

The problem of placebo-controlled trials allowing rescue medication: how to assess efficacy in rhinosinusitis and allergic rhinoconjunctivitis when both groups receive active rescue treatment

Julia Vent^a and Walter Lehmacher^b

^aDepartment of Otorhinolaryngology, Head and Neck Surgery, University Hospital of Cologne and ^bInstitute of Medical Statistics, Informatics and Epidemiology (IMSIE), Medical Faculty, University of Cologne, Cologne, Germany

Correspondence to Julia Vent, MD, PhD, Department of Otorhinolaryngology, Head and Neck Surgery, University Hospital of Cologne, 50924 Cologne, Germany
Tel: +49 221 478 4750; fax: +49 221 478 4793;
e-mail: Julia.Vent@uk-koeln.de

Current Opinion in Allergy and Clinical Immunology 2009, 9:222–227

Purpose of review

The assessment of pharmacological efficacy of rhinosinusitis and rhinoconjunctivitis has been performed in various studies involving standardized measures such as symptom scores, various health-related quality of life scales, the sino-nasal outcome test, or the rhinoconjunctivitis quality of life questionnaire. However, the methods and study designs are heterogeneous, and a methodological standard is still lacking. The aim of this review was – from a clinical and a methodological point of view – to evaluate placebo-controlled trials permitting the use of rescue medication and to provide insight into the specific challenges of such clinical trials. Using studies published within the past 24 months as examples, we here demonstrate the problems in evaluating and interpreting these studies and describe possibilities for refining the methodology in order to improve future study designs.

Recent findings

The World Allergy Organization has recently published recommendations for the standardization of clinical trials on respiratory allergy in which methodological aspects regarding the use of rescue medication are discussed.

Summary

This study summarizes the recommendations for study design, outcome measures, and statistical analysis to be used in planning, performing, and analyzing clinical studies on rhinoconjunctivitis and rhinosinusitis that allow the use of rescue medication.

Keywords

methods, rescue medication, rhinoconjunctivitis, rhinosinusitis, treatment

Curr Opin Allergy Clin Immunol 9:222–227
© 2009 Wolters Kluwer Health | Lippincott Williams & Wilkins
1528-4050

Introduction

The US Food and Drug Administration (FDA) of the United States Department of Health and Human Services regularly publishes regulations for trials and guidelines for good clinical practice (<http://www.fda.gov/oc/gcp/>) in order to protect humans. The European Agency for Evaluation of Medical Products (<http://www.emea.europa.eu/> as well as <http://www.emea.europa.eu/pdfs/human/press/pos/1742401en.pdf>) and its Committee for Proprietary Medicinal Products (CPMP) also provide guidelines, requirements, and standards for clinical trials. These requirements address the necessity of rescue medication in double-blind, randomized, and placebo-controlled studies.

However, rescue medication use influences the evaluation of symptoms, and thus complicates the interpretation of results. Here, we discuss the pitfalls and strengths that need to be taken into account in assessing

placebo-controlled trials permitting active rescue medication.

Methodological aspects

The World Medical Association (WMA) has developed the Declaration of Helsinki as a statement of ethical principles for medical research involving humans. Study design, conduct, and analysis should always follow these principles of good clinical practice as published online.

The study designs discussed here involve placebo-controlled trials with active rescue medication. The position of the WMA regarding the use of placebos is laid down in paragraph 32 of chapter C of the current Declaration of Helsinki (<http://www.wma.net/e/ethicsunit/helsinki.htm>): ‘The benefits, risks, burdens and effectiveness of a new intervention must be tested against those of the best current proven intervention, except in the following circumstances:

- (1) The use of placebo, or no treatment, is acceptable in studies where no current proven intervention exists; or
- (2) Where for compelling and scientifically sound methodological reasons the use of placebo is necessary to determine the efficacy or safety of an intervention and the patients who receive placebo or no treatment will not be subject to any risk of serious or irreversible harm. Extreme care must be taken to avoid abuse of this option.'

Statistical power

A study should include a large enough number of individuals in each treatment arm, and the calculation of this number is vital in order to achieve adequate power. However, here is where the dilemma begins. Splawinski and Kuzniar [1] address the controversy in their study. 'Switching from a superiority type of trial with placebo to a noninferiority trial with an active-control – following the latest revision of Declaration of Helsinki – is in practice switching from the violation of the uncertainty principle to uncertainty of results. Given human and financial resources, it appears an academic question as to which is more unethical: to violate patients' rights or to produce results without scientific value.' The latter may also be the case in placebo-controlled trials with effective rescue medication.

