Table E4. Risk of bias ratings, part 4

| Author, Year  Trial name | Harms assessed using valid and reliable measures, implemented consistently across all study participants? | Potential outcomes pre-specified by researchers? Are all pre-specified outcomes reported? | [For observational studies only] Important confounding and modifying variables taken into account in the design and/or analysis? | Risk of bias | Comments |
| --- | --- | --- | --- | --- | --- |
| Babamoto et al., 20091  NR | NA | No | NA | High | Higher rates of attrition in standard care (50%) and case management(43%) groups compared to CHW group (28%); could be the reason why adherence worsened in standard care and case management groups; differences in groups at baseline, no blinding, single-question self-report adherence measure |
| Bender et al., 20102 NA | Yes | Yes | NA | Medium | Few baseline characteristics measured so difficult to evaluate the success of randomization; Recruitment occurred through ads in newspapers: the self-selection may have resultant in disproportionately large gains |
| Berg et al., 19973  NA | Yes | Yes | NA | Medium | Method NR or inadequately reported |
| Berger et al., 20054 NA | Unclear or NR | yes |  | Medium | The danger of social desirability bias may be high due to self-report persistence measure. It is also unclear whether the outcome assessors were blinded to the random status of the patients. |
| Bogner et al., 20106 NA | Unclear or NR | Yes | NA | Low | The study uses ITT analysis and clearly describes potential outcomes, their measures, and rationale for using these measures. The main concern is that several key procedures are not clearly described or reported, such as how randomization was conducted and whether outcome assessors were properly blinded to participants' treatment assignments. On the other hand, blinding participants or providers in this study was probably not feasible because of the nature of the intervention and its clear distinction from the usual care treatment. This study has a low risk of bias because the strengths of the study design, such as the 0% attrition rate and use of the MEMS adherence measure, seem to outweigh the uncertainties. |
| Bogner et al., 20085  NA | NA | Yes | NA | Medium | No information on randomization and allocation concealment; unclear whether outcome assessors were blinded |
| Bosworth et al., 20057 V-STITCH | NA | Yes | NA | Medium | Unclear if outcome assessors blinded; baseline adherence not stratified by intervention vs. control group; self-report adherence measures |
| Bosworth et al., 20088 TCYB  Bosworth et al., 20079 TCYB Methods paper | NA | Yes | NA | Medium | This study only reports preliminary 6 month results; details of study that would help with quality assessment were not been reported (i.e., randomization, blinding, etc.) |
| Brown et al., 200810 | NA | No | N/A | High | randomization, intervention, and I/E criteira varied by site (e.g., one site randomized w/n disease severity strata); med adherence measure not pre-defined |
| Capoccia et al., 200411 NA | NA | Yes | NA | Medium | Risk of bias: medium: the clinical pharmacist not only did the intervention but was involved in screening patients for eligibility, and measure of adherence is self-reported; unclear to what extent the intervention is standardized and whether protocol was maintained; possible Hawthorne effect |
| Carter et al., 200812  NA | Unclear or NR | Yes | NA | High | This study received a high risk of bias rating because the investigators suggest their attempts to keep physicians and enrolled patients blinded did not work. Physicians were able to refer patients to the study, which introduces risk of nondifferential selection bias. It also was not clear if the investigators used allocation concealment. Still, there were several strengths, including ITT analysis, good randomization, blinding of outcome assessors, low attrition, and use of a good adherence measure. |
| Carter et al., 200913 NA | Unclear or NR | Yes | NA | Medium | Medication adherence was measured with a self-report questionnaire, which may introduce information bias. It is unclear whether allocation concealment was used or whether blinding was used at all. |
| Chernew et al., 200814 NA | NA | Yes | Partial (some variables were taken in to account) | Medium | There were differences between the intervention and comparison group. The investigators did little to control for these differences. The possibility of unmeasured differences also cannot be ruled out. In addition, the sample varied over time and this is not described in sufficient detail to permit an assessment of potential impact on findings. |
