| **Study, First Author, Year** | **Were harms prespecified and defined?** | **Were ascertainment techniques for harms adequately described?** | **Were ascertainment techniques for harms** **equal, valid, and reliable?** | **Was duration of followup adequate for harms assessment?** | **Harms quality rating** | **Comments** |
| --- | --- | --- | --- | --- | --- | --- |
| ACST, Halliday, 200418 Halliday, 201019 den Hartog, 201320 Halliday, 199421 Halliday, 199522 | Yes | Yes | Yes for death or major stroke, perhaps less so for minor stroke and myocardial infarction (without masking of providers making the initial assessments) | Yes | Fair | For perioperative morbidity, still no masking of initial outcome assessors; may introduce bias (some incentive to underreport harms for surgeons doing the procedure, as the design paper explains that those with unacceptably high morbidity and mortality may be asked not to enter any more patients) |
| ACAS, ACAS Study Group, 199523 Baker, 200024 Young, 199625 | Yes | Yes | Yes | Yes | Good | For perioperative morbidity, still no masking of initial outcome assessors; may introduce bias (some incentive to underreport harms for surgeons doing the procedure) |
| VACS, Towne, 199026 Hobson, 199327 Hobson 198628 | Yes | Yes | Yes | Yes | Good | For perioperative morbidity, still no masking of initial outcome assessors; may introduce bias (some incentive to underreport harms for surgeons doing the procedure) |

Good: Meets all criteria: comparable groups are assembled initially and maintained throughout the study (followup ≥80%); reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; all important outcomes are considered; and appropriate attention to confounders in analysis. In addition, for RCTs, intention to treat analysis is used.

Fair: Any or all of the following problems occur, without the fatal flaws noted in the "poor" category: generally comparable groups are assembled initially but some question remains whether some (although not major) differences occurred with followup; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention to treat analysis is done for RCTs.

Poor: Any of the following fatal flaws exists: groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking outcome assessment); and key confounders are given little or no attention. For RCTs, intention to treat analysis is lacking.

**Abbreviations:** KQ = key question; RCT = randomized, controlled trial.