Are You "Tilting at Windmills" or Undertaking a Valid Clinical Trial?

John D. Steeves, Jose Zariffa, and John L.K. Kramer

ICORD (International Collaboration On Repair Discoveries), University of British Columbia (UBC) and Vancouver Coastal Health,
Blusson Spinal Cord Centre, Vancouver General Hospital, Vancouver, Canada.

In this review, several aspects surrounding the choice of a therapeutic intervention and the conduct of clinical trials are discussed. Some of the background for why human studies have evolved to their current state is also included. Specifically, the following questions have been addressed: 1) What criteria should be used to determine whether a scientific discovery or invention is worthy of translation to human application? 2) What recent scientific advance warrants a deeper understanding of clinical trials by everyone? 3) What are the different types and phases of a clinical trial? 4) What characteristics of a human disorder should be noted, tracked, or stratified for a clinical trial and what inclusion /exclusion criteria are important to enrolling appropriate trial subjects? 5) What are the different study designs that can be used in a clinical trial program? 6) What confounding factors can alter the accurate interpretation of clinical trial outcomes? 7) What are the success rates of clinical trials and what can we learn from previous clinical trials? 8) What are the essential principles for the conduct of valid clinical trials?

Key Words: Human, research, protocol, guidelines, translation, neurology

INTRODUCTION

"Tilting at windmills" is an old English expression, which means attacking imaginary enemies, fighting unwinnable or futile battles. The phrase is derived from a passage from the Spanish novel of Miguel de Cervantes, entitled *Don Quixote*, which was written in 1605. In Chapter VIII of Part 1, Don Quixote wishes to fight windmills that he imagines to be giants, much to the bewilderment of his faithful and more grounded companion, Sancho Panza (Fig. 1). The word "tilt" here means to fight or joust. Today, the phrase is sometimes used to describe a delusional activity where the goal (prize or adversary) is incorrectly perceived or emotionally idealized. In literature the protagonist's foolish course of action is a misguided effort to attain the prize at all cost. A comprehensive review of clinical trials could also be interpreted as 'tilting at windmills' as the topic is very complex; we are not delusional enough to think we can provide any more than a sketch of such a substantial subject.

When evaluating our scientific and clinical trial data, we have to ask ourselves as researchers whether we sometimes act emotionally and irrationally, like Don

Quixote, or do we take a more skeptical approach, like Sancho Panza. How do we guard against our intrinsic emotion blinding us to the logical and objective evaluation of a scientific discovery as a treatment for a human disorder? First, every scientist and clinician must accept that they are biased about their research, which is why regulatory guidelines have been established for the valid conduct of human studies investigating novel therapeutic interventions. No matter how passionately we pursue a goal or persistently try to persuade our colleagues that our intervention is the best treatment, the final determination of its value is in the hands of others. If the data is incomplete or the study is experimentally flawed, we will end up looking like Don Quixote tilting at windmills!

Clinical trial guidelines are not new; they have been with us in various forms since the advent of the written word. Some of the best records come from Ibn Sīnā (Avicenna) who was a famous Persian physician and philosopher. At the height of the Islamic Golden Age, in 1025, he included a detailed set of guidelines for the testing of drugs in his book, The Canon of Medicine. This work was subsequently translated into Latin and still being used as a textbook in European medical schools, 600 hundred years later. One of the earliest documented clinical trials was that of James Lind in 1747 where he used clinical comparisons between experimental and control substances to show that scurvy could

Fig. 1. Copy of the Pablo Picasso sketch of Don Quixote (1955), copyright holder unknown, but presumed to be copyrighted. First appeared in August 18-24 issue of the French weekly journal Les *Lettres Françaises* in celebration of 350th anniversary of the first part of Cervantes' novel *Don Quixote*. Used here in a completely non-profit and non-commercial manner for scholarly research and teaching purposes, under "fair use or fair dealing" exemptions.

be prevented and treated by supplementing the diet with citrus fruit such as limes or lemons, but not by other acids.[1] Parenthetically, this early clinical trial was conducted on board a ship where scurvy had taken hold of many of the crew.

Thus, valid clinical trial protocols have been around for a long time and can be conducted in even the most inhospitable circumstances. Therefore, it is probably not asking too much of anyone in the 21st century to adopt the fundamental precepts for the design of an objective clinical trial, as well as all data being analyzed in a valid and unbiased manner. This is easier said than done and a number of inadvertent mistakes can result without some initial guidance as to the best way to navigate what can be a tortuous trail. What follows is a very brief review of some of the central issues associated with clinical trials. Our review is populated with examples from a number of disorders, although illustrations from spinal cord injury (our field of interest) predominate, as it is only within this field that we can be confident that the chosen example is representative of a more universal issue.

But probably the first question to tackle is the actual need for undertaking a clinical trial in the first place. It is not unusual to hear that a good observational study, case report, or "open-label" trial can be just as valuable as a rigorous randomized controlled trial (RCT). Perhaps this is true, but a couple of difficult criteria have to be satisfied for this approach to be accepted. First, the outcome of the experimental therapy has to be strongly "face evident" (i.e. extremely obvious), such as the use of parachutes when jumping from high altitudes.[2] It is true that no one has conducted an RCT on the use of parachutes to avoid the consequences of gravity, however several controlled comparison studies have been subsequently undertaken to determine the best design of later parachutes. Second, the assignment of subjects to different groups within an observational study, the observation of outcomes, as well as any data analysis, must be free of even unintended bias by the clinicians. This is unlikely to be the case as clinicians are human and subject to unintended bias like everyone else (see above).

Thus the strongest level of clinical evidence is from multicenter, double-blind RCTs. An RCT is considered the most reliable evidence because the protocols used during the conduct of an RCT minimize the risk of confounding factors influencing the results (including investigator bias or a subject's expectation). As a result, the findings generated by RCTs are likely to be closer to the true effect than the findings generated by other research methods, including observational studies or clinical case reports.[3] Nevertheless,

there are instances when conducting an RCT may be unethical or virtually impossible, such as the case for many surgical interventions and then observational studies might be the only available approach, but these studies should still rely on blinded assessors for the measurement of outcomes and data analysis.

WHAT CRITERIA SHOULD BE USED TO DETERMINE WHETHER A SCIENTIFIC DISCOVERY OR INVENTION IS WORTHY OF TRANSLATION TO HUMAN APPLICATION?

Clinical trials are expensive, but more importantly no one wishes to waste patient resources on a scientific discovery that has little or no legitimacy going forward to human application. Thus, to generate greater confidence within the research community, preclinical research programs need to validate their discoveries through the use of high quality preclinical protocols that incorporate blinded assessments, adequate power, and "functional" outcome measures similar to those used in a human study.

