ICARE PROTOCOL AND ANALYSIS PLAN

Study Title: A New Approach to Vision Therapy Based on Naturalistic 3-D Computer Gaming

Study ID: ICARE2B

Document Date: 8/30/2023

NCT number: NCT03066063

The research team will execute two concurrent studies within a common framework. The first will compare iCare to standard of care occlusion therapy (patching) for amblyopia in patients ages 4 to 18 (inclusive). The second will compare iCare to a standard of care for home-based CI therapy (pencil pushups) for convergence insufficiency patients ages 8 to 18 (inclusive). In the context of this study, the term iCare refers to either the amblyopia or CI video game module as appropriate for the respective study group.

Both the amblyopia and CI studies will follow the same basic approach. Each study will last 12 weeks, and participants will return for follow-up visits at Weeks 4, 8, and 12. Patients will receive full visual exams to establish baseline data upon enrollment, and again at Week 12 to record post-treatment data. During the interim visits (Weeks 4 and 8), patients' visual function will be assessed by a study optometrist, and any changes to the iCare configuration will be addressed as needed. As iCare is a home-based therapy, participants must have access to a personal computer and the Internet. For patients without access, the research team will provide loaner computers and/or pre-paid Verizon wireless mobile hotspots for Internet connectivity.

The proposed clinical studies will use the provider-as-investigator model that involves a minimum of 10 study sites and approximately 100 participants (see statistical power analysis below). With this approach, the research team will recruit optometry and vision therapy clinics and practices from around the country to support study execution under the training, guidance, and supervision of the iCare team. UND will serve as the "coordinating center" [48], enrolling, training, and overseeing the study sites and providing Institutional Review Board (IRB) governance. The UND team has experience in successfully serving in this multi-center coordination role in related research (see Biosketches). Enrolled providers (known hereafter as *investigator-providers*) will, in turn, recruit participants from their own patient rosters. This model serves several important purposes:

(1) Provides a geographically diverse pool from which to draw participants, ensuring research isn't skewed to local demographics (study cohort more closely matches the U.S. population);
(2) Expands the recruitment base, addressing the systemic challenge of meeting recruitment goals in research involving children [49];
(3) Implements a National Research Council-recommended strategy of involving community-based physicians in clinical trials in order to promote translation of research into clinical practice [1]; and
(4) Ensures that study procedures will easily be incorporated into practice [1] and that usability data gathered during study execution are valid.

The proposed effort is a Continuing Renewal application, with commercialization of the iCare product as its ultimate goal. By putting the iCare system directly in the hands of practicing optometrists and vision therapists, the research team can solicit feedback on features, commercialization models, and patient experiences. This will help ensure rapid transition of the test system to commercial adoption at the conclusion of the research effort.

To ensure protection of human subjects, integrity of data, and reliability of results, the UND-Barron Associates team will: provide rigorously-complete step-by-step checklists for all procedures; conduct webinar-based training on detailed protocols; and promulgate specific instructions on data protection, patient safety, and research ethics. All investigator-providers will complete Collaborative Institutional Training Initiative (CITI) [50] training and be registered with the UND IRB. During interactive online video training sessions, iCare team members, Drs. Clark-Gelburd and Biberdorf, will provide detailed instructions regarding study protocols to investigator-providers, including consent administration, inclusion/exclusion criteria, and assessments. Barron Associates, Inc. will provide thorough instructions on installation and use of the iCare system, the Provider Dashboard web app, and the parental-support web app. These will also be documented in printed manuals and online "how-to" videos. Each training webinar is expected to last 4 hours, with one-on-one "just-in-time" training available at any time. Follow-up teleconferences with the investigator-providers during study execution will reinforce patient protections and protocol adherence, as well as solicit recommendations on feature enhancements and integration of iCare into provider workflows.

The amblyopia and convergence insufficiency branches of the study will run concurrently at the participating

study sites. Study details specific to each are described below.

