

Inadequate planning and reporting of adjudication committees in clinical trials: recommendation proposal.

Agnes Dechartres, Isabelle Boutron, Carine Roy, Philippe Ravaud

▶ **To cite this version:**

Agnes Dechartres, Isabelle Boutron, Carine Roy, Philippe Ravaud. Inadequate planning and reporting of adjudication committees in clinical trials: recommendation proposal.. Journal of Clinical Epidemiology, 2009, 62 (7), pp.695-702. 10.1016/j.jclinepi.2008.09.011 . inserm-00355577

HAL Id: inserm-00355577

https://inserm.hal.science/inserm-00355577

Submitted on 23 Jan 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Inadequate planning and reporting of adjudication committees in clinical trials: recommendation proposal

Agnes Dechartres, M.D.[1], Isabelle Boutron, M.D. PhD. [1], Carine Roy, MS.C.[1], Philippe Ravaud, M.D. Ph.D.[1].

1 INSERM U738, Paris, France ; Université Paris 7 Denis Diderot, UFR de Médecine, Paris, France ; AP-HP, Hôpital Bichat, Département d'Epidémiologie, Biostatistique et Recherche Clinique, Paris, France ;

Corresponding author: Agnes Dechartres, MD.,

Département Epidémiologie, Biostatistique et Recherche Clinique

Secteur Claude Bernard, Hôpital Bichat Claude Bernard,

46, rue Henri Huchard, 75877 Paris, cedex 18, France

Tel: (33) 6 88 67 83 97 / Fax: (33) 1 40 25 67 73

E-mail: adechartres@gmail.com

Conflict of interest: none

Funding: The authors did not receive any specific funding for this study.

ABSTRACT (196 words)

Objectives: Adjudication committees (ACs) are recommended in randomized controlled trials (RCTs) to standardize the assessment of events. We aimed to assess the reporting and functioning of ACs (synonyms: clinical event committee, endpoint committee) in clinical trials.

Study Design and Setting: We searched 5 high-impact-factor medical journals for reports of RCTs with clinical event endpoints published between January 1, 2004 and December 31, 2005.

Results: ACs were reported in 33.4% of the 314 reports of RCTs. ACs were reported in 29.6% of trials with low risk of misclassification (i.e., "hard" main outcome), in 47.5% of trials with medium risk of misclassification (i.e., subjective main outcome and intervention delivered in a blinded fashion) and in 31% of trials with high risk of misclassification (i.e., subjective main outcome without intervention delivered in a blinded fashion). Selected cases to be adjudicated consisted largely of events identified by site investigators (93.3%). Data provided to the AC was reported for 47.4% of ACs.

Conclusion: Reporting of ACs is not fitted to the risk of biased misclassification. Important aspects of the functioning of ACs are insufficiently reported or raise methodological issues. We propose some recommendations for planning and reporting ACs in clinical trials.

Key words: Randomized Controlled Trials, Adjudication committee, Clinical event committee, Endpoint committee, Classification bias, Recommendations.

Running title: Adjudication Committees in Randomized Controlled Trials

Word count: 2997 words

INTRODUCTION

The main purpose of a randomized controlled trial is to obtain a valid estimate of the treatment effect. The process of outcome assessment has a direct impact on the study results[1]. Determining whether a patient has reached an event may be difficult if the decision involves some subjectivity or when the endpoints require the application of a complex definition. Moreover, when the intervention is not delivered in a blinded fashion, the risk of ascertainment bias is high[2].For this reason, the Food and Drug Administration (FDA)[2] and the European Medicine Agency (EMEA)[3] recommend assessment of events by adjudication committees (ACs) in guidelines published in November 2001 and July 2005, respectively.

