The YODA Project
Research Proposal Review

The following page contains the final YODA Project review approving this proposal.

The YODA Project
Research Proposal Review - Final
(Protocol #: 2016-0765)

Reviewers:

- ☒ Nihar Desai
- ☐ Cary Gross
- ☒ Harlan Krumholz
- ☒ Richard Lehman
- ☒ Joseph Ross

Review Questions:	**Decision:**
1. Is the scientific purpose of the research proposal clearly described?	Yes
2. Will request create or materially enhance generalizable scientific and/or medical knowledge to inform science and public health?	Yes
3. Can the proposed research be reasonably addressed using the requested data?	Yes, or it's highly likely
4. Recommendation for this data request:	Approve

Comments:

No additional comments.

The YODA Project
Research Proposal Review

Revisions were requested during review of this proposal.
The following pages contain the original YODA Project review and the original submitted proposal.

The Yale University Open Data Access (YODA) Project
Yale University Center for Outcomes Research and Evaluation (CORE)
1 Church Street, Suite 200, New Haven, CT 06510

The YODA Project
Research Proposal Review - Revisions Requested
(Protocol #: 2016-0765)

Reviewers:

- ☒ Nihar Desai
- ☐ Cary Gross
- ☒ Harlan Krumholz
- ☒ Richard Lehman
- ☒ Joseph Ross

Review Questions:	Decision:
1. Is the scientific purpose of the research proposal clearly described?	No
2. Will request create or materially enhance generalizable scientific and/or medical knowledge to inform science and public health?	Unsure, further clarification from requestor is needed
3. Can the proposed research be reasonably addressed using the requested data?	Unsure, further clarification from requestor is needed
4. Recommendation for this data request:	Not Approve

Comments:

The authors just report that they are looking to identify the most "efficient" analytic methods but I remain unclear about how they determine what is "efficient." How will they determine if they have identified an improved approach? How is this defined prospectively? What is the gold standard?

In addition, in the dissemination section, they say the manuscript will be written by a "publication committee." I think this should be further clarified as well.

It is not clear to me what the stated objective of determining the "most efficient cognitive measurement and analysis method for cognition and dementia" really means. I do not think this can be done by data analysis from past clinical trials - it would require prospective clinical study.

The investigators need to clarify (a) their objective and (b) the analysis method.

Principal Investigator

First Name: Philip
Last Name: Bath
Degree: BSc MB BS MD FRCPath FRCP FESO FAHA DSc
Primary Affiliation: University of Nottingham
E-mail: polly.scutt@nottingham.ac.uk
Phone number: 01158231769
Address: Clinical Sciences Building
City Hospital Campus, Hucknall Road
City: Nottingham
State or Province: Nottinghamshire
Zip or Postal Code: NG5 1PB
Country: United Kingdom

2016-0765

General Information

Key Personnel (in addition to PI): **First Name:** Polly
Last name: Scutt
Degree: MSc Medical Statistics
Primary Affiliation: University of Nottingham

Are external grants or funds being used to support this research?: No external grants or funds are being used to support this research.

conflict_of_interest_philip_bath.pdf
conflict_of_interest_polly_scutt.pdf

Certification

Certification: All information is complete; I (PI) am responsible for the research; data will not be used to support litigious/commercial aims.
Data Use Agreement Training: As the Principal Investigator of this study, I certify that I have completed the YODA Project Data Use Agreement Training

Associated Trial(s): NCT00253201 - Efficacy, Tolerability and Safety of Galantamine in the Treatment of Alzheimer's Disease
NCT00304629 - Long Term Safety and Efficacy of Galantamine in Alzheimer's Disease (Extension INT-8)
NCT00253227 - Galantamine in the Treatment of Alzheimer's Disease: Flexible Dose Range Trial
Long Term Safety and Efficacy of Galantamine in the treatment of Alzheimer's Disease
Long Term Safety and Efficacy of Galantamine in the treatment of Alzheimer's Disease
What type of data are you looking for?: Individual Participant-Level Data, which includes Full CSR and all supporting documentation

Research Proposal

Project Title

Improving the statistical analysis of cognitive outcomes in randomised controlled trials: The 'Optimising the Analysis of Cognition' Collaboration

Narrative Summary:
Over 800,000 people suffer with dementia in the UK. Despite being common; costly in economic terms and devastating to patients and their families, the evidence base for the treatment of dementia is small. This may be because statistical analyses used are suboptimal. OA-Cog aims to identify the most efficient analysis technique for cognition outcome data and dementia in randomised controlled trials. Chief investigators of RCTs with cognitive outcome date are asked to share their individual patient data. Data are then analysed using various endpoints and statistical methods in order to identify which is the most efficient. Approaches for dealing with patients who die will also be addressed.