Similarly, most scientists would agree that the optimal pre-clinical translation process would include independent replication of promising preclinical strategies.[4,5] Furthermore, independent replication efforts might involve the use of slightly different preclinical models (e.g., different severities of the disorder) or variations of the treatment paradigm. This would establish both the relevance and robustness of the initial discovery. In addition, finding similar beneficial outcomes of the experimental intervention in different animal species would demonstrate the fundamental nature of the therapeutic target and increase the likelihood that the treatment would also benefit humans. Finally, all to often, scientists determine an intervention has benefit when it is administered immediately after experimentally creating the disorder in the preclinical animal model (e.g., ischemic stroke). Unfortunately, patients can rarely be treated within minutes of a sudden trauma or illness. Thus, if a therapy purports to be a useful acute treatment, it should be demonstrated to have benefit in animals when administered within a clinically relevant time frame (e.g., several hours after the appearance of any acute symptoms). If an intervention is suggested to benefit people living with a chronic condition, then it should be shown to be effective in a chronic animal model.[6]

As outlined in Table 1, there are several important traits of an experimental treatment that must be established prior to

Table 1. Characterization of an Experimental Treatment Using Animal Models Prior to Human Study is Necessary and Includes Traits that Often form an Iterative Cycle of (Bench-Bedside-Bench-etc.) for the Ongoing Development of Therapies

Therapeutic trait	Approach	Outcome
Temporal "window of opportunity"	Test different time limits for provision of therapeutic strategy after experimentally creating disorder in animal model	Determine early and late time points for the therapeutic having a benefit in animal model
Formulation	Examine different: forms of drug, types of cells, or rehabilitation training programs	Identify which formulation provides best actions with minimal adverse effects
Route of administration	Investigate what is the most effective route to achieve therapeutic benefit (e.g., intravenous, intraparenchymal)	Determine best potential route for administering intervention to a human subject
Dose	Test different levels (dose response) of therapeutic (e.g., drug doses, cells numbers, or length and number of rehabilitation sessions) (also see adverse effects)	Define the optimal dose that will potentially achieve a meaningful benefit, without unwanted side effects or adverse events
Adverse effects	Determine doses that cause unwanted side effects or adverse events	Identify tolerable and maximal doses of therapeutic
Fate of drug, transplanted cells, or rehabilitation strategy	Drugs- pharmacodynamics (action of drug on body) and pharmacokinetics (actions of body on drug) absorption, diffusion, metabolism, excretion Cells- integration with host tissue, distribution, survival, tumorigenicity Rehabilitation- discover persistence of rehabilitation benefit	Outline the expected duration (length) of benefit that may be achieved through the administration of the therapy Obtain a better understanding of how the body interacts with the therapeutic intervention
Mechanism of action	Identify target or biological action of the intervention (e.g., altered biochemical pathway)	Provides information for development of subsequent (next generation) therapeutics

Rehabilitation strategies (activity-dependent training programs) should undergo similar characterization.

beginning a human study. Satisfying these requirements is often perceived as scientifically uninteresting, but without reasonable answers for these characteristics, it is doubtful that any regulatory agency or institutional review board of a hospital or university would consider a human study as safe or adequately justified. To assure objectivity and provide confidence to trial investors, many of these therapeutic traits are often tested by an independent contract research organization.

Nevertheless, some experimental interventions enter into clinical trials without being directly studied in a preclinical animal model of the human disorder. This may occur, for example, if the treatment has a history for clinical use to treat a different disorder, but involves a related therapeutic target. The advantages for such a translational path include a prior understanding of the safety and toxicology of the treatment in humans. Thus when translated to the new disorder, there is reduced risk that the intervention might result in any adverse events or detrimental side effects (Fig. 2). Nevertheless, off-label prescriptions must serve patient needs better than conventional alternatives and must be supported by human study evidence for the "new" clinical disorder (i.e., a separate clinical trial).

WHAT RECENT SCIENTIFIC ADVANCE WARRANTS A DEEPER UNDERSTANDING OF CLINICAL TRIALS BY EVERYONE?

The biomedical world changed dramatically with the confirmation that stem and progenitor cells can be found within all adult body tissues, including the brain and spinal cord. Furthermore, it was discovered that many adult human somatic cells could be induced to become a variety of pluripotent stem and progenitor cells.[7-9] After transplantation of these pluripotent cells, it is hypothesized that they can differentiate into a number of adult cell phenotypes. This promises an infinite source of cells to repair damaged tissues and organs. However, except for some hematopoietic disorders or the repair of damaged bone, a great deal of basic science still needs to be completed on the control of differentiation and transplant survival before these technologies become realistic clinical treatment options.

The public media reports for the potential wonders of cell transplants have captivated worldwide interest and generated emotion and hope in people living with a chronic physical disability. With increased access to scientific and medical information via the Internet, patients are now very well informed. Consequently, all scientists and clinicians need to answer a wider scope of questions, ranging from fundamental cell biology to the many claims for clinical benefits from a seemingly endless number of treatment options. In brief, all biomedical researchers and health care professionals must now be competent knowledge translators. This means scientists and clinicians must be able to critically evaluate the quality and strength of evidence provided by different types of preclinical discoveries and human studies, as well as their realistic prospect for clinical application.

Development of the basic cell culture technologies required for cell transplants is relatively inexpensive and involves only modest biological science expertise. Thus, there has been a rapid expansion of independent "for-profit" clinics offering such treatments. The attractiveness of these transplantation practices has understandable appeal to desperate patients when presented as a "cure", but the suppliers are currently exploiting their hopes for financial reward. The
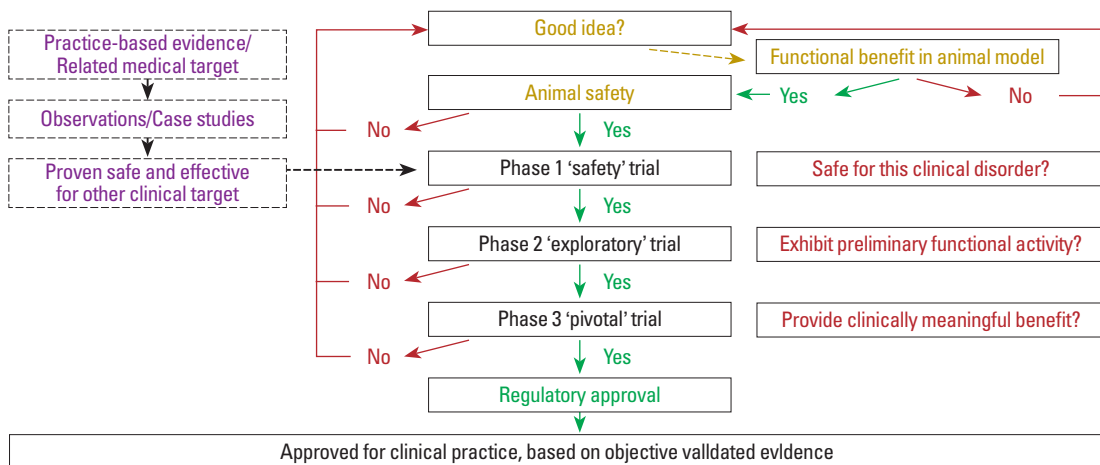

Fig. 2. Summary chart of a translational path for a therapeutic intervention from scientific discovery to approval for clinical use.

claims of beneficial outcomes are almost always unsubstantiated (i.e. not validated by an objective clinical trial) and often rely on anecdotal testimonials from hopeful patients or biased practitioners who stand to personally profit.

At this time, cell transplants are not without risk. There is not sufficient knowledge about the means to manage the possible fates of most cells after transplantation. Will some transplanted cells undertake uncontrolled differentiation and form cancerous tumors? Thus, a detailed scientific understanding of the potential risks and limited evidence for possible benefits should be completed prior to providing informed consent for a cell transplant procedure.[10] In addition, subjects are sometimes reluctant to enroll in an RCT as they do not want to find out after the trial that they were part of the control group. It is important to emphasize that any participant within the control group is receiving the current standard of clinical care (i.e., best practices) and should the experimental treatment have deleterious effects, it is better to have been a control participant (Box 1).