The EMEA defines an AC as a committee consisting of clinical experts in a specific clinical area whose aim is to harmonize and standardize endpoint assessment. Synonyms are "clinical endpoints committee," "clinical event committee," and "panel review committee." The importance of such committees has been outlined in several studies[1, 4-8] showing that the classification of events changed in about 20% to 30% of cases after assessment by an AC. These modifications could have an important impact on treatment effect estimates, as demonstrated by Naslund et al[6], who showed that a 4 member AC, after examining case report forms transmitted by local investigators, corrected misinterpretations in 28.3% of cases leading to significantly different results from the preliminary results provided by site investigators.

There is no recommendation on how ACs should process to ascertain endpoints and little is known about what is reported in RCTs concerning the functioning of ACs (e.g., number of members in the AC, how cases to adjudicate are selected, and reviewing process). We systematically reviewed RCTs published in 2004 and 2005 in 5 high-impact-factor general medical journals to assess the reporting of ACs to ascertain clinical event endpoints

and to describe the reported process of adjudication.

MATERIAL AND METHODS

Data sources and searches

We performed a computerized search in MEDLINE (via PUBMED) to identify all reports of RCTs published in 2004 and 2005 in five high impact factor general medical journals (*New England Journal of Medicine, Lancet, JAMA, Annals of Internal Medicine, BMJ*) from January 1, 2004 to December 31, 2005 with "Randomized controlled trials" as limit. Our goal was not to be exhaustive but, rather, to raise awareness of methodological issues concerning ACs. We chose these 5 journals because: 1) They publish a high number of RCTs in many medical areas; 2) They have a high impact factor which is a good predictor of high methodological quality of journal articles[9]; 3) They are also considered as having a higher quality of reporting than the others[10, 11].

Study selection

One of us (A.D.) selected potentially relevant articles after screening titles, abstract and material and methods sections. Articles were included if the study was a RCT having an event as a primary or secondary outcome. An event was defined as an outcome that may or may not occur for each subject in the RCT, anytime during the study period. Trials reported as phase 1, 2 or pilot studies or assessing the efficacy or safety of diagnostic or screening procedures were excluded, as were subgroup analyses, secondary analyses and follow-up studies from an RCT. Articles were screened for duplicate publication (i.e., the same trial described in several articles), and only the trial with the main results (i.e., the article reporting the results for the primary outcomes) was selected. When articles referred to a former publication for methodology (e.g., publication of the protocol), this publication was searched and also evaluated.

Data extraction

We used a data collection form that had been developed after an extensive bibliography on the subject and previously tested by 2 reviewers (A.D, I.B) on a random sample of 10 reports published in 2003. One reviewer (A.D.) independently completed all the data extractions. A second member of the team (I.B.) reviewed a random sample of 30 articles to assess inter rater agreement. In case of discrepancies between the abstract and the full text, the reviewers relied on the full text. The reviewers were not blinded to the journal name and authors.

Data were obtained from each article as specified below:

1. Characteristics of the selected articles, including medical area, funding sources (public, private, both public and private, or unclear), type of treatment assessed (pharmacological, nonpharmacological, or both), number of centers involved, and sample size. We also assessed whether patients, care providers and outcome assessors were reported as blinded to the treatment arm.
2. Description of the primary and secondary outcomes: Events were classified in several categories derived from previous studies[12-14]: death from all causes, death from a specific cause (e.g., cardiac death), non-fatal medical events involving a complex definition (e.g., myocardial infarction, stroke), non-fatal medical events whose diagnostic relies only on radiological tests (e.g., stent stenosis, myocardial revascularization), non-fatal medical events whose diagnosis relies only on biological tests (e.g., diabetes), therapeutic decision (e.g., angioplasty, blood transfusion). We also checked whether events were of the same medical area or needed the expertise of physicians from different fields such as myocardial infarction and stroke.
3. Reporting and functioning of ACs: for each selected article, we looked for the reporting of an AC by searching all possible synonyms (e.g., adjudication committee, endpoint committee, clinical event committee, panel review committee) in the article

as well as in the appendix and acknowledgement sections, the online extra material and the published protocol if possible. For each article reporting an AC, we noted

