Acta Oncologica, 2012; 51: 890–896

informa healthcare

ORIGINAL ARTICLE

Non-inferiority trials in breast and non-small cell lung cancer: Choice of non-inferiority margins and other statistical aspects

EVERARDO D. SAAD[1] & MARC BUYSE[2]

[1]*Dendrix Research, Sao Paulo, Brazil, and* [2]*International Drug Development Institute, Louvain-la-Neuve, Belgium*

Abstract

Background. Determining the non-inferiority margin is an essential step in the design and interpretation of non-inferiority trials, and this margin should be preferably justified on clinical and statistical grounds. *Methods.* After a PubMed search for phase III trials in advanced breast cancer (BC) or non-small cell lung cancer (NSCLC) published between January 1998 and December 2009 in 11 leading journals, non-inferiority trials were selected by manual search of the full papers. *Results.* Twenty-four of 195 trials had a primary non-inferiority hypothesis. When the two six-year study periods were compared, there were time trends within BC and NSCLC, with most non-inferiority trials in BC reported in the first six-year period, and vice-versa for NSCLC. The median sample size was larger for non-inferiority than superiority trials ($p = 0.01$). The choice of a non-inferiority margin was reportedly justified in only five cases. Non-inferiority trials were more likely than superiority trials to yield positive results ($p < 0.001$), as were trials in breast cancer ($p = 0.02$). *Conclusions.* Non-inferiority margins for cancer trials appear to be chosen mostly on historical grounds. Since nearly three-quarters of non-inferiority trials achieve their primary objective, the extent to which the choice of margins has influence on trial results remains to be determined.

Randomized studies, especially phase III trials, are the gold standard approach for evaluating therapeutic interventions in medicine. The goal of phase III trials is to compare two or more treatments in order to find the one with greater utility in clinical practice. Utility may be reflected by higher efficacy or quality of life, lower toxicity, or greater convenience to the patient or health-care system. A trial with the primary objective of showing that the experimental arm is superior to a control arm is commonly termed a superiority trial. Although most phase III trials aim to demonstrate that an experimental treatment is superior to a control intervention, non-inferiority phase III trials assess whether a more convenient, less toxic or cheaper intervention is at least as efficacious as an existing standard of care, given clinically accepted limits [1,2]. Such limits, termed non-inferiority margins, bear important implications on design, sample-size estimation, analysis, and clinical application of results of non-inferiority trials. In fact, it has been pointed out that 'determining the non-inferiority margin is the single greatest challenge in the design, conduct, and interpretation of non-inferiority trials' [3].

Efficacy determination in non-inferiority trials may involve both a statistical comparison of the study drug with a concurrent active control and historical evidence that the active control is superior to placebo. Moreover, in non-inferiority trials conducted in some medical fields, a placebo group as an internal control is desirable in order to assure what is termed assay sensitivity [4,5]. In oncology, direct comparisons with placebo or best supportive care (BSC) are not always feasible from an ethical standpoint; therefore, even historical data are often unavailable for indirect comparisons with placebo or BSC in the design of non-inferiority trials of novel cancer therapies [3,6,7]. As a result, adequate choice of the non-inferiority margin is critical to obtaining a meaningful answer in a non-inferiority cancer trial. According to the Consolidated Standards of Reporting Trials Group, 'The margin of non-inferiority or equivalence should be specified, and preferably justified on clinical grounds' and 'its

Correspondence: E. D. Saad, Dendrix Research, Rua Joaquim Floriano, 72/24, 04534-000, Sao Paulo, Brazil. Tel: + 55 11 31687088. Fax: + 55 11 31671148. E-mail: everardo@dendrix.com.br

(Received 29 March 2012; accepted 30 May 2012)

ISSN 0284-186X print/ISSN 1651-226X online
DOI: 10.3109/0284186X.2012.702924

relation to the effect of the reference treatment relative to placebo in any previous trials should be noted' [2]. Regulatory agencies in Europe and the US have issued guidance documents on the choice of the non-inferiority margins, with details on key considerations driving the choice [7], and mathematical approaches to calculate a margin based on prior trials when possible [3].

The objective of the current study was to assess the design features, especially with regard to the justification for non-inferiority margins, of recent non-inferiority trials in the two tumor types with the largest number of phase III trials: breast cancer (BC) and non-small cell lung cancer (NSCLC).

Methods

Selection of trials

We have described our trial selection strategy in a previous study on advanced BC [8], and we applied the same strategy to NSCLC. In brief, we searched PubMed for articles reporting the efficacy results of phase III trials on systemic antineoplastic treatments for advanced BC or advanced NSCLC published between January 1998 and December 2009 in 11 leading medical journals (*Annals of Oncology*, *Breast Cancer Research and Treatment*, *British Journal of Cancer*, *Cancer*, *Clinical Cancer Research*, *European Journal of Cancer*, *Journal of Clinical Oncology*, *Journal of the National Cancer Institute*, *The Lancet*, *Lancet Oncology*, and *The New England Journal of Medicine*). We made no attempt to assess the quality of reporting, as done previously by other investigators, nor have we tried to control for publication bias. The aim of our study was to analyze the results of trials that are likely to influence clinical and statistical practice because of their publication in widely read journals. Of note, we excluded randomized phase II trials, studies on high-dose chemotherapy, papers reporting combined analyses of two or more separate trials, and studies only on correlative biology or prognostic factors. Non-inferiority or equivalence trials were selected by manual search of the full paper of all eligible trials and analyzed in detail alongside superiority trials.

