Looking for more than hot air: how experimental design can enhance clinical evidence for hyperbaric oxygen therapy

Adam T. Biggs*, Lanny F. Littlejohn
Naval Special Warfare Command, Coronado, CA, USA

*Correspondence to: Adam Biggs, Adam.t.biggs.mil@socom.mil.

Abstract

Hyperbaric oxygen therapy is emerging as a potential treatment for critical medical and psychological issues, including mild traumatic brain injury, traumatic brain injury, and post-traumatic stress disorder. Based on the promising results from numerous case studies, randomized clinical trials generated conflicting interpretations despite frequent improvements in patient symptoms. The primary debate concerns whether the therapeutic benefits could be attributed to placebo effects or sham conditions that actually induce a therapeutic state. In part, the contention has been exacerbated by experimental designs which could not properly account for extraneous variables, such as the potential for differing patient expectations to influence the outcome. The current discussion addresses these methodological challenges that complicate any determination of clinical significance due to experimental design. These challenges include: 1) not properly addressing or controlling patient expectations prior to the experimental sessions; 2) the challenge of experimental masking in clinical designs that require pressurized environments; 3) patient subjectivity in the primary dependent variables; 4) potential fluidity in patient symptoms or data, such as regression to the mean; and 5) the potential for nocebo effects to exaggerate treatment benefits by lowering performance expectations during pre-treatment assessments. Each factor provides an influential means by which placebo effects could complicate results and prevent the combined data from reaching a threshold of clinical significance. The discussion concludes with methodological best practices with which future research could minimize placebo effects and produce more conclusive results.

Key words: best practices; brain health; cognitive; hyperbaric oxygen therapy; mild traumatic brain injury; placebo; post-traumatic stress disorder; research design

doi: 10.4103/2045-9912.337992

How to cite this article: Biggs AT, Littlejohn LF. Looking for more than hot air: how experimental design can enhance clinical evidence for hyperbaric oxygen therapy. Med Gas Res. 2022;12(4):116-124.

INTRODUCTION

Hyperbaric oxygen therapy (HBOT) has received growing attention in recent years due to the potential therapeutic benefits of this technique. In addition to numerous uses already approved by the U.S. Food and Drug Administration or covered by health insurance policies such as TRICARE, some more controversial applications are attracting the bulk of the increased attention – both in empirical research and among the media. These uses specifically involve HBOT as a treatment for mild traumatic brain injury (mTBI), traumatic brain injury (TBI), and post-traumatic stress disorder (PTSD). Some researchers argue that this finding is evidence of a placebo effect, whereas other researchers argue that the “sham” conditions are actually inducing therapeutic conditions and therefore are not actually sham conditions. This debate underscores the importance of experimental design and medical ethics in eliminating placebo effects as a possible explanation of the benefits. Specifically, how do the experimental designs attempt to control for placebo effects, how do these designs in turn impact the results, and how close does the combined evidence come to meeting a threshold of clinical significance? More importantly, what can future research do to establish more concrete and reliable conclusions about these alternative clinical uses of HBOT?

The current discussion will take several steps to explore the existing challenges within alternative applications of HBOT. First, we will establish boundaries for the debate by discussing the controversy surrounding HBOT as a clinical treatment for mTBI, TBI, and PTSD. Second, the discussion will describe some of the basic methodological complications in research design. Third, we will address five distinct challenges that arise in distinguishing placebo effects from true clinical significance due to experimental design in hyperbaric research. These combined challenges in experimental design and result interpretation will then be compared among several existing studies that examined HBOT as a treatment for one of the three alternative uses. Finally, we will conclude with some best practices that should assist future research initiatives in addressing the potential for placebo effects as a possible explanation for any therapeutic benefits.

SEARCH STRATEGY

Search terms involved using the terms “hyperbaric oxygen therapy” with either “post-traumatic stress disorder” or “traumatic brain injury.” Primary search database was PubMed between 2000 and 2020. However, this literature review is exclusive to the first topic to establish the clinical treatment literature of HBOT for TBI/PTSD.