- Methods for selecting cases to adjudicate: we reported whether the AC assessed the endpoints for all patients included in the RCT or only patients suspected of having an event according to site investigators or whether other methods for tracking events such as national registries or development of a specific computer algorithm were used.
- Type of information provided to the AC: we looked whether the information included the complete medical file for each patient, only some elements of the file or a standardized case report form; we also noted whether results of different tests were reported as being provided to the ACs.
- Composition of the AC: we checked the size of the AC, the name of members, their field of skill and their training before adjudication began; we also noted whether the members of the AC were reported to be independent of the study or blinded to the treatment arm.
- Reviewing process of the AC: We searched whether the outcomes to adjudicate were defined and reporting on the number of members reviewing each case and the process used to reach consensus. We checked for the reporting of a comparison between the results of the AC evaluation and the site investigators' evaluation.

Data analysis

We classified all trials according to the risk of biased misclassification. This classification was based on previously published works[12-14]. Briefly, we assumed that the risk of misclassification depends on the type of events: whether or not it involves some subjectivity and on the blinding status of both participants and care providers[15-17]. Table 1 shows the

definitions of low, medium and high risk of biased misclassification. Descriptive statistics included frequencies and percentages. Inter rater agreement between the two assessors (A.D and I.B) was assessed with the use of Kappa coefficient. All analyses involved SAS version 9.1 (SAS Institute).

RESULTS

Characteristics of the selected articles

A flowchart of the selection of articles is reported in the figure. Briefly, the electronic search yielded 636 citations. From this list, 338 potentially relevant articles were selected after screening titles and abstracts, and, finally, 314 articles were selected after the full text was read. Inter rater agreement for the data extraction was overall correct: the lower kappa value was 0.6 (95% CI: 0.3, 0.9) and the higher, 1.

An AC was reported in 105 articles (33.4%). Table 2 shows the reporting of ACs according to the characteristics of the trial: ACs were described in 81.3% of cardiovascular trial reports, 28.6% of neurological trial reports, 15.4% of hematologic or oncologic trial reports and 11.1% of gynecology or obstetric trial reports. An AC was reported in 41.8% of reports of trials with private funding support and 21.1% of those with public funding.

Types of primary outcomes adjudicated

An AC was reported to assess the primary outcomes in 84 articles (80%). The adjudicated primary outcomes was a composite endpoint for 53 articles (63.1%).Table 3 shows the type of events adjudicated as primary outcomes. The main primary composite outcomes that were adjudicated were: specific cause of mortality or non fatal medical event (22.6%), all cause mortality or non fatal medical event (20.7%) and specific cause of mortality or non fatal medical even or therapeutic decision (18.9%). Adjudicated non-composite primary outcomes included mainly non fatal medical events (57.7%).

Reporting of ACs according to the risk of biased misclassification

When the risk of misclassification as defined in table 1 was low, the rate of reported AC was 29.6%; when the risk of misclassification was medium, the rate of reported AC was 47.5% and when the risk of misclassification was high, the rate of reported AC was 31.0%. In cardiology, ACs are used in 81.1% of studies having a low risk of biased misclassification, in 86.2% of studies with a medium risk of biased misclassification and in 83.9% of studies with a high risk. In the other medical areas, ACs were reported in 11.4%; 31.4% and 16.2%, respectively.

Functioning of the ACs

Among the 105 articles reporting an AC, 97 articles reported one AC, 4 reported 2 ACs, 3 reported 3 ACs and 1 reported 4 ACs. The functioning of the AC was consequently assessed for 118 ACs.

Methods for selecting cases to adjudicate and information provided to the AC (Table 4)

The method of selecting cases to adjudicate was reported for 88.1% of ACs described. Methods for selecting cases provided to the AC consisted largely of suspected events identified by site investigators (n=97, 93.3%), reported to be blinded in 61.9% of cases. Other methods included use of national registries of death (n=9, 8.6%), routine electrocardiography screening by core laboratories (5, 4.8%), adjudication of all patients (n=1, 1%) and use of a specific computer algorithm (n=1, 1%).