Patient population

The patient population should be carefully chosen to include monosensitized patients in allergy study designs or patients with acute bacterial infections in hypotheses involving infectious rhinosinusitis [2**].

The inclusion and exclusion criteria should restrict the study population as far as possible to a patient group homogeneous in their presenting disease. Furthermore, in terms of concomitant rescue medication and other chronic diseases, the criteria should be well defined in order to identify adverse effects and hypothesis-related symptoms precisely.

Active treatment dose and rescue medication

As described by Clark and Schall [3**], the main problem of using rescue medication is that alleviated and reduced symptoms can bias the difference in outcome between the placebo and the active treatment group. They further describe trials that assess the different outcome measures of rhinoconjunctivitis regarding the power to discriminate between effective treatment and placebo. They then calculate the degree of effect of medication when applied to data from a clinical trial of immunotherapy for allergic rhinoconjunctivitis. In their study, they also present a formula for assessing the efficacy of immunotherapy for

allergic rhinoconjunctivitis and emphasize that it is crucial to combine symptom and rescue medication scores. However, there is no consensus on how the different symptomatic treatment modalities such as oral, nasal, or ocular antihistamines, oral or nasal glucocorticosteroids, nasal mast cell stabilizers, and oral or nasal decongestants should be weighted with regard to each other or to the symptoms they suppress. A scoring system as used in a recent publication by Ott *et al.* [4] is completely arbitrary and may fail to comply with the symptom score used. Other authors use different scores for rescue medications as shown in Table 1 [4,5,6*,7–11,12*,13–16].

Furthermore, it is difficult to define the optimum dose when rescue medication is used [8]. Thus, it is crucial to instruct the patients to use the rescue medication only if absolutely necessary. It is possible to provide single-blind rescue medication as well as open-label rescue medication as published by Durham *et al.* [10]. The patients in that study were provided with stepwise rescue medication in which the first step involved a single-blinded 5 mg desloratadine oral tablet followed by budesonide nasal spray and oral prednisone in case of allergic symptoms. The investigators used a daily score of rescue medication taking into account which medication was applied and combined this with a general question ('How has your hay fever been today?') to assess symptoms. In this well designed large-scale trial, the primary efficacy analysis focused on the reduction in symptom scores alone. The therapeutic effect during the pollen season was only moderate ($P=0.071$). However, there was a significant reduction in rescue medication use ($P=0.047$) in the active treatment group compared with the placebo group. There is a trade-off between symptom scores and rescue medication scores in the rhinoconjunctivitis quality of life questionnaire (RQLQ) [17**]. Here, the effect was significant ($P=0.02$), regardless of the greater amount of symptomatic rescue medication taken by the placebo patients. This fact has already been described in one of the first trials measuring quality of life in allergic rhinitis [18]. In the above-mentioned trial by Ott *et al.* [4], however, the opposite occurred. The combined score of symptom reduction and rescue medication use was regarded as the primary outcome parameter. In this case, rescue medication use was only numerically lower in the active treatment group than in the group receiving placebo drops for sublingual immunotherapy (SLIT) treatment; yet, this difference never reached statistical significance. This blurred the excellent symptom reduction to such an extent that in the follow-up year after treatment, the null hypothesis could not be rejected.

The problematic use of concomitant rescue medication for symptom relief along with the study medication is also well demonstrated in the sinusitis trial number 11566 by Bayer Healthcare Pharmaceuticals [6*]. This