FRAMING THE DEBATE: CONTROVERSY AND THE NEED FOR CLINICAL SIGNIFICANCE

There is an obvious diagnostic difference between these three...
conditions. Both mTBI and TBI are related injuries that differ primarily in severity, whereas PTSD encompasses a wider range of issues from a psychological moral injury to high comorbidity with physical injury. The controversy does not arise due to the disparities in symptom profiles. Instead, the concern is in establishing a causal and clinically significant explanation for the observed improvement. Specifically, why would hyperbaric oxygen improve these conditions?

There is a potential mechanistic explanation for mTBI and TBI where the combination of increased pressure and increased oxygen creates a neuroprotective or regenerative environment by allowing dissolved oxygen in the bloodstream to penetrate deeper into brain tissues with damaged cerebral vasculature. However, it is more difficult to assign a mechanistic explanation for any improvements in PTSD following HBOT treatments. The high comorbidity between PTSD and mTBI/TBI could mean improvements in PTSD symptoms because HBOT treatments facilitate recovery by addressing issues arising from mTBI/TBI. Inconsistent mechanistic explanations are a primary problem driving controversy because it is difficult to assign theoretical or causal explanations to any observed improvements. The second challenge involves inconsistencies in the empirical evidence. For example, it is possible to find randomized controlled trials (RCTs) with conclusions that HBOT is ineffective at treating post-concussive symptoms and another trial demonstrating that it is effective at treating post-concussive symptoms. A disputed theoretical explanation would be problematic in and of itself. When combined with conflicting empirical evidence, it creates some controversy around any medical applications of treatment.

Although these implications are noted, the discussion will cover all three controversial applications concomitantly for the remainder of the discussion as the issues of clinical relevance and experimental design affect all three topics. Across these three health issues, multiple arguments could explain discrepancies among the empirical literature. Each position has some implication for experimental design with the caveat that establishing clinical significance is a critical concern for all four positions.

The first two arguments both offer different positions to explain why HBOT treatments produce large, but often statistically non-significant improvements between active treatment versus control groups. One position suggests that both groups are demonstrating a placebo effect, whereas the alternative position suggests that both active and sham treatments are inducing a therapeutic benefit (e.g., the sham is not a sham argument). Each account offers a different causal explanation for observed improvements, and the two positions are mutually exclusive. These competing positions also highlight the need for sound experimental design. For example, there is no way to make causal inferences between these positions if the study is observational and only applies an active treatment. Placebo effects would suggest that there is no real improvement and participants only appear to improve after the treatment because they expect to improve following medical intervention. Without sound recruitment strategies and expectation management, among others, it is difficult to discount placebo effects as an explanation. The sham-is-not-a-sham argument would posit that the benefits were observed because there is a real therapeutic benefit to these alternative HBOT applications. Without tight experimental controls between treatment conditions or the establishment of a dose-response curve, an observational or single-condition study cannot make robust causal claims about any improvements due to an intervention. As such, both arguments demonstrate the need for rigorous experimental design to ensure that experimental outcomes can make causal claims about alternative applications of HBOT.

The remaining two positions likewise have implications for experimental design, albeit in very different ways. Treatment heterogeneity suggests that many different HBOT treatment regimens composed of different durations and pressures – combined with an already heterogeneous symptom profile between cases of mTBI, TBI, and PTSD – contribute to the discrepancies as observed across studies. More experimental control and eventually more evidence are needed to help solve these issues. In short, the heterogeneity position suggests in part that it is simply too early in the research process to make causal conclusions without a more diverse empirical body of literature to support a symptomatically diverse set of alternative treatments. Meanwhile, the facilitation argument suggests that there is no direct effect of HBOT on mTBI, TBI, or PTSD symptoms. The observed benefits are indirect and improve the general quality of life rather than having any direct or physiological consequence upon the patient. For example, PTSD may have comorbid conditions that induce cognitive impairment. Addressing cognitive impairments for simultaneously present conditions smooths the recovery process by removing complications. This argument could be addressed with better data analyses to control for statistical irregularities that may arise through self-report variables or by accounting for covariates when designing data collection.

Despite the various explanations for the theoretical and empirical inconsistencies across the collected evidence, there is one common issue critical to all interpretations: clinical significance. This issue raises the question of whether the treatment can be ethically administered to a patient. However, there is no single, fixed definition of clinical significance, although there are general issues that impact most determinations of clinical significance. One approach recommends three criteria for determining a clinically significant change: 1) internal validity, which addresses whether change can be attributed to treatment rather than outside variables; 2) probability, which addresses whether chance is real or attributable to random chance as determined by statistically significant differences; and 3) practical significance, which addresses whether the change is important rather than trivial to patient care. We will describe clinical significance throughout the remainder of this discussion with these three criteria in mind.