Most countries in the developed world have regulatory agencies with established criteria for the safe and effective evauation of an experimental therapeutic within a valid clinical trial program. Unfortunately, regulatory requirements are only as good as their enforcement and this has not been uniform across the globe. To provide some objective assistance to what is a complex series of decisions weighing the possible risks and benefits for a human study, a number of clinical trial guidelines have been created for a number of disorders and published by international panels of scientists and clinicians, including CONsolidated Standards Of Reporting Trials (CONSORT). The CONSORT Statement is an evidence-based, minimum set of recommendations for reporting RCTs. It offers a standard way for authors to prepare reports of trial findings, facilitating their complete and transparent reporting, and aiding their critical appraisal and interpretation. The CONSORT Statement comprises a 25-item checklist and a flow diagram, along with some brief descriptive text. The checklist items focus on reporting how the trial was designed, analyzed, and interpreted; the flow diagram displays the progress of all participants through the trial (all information is freely available for download: www.consort-statement.org/consort-statement).

As but one example of the breadth and depth of trial guidelines, recent publications have outlined the requirements for clinical trials involving subjects with spinal cord injury (SCI).[11-14] The issues being considered ranged from: 1) the degree of spontaneous recovery after SCI, which is neces-

---

Box 1- TGN1412: How NOT to start a phase 1 (safety) clinical trial

TGN1412 is the working name of a humanized monoclonal antibody, which was developed by TeGenero Immuno Therapeutics, tested by Parexel and manufactured by Boehringer-Ingelheim. TGN1412 was predicted to activate T cells. In its first human phase 1 clinical trials in March 2006, it caused severe inflammatory reactions and catastrophic systemic organ failure in all of the 6 healthy subjects who received the humanized monoclonal antibody, despite a supposed sub-clinical dose of 0.1 mg per kg; some 500 times lower than the dose found safe in animals, including non-human primates.

The antibody was given by intravenous infusion with an interval of around 10 minutes between patients. Thus all 6 subjects received the experimental antibody within one hour. A more conventional phase 1 study protocol examines the response in one individual at a time for a sufficiently appropriate period to assure there are no adverse events before administering the therapeutic to the second subject. If Parexel had undertaken this approach, only one volunteer subject would have suffered. Although the differences between the preclinical animal responses and the human reaction have not been completely explained, it would seem that there are some subtle, but significant differences between the pathways for activating T cells in non-human primates and humans which are beyond the scope of this review. TeGenero has ceased business and, 5 years later, all human monoclonal antibody clinical trials are still banned in the UK.

sary to set reasonable clinical endpoint thresholds,[11,15,16] 2) approaches for selecting accurate, sensitive and reliable outcome measures,[12,15] 3) discussion of numerous inclusion/ exclusion criteria and important ethical considerations,[13] as well as 4) various trial designs and protocols.[14] In addition, the same authors created a document written for the general public and allied health care professionals, which outlined what a patient should ask of any investigator before accepting an experimental intervention, such as a cell transplant (available at www.icord.org).

WHAT ARE THE DIFFERENT TYPES AND PHASES OF A CLINICAL TRIAL?

A detailed understanding and appreciation of the clinical trial process can be obtained from various national and international regulatory agencies such as *www.clinicaltrials.gov*. There are a number of different types of clinical trials and each has a distinct goal.

1) Diagnostic trials are conducted to find better tests for diagnosing a particular disease or disorder and are usually conducted with people showing symptoms of that condition. For example, if researchers wanted to compare the diagnostic performance of magnetic resonance spectroscopy to that of magnetic resonance imaging (MRI) for a particular disease, they would conduct a diagnostic trial.

2) Screening trials test the best way to detect certain diseases or health conditions. A screening trial may complement or overlap with a diagnostic trial.

3) Prevention trials look for ways to prevent illness or injury in people who have never had the disease and the approaches may include: vaccines, vitamins, exercise, or lifestyle changes.

4) Quality of Life trials explore ways to improve access, opportunity and integration within the community for individuals with a chronic illness or disability.

5) Treatment trials are the most familiar and test experimental interventions designed to improve outcomes for patients whether the therapeutic is a drug, cell transplant, surgical procedure, assistive device, or rehabilitation strategy. However, this does not imply that all these different treatment categories have identical regulatory demands for approval.

Each phase of a clinical trial program also has distinct goals and thus different parameters, protocols, outcome measures, and endpoints that can govern the conduct for each stage of investigation. Nevertheless, safety is continuously monitored throughout all subsequent trial phases. The overall translational path for an experimental intervention through the various stages of preclinical and clinical study might be summarized as shown in Fig. 2.

Phase 1

Trials are centered on the initial exploration of safety, and in the case of a drug, or cell transplant often include an evaluation of the responses to different therapeutic doses. Moreover, in the case of a drug, the purpose of this phase is to characterize its pharmacokinetic behavior and determine the metabolic profile. Typically, this is described using the acronym, 'ADME' which stands for 'Absorption, Distribution, Metabolism and Excretion. The primary goal of a Phase 1 program is to determine if a projected therapeutic blood level, based on preclinical data, can be reached without evidence of harmful side effects in the healthy human or targeted population (elderly, pediatric, renal failure, etc.). Phase 1 trials will sometimes attempt to collect pilot data on functional outcomes, primarily to justify continuation, particularly funding, for the trial program.

Phase 2

Trials are still exploratory with a focus on the preliminary demonstration of functional biological activity and/or functional benefit of the intervention. They will usually compare a number of different biological, clinical or functional outcomes to determine which endpoint most reliably measures a clinically meaningful outcome in a sensitive and accurate manner. Moreover, these studies reveal information on the therapeutic response variability seen with a given measurement and therefore, provide a basis to calculate the study size needed to perform a definitive (Phase 3) study in a larger population. In addition, for drug trials, an attempt is often made to correlate a given blood level with a therapeutic outcome by dose-ranging studies where the lowest dose projected to have efficacy is tested and doses are increased in subsequent cohorts to find the best balance between benefit and risk.

Phase 3

Trials are the pivotal confirmatory studies where an intervention must demonstrate benefit in a clinically meaningful manner. This efficacy is then weighed against the associated risks before the relevant regulatory body can approve it. A clinical meaningful benefit for a therapeutic can be difficult to define (see below), especially for disorders with low

incidence such as SCI or disorders where no therapeutic intervention has been established as the approved standard of care (i.e., the 'gold' standard). Finally, Phase 3 studies help to create a more substantial safety database by providing information about potential adverse events or side effects in a large number of patients.

After regulatory approval and adoption of the intervention as a standard of clinical practice, most interventions enter a surveillance period where the greatly increased exposure of a more heterogeneous array of patients to the treatment will enable detection of less frequent adverse events. At this stage, it is also possible to perform Phase 4 (post-market) use studies that continue the examination for optimal treatment approaches, and safety in a more controlled way, including additional forms of efficacy and interactions with other treatments the patient may also be receiving (e.g., drugs).

WHAT CHARACTERISTICS OF A HUMAN DISORDER SHOULD BE NOTED, TRACKED, OR STRATIFIED FOR A CLINICAL TRIAL AND WHAT INCLUSION/EXCLUSION CRITERIA ARE IMPORTANT TO ENROLLING APPROPRIATE TRIAL SUBJECTS?