Table 1 Outcome measures of clinical studies and scores used

Author	Medication tested	Symptom score	Rescue medication
Wahn <i>et al.</i> [5]	5-grass-pollen tablet	RTSS (0–3)	0, no medication; 1, oral antihistamine; 2, intranasal corticosteroid; 3, oral corticosteroid
Bayer Company study (www.ClinicalStudyResults.org) [6*] Dahl <i>et al.</i> [7]	Avelox (Bayer Healthcare Pharmaceuticals, Montville, New Jersey, USA) Grass allergen tablet (GRAZAX; ALK-Abelló A/S, Hørsholm, Denmark)	SNOT-16, AIA Rhinoconjunctivitis symptoms (scale from 0 to 3): 0, no symptoms 1, slight symptoms 2, moderate symptoms 3, severe symptoms RTSS; quality of life (RQLQ); 6 individual symptom scores	Recorded: use of oral antibacterials; oral steroids ^a ; nasal preparations; analgesics Desloratadine 5 mg once daily >score 6; budesonide nasal spray (as much as 32 mg; two puffs per nostril twice daily) >score 1 per spray; prednisone (as much as 50 mg once daily) >score 1.6 per 5 mg
Didier <i>et al.</i> [8,9]	Oralair	Rhinoconjunctivitis symptom scores; quality of life; number of well days	Proportion of days with rescue medication: Oral antihistamine (cetirizine) Intranasal corticosteroid (mometasone furoate) Step one: single-blind (loratadine or placebo); step two: open label: budesonide nasal spray and oral prednisone for rhinoconjunctivitis symptoms and salbutamol inhaler, fluticasone inhaler, and oral prednisone for asthma symptoms Topical decongestant
Durham <i>et al.</i> [10] and Lima <i>et al.</i> [11]	Grass allergen tablet (GRAZAX)		
Ciebiada <i>et al.</i> [12*]	Montelukast alone or in combination with levocetirizine	HRQL RQLQ adapted for Polish population from 0 to six points; night-time symptom score rated from 0 to 3	
Meltzer <i>et al.</i> [13]	Mometasone furoate nasal spray, amoxicillin, and placebo	MSS a.m./p.m. (MSS >sum of the scores for rhinorrhea, postnasal drip, nasal congestion/stuffiness, sinus headache, and facial pain/pressure/tenderness on palpation over the paranasal sinuses); TSS (TSS >sum of the MSS and cough score assessed on the same scale); 0 (none), 1 (mild), 2 (moderate), or 3 (severe)	None
Williamson <i>et al.</i> [14]	500 mg of amoxicillin 3 times per day for 7 days and 200 µg of budesonide in each nostril once per day for 10 days	Symptom diaries (duration and severity); TSS score; 11 symptom variables assessed on 7-point Likert scales	None
Ott <i>et al.</i> [4]	Ecrt study	Daily diary cards (rhinitis, conjunctivitis, and asthma) VAS; symptom score graded 0–3 (no, mild, moderate, severe)	Rescue medication score (one to six points depending on medication): 1 tablet of corticosteroid, 6 points 1 tablet of antihistamine, 4 points 1 puff of inhalant β agonist, 4 points 2 corticosteroid eye drops, 2 points 2 corticosteroid nose spray puffs, 2 points 2 antihistamine eye drops, 1 point 2 antihistamine nose spray puffs, 1 point Recorded on telephone contact; report of change in concomitant rescue medication None
Ibanez <i>et al.</i> [15]	Grazax pediatric	Daily telephone contacts for adverse events; daily individual diary	
Bousquet <i>et al.</i> [16]	VAS – ARIA guidelines	ROLO; VAS on severity and duration of symptoms	

AIA, aspirin intolerant asthma; ARIA, allergic rhinitis and its impact on asthma; HRQL, health-related quality of life; MSS, major symptom score; RQLQ, rhinoconjunctivitis quality of life questionnaire; RTSS, rhinoconjunctivitis total symptom score; SNOT-16, 16-item sino-nasal outcome test; TSS, total symptom severity; VAS, visual analog scales.

^a Bayer Company Study 404_0.

Table 2 Summary of the World Allergy Organization recommendations for clinical research

Item	Description
Study design	Randomized, double blind, placebo controlled
Randomization	Computer-generated list. Block randomization (the method must be clearly stated)
Power	Power not <90%, α 1 type error not greater than 0.05
Patients' selection	Ideally monosensitized. If polysensitized, the causal role of the allergen used for IT must be clearly ascertained
Disease	Rhinitis (rhinoconjunctivitis) and/or asthma Diseases must be defined and classified according to GINA and ARIA criteria
Exclusion criteria	Those recommended in the ARIA and WHO guidelines
Extract	Standardized vaccines must be used when available. The cumulative and maintenance dose (possibly in micrograms of allergen) must be recorded
Concomitant medications	Rescue medications (as needed) must be prescribed to all patients and recorded on the diary card
Clinical assessment	Total score of symptoms (nasal + ocular + bronchial) must be recorded daily on the diary card Each single symptom graded from 0, absent to 3, severe
Primary outcome	Total clinical score calculated as the sum of symptom score as well as drugs usage also calculated as a score
Secondary outcomes	Scores of individual symptoms separately and changed in drug usage Quality of life assessment Safety (total number of adverse events and SIT-related side effects) Functional measures (e.g. lung function and rhinometry) Cost-effectiveness (e.g. QALY)
Surrogate markers	Skin tests Nasal, ocular, or bronchial challenge Antibody measurement Inflammatory parameters
Dropouts/withdraws	Dropouts and the reasons for must be reported

