Power examples:

[Weight Loss, Glycemic Control, and Cardiovascular Disease Risk Factors in Response to Differential Diet Composition in a Weight Loss Program in Type 2 Diabetes: A Randomized Controlled Trial | Diabetes Care | American Diabetes Association](https://diabetesjournals.org/care/article/37/6/1573/29953/Weight-Loss-Glycemic-Control-and-Cardiovascular?utm_source=chatgpt.com)  
Power calculations were based on data from a previous clinical trial of the weight loss program ([**8**](javascript:;)) and biochemical laboratory data in individuals with diabetes ([**4**](javascript:;)). Using mean (SD) effect sizes of 6.8 (8.8) in the intervention groups and 2.0 (7.2) in the control group, there was 90% power for the primary aim with 75 participants per arm and a dropout rate of up to 20%. There was also 90% power to discern between-group HbA1c differences of 0.5% (6 mmol/mol). **[verified by PASS]**

[Whole-body hypothermia in mild neonatal encephalopathy: protocol for a multicentre phase III randomised controlled trial - PMC](https://pmc.ncbi.nlm.nih.gov/articles/PMC11256637/)

The Bayley-IV Cognitive Scale Composite score has a normative mean of 100 and SD of 15. To detect a clinically important minimum difference of 5 points (0.3 SD), at a 0.05 significance level and 90% power, we would need 191 neonates per group, 382 in total. This increases to 426, after allowing for a conservative 10% drop-out rate (Table [2](https://pmc.ncbi.nlm.nih.gov/articles/PMC11256637/#Tab2)). The total duration of the trial is 66 months which will include a six-month trial set up period, 30 months of recruitment, and outcome assessments at the age of 24 (*±* 2) months. **[verified by PASS]**

[Thoracentesis to alleviate pleural effusion in acute heart failure: study protocol for the multicentre, open-label, randomised controlled TAP-IT trial - PMC](https://pmc.ncbi.nlm.nih.gov/articles/PMC10806591/)

For estimation of sample size, we assumed a t-test of superiority and found a total of 126 participants required to detect a difference of 3 days in the primary endpoint with an α of 0.05 and a power of 90%. This assumes that participants assigned to a strategy with referral to up-front thoracentesis in addition to pharmacological therapy will have 85 days alive and not hospitalised during the 90 days after randomisation while participants assigned to pharmacological therapy alone will have 82 days, with a shared SD of 5 days, and in-hospital mortality of 5% in both groups. **[verified by PASS]**

[Study protocol: a multicentre, open-label, parallel-group, phase 2, randomised controlled trial of autologous macrophage therapy for liver cirrhosis (MATCH) - PMC](https://pmc.ncbi.nlm.nih.gov/articles/PMC8576470/)

To detect a difference in the baseline to 90-day change in MELD score of 1 SD using a two-sided, two-sample test with a 5% level of significance, a sample size of 23 per group to detect the same level of difference with 90% power is required. All analyses will be carried out on an intention to treat basis, retaining participants in their randomised treatment groups irrespective of the treatment received. Adverse event (AE) data will be presented by treatment received. **[verified by PASS]**

[Randomized phase III trial evaluating motivational interviewing and text interventions to optimize adherence to breast cancer endocrine therapy (Alliance A191901): the GETSET protocol - PMC](https://pmc.ncbi.nlm.nih.gov/articles/PMC10568920/)

Enhanced usual care adherence at 12 months is assumed to be 69%, based on our preliminary studies of pharmacy refill data from claims in Medicare, Medicaid, and private insurance populations, using weighted estimates of 12-month adherence among Black women < 50 years old, non-Black women < 50 years old, Black women ≥ 50 years old, and non-Black women ≥ 50 years old. We hypothesize 81% adherence for the TMR-only and MI-only intervention arms based on preliminary data from the GETSET MI counseling pilot intervention [[37](https://pmc.ncbi.nlm.nih.gov/articles/PMC10568920/#CR37), [45](https://pmc.ncbi.nlm.nih.gov/articles/PMC10568920/#CR45)]. To maintain a primary family-wise two-sided type I error rate of*α* = 0.05, the two chi-square tests will be conducted with nominal *α* = 0.025. At a 2.5% two-sided significance level, the study will have at least 80% power to detect differences in adherence (81% in single component intervention arms vs 69% in enhanced usual care) if 263 evaluable patients are randomized each to the TMR-only, MI-only, and enhanced usual care arms. Because we anticipate ineligible and loss to follow-up of approximately 11%, consistent with other clinical trials of cancer survivors, we will over-recruit, rendering a recruitment target of 1180 randomized patients total (295 per arm). Power Analysis and Sample Size (PASS) software (PASS v15, NCSS, Kaysville, UT) was used to conduct the power analysis. **[verified by PASS; normal approximation; Fisher’s Exact Test]**

