Supplementary Materials. Appendix 1

LCMS analysis was carried out in both negative and positive ion modes at 2 time points: 15 weeks and 20 weeks' as described by Dunn et al [19]. Deproteinized samples were prepared for UPLC-MS analysis by reconstitution in 70 L of high-performance liquid chromatography grade water followed by vortex mixing (15 seconds), centrifugation (11 337 g, 15 minutes), and transfer to vials. Samples were analyzed by an Acquity UPLC (Waters Corp) coupled to a hybrid LTQ-Orbitrap mass spectrometry system (Thermo Fisher Scientific) operating in electrospray ionization mode (Manchester University, 2012). All data related to technical or biological replicates are acquired by analysis of the Quality Control (QC) samples. These QC samples are prepared from a pooled sample and so multiple samples are prepared which provides information on both variations associated with multiple preparation of the same sample but also technical variation through analysis of the same sample.

Data were acquired over 3 analytical batches each consisting of 145 sample injections. The biologically identical QC samples were created to assess analytical precision and model systematic measurement bias. These samples were periodically injected throughout the analytical run following standard practice [19-21]. At the start of each analytical batch, a solvent blank, matrix blank and ten conditioning QC samples were analysed. A pooled QC was then analysed every 5th sample. After all the complete set of raw Thermo MS files had been acquired, they were grouped into metabolite features, and peak areas quantified for each sample, using the XCMS algorithm[22]. XCMS peak detection and alignment were completed using the following parameters: method = 'centWave', ppm = 10, peak width = 5 – 20, snthresh = 6, mzdiff = 0.01, retention time correction method = 'obiwarp', and mzdiff = 0.01. For each metabolite feature in turn, any within or between batch systematic bias in peak area w.r.t. injection order, was corrected using the Quality Control-Robust Spline Correction (QC-RSC) algorithm [23]. All metabolite features were annotated according to level 2 of the MSI reporting standards applying PUTMEDID_LCMS, as previously described [24].