| Choudhry et al., 201015 NA | NA | Yes | Partial (some variables were taken in to account) | Medium | The investigators were unable to account for other interventions/exposures that could have affected the results. They also did not provide a rationale for how they set their medication adherence threshold of 80%, so this could lead to measurement bias. A lot of important information needed for quality assessment was not reported, such as attrition and whether ITT analysis was used. |
| Choudhry et al., 201116 MI FREEE | Yes | Yes |  | Low |  | |
| Esposito et al., 199517  NA | NA | yes |  | high | Very small sample and study arms differ in several characteristics. There were no statistical analyses of results. |
| Fortney et al., 200718  TEAM (Telemedicine Enhanced Antidepressant Management) | NA | Yes | NA | High | Medium / high - patient characteristics are similar; no information on characteristics of the clinics except that 5 clinics had on-site mental health providers (i.e. social workers); unclear how resources and intensity of interactions with healthcare personnel aside from PCPs affected results; telemedicine appears to have been used at low rate (specific rate not reported); also study only conducted in clinics that had telemedicine equipment-- possible that these clinics are not generalizable to other clinics. Increased risk of bias from self-reporting of adherence info. Finally, p-values not reported with unadjusted estimates; they are provided with adjusted estimates, but unclear what covariates were included in the model. Also, not sure that this is truly an ITT analysis b/c adherence analysis only included subsample of patients with an active antidepressant prescription, and not reporting antidepressant discontinuation as a result of PCP instruction.  col S: cut-off determined not by clinical evidence; authors cite comparability to other studies as rationale for cutoff |
| Friedman et al., 199619 NA | NA | Yes | NA | Medium | Both groups started out with a very high adherence rate; only data from those who completed study were used for analyses; article did not report the average number of calls made by the intervention group. |
| Fulmer et al., 199920 NA | NA | yes |  | Medium | SF-36 and MLHF may have been affected by social desirability bias in the intervention groups more than the control as the article implies that the daily reminders were administered by the same RA who collected follow-up data |
| Grant et al., 200321 NA | NA | Yes | NA | Medium | Use of self-report by the interventionist as adherence measure and other lack of blinding and high attrition before intervention administers make risk greater than LOW but not high b/c randomization appears to have been done well and most attrition occurred same in both arms and was before intervention |
| Gould et al., 201192 | NA | Yes | N/A | High | Baseline characteristics not reported at all; differential attrition apparent- much lower drop-out rate in usual care groups than both intervention groups; method of randomization could be subverted easily and concealment broken easily; non-ITT analysis. |
| Guthrie et al., 200122  First Myocardial Infarction (MI) Risk Reduction Program | NA | Yes | NA | Medium | Very high attrition; medication adherence measure is not a validated measure; many quality measures unclear/NR |
| Hoffman et al., 200323  NA | NA | Yes | NA | Low | Comments: Zip codes of physicians were randomized, and then alternatingly assigned to each arm; No reporting of attrition but ITT analysis conducted. |
| Hunkeler, et al., 200024 | Yes | Yes | NA | High | Authors changed randomization scheme midway through the project to include a third active intervention group; results combined both active intervention groups and compared against usual care. It is unclear whether the absence of difference between usual care and active intervention can be explained by effects in opposite directions for the two embedded interventions arms within the active comparator. |
| Hunt et al., 200825  NA | NA | Yes | NA | Medium | There was high attrition in both groups, no ITT analysis, adherence thresholds not described (e.g. what is "high adherence"?) however randomization methods were good, and the study showed no difference between groups therefore this study was given a medium risk of bias instead of a high risk of bias. |
| Janson et al., 200326 NA | NA | Yes | NA | Medium | Methods NR in detail; adherence was measured primarily through diary but also collected with medication monitors; in case of discrepancy between diary and monitor, used monitor data; unclear why didn't exclusively use monitor data and extent to which monitor and self-report were different |
| Janson et al., 200928 NA | NA | Yes | NA | Low | Only difference is in peak flow and Latino ethnicity—but essentially groups were similar; baseline characteristics of intervention and control clinicians not reported. Note that results reported in the abstract somewhat misleading in that they don't focus on comparison of intervention and control arms across follow-up period despite the fact that the goal of the intervention was to increase long-term adherence. |