For our purposes, we do not consider each component as a checked box or otherwise all-or-none criterion. Each aspect represents something important that should be considered in context when evaluating clinical significance. Each aspect can also be multi-faceted. Probability, as one example, might establish statistical significance ($P < 0.05$), but the observed statistical effect size provides further critical information that could be useful in establishing clinical significance. Comprehensive outcome evaluations may not always be possible,
and these logistical complications will likely be the limiting factor in achieving a comprehensive evaluation of clinical significance. As such, our interpretation of clinical significance is that better patient outcomes will occur as each component is better represented in any conclusions – that is, representing one component in conclusions is good, two is better, and multi-faceted representations of all three is best. The result is a continuum-like interpretation of clinical significance rather than dichotomous interpretations of having or not having a component.

Consider three questions when evaluating clinical significance (Figure 1). First, is it real? This question addresses issues of internal validity and whether causal implications can be drawn from the situation. Second, is it robust? This question addresses the statistical properties of a finding or effect, including factors such as statistical significance and effect size. Third, is it relevant? This question addresses practical significance and whether the outcome is going to have a meaningful impact. More importantly, consider how this approach allows the three components to be compared when a component is missing. An effect can be real and robust, but without relevance, there is no practical significance to its application. Administering adrenalin will have a substantial impact on heart rate, although there is no reason to administer adrenaline as a treatment for mTBI. An effect can be real and relevant, but without a robust impact, there is a low probability that it will help. There is evidence that tea can help improve cognition, although recommending tea is unlikely to help someone overcome the cognitive deficits of TBI. Finally, an effect can be robust and relevant, but without being able to determine causal relationships, the findings would be unreliable or inconsistent. This latter possibility best encapsulates the debate about using HBOT for mTBI, TBI, or PTSD that RCTs have demonstrated robust and relevant impacts on patients – yet not all studies have demonstrated such improvement.

Figure 1: How do various components of clinical significance overlap?
Note: Each circle in the Venn diagram asks a critical question related to clinical significance, and the areas of two overlapping questions identify the shortcoming when the third component is not addressed.

**THE CURRENT DEBATE: FROM CASE STUDIES TO RANDOMIZED CLINICAL TRIALS**

HBOT is an approved therapeutic technique for multiple conditions, including the TRICARE-covered applications of decompression sickness, carbon monoxide poisoning, air or gas embolism, and profound blood loss when transfusion cannot be accomplished. It is important to reiterate this point because HBOT is offered as a covered therapeutic option for several medical issues – not as a purely experimental or novel treatment technique. The novelty arises because some early evidence, particularly from case studies, demonstrated patient improvement following brain injury for individuals who received HBOT. Case reports have incorporated supplemental, archival, or observational studies, although each instance documented significant improvement for individuals suffering from mTBI, TBI, or PTSD.

Case study data should be carefully considered and evaluated, but in developing a novel medical therapy, case studies must be weighed in light of what they are – extremely well-documented anecdotes. When describing these studies, the term “case studies” could be broadly construed to incorporate both a case study and a case series. From case reports and case series, despite their uncontrolled nature, there remained significant questions about the repeatability and reliability of the findings given that these observational studies are inherently limited in scope. Their medical purposes can be broadly divided into two areas. The first medical usage of case studies is to be considered as instructive examples. Established physicians can gain a better understanding of clinical diagnoses or treatments for which they might be unfamiliar, thereby making them a continuing education tool for medical professionals. Another instructive use is for more junior medical personnel, who use case studies as a training vehicle to consider evidence, analyze a situation, and come to a conclusion in light of all known factors associated with a case. The medical student can then apply lateral thinking should the individual encounter a similar case with similar variables at some point in the future. These possibilities describe the instructive uses of case studies. The second medical usage of case studies is oriented more toward research, where atypical or exploratory examples provide occurrences where patient health was impacted by conditions or procedures atypical to the standard practice of care. In this instance, case studies point to a potential application of some treatment or procedure to a safe and reliable future usage. However, case studies alone can never fully satisfy the threshold of clinical significance – no matter how many case studies are included.