In the translation from preclinical animal experiments to human studies there are a number of considerations to contemplate. Most categories of human disorders are not homogeneous in their origin, presentation, or course of progression. Potential human participants in a clinical trial are unlikely to be as homogeneous a group as the preclinical animals used in an experimental laboratory. Scientists can often create a good approximation of a disorder that may match the human disorder in several aspects, but it is a rare situation where a preclinical animal model can completely mimic a human disorder.

For example, special impact devices have allowed scientists to create a SCI in animal models within the same relative time frame for most traumatic human spinal injuries (<10 msec), but to date, these devices cannot recreate the mechanical damage that may simultaneously compress, dislocate and distract the spinal cord along 3 different physical axes. In addition, a human spinal injury can be acquired from a variety of non-traumatic incidents, including: infection, vertebral stenosis, or spinal tumors. Even though many people think of SCI as relatively homogeneous, it can

be heterogeneous in terms of:

-Level of injury along the rostrocaudal length of the cord (from ventilator-dependent, high cervical SCI to ambulatory, cauda equina injury)

-Severity of injury (from incomplete spinal damage to complete sensory and motor loss below the level of injury)

-Progress since the onset of sensory and motor disruption, including the degree of ongoing deterioration or spontaneous recovery from the acute stage, through sub-acute, to more chronic time points.

Therefore, when initially examining an experimental intervention it is prudent to recruit subjects that best approximate the type of disorder that was created in the preclinical studies. In short, it is wise to minimize the variability between study subjects. If we use SCI as an example, it is possible to enroll subjects that range from complete sensorimotor loss to a minimal loss of sensory and motor function. However, it would be no surprise that people with less severe forms of SCI are capable of more spontaneous recovery than those individuals with complete sensorimotor loss. Thus, the potentially different outcomes from heterogeneous subjects during a clinical trial may obscure, offset, or cancel out detection of any treatment effect for the therapeutic intervention. Consequently, the investigators might conclude erroneously that there is no treatment effect. A robust, sensitive and accurate classification system for inclusion and exclusion from a trial is important to enrolling appropriate study subjects.

The greater the variation in the subjects that are recruited to a study, the higher the requirement for stratification and thus, the lower the statistical power (significance) of any analyses. The only way to overcome subject heterogeneity is to increase the number of enrolled participants or start again; either choice will mean more time and money. For the same reasons, it follows that in every possible respect, a (placebo) control group should be matched to the composition of the experimental study group. Once a therapeutic intervention has been validated as beneficial for a homogeneous target population, there will be ample opportunity (and funding available) to subsequently study other forms of the disorder.

In addition to homogeneity of study subjects, there are other criteria for including and excluding the enrollment of subjects to a clinical study.[13] Needless to say, these criteria will vary as a function of the disorder being evaluated; however, there are some general principles for all clinical studies. The hypothesized action of the candidate therapy will define the time points for its administration and thus influ-

ence which subjects are appropriate to recruit to the study. For example, tissue Plasminogen Activator (tPA) is a protein involved in the breakdown of blood clots. It catalyzes the conversion of plasminogen to plasmin, the major protein responsible for the breakdown of blood clots. tPA is used to treat ischemic stroke within a few hours after the onset of stroke symptoms (usually <3-4.5 hours), but contraindicated in hemorrhagic stroke where it would exacerbate outcomes. This is a challenging "window of opportunity" for the administration of tPA since many patients do not seek medical assistance quickly, and/or an MRI cannot be completed within this time frame to rule out hemorrhagic stroke. As a consequence, the majority of people suffering an ischemic stroke do not receive tPA.

The need to rapidly identify and accurately screen suitable participants is a major challenge to recruit a sufficient number of participants for any acute clinical trial. Regardless of the clinical target, the list of inclusion and exclusion criteria can quickly reduce the number of qualified subjects and it is not unusual to enroll less that 10% of the initially eligible participants (Fig. 3). But, should the diagnosis for the severity of the disorder be difficult to correctly determine, you could enroll an inappropriate participant that will distort trial outcomes.

The chronic stage of some disorders, such as SCI, yields a fairly static functional baseline where there is little ongoing functional improvement beyond the first year after injury. As a result, some chronic studies have a broader time range for the recruitment of subjects with varying survival times. The low functional variability in chronic activities of daily living facilitates the detection of a treatment effect using a smaller sample size. Conversely, it is difficult to biologically influence damaged tissues or cellular actions in a positive manner at chronic time points.

Most clinical trials have lower and upper age limits for subject enrollment. Lower limits are often set at the legal age for adulthood (often 18 years of age) as this facilitates ethical considerations surrounding direct informed consent being adequately provided, understood, and agreed to by the potential study participant. The upper limit (usually 55-70) is often influenced by two major considerations. First, older study participants may not have the same rate for spontaneous recovery and their diminished capacity may not align appropriately with younger participants, leading to inconsistent trial outcomes. Aging is often accompanied by pre-existing or co-existing medical conditions that may also influence or alter the outcomes of a clinical trial. The longer and more physically active lifespans of many people means careful attention must be maintained in matching study participants in terms of age and functional capacity.

There is growing consensus that active physical and occupational rehabilitation benefits faster and greater recovery after most health challenges. In short, it can be argued that rehabilitation training is a necessary component for maximizing recovery after any surgical or drug intervention. Nevertheless, augmented rehabilitation has its own intrinsic benefits and conversely physical inactivity can worsen functional improvement. Thus, when undertaking a clinical trial of a therapeutic intervention, it is possible that rehabilitation effort can be an independent variable that confounds the accurate determination of how much benefit was due to the therapy or to the accompanying rehabilitation. Without more rehabilitation RCTs providing the necessary wisdom for the "yin and yang" of active physical and occupational training, rehabilitation will remain a double-edged sword. Nonetheless, it is important to attempt to standardize rehabilitation during a clinical trial and at a minimum activity should be monitored and tracked over the course of the study.

Depending on the therapeutic target, gender may or may not be a confounding factor during the course of a clinical study, but pregnant women or women unwilling to use double barrier methods to avoid pregnancy during the course of a clinical trial are usually excluded. The known risks for an adult drug dose being delivered to an unborn fetus (Box 2) are the obvious reason and also exclude adolescent children from most initial treatment trials. Should a therapeutic be approved as a beneficial standard of care, a subsequent trial can be initiated for these populations.
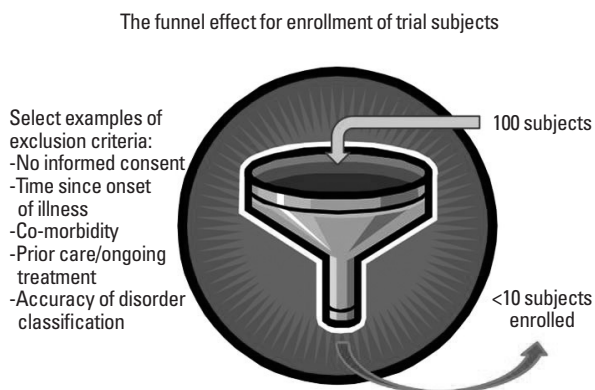
The funnel effect for enrollment of trial subjects

Select examples of exclusion criteria:
-No informed consent
-Time since onset of illness
-Co-morbidity
-Prior care/ongoing treatment
-Accuracy of disorder classification

100 subjects

<10 subjects enrolled

Fig. 3. The necessary inclusion and exclusion criteria to qualify as a trial subject can create a funnel effect that substantially reduces the number of trial participants that can enroll in a human study. This can be a very difficult challenge for rare (small incidence) disorders and to generate sufficient participants may require recruitment from multiple center trials at even the earliest phases of a study.

