Paper Critique - Tai Chi

Dominic LaRoche

October 29, 2015

1. The purpose of this research is to determine whether Tai Chi improves physical function and quality of life as compared to usual care and an exercise control. The authors hypothesize that Tai Chi exercise will lead to fewer falls, improved physical function, and better quality of life as compared to usual care or an established strength and range of movement exercise program for older adults (Silver Sneakers).
2. The authors claim the study design is a 3-group, single blind, randomized controlled trial.
3. I agree that the study is a 3-group , single blind, randomized controlled trial.
4. This was a randomized trial because allocation to one of three groups was random. This was a single blind randomized trial because the research staff assessing the outcome measures were blinded to group assignment. This is a controlled trial because and there was a control group of usual care (two controls if you count the exercise control).
5. Yes, this is hypothesis driven research so an analytic design is necessary. Moreover, a randomized controlled trial is appropriate because all three treatment groups are ethical, i.e. there is no reasonable expectation of undue harm to any of the participants, and the outcome measures can all be obtained over a relatively short time scale.
6. The target population is stroke survivors older than 50 with at least 3 months since a stroke.
7. The inclusion criteria:
   * Stroke survivor
   * 50 years of age or older
   * 3 months post-stroke
   * living in the greater Tucson area

The exclusion criteria:

* No disability
* Severe disability (requiring constant nursing care)
* Serious medical condition that would interfere with study participation (e.g. cancer)

1. Subjects were recruited from the community through a combination of media advertisements including radio, newspapers and brochures at community centers, outpatient rehabilitation centers, and physician offices.
2. There were 2 primary outcomes in the study: physical function and quality of life. The first outcome, physical function, was assessed in two ways. The Short Physical Performance Battery (SPPB) was used to asses overall physical function such as balance, and lower body strength. The investigators also collected information on patient-reported falls and near falls. A fall was defined as an event in which a subject ended up in the ground when they did not expect to. A near fall was defined as an event in which a subject recovered their balance without falling.  
   Quality of life was broken into 3 components: 1) general quality of life as assessed by the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36); 2) depression was assessed with the Center for Epidemiological Studies Depression Scale (CES-D); and 3) sleep quality was assessed by the Pittsburgh Sleep Quality Index (PSQI).
3. The three treatment groups represent the different exposures. The Tai Chi group was exposed to a 1 hour Yeng-style 24-posture short-form Tai Chi class developed by Fei 3 times per week for 12 weeks. The Silver Sneakers (exercise control) group was exposed to a 1 hour fitness program for older adults 3 times per week for 12 weeks. The usual-care control group was exposed to written materials and resources describing community-based physical activity programs suitable for older adults and a weekly phone call to inquire of their health status and provide personal attention (attentional control).
4. There are several sources of potential bias in this study. Since the intervention of interest requires physical activity it is possible that only those participants which are relatively healthy will able to complete the intervention throughout the study. If unhealthy participants in the exercise-based interventions drop out of the study then the remaining participants will be biased towards being healthy. This would lead to inflated measures of physical function and quality of life as compared to the control which would be less likely to lose the participants in poor health. Ultimately this would bias the results away from the null and towards finding a significant difference.

Another potential source of bias is the lack of blinding of the participants. Although impossible to avoid, not having participants blinded to their treatment group introduces the potential for a placebo effect for the two active interventions. This is especially true of difficult to measure outcomes like depression; participants may feel less depressed simply because they feel luck to have been allocated to the active intervention. This would bias the results away from the null.

Both of the active treatments involved much more time (3 hours per week compared to 1 phone call per week) interacting with a care provider. This interaction by itself could potentially improve the quality of life for participants without the need for any actual exercise or Tai Chi.

Another source of potential bias could have arisen in the allocation of individuals to treatment groups. Although the allocation was random, with small sample sizes unbalanced allocation can arise between treatment groups. In this study we see a higher proportion of white, married, college educated participants. All of these factors might be reasonably correlated with improved recovery from a stroke and therefore might bias the results away from the null. I should also note that the proportion of college graduates (79.3%) is higher than I would expect in the general target population which might reduce the generalizability of the results. This is likely a result of the recruitment procedure which would bias towards individuals interested in the intervention.

The authors used last-observation-carried-forward to replace missing data. This is known to bias the results but is not guaranteed to bias in any one direction (although typically this biases away from the null). They should have used multiple imputation or maximum likelihood estimators instead.

Finally, performing an intention-to-treat analysis could potentially bias the results towards the null since there were more drop-outs in the two intervention arms than in the usual care arm. If the drop-outs were not related to the intervention then this would bias the results towards the null. If the drop-outs were related to the intervention then the intention-to-treat analysis would eliminate the bias described in the first paragraph of this question.

1. The authors used a time-by-treatment interaction within an ANOVA to determine the association of treatment group with improvements in physical function and quality of life. The authors report statistics associated with testing for a difference in fall rates between the 3 groups.
2. As I mentioned earlier, due to the missing data, the authors should have used maximum likelihood based random-effects model to evaluate the treatment effect in order to avoid bias in the estimators. I also would have preferred to see relative risks and confidence intervals for pair-wise comparisons of the three treatment groups rather then just the test.
3. The authors claim that both Tai Chi and Silver Sneakers treatment groups experienced larger gains in aerobic endurance as compared with the control (5.9 more steps for Tai Chi, 16.3 more steps for Silver Sneakers, and 4.4 more steps for control). Additionally, the authors claim that Tai Chi lead to fewer falls than the other treatment groups.
4. The authors claim to have completed a sample size calculation and claim that the study may have been under-powered, although they do not provide any details about the sample size calculation. It is difficult to determine whether the study was under-powered because I don't know if the observed effect sizes were clinically meaningful.
5. I think the design of the study was a strength. Including a exercise and attentional control were both strengths of the design and the outcome measures were appropriate. The data analysis was a major weakness because some of the methods were inappropriate and the results were difficult to interpret as a result. The results table included standard deviations where it should have reported means and standard errors from the model, this makes it difficult to compare treatment groups. Also, a plot showing the time by treatment interaction for the two significant results would have aided interpretation.
6. No. Specifically, the claim that both Tai Chi and Silver Sneakers were associated with an increase in aerobic endurance (as measured by # of steps) may not be true. The claim is made based on the significant F-test associated with the time by treatment effect. However, this F-test only verifies that there was a difference between the groups and does not ascribe significance to any single pair-wise comparison. In order to determine if Tai Chi improved aerobic endurance compared to the control they would have had to construct a contrast to estimate the difference between the two time by treatment effects. Also, a marginally significant (p=0.05) difference in fall rates from a post-hoc test is not terribly convincing. However, it does appear that the sample size was reasonably small so the study may indeed have been under-powered. The final claim that Tai Chi should be investigated further for fall prevention is justified given the marginally significant result with respect to falls.