Case studies inherently cannot address internal validity because there is not enough experimental control to unambiguously assign causality to the outcome, although they can identify likely factors that contributed to the outcome for future investigations. Likewise, case studies cannot address probability or random chance as there is not enough evidence to conduct statistical comparisons and no control group for comparison. Observational studies could conduct some pre-treatment and post-treatment assessments, but there would not be a control group to which the results could be contrasted. However, case studies excel at documenting practical significance as they address the change relevant to a single patient. The focus on a single individual or a few individuals allows for a level of tangible evaluation not always appreciated in abstractions taken from statistical inference. As such, case studies are a vital research tool indicating future directions in research as well as being a vital instructional technique, even
if case studies alone will not likely satisfy a multi-faceted definition of clinical significance.

After considering the case study evidence, these observations regularly demonstrate substantial improvements for cases of mTBI, TBI, and PTSD following HBOT treatments. In part, these combined case studies prompted RCTs that could elevate the therapeutic benefits while controlling for some of the extraneous variables. This approach would then allow for an evaluation of clinical significance that could address internal validity and probability in addition to practical significance. To date, several RCTs have been conducted on HBOT treatments to evaluate the therapeutic benefit, albeit the results have not provided consistent evidence. Several RCTs demonstrated a significant improvement in patient condition following HBOT treatments, whereas other RCTs have demonstrated no difference between active treatment and sham or control conditions. These conflicting results across multiple studies have raised questions not about the efficacy of the treatment, but rather the causality of individual improvement. Patients improved during HBOT treatments in almost every condition and manipulation. However, the source of the therapeutic benefit remains the disputed point. Proponents of the treatment argue that the sham conditions induced a therapeutic state because the increased pressure and/or oxygen produced increased levels of dissolved oxygen in the bloodstream, which prompted a cascade of metabolic changes. Alternatively, some researchers have challenged this interpretation as they claim the evidence merely suggests that the benefits of HBOT cannot outperform a sham condition or produce a dose-response improvement with the treatment.

This contention is central to any evaluation of HBOT evidence. Specifically, there is little question that even sham pressurizations produced improvements in these studies, but controlled intervention is still intervention. If the evidence cannot differentiate between control and active treatment condition, is it really the treatment producing the improvement or simply the act of intervention that produces improvement? The latter could describe a placebo effect, which can be pervasive in psychological and medical studies, and HBOT studies touch on both topics. This possibility makes placebo effects the primary concern in evaluating the clinical significance of HBOT treatments. As such, the discussion will now turn to how experimental design can impact the results and produce both placebo and nocebo effects that might complicate any experimental results and interpretations regarding clinical significance.

The Greatest Hurdle to Clinical Significance: Placebo Effects or Causal Improvement?
Placebo effects are psychobiological occurrences in both clinical and laboratory settings where patient improvement happens as a consequence of individual expectations – whether or not a physical placebo is actually provided. This potential raises significant ethical concerns, and while those concerns are duly noted, they are discussed more thoroughly elsewhere. For this discussion, placebo effects are repeatable occurrences that could complicate the conclusions of any scientific finding if proper controls are not implemented. As it pertains more specifically to HBOT, the placebo debate centers upon the competing arguments that: 1) the improvements are merely placebo effects from individuals who respond well to the intervention, or 2) the “sham” conditions often provided actually induce a therapeutic state due to the increased pressure and/or oxygen combination. Here we address five potential challenges that could complicate any causal interpretations of the evidence in HBOT studies (Figure 2). Notably, each challenge could obfuscate interpretations by creating perceived and often measurable benefits. The observed indicators that a given intervention is effective should be based upon the steps taken to establish clinical significance, which has been discussed previously. These methodological challenges are presented as complications to prevent confusion in the conclusion – not as steps to prove efficacy.

![Figure 2: A flowchart of questions to ask when determining whether placebo effects are likely in the experimental design.](Image)

The first and most common methodological issue regarding placebo effects involves the expectation of a benefit. This possibility operates on the presumption that an expectation of effect is one of the most powerful vehicles to cause a placebo response. In essence, any individual expecting medicine to work will exhibit improvements no matter what they are given. Participant expectations demand experimental control to prevent such individual differences from interfering with causal interpretations of the empirical evidence. Still, the first issue to note is that placebos can influence medical or physiological studies as much as psychological studies. A notable example here comes from cognitive training initiatives. Expectations can produce such robust effects that individuals could improve performance on fluid intelligence tests after an hour of sham cognitive training simply because they believed the training to be effective. In this case, participants were observed to be “smarter” following training simply because they believed the sham training would help make them smarter. It is an extreme example as to how expectation could exacerbate placebo effects, but effort and expectation can influence training or treatment, which in turn influences the perceived benefits of any training or treatment.

