four-week evaluation period consisting of data analysis, feedback of results to patient and physician, and trial evaluation occurred. The study

population consisted of adult patients with generalized MG as confirmed by a positive serological test for acetylcholine receptor antibodies whose disease status had not progressed with prior treatments of pyridostigmine alone or in combination with low-dose prednisone or other immunosuppressive drugs and who were able to tolerate ephedrine. A sample size of four patients was used (Figure 1), which resulted in a study that was 77% powered. Efficacy was assessed based upon the performance of ephedrine versus placebo as measured by the Quantitative Myasthenia Gravis (QMG) score. In addition to secondary objectives pertaining to treatment effect, acceptability of N-of-1 trial design and execution was also assessed (Vrinten et al, 2015).

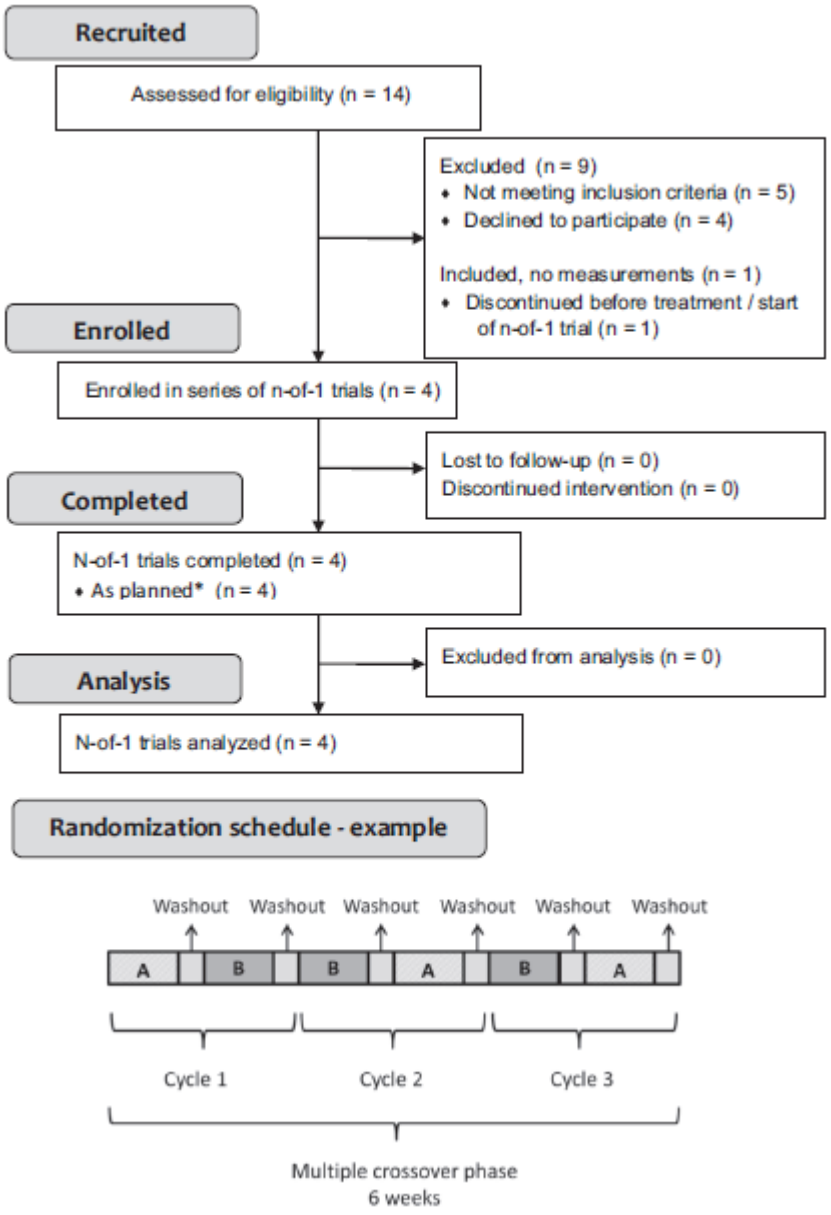

Figure 1: CONSORT Flow Diagram for N-of-1 trial

CONSORT Flow Diagram for N-of-1 trial to assess the use of ephedrine as an add-on treatment to acetylcholinesterase inhibitors for the treatment of Myasthenia Gravis (Lipka et al, 2017).

As a result, a mean improvement in QMG score was observed (increase of 1.0) in the ephedrine as add-on treatment group versus placebo (Lipka et al, 2017). This

improvement was significant in terms of both the trial average treatment effect and the population treatment effect. A trial average treatment effect on secondary endpoints was also seen upon administration of ephedrine; however, no treatment effect was seen at a population level. Further, the trial showed the treatment to have a good safety profile, with most adverse events occurring on days the subjects received treatment. After the trial was completed, three of the four patients decided to continue with ephedrine as an add-on treatment in an open-label phase (Lipka et al, 2017).

As an assessment of the trial design and execution, different standpoints were examined to determine the feasibility of N of 1 trials (Weinreich et al, 2017). While no time was saved in terms of IRB approval procedures and trial registration as compared to a larger clinical trial, the trial overall was reported as easy to conduct. This study also determined that "a series of N-of-1 trials can be a very effective study design to detect even a small effect in a small patient population by replacing the large variance between patients in standard RCTs with smaller variance within individual patients" (Lipka et al, 2017). While a small number of single patient trials was performed (four), performing statistical analyses to examine both population and trial average effects allowed for extrapolation to the entire population suffering with MG.

In another study, Nikles et al, a group from Australia, assessed the feasibility of the N-of-1 trial design approach to evaluate palliative care treatment for advanced cancer patients. Many patients who are suffering from this disease experience xerostomia, or dry mouth. While there are treatments available to increase saliva production, Nikles et al wanted to investigate whether pilocarpine, a cholinergic agent used to treat open-angle glaucoma, would be an effective means of treatment of xerostomia. The group focused on

a N-of-1 trial design rather than a randomized controlled design due to the small patient population that is affected. With a half-life of less than an hour, pilocarpine is a good candidate for N-of-1 trials (Nikles et al, 2015). This study focused on buccal administration of the drops, as a pill form of Pilocarpine was not available.

Adults with advanced cancer and a dry mouth score of greater than 3 on an 11-point scale were enrolled in the study and were given both pilocarpine and placebo drops, each to be administered three days at a time for three cycles (Figure 2) (Nikles et al, 2015. The primary endpoint of this study examined average xerostomia in a 24-hour period, measured by a numerical rating scale (NRS) for dry mouth. Secondary outcomes were also examined to assess the ease of administration of the medication, looking at dysphagia and dysgeusia. Subjects were randomized and all participants in the trial, including trial staff, were blinded. As the main purpose of this study was to assess feasibility of the N-of-1 design with this treatment, only 20 participants were required. However, 70 patients were needed in order to measure a definitive effect (Nikles et al, 2015).

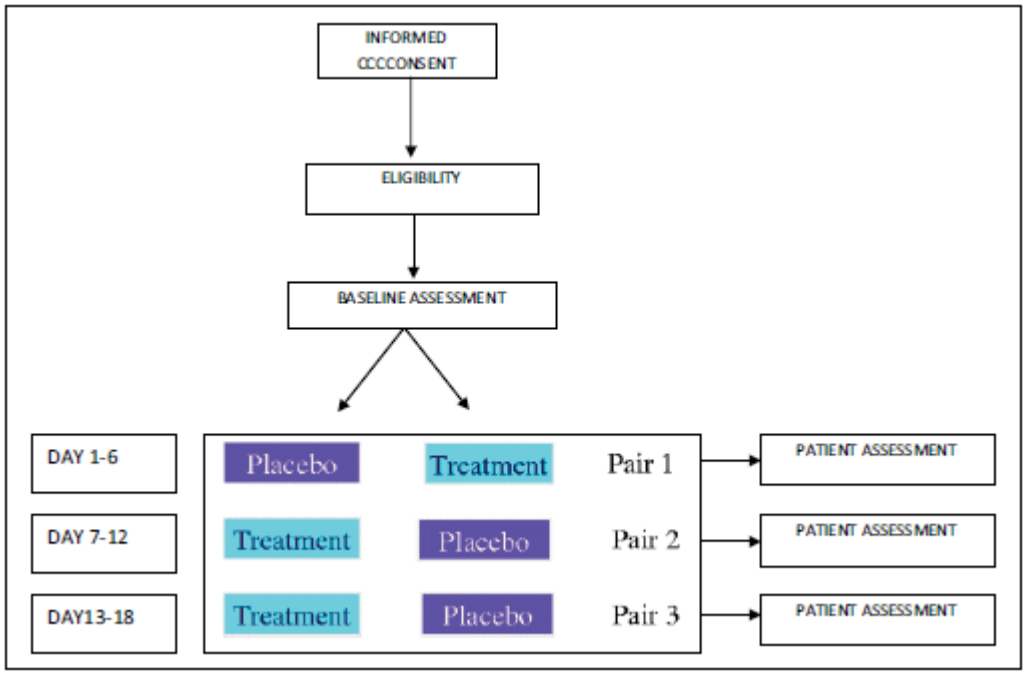

Figure 2: Study schema for N-of-1 trial

Study schema for N-of-1 trial assessing pilocarpine drops for the treatment of xerostomia in patients with advanced cancer (Nikles et al, 2015).