Concomitant medications or treatments can also constitute a criterion for exclusion of a subject from a study. Minimally, all medications and alternative treatments (e.g. acupuncture, herbal medicines) must be documented and tracked over the course of a study. Previous participation in a clinical trial may also constitute a reason for exclusion as the previous intervention might contaminate the results of the current trial (e.g. synergistic or antagonistic therapeutic effects). Prior to enrollment, there has to be sufficient evidence that the study subject has not retained any residual effects from a previous therapeutic intervention. This does not mean that the patient cannot receive the treatment once it has completed a clinical trial program and been approved for use (as long as the patient still qualifies for the approved use of the treatment). More detailed discussions of subject inclusion/exclusion criteria and ethics can be found elsewhere.[13,17]

WHAT ARE THE DIFFERENT STUDY DESIGNS THAT CAN BE USED IN A CLINICAL TRIAL PROGRAM?

Numerous guidelines for the general conduct of any clinical trial have been developed, and readers are encouraged to make themselves familiar with these teachings, especially those developed by the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (available at www.ich.org). In addition, any clinical trial must minimally adhere to the ethics guidelines of the Declaration of Helsinki, the Belmont Report, and the standards of the host country (Box 3). Where these ethical standards come into conflict, it is the responsibility of the investigators to work with review committees and regulatory agencies to determine whether the study can be undertaken or continued.[18]

Besides the ethical use of study subjects, one of the primary goals in designing a valid clinical trial program is the removal of any bias in the outcomes of the results. Bias is defined here as the systematic tendency of any factors associated with trial design, conduct, analysis, and interpretation of results to cause the estimate of a treatment benefit to deviate from its true value. It must be remembered that it is relatively easy to get positive responses from subjects when they know they have recently been treated with an experimental therapy, and when they expect or hope for a benefit (the so-called 'placebo effect'). Thus, trials with randomized controls and blinded assessments are necessary standards to remove investigator and subject bias. Before a study commences, all trial protocols should be registered with appropriate regulatory agencies where the study is to be conducted (e.g. http://prsinfo.clinicaltrials.gov). Needless to say all clinical trial results, whether positive or negative, should be submitted for publication.

There are varying degrees of blinding, starting with 'Open

Box 2- Thalidomide: Why pregnant women are not enrolled in clinical trial programs

Thalidomide, launched by Grünenthal on 1 October 1957, was found to act as an effective tranquilizer and painkiller and was also found to be an effective antiemetic, which had an inhibitory effect on "morning sickness". Thousands of pregnant women took the drug to relieve their symptoms. At the time of the drug's development it was thought unlikely that any drug could pass from the mother across the placental barrier and harm a developing fetus. Of course, everyone today knows this is not true.

Thalidomide was sold in a number of countries across the world from 1957 until early 1962. Canada was the last of 46 countries to have thalidomide withdrawn from the market after it was found to be a major cause of fetal deaths or teratogenic birth defects (especially of the limbs). It is not known exactly how many victims of the drug there have been, although estimates range up to 20,000. Since then thalidomide has been found to be a valuable treatment for a number of medical conditions, including multiple myeloma. Thalidomide is being prescribed again in a number of countries, although its use remains controversial as to how to protect pregnant women from being administered the drug.

Label', wherein the identity of the treatment is known to both the investigators and participants. 'Open Label' protocols have been used in the study of both pharmacological, surgical and rehabilitation interventions. They should generally be reserved for Phase 1 trials.

The next level is a 'Single Blind' study where either the clinical investigator (examiner) or the subject, but not both, are blinded. For trials where a surgical intervention is part of the experimental protocol, it may be necessary for the surgeon to know what is being undertaken in that subject. However, it is preferred that the patient remain blinded to the treatment received (for both experimental and control groups), although this is not always possible. Ethical or legal difficulties may interfere with the use of blinding when it entails sham operative procedures. Nevertheless, independent outcome examiners should and can always be used and remain blinded to the treatment provided. This may also require audio monitoring to assure that a subject does not disclose to the examiner which treatment arm they were assigned. Using techniques such as identical bandaging of the overlying skin during assessments can also facilitate the blinding of surgical outcomes to independent examiners.

Finally, in an optimal 'Double Blind' design, neither the participating trial subject, or investigators, trial staff, and sponsors should be aware of the treatment received by the subject. Ideal blinding would ensure that the treatments couldn't be distinguished by subjective experience, appearance, timing, and delivery method by anyone. This should be maintained throughout the conduct of the entire trial from determination of eligibility through evaluation of all endpoints, and requires full compliance of the subject. Double Blind design has been used in numerous pivotal Phase 3 trials.

Randomization in the assignment of trial participants to the different study arms (groups), including a placebo control group, is done to reduce bias in the outcomes of a trial. This provides a valid statistical basis for the evaluation of any treatment effect. External independent randomization tends to produce study groups where the distribution of any known or unknown independent variables (confounding influences) are similar and representative of the overall patient population. Randomization should only be performed after the eligibility of the subject to participate in the trial

---

Box 3- Tuskegee syphilis study: An unethical clinical trial that generated new regulations to protect human study participants

The Tuskegee syphilis study was an infamous clinical study conducted between 1932 and 1972 in Tuskegee, Alabama by the U.S. Public Health Service to study the natural progression of untreated syphilis in poor, rural black men who thought they were receiving free health care from the U.S. government. Investigators enrolled in the study a total of 399 impoverished, African-American field workers who had previously contracted syphilis before the study began. They were never told they had syphilis, nor were they ever treated for it. The victims of the study included numerous men who died of syphilis, wives who contracted the disease, and children born with congenital syphilis.

The 40-year study was controversial for reasons related to ethical standards; primarily because researchers knowingly failed to treat patients appropriately after the 1940s validation of penicillin as an effective cure for syphilis. The ethical decision should have been to stop the study and provide all the participants with antibiotic treatments. Instead, the Tuskegee scientists continued the study without treating any participants and withholding penicillin and information about it from the patients.

Revelation of this unethical study in 1972 led to an immediate halt of the study and major changes in regulations for the protection of participants in clinical studies. Now studies require detailed informed consent from participants before they enroll in a human study. The Tuskegee syphilis study led to the 1979 Belmont Report and the establishment of the Office for Human Research Protections (OHRP). It also led to the creation of Institutional Review Boards (IRBs) for the protection of human subjects in studies involving human subjects.

has been established. Randomization schemes for multi-center trials should be centrally organized and 'blocked' by each participating center, such that all participants within a center are also randomized to the different study groups.

There are numerous variations in the design of clinical trials, but only the most prevalent trial design configurations are provided below. Parallel group design is the most common clinical trial design for pivotal Phase 3 trials. Subjects are randomized (often in equal numbers) to one or more treatment arms, each testing a different treatment or combination of treatments. The treatments might include the investigational product at one or more doses, and one or more control conditions such as a placebo and/or an active comparator, which is a current (standard of care) treatment being provided to both the experimental and control groups. There are many variations for a parallel group design including superiority or non-inferiority trials.

As indicated above, superiority trials examine the new experimental therapeutic against a comparator, often the current standard of treatment, with the intent to show the new therapy is better. Often as not, the trial may only demonstrate equivalency or the non-inferiority of the experimental treatment. Then the decision whether to license the new treatment is usually determined on other criteria, such as less side-effects, lower dose, less cost, longer action, etc.; all of which requires additional data. Nevertheless, the assumptions underlying the parallel group design are less complex and more robust than those of other designs.