ARIA, allergic rhinitis and its impact on asthma; GINA, Global Initiative for Asthma; IT, immunotherapy; QALY, quality-adjusted life years. Data from Canonica *et al.* [2^{••}].

trial used standardized questionnaires to assess the patients' symptoms; however, it failed to provide a standardized rescue medication protocol. In the end, there was a much higher consumption of rescue medications in the placebo group compared with the active treatment group, skewing the results to such a degree that the null hypothesis ('moxifloxacin has the same effect as placebo in acute bacterial sinusitis') had to be accepted. In fact, a major proportion of the patients in the 'placebo' group had received a treatment modality that is recommended by several current guidelines as first-line symptomatic treatment for sinusitis [19,20^{••}]. Would a proper protocol with a stepwise administration of rescue medication based on symptom severity have changed the outcome of the study towards rejection of the null hypothesis?

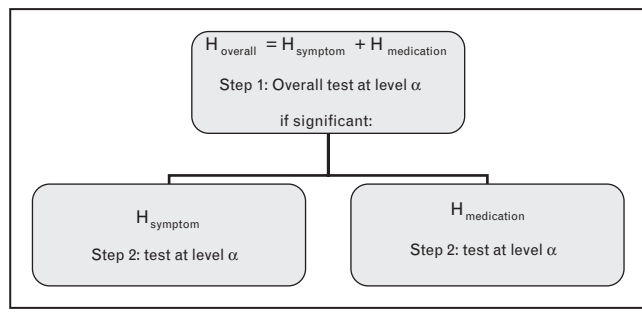
In another large-scale rhinosinusitis trial [13], a similar phenomenon was observed. Ten days of treatment with 500 mg amoxicillin capsules and placebo nasal spray three times a day did not alter the patients' situation more than treatment with placebo capsules and placebo nasal spray. However, one must realize that the 'placebo' nasal spray, applied twice daily, contained the preservative benzalkonium chloride in a solution of alcohol that, in addition to the rinsing effect, may have had an antimicrobial effect on the mild rhinosinusitis these patients were suffering from.

Canonica *et al.* [2^{••}] describe in their study on World Allergy Organization (WAO) guidelines for clinical studies published in 2007 that the immunotherapy of allergic rhinoconjunctivitis has a potential effect on both the symptom scores and the use of rescue medication. They

suggest that a combination of both outcome measures rather than the two variables separately should be taken into consideration in order to provide more clinically relevant information. Furthermore, patients would profit from thorough instructions on the minimal use of rescue medication. The positive effect of rescue medication on symptom relief as well as the need for standardized recording of both symptoms and quality of life should be taken into account when designing, executing, and interpreting placebo-controlled studies (Table 2).

Statistical aspects

There can be no doubt that a univariate consideration does not reflect the effects observed in placebo-controlled trials with rescue medication. Clark and Schall [3^{••}] propose weighing rescue medication and symptom scores equally, giving both a maximum daily value of three points. They could demonstrate that this was reasonable in one double-blind, placebo-controlled study they had analyzed in which oral antihistamines and nasal and oral glucocorticosteroids were administered. Would this also work under other circumstances? Rohmel *et al.* [21] presented a more general approach based on the combination of two different endpoints as proposed by O'Brien [22] and Läuter [23]: for each of both endpoints x and y , the data of the two pooled samples are Z transformed, that is, the symptom score (x) data are transformed to $x' = (x - x_{\text{mean}})/s_x$, where x_{mean} is the overall mean of the pooled x data and s is the standard deviation of the pooled x data; the same transformation is done for the rescue medication (y) data with $y' = (y - y_{\text{mean}})/s_y$. Then, for each patient n , the sum of two Z scores $x'n + y'n$

Figure 1 Multiple test procedure with two steps

is calculated, and these two samples of sums are compared by a standard *t*-test. As a nonparametric alternative for each endpoint, the data are transformed separately into ranks R_x and R_y . Then, for each patient, the sum of two ranks is calculated, and these two samples of individual rank sums are compared by a Wilcoxon rank test. These sums of Z scores and ranks are fairly independent of the original scales of the symptom and rescue medication scores, and this procedure can be applied if their relative magnitude cannot be assessed.