While 20 subjects were recruited for this study, withdrawals from the trial (not related to the medication) resulted in only five participants completing all three cycles; however, one of the five subjects did not have a complete data set (Nikles et al, 2015). Of the four subjects, there was a 50 percent response rate based on xerostomia NRS measurements. Side effects of each treatment were similar, thus there was no indication that trial participants were unblinded in any manner. However, due to positive response to the medication, participants were able to correctly identify which medication they received in most cases (Nikles et al, 2015). Despite there being issues with withdrawals from this feasibility study, results could still be used to assess design and conduct of

future proposed trials involving pilocarpine to combat dry mouth in palliative care situations.

Another palliative care situation in which N-of-1 trials were used to assess treatment in cancer patients was the use of Methylphenidate hydrochloride (MPH) to treat fatigue (Mitchell et al, 2015). Fatigue is a common side effect of cancer and subsequent treatment and can affect the patient's quality of life. Sixty to ninety percent of patients with advanced cancer suffer from fatigue. This group examined the effect of MPH, a central nervous system stimulant, on reversing fatigue. MPH stimulates the central nervous system by increasing extracellular dopamine levels through blocking the dopamine transporter in the presynaptic membrane. There have been previous studies involving MPH as a treatment for fatigue, some showing improvements and others showing no effect. Due to a small sample size, RCTs are difficult to conduct for this population. Thus, Mitchell et al explored the use of N-of-1 trials to evaluate the use of MPH to treat fatigue in patients with advanced cancer. While data are generated for each individual trial, aggregation of data from all trials can occur in order to determine an effect on a population level, thereby becoming more similar to a RCT. This N-of-1 study was conducted to determine whether there was a population effect of treatment with MPH compared with placebo in decreasing fatigue in patients with advanced cancer. The assessment of the feasibility of conducting the N-of-1 trials in a palliative care situation was a secondary objective (Mitchell et al, 2015).

The trial was designed to compare MPH treatment to placebo in a double-blind, crossover fashion (Mitchell et al, 2015). Three cycles of drug administration periods occurred, with each drug being administered for 3 days per cycle. Subjects were enrolled

in the trial if they were a minimum of 18 years old, had a diagnosis of advanced cancer, and a screening fatigue score of at least 4 out of 10 and had a stable treatment regimen that was unlikely to influence fatigue. Treatment order was randomized for each N-of-1 trial and fatigue was measured via a daily diary using the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) subscale and the Wu Cancer Fatigue Scale (WCFS). Subjects were defined as responders if MPH administration was favored over placebo in all three cycles. In order to have the study powered similarly to what would be necessary for a RCT to achieve statistical significance, a minimum of 21 subjects were needed to complete the trials. While 43 patients were recruited, due to withdrawals, only 24 completed all three cycles, thereby having enough patients enrolled to provide statistical significance (Figure 3).

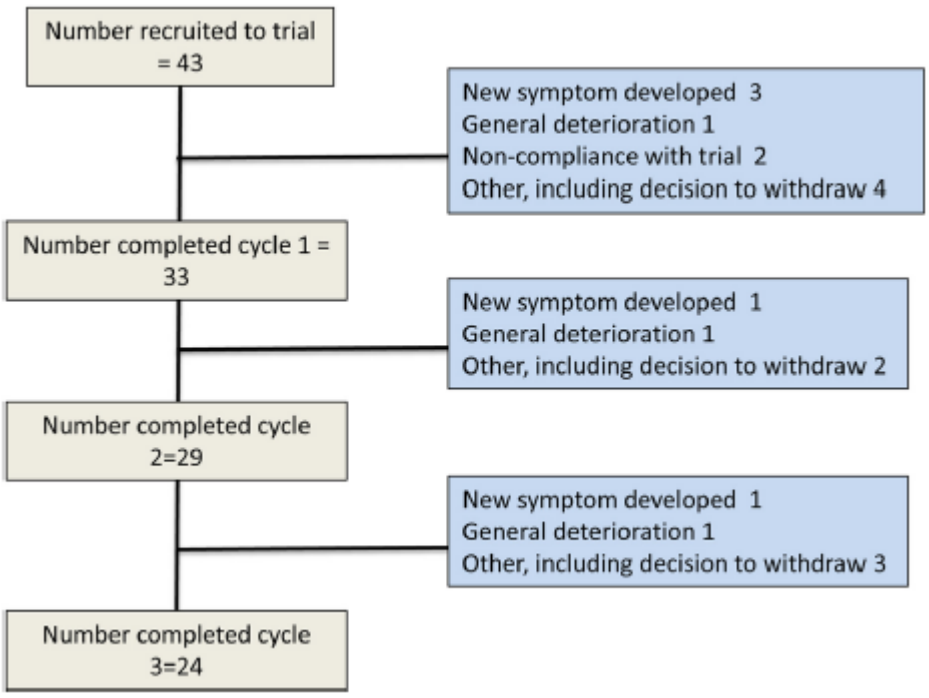

Figure 3: Withdrawal schema for aggregated N-of-1 study

Withdrawal schema for the aggregated N-of-1 study to determine the effect of MPH as a treatment for fatigue in cancer patients (Mitchell et al, 2015).

Data were analyzed two ways: the first was to assess treatment responses for individual subjects based on the FACIT-F subscale, while the second was to analyze a possible population effect by aggregating results from all N-of-1 trial patients (Mitchell et al, 2015). This aggregated analysis was performed using Bayesian statistical methods. A positive response was determined from a population perspective when "the posterior probability of the mean difference favoring the treatment exceeded 0.975" (Mitchell et al, 2015). Likewise, a negative response was determined when this value was less than 0.25 (Mitchell et al, 2015).

From a population standpoint, no difference between MPH and placebo was seen using either the FACIT-F scale or the WCFS scale (Mitchell et al, 2015). However, eight subjects did see an individual improvement with both scores, seven saw an improvement with only the FACIT-F score, and three participants saw an improvement in fatigue using only the WCFS score. Only one subject experienced a decrease in symptoms using both scales (Mitchell et al, 2015).

While this trial was able to demonstrate that MPH does not have an effect on fatigue for patients with advanced cancer as compared to placebo at the population level, it does reveal that MPH was effective for certain individuals. It also demonstrates that a population level analysis can be performed with aggregated data from N-of-1 clinical trials and can reach the same level of power generated by a RCT. This type of trial and analysis demonstrates the application of N-of-1 trials in clinical development or for the

purposes of off-label studies and validates their use as compared to RCTs in terms of statistical power.

Recently, an N-of-1 study was used during the clinical development process of a drug that initially gained market approval in 2012 and was approved for expanded use in more specific patient populations in 2017. In September 2012, an N-of-1 study was initiated by Vertex Pharmaceuticals in Cambridge, MA to study the effects of ivacaftor on lung function in patients with cystic fibrosis, residual CFTR (cystic fibrosis transmembrane conductance regulator) function, and a FEV1 (forced expiratory volume in one second) of greater than or equal to 40% predicted (Pilot Study, 2012). This was a Phase 2 crossover study enrolling 24 participants in a randomized, double-blind fashion. Patients received either placebo or ivacaftor over two week periods (one cycle), followed by a minimum of a four week washout period between cycles. Four study arms with different randomization schemes were included in the trial design. Most patients participated in all seven periods included in the study. An eight-week open-label period was included at the end of the study in which patients received study drug only. The full data analysis set included all participants who received at least one dose of either ivacaftor or placebo (Pilot Study, 2012). While ivacaftor (KALYDECO®) was already approved when this study began, it is a strong example of how N-of-1 trials can be used to obtain further indications for the product label.

Methods of Statistical Analysis for N-of-1 Trials: Applications at the Individual and Population Levels

Randomized controlled trials typically use frequentist, or classical, statistics for study design and data analysis. Frequentist statistics refer to hypothetical frequencies that

are predicted using an assumed statistical model (Greenland et al, 2016). This type of statistics is used to determine an observed significance level ("P" value) to evaluate the certainty or uncertainty level of a test hypothesis expressed in terms of frequency probability values. This value acts as a statistical summary of the observed data predicted by a statistical model of assumptions and measures the distance between these two values (Greenland et al, 2016). While most commonly used for the analysis of drug efficacy for RCTs, these methods of analysis are applicable for use only for trials with a certain sample size and using certain models (Wakefield, 2013). In the case of N-of-1 studies, other statistical methods are needed to analyze studies with a small sample size number and should be able to be further expanded to examine an effect at both the individual and population level. For this type of analysis, Bayesian statistics can be used.

Bayesian statistics apply probabilities to statistical problems and allow for the incorporation of prior knowledge to these probabilities (Gupta, 2012). Parameters are treated as random variables, and uncertainty surrounding them is expressed as probability distributions. Therefore, statements can be made regarding whether the parameter falls within a certain interval (Swaminathan et al, 2014). For example, in frequentist statistics, the mean either is or is not a fixed value, whereas in Bayesian statistics, the mean can be greater than a certain value (Rindskopf, 2014). Bayesian models are based upon prior distribution, or prior knowledge of a parameter. This information is used in the statistical model to determine the posterior distribution, or the resulting analysis (Eguchi et al, 2008). Mathematical methods are utilized to calculate the possibility of a future event based upon knowledge of a past event. While frequentist statistics incorporate prior knowledge at the clinical trial design stage only, Bayesian statistics use this information

during the design phase, clinical trial, and also at the analysis stage (Gupta, 2012). There are five different components of Bayesian statistics: the likelihood principle, posterior probabilities, predictive probabilities, exchangeability of trials, and decision rules. Under the likelihood principle, information from two sources is used: sample data and prior distribution. This differs from the frequentist statistics method in that only sample data is used. Posterior probabilities represent prior knowledge from previous distributions, while predictive probabilities create predictive distributions which assist with making clinical decisions and predicting clinical outcomes. Exchangeability of trials allows for the current trial to "borrow strength" or combine results from previously conducted trials in order to attain estimates of different parameters of interest for the trial at hand (Gupta, 2012). Finally, decision rules in Bayesian statistics are used to determine an action to take given the observations made.