Scientific Abstract:
Background: Over 800,000 people suffer with dementia in the UK. The evidence base for the treatment of dementia is small. One reason for this may be that the measures used to assess cognition in clinical trials are not sensitive to change and/or the analyses used are suboptimal.
Objective: OA-Cog aims to identify the most efficient cognitive measurement and analysis method for cognition and dementia in randomised controlled trials including patients with or at risk of vascular dementia or Alzheimer's disease.
Study Design: Chief investigators of randomised controlled trials with cognitive outcome assessments are asked to share individual patient data from their trials. Variables requested include baseline prognostic factors, treatment group, cognitive measures (e.g. Mini Mental State Examination (MMSE)) and other outcome measures (e.g. death, dementia).
Data from shared trials will be reanalysed using various statistical methods to identify which is most efficient.
Participants: Participants with or at risk of vascular dementia or Alzheimer's disease will be included.
Main outcome measure: A number of global cognition outcomes will be analysed (e.g. MMSE).
Statistical analysis: Data will be analysed using various endpoints (e.g. mean MMSE score at end of trial, MMSE score as a gradient over time) and statistical methods (e.g. Wilcoxon rank-sum test, repeated measures ANOVA) in order to identify the most efficient. Approaches for dealing with missing data and in particular missing data due to death be addressed; currently, such patients are ignored from analyses.

Brief Project Background and Statement of Project Significance:
Over 800,000 people suffer with dementia in the UK [1]. Despite being common; costly in economic terms to society and devastating to patients and their families, the evidence base for the treatment of cognitive decline and dementia is small. One reason for this may be that the measures used to assess cognition in clinical trials are not sensitive to change and/or the analyses used are suboptimal. OA-Cog aims to assess this hypothesis.
Taking a commonly used global measure of cognition, the mini mental state examination (MMSE), several approaches for analysis of this outcome have been used in completed trials:
1. Comparison of proportions with cognitive decline, i.e. MMSE<24 [2, 3]
2. Comparison of mean MMSE on treatment [4, 5, 6]
3. Comparison of change in MMSE from baseline to on treatment [7, 4, 8, 9, 10]
4. Comparison of change in MMSE assessed as a gradient

From a statistical perspective, methods which utilise multiple measurements, including at baseline, are likely to be more powerful (efficient) than those which simply compare a measurement near to or at the end of trial. However, this approach does not address all other problems:
1. Death. Patients who die are usually ignored from analyses of cognition in contrast to trials assessing functional outcome after stroke, or vascular prevention, where death is included in the main outcome. However, excluding death may have unintended effects:
a. Many treatments, which reduce cognitive decline and dementia, may also reduce death providing death from dementia explains a large proportion of the deaths (or at least that the intervention reduces another important cause of death, e.g. from vascular events); here, ignoring death reduces the power of the analysis.
b. Alternatively, a treatment which reduces cognitive decline and dementia but increases death might appear to be safe and effective since the hazard would not be included in the primary outcome. This scenario is present for thrombolysis in acute stroke where alteplase reduces disability/dependency and increases death [11]. Since the latter effect is smaller than the former, the combined outcome of death or dependency is reduced with thrombolysis.

2. Missing data. Where missing data on cognition is common, drop-outs may not occur equally in the treatment groups, thereby leading to potential over or under estimation of treatment effect; this problem was studied in the SHEP trial [12].

This project could influence the design and analysis of future trials assessing cognition. If a more efficient analysis method is adopted in future trials then a smaller sample size would be required to detect a significant treatment effect. This would benefit patients as fewer would need to be exposed to interventions that may be harmful. Smaller trials would also benefit research funding bodies and healthcare professionals as they would cost less and be less demanding in terms of resources. This would mean that research funding bodies would be able to fund more much needed trials in this area to improve the evidence base.

Specific Aims of the Project:
• To identify the most efficient statistical analysis method for analysing global cognition data (e.g. MMSE, ADAS-Cog) and dementia from randomised controlled trials including patients with or at risk of vascular dementia or Alzheimer's disease.
• To identify the most efficient cognitive measurement for use in randomised controlled trials including patients with or at risk of vascular dementia or Alzheimer's disease.
• To identify the most appropriate missing data technique for use with cognitive data from randomised controlled trials including patients with or at risk of vascular dementia or Alzheimer's disease.
• To assess how patients who have died should be included in the analysis.
• To assess the implication of choosing particular methods of analysis on trial sample size
• To assess the implication of adjusting analyses for baseline prognostic factors on sample size

What is the purpose of the analysis being proposed? Please select all that apply. Participant-level data meta-analysis
Participant-level data meta-analysis will pool data from YODA Project with other additional data sources

Research Methods

Data Source and Inclusion/Exclusion Criteria to be used to define the patient sample for your study:
Eligible trials include randomised controlled trials in patients with or at risk of vascular dementia or Alzheimer's disease with on treatment global cognition data. Trials of any intervention will be considered for the project. Trials included in the project will either be positive (i.e. showing statistical evidence of benefit), negative or from a meta-analysis showing statistically significant benefit or harm. There are no individual patient inclusion/exclusion criteria.

