The text below is a direct copy of what is written in the iFRAP protocol. It forms version 1.0 of the analysis plan

1 STATISTICS AND DATA ANALYSIS

1.1 Sample size calculation

This study is powered to detect a between group effect size of at least 0.4 in the primary outcome at 2-week follow-up, with 2-tailed 5% significance and 80% power. With an estimated standard deviation of 15,[75, 76] this translates to minimum clinically important difference (MCID) of 6 points on the DCS (scale range 0 – 100) – a difference considered by the study team to be meaningful and one that produces an effect size in the range of meaningful effect sizes recommended by the authors of the tool.[1] To achieve an effect size of 0.4 between the study arms, we plan to randomise 328 patients. This recruitment target assumes that approximately 32% of patients will not receive a treatment recommendation (hence for whom the primary outcome is not relevant) and for 10% loss to follow-up in the primary outcome at 2-weeks; our target at 2-week follow-up is therefore 200 i.e. 100 per arm.

The figure of 68% of patients receiving a treatment recommendation is derived from estimates that 40% of patients at Stoke will receive a drug recommendation and 90% at Oxford and Portsmouth[1] and that Stoke will contribute to 43% of recruitment.

1.2 Planned recruitment rate and feasibility

We estimate that an average FLS identifies approximately 120 patients per month for FLS appointments, 90% will meet the eligibility criteria (N = 108) and that 20% will consent to the study (N=21 per month). Enrolling two FLSs to the trial, who each recruit for 5-months and a further site which recruits for 8 months yield a potential sample size of (N= 378), which would be sufficient pool of patients to achieve our randomised sample size of 327 participants.

1.3 Statistical analysis plan

The full statistical analysis plan will be written and agreed with the Trial Steering Committee prior to analysis, hence only an outline of the analysis is described below.

1.3.1 Summary of baseline data and flow of patients

A CONSORT flow diagram will be produced to describe the number of participants included in the study at the multiple stages of recruitment and follow-up. Any reported serious adverse events and protocol deviations will be reported throughout the study.

---

[1] These are different because of different service models. Oxford and Portsmouth assess all patients before consultations so in theory, all patients seen in consultation should be offered drug recommendation and be able to complete primary outcome. We have estimated there may be a small number of consultations in which a drug recommendation is not given for unforeseen reasons (10%). In Stoke, patients are assessed in the appointment meaning low risk patients are not offered treatment. Source of 40% figure - FLS-DB National Audit – most conservative from range 40-55%

Descriptive statistics (means, standard deviations, medians, inter-quartile ranges and numbers and percentages as appropriate) will be used to describe the baseline characteristics of participants at each stage of recruitment and follow-up, and by blinded treatment arm, to assess if there is any evidence of selection bias and to evaluate the success of the randomisation procedure.

1.3.2 Primary outcome analysis

The primary outcome analysis will be conducted blind to treatment arm, on an intention-to-treat basis. Analysis of covariance (ANCOVA) will be used to analyse the primary outcome (the total DCS score) at the 2-week primary endpoint, by comparing the mean outcome in each treatment arm, after adjustment for any pre-specified baseline covariates. The results of the primary analysis will be presented as means and 95% confidence intervals.

1.3.3 Secondary outcome analysis

Treatment effects for secondary outcomes measured at a single follow-up time-point will be explored using similar methods to the primary outcome analysis, but with ANCOVA, logistic and ordinal regression used as appropriate for continuous, binary, and ordinal outcomes. For outcome measures collected at more than one time-point, linear mixed models will be used to model change in the outcome over time. Results will be presented either as mean or percentage differences/odds ratios alongside their associated 95% confidence intervals.

Descriptive statistics (numbers and percentages) will be used to describe patients experience of their FLS appointment and exploratory analysis conducted to explore whether patients' perception of their fracture risk changes following the intervention, and whether their post intervention perception of risk is more in line with their predicted fracture risk (as calculated by their FRAX score). In addition, we will also explore whether patients' level of worry about falls and fractures changes following their FLS consultation and whether such changes are similar in both trial arms.

1.3.4 Adjusted analysis

Covariates included in an adjusted analysis will be specified a priori in the analysis plan, but are likely to include the baseline in the outcome of interest (if this is measured for the outcome of interest) and the stratification variable used in the randomisation process, along with other key baseline variables of interest e.g. age, health literacy and socio-economic status. The adjusted model will be the primary model used in the study.

1.3.5 Subgroup analyses

Any subgroup analyses completed in the trial will be exploratory and will be stated in the analysis plan *a priori* before the main trial analysis is undertaken.

1.3.6 Sensitivity analyses

A sensitivity analysis of the primary outcome model will be conducted to explore whether study conclusions change when outcome variation between FLS clinicians is accounted for in the model. This will be achieved by adapting the ANCOVA model in section 1.3.2 into a mixed model framework and incorporating a random effect term to represent the clinician who treated the patient. The magnitude of the treatment effect from this model will then be compared to that from the primary analysis in the study.

1.3.7 Interim analysis and criteria for the premature termination of the study

No interim analysis is planned to be undertaken during the study to assess effectiveness.

1.4 Subject population

We will analyse all participants who have been randomised to the study on an intention-to-treat basis, except for those outcome measures collected only for participants who receive a drug recommendation. These latter outcomes will be evaluated on an intention to treat basis but only for patients who have received a drug recommendation from the treating clinician.

1.5 Procedure(s) to account for missing or spurious data

We will consider using multiple imputation to impute the patient-level missing data if the missing data rate is greater than 5% for at least one patient-level outcome or predictor of interest [77]. If multiple imputation is used, this will be regarded as the primary analysis over a complete-case analysis.

1.6 Other statistical considerations

Logistic regression will be used to develop a predication model to predict treatment initiation at 3-month follow-up in those participants who were given a drug recommendation. Candidate predictors will be listed *a priori* in the analysis plan, but are likely to include: age, sex, fracture risk, self-reported receipt of osteoporosis diagnosis, consultation length, consultation modality (face-to-face vs telephone), receipt of a DXA scan, level of health literacy and socioeconomic status. In addition, we will use linear regression to explore whether factors (such as age, sex, health literacy and socioeconomic status) are associated with the primary outcome of interest (the DCS).

1.7 Economic evaluation

The economic evaluation will comprise a within-trial cost-effectiveness and value of information (VoI) analysis to determine whether the iFraP intervention is cost-effective compared with usual care. A detailed record of all the resources (i.e. financial, staff, equipment, training, etc.) required to set-up and deliver the iFraP

intervention will be made. Resource use information will be obtained from patient self-reports (see sections 7.3.4 and 11.10) and the medical record review.

We will conduct a trial-based Bayesian value of information analysis to identify the main sources of uncertainty regarding the value for money of iFraP when compared with usual care.