Concerning HBOT treatment, this issue means that participant expectations should be managed or assessed prior to the
Participants breathe room air under pressurized conditions. The same increase in inspired oxygen to the patient. Instead, pressure as a causal mechanism remains an understudied issue. This approach addresses both the experimental masking and dissolved oxygen concentration issues but does not address the potential therapeutic condition of increased pressure being a mechanism of action. Due to the inherent realities of physiological functioning, it is difficult to disentangle these factors as increased pressure and increased oxygen both provide a potential causal mechanism for therapeutic improvement. However, it is notable that previous research has primarily focused on controlling for dissolved oxygen concentration in the bloodstream while mimicking pressure. The assumption thus seems to be that oxygen concentration is the primary mechanism of action with pressure being a supplemental condition to affect dissolved oxygen. Providing more concrete experimental control of pressure conditions and the role of pressure as a causal mechanism remains an understudied issue regularly raised by proponents of HBOT to describe why previous sham conditions have been inappropriately employed for experimental control.

A third methodological issue regarding placebo effects involves the subjectivity of the dependent variables, and this issue extends the challenge of expectations. Placebo effects approach the efficacy of treatment effects when the disorder or impairment in question is amenable to a placebo. However, the opportunity to express expectations through subjective information represents the extreme combination of these challenges. Subjective dependent variables allow for an individual to unconsciously express and alter the results through unconscious means. For example, any variable directly assessing pain inherently asks a subjective question of the patient as pain is in and of itself a subjective assessment.

The solution to this methodological challenge is to incorporate objective metrics into experimental designs where possible. It is more difficult to observe placebo effects when using objective dependent variables, thereby making objective variables preferable over subjective variables, although this approach has its limitations. Objective metrics could provide discrete and measurable outcomes for analyses while remaining pliable to expectations and effort. Another solution is to use third-party metrics such as an independent rater who is blinded to the study goals. This approach would limit how patient expectations could influence subjective responses as they pertain to symptoms.

A fourth methodological issue addresses a subset of potential factors via inherent fluidity and changes in symptoms or data, such as regression to the mean. Unfortunately, these challenges are not new, and they affect all results dating back to early demonstrations of the placebo effect. Modern analyses of this classic work, almost ironically, indicated that there was no evidence of placebo effects in the studies cited.
by the original work describing the existence of a placebo effect. Among these alternative explanations that could explain these discrepancies, one aspect particularly relevant to mTBI, TBI, and PTSD is the potential of regression toward the mean. The general idea is that some improvement should be expected among participants selected for abnormality, in this case, taken to mean selected for impairment, as they naturally return to more normal functioning over time. The problem is particularly important for brain injuries where a primary recommendation is a rest, thereby anticipating some improvement over time as the expectation.

The solution depends largely upon the experimental design, but mostly, the means to address the issue require some sort of control group for comparison. In practice, no-contact control conditions are not generally considered to be advisable in experimental design to discount placebo effects, but with regard to HBOT, a no-contact control might be advisable to discount any natural healing or regression to the mean that could easily be confused as a placebo-like effect. These groups would describe changes in the observed patient symptoms and identify what would happen with no intervention or the extent of natural healing as it would occur. Alternatively, as some studies have done, another approach is to measure the time since injury. If the issue is chronic or if substantial time has passed since the injury or incident, then it is unlikely to expect continued natural healing. Still, if nothing else, studies should track the time since injury as well as injury severity as covariates about potential improvements.

Finally, a fifth methodological challenge to address has received almost no attention from HBOT investigations – a nocebo effect. Whereas placebo effects arise because there is some expected improvement, a nocebo effect represents the possibility that patient condition will decline because they expect the intervention to fail or expect to produce poor results. This problem does not contradict pre-treatment to post-treatment improvements observed during HBOT investigations, but rather the issue addresses the scale of observed improvements. Specifically, some extremely large effect sizes have been observed following HBOT treatments. Poor performance during pre-treatment assessments could lead to larger observed effects or at least larger potential effects at post-treatment assessments.