[DART: diagnostic-CT-enabled planning: a randomized trial in palliative radiation therapy (study protocol) - PMC](https://pmc.ncbi.nlm.nih.gov/articles/PMC9733349/)

Based on a review of recent same-day treatments at the London Regional Cancer Program (LRCP), the estimated TIC in the standard arm of this trial will be 4.8 hours, with a standard deviation of 2. We hypothesize that in the experimental arm, the estimated TIC will be 2.5 hours, with the same standard deviation. Using a two-sided, two-sample t-test, with a 1:2 randomization, alpha of 0.05 and power of 80%, adjusting for 10% dropout, 33 patients are required, 11 in the standard arm and 22 in the experimental arm. **[verified by PASS]**

[A randomized controlled trial of the effects of a prudent diet on cardiovascular risk factors, gene expression, and DNA methylation - the Diet and Genetic Intervention (DIGEST) Pilot study | BMC Nutrition | Full Text](https://bmcnutr.biomedcentral.com/articles/10.1186/s40795-016-0074-6?utm_source=chatgpt.com)

With 120 participants completing each intervention and control treatment period in a crossover design, with adjustment for multiple-testing of 3 *pre-specified* primary outcomes, we will have at least 90 % power to detect a 0.75 % difference in methylation of CDK2NB, and a 10 % difference in LDL-C. We will have 53 % power to detect a 25 % difference in gene expression.

[Effects of Low-Energy Diet or Exercise on Cardiovascular Function in Working-Age Adults With Type 2 Diabetes: A Prospective, Randomized, Open-Label, Blinded End Point Trial | Diabetes Care | American Diabetes Association](https://diabetesjournals.org/care/article/43/6/1300/35663/Effects-of-Low-Energy-Diet-or-Exercise-on?utm_source=chatgpt.com)

The trial sample size calculation was determined according to published pilot data from our group ([**5**](javascript:;)). To detect a between-group difference in PEDSR of 0.2 s−1 postintervention, at least 21 participants with T2D completing each of the three trial arms were needed to provide 80% power at α = 0.025 (to allow for two primary comparisons, i.e., MRP vs. routine care and exercise vs. routine care). Assuming a maximum dropout rate of 30%, we targeted recruitment of 30 patients per group at baseline.

[Effectiveness of a positive psychology and mindfulness-based app on mental health for parents of children with a neurodevelopmental disorder: study protocol of a pragmatic international randomized controlled trial - PMC](https://pmc.ncbi.nlm.nih.gov/articles/PMC11210113/)

The required sample size was computed with RMASS2 software for two-group longitudinal designs with attrition [[39](https://pmc.ncbi.nlm.nih.gov/articles/PMC11210113/#CR39)]. Given a required power of 80% and a two-sided significance level of 5%, 106 participating parents will need to be recruited in each arm to detect a linear trend effect over the first three time points (a group by linear time interaction from T0 to T2) resulting in at least a moderate effect size difference between the groups (Cohen’s *d* = 0.5) at 4 months after baseline. So, at least 212 parents must be included in this RCT.

This sample size estimation is based on the assumption that the repeated measurements follow a first-order autoregressive (AR1) variance–covariance structure (starting at *ρ* = 0.5 between repeated measurements) and an assumed attrition of 10% between both the baseline (T0) and post-intervention (T1) and between the post-intervention (T1) and 4 month follow-up (T2) measurements. Given the relatively low number of involved countries (five) in the trial and the self-help nature of the intervention, no relevant clustering between participants at the country level is anticipated in the sample size calculation.

[Intensive Blood Glucose Control and Vascular Outcomes in Patients with Type 2 Diabetes | New England Journal of Medicine](https://www.nejm.org/doi/full/10.1056/NEJMoa0802987)

The ADVANCE trial was originally designed to have a statistical power of 90% to detect a relative risk reduction of 16% or more for intensive control, as compared with standard control, for each of the primary outcomes, with the use of a two-tailed test with an alpha level of 5%.

[Intensive vs Standard Blood Pressure Control and Cardiovascular Disease Outcomes in Adults Aged ≥75 Years: A Randomized Clinical Trial | Cardiology | JAMA | JAMA Network](https://jamanetwork.com/journals/jama/fullarticle/2524266?utm_campaign=articlePDF&utm_medium=articlePDFlink&utm_source=articlePDF&utm_content=jama.2016.7050)

Power to detect a 25% treatment effect for the primary outcome within the subgroup of participants aged 75 years or older was estimated assuming an enrollment of 3250. With a 2-year recruitment period, maximum follow-up of 6 years, and annual loss to follow-up of 2%, power was estimated to be 81.9%, assuming an event rate of 3.25% per year in the standard treatment group (Appendix B in [Supplement 1](https://jamanetwork.com/journals/jama/fullarticle/2524266?utm_campaign=articlePDF&utm_medium=articlePDFlink&utm_source=articlePDF&utm_content=jama.2016.7050#note-JOI160061-1)).

