

1. STATISTICAL CONSIDERATIONS

1.1. Statistical Analysis

A Chi Square Test of Independence will be applied to assess whether the prevalence of actionable genotypes is the same in the IPTIP cohort as in existing control datasets (ie. TARDIS and Vanderbilt datasets).²⁴ The criterion for an actionable variant is as previously detailed (Section 7.0). The primary outcome measure between the IPTIP-I and IPTIP-O cohorts will be compared using the same statistical approach. Statistical differences in genotype based on demography, or recruiting site, will be tested via ANOVA or independent t-test. Logistic regression analysis will be used to determine the predictive value of patient variables, such as age, co-morbid status, or ethnicity, for exposure medicines where pharmacogenetic guidelines exist. Statistical support will be provided throughout the trial by an independent trial statistician, Mr Duncan Stoddard.

1.2. Sample Size:

Sample Size: Determined based on the primary outcome and calculated with assistance from the NIHR Research Design Service and an independent statistician, Duncan Stoddard. Sample size is calculated using the formula $n = Z^2 x P x (1 - P) / e^2$ where z is value from standard normal distribution corresponding to desired confidence level (Z=1.96 for 95% CI), P is expected true proportion, e is desired precision.

For the inpatient arm, a previous estimate from Vanderbilt, USA, suggests that 40% of patients had a genetic variant related to a medication that they were prescribed. Given the diversity of the population presenting to MFT and the relative over-medicalisation of healthcare in the USA, a lower prevalence of 0.3 will be chosen to represent P which, with a desired precision of 0.05 and a CI of 95%. This results in a required sample size of **323**.

For the outpatient arm, feasibility trials in Vanderbilt, USA, and the Netherlands suggest that approximately 30% of patients have a genetic variant which would impact prescribing of a medication they were actually prescribed.^{26,27} Therefore, a P of 0.3 will again be used, assuming the same precision and CI as above, generating a required sample size of **323**.