According to Clark and Schall [3**] or O'Brien [22] and Läuter [23], all these bivariate statistics combine the symptom and rescue medication scores to a global efficacy measure. In case of the superiority of one of the two treatment groups, one cannot determine whether one single endpoint is superior or whether both are superior. Therefore, a two-step multiple testing procedure [24] can be useful. In the first step, the bivariate global test is performed at significance level α ; if this overall test is significant, then both single endpoints can be simultaneously tested at level α . This procedure controls the multiple significance level α and can be applied after the O'Brien [22] and Läuter [23] tests or similarly to the Clark and Schall [3**] tests (Fig. 1). Röhmle *et al.* [21] suggested further that the superiority of a single endpoint should only be accepted if the other endpoint is at least noninferior; for example, the high superiority of the symptom score is not acceptable in the case of an inferiority of the rescue medication score.

With this calculation, a more flexible strategy is presented that will also work when few placebo patients take higher amounts of rescue medication, which may otherwise falsify results. As long as there is no consensus on which rescue medication to administer and how to accurately combine this with the symptom scores, this adaptive procedure may represent the best solution to deal with bivariate or multiple endpoints. Otherwise, the risk will remain that one or the other outcome parameter will distort the results.

Future directions

Long-term studies will require the meticulous recording of rescue medication use in order to assure homogeneity. Electronic systems such as the Medication Event Monitoring System (MEMS) may aid in recording the consumption of rescue medication objectively. Symptoms and the use of rescue medications should be assessed in a consistent manner, as both are primary outcome measures. It appears vital to find a consensus on a standard procedure for assessing symptoms and evaluating the use of rescue medication in order to draw correct conclusions from a study.

Conclusion

It has been shown that the use of rescue medication plays an important role in the interpretation of results. Multiple questionnaires and symptom scores as well as the amount of rescue medication used need to be taken into account individually, but most of all in the combined clinical picture to unveil the results of major impact [2**,3**].

At present, the recommendations for assessing clinical trials on immunotherapy are summarized and regularly updated by the WAO taskforce. For rhinosinusitis, in a likewise manner, special awareness of the use of rescue medication needs to be raised by the investigator as well as the patients who require special instructions.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

Additional references related to this topic can also be found in the Current World Literature section in this issue (p. 286).

- 1 Splawinski J, Kuzniar J. Clinical trials: active control vs placebo – what is ethical? *Sci Eng Ethics* 2004; 10:73–79.
- 2 Canonica GW, Baena-Cagnani CE, Bousquet J, *et al.* Recommendations for standardization of clinical trials with allergen specific immunotherapy for respiratory allergy. A statement of a World Allergy Organization (WAO) taskforce. *Allergy* 2007; 62:317–324.
- 3 Clark J, Schall R. Assessment of combined symptom and medication scores for rhinoconjunctivitis immunotherapy clinical trials. *Allergy* 2007; 62:1023–1028.
- 4 Ott H, Sieber J, Brehler R, *et al.* Efficacy of grass pollen sublingual immunotherapy for three consecutive seasons and after cessation of treatment: the ECRIT study. *Allergy* 2009; 64:179–186.
- 5 Wahn U, Tabar A, Kuna P, *et al.* Efficacy and safety of 5-grass-pollen sublingual immunotherapy tablets in pediatric allergic rhinoconjunctivitis. *J Allergy Clin Immunol* 2009; 123:160–166.e3.
- 6 BSP AG Germany/Bayer Healthcare Pharmaceuticals: prospective, multi-center, randomized, double-blind, placebo controlled trial to evaluate the efficacy and safety of moxifloxacin 400 mg QD for 5 days versus placebo in the treatment of acute bacterial sinusitis.

A perfectly designed study serving as an example on how important unified measures for rescue medication are in order to come to a valid conclusion.