Bayesian statistics are attractive to use for small clinical trials, as trials with a small population may not be capable of generating all evidence necessary for effective hypothesis testing as is needed with frequentist statistics (Medicine, 2001). Further, because of the small sample size, use of prior knowledge is practical as it improves the estimates for each case (Rindskopf, 2014). Bayesian statistics also provide more direct answers (Lee et al, 2012) and flexible probability statements (Rindskopf, 2014). Because of its applicability to small clinical trials, Bayesian analysis is extremely relevant to N-of-1 clinical trials and is continuing to be explored for their use.

While analyses can be performed for individual patients involved in N-of-1 studies to determine the effect of interventions on a single subject level, there is likewise a need to determine a population effect. This can be achieved through meta-analysis.

Meta-analysis is a method of providing a quantitative summary of the effect of an intervention based on pooled data analysis from multiple clinical trials (Sutton et al, 2001). There are two commonly used methods for meta-analysis: fixed effects and random effects models. The fixed effects method functions under the assumption that "each observed individual study result is estimating a common unknown overall pooled effect," while the random effects model assumes that "each individual observed study result is measuring its own unknown underlying effect, which in turn are estimating a common population mean" (Sutton et al, 2001). Because the random effects model measures effects from both the individual and population levels, both intra- and inter-study variability can be assessed and is typically used (Sutton et al, 2001). This type of analysis can be applied by combining data and analyses from multiple clinical trials in order to obtain evidence for statistical significance at the population level (Gupta, 2012). The concept of meta-analysis is easily demonstrated with use of Bayesian statistics due to exchangeability of trials; Bayesian analysis can apply prior distributions from other trials to increase validity. However, any prior information used should only occur during the planning of the trial design and should be based on empirical evidence rather than opinion so as to avoid introduction of potential bias or call into question the validity of results obtained from the trial (Gupta, 2012).

Several groups have demonstrated that meta-analysis of N-of-1 studies can be used to make population level claims from individual patient studies. By combining analyses, additional subgroups within the data can be explored along with specific patient factors contributing to the effect (Punja, Xu et al, 2016). Meta-analyses of individual N-of-1 data is important because it allows for the investigation of a group mean effect in

addition to the individual patient effect which is inherently present in the trial design, differing from RCTs whose analysis functions to treat all patients similarly and does not take into account the individual effects a drug may have on a patient.

Further, meta-analysis including both N-of-1 and RCTs was investigated by Punja et al for two interventions to treat pediatric ADHD. N-of-1 data used in this study were acquired through database searches of previously conducted studies using amphetamines and methylphenidate as treatment (Punja, Schmid et al, 2016). RCT data was obtained through a database search from a previously conducted review of amphetamine versus placebo. To perform the meta-analysis, means for both placebo and intervention were calculated for each N-of-1 trial before being combined using a random effects model to produce an overall standard error. Likewise mean of placebo and intervention were also calculated for each included study. These data were then combined under a random effects model using standard error (Punja, Schmid et al, 2016).

Data from the meta-analysis were examined as RCT data alone and RCT and N-of-1 trial data combined (Punja, Schmid et al, 2016). As a result of their analysis, the group was able to achieve higher precision in relation to their effect estimates and increase the power of the study by including more data points. While this study focused on a method of analysis for N-of-1 and RCT data, it also reinforced the importance of meta-analysis of N-of-1 data to estimate a population effect of an intervention.

Reporting N-of-1 Clinical Trial Data

Guidelines have been established for the reporting of data and results from randomized controlled trials and are widely used in publications. These guidelines, called CONSORT statements, are prepared at the conclusion of the clinical trial by statisticians.

CONSORT, or Consolidated Standards of Reporting Trials, statements are comprised of a checklist and diagram that are intended to provide information pertaining to the trial, such as design, management, data analysis, and validity of the results (Schulz et al, 2011). CONSORT statements were put into place in order to standardize and improve reporting of RCTs by suggesting the minimum amount of information required for accurate reporting (Schulz et al, 2011). These statements are only effective if the above information is reported accurately by a sponsor.

The CONSORT statement was initially developed in 1996 as a result of poor reporting of clinical trial results (Schulz et al, 2011). Since its implementation, it has been revised twice, focusing on two group, randomized, parallel studies, accounting for the most common trial type. CONSORT statements consist of two items: a flow chart describing enrollment, allocation, follow-up, and analysis (Figure 4), and a checklist of information to include for reporting/publishing purposes (Figure 5) (Schulz et al, 2011).

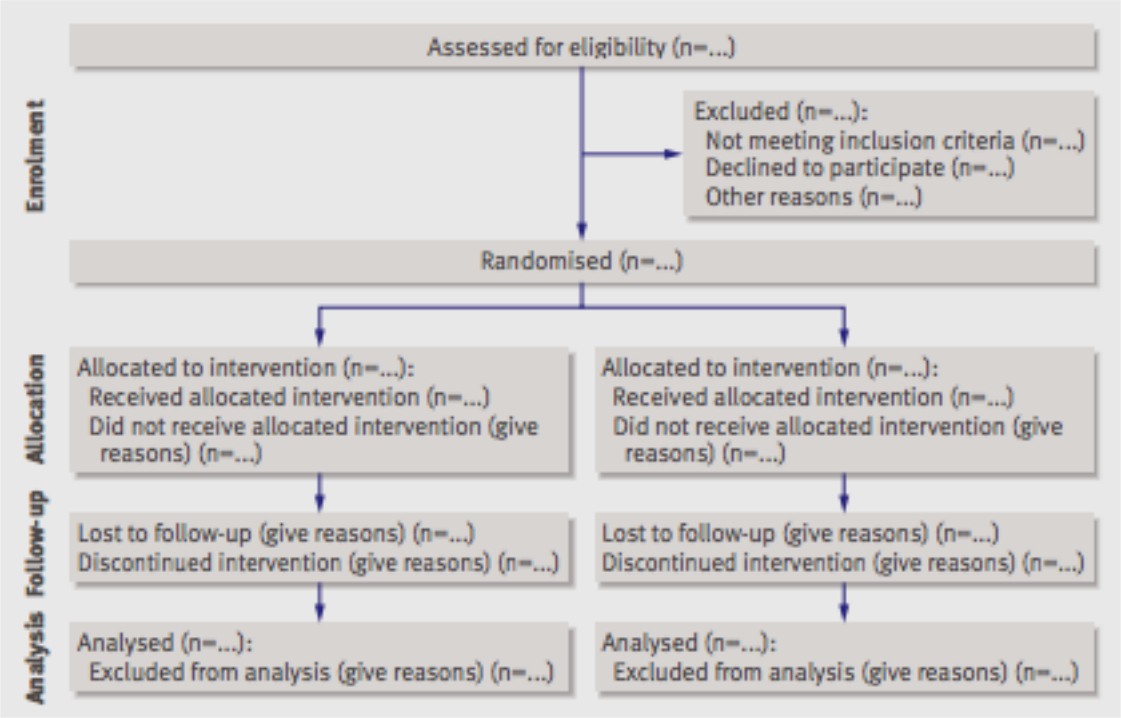

Figure 4: CONSORT statement flow diagram

Diagram outlining design, execution, and analysis of clinical trial, to be completed after trial has been completed (Schulz et al, 2011).

Because this is not the only type of clinical trial design that can be executed, reporting practices for other types of trials have been outlined in the form of extensions. In addition to trial design, extensions are also available for certain interventions (for example: non-pharmacologic treatment interventions and acupuncture) and data (for example: harms and abstracts) (CONSORT, n.d.). While there are official extension statements available, the list is constantly being developed to provide guidance where it is needed.

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item
Title and abstract		
	1a	Identification as a randomised trial in the title
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts[21 35])
Introduction		
Background and objectives	2a	Scientific background and explanation of rationale
	2b	Specific objectives or hypotheses
Methods		
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons
Participants	4a	Eligibility criteria for participants
	4b	Settings and locations where the data were collected
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed
	6b	Any changes to trial outcomes after the trial commenced, with reasons
Sample size	7a	How sample size was determined
	7b	When applicable, explanation of any interim analyses and stopping guidelines
Randomisation:		
Sequence generation	8a	Method used to generate the random allocation sequence
	8b	Type of randomisation; details of any restriction (such as blocking and block size)
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how
	11b	If relevant, description of the similarity of interventions
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses
Results		
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome
	13b	For each group, losses and exclusions after randomisation, together with reasons
Recruitment	14a	Dates defining the periods of recruitment and follow-up
	14b	Why the trial ended or was stopped
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[28])
Discussion		
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses
Generalisability	21	Generalisability (external validity, applicability) of the trial findings
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence
Other information		
Registration	23	Registration number and name of trial registry
Protocol	24	Where the full trial protocol can be accessed, if available
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration[13] for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials,[11] non-inferiority and equivalence trials,[12] non-pharmacological treatments,[32] herbal interventions,[33] and pragmatic trials.[34] Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

Figure 5: CONSORT statement checklist

CONSORT statement checklist, describing pertinent information that should be included in clinical trial reporting for consistency purposes (Schulz et al, 2011).