The information provided to the AC was reported for 56 ACs (47.4%). This information included all medical files (n = 14; 25%), a standardized case report form (n = 13; 23.2%), results of tests and procedures performed (n = 17; 30.3%) and autopsy reports (n = 8; 14.3%).

Composition of the AC (Table 4)

The composition of the AC was reported for 89% of ACs. The median number of adjudicators was 3 (interquartile range: 3, 6). Medical skill were reported for 35 ACs (33.3%); of them, 10 ACs had members from different medical skills. For 6 ACs (5.7%), members were reported to be trained. For 89.3% of ACs, members were reported to be independent to the study and for 52.4% blinded to the treatment allocated. Committee member names were reported in the acknowledgements or appendix (83.8%) or as the authors of the article (32.4%).

Reviewing process (Table 4)

For 55 ACs (46.6%), all outcomes adjudicated were defined, for 38 (32.2%) only some outcomes were defined and for 25 (21.2%) none was defined.

Information related to the reviewing process was reported for 28.8% of the ACs. For 14 ACs (41.2%), the whole AC reviewed each case. For 21 ACs (61.8%), the method to reach consensus was reported, and for 12 (35.3%), a consensus was reached by the whole AC.

Methods for assessing reliability of adjudication were reported for 12 ACs (10.4%) and included verification of data transmitted to the AC by independent monitors (n=7) and readjudication of a certain number of cases (n=4). For 7 ACs (9.6%), results of the adjudication process were provided to the data safety monitoring board.

A comparison between event rates according to the AC and to site investigators was reported for only 7 ACs. For all of them, adjudicated outcomes were outcomes suspected by site investigators. ACs did not confirm the outcomes suspected by site investigators in a median of 21% of cases (IQR: 18, 25).

DISCUSSION

To our knowledge, this is the first survey identifying all RCTs with clinical events endpoints in high impact factor journals to assess the epidemiology and reporting of ACs. Even though the use of ACs is recommended to harmonize and standardize endpoint assessment, we found reporting of ACs in only 33.4% of reports of RCTs with clinical event endpoints. Our results also highlight that some important aspects of the functioning of ACs are insufficiently reported or raise methodological issues.

We classified all the selected trials according to the risk of endpoint misclassification as defined by the type of event and by the blinding status of participants and care providers. Our results highlight that, when the risk of misclassification is low, nearly one third of articles reported an AC. When the risk of misclassification is medium, the reporting of AC increases but, surprisingly, when the risk of misclassification is high, ACs are reported for only one third of articles. On the contrary, in cardiovascular trials, ACs are frequently reported whatever the risk of biased misclassification. These results underscore that ACs could be used more frequently especially when the risk of misclassification is high and more efficiently since they are expensive and time-consuming[18]. ACs were found to be more often reported in trials with private support than in trials with public support. This result can be explained by the high prevalence of ACs in cardiovascular trials and the major private funding of these trials.

Our results also highlight that some important aspects of the functioning of ACs are insufficiently reported like details of the information provided to the AC. Space restriction required by some journals could be a reason for not reporting an AC and details concerning its functioning. Nevertheless, we believe that the description of an AC could be summarized briefly in one sentence in the article and other important details like the composition of the AC could be reported in an online appendix of the article. Information on how the AC