- 7 Dahl R, Kapp A, Colombo G, *et al.* Sublingual grass allergen tablet immunotherapy provides sustained clinical benefit with progressive immunologic changes over 2 years. *J Allergy Clin Immunol* 2008; 121:512–518.
- 8 Didier A, Malling HJ, Worm M, *et al.* Optimal dose, efficacy, and safety of once-daily sublingual immunotherapy with a 5-grass pollen tablet for seasonal allergic rhinitis. *J Allergy Clin Immunol* 2007; 120:1338–1345.
- 9 Didier A, Melac M, Montagut A, *et al.* Agreement of efficacy assessments for five-grass pollen sublingual tablet immunotherapy. *Allergy* 2009; 64:166–171.
- 10 Durham SR, Yang WH, Pedersen MR, *et al.* Sublingual immunotherapy with once-daily grass allergen tablets: a randomized controlled trial in seasonal allergic rhinoconjunctivitis. *J Allergy Clin Immunol* 2006; 117:802–809.
- 11 Lima MT, Wilson D, Pitkin L, *et al.* Grass pollen sublingual immunotherapy for seasonal rhinoconjunctivitis: a randomized controlled trial. *Clin Exp Allergy* 2002; 32:507–514.
- 12 Ciebiada M, Gorska Ciebiada M, Kmiecik T, *et al.* Quality of life in patients with persistent allergic rhinitis treated with montelukast alone or in combination with levocetirizine or desloratadine. *J Investig Allergol Clin Immunol* 2008; 18:343–349.
- This recent study focuses on evaluating quality of life and can serve as a good example for clinical assessment.
- 13 Meltzer EO, Bachert C, Staudinger H. Treating acute rhinosinusitis: comparing efficacy and safety of mometasone furoate nasal spray, amoxicillin, and placebo. *J Allergy Clin Immunol* 2005; 116:1289–1295.
- 14 Williamson IG, Rumsby K, Bengt S, *et al.* Antibiotics and topical nasal steroid for treatment of acute maxillary sinusitis: a randomized controlled trial. *JAMA* 2007; 298:2487–2496.
- 15 Ibanez MD, Kaiser F, Knecht R, *et al.* Safety of specific sublingual immunotherapy with SQ standardized grass allergen tablets in children. *Pediatr Allergy Immunol* 2007; 18:516–522.
- 16 Bousquet PJ, Combescure C, Neukirch F, *et al.* Visual analog scales can assess the severity of rhinitis graded according to ARIA guidelines. *Allergy* 2006; 62:367–372.
- 17 Petersen KD, Kronborg C, Gyrd-Hansen D, *et al.* Quality of life in rhinoconjunctivitis assessed with generic and disease-specific questionnaires. *Allergy* 2008; 63:284–291.
- A thorough evaluation of various questionnaires (RQLQ and 15-dimensional instrument on quality of life).
- 18 Bousquet J, Duchateau J, Pignat JC, *et al.* Improvement of quality of life by treatment with cetirizine in patients with perennial allergic rhinitis as determined by a French version of the SF-36 questionnaire. *J Allergy Clin Immunol* 1996; 98:309–316.
- 19 Stuck BA, Bachert C, Federspil P, *et al.* Rhinosinusitis guidelines of the German society for otolaryngology, head and neck surgery. *HNO* 2007; 55:758–777.
- 20 Thomas M, Yawn BP, Price D, *et al.* EPOS primary care guidelines: European Position Paper on the Primary Care, Diagnosis and Management of Rhinosinusitis and Nasal Polyps 2007 – a summary. *Prim Care Respir J* 2008; 17:79–89.
- This publication is a standard reference for rhinology and any rhinitis-related issues.
- 21 Rohmel J, Gerlinger C, Benda N, Lauter J. On testing simultaneously non-inferiority in two multiple primary endpoints and superiority in at least one of them. *Biom J* 2006; 48:916–933.
- 22 O'Brien PC. Procedures for comparing samples with multiple endpoints. *Biometrics* 1984; 40:1079–1087.
- 23 Läuter J. Exact T and F tests for analysing studies with multiple endpoints. *Biometrics* 1996; 52:964–970.
- 24 Lehmacher W, Wassmer G, Reitmeir P. Procedures for two-sample comparisons with multiple endpoints controlling the experiment wise error rate. *Biometrics* 1991; 47:511–521.