Report Supplementary Material 7:

Study 9 additional explanation of statistical methods for sensitivity analyses

For the primary analysis, additionally we undertook simulations to examine the bias in the treatment effect estimates arising from the interim analysis decision to close the two less promising of the four trial arms. Matching the intended study design and the actual study participant attrition as closely as possible we simulated the trial recruitment, attrition, and baseline, endpoint, and follow-up measures. We then compared estimates according to the analysis method of the SAP that were obtained from analyses that (i) used the data available with all four trial-arms open to recruitment till the end of the study and (ii) used the data available from the four-arms up to the date of the interim analysis, after which data from subsequent recruitment to the two arms with the least promising treatment effect estimates was removed. Two simulations were undertaken: a null scenario where all treatment effects were zero, and a second scenario where the treatment effects corresponded to those from the SAP specified primary analysis, unadjusted for the interim recruitment closures. For the GAS primary outcome, 1000 simulations estimated the bias for the contrasts of interest. All were sufficiently small to allow us to present the results from the unadjusted analyses of the prespecified SAP.

To estimate the effect of actually receiving treatment, an instrumental variable approach was used to estimate the local average treatment effect per hour of CR on the primary outcome compared to TAU, using an assumption of a common effect per hour of active CR. The model featured site and baseline GAS as covariates in both stages of the IV regression and used randomisation assignment as the instrument.

Also, sensitivity analyses were carried out for i) clustering effects within site; the intention was to account for any clustering effects within the Group arm, but we did not record group membership within the database and so site clustering was used as a proxy (ii) The primary outcome using only data from follow-up visits (post-therapy and 6M post therapy) that occurred within the protocol specified visit windows (iii) to assess the existing missing at random assumption that is made by the primary analysis model. Two different approaches were used to examine the missingness assumption; first we examined whether key demographic variables predicted primary outcome missingness. Secondly, we used last observation carried forward for the 6-month (post-therapy) visit carrying forward observations from the post-therapy visit for missing observations, using a linear regression model for GAS T-score with the same covariates as above (iv).

Additionally, to check the robustness of results against departures from the missing at random assumption, we used the rctmiss package in Stata 15 to carry out the primary analysis under a range of assumptions about the missing data, with data and missingness modelled jointly using a pattern-mixture model. Regression models with outcomes modelled separately were used instead of the gsem command (in order for rctmiss to work), with the same covariates included. The scenarios estimated treatment effects (with confidence intervals) for different values of delta, where delta is defined as the difference in the mean of the unobserved values of the outcome from the mean of the adjusted observed values.