Recruiting underserved communities to a self-management intervention for people with Parkinson's (PD-Care SWAP)

Statistical Analysis Plan

Version 0.2
26 November 2024

1 Introduction

1.1 Purpose and scope of the statistical analysis plan

This statistical analysis plan (SAP) contains details of the statistical analyses for the PD-CARE trial Study Within a Project (SWAP) to explore if we were able to recruit a diverse sample for the trial and if intervention effectiveness is different for those in underserved groups. This SAP covers the quantitative aspects of this SWAP. Additional work exploring the impact of the intervention on these groups, including qualitative interviews as part of the process evaluation are reported elsewhere.

This Sap should be read as an appendix to the main trail SAP (V1.3).

1.2 Authorship

This SAP has been written by Tasmin Rookes (TR), based on the main trial SAP written by Mariam and Gareth.

2 Trial Population

2.1 Recruitment and retention

We will explore recruitment and retention based off the specific underserved group characteristics, to see if there are systematic differences in those who withdraw or are lost to follow-up based on these characteristics.

2.2 Baseline Characteristics

We will summarise the baseline underserved groups characteristics, separated by study arm. Categorical variables will be reported as counts and percentages. Continuous variables will be summarised as means and standard deviations (SD) or medians and interquartile ranges as appropriate depending on the distribution of the data.

These underserved group characteristics will be:

- Ethnic minority status
- Age left full time education
- Socioeconomic status (indices of multiple deprivation)
- Rurality

We will compare the demographics of this sample with the wider population, using census data to explore if we have recruited a representative sample of people from underserved groups. These characteristics include age, gender, ethnicity, age left education (if available), index of multiple deprivation (IMD – decile or quintile), and rurality (if available). One-sample tests of means/proportions will be conducted to determine is there are statistical differences between the trial sample and the wider population.

3 Statistical analysis plan

3.1 Primary outcome analysis

The primary outcome is PDQ-39 score at 12 months. We will conduct mixed model regression including the PDQ-39 scores and the underserved group variables with an interaction term between the intervention group variables to enable estimation of the intervention effect at 12 months depending on if participants in part of an underserved group or not. These will be done for ethnicity, education, deprivation, and rurality as individual measures. We will then combine these variables to create two groups of people considered to be underserved (any combination of the above four variables) and those who are not. This model will also adjust for baseline PDQ-39 score, age and gender. All analyses will be performed on an intention-to-treat basis and all modelling assumptions will be checked (e.g. using residuals). The effect of intervention and corresponding 95% confidence interval will be reported.

3.2 Secondary outcome analysis

The effect of the intervention on some secondary outcomes will be assessed as exploratory analysis using appropriate two-level regression models, i.e. linear mixed models for numerical outcomes. All models will be adjusted for age and baseline outcome (if available) using fixed effects. The effect of intervention and corresponding 95% confidence interval will be reported., though P-values will not be reported for secondary analyses. These analyses will be considered supportive.