If individualized patient trials are to be used more frequently and used as a means for analyzing market pharmaceuticals, standards surrounding the reporting of data and trial management should be similar to the reporting of RCTs. However, to date, there has been inconsistencies in reporting of N-of-1 trials in terms of information disclosed (Vohra et al, 2015). As a result, an add-on to the CONSORT statement has been developed for N-of-1 trial reporting, called CONSORT extension for reporting N-of-1 trials (CENT) 2015. These guidelines provide direction for the complete reporting of both individual patient trials and a series of multiple N-of-1 trials (Vohra et al, 2015). Of the 25 items included in the CONSORT 2010 checklist, the CENT guidelines expands on 14. Many of the items expanded upon in the extension provide clarification regarding information pertaining specifically to N-of-1 design, such as rationale for using the design, a more specific methods section, and further details on randomization and statistical analysis methods. Changes and additions to CONSORT in the form of CENT are found below (Figures 6 and 7).

Section/Topic	CONSORT 2010 No	CONSORT 2010 Item	CENT 2015 No	CENT 2015 Item
Title and abstract				
	1a	Identification as a randomised trial in the title	1a	Identify as an "N-of-1 trial" in the title *For series:* Identify as "a series of N-of-1 trials" in the title
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	1b	For specific guidance, see CENT guidance for abstracts (table 2)
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	2a.1	
			2a.2	Rationale for using N-of-1 approach
	2b	Specific objectives or hypotheses	2b	
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	3a	Describe trial design, planned number of periods, and duration of each period (including run-in and wash out, if applicable) *In addition for series:* Whether and how the design was individualized to each participant, and explain the series design
	3b	Important changes to methods after trial start (such as eligibility criteria), with reasons	3b	
Participant(s)	4a	Eligibility criteria for participants	4a†	Diagnosis or disorder, diagnostic criteria, comorbid conditions, and concurrent therapies. *For series:* Same as CONSORT item 4a
	4b	Settings and locations where the data were collected	4b†	
			4c	Whether the trial(s) represents a Research Methods & Reporting study and if so, whether institutional ethics approval was obtained
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	5	The interventions for each period with sufficient details to allow replication, including how and when they were actually administered
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	6a.1	
			6a.2	Description and measurement properties (validity and reliability) of outcome assessment tools
	6b	Any changes to trial outcomes after the trial commenced, with reasons	6b	
Sample size	7a	How sample size was determined	7a	
	7b	When applicable, explanation of any interim analyses and stopping guidelines	7b	
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence	8a	Whether the order of treatment periods was randomised, with rationale, and method used to generate allocation sequence
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	8b	When applicable, type of randomisation; details of any restrictions (such as pairs, blocking)
			8c	Full, intended sequence of periods
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	9	
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	10	
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	11a	
	11b	If relevant, description of the similarity of interventions	11b	
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	12a	Methods used to summarize data and compare interventions for primary and secondary outcomes
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	12b	*For series:* If done, methods of quantitative synthesis of individual trial data, including subgroup analyses, adjusted analyses, and how heterogeneity between participants was assessed *(for specific guidance on reporting syntheses of multiple trials, please consult the PRISMA Statement)*
			12c	Statistical methods used to account for carryover effect, period effects, and intra-subject correlation

Figure 6a: Adaptations made to CONSORT 2010 statement

CENT 2015 for N-of-1 clinical trials was developed based upon CONSORT 2010 guidelines for randomized controlled trials. Differences between the two guidelines are shown (Vohra et al, 2015).

	CONSORT 2010		CENT 2015	
Section/Topic	**No**	**Item**	**No**	**Item**
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	13a.1	Number and sequence of periods completed, and any changes from original plan with reasons
			13a.2	*For series:* The number of participants who were enrolled, assigned to interventions, and analysed for the primary outcome
	13b	For each group, losses and exclusions after randomisation, together with reasons	13c	*For series:* Losses or exclusions of participants after treatment assignment, with reasons, and period in which this occurred, if applicable
Recruitment	14a	Dates defining the periods of recruitment and follow-up	14a†	
	14b	Why the trial ended or was stopped	14b	Whether any periods were stopped early and/or whether trial was stopped early, with reason(s).
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	15†	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	16	For each intervention, number of periods analysed. *In addition for series:* If quantitative synthesis was performed, number of trials for which data were synthesized
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	17a.1	For each primary and secondary outcome, results for each period; an accompanying figure displaying the trial data is recommended.
			17a.2	For each primary and secondary outcome, the estimated effect size and its precision (such as 95% confidence interval) *In addition for series:* If quantitative synthesis was performed, group estimates of effect and precision for each primary and secondary outcome
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	17b	
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	18	Results of any other analyses performed, including assessment of carryover effects, period effects, intra-subject correlation *In addition for series:* If done, results of subgroup or sensitivity analyses
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	19	All harms or unintended effects for each intervention. *(for specific guidance see CONSORT for harms)*
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	20	
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	21	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	22	
Other information				
Registration	23	Registration number and name of trial registry	23	
Protocol	24	Where the full trial protocol can be accessed, if available	24	
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	25	

*It is strongly recommended that this checklist be read in conjunction with the CENT 2015 Explanation and Elaboration[24] for important clarification on the items. The copyright for CENT (including checklist) is held by the CENT Group and is distributed under a Creative Commons Attribution (CC-BY 4.0) license.
†Caution should be taken when reporting potentially identifying information pertaining to CENT items 4a, 4b, 14a, and 15.

Figure 6b: Adaptations made to CONSORT 2010 statement (continued)

CENT 2015 for N-of-1 clinical trials was developed based upon CONSORT 2010 guidelines for randomized controlled trials. Differences between the two guidelines are shown (Vohra et al, 2015).

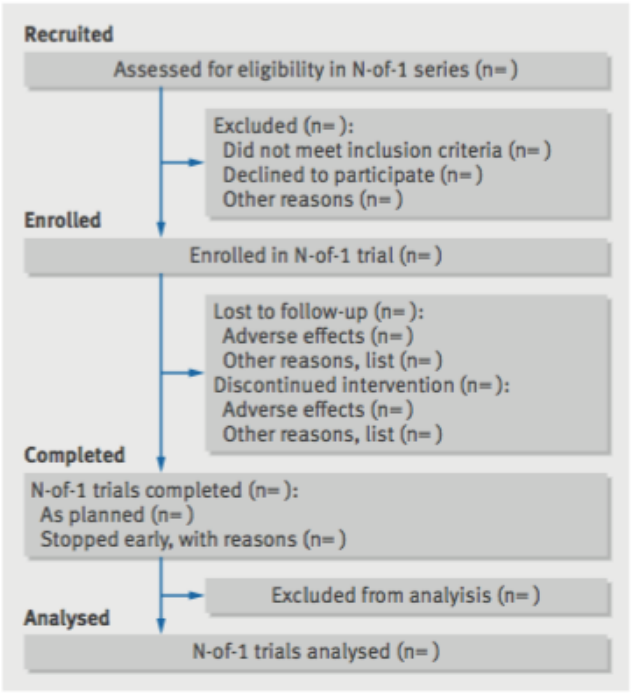

Figure 7: CONSORT CENT diagram

Example of CONSORT CENT diagram for N-of-1 trial design from Vohra et al, 2015.

Establishing reporting guidelines for N-of-1 trials is a result of the gaining of their popularity within the research community. As there is more structure and validity surrounding the trials, this helps make a stronger case for the reputability of performing them. While establishing these guidelines for individualized patient trials is not necessary for market approval, it is necessary to report quality clinical data, and the institution of this extension for N-of-1 design will allow sponsors or researchers to do so, providing the same high quality level of reporting as Randomized Controlled Trials.

Regulatory Considerations for N-of-1 Clinical Trials

Regulatory pathways for drugs investigated though use of randomized controlled trials are well established and utilized. As RCTs are the most common clinical trial, both FDA and sponsors have the most experience with drug approval using these types of trials. While this process is typically straight-forward with large trials, it becomes more challenging in the case of rare diseases. Despite a small sample number, drugs for rare diseases are still approved, yet they need to be considered differently than standard RCTs, as there are many differences. First, there is limited opportunity to conduct and perform repeatability due to small patient numbers. Another common issue with small sized clinical trials is the lack of power of a study due to small sample size. This is expressly the case with clinical trials for rare disease drugs, as the number of subjects with the disease is low and patients may be dispersed around the world, therefore causing difficulty with enrollment and the inability to attain a well powered study.

Despite these challenges, drugs for rare diseases do gain approval. The FDA has approached drugs for the rare disease space by recognizing that drug development for rare diseases may not be as straightforward as drug development for common diseases (FDA draft guidance for industry (Rare Diseases), 2015). Flexibility in their regulatory requirements for drugs for rare diseases may be provided contingent upon the trials proving effectiveness and being adequate and well-controlled (FDA guidance for industry (Rare Diseases), 2015). While there are not currently any regulatory guidances for N-of-1 trials, they could be developed if there is a need. The FDA has provided guidances for trials with small populations and outlines important factors for filing INDs and conducting trials. Well-designed, adequate clinical trials should always be conducted in order to be considered legitimate, even for trials with small populations. Prior to

submitting an IND, as much knowledge about the disease as possible should be gathered though natural history studies and translational research (Pariser, 2014).