If participants are recruited into a study under the presumption that they are injured or impaired, their pre-test performance evaluation may yield suboptimal results because they expect to be performing poorly or responding with significant pain or impairment. This possibility does not disregard severe medical or psychological impairments that accompany mTBI, TBI, or PTSD – merely that, when patients know they are being evaluated for these reasons, poor performance on pre-treatment evaluations may be further exacerbated by patient expectations to perform negatively. In turn, patient symptoms or cognitive performance scores are likely to be near the floor, which allows for large potential performance improvements. A nocebo effect is thus the flip side to expectations, where rather than improving after treatment because the patient expects the treatment regimen to work, a patient unconsciously performs worse due to their expectations during pre-training assessment. Whereas measuring expectations regarding HBOT prior to treatment as a way to control for placebo effects induced by treatment expectations, one solution here would be to measure expectations about the impairments due to the medical or psychological issues (e.g., TBI, PTSD). Again, there are significant cognitive impairments often associated with brain injuries and/or PTSD. The point here is merely that assessments could be driven to floor effects or further exacerbated to poor performance during pre-training assessments by expectations.

A final thought is that the challenges presented here are not orthogonal or otherwise independent constructs. All five challenges can co-exist and interact in ways that exacerbate the problems of the others. For example, the first challenge discussed the expectation of a benefit, whereas the fifth challenge discussed the potential for nocebo effects in baseline testing. Both descriptions address how expectations could impact performance, outcomes, and interpretations, yet each impact could occur in different ways as nocebo effects impact baseline testing and expectations of improvement might have a greater impact on the post-treatment testing. These challenges are so interconnected when considering placebo effects that the best experimental designs should have some method for controlling all of them.

**Demonstrations of Placebo Effects**

To illustrate some challenges of placebos in experimental design with HBOT, we provide two hypothetical examples. The first is a case study example, and the second is an observational study. Each should provide a tangible construct for placebo issues. These examples are hypotheticals and not actual studies but provided for demonstration purposes to explore weaknesses in experimental design.

In the first example, a patient receives a concussive injury following a car crash. The doctor identifies the hallmark signs and symptoms of TBI and begins to plan a treatment. Part of the recommendations includes recent evidence that a colleague has seen a remarkable improvement in patient condition if HBOT can be administered within 72 hours of the injury. The scientific consensus is uncertain around this experimental approach, but the doctor – coming from a top-tier medical program – believes the work to be cutting edge. After discussion, the patient agrees and HBOT is administered. The patient reports significant reductions in symptoms, and the doctor reports the case study in a peer-review journal as a success.

This case study could have prompted placebos in numerous ways without explicitly producing an ethical violation on behalf of the doctor or hospital. Foremost, the doctor creates an expectation of robust improvement. Although as commonly stated that the plural of anecdote is not evidence, anecdotal evidence remains a persuasive potential to change hearts and minds. Scientific research can often be too abstract, whereas the visible improvements within an individual condition are highly salient. Expectations create the first possibility for a placebo. The second aspect is the lack of causality. A case study cannot truly assign a causal explanation to the outcome because this individual case could be impacted by too many factors. There is no control group, no randomization, and no guarantee that latent variables affected the outcome.
Case studies can be useful in an exploratory sense, yet their practical advancement of medical science is limited. The self-report nature of the demonstrated improvement also carries multiple problems. Patients are likely to be impaired immediately following the injury, and this case study design cannot discount naturalistic healing as part of the process. That is, the patient would likely have improved some with little or no intervention. The naturalistic healing cannot be separated from the experimental intervention in this design. Finally, the self-report nature of symptoms interacts with expectations and potentiates the change. This combination means that an impaired patient will begin closer to ceiling effects in symptom severity with the post-intervention improvement being attributable in part to the intervention, naturalistic healing, and implicit alterations in self-report due to expectations. The end result is as likely to be a placebo effect as a reliable and clinically significant improvement.