Crossover designs consist of randomly assigning subjects to a sequence of two or more treatments (e.g., placebo control first and then the experimental therapeutic or vice versa). Hence, the subject acts as his or her own control for the evaluation of benefits from the experimental treatment. To make valid assessments for the functional efficacy of a treatment using a crossover design, the clinical outcome measure (primary endpoint) must have a very stable (unchanging) baseline value before application of the experimental treatment and any subsequent measurements. Because the functional capacities of a person with an acute or sub-acute disorder can vary dramatically over a short period of time, this type of design should be restricted to studies of chronic illness, where the functional capacity to be assessed is expected to remain stable and unchanging. Thus, a crossover trial in patients with a progressively deteriorating disorder is not advised unless the rate of deterioration has a longer time frame then the therapeutic being tested.

In addition, the effects of the experimental treatment should be evident within the treatment period, and reverse following removal of the treatment. One important concern for a crossover design is the possibility of residual effects (i.e., a carryover influence) from either the experimental treatment or the placebo control. A lingering effect might alter the subject's response after crossing over to the opposite treatment arm. Thus, the 'washout' time period between treatment conditions should be sufficiently long to allow the complete reversibility of any detectable treatment effect. Nevertheless, an advantage of a crossover design is usually a reduction in the number of subjects needed to achieve statistical significance. Once again, the trial should employ examiners 'blinded' to the treatment the subjects are receiving during each stage of a crossover study.

An independent Data Safety Monitoring Board or Independent Data Monitoring Committee can use a group sequential design to facilitate an interim analysis. It involves any statistically valid analysis intended to compare treatment arms with respect to safety or efficacy at any time before the completion of the trial. Such an approach can be used to identify and quickly discontinue an unsafe trial, stop a trial when the treatment has demonstrated dramatic benefit, or identify an ineffective treatment dose with the subsequent randomization of subjects to a potentially more effective dose. Utilization of such a design requires a declaration, at the beginning of a trial, that an interim analysis will be undertaken, and the use of prospectively defined criteria (thresholds) for the early termination due to either safety or efficacy reasons.

Factorial Designs involve the testing of two or more treatments simultaneously for possible synergistic or antagonistic effects (combination treatments). In the simplest 2×2 form, subjects are allocated to one of four groups: A alone, B alone, both A and B, or neither A nor B. In 'Add-on' Trials, the test treatment and placebo are added to a common standard therapy.

Since there is usually an ongoing debate within any research community over the improved benefits that will be achieved with a combination of treatments, it is sometimes suggested that a factorial clinical trial does not need to include separate study groups that individually examine only one element of the combination therapy. To undertake this expedited approach requires strong preclinical animal data that the combination therapy is effective. Alternatively, the isolated administration of the individual components in previous clinical trials may have only demonstrated a modest effect that was sub-threshold to achieving any clinically

meaningful benefit. Preclinical safety and toxicity data would also be required to indicate that the components of the therapy are safe, when administered separately or when they are combined.

WHAT CONFOUNDING FACTORS CAN ALTER THE ACCURATE INTERPRETATION OF CLINICAL TRIAL OUTCOMES?

A list of several potential confounding factors that could jeopardize the accurate interpretation of the outcomes in a clinical trial is provided in Table 2. Some of these influences are unknown and/or beyond the control of the investigators conducting the trial. Depending on the start and end dates of the trial, subjects may have already been influenced by a variety of emergency, primary care or rehabilitation treatments. There may be concomitant damage to a varying number of body tissues or cellular functions. For longer-term studies, the amount of rehabilitation effort expended by study participants can be highly divergent. It is unethical to tell a person they cannot engage in exercise or active rehabilitation. Thus in many instances, the choices are either to provide all subjects with uniform high levels of ongoing care or rehabilitation, which is expensive and sometimes futile. Alternatively, any differences can be carefully tracked and documented for

subsequent analysis of their influence on trial outcomes. In any case, these decisions should be declared prospectively prior to the commencement of the study.

There are, however, a number of confounding factors that any investigator can control so they do not alter the accurate interpretation of clinical trial outcomes (Table 2). With the use of appropriate study designs and protocols the influence of these confounders can be mitigated and/or marginalized.

One issue that is always a concern is the choice of appropriate outcome measures and how often and for how long the outcome measures should be assessed. The specific outcome measures and trial endpoints can and should change as the trial program progresses through the various phases. The objectives of Phase 1 trials can be quite varied, from the initial exploration of tolerability, through study of human pharmacokinetics and metabolism, to identification of the maximum safely tolerated dose of a candidate therapeutic.[12,14] Participant's who choose to take part in a Phase I trial may experience significant risks with a limited probability of receiving individual benefit.

General Phase 1 trial safety outcome measures include: ongoing assessment of standard vital signs, physical examination data (e.g., temperature, respiration, heart rate, and blood pressure), clinical laboratory tests (e.g., hematology and urine analysis), as well as the appearance of any systemic adverse event (observed or reported by a trial subject). Depending on the therapeutic drug or cell line being evaluated

Table 2. Confounding Factors that Could Alter the Accurate Interpretation of Clinical Trial Outcomes. These Factors May or May Not Be Controlled by the Study Investigator, but Should be Carefully Noted Over the Course of the Study and Taken Into Consideration during the Analysis of Data

Possible confounding factor(s)	Can this confounder always be controlled by study investigator (Y/N)
Prior emergency or primary care, intensive care treatment and ongoing patient management	N
Surgical procedures not related to therapy (and timing of surgery)	N
Damage to other organs or subsequent medical complications after onset of disease or disorder	N
Onset, type, and extent of rehabilitation activity in which subject engages during trial	N
Unsuitable study protocol design or statistical analysis of trial results	Y
Inappropriate inclusion and exclusion criteria for subject to participate in trial (e.g. different: distinct types of disorder, severities of disorder, or time since onset of disorder)	Y
Lack of "blinded" randomization of subjects and randomization of the treatment allocation sequence	Y
Lack of appropriate control subjects, which match criteria of experimental treatment group	Y
Investigator or subject expectations and bias, because they know the type of treatment being provided to the subject	Y
Inappropriate, insensitive or unreliable clinical outcome tools and/or primary outcome measure (clinical trial endpoint)	Y
Lack of independent "blinded" assessments of trial outcome measures	Y
Poor intra- and inter-rater reliability of outcome assessments (lack of ongoing training of trial assessors)	Y
Lack of sufficient follow-up assessments (sometimes 6-12 months after trial completion)	Y

and the route of administration, other Phase 1 safety outcome measures may include the evaluation of unintended effects on the target tissue for the therapeutic or other body tissues, including infection, inflammation, or immune reactions.

Combined Phase 1/2 trials, where safety and bioactivity of the therapeutic are evaluated together can often occur when the Phase 1 trial cannot involve healthy subjects and is restricted to people having the clinical disorder. It is possible for SCI clinical trials to be designed in this manner. Nevertheless, the data from such a combined Phase 1/2 trial must be able to satisfy the essential outcomes for each respective trial phase. The preliminary Phase 2 data are collected to develop a preliminary sense of potential efficacy and to assist in the identification of appropriate outcome measures to be used in subsequent properly powered Phase 2 or 3 trials.