Another type of trial for which the FDA provides guidance is adaptive design clinical trials. Adaptive design clinical trials allow for changes to occur to at least one aspect of trial design as a result of interim analyses of the clinical trial data. Alterations to trial design, such as dosing schedule or amount of dose, can occur if necessary. The need for adaptive clinical trial design came about in order to enhance study efficacy in a regulated manner (FDA draft guidance for industry (Adaptive Trials), 2010). The adaptive trial must be prospectively designed, outlining when interim analyses will occur and whether they will be blinded or unblinded. The prospective nature of the adaptive design minimizes the potential for introduction of bias into the study by an unblinded analyst. Thus, potential changes are outlined prior to unblinded analysis and no other changes to the study may take place if they were not initially included. The clinical trial protocol and statistical analysis plan (SAP) must both be written prospectively (FDA draft guidance for industry (Adaptive Trials), 2010). The adaptive design guidance was released by the FDA, as they saw a need for it for the gaining of product licensure. This is applicable to the potential use of N-of-1 studies, as it exemplifies how the FDA creates guidances and adapts regulations to the needs of the field.

How N-of-1 Can Benefit Drug Approval

Santhera Pharmaceuticals is a Switzerland-based company developing medicines to address mitochondrial disorders along with other rare diseases. Specifically, they focus on developing drugs for neuromuscular diseases that do not have treatment options. Branded as Raxone, idebenone is a drug developed to treat patients with Duchenne

Muscular Distrophy (DMD), an X-linked genetic disease developing mainly in boys at age 3-5. Symptoms of DMD are progressive muscle weakness caused by the loss of dystrophin, a protein in muscle cells (Santhera Pharmaceuticals, n.d.). The loss of dystrophin ultimately leads to mitochondrial dysfunction and therefore reduced energy production in muscle cells, causing weakness and loss of muscle tissue. Respiration is affected over time as the disease progresses, leading to respiratory insufficiency, or a deterioration in respiratory function, and ultimately death. Because of this, patients suffering with DMD are more prone to respiratory infections and may need support in order to breathe (Santhera Pharmaceuticals, n.d.).

Raxone reverses decreased energy production by mitochondria by acting as an electron carrier, which provides energy to the mitochondria in the form of electrons (DuchenneConnect, 2017). The increase in energy production by mitochondria will in turn lessen pulmonary fatigue and loss of function (DuchenneConnect, 2017). While glucocorticoids are most commonly prescribed to slow the loss of pulmonary muscle weakness and function, not all patients are responsive to the treatment, and side-effects restrict their use, specifically in non-ambulatory patients (Buyse et al, 2015).

In July 2009, Santhera began a Phase 3, double-blind, randomized, placebo-controlled study examining the safety and efficacy of administration of Raxone to males ages 10-18 who suffer from DMD and had not taken glucocorticoid therapy in the prior 12 months (Phase III, 2009). Patients were enrolled at multiple sites across Europe and North America and were administered either placebo or idebenone three times daily (total daily dosing of 900mg per day). The primary endpoint of this study measured change in respiratory muscle strength from week 52 as compared to baseline using a measurement

of Peak Expiratory Flow (PEF) as percentage predicted (PEF%p) by hospital-based spirometry (Phase III, 2009). Efficacy was also evaluated at three other time points: week 13, week 26, and week 39. Subjects were randomized 1:1, with 31 patients receiving idebenone and 33 receiving placebo (Buyse et al, 2015).

After 52 weeks, idebenone showed a statistically significant increase in muscle strength over placebo (Buyse et al, 2015). While a randomized controlled trial was used in this instance, an N-of-1 design could have been applied to this case. By enrolling subjects into individual trials, a direct comparison of drug and placebo could have been made for each subject, eliminating the need for half of the trial population to receive placebo only and not benefit from the administration of a drug. Ethically, the duration of administration of placebo in this case causes concern due to the advanced state of the children suffering from DMD (Buyse et al, 2015). Further, sample size was a limitation of the study, as the patient population selected was comprised of children who were not taking glucocorticoids as treatment. By using Bayesian statistical analyses and repetition of treatment within an individual, a smaller sample size would be able to power the study sufficiently enough due to multiple measurements made in each individual. Because the effects of the drug are short-lasting, it makes for a good candidate for observation in N-of-1 as there can be a short washout period. While this RCT did prove to be successful in demonstrating an effect of Raxone, individual patient trials could have been a more ethically appealing alternative.

Re-introduction of Banned Drugs to the Market: the Role of N-of-1

There are instances where FDA removes drugs from the market due to safety issues. However, there have not been many occasions of drugs being reinstated after

being removed. In 2002, the first drug to be placed back on the market by FDA was Lotronex (alosetron hydrochloride), a drug marketed by GlaxoSmith Kline (GSK) to treat severe irritable bowel syndrome in women who suffered from intense diarrhea symptoms (Overview of Selected Issues, 2002). On February 9, 2000, Lotronex was introduced to the market. Despite its projections to be a blockbuster drug, within eight to nine months, the drug was voluntarily removed from the market by its sponsor due to safety concerns. Patients were reporting severe complications, including constipation and ischemic colitis, and, as a result, deaths occurred. Soon after the removal of this product, patients began expressing the necessity for the drug to control the symptoms of their condition, and on April 9, 2001, a petition was brought to the FDA by the Lotronex Action Group to have it brought back to market (Overview of Selected Issues, 2002). One year later, the FDA called an advisory committee meeting to discuss the safety and risk-management and advise them on whether the drug should be reintroduced to the market.

After Lotronex was removed from the market, new clinical data, along with adverse events, from ongoing or discontinued trials were collected and presented to the FDA advisory panel in an effort to convey all information to make the best decision for the future of the drug and patients. FDA advisory panels provide FDA with advice from independent experts in the field regarding approvals or issues with FDA regulated products (What is an FDA Advisory Committee, 2017). Advisory panels are employed upon the FDA's discretion when they believe it is in the best interest of the public to have an open forum relating to the drug or medical device in question (Rettig, 1992). The advisory committee, generally comprised of a chair and other general members along with a consumer, industry member, and occasionally a patient representative, functions to

supplement FDA expertise with external experts in the field and to assist the Agency in remaining current with trends and technologies in the industry, thereby allowing for a valid and credible decision process for a drug or device (Rettig, 1992). While the committee can provide recommendations to the FDA, decisions are ultimately left to the Agency (What is an FDA Advisory Committee, 2017). Advisory committees are typically consulted by the FDA for drug approvals, and in this instance, due to original data presented, the advisory committee initially unanimously recommended the approval of Lotronex (Overview of selected issues, 2002).

After withdrawing Lotronex from the market, there were many discussions between GSK and FDA regarding the reintroduction of the drug to certain IBS patients. A second advisory committee meeting was held in 2002, during which new data was presented. It was agreed upon by both parties to attempt reintroduction of the drug with restricted market access through the review of clinical data from the trials ongoing when the drug was removed from the market. While it was beneficial that there were continuing trials from which data could be collected for further analysis, an alternative to this could have been the use of N-of-1 clinical trials to determine whether Lotronex works for patients. It was unknown why some patients responded to the drug poorly and some did not. A study such as this could have helped with re-introduction by allowing both the doctor and patient to be involved in the administration of medication and assessment of potential side effects or efficacy. The FDA advisory panel meeting resulted in recommending the reinstatement of the drug. N-of-1 trials could have helped inform them as well, and could also serve a purpose for post-marketing approval. Ultimately, the drug was brought back to market with the restrictions that doctors were required to enroll

in a program by GSK that required them to self-attest that they knew how to prescribe the medication and that patients were in complete understanding of the potential side effects.

Other Applications: N-of-1 in Oncology

With the increase in popularity of personalized medicine, new advances have been made, particularly in the oncology field. It is becoming more and more recognized that, due to the stratification of patients into different classifications of cancer subtypes, there will be increased difficulty in enrolling enough subjects in trials in order to have sufficient power and prove efficacy. The rise of personalized medicine calls for clinical trials that address the individual directly rather than through generalization. Genetic screening of tumors is becoming a more routine practice and is essential to determining the correct treatment for an individual. While this important development is changing the way treatment decisions are made, it is further complicating and stratifying the patient population, thus making it more challenging in terms of trial design and enrollment.

Because of these reasons, there is an interest in exploring N-of-1 clinical trials in the oncology field. N-of-1 can be used in this case as they were designed, comparing standard of care (or placebo) with a new drug of interest. These treatment options could be altered, one after the other, with washout periods in between. If there is a surrogate endpoint identified, this can be the endpoint measured for each treatment. However, measuring time to subsequent progression of the disease is also being explored (Markman, 2016).

N-of-1 trials are also being utilized in a less regimented manner in regards to cancer. Whereas traditionally they are designed to alternate between two or more treatments in order to measure effects, they can additionally be thought of as simply an

assessment of the drug on an individual (without a comparator). Due to the nature of oncology trials, adaptations are made when N-of-1 design is used for cancer patients (Collette et al, 2015). While treatments may be alternated, a treatment would not be repeated if it was initially found to not work. Ethically, if a treatment is working for a patient, it would be challenging for a clinician to justify removing the patient from that treatment (Collette et al, 2015). Further, repeatability of drugs may not be supported, as it would not be ethical to administer a drug to this population that is known not to work. Moreover, randomization and fixed duration of treatment may be excluded from trial design criteria, as they are often not included in oncology trial design due to the severity of the patient's disease state. Despite these challenges, N-of-1 trials may be used more frequently as the push for personalized medicine evolves.

Economical Assessment of N-of-1 Trials

The economical impact of N-of-1 clinical trials has not been assessed in great depth. However, it has been examined to some extent in the context of clinical practice. To evaluate the economic impact of healthcare interventions, cost-effective analyses (CEAs) are typically performed (Nikles and Mitchell, 2015). A cost-effective analysis is a method of evaluating health interventions for potential benefits as compared to the cost of the intervention (Jameson et al, 2006). This type of analysis is beneficial in that it establishes a relationship between the financial and scientific aspect of interventions. The cost-effectiveness calculation incorporates "dividing the cost of an intervention in monetary units by the expected health gain measured in natural units" (Jameson et al, 2006). While a health gain could be measured in terms of number of lives saved, other examples of health measurements are quality-adjusted life years (QALY) or disability-

adjusted life year (DALY). Depending on the intervention and the goal it is trying to achieve, the health gain selected can greatly affect the outcome of the analysis (Nikles and Mitchell, 2015).