Amblyopia study

Children (ages 4-18) with amblyopia and meeting study inclusion/exclusion criteria will be recruited according to IRB-approved procedures. Informed consent will be sought from the parents or legal guardian with assent from the child. Enrollees will be randomized to one of two groups. A standard of care group will be prescribed patching for two or six hours per day, seven days per week. Standard of care group participants with best-corrected vision worse than 20/200 in their amblyopic eye will be prescribed six hours of patching daily, while participants whose best-corrected visual acuity is 20/200 or better will be prescribed two hours of patching daily. Participants assigned to an experimental group will play iCare vision therapy games for approximately 20 minutes per day, five days per week. Assignments will be made via a block randomization scheme based on a random permutation. With regards to age, it is expected that amblyopia patients between ages 4 to 6 years may experience up a one line (0.1 logMAR) greater improvement than those over 6 years of age [51]. For this reason, assignments within the amblyopia patient cohort will be age-balanced.

The primary investigational hypothesis is that the improvement in best-corrected visual acuity for participants in the iCare experimental group will not be inferior to the level of improvement observed in the standard of care (patching) group. An ETDRS (Early Treatment Diabetic Retinopathy Study) logMAR (logarithm of the minimum angle of resolution) chart [52] will be used as the primary assessment of visual acuity (VA). To reduce test-retest-variability, scoring will be based on an interpolated logMAR value (single-letter score) based on all measurements via the formula, $VA = 1.1 - T_C L_V$, where T_C is the total number of correctly read letters and L_V is the logMAR value of each letter on the chart (0.02 for ETDRS) [53]. In children 4-6 years old, a Lea Symbols chart [46] will also be used, with the best VA score from the two tests used for primary analyses.

The pre- to post-intervention change in best-corrected vision (Δ logMAR) at 12 weeks will serve as the primary efficacy endpoint. In terms of hypothesis testing, the primary null hypothesis will correspond to a non-inferiority test in which the null hypothesis will be that the mean of the Δ logMAR distribution for the iCare experimental study population is greater than that of the patching experimental study population (indicating inferior average outcome), while the alternative hypothesis will be that the mean Δ logMAR for the iCare experimental study population is equal to or more negative than the mean Δ logMAR patching experimental study population (indicating iCare therapy is non-inferior to patching therapy with respect to improving visual acuity). The null hypothesis will be rejected if the upper 95% confidence limit for the between-group difference (i.e. iCare – patching) in mean Δ logMAR does not exceed the non-inferior threshold of one line on an ETDRS chart (approximately 0.1 logMAR). Analysis will be performed by way of a linear mixed model (LMM) in which the pre- to post-intervention change (Δ logMAR) will be the LMM dependent variable, while the intervention (iCare vs. patching) will serve as the primary LMM independent variable and patient age, gender, pre-intervention logMAR in the amblyopic eye, and logMAR in the fellow eye will all serve as secondary LMM independent variables. These secondary variables will be utilized to adjust for any between-intervention disparities in age, gender, pre-intervention logMAR in the amblyopic eye, and logMAR in the fellow eye. Inclusion of these secondary independent variables in LMM will also allow assessment of the impacts on these characteristics on outcomes. Additionally, a stratification variable “study site” will be incorporated into the LMM as a random-effect to account for within study-site measurement correlation as well as between study-site variability.

Secondary analyses will test, on a per intervention basis, the null hypotheses that the mean Δ logMAR for each treatment group mean (iCare and patching) is greater than or equal to zero, with the alternative hypothesis that the mean Δ logMAR of the respective study-population is less than zero (indicating improved acuity associated with the respective treatment between pre-/post-testing). Additionally, systematic relationships between Δ logMAR and age, gender and per-intervention logMAR will be examined in secondary factorial analysis.