| Janson et al., 201027  NA | NA | Yes | NA | High | Patients were blinded to treatment group by providers were not; no info. Given describing provider characteristics or info about their inclusion. Clinic does NOT use electronic medical records; clinicians are the unit of randomization (and their panel of patients considered in either G1 or G2), but patients are often seen by different clinicians for follow-up visits |
| Johnson et al., 200630  NR | NA | Yes | NA | Medium | Attrition is very high and doesn't appear this was an ITT analysis, study does not stratify n analyzed by intervention vs. control group; whether there are differences in baseline characteristics is also unclear, so much is unknown about quality metrics, difficult to assess if medium vs. high risk of bias |
| Johnson et al., 200629 NR | NA | Yes | NA | Medium | Difficult to tell since many elements not reported |
| Johnston et al., 200031  NA | Unclear or NR | Yes | NA | High | Multiple potential sources of bias, unclear how randomized, non-blinded, outcome measure for adherence unclear. |
| Katon et al., 199633 NA | NA | Yes | NA | Medium | Unclear how many patients from each group were analyzed for some of the health outcomes. The adherence outcomes, 50% or more reduction in depressive symptoms, and patient satisfaction were done by ITT analysis; other outcomes used 141 patients who completed 2 follow up, but the study does not report information about how many in each group were included in these analyses. |
| Katon et al., 200136 NA  Ludman et al., 200337 NA  Van Korff et al., 200338 NA | NA | Yes | NA | Medium | Allocation concealment unclear; although rate of attrition for medication adherence outcome is low overall (differential rate unspecified), differential rates of attrition between arms for health outcomes of 6.2% in the intervention arm and 12.5% in the control arm |
| Katon et al., 200439  Pathways | NA | Yes | NA | High | Intervention based on IMPACT intervention (which is referenced) but nature of contact between nurses and patients not well described. Approx 20% of participants from each group dropped out; unclear if characteristics of participants who dropped out differed by group. The intervention itself includes prescriptions for AD, but only for some patients, so the outcome of adherence is endogenous to the intervention. In this context, it is impossible to attribute the change in refills to improvement in adherence; the change could just be the result of initiation of the new drug prescribed. The measure does not take into account number of prescriptions or number of medications. |
| Katon et al., 199532  NA | NA | Yes | NA | Medium | Results for medication adherence are not presented for the entire sample; they are presented for major and minor depression, the strata within which the strata were randomized. The strata, however, were constructed based on SCL depression scores, but the analysis was presented based on IDS scores that became available after randomization. The difference between randomization groups and analysis groups is unclear. |
| Katon et al., 199934  NA  Katon et al., 200235  NA | NA | Yes | NA | Medium | 70% of participants completed all follow-up assessments; ITT analysis conducted but only the 82% who were enrolled in HMO for at least 3 of 5 6-month periods and were included in adherence & cost analyses; Adequate dosage guidelines justified, but thresholds for medication adherence not supported |
| Laramee et al., 200340  NA | NA | Yes | NA | High | Attrition is extremely high and uncertain how many participants were analyzed for med adherence outcomes; given problems with randomization, would consider high. |
| Lee et al., 200641 FAME | NA | Yes | NA | Medium | Different measurement method and frequency between intervention and control group for 14 month outcomes, no blinding |
| Lin et al., 200642  NA | Unclear or NR | Yes | NA | Medium | The adherence measure in this study, computerized pharmacy refill records, was vulnerable to bias. It only measured medication refills, not actual usage by participants. As a result, it may have overestimated or even underestimate adherence rates. Data for diabetes self-management behaviors may have been affected by information bias, since they were based on self-report. |
| Maciejewski et al., 201043 NA | NA | Yes | Partial (some variables were taken in to account) | Medium | Several important factors not considered in analysis controlling for covariates, including ethnicity/race and income. The study used several measures to reduce the risk of bias due to confounding, in particular propensity score matching. | |