Another example involves open enrollment in an ongoing exploratory therapy. The design is conceived as an open-label, single-arm study where experimental care is made available to patients through a third-party, non-profit organization. In this specific example, the third party is a care advocacy group that specializes in treatment for wounded veterans. They work with various facilities to cover the costs of compassionate care avenues for people who had little success with standard treatment regimens. The study asks numerous questions about symptoms at pre-test, has the individual conduct a few cognitive tests, and then applies the same assessments post-intervention immediately after the final treatment.

Once again, this design opens up the potential for placebo effects despite no ethical violations. Moreover, all personnel involved are highly motivated to care for these wounded personnel, remaining optimistic in the face of little improvement. This motivation could contribute to placebo effects during intervention. Foremost, as with the case study example, it would be difficult to assign causality without a control group of some kind. A quasi-experimental design could help address this issue if some individuals are given the treatment initially and some participate in a masked control session designed to mimic the treatment, and then receive the treatment afterward. Both groups would receive the treatment at different times, which would help one group function as a control group despite eventually receiving the treatment. This design would be especially necessary in the case of third-party care advocates, who would be unwilling to support a control group where patients do not receive experimental care. However, this benevolent intention also creates a different environment that would be psychologically beneficial for the patient. Given the comorbidity of psychological issues in TBI/PTSD diagnoses, a wounded veteran may greatly benefit from such focused care and benevolent attention that this environment itself becomes psychologically therapeutic independently of the experimental treatment. The post-intervention assessment unfortunately has some of the same issues as the case study too, where a participant may report improvement because they enjoyed the intervention far more than other standard courses of care that produced no benefit. Likewise, this participant would come to the pre-intervention assessment with likely well-formed expectations about performance following brain injury. It creates the potential for nocebo effects at pre-test that might lower overall scores such that post-test scores could be significantly higher simply due to the removal of the nocebo effect.

**Best Practices in Experimental Design to Advance the Evidence**

Taken together, the five methodological challenges described here could independently or concomitantly contribute to a placebo-like effect in HBOT treatments. There are additional complications that could arise, although these five issues sufficiently illustrate the depth of challenge in demonstrating clinical significance when placebo effects can be so pervasive. The goal of additional research should be demonstrating clinical significance with HBOT interventions by addressing these placebo potentials or at least reducing their potential interference. To that end, there are several recommendations as to how experimental techniques could advance toward more comprehensive support for or against HBOT techniques as a therapeutic intervention.

**Case studies or observational reports are suboptimal efforts at this point**

There are sufficient case studies to warrant further investigations, and even accumulated observational evidence— in essence, a large collection of case studies— will not refute RCTs or other studies with more careful experimental control. The doubt is not in observable improvement, which has been demonstrated. The doubt is in the causal mechanism, which remains hotly contested. The exception to this point involves rare situations or complications, where case study evidence remains hotly contested. The exception to this point involves rare situations or complications, where case study evidence is among the most practical means of advancing scientific knowledge. Case studies and observational evidence should be the exceptions, not the rule, once a scientific discussion reaches a certain point of maturity, and HBOT treatments have reached the said point.

**Quasi-experimental design could be explored to great benefits**

In lieu of case studies or observational evidence, there are additional experimental designs that can and should be explored to greater benefit. Organizations already conducting HBOT interventions could address their evidence in terms of symptom severity prior to engaging in the treatment. By separating patients into categories based on pre-existing conditions prior to HBOT treatment, this approach provides a means of expanding the scientific contribution beyond simple observation. The categorical approach could be applied to a series of pervasive issues affecting the current body of evidence, including patient expectations and the severity of symptoms. Likewise, active and sham treatments could be applied to the same patient without denying treatment to anyone. For example, one group can receive a series of active treatment sessions followed by a series of sham treatment sessions, whereas another group can receive a series of sham treatment sessions followed by a series of active treatment sessions. This quasi-experimental design would allow organizations to provide treatment while also advancing the scientific contribution.
However, the continuing debate centers upon the causal mechanism by which the therapeutic benefit is observed – mostly notably due to the contradicting points that: 1) active interventions do not outperform sham, and 2) sham conditions could be active interventions themselves. No resolution to this debate will be achieved by simply applying more of the same evidence collected previously. Instead, researchers are advised to explore additional experimental designs to provide novel contributions regarding HBOT interventions. The goal should be reaching a level of clinical significance, which would allow health care providers to more definitely support or refute the technique. To that end, additional experimental evidence should address one or more of the concerns rose here, but there is a single, most important question that should be asked of every experimental design for HBOT studies in the near future: Does this new evidence contribute to the discussion of clinical significance for HBOT?