classifies events is important for readers to judge the risk of ascertainment bias since reporting an AC does not hamper from bias. Actually, some details on ACs that were well reported in the reports of RCTs raise methodological issues. Methods for selecting cases assessed by the AC mainly rely on the events identified and reported by site investigators. However, site investigators may fail to identify all events or can be biased if investigators are not blinded to the treatment allocated. The rates of events can vary considerably depending on whether committees merely confirm events identified by site investigators or whether they adjudicate all patients[5, 6]. Adjudicating all patients is difficult to achieve because of time and cost. However, some RCTs used systematic screening of patients by routine check of biochemical markers and electrocardiographic analyses by core laboratories[6, 19-21], which allow for identifying myocardial infarction unsuspected by site investigators. Some studies have also developed a specific computer algorithm to identify events[1, 22]. Mahaffey et al[1] showed that, using a computer algorithm, 270 cases of myocardial infarction (5.2% of all patients enrolled) were not identified by site investigators and 134 cases (2.6%) identified by site investigators were not confirmed by the AC.

Information provided to the ACs to adjudicate events was insufficiently reported. Petersen et al[23] outlined the limits of medical records and death certificates because of inconstancies and omissions in the reporting of important details. Several sources of information provided to the AC could be well-kept medical files and results of routine tests and procedures. Moreover, in some studies, ACs could ask site investigators about additional information when members judge that insufficient data to draw conclusions is provided[6, 18, 24].

The optimal number of adjudicators is not known[6]. Walter et al[18] described a trade-off between the accuracy of adjudicated results and the financial effort required for the adjudication process. Using different models, the authors showed that a size of 3 was a good compromise to allow for a majority opinion. The field of medical skill was reported for only

one-third of ACs, and for most trial reports, we could not assess the ability of adjudicators to classify an event. In more than 50% of articles, outcomes adjudicated were of different medical areas, and we supposed that members of the AC had different medical skills. Because some events are more difficult to adjudicate than others[24], training or education of adjudicators is also important to report, thus allowing for high consistency of results. In the same way, agreement between adjudicators[25-27] could be assessed to verify that additional training is not needed. Finally, although blinding of the AC is an important issue, it was reported in only half of the articles.

Little information has been published on the process of an AC. Nevertheless, some studies reported the need for consensus among adjudicators to resolve disagreements[18, 23]. To verify the integrity of their event rates, some studies used methods of quality assurance such as verification of all data collected by site investigators by independent monitors and readjudication of a certain number of cases to compare with their previous results [4, 28-30]. All these findings highlight that adjudication of events in RCTs is complex and the functioning of ACs raises some issues. We propose some recommendations for the planning and reporting of ACs in RCTs (Table 5).

Our study presents some limitations. First, our search was restricted to articles published in 5 high-impact-factor medical journals and may not be representative of the entire literature. Second, discrepancies may exist between real applications and reported methods. Some deficiencies may simply appear because of poor reporting, which does not necessarily mean that the methods were not applied[31, 32].

In conclusion, ACs could be used more efficiently since they are expensive and time consuming. Our findings highlight the need for improved planning and reporting of ACs in

RCTs having an event as the primary outcome. The practical recommendations may help researchers enhance the quality of clinical event assessment.

REFERENCES

[1] Mahaffey KW, Roe MT, Dyke CK, Newby LK, Kleiman NS, Connolly P, et al. Misreporting of myocardial infarction end points: results of adjudication by a central clinical events committee in the PARAGON-B trial. Second Platelet IIb/IIIa Antagonist for the Reduction of Acute Coronary Syndrome Events in a Global Organization Network Trial. Am Heart J. 2002 Feb;143(2):242-8.

[2] U.S. Food and Drug Administration. Guidance for Clinical Trial Sponsors On the establishment and Operation of Clinical Trial Data Monitoring Committees (2001). Available at http://www.fda.gov/CBER/gdlns/clindatmon.pdf.

[3] European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP) (2005). Available at http://www.emea.europa.eu/pdfs/human/ewp/587203en.pdf

[4] Ives DG, Fitzpatrick AL, Bild DE, Psaty BM, Kuller LH, Crowley PM, et al. Surveillance and ascertainment of cardiovascular events. The Cardiovascular Health Study. Ann Epidemiol. 1995 Jul;5(4):278-85.