Based on the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT) event rates adjusted downward approximately 50% for temporal changes in CVD risk factors and improved therapy, a sample size of 9250 provides approximately 90% power to detect a 20% effect on the primary composite endpoint of CVD mortality and non-fatal MI, ACS, stroke, and heart failure. The annual event rate used in this calculation was 2.2%. Recruitment of a subgroup of 4300 participants with CKD provides 80% power to detect a 20% effect on the same CVD composite endpoint. The probable dementia component of the MIND study will provide 80% power to detect a 15% reduction in the incidence of dementia, 2800 SPRINT-MIND participants will provide ample power to detect a 20% reduction in the rate of decline in cognitive function 162 between the two arms (more intensive vs. less intensive blood pressure control). In 163 addition, MRI testing to detect differences in small vessel ischemic disease and total 164 brain volume will provide 80% and 90% power, respectively, between the two strategy 165 groups in SPRINT.

[The ProBio trial: molecular biomarkers for advancing personalized treatment decision in patients with metastatic castration-resistant prostate cancer - PMC](https://pmc.ncbi.nlm.nih.gov/articles/PMC7318749/)

We have selected the threshold values for graduation of a treatment-biomarker combination or stopping for futility based on extensive simulation studies, since operating characteristics cannot be easily calculated for complex platform trials [[7](https://pmc.ncbi.nlm.nih.gov/articles/PMC7318749/#CR7)]. The calibration of those thresholds has been performed to control the type-I error and assure an adequate power for graduating treatment-biomarker combinations.

In the simulations, we assumed multiple scenarios ranging from no differences in treatments in any of the biomarker combinations to treatments prolonging the mean PFS by 5 to 10 months. In terms of sample size, the average number of participants in a treatment-biomarker signature combination ranged from 70 to 95 to achieve graduation depending on the assumed scenario. The average time in which effective signature-treatment combinations remained in the trial ranged from 21 to 30 months. Given the multiplicity of therapies, we controlled the overall type-I error to be lower than 30% (10% for the individual drugs), with varying power figures up to 80% for graduating treatment-biomarker combinations. The choice of an adequate alpha level in the independent confirmatory trial will ensure an overall type-I error below 5% (such as alpha = 15%, overall type-I error = 15%·30% = 4.5%).

A comprehensive description of the simulation study will be published in a future manuscript, detailing the statistical aspects of the trial. A summary of the simulations’ results can be found in the protocol. A web interface to the simulations is available at <http://alessiocrippa.com/shiny/probio_dsmb/>.

[Effect of active vitamin D treatment on development of type 2 diabetes: DPVD randomised controlled trial in Japanese population | The BMJ](https://www.bmj.com/content/377/bmj-2021-066222?utm_source=chatgpt.com)

In the original trial design, approximately 750 patients were needed on the basis of the following assumptions: 8.4% per year incidence of diabetes in the placebo group, participant accrual period of 2.3 years, study duration of 5.3 years, and a 7% dropout rate. The study had 80% power to detect a 36% lower rate of the primary endpoint in the active vitamin D group than in the placebo group, with a two sided type I error of 0.05.

[multi-centre, double-blind, randomized, placebo-controlled trial to evaluate the effectiveness and safety of ramelteon for the prevention of postoperative delirium in elderly cancer patients: a study protocol for JORTC-PON2/J-SUPPORT2103/NCCH2103 | Japanese Journal of Clinical Oncology | Oxford Academic](https://academic.oup.com/jjco/article/53/9/851/7203936?login=true)

The proportion of delirium in patients aged ≥75 years is assumed to be 24% in the ramelteon group and 36% in the control group. With a two-sided alpha error of 0.05 and power of 0.90, the required sample size is calculated at 610 patients. **[verified by PASS]** We therefore plan to enrol 678 patients aged ≥75 years (339 per group) to allow for a 10% dropout. The proportion of delirium in patients aged ≥65 years is assumed to be 22% in the ramelteon group and 33% in the control group. With a two-sided alpha error of 0.05 and power of 0.90, the required sample size is calculated at 690 patients. Therefore, 766 patients (383 per group) would be required to allow for 10% dropout. We thus plan to enrol 88 patients aged 65–74 years in addition to the 678 patients aged ≥75 years. The study will continue to enrol both age groups until the planned enrolment of 88 patients in the 65- to 74-year-old group and 678 patients in the ≥75-year-old group is met.