Statistical power analysis for the amblyopia study is based on selecting N such that there is at least an 80% probability of rejecting the primary null hypothesis (when the alternative hypothesis is true) that the mean Δ logMAR of the iCare experimental study population minus the mean Δ logMAR of the patching experimental study population exceeds the non-inferior threshold. Assuming a standard deviation in acuity improvement between subjects of approximately 2 lines on the ETDRS chart (0.2 logMAR – consistent with observed variability in previous studies [51], [26]), at least 50 patients must complete the amblyopia study (25 each in the iCare and patching groups) to achieve adequate statistical power. Thus, if the true mean of the Δ logMAR distribution for the iCare group is less than or equal to (better than) the true mean of the Δ logMAR distribution for the patching group, there is an 80% chance that the null hypothesis of iCare amblyopia therapy inferiority to patching amblyopia therapy will be correctly rejected. To account for potential dropouts, up to 64 total patients will be consented into the amblyopia study.

Convergence insufficiency study

The team will perform a clinical study comparing use of iCare by an experimental group to a standard of care (pencil pushups) control group for convergence insufficiency by patients ages 8 to 18 (inclusive). Pencil pushups are chosen as the control condition as it is the home-based therapy that is most often prescribed by both ophthalmologists and optometrists [33], [34]. Note that convergence insufficiency is rarely diagnosed and treated before age eight, thus the age range is different than for the amblyopia cohort. Informed consent will be sought from the parents or legal guardian with assent from the child. Enrollees will be assigned via a random permutation scheme to one of two groups: (1) the iCare experimental group, in which participants will be prescribed vision therapy gameplay for approximately 20 minutes per day, five days per week; and (2) a dosage-matched control group, which will be prescribed pencil pushups for 20 minutes per day, five days per week.

The primary investigational hypothesis is that improvement in positive fusional vergence base-out breakpoint (BOB) is greater for participants in the experimental group than it is for the control group. Convergence amplitudes will be measured with a base out prism bar with the patient focused on a target at 40 cm [54]. BOB will be recorded as prism magnitude at which fusion can no longer be maintained. Study investigators will also collect near point of convergence (NPC) data using an Astron International Accommodative Rule. Additionally, participants will be assessed via a CI Symptom Survey [32], a 15-item, 5-level Likert-scale. Individual responses will be assigned a value from 0-4 and summed to produce a score that can be compared to normative data [55].

The pre- to post-intervention change in BOB (ΔBOB) at 12 weeks will serve as the primary efficacy endpoint for the Phase IIB CI study. Analysis will employ a LMM with pre- to post-intervention change in BOB as the dependent variable and intervention group (iCare or control) as the primary independent variable. Additionally, the model will include age, gender, pre-intervention BOB, and study site assignment as secondary independent variables, allowing the LMM to account for variability due to these factors and enabling factorial analysis of their impact on outcomes.

The primary null hypothesis will be that the mean of the ΔBOB distribution is the same for the iCare and control groups, with the alternative hypothesis that the mean ΔBOB is greater (i.e. indicating greater improvement) for the iCare group. A one-sided $p \leq 0.025$ decision rule will be utilized to test the null hypothesis. Secondary null hypotheses will test, on a per intervention basis, whether the mean ΔBOB of the study-population is equal to zero, with the alternative that the respective study-population ΔBOB is greater (indicating improved functional vergence). Additionally, relationships between ΔBOB and age, gender, and pre-intervention BOB will be systematically examined employing the LMM. NPC data will be analyzed in an identical manner.

Statistical power analysis for the CI study is based on selecting N such that there is at least an 80% probability of rejecting the primary null hypothesis when the alternative hypothesis is true. With N=50 (25 patients per group) the underlying minimum detectable difference in the mean positive fusional vergence in convergence is 7.3 prism diopters (conservatively assuming a standard deviation of 9 prism diopters [32]). In a previous randomized controlled trial, office-based CI vision therapy was found to produce a mean difference in improvement after 12 weeks of greater than 10 prism diopters [32]. The Phase IIB CI study is therefore sufficiently powered. To account for potential dropouts, up to 64 total patients will be consented into the CI study.