Cost-effectiveness analyses can be performed for both N-of-1 trials and traditional trials in order to determine if the costs associated with the intervention are supported by the health benefits. When performing a cost-effectiveness analysis for a pharmaceutical, costs per patient and per trial phase are taken into account (Nikles and Mitchell, 2015). There are slight differences in cost-effectiveness calculations between conventional clinical trials and N-of-1 trials, mainly due to the fact that N-of-1 trials consider each intervention arm for each patient. Thus, the number of times a treatment is administered is compiled and accounted for within the analysis. Incremental analyses are completed in order to determine the difference in total cost between interventions (Nikles and Mitchell, 2015).

There are typically three metrics used to complete an incremental cost-effectiveness analysis for health interventions: incremental cost, incremental effect, and incremental cost effectiveness ratio (Nikles and Mitchell, 2015). Incremental cost compares the cost of the intervention to the cost of the comparator while incremental effect compares the effect of the intervention to that of the comparator. These are both important metrics that need to be assessed. Finally, the incremental cost-effectiveness ratio (ICER) is calculated to assess the additional cost per unit of the intended outcome. While this analysis can be performed for both RCTs and N-of-1 studies, there are differences in methods. Similar to statistical analyses for clinical trial data, cost-effectiveness analyses for RCTs are performed using the mean differences in costs and

health outcomes. However, for N-of-1 studies, the analysis is completed for each intervention arm for each individual patient. In cases where multiple crossovers occur, these results can be summed for each individual. Likewise, ICER calculations are similar in that they relate to the individual (Nikles and Mitchell, 2015). A complete list of all equations can be found in Table 1.

Table 1: Cost-effectiveness analysis equations

Equation	RCT	N-of-1
Incremental Cost	$Cost_{Int} - Cost_{Comp}$	$\frac{\Sigma Cost_{Int}}{n_{Int}} - \frac{\Sigma Cost_{Comp}}{n_{comp}}$
Incremental Effect	$Effect_{Int} - Effect_{Comp}$	$\frac{\Sigma Effect_{Int}}{n_{Int}} - \frac{\Sigma Effect_{Comp}}{n_{Comp}}$
ICER	$\frac{\Delta Cost}{\Delta Effect}$	$\frac{\Delta Cost_i}{\Delta Effect_i}$

Cost-effectiveness analysis equations for randomized controlled trials (RCTs) and N-of-1 clinical trials adapted from Nikles and Mitchell, 2015.

For decision making purposes, N-of-1 decision criteria revolve around the individual patient as opposed to the group and analyzes whether incremental gains are attainable at an acceptable level of additional cost to the patient (Nikles and Mitchell, 2015). This can then be applied to the entire cohort of patients to determine the proportion of patients whose benefit was achieved at an acceptable cost. Traditional RCT cost effectiveness decisions are examined by way of whether the cost per unit effect is

below an acceptable level (typically the cost-effectiveness threshold) that the decision maker is willing to pay (Nikles and Mitchell, 2015).

In analyses for both applications, uncertainty levels need to be examined, quantified, and applied (Nikles and Mitchell, 2015). For randomized controlled trials, uncertainty is indicated in the mean and 95% confidence intervals calculated for average patient response. Following the trend for individualized patient analyses, uncertainty is with respect to the ICER estimate for each individual. However, multiple crossovers are needed for N-of-1 studies in order to achieve sufficient data points to estimate confidence levels for each individual in the trial (Nikles and Mitchell, 2015). Overall, similar to RCTs, interventions evaluated through the use of N-of-1 trials can be assessed for their economic impact on a per patient basis to establish benefits of the treatment from a cost perspective.

Chapter IV.

Discussion

The Food and Drug Administration has been a reputable regulatory body since its institution. Without this agency, there would be a lack of regulation for pharmaceuticals, medical devices, food, and veterinary products. From inception, randomized controlled trials have been an established means of assessing safety and efficacy for pharmaceuticals, and without them, there would be a lack of quality for these products. Despite the fact that RCTs are so well established and are considered the gold standard for assessment of a pharmaceutical intervention, there are shortcomings associated with them that cause weaknesses in their implementation.

One benefit of RCTs is that they are typically sufficiently powered if enrollment specifications are achieved. However, in the rare disease space or for a rare patient population, it is often difficult to enroll enough patients for a well-powered study. In this event, it is challenging to generate meaningful data that can be applied to the entire patient population due to lack of statistical power. Further, there is a growing trend towards personalized medicine, where treatments are focused on the individual and are tailored towards each individual patient. It is becoming a more common belief that personalized medicine is critical for future development of drugs, as more and more treatments are being developed and administered based on genetic profiles. This drive for personalization elicits the need for clinical trials that are designed to assess responses from the individual rather than a group as a whole; the need for external validity is great

if personalized medicine will become conventional in the future. For this to happen, clinical trials need to be designed in such a way that their focus is on the response of an individual rather than a group of people. N-of-1 trials are the type of clinical trial design that is needed, as they can be customized to the individual patient and can assess treatment options for individuals, yet can also be generalized to the overall patient population for that specific disease. Major areas of importance in the reputability of trial design in the context of N-of-1 clinical trials have been explored, and while this type of trial design has not gained substantial popularity thus far, it could play a very important role in the future of pharmaceutical approval.

Because N-of-1 trials have been utilized in a non-pharmaceutical setting, most aspects of trial design have already been established and can be applied seamlessly to the pharmaceutical industry. Many properties of randomized controlled trials have been incorporated into the design of N-of-1 trials to ensure that they are adequate and well-controlled, which is a requirement for good clinical trial design. Randomization and blinding still occur despite each trial only enrolling one patient. While washout periods can be included to ensure remnants of one medication do not alter the effects of the next, they are not always necessary, as the half-life of interventions used in N-of-1 is short. Further, reporting of N-of-1 data has been given CONSORT guidelines in the form of extension guidelines to ensure the same quality data that is reported for RCTs is reported for N-of-1. Consistency in this reporting will further strengthen the validity of N-of-1 trials and will help them to become more reputable in the field.

There are, however, limitations to their use. These trials are not suitable for acute conditions or diseases, as they do not remain stable enough to switch back and forth

between treatments. They also are not suitable for treatments that will alter the disease state of the individual. In both instances, there would be no way to accurately measure the effects of more than one intervention if the disease state was constantly changing. Further, N-of-1 trials are not ideal for pharmaceuticals with long lasting effects, as determining which medication was causing an effect would be challenging. While washout periods could be used, these could end up being lengthy, therefore extending the overall length of the trial. One advantage of N-of-1 trials is that they provide results to patients faster than if they were participating in a RCT. In order to effectively treat the individual, a quick turnaround time is desirable. Finally, N-of-1 trials must be designed properly with interventions repeated in an individual subject enough times to attain proper statistical power in order for results to be considered reputable.

From a statistical perspective, N-of-1 trials can be accurately analyzed through Bayesian statistics. The use of this type of statistics considers prior knowledge that can be applied to the design and analysis, therefore strengthening the strategy and conclusions of the study. Bayesian statistics also consider small sample sizes, making it an ideal statistical method for N-of-1 clinical trials. Importantly, data generated from individual trials can be aggregated and meta-analyzed to provide a generalized effect at the population level, providing studies with strong external validity but also demonstrating effect claims similar to RCTs.

Despite having multi-purpose functions, the primary focus using Bayes theorem should be on analyzing the data at the individual patient level prior to being expanded to a population, otherwise there is not a strong case for performing N-of-1 trials over RCTs. While there are many different analysis models, investigators using statistical analysis for

clinical trial data should be cognizant of determining a statistical analysis method prior to conducting the trial so as to keep internal validity high; any changes performed during the clinical trial could affect accuracy of the data. Further, Bayesian statistics are currently used for adaptive clinical trials, as these trials incorporate many changes throughout their occurrence. Bayesian statistics are important in this case, as prior distributions are generated throughout the course of the trial, and these distributions can influence the future design of the trial. As there is a precedent for using Bayesian statistics for adaptive trials, it can seamlessly be employed for N-of-1 design due to their similarities. Because of the increased trend of examining data for an individual patient and the strength of using Bayesian analysis, this method of evaluation should be used for analysis of N-of-1 clinical trials when possible and should be considered an acceptable method of testing to prove market achievability for pharmaceutical products.

In terms of regulatory, the FDA does not currently have guidances for conducting N-of-1 trials or for using them for drug approval purposes. However, it is clear that the FDA is adapting and updating their guidances and tolerability for different types of studies that are non-traditional. Changes and additions to FDAs regulations and guidelines represent an evolution of change in the industry which has been present since its commencement. Because of this flexibility within the Agency, it would be possible to draft and release a guidance to use N-of-1 trials as part of clinical study phases. Single patient INDs are already included in the Code of Federal Regulations, thus representing the Agency's acknowledgement of the importance of individual patient studies. As this type of trial gains more popularity, the involvement of FDA in regards to design and implementation will increase and there may be a need to release a guidance. This need

will become critical in order to ensure adequate, well-designed N-of-1 studies are conducted in order to attain market approval.