[Investigator choice of standard therapy versus sequential novel therapy arms in the treatment of relapsed follicular lymphoma (REFRACT): study protocol for a multi-centre, open-label, randomised, phase II platform trial | BMC Cancer | Full Text](https://bmccancer.biomedcentral.com/articles/10.1186/s12885-024-12112-0)

Based on feasibility assessments a total sample size for all three rounds of 284 patients (95 control + 189 novel arm patients in total) was deemed appropriate. R1 will treat 126 patients (63 patients per arm, 1:1 randomisation). R2 and R3 will treat 63 patients in the novel arm and 16 in the control arm (4:1 randomisation). In order to make the most efficient use of patients’ contributions and reduce the number of control arm patients required in R2 and 3, data from patients recruited to previous control arms will be incorporated into subsequent rounds using power priors [[24](https://bmccancer.biomedcentral.com/articles/10.1186/s12885-024-12112-0#ref-CR24)]. Bayesian operating characteristics were used to calculate the probability that the PET-CT CMR rate in the novel arm is greater than a given value, under predefined conditions. A more detailed description of the Bayesian methodology using power priors is described elsewhere (manuscript submitted), sample size determinations can be found in the predefined statistical analysis plan (Supplementary Appendix [10](https://bmccancer.biomedcentral.com/articles/10.1186/s12885-024-12112-0#MOESM10)).

[Impact of a pharmacy-led screening and intervention in people at risk of or living with chronic kidney disease in a primary care setting: a cluster randomised trial protocol - PMC](https://pmc.ncbi.nlm.nih.gov/articles/PMC10748882/)

Sample size selection and estimation was performed based on the CKD diagnosis. Using the GRT sample size calculator,[35](https://pmc.ncbi.nlm.nih.gov/articles/PMC10748882/#R35) it is estimated that the intervention will increase diagnosis from the current CKD prevalence of 10% to between 15% and 20%, an increase of 5%–10%. Sample size is estimated based on a 5% increase in diagnosis of CKD. With 80% power, alpha of 0.05 and intraclass correlation coefficient (ICC)=0.05, at least 57 clusters are required per arm. Based on our sampling model, cluster size has been increased to 61 per arm to account for attrition. With an average number of 30 participants per cluster, the total sample size to be targeted is 3660. Based on previous study[11](https://pmc.ncbi.nlm.nih.gov/articles/PMC10748882/#R11) that showed a reduction of inappropriate medications in CKD from 32.3% to 17.5% (an approximate reduction of 15%), with 90% power, alpha=0.05 and ICC of 0.05 and a cluster size of 30 per cluster, 19 clusters in each arm are required (38 clusters in all) with a sample size of 1140 for the other primary outcome, that is, change in inappropriate medication use. The sample size required for the diagnosis outcome is, therefore, sufficient to cover the sample size required for the medication changes outcome.

[Effectiveness of group arts therapies (art therapy, dance movement therapy and music therapy) compared to group counselling for diagnostically heterogeneous psychiatric community patients: study protocol for a randomised controlled trial in mental health services (the ERA study) - PMC](https://pmc.ncbi.nlm.nih.gov/articles/PMC10464011/)

We have designed the trial to detect a treatment effect of 0.5 standard deviations on the primary health outcome, i.e. the level of psychological distress, as measured by the Brief Symptom Inventory Global Severity Index. In a sample of 378 patients from a UK psychiatric outpatient population, the mean GSI was 1.65 with a standard deviation of 0.81 [[40](https://pmc.ncbi.nlm.nih.gov/articles/PMC10464011/#CR40)]. An effect of 0.5 standard deviations would therefore represent a difference of 0.4 on the GSI. We assume clustering of outcomes of patients treated in the same therapy group. In the NESS trial on group body psychotherapy [[10](https://pmc.ncbi.nlm.nih.gov/articles/PMC10464011/#CR10)], the ICC for therapy groups of 10 patients varied for different outcomes, but did not exceed 0.01 (which applied to the primary outcome). We assumed, conservatively, an ICC of 0.1. We assume a drop-out rate of 15% by the end of the study, so that if we allocate 10 patients on average to each therapy group we will end up with clusters with 8.5 patients on average. Assuming a coefficient of variation of cluster size of 0.5, we will need 200 patients before drop-out in 20 clusters in each arm to achieve 90% power at the 5% significance level [[41](https://pmc.ncbi.nlm.nih.gov/articles/PMC10464011/#CR41)]. Allowing, conservatively, for the additional loss of one full cluster in each arm, we plan to recruit a total of 2 × 210 patients.