Main Outcome Measure and how it will be categorized/defined for your study:
A number of outcomes will be analysed as the aim of the study is to identify the most efficient for the analysis of cognition in patients with or at risk of Alzheimer's disease or Vascular dementia. Outcomes of interest to the study include:
Continuous outcomes:
• Mini-mental state examination (MMSE)
• Alzheimer's disease assessment scale – cognitive sub-score (ADAS-Cog)
• Montreal cognitive assessment (MoCA)
• Addenbrooke's cognitive examination – revised (ACE-R)
• Clinical Dementia Rating Sum of Boxes (CDR-SB)
• Trail making test
• Finger tapping test, new dot test and maze tracing test
• CDR and category fluency test
• Modified Boston naming test
• Raven's coloured matrices
• Telephone interview for cognitive status
Ordinal outcomes:
• Ordinal cognition which combines impairment and dementia: normal score (MMSE>28) / mild impairment (MMSE 23 – 28), moderate impairment (MMSE 10 – 22) / severe impairment (MMSE < 10) / dementia
Binary outcomes:
• Dementia (yes, no) and time to diagnosis
• Death (yes, no) and time to death.

Main Predictor/Independent Variable and how it will be categorized/defined for your study:
The main predictor variable will be randomisation allocation. Randomisation will be categorised as active versus control (as defined in each trial protocol). For factorial and partial factorial trials each comparison will be treated as a separate trial.

Other Variables of Interest that will be used in your analysis and how they will be categorized/defined for your study:
Other variables which Chief investigators of trials have been asked to share include:
Demographics and risk factors: age, sex, history of hypertension, smoking status, hypercholesterolemia, previous stroke, previous MI, diabetes mellitus, history of cognitive impairment, history of dementia, atrial fibrillation, education and family history of dementia.
Baseline clinical data: blood pressure, cholesterol, glucose, ECG, stroke type and severity (if applicable), global cognition scale (e.g. MMSE, ADAS-Cog) score, APOE ?4 carrier, mood/depression (yes/no or scale and score) and functional status (e.g. modified Rankin scale).
Trial characteristics: randomisation (active, control), randomisation date, start date of intervention, indication for trial (e.g. Alzheimer's disease, mild cognitive impairment, stroke, hypertension) and length of follow up.
Other non-cognitive outcomes measures: functional outcome (e.g. modified Rankin scale), quality of life, mood/depression (yes/no or scale and score) and stroke during trial.

Statistical Analysis Plan:
Individual patient data received will be reformatted and merged into one central database with standard variable names and coding. Each trial will be given a unique identifier.
The following analyses will be performed on the MMSE at the end of the trial ("end of trial" is defined as the last visit post baseline where >80% of participants have a score or the post baseline with the highest percentage of non-missing data if no post baseline visits have data for >80% of participants):
1. Proportions with cognitive decline and dementia: chi-square test; binary logistic regression.
2. Binary cognition with only one cut off point (MMSE ?16 vs >16): chi-square test; binary logistic regression.
3. Mean score at end of treatment: two-sample t-test (parametric)
5. Median score at end of treatment: Wilcoxon (Mann-Whitney U) test (non-parametric)
6. Mean change in score from baseline: two-sample t-test
7. Gradient of score (needs score at multiple time points): repeated measures ANOVA
8. Ordinal cognition (5 level) which combines impairment and dementia: normal score (MMSE>28) / mild impairment (23 – 28) / moderate impairment (10-22) / severe impairment (<10) / dementia: ordinal logistic regression
9. Ordinal cognition (4 level) which combines impairment and dementia: normal score (MMSE>23) / impairment (MMSE<23) / dementia: ordinal logistic regression.
Cognitive data from each dataset will be analysed using every statistical method mentioned above. The results of the analyses will be ordered within each dataset and given a rank, with the lowest rank given to the method which produced the most significant result (the smallest p-value) in that dataset. A 2-way analysis of variance test will be used to see on average which statistical method produced the lowest ranks. The methods will be ordered in terms of their efficiency in identifying treatment effects using Duncan's multiple range test.
The validity and reliability of the results will be assessed by comparing statistical methods within subgroups of trials sharing similar characteristics, e.g. AD trials. The statistical assumptions underlying each method will be assessed within each dataset to see if and how often these assumptions are violated.
The effects of different statistical methods on sample size will be established by calculating the sample size for each dataset based upon each statistical method. The methods will then be ranked using the same methodology as for finding the most efficient statistical analysis method.
The effect of adjusting for prognostic factors on sample size will be established by calculating p-value for each dataset from both the adjusted and unadjusted analyses. The reduction in sample size will then be calculated using the mean adjusted and unadjusted p-values.
To identify the most appropriate missing data technique for use in trials assessing cognition a random subset of patients will be removed from each dataset. Each missing data technique will then be applied in turn to each dataset. The most appropriate missing data technique will be the one which most closely matches the data that was removed. This process will be repeated a number of times removing a different random subset of patients to make sure the findings are robust. This method will compare techniques for dealing with missing data assuming the data is MCAR, however, this may not be the case. The process above will be repeated but simulating different mechanisms of missing data.
The most appropriate missing data technique to use with patients who die may be different from the most