During a Phase 2 study, an exploratory evaluation of efficacy becomes more prominent, with the objective of determining potential effect size and variability of an experimental therapy in comparison to a useful control group. Information is gained regarding choice of optimal end points for a larger Phase 3 confirmatory trial of efficacy. During a Phase 2 trial, additional information is also obtained regarding safety.

The preferred Phase 2 design would be a RCT where each participant is recruited prospectively and randomly assigned to either the experimental or control arm of the study and where the investigators and the participants are blinded to which study arm they have been assigned. If available, Phase 2 trials could employ surrogate end points, which are expected to be accurate predictors of functional improvement, to estimate presumed effective doses, and to allow trials of shorter duration and smaller size to be conducted. Therapies conceived as early interventions or acute stage treatments are likely to be administered within days of the trauma or onset of the disorder and it is important that the outcome tools have the ability to accurately and sensitively track meaningful changes across a broad chronological timeframe.

Phase 3 (therapeutic confirmatory) trials are generally the definitive or pivotal clinical trial phase and typically undertaken as a RCT. The object is to confirm the preliminary evidence obtained at the Phase 2 stage by demonstrating a statistically significant and clinically meaningful benefit of the therapeutic in a wider group of subjects across multiple study centers (see below).

An interesting consideration for Phase 2 and 3 outcome measures is the situation where the technique for administering the experimental therapeutic intervention, such as surgery or a concomitant surgical treatment might have an independent therapeutic value that is separate from the drug, cells, or device that is implanted. When is the best time to make a baseline (starting) outcome assessment, before or after the surgical procedure? The most prudent approach would be to examine baseline capacities, both before and after surgery, to determine if the surgery alone might have provided a beneficial outcome. The number of times over the course of the study that an interim outcome assessment is performed before the endpoint evaluation is dependent on the presumed mode of therapeutic action and the duration of the study, but is usually no less than 2-3 interim examinations. More frequent assessments may be made at earlier study time points, but the number of evaluations depends on numerous factors, not the least of which are the demands placed on the subjects and the cost for completing each assessment.

Detailed reasoned explanations on specific clinical trial procedures and conventions have been written for every major clinical disease or disorder (e.g., SCI[14]) and should be reviewed prior to engaging in a clinical trial. It is important that all trial investigators understand the reasoning behind the specific trial protocol and agree to abide by the guidelines for the study.[19]

WHAT ARE THE SUCCESS RATES OF CLINICAL TRIALS AND WHAT CAN WE LEARN FROM PREVIOUS CLINICAL TRIALS?

A recent Reuters News article from February 14, 2011 (http://www.reuters.com/article/2011/02/14/us-pharmaceuticals-success-idUSTRE71D2U920110214?pageNumber=2) quotes several sources, from 2004 through 2010, that suggest the overall success rate for drugs moving from early stage Phase 1 clinical trials to FDA approval is about one in 10, down from one in five in previous years. BIO and BioMed-Tracker, which collects data on drugs in development, conducted the study. Some 63 percent of drugs in Phase 1 testing did advance to Phase 2, but only 33 percent of Phase 2 drugs made it to Phase 3. Approval applications were filed for 55 percent of the drugs that made it through the completion of pivotal Phase 3 study, and 80 percent of those gained eventual approval, although only half were approved after their initial FDA review. There are a lot of considerations in the decision to advance to the next phase of a clinical trial program and even when the trial data is positive, it does not guarantee

advancement through a program. Corporate decisions are based on risk-benefit analysis, which may often be the ratio of financial benefits to risks and not patient benefits or risks!

Nevertheless, the wisdom gained from "failed" clinical trial programs can provide valuable information and guidance for the next generation of human studies. Thus, the thoughtful publication of all clinical trial outcomes can always provide lessons, such as how the next study might conduct their trial in a more informed manner (i.e., what to do and what not to do). Prior clinical trial data is also valuable for understanding the "natural history" of any disorder. Knowledge of the biological and functional capacities, limitations or variability within a population of individuals can establish reasonable clinical outcome (endpoint) thresholds that must be exceeded by any experimental therapeutic to demonstrate both a statistical and a meaningful benefit in the next trial.[11,15]

Previous trial data can also help inform the research community about what might be a minimal clinically important difference (MCID) when conducting a trial for that target population. MCID is an important emerging concept that can be difficult to define, especially for neurological disorders. In brief, MCID is the smallest difference in an outcome measure that is perceived as beneficial (clinically meaningful) and would lead to a change in the patient's medical management or ongoing lifestyle (activities of daily living), assuming an absence of excessive side effects (risks) or costs.[20]

In determining a MCID for an outcome measure, several ingredients are needed, including: 1) an indicator that a change has been facilitated by the experimental therapeutic (i.e., a statistical difference exists between the experimental and control groups), 2) the observed change has value in terms of the patient achieving a functional benefit, 3) an appropriate method to determine the MCID threshold level for the outcome measure, and 4) a reasonable percentage difference between the experimental and control groups for those subjects achieving the MCID outcome threshold (i.e. responding).

For example, it may not be sufficient to demonstrate a drug reduces cholesterol level (the surrogate outcome); it may be necessary to show that the drug actually reduces the incidence of coronary artery disease. In the case of SCI, it may not be sufficient to show that a patient has gained some motor function from a treatment unless the neurological change can be related to an improvement in functional activities of daily living, such as improved independence for activities of self-care. It can be readily appreciated that setting a MCID for any disorder can be challenging and depend on many factors, including: 1) the stage of the disorder being tested, 2) the magnitude of any spontaneous improvement (natural history), 3) the therapeutic target of the intervention, 4) the primary outcome measure being used in the study, and 5) the benefits achieved by any previous treatment (i.e. current standard of care). Given the potential magnitude of any consensus statement about what constitutes a MCID for a given clinical trial, most clinical specialties have struggled to reach agreement on what is operationally meant by the terms "Important" and "Meaningful". Perhaps it is equally important to also consider "for what" benefit and "to whom" does it matter. Is this decision the right of patients alone or do patients in consultation with the investigators better define a MCID, as the latter better represents the overall interests of society?

Recently, the World Health Organization established an International Classification of Functioning, disability and health (abbreviated as ICF) to better represent the continuum of disability domains, including evolving patient goals, barriers to patient independence and community participation, as well as how these issues should best be tracked. ICF is a taxonomic structure or bio-psycho-social model integrating medical and social influences on disability (Fig. 4). Although many biological and medical investigators focus on improving body function after damage, it is not as meaningful to the patient unless the treatment enables a more independent lifestyle and increases the individual's community participation. Clinically meaningful recovery after a disorder or disease benefits from complementary interventions. An improvement in body functions and/or the capacity for independent activities increases the performance and participation of a person within their community.

SUMMARY: WHAT ARE THE ESSENTIAL PRINCIPLES FOR THE CONDUCT OF VALID CLINICAL TRIALS?

