Statistical analysis plan (SAP) for «Motivational interviewing in long-term sickness absence»

This plan will cover items not covered in the published protocol article *Motivational interviewing in long-term sickness absence: study protocol of a randomized controlled trial followed by qualitative and economic studies* (BMC Public Health. 2018 Jun 18;18(1):756) or at ClinicalTrials.gov (NCT03212118).

SAP version: March 30th 2020

Written by: Lene Aasdahl

Statistical interim: No interim analyses are planned.

Timing of analyses: The main analyses will take place 12 months after the last participant is included in the study. Register data for use of medical benefits (sick leave, work assessment allowance and disability pension) will be collected for all participants for 12 months follow-up after inclusion in the study.

Data sources:

Sick leave data: will be measured using data from the National Social Security System Registry, where all individuals receiving any form of benefits in Norway are registered by their social security number. The data consists of registrations of medical benefits from four different sources; sick-leave payments, sick leave certificates, work assessment allowance and disability pension. Sickness absence will registered both as number of days per month and as a dichotomous measure of whether or not the participant is registered on sick leave that month. By combining information from the different medical benefits we will calculate days on medical benefits (according to a 5-day work week) for every month during follow-up. Time on graded sick leave will be transformed to whole workdays. Days receiving sick-leave payment and work assessment allowance will be adjusted for employment fraction, including a graded disability pension at inclusion. Any increase in disability pension during follow-up will be counted as sick leave. Monthly intervals (rather than exact dates) will most likely be used in order to include all relevant sick leave benefits in the same measure, as exact dates are not available for payments and the long-term benefits. If this data is available in the future, exact days will be used if possible.

Questionnaires: self-reported web-based questionnaires are answered at five time points: before randomization (T1), before the intervention (T2), after the intervention (T3), at 6 months (T4) and 12 months after inclusion (T5).

Primary outcome:

Total number of sickness absence days during 12 months after enrollment in the study (i.e. after randomization) based on registry data.

Secondary sick leave/work outcomes:

1. The time until full sustainable RTW, i.e. at least 4 weeks without relapse during 12 months of follow up, based on registry data.
2. Probability of working (i.e. not receiving medical benefits) each month during 12 months of follow up, measured as repeated events, based on registry data.

Other secondary outcomes:

See published protocol article and ClinicalTrials registration

Statistical analyses:

Number of sickness absence days will likely not be normally distributed and compared by non-parametric methods. To analyze time to sustainable RTW, Kaplan Meier curves will be estimated and compared with the log rank test. To estimate HR for RTW, we will use the Cox proportional hazard model and the Efron method for ties. Time will be calculated as the number of months until RTW, and participants censored at full sustainable RTW or end of follow-up. Analyses will be performed unadjusted and adjusted for age, gender, education, main diagnosis and length and type (full/partial) of sick leave at inclusion. Probability of working each month during follow-up will be measured as repeated events and analyzed with logistic General Estimating Equations (GEE). Effect differences for other secondary outcomes will be analyzed with linear mixed models and non-parametric methods, when appropriate.

The main analyses will be performed in line with the intention-to-treat principle. In addition, we will perform per protocol analyses (participants receiving both sessions). Subgroup analyses will be performed if sufficient power for age, gender, diagnoses for sick leave, occupational category and length of previous sick leave.

Precision will be assessed using 95% confidence intervals. P-values (two-tailed) <0.05 will be considered statistically significant.

This statistical analyses plan was written in the early days of the Covid-19 pandemic. There probably will be a need for further sensitivity analyses and maybe even adjustments to the planned analyses. This will depend on the development of the pandemic.