appropriate technique used for other types of missing data. For this reason the following approaches for analysing patients who die will also be assessed:
1. Assign worst state, i.e. MMSE=-1
2. Last value when alive carried forward (needs scores at multiple time points)
3. Gradient when alive
4. Ordinal cognition (which includes death)
All analyses will be carried out in SAS (version 9.3).

Project Timeline:
Data collection will continue until third quarter 2016. Coding of analyses has begun and will continue until second quarter 2017. Manuscripts will be written once coding and interpretation of analyses has been completed. Once draft manuscripts have been written, they will be sent to collaborators for comments and then submitted for publication.

Dissemination Plan:
The results of analyses for individual trials will not be published (since the Collaborators have already done this); rather, the results of different analysis methods across the trials will be compared and published. As a result, individual trials will not be identifiable. The results will be published under the banner of OA-COG with the collaboration listed by name. Collaborators will be listed in the acknowledgments by trial. The papers will be written by a 'Publication Committee' and then distributed to all Collaborators for comment, interpretation, changes, additions etc. The results will be submitted to major dementia conference(s) for presentation, and submitted for publication in major quality peer-reviewed journal(s). Data will not be used for any purpose other than to do with OA-COG.

Bibliography:
[1] Prince, M. et al., Dementia UK: - overview. 2014.
[2] Tzourio, C., et al., Effects of blood pressure lowering with perindopril and indapamide therapy on dementia and cognitive decline in patients with cerebrovascular disease. Arch Intern Med, 2003. 163(9): p. 1069-75.
[3] Heart Protection Study Collaborative Group, MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20536 high-risk individuals: a randomised placebo-controlled trial. Lancet, 2002. 360: p. 7-22.
[4] Prince, M.J., et al., Is the cognitive function of older patients affected by antihypertensive treatment? Reuslts from 54 months of the Medical Research Council's treatment trial of hypertension in older adults. British Medical Journal, 1996. 312: p. 801-5.
[5] Simons, M., et al., Treatment with simvastatin in normocholesterolemic patients with Alzheimer's disease: A 26-week randomized, placebo-controlled, double-blind trial. Ann Neurol, 2002. 52: p. 346-350.
[6] Sparks, D.L., et al., Atorvastatin for the treatment of mild to moderate Alzheimer disease. Arch Neurol, 2005. 62: p. 753-757.
[7] Applegate, W.B., et al., Impact of the treatment of isolated systolic hypertension on behavioral variables. Arch Intern Med, 1994. 154.
[8] Skoog, I., et al., Effect of baseline cognitive function and antihypertensive treatment on cognitive and cardiovascular outcomes: study on COgnition and Prognosis in the Elderly (SCOPE). American Journal of Hypertension, 2005. 18: p. 1052-59.
[9] Shepherd, J., et al., Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. Lancet, 2002. 360: p. 1623-1630.
[10] Muldoon, M.F., et al., Randomized trial of the effects of simvastatin on cognitive functioning in hypercholesterolemic adults. The American Journal of Medicine, 2004. 117(11): p. 823-9.
[11] Wardlaw, J.M., et al., Thrombolysis for acute ischaemic stroke. Cochrane Database Systematic Review, 2003(3).
[12] Di Bari, M., et al., Dementia and disability outcomes in large hypertension trials: lessons learned from the systolic hypertension in the elderly program (SHEP) trial. American Journal of Epidemiology, 2001. 153(1): p. 72-78.