In brief, human studies without appropriate control data and blinded assessments are NOT pivotal clinical trials and will not validate an intervention as a treatment. Payment (by a subject) for an experimental treatment automatically means it is NOT a clinical trial, as the investigator has a financial incentive (i.e., a bias). It is also more difficult (but not impossible) to develop an effective trial design and protocol for a disorder, which has yet to validate a treatment with a clear clinical benefit. In short, it is easier to develop
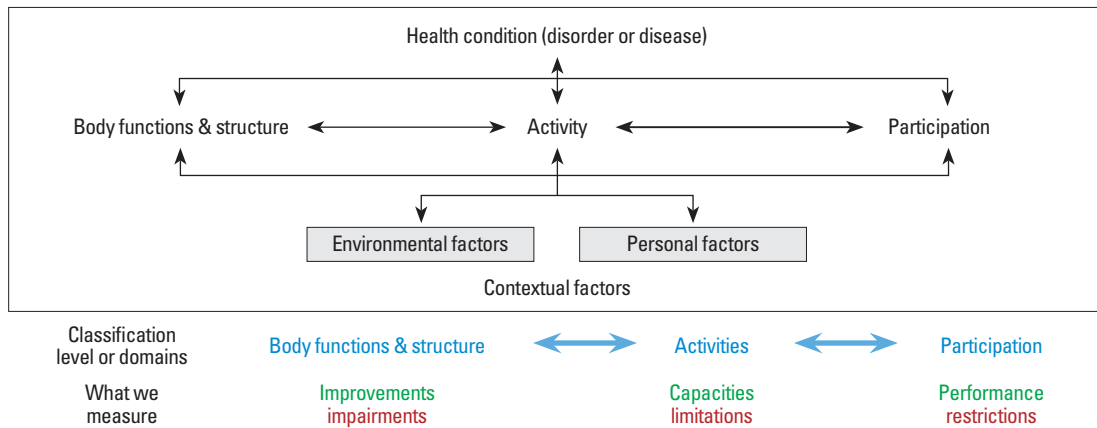
Fig. 4. International Classification of Functioning, Disability and Health (abbreviation ICF) is a taxonomic structure or bio-psycho-social model (from WHO) integrating medical and social influences on disability. Therefore, meaningful recovery is a continuum across the ICF Domains where we measure: improvements (or impairments) in the function of body tissues, capacities (or limitations) in the daily activities of the individual, and ultimately, at the highest and most complex level, the person judges their performance (or restrictions) to participate within their community. To date, participation outcomes have not been routinely used as a primary outcome in a clinical trial.

and validate new therapies when you can compare the new experimental treatment against a pre-existing comparator.

The optimal foundations for an ideal clinical trial program are difficult to guarantee, but planning how to achieve some of these is instructive. The considerations may include:

-Compelling preclinical data (independently replicated) with an understanding of the mechanism of therapeutic action so as to generate enthusiasm and financial support to withstand any small setbacks in the early trial phases. Financing clinical programs is difficult and usually limits the opportunities for trial and error learning.

-Sensitive and reliable outcome measures to show an objectively measurable effect of the treatment.

-Prospectively designed criteria for establishing clinically meaningful effects (clinical endpoints) sufficient to convince others that there is a real benefit to offset the development costs and risks of the intervention.

-A plan for dealing with partial success (e.g., small therapeutic benefit), with an understanding that there may be no practical way back or forward from a partial success in a pivotal Phase 3 trial.

-An understanding that randomized studies with appropriate placebo controls provides the strongest evidence for validating any clinical intervention.

-A clear understanding of the natural history of the clinical disorder, including the spontaneous (untreated) recovery that might be expected.

-Enrolment criteria that can be set to provide a pool of participants sufficiently homogeneous to show a treatment effect with the chosen endpoints.

-Appropriate statistical methods to ensure that the protocol

utilizes patients in the most ethical and effective manner. Unfortunately, space limitations does not permit discussion of the emerging analytic approaches, including Bayesian and Responder analyses, which may be more suitable to some trials than the conventional "comparison of the means" statistics.

REFERENCES

1. Bown SR. Scurvy: How a Surgeon, a Mariner, and a Gentleman Solved the Greatest Medical Mystery of the Age of Sail. New York: Thomas Dunne Books (St. Martin's Press); 2004.
2. Smith GC, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ 2003;327:1459-61.
3. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:

924-6.

4. Steeves JD. Considerations for the Translation of Pre-Clinical Discoveries and the Conduct of Valid Spinal Cord injury Clinical Trials. In: Lin VW, editor. Spinal Cord Medicine. 2nd ed. New York: Demos; 2010. p.930-8.

5. Kwon BK, Okon EB, Tsai E, Beattie MS, Bresnahan JC, Magnuson DK, Reier PJ, et al. A Grading System To Evaluate Objectively the Strength of Pre-Clinical Data of Acute Neuroprotective Therapies for Clinical Translation in Spinal Cord Injury. J Neurotrauma 2010 Oct 18. [Epub ahead of print]

6. Tetzlaff W, Okon EB, Karimi-Abdolrezaee S, Hill CE, Sparling JS, Plemel JR, et al. A systematic review of cellular transplantation therapies for spinal cord injury. J Neurotrama 2010; doi:10.1089/neu.2009.1177.

7. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006;126:663-76.

8. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007;131:861-72.

9. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science 2007;318:1917-20.

10. Blight A, Curt A, Ditunno JF, Dobkin B, Ellaway P, Fawcett J, et al. Position statement on the sale of unproven cellular therapies for spinal cord injury: the international campaign for cures of spinal cord injury paralysis. Spinal Cord 2009;47:713-4.

11. Fawcett JW, Curt A, Steeves JD, Coleman WP, Tuszynski MH, Lammertse D, et al. Guidelines for the conduct of clinical trials for spinal cord injury as developed by the ICCP panel: spontaneous recovery after spinal cord injury and statistical power needed for therapeutic clinical trials. Spinal Cord 2007;45:190-205.

12. Steeves JD, Lammertse D, Curt A, Fawcett JW, Tuszynski MH, Ditunno JF, et al. Guidelines for the conduct of clinical trials for spinal cord injury (SCI) as developed by the ICCP panel: clinical trial outcome measures. Spinal Cord 2007;45:206-21.

13. Tuszynski MH, Steeves JD, Fawcett JW, Lammertse D, Kalichman M, Rask C, et al. Guidelines for the conduct of clinical trials for spinal cord injury as developed by the ICCP Panel: clinical trial inclusion/exclusion criteria and ethics. Spinal Cord 2007;45:222-31.

14. Lammertse D, Tuszynski MH, Steeves JD, Curt A, Fawcett JW, Rask C, et al. Guidelines for the conduct of clinical trials for spinal cord injury as developed by the ICCP panel: clinical trial design. Spinal Cord 2007;45:232-42.

15. Steeves JD, Kramer JK, Fawcett JW, Cragg J, Lammertse DP, Blight AR, et al. Extent of spontaneous motor recovery after traumatic cervical sensorimotor complete spinal cord injury. Spinal Cord 2011;49:257-65.

16. Zariffa J, Kramer JL, Fawcett JW, Lammertse DP, Blight AR, Guest J, et al. Characterization of neurological recovery following traumatic sensorimotor complete thoracic spinal cord injury. Spinal Cord 2011;49:463-71.

17. Illes J, Reimer JC, Kwon BK. Stem Cell Clinical Trials for Spinal Cord Injury: Readiness, Reluctance, Redefinition. Stem Cell Rev 2011 Apr 8. [Epub ahead of print].

18. Donovan WH. Ethics, health care and spinal cord injury: research, practice and finance. Spinal Cord 2011;49:162-74.

19. Ziebland S, Featherstone K, Snowdon C, Barker K, Frost H, Fairbank J. Does it matter if clinicians recruiting for a trial don't understand what the trial is really about? Qualitative study of surgeons' experiences of participation in a pragmatic multicentre RCT. Trials 2007;8:4.

20. Gatchel RJ, Lurie JD, Mayer TG. Minimal clinically important difference. Spine (Phila Pa 1976) 2010;35:1739-43.