While there are still structures that need to be established in order for N-of-1 trials to be viewed as suitable clinical trials for drug approval, there are instances where they can be used. N-of-1 trials are not at the point where they alone can be used to achieve market approval of a drug for all circumstances; however, the trend could be moving in that direction. Their use in the rare disease and small patient population spaces will most likely increase as more and more trial sponsors realize their power and benefits. The N-of-1 study used to support the approval of additional indications for KALYDECO® (ivacaftor) by Vertex Pharmaceuticals is representative of how N-of-1 trials can be integrated into the pharmaceutical development process. This study enrolled a small number of participants and achieved an adequate, well-controlled N-of-1 design though blinding and randomization along with statistical power.

Aside from playing a role in obtaining market approval for pharmaceuticals, N-of-1 clinical trials can be used in other applications. They could act as a powerful tool to influence additional clinical claims or new clinical indications from a drug that is already on the market. Being that they are small studies, they could be used for Phase 4 post-marketing approval studies to determine whether a treatment will work for a new indication before proceeding to an expensive randomized controlled trial. They can also be conducted to be used as supporting evidence for FDA advisory committee meetings to strengthen the advisory panel's recommendations of drug approval or usage. This could be very beneficial, as they would provide the panel with data regarding individuals, which could be an important factor in their decision process. Further, comparator studies

may benefit from the use of N-of-1 trial design, as the individual patient will be comparing both interventions and a direct comparison can be made with the same individual. N-of-1 studies could also be influential in the approval of generic drugs. While clinical studies are not typically needed for generics, there are cases where generics do need to be tested for efficacy purposes, and a study such as N-of-1 could be the solution for this issue.

Finally, in addition to benefitting the drug approval and marketing process, N-of-1 clinical trials could serve to help restore public trust in the pharmaceutical industry. A portion of the public believes that the pharmaceutical industry is only trying to make a profit. While the cost of drugs can be expensive, this is not the image the industry desires. Moreover, the fact that many drugs do not work for countless patients does not benefit the industry. N-of-1 trials can restore faith in pharmaceutical companies because the patient is directly involved. This has been established through the use of N-of-1 design by doctors in an attempt to find the best treatment possible for a patient. The patient feels that he is involved in the process and ultimately has the ability to provide input for the treatment he will receive. N-of-1 design of clinical trials can serve as a means of regaining pubic trust in the industry by demonstrating their power of treating and drawing conclusions regarding the patient as an individual.

N-of-1 clinical trials are necessary to explore for the use of approving and marketing pharmaceutical products. Not only can they assist in this process, they can also help strengthen the public image of the industry. While they may not be able to be used on their own entirely at this point in time depending on the disease or indication in question, they should be investigated further in the pharmaceutical setting to more

concretely establish their potential benefit to the industry. With the increase of personalized medicine, the use of personalized clinical trials should be justified, and while this may not currently be entirely feasible, the industry should trend towards this for the future.

References

About FDA: History. (2015). Retrieved from https://www.fda.gov/aboutFDA/WhatWeDo/History/

Akobeng, A. K. (2005). Understanding randomised controlled trials. BMJ Journals. Retrieved from http://adc.bmj.com/content/90/8/840

Arehart, E., Stitham, J., Asselbergs, F. W., Douville, K., MacKenzie, T., Fetalvero, K. M., … Hwa, J. (2008). Acceleration of cardiovascular disease by a dysfunctional prostacyclin receptor mutation, potential implications for COX-2 inhibition. *Circulation Research*, *102*(8), 986–993. http://doi.org/10.1161/CIRCRESAHA.107.165936

Augustine, E. F., Adams, H. R., & Mink, J. W. (2013). Clinical Trials in Rare Disease: Challenges and Opportunities. Journal of Child Neurology, 28(9), 1142–1150. http://doi.org/10.1177/0883073813495959

Biopharmaceutical Industry-Sponsored Clinical Trials: Impact on State Economies. (2015) (1st ed., p. 6). Retrieved from http://phrma-docs.phrma.org/sites/default/files/pdf/biopharmaceutical-industry-sponsored-clinical-trials-impact-on-state-economies.pdf

Bothwell, L. E., Greene, J. A., Podolsky, S. H., & Jones, D. S. (2016). Assessing the gold standard — lessons from the history of RCTs. *N Engl J Med, 374*(22), 2175-2181. doi:10.1056/NEJMms1604593

Brody, Tom. (2016) *Clinical Trials.* San Diego, CA: Elsevier.

Buyse, G. M., Voit, T., Schara, U., Straathof, C. S. M., D'Angelo, M. G., Bernert, G., . . . Meier, T. (2015). Efficacy of idebenone on respiratory function in patients with Duchenne muscular dystrophy not using glucocorticoids (DELOS): a double-blind randomised placebo-controlled phase 3 trial. *Lancet, 385*(9979), 1748-1757. doi:10.1016/s0140-6736(15)60025-3

Center for Drug Evaluation and Research. (2005). Postmarket Drug Safety Information for Patients and Providers - Questions and Answers FDA Regulatory Actions for the COX-2 Selective and Non-Selective Non-Steroidal Anti-inflammatory drugs (NSAIDs). Retrieved May 16, 2017, from https://www.fda.gov/drugs/drugsafety/postmarketdrugsafetyinformationforpatientsandproviders/ucm106148.htm

Center for Drug Evaluation and Research. (2016). Information for Consumers (Drugs) - Orphan Products: Hope for People With Rare Diseases. Retrieved May 17, 2017, from https://www.fda.gov/Drugs/ResourcesForYou/Consumers/ucm143563.htm.

Chen X, Chen P (2014). A Comparison of Four Methods for the Analysis of N-of-1 Trials. PLoS ONE 9(2): e87752. doi:10.1371/journal.pone.0087752

Clay, R. A. (2010, September). More than one way to measure. Retrieved May 17, 2017, from http://www.apa.org/monitor/2010/09/trials.aspx

Collette, L., & Tombal, B. (2015). N-of-1 trials in oncology. *Lancet Oncol, 16*(8), 885-886. doi:10.1016/S1470-2045(15)00062-5

Commissioner, O. O. (2013). Designating an Orphan Product: Drug and Biological Products - Orphan Drug Act - Relevant Excerpts. Retrieved May 17, 2017, from https://www.fda.gov/forindustry/developingproductsforrarediseasesconditions/howtoapplyfororphanproductdesignation/ucm364750.htm

CONSORT Transparent Reporting of Trials. (n.d.). Retrieved from from http://www.consort-statement.org/

Dean, L. Celecoxib Therapy and CYP2C9 Genotype. (2016). In: Pratt V, McLeod H, Dean L, et al., editors. Medical Genetics Summaries [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2012-. Available from: https://

DuchenneConnect. (2017). RAXONE® - Phase 3 Study Assessing the Efficacy, Safety and Tolerability of Idebenone in Patients with Duchenne Receiving Glucocorticoid Steroids (SIDEROS). Retrieved from https://www.duchenneconnect.org/clinical-trials/research-faqs/652-raxone-phase-3-study-assessing-the-efficacy-safety-and-tolerability-of-idebenone-in-patients-with-duchenne-muscular-dystrophy-receiving-glucocorticoid-steroids-sideros.html

Eguchi, T. (2008). An Introduction to Bayesian Statistics Without Using Equations. Marine Turtle Newsletter, 122, 1-5.

FDA draft guidance for industry. (2010) Adaptive Design Clinical Trials for Drugs and Biologics

FDA draft guidance for industry. (2015). Rare Diseases: Common Issues in Drug Development.

Food and Drug Administration. (2016) Code of Federal Regulations, Title 21, Volume 5, Part 312.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power:

a guide to misinterpretations. *Eur J Epidemiol, 31*(4), 337-350. doi:10.1007/s10654-016-0149-3

Gupta, S. K. (2012). Use of Bayesian statistics in drug development: Advantages and challenges. *International Journal of Applied and Basic Medical Research*, *2*(1), 3–6. http://doi.org/10.4103/2229-516X.96789

Hay, M., Thomas, D. W., Craighead, J. L., Economides, C., & Rosenthal, J. (2014). Clinical development success rates for investigational drugs. Nature Biotechnology, 32(1), 40-51. doi:10.1038/nbt.2786 [doi]

HER2 Testing for Breast Cancer. (2015). Retrieved May 17, 2017, from http://www.cancer.net/research-and-advocacy/asco-care-and-treatment-recommendations-patients/her2-testing-breast-cancer

Herper, M. (2005). David Graham On The Vioxx Verdict. Retrieved May 16, 2017, from https://www.forbes.com/2005/08/19/merck-vioxx-graham_cx_mh_0819graham.html

Higgins JPT, Green S (editors). (2011). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0. The Cochrane Collaboration. Available from www.handbook.cochrane.org.

Jamison DT, Breman JG, Measham AR, et al., editors. (2006) Priorities in Health. Washington (DC): The International Bank for Reconstruction and Development / The World Bank; Chapter 3, Cost-Effectiveness Analysis. Available from: https://www.ncbi.nlm.nih.gov/books/NBK10253/

Junod, S. W., Ph.D. (n.d.). Overviews on FDA History - FDA and Clinical Drug Trials: A Short History. Retrieved May 17, 2017, from https://www.fda.gov/aboutfda/whatwedo/history/overviews/ucm304485.htm

Kinch, M. S., Haynesworth, A., Kinch, S. L., & Hoyer, D. (2014). An overview of FDA-approved new molecular entities: 1827–2013. Drug Discovery Today, 19(8), 1033-1039. doi:http://dx.doi.org.ezp-prod1.hul.harvard.edu/10.1016/j.drudis.2014.03.018

Kravitz RL, Duan N, eds, and the DEcIDE Methods Center N-of-1 Guidance Panel (Duan N, Eslick I, Gabler NB, Kaplan HC, Kravitz RL, Larson EB, Pace WD, Schmid CH, Sim I, Vohra S) (2014). Design and Implementation of N-of-1 Trials: A User's Guide. AHRQ Publication No. 13(14)-EHC122-EF. Rockville, MD: Agency for Healthcare Research and Quality; January. http://www.effectivehealthcare.ahrq.gov/N-1-Trials.cfm.