| Mann et al., 201044 The Statin Choice | NA | Yes | NA | Medium | The combination of risk of bias for the outcome measure by arm and lack of any reporting of attrition or ITT analysis - CW: There is not enough information to determine the answers for many of the quality questions, so in the absence of information to say for sure, this would probably have a medium risk and not a high risk of bias. |
| Martin et al., 201145  HARP | NA | Yes | N/A | High | very high attrition without reports on n analyzed from each group; non-ITT analysis |
| Montori et al., 201146 NA | NA | Yes |  | Low |  | |
| Mundt et al., 200147  NA | NA | Yes | NA | High | There was a high attrition rate in both groups (73.8% of intervention group completed all three follow up calls, and 66.9% of control group completed all three calls); the medication compliance analysis excluded 75 out of 246 (30%) patients (33 intervention and 42 control patients), the text explains that patients were excluded because they had prescription refill records in excess of 15 days (25), no prescription records (3), or a single prescription fill (26). These post-hoc exclusions (for reasons of the adequacy of prescription fill data) could result in unaccounted-for differences between the originally randomized arms. No sensitivity analysis was reported to indicate how the excluded group compared to the subgroup retained in the analysis. |
| Murray et al., 200748 NA | Yes | Yes | NA | Low | NA |
| Nietert et al., 200949  NA | NA | No | NA | Medium | The randomization method was effective, and the sample size seemed adequate. On the other hand, 2 of the 9 study locations had no refill data for the first 5 months of the study, and gender information was missing for the study sample. Also, race, education, and income data were all based on population-level data in each patient's zip code of residence, rather than each individual's information. Assuming that this group-level data also applies to the sample size leaves room for bias. Finally, it was unclear whether the adherence measure in this study, time-to-refill, is valid and reliable. |
| Odegard et al., 200550  NA | NA | Yes | NA | High | Not randomized by clinic, patient level randomization not described, high attrition in control group (20%) (Intervention group was 10 %); Not just greater attrition in control group, but many fewer were randomized to control group. |
| Okeke et al., 200951 NA | Unclear or NR | Yes |  | Medium | It is unclear whether treatment arm was concealed from medical provider or from study staff assessing outcomes. |
| Park et al., 199652  NA | yes | yes |  | high | The pharmacists delivering the intervention were responsible for recruiting, consenting, randomizing, intervening, and collecting data on all patients. Providers were not blinded. Sample size was small and far more control patients than study patients had controlled BP. |
| Pearce et al., 200853  Cardiovascular Risk Education and Social Support (CaRESS) Trial | Unclear or NR | Yes | NA | Medium | There is a medium risk of bias for several reasons. There is potential information bias because medication adherence was measured using a self-report questionnaire instead of an objective measure like MEMS. Confounding by health insurance status is unlikely but possible, since there were significant between-group differences in this variable at baseline. Also, the power of the study to avoid type II errors was limited because of insufficient recruiting. |
| Planas et al., 200954  NR | NA | Yes | NA | High | Small sample size (40 for adherence outcomes), high attrition; number of medications at baseline not accounted for; baseline characteristics appear to differ for ethnicity and BMI |
| Powell et al., 199555  NA | NA | Yes | NA | Medium | The investigators did not take baseline disease co-morbidities into account (potential confounder), and their method of deducing their subjects' disease states based on the drug prescribed seems prone to bias, as well. For example, what if a large group of patients received their medications for off-label usage? Too little information is provided about blinding and allocation concealment, so it wasn't possible to rate the study on these traits. |
| Powers et al., 201156 NA | NA | Yes | N/A | Medium | blinding and randomization methods unclear; using self-reported measure for adherence | |
| Pyne et al., 201157 HIV Translating Initiatives for Depression Into Effective Solutions (HITIDES) | NA | Yes | NA | Medium | Low rates of attrition for the overall intervention study, but low response rates for measuring outcomes. Risk of Hawthorne effect; validity of outcome assessment unlikely to vary by study group |