[5] Ferguson JJ, Califf RM, Antman EM, Cohen M, Grines CL, Goodman S, et al. Enoxaparin vs unfractionated heparin in high-risk patients with non-ST-segment elevation acute coronary syndromes managed with an intended early invasive strategy: primary results of the SYNERGY randomized trial. Jama. 2004 Jul 7;292(1):45-54.

[6] Naslund U, Grip L, Fischer-Hansen J, Gundersen T, Lehto S, Wallentin L. The impact of an end-point committee in a large multicentre, randomized, placebo-controlled clinical trial: results with and without the end-point committee's final decision on end-points. Eur Heart J. 1999 May;20(10):771-7.

[7] Mahaffey KW, Harrington RA, Akkerhuis M, Kleiman NS, Berdan LG, Crenshaw BS, et al. Disagreements between central clinical events committee and site investigator

assessments of myocardial infarction endpoints in an international clinical trial: review of the PURSUIT study. Curr Control Trials Cardiovasc Med. 2001 Jul 17;2(4):187-94.

[8] Mahaffey KW, Harrington RA, Akkerhuis M, Kleiman NS, Berdan LG, Crenshaw BS, et al. Systematic adjudication of myocardial infarction end-points in an international clinical trial. Curr Control Trials Cardiovasc Med. 2001 Jul 17;2(4):180-6.

[9] Lee KP, Schotland M, Bacchetti P, Bero LA. Association of journal quality indicators with methodological quality of clinical research articles. Jama. 2002 Jun 5;287(21):2805-8.

[10] Mills EJ, Wu P, Gagnier J, Devereaux PJ. The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. Contemp Clin Trials. 2005 Aug;26(4):480-7.

[11] Mills E, Wu P, Gagnier J, Heels-Ansdell D, Montori VM. An analysis of general medical and specialist journals that endorse CONSORT found that reporting was not enforced consistently. J Clin Epidemiol. 2005 Jul;58(7):662-7.

[12] Boutron I, Estellat C, Guittet L, Dechartres A, Sackett DL, Hrobjartsson A, et al. Methods of blinding in reports of randomized controlled trials assessing pharmacologic treatments: a systematic review
Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. PLoS Med. 2006 Oct
Feb;3(10):e425.

[13] Boutron I, Guittet L, Estellat C, Moher D, Hrobjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. PLoS Med. 2007 Feb;4(2):e61.

[14] Boutron I, Tubach F, Giraudeau B, Ravaud P. Blinding was judged more difficult to achieve and maintain in nonpharmacologic than pharmacologic trials. J Clin Epidemiol. 2004 Jun;57(6):543-50.

[15] Devereaux PJ, Bhandari M, Montori VM, Manns BJ, Ghall WA, Guyatt GH. Double blind, you have been voted off the island! Evid Based Ment Health. 2002 May;5(2):36-7.

[16] Devereaux PJ, Manns BJ, Ghali WA, Quan H, Lacchetti C, Montori VM, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. Jama. 2001 Apr 18;285(15):2000-3.

[17] Montori VM, Bhandari M, Devereaux PJ, Manns BJ, Ghali WA, Guyatt GH. In the dark: the reporting of blinding status in randomized controlled trials. J Clin Epidemiol. 2002 Aug;55(8):787-90.

[18] Walter SD, Cook DJ, Guyatt GH, King D, Troyan S. Outcome assessment for clinical trials: how many adjudicators do we need? Canadian Lung Oncology Group. Control Clin Trials. 1997 Feb;18(1):27-42.

[19] Schnitzer TJ, Burmester GR, Mysler E, Hochberg MC, Doherty M, Ehrsam E, et al. Comparison of lumiracoxib with naproxen and ibuprofen in the Therapeutic Arthritis Research and Gastrointestinal Event Trial (TARGET), reduction in ulcer complications: randomised controlled trial. Lancet. 2004 Aug 21-27;364(9435):665-74.

