

The quality of RCTs is assessed using the following criteria outlined in CRD's *Guidance for Undertaking Reviews in Health Care*:³⁴

- Was the method used to assign participants to the treatment groups really random? (Computer-generated random numbers and random number tables will be accepted as adequate; inadequate approaches will include the use of alternation, case record numbers, birth dates and days of the week.)
- Was the allocation of treatment concealed? (Concealment will be deemed adequate when randomisation is centralised or pharmacy controlled or when the following are used: serially numbered identical containers, on-site computer-based systems in which the randomisation sequence is unreadable until after allocation, other approaches with robust methods to prevent foreknowledge of the allocation sequence to clinicians and patients. Inadequate approaches will include the use of alternation, case record numbers, days of the week, open random number lists and serially numbered envelopes, even if opaque.)
- Was the number of participants who were randomised stated?
- Were details of baseline comparability presented in terms of treatment-free interval, disease bulk, number of previous regimens, age, histology and PS?
- Was baseline comparability achieved in terms of treatment-free interval, disease bulk, number of previous regimens, age, histology and PS?
- Were the eligibility criteria for study entry specified?
- Were any co-interventions identified that may influence the outcomes for each group?
- Were the outcome assessors blinded to the treatment allocation?
- Were the individuals who administered the intervention blinded to the treatment allocation?
- Were the participants who received the intervention blinded to the treatment allocation?
- Was the success of the blinding procedure assessed?
- Were at least 80% of the participants originally included in the randomisation process followed up in the final analysis?
- Were the reasons for withdrawals stated?
- Is there any evidence to suggest that the authors measured more outcomes than they reported?
- Was an ITT analysis included?

Items are graded in terms of ✓X (item properly addressed), X (item not properly addressed), ✓X (item partially addressed), NS (unclear or not enough information) or NA (not applicable).