| Rich et al., 199658 NA | NA | Yes | NA | Medium | A few significant/borderline differences between groups: 1) age (older in treatment group) p=0.029 2) heart rate (higher in treatment) p=0.004 3) serum cholesterol (higher in treatment) p = 0.052 Analysis did not control for differences |
| Rickles et al., 200559 NA | NA | Yes | NA | Medium | Col H: baseline characteristics similar except for intervention group had more people with past history of psychiatric meds; not adjusted for in the analysis  col p: main analysis is not intent to treat; however, noted that with ITT analysis, no sign. difference across study arms on adherence measures at 6 mos. Risk of bias: Medium -- no blinding in the study; numbers were small and ITT analysis showed no effect; also authors chose to use 1-sided statistical tests; if used 2-sided test, unclear if non-ITT results would still be statistically significant; unclear if the much higher proportion of previous psychiatric meds in the intervention arm resulted in a group that was more resistant to the intervention, which may explain the lack of effect of the intervention |
| Rodin et al., 200960  NA | NA | Unclear or NR | No (Not accounted for or not identified) | High | The investigators did not control for any potential confounding variables in their analyses. This, compounded by the differences at baseline between the intervention and control groups, resulted in the high risk of bias rating. |
| Ross et al., 200461  NR | NA | Yes | NA | Medium | Providers did not know which patients enrolled in study unless they received communication from patient using SPPARO so no protocol to keep providers blinded; difference in 12-month attrition between groups ~10%; small n |
| Rudd et al., 200462 NA | NA | Yes | NA | Medium | Randomization method unclear, baseline adherence not reported, unclear if ITT analysis |
| Rudd et al., 200963 NA | Unclear or NR | Yes |  | Low | Adherence was measured only through self-report. |
| Ruskin et al., 200464  NA | NA | Yes | NA | High | Possible detection bias from failure to validate adherence threshold & reduced power to detect statistical differences in adherence due to overall attrition. Possible risk of contamination because same providers delivered treatment in both intervention groups (although treatment goals were identical between groups). Also, authors raise concern that adjustment for medical comorbidities was insufficient. The study had 12 post-randomization exclusions from 131 randomized, an additional 46 patients dropped out of the adherence analysis, leaving 56% of the original randomized sample. The adherence analysis is not based on intention-to-treat. The 70% cutoff for the dichotomous outcome of adherence is not supported by evidence. There was a possible Hawthorne effect. |
| Schaffer et al., 200465  NA | NA | No | NA | Medium | Inclusion and exclusion criteria not described; small sample size likely limited ability to test differences across groups |
| Schectman et al., 199466 NA | NA | Yes | NA | Medium | No reports on method of randomization; very high attrition >20% in niacin >30% in BAS and non-ITT analysis done (only subjects maintained on drug for 2 months analyzed- see Table 3); follow-up time to outcomes extremely short- only 2 months |
| Schneider et al., 200867 NA | Unclear or NR | Yes |  | Low |  |
| Schnipper et al., 200668 NA | Unclear or NR | yes |  | Low |  |
| Shu et al., 200969  NA | Unclear or NR | Yes |  | High | This study was a post-hoc analysis of an RCT with different outcomes from adherence. Additional details on study quality may be reported in another article: Solomon DH, Polinski JM, Stedman M, et al. Improving care of patients at-risk for osteoporosis: a randomized controlled trial. JGIM 2007; 22(3):362-367. |
| Simon et al., 200670 NA | NA | Yes | NA | Medium | Risk of bias: Medium: assessed success of baseline randomization using few characteristics; characteristics of psychiatrists unknown; The adherence measure is weak b/c prescription refills could be missing for 1/2 of study time (3 months) and person could still be considered perfectly adherent if adherent for another 3 months  Other comments:  col H: few baseline characteristics recorded; usual care group was sign. older than intervention groups: the adherence measure is filled prescriptions for at least 90 days of continuous antidepressant treatment at a minimally adequate dose - specific doses for specific meds - doses appear to be derived clinically but not referenced as mentioned above, could be nonadherent for half of follow-up time but still considered adherent. |
| Sledge et al., 200671  NA | Unclear or NR | No |  | Medium | Adherence was not a main aim of the study and was not reported in the results. |
| Smith et al., 200872 NR | NA | Yes | NA | Medium | One site was randomized by patient instead of practice; contamination could have underestimated effect of intervention |