[20] Mitchell LG, Andrew M, Hanna K, Abshire T, Halton J, Anderson R, et al. A prospective cohort study determining the prevalence of thrombotic events in children with acute lymphoblastic leukemia and a central venous line who are treated with L-asparaginase: results of the Prophylactic Antithrombin Replacement in Kids with Acute Lymphoblastic Leukemia Treated with Asparaginase (PARKAA) Study. Cancer. 2003 Jan 15;97(2):508-16.

[21] Dormandy JA, Charbonnel B, Eckland DJ, Erdmann E, Massi-Benedetti M, Moules IK, et al. Secondary prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (PROspective pioglitAzone Clinical Trial In macroVascular Events): a randomised controlled trial. Lancet. 2005 Oct 8;366(9493):1279-89.

[22] Cuenca-Estrella M, Moore CB, Barchiesi F, Bille J, Chryssanthou E, Denning DW, et al. Multicenter evaluation of the reproducibility of the proposed antifungal susceptibility testing method for fermentative yeasts of the Antifungal Susceptibility Testing Subcommittee of the European Committee on Antimicrobial Susceptibility Testing (AFST-EUCAST). Clin Microbiol Infect. 2003 Jun;9(6):467-74.

[23] Petersen JL, Haque G, Hellkamp AS, Flaker GC, Mark Estes NA, 3rd, Marchlinski FE, et al. Comparing classifications of death in the Mode Selection Trial: agreement and disagreement among site investigators and a clinical events committee. Contemp Clin Trials. 2006 Jun;27(3):260-8.

[24] Eisenbud R, Assmann SF, Kalish LA, van Der Horst C, Collier AC. Differences in difficulty adjudicating clinical events in patients with advanced HIV disease. J Acquir Immune Defic Syndr. 2001 Sep 1;28(1):43-6.

[25] Cleland JG, Daubert JC, Erdmann E, Freemantle N, Gras D, Kappenberger L, et al. The effect of cardiac resynchronization on morbidity and mortality in heart failure. N Engl J Med. 2005 Apr 14;352(15):1539-49.

[26] Graafsma YP, Prins MH, Lensing AW, de Haan RJ, Huisman MV, Buller HR. Bleeding classification in clinical trials: observer variability and clinical relevance. Thromb Haemost. 1997 Oct;78(4):1189-92.

[27] Curb JD, McTiernan A, Heckbert SR, Kooperberg C, Stanford J, Nevitt M, et al. Outcomes ascertainment and adjudication methods in the Women's Health Initiative. Ann Epidemiol. 2003 Oct;13(9 Suppl):S122-8.

[28] Anderson GL, Limacher M, Assaf AR, Bassford T, Beresford SA, Black H, et al. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. Jama. 2004 Apr 14;291(14):1701-12.

[29] Albers GW, Diener HC, Frison L, Grind M, Nevinson M, Partridge S, et al. Ximelagatran vs warfarin for stroke prevention in patients with nonvalvular atrial fibrillation: a randomized trial. Jama. 2005 Feb 9;293(6):690-8.

[30] Halperin JL. Ximelagatran compared with warfarin for prevention of thromboembolism in patients with nonvalvular atrial fibrillation: Rationale, objectives, and design of a pair of clinical studies and baseline patient characteristics (SPORTIF III and V). Am Heart J. 2003 Sep;146(3):431-8.

[31] Soares HP, Daniels S, Kumar A, Clarke M, Scott C, Swann S, et al. Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. Bmj. 2004 Jan 3;328(7430):22-4.

[32] Eliakim R, Sharma VK, Yassin K, Adler SN, Jacob H, Cave DR, et al. A prospective study of the diagnostic accuracy of PillCam ESO esophageal capsule endoscopy versus conventional upper endoscopy in patients with chronic gastroesophageal reflux diseases. J Clin Gastroenterol. 2005 Aug;39(7):572-8.