If so, then does it address multi-faceted aspects of clinical significance (internal validity, statistical significance, and practical significance) or is it intended to make a more incremental contribution? If not, how could the design be altered so that the study does contribute to clinical significance? Replication, while warranted for good scientific practice, will not provide either necessary or sufficient evidence to resolve the current HBOT debate. Researchers should instead focus on advancing the scientific evidence to directly answer the question of clinical significance.

**References**

1. TRICARE. TRICARE Policy Manual 6010.60-M. 2015.
2. Peterson K, Bourne D, Anderson J, Boundy E, Helfand M. Evidence brief: Hyperbaric oxygen therapy (HBOT) for traumatic brain injury and/or post-traumatic stress disorder. VA Evidence Synthesis Program Evidence Briefs. Washington (DC): Department of Veterans Affairs (US); 2011.
3. Hyperbaric Oxygen Therapy: Don’t Be Misled. Washington, DC: FDA Consumer Health Information, U.S. Food and Drug Administration. 2013.
4. Miller RS, Weaver LK, Bahraini N, et al. Effects of hyperbaric oxygen on symptoms and quality of life among service members with persistent postconcussion symptoms: a randomized clinical trial. *JAMA Intern Med.* 2015;175:43-52.
Simon R. Are placebo-controlled clinical trials ethical or needed when... Holbrook TL, Galarneau MR, Dye JL, Quinn K, Dougherty AL. Mor... Beecher HK. The powerful placebo. Carey JC. The importance of case reports in advancing scientific... Harch PG. Department of Defense trials for hyperbaric oxygen and... Wampold BE, Imel ZE, Minami T. The story of placebo effects in med... Reynolds T. The ethics of placebo-controlled trials. Wright JK, Zant E, Groom K, Schlegel RE, Gilliland K. Case report:... Stoller KP. Hyperbaric oxygen therapy (1.5 ATA) in treating sports re... Foroughi CK, Monfort SS, Paczynski M, McKnight PE, Greenwood... Weaver LK, Churchill SK, Bell J, Deru K, Snow GL. A blinded trial... Wampold BE, Minami T, Tierney SC, Baskin TW, Bhati KS. The pla... Shytle RD, Eve DJ, Kim SH, et al. Hyperbaric oxygen for mTBI-... Colloca L, Finniss D. Nocebo effects, patient-clinician communic... Gama Sosa MA, De Gasperi R, Janssen PL, et al. Selective vulner... Matchett GA, Martin RD, Zhang JH. Hyperbaric oxygen therapy... Behrens, et al. Measures of clinical significance. Kraemer HC, Morgan GA, Leech NL, Gliner JA, Vaske JJ, Harmon RJ. Measures of clinical significance. J Am Acad Child Adolesc Psychiatry. 2003;4:1524-1529. Bain Barbara A, Dollaghan Christine A. The notion of clinically significant change. Lang Speech Hear Serv Sch. 1991;22:264-270. Hindmarsh J, Quinan PT, Moore KL, Parkin C. The effects of black tea and caffeine on cognitive function and postural control: a randomized controlled trial. J Clin Nurs. 2020;29:11-20. Wampold BE, Imel ZE, Minami T. The story of placebo effects in medicine and psychotherapy from randomized clinical trials. J Clin Psychol. 2005;61:835-854. Wampold BE, Imel ZE, Minami T. The story of placebo effects in medicine: evidence in context. J Clin Psychol. 2007;63:379-390; discussion 405-408. Beecher HK. The powerful placebo. J Am Med Assoc. 1955;159:1602-1606. Kienle GS, Kiene H. The powerful placebo effect: fact or fiction? J Clin Epidemiol. 1997;50:1311-1318. McDonald CJ, McGlashan TH, Petrus Romanoski, McCabe GP, Jr. How much of the placebo effect is really statistical regression? Stat Med. 1983;2:417-427. Boot WR, Simons DJ, Stothart C, Stutts C. The pervasive problem with active control groups in randomized controlled trials. Mil Med. 2010;165:7470-7474. Wampold BE, Imel ZE. The place of placebo. J Cogn Enhanc. 2012;6:785-789.