| Solomon et al., 199873 NA  Gourley et al., 199874 NA | NA | Unclear or NR | NA | Medium | Difficult to fully assess quality given many items unknown; attrition unclear so can't tell if ITT analysis done, lack of masking of participants and outcome assessors, etc. |
| Stacy et al., 200975 NA | NA | Yes | NA | Medium | Non-ITT analysis, not sure if randomization was adequate; certain exclusions made after randomization occurred creating a population that is already fairly adherent and motivated to take their statins |
| Stuart et al., 200376  NA | NA | No | NA | High | Methods, data, results inadequately reported. High attrition rates (50%) in at least one arm, other attrition rates NR, no results reported in text, unclear if results addressed high attrition rate. |
| Taylor et al., 200377 NA | NA | yes |  | Medium | There are many aspects of the randomization and data collection procedures that are not reported, and the compliance outcome was assessed by self-report. |
| Vivian et al., 200278 NA | NA | Yes | NA | Medium | Compliance measured monthly in intervention group; only measured at baseline and at 6 months for control group; small n |
| Waalen et al., 200979  NA | Unclear or NR | Yes |  | Medium | It is unclear whether treatment arm was concealed from study staff assessing outcomes. The authors also report an independent HMO-wide program to improve osteoporosis treatment which would have impacted only the control arm. |
| Wakefield et al., 200880  NA | Unclear or NR | Yes | NA | High | High differential attrition at 180 days in videotelephone group, baseline differences between control and intervention groups in changes to medications at discharge and understanding regimen; approximately 2.6 video calls (out of 14) were transitioned to telephone calls due to technical errors; single question, non-validated assessment of adherence. |
| Wakefield et al., 200981  NA | Unclear or NR | Yes | NA | High | High differential attrition at 180 days in videotelephone group, baseline differences between control and intervention groups in changes to medications at discharge and understanding regimen; approximately 2.6 video calls (out of 14) were transitioned to telephone calls due to technical errors; single question, non-validated assessment of adherence. |
| Wakefield et al., 201182 NA | NA | Yes | N/A | Medium | measure of medication adherence is weak and data not reported | |
| Weinberger et al., 200283 NA | NA | Yes | NA | Low | Information on allocation concealment and blinding concealment not reported; study used only self-report measures of adherence |
| Weymiller et al., 200784 Statin Choice Randomized Trial  Jones et al., 200985 Statin Choice Randomized Trial | Unclear or NR | Yes | NA | Medium | In the Weymiller and Jones articles, the investigators did a commendable job of protecting the internal validity of their study data by computerizing randomization and provider allocation, blinding participants and outcome assessor to group assignments, and ITT analysis. Unfortunately, baseline adherence rates were not calculated, and the only measure of adherence was a single self-report "Yes/No" item, which could introduce information bias. |
| Williams et al., 200486  IMPACT (Improving Mood–Promoting Access to Collaborative Treatment) | NA | Yes | NA | High | Ceiling effect on baseline adherence measure makes it impossible to assess whether lack of difference at follow-up is an artifact of measurement of adherence. |
| Williams et al., 201087  NA | NA | Yes | NA | Low | Col J: providers were the target of the intervention - they were not blinded; unclear if patients were blinded. Physicians were given access to data, but most physicians did not use the data. Like an effectiveness trial to see whether intervention would be taken up by physicians. |
| Wilson et al., 201088  Better Outcomes of Asthma Treatment (BOAT); note that there is online supplemental material for methods and timeline | Yes | Yes | NA | Medium | No ITT analysis; included participants with complete data for the entire year of analysis; Computer-based adaptive randomization algorithm used to ensure concealment and better-than-chance balance among the three groups for baseline characteristics; inclusion criteria somewhat vaguely described |
| Wolever et al., 201089  NA | NA | Yes | NA | Medium |  |
| Zeng et al., 201090  NA | NA | Unclear or NR | Partial (some variables were taken in to account) | High | Analyses used different numbers of control group patients (e.g. PDC included 710 total (71 cases, 639 controls). The intervention group was limited to patients at one clinic. Not clear why that clinic was selected. |
| Zhang et al., 201091 NA | NA | Unclear or NR | Yes | Medium | Comparison group differed from intervention groups. Propensity scores may not adequately adjust for all potential confounders. |