Krumholz, H. M., Ross, J. S., Presler, A. H., & Egilman, D. S. (2007). What have we learnt from Vioxx? . *BMJ : British Medical Journal*, *334*(7585), 120–123. http://doi.org/10.1136/bmj.39024.487720.68

Lee, J. J., & Chu, C. T. (2012). Bayesian clinical trials in action. *Stat Med, 31*(25), 2955-2972. doi:10.1002/sim.5404

Lipka, A. F., Vrinten, C., van Zwet, E. W., Schimmel, K. J., Cornel, M. C., Kuijpers, M. R., . . . Verschuuren, J. J. (2017). Ephedrine treatment for autoimmune myasthenia gravis. *Neuromuscul Disord, 27*(3), 259-265. doi:10.1016/j.nmd.2016.11.009

Lillie, E. O., Patay, B., Diamant, J., Issell, B., Topol, E. J., & Schork, N. J. (2011). The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? Personalized Medicine, 8(2), 161–173. http://doi.org/10.2217/pme.11.7

Lipsky MS, Sharp LK. (2001). From idea to market: the drug approval process. J Am Board Fam Pract.; 14(5):362-7. Review. PubMed PMID: 11572541.

Mandal, A. (2015). History of Thalidomide. Retrieved May 17, 2017, from http://www.news-medical.net/health/History-of-Thalidomide.aspx

Markman, M., Kramer, K., Alvarez, R. H., Weiss, G. J., Ahn, E., & Daneker, G. W. (2016). Evaluating the Utility of a 'N-of-1' Precision Cancer Medicine Strategy: The Case for 'Time-to-Subsequent-Disease Progression'. *Oncology, 91*(6), 299-301. doi:10.1159/000450682

Medicine, I. o. (2001). *Small Clinical Trials: Issues and Challenges*. Washington, DC: The National Academies Press.

Mitchell, G. K., Hardy, J. R., Nikles, C. J., Carmont, S. A., Senior, H. E., Schluter, P. J., . . . Currow, D. C. (2015). The Effect of Methylphenidate on Fatigue in Advanced Cancer: An Aggregated N-of-1 Trial. *J Pain Symptom Manage, 50*(3), 289-296. Pannucci, C. J., & Wilkins, E. G. (2010). Identifying and Avoiding Bias in Research. *Plastic and Reconstructive Surgery*, *126*(2), 619–625. http://doi.org/10.1097/PRS.0b013e3181de24bc

Nikles and Mitchell, editors. (2015). The Essential Guide to N-of-1 trials in Health. Dordrecht, The Netherlands: Springer. doi:10.1007/978-94-017-7200-6

Nikles, J., Mitchell, G. K., Hardy, J., Agar, M., Senior, H., Carmont, S. A., . . . Currow, D. (2015). Testing pilocarpine drops for dry mouth in advanced cancer using n-of-1 trials: A feasibility study. *Palliat Med, 29*(10), 967-974. doi:10.1177/0269216315585473

Overview of Selected Issues – Lotronex tablets. (2002). Gastrointenstinal Drugs Advisory Committee Drug Safety and Risk Management Subcommittee Background Package.

Pariser, Anne. (2014). Rare Disease and Clinical Trials [PowerPoint slides]. Retrieved from https://www.fda.gov/downloads/Drugs/NewsEvents/UCM440797.pdf

Phase III Study of Idebenone in Duchenne Muscular Dystrophy (DMD) (DELOS). (2009). Retrieved from https://clinicaltrials.gov/ct2/show/NCT01027884?term=raxone&rank=4) (Identification No. NCT01027884)

Pilot Study Testing the Effect of Ivacaftor on Lung Function in Subjects With Cystic Fibrosis and Residual CFTR Function. (2012). Retrieved from https://clinicaltrials.gov/ct2/show/NCT01685801 (Identification No. NCT01685801)

Punja, S., Schmid, C. H., Hartling, L., Urichuk, L., Nikles, C. J., & Vohra, S. (2016). To meta-analyze or not to meta-analyze? A combined meta-analysis of N-of-1 trial data with RCT data on amphetamines and methylphenidate for pediatric ADHD. *J Clin Epidemiol, 76*, 76-81. doi:10.1016/j.jclinepi.2016.03.021

Punja, S., Xu, D., Schmid, C. H., Hartling, L., Urichuk, L., Nikles, C. J., & Vohra, S. (2016). N-of-1 trials can be aggregated to generate group mean treatment effects: a systematic review and meta-analysis. *J Clin Epidemiol, 76*, 65-75. doi:10.1016/j.jclinepi.2016.03.026

Rettig RA, Earley LE, Merrill RA, editors. (1992). Institute of Medicine (US) Committee to Study the Use of Advisory Committees; Food and Drug Administration Advisory Committees. Washington (DC): National Academies Press (US); 3, The FDA Advisory Committee System. Available from: https://www.ncbi.nlm.nih.gov/books/NBK236088/

Ricciotti, E., & FitzGerald, G. A. (2011). Prostaglandins and Inflammation. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *31*(5), 986–1000. http://doi.org/10.1161/ATVBAHA.110.207449

Rindskopf, D. (2014). Nonlinear Bayesian analysis for single case designs. *J Sch Psychol, 52*(2), 179-189. doi:10.1016/j.jsp.2013.12.003

Rothwell, P. M. (1995). Can overall results of clinical trials be applied to all patients? doi:http://dx.doi.org.ezp-prod1.hul.harvard.edu/10.1016/S0140-6736(95)90120-5

Santhera Pharmaceuticals. (n.d.). Duchenne muscular dystrophy. Retrieved from http://www.santhera.com/patients/our-research/duchenne-muscular-dystrophy

Schork, N. J. (2015). Personalized medicine: Time for one-person trials. Nature, 520(7549), 609-611. doi:10.1038/520609a [doi]

Schulz, K. F., Altman, D. G., & Moher, D. (2011). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Int J Surg, 9*(8), 672-677. doi:10.1016/j.ijsu.2011.09.004

Scuffham PA, Yelland MJ, Nikles J, Pietrzak E, Wilkinson D. (2008) Are N-of-1 trials an economically viable option to improve access to selected high cost medications? The Australian experience. Value Health. 11(1):97-109. doi: 10.1111/j.1524-4733.2007.00218.x. PubMed PMID: 18237364.

Sutton, A. J., & Abrams, K. R. (2001). Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res, 10*(4), 277-303. doi:10.1177/096228020101000404

Swaminathan, H., Rogers, H. J., & Horner, R. H. (2014). An effect size measure and Bayesian analysis of single-case designs. *J Sch Psychol, 52*(2), 213-230. doi:10.1016/j.jsp.2013.12.002

The American Cancer Society medical and editorial content team (2015). Off-label Drug Use. American Cancer Society. https://www.cancer.org/treatment/treatments-and-side-effects/treatment-types/chemotherapy/off-label-drug-use.html

The Drug Development Process – Step 3: Clinical Research. (2005). Retrieved May 17, 2017, from https://www.fda.gov/forpatients/approvals/drugs/ucm405622.htm

The Principles of ICH GCP. (n.d.). Retrieved May 28, 2017, from http://ichgcp.net/2-the-principles-of-ich-gcp-2

Tufts Center for the Study of Drug Development. (2016). Retrieved May 17, 2017, from http://csdd.tufts.edu/news/complete_story/tufts_csdd_rd_cost_study_now_published

Vohra, S., Shamseer, L., Sampson, M., Bukutu, C., Schmid, C. H., Tate, R., . . . Moher, D. (2016). CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. *J Clin Epidemiol, 76*, 9-17. doi:10.1016/j.jclinepi.2015.05.004

Vrinten, C., Lipka, A. F., van Zwet, E. W., Schimmel, K. J., Cornel, M. C., Kuijpers, M. R., . . . Verschuuren, J. J. (2015). Ephedrine as add-on therapy for patients with myasthenia gravis: protocol for a series of randomised, placebo-controlled n-of-1 trials. *BMJ Open, 5*(7), e007863. doi:10.1136/bmjopen-2015-007863

Wakefield, J. (2013). Frequentist Inference *Bayesian and Frequentist Regression Methods* (pp. 27-83). New York, NY: Springer New York.

Weinreich, S. S., Vrinten, C., Kuijpers, M. R., Lipka, A. F., Schimmel, K. J. M., van Zwet, E. W., . . . Cornel, M. C. (2017). Aggregated N-of-1 trials for unlicensed medicines for small populations: an assessment of a trial with ephedrine for myasthenia gravis. *Orphanet J Rare Dis, 12*(1), 88. doi:10.1186/s13023-017-0636-y

What is an FDA Advisory Committee? (2017). Retrieved from https://www.fda.gov/AboutFDA/Transparency/Basics/ucm222191.htm

Zarghi, A., & Arfaei, S. (2011). Selective COX-2 Inhibitors: A Review of Their Structure-Activity Relationships. *Iranian Journal of Pharmaceutical Research : IJPR*, *10*(4), 655–683. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3813081/