Abstraction of data and definitions used

For each trial, we abstracted data on its type (non-inferiority vs. superiority), number of patients and arms, journal and year of publication, treatment line, primary endpoint, and trial positivity, as ascertained by the original authors. Of note, some studies were termed equivalence trials, although most assessed non-inferiority (i.e. they had a one-sided hypothesis); we therefore analyzed non-inferiority and equivalence trials together. We analyzed the primary efficacy endpoint for each trial following the authors' definitions in the papers. When not explicitly stated by the authors of the reports, we considered the primary endpoint as the one used for sample-size calculation or the first endpoint cited in the 'Methods' or 'Results' section of each paper. We investigated the statistical aspects relating to sample-size calculation, including those related to the definition of non-inferiority margins. We considered that such margin was justified when authors gave an explicit reason for their choice, regardless of whether specific methods or references were cited. Otherwise, we considered that the margin had not been justified by authors. In a previous study on breast cancer [9] we have shown that progression-free survival (PFS), time to tumor progression (TTP) and time to treatment failure (TTF) are often used interchangeably by investigators, notwithstanding their different definitions. For the sake of simplicity, therefore, we analyzed these three endpoints collectively, and refer to them collectively as PFS.

Statistical analysis

We used the MedCalc software (Mariakerke, Belgium) for statistical analysis and computation of 95% confidence intervals (CIs). We performed exploratory analyses by comparing superiority and non-inferiority trials with regard to the number of patients using the Mann-Whitney test, and used the χ^2-test or Fisher's exact test as appropriate to compare proportions.

Results

General features of trials

Our search retrieved 93 trials on BC and 102 on NSCLC (a list of such articles is available upon request). Twenty-four of these 195 trials (12.3%; 95% CI 8.0–17.8%) had a primary non-inferiority hypothesis (12 in BC and 12 in NSCLC). There was no trend in the proportion of non-inferiority trials when the two six-year periods were compared (11.5% vs. 13.0%; $p = 0.83$). However, there were opposing time trends within BC and NSCLC, with eight of 12 non-inferiority trials in BC reported in the first six-year period, and 10 of 12 NSCLC trials reported in the second.

The median number of patients per arm for non-inferiority and superiority trials were 250 (range, 80–835) and 160 (range, 34–846) patients, respectively ($p = 0.01$). The primary endpoint for non-inferiority trials was overall survival (OS), PFS and response rate (RR) in 10, five and nine cases, respectively, whereas the corresponding figures for superiority trials were 75, 51 and 34 (in addition, quality of life was the primary endpoint in nine cases, and time to response or clinical benefit rate in one case each among superiority trials). Table I summarizes these and other selected features of the trials analyzed.

Non-inferiority margins

Table II displays selected features of the 24 non-inferiority trials, all of which reported the non-inferiority margin [references are available online at http://informahealthcare.com/10.3109/0284186X.2012.702924]. Among trials with time-to-event endpoints (n = 11), all used the lower limit of the 95% CI for the hazard ratio for the primary endpoint as the non-inferiority margin. For those 11 trials, the non-inferiority margin ranged from 0.75 to 0.86, with mean and median of 0.8. For 12 of the trials with a categorical endpoint (RR in nine cases and one-year OS in three), the non-inferiority margin corresponded to a lower limit of the 95% CI for the difference in proportion between groups, and this margin ranged from 10% to 25%, with mean of 16% and median of 15%. In one trial with one-year OS as the primary endpoint, an equivalence limit of 2% was reportedly used.

Table I. Summary features of non-inferiority and superiority trials.

	N or value		
Characteristic	Non-inferiority trials (n = 24)	Superiority trials (n = 171)	p-value
Journal of publication			0.32
Journal of Clinical Oncology	13	103	
Annals of Oncology	5	20	
British Journal of Cancer	0	13	
Other journals*	6	35	
Patients per arm, median (range)	250 (80–835)	160 (34–846)	0.01
Treatment line			0.48
First only	15	122	
First and others, including maintenance therapy	3	23	
Others except first, excluding maintenance therapy	6	26	
Primary end point			0.16
Overall survival	10	75	
Progression-free survival	5	51	
Response rate	9	34	
Others	0	11	

*European Journal of Cancer (n = 9), Breast Cancer Research and Treatment (n = 7), Cancer (n = 7), The New England Journal of Medicine (n = 7), The Lancet (n = 5), Journal of the National Cancer Institute (n = 3), Lancet Oncology (n = 3).

The choice of a non-inferiority margin was justified in the publication for only five of the 24 trials. In three of these cases, the margin was determined on the basis of previous direct comparisons between the control group used in the current trial and either placebo or active treatment, in both cases using the hazard ratio for OS. In the other two cases, the margin was determined on the basis of theoretical considerations regarding a maximum decrease in RR that would not be associated with decreased survival or to the risk of death that was deemed worth the toxicity associated with one of the treatments. For the other 19 trials, the non-inferiority margin was presented with no explicit clinical or statistical justification for its choice.

Other statistical features of non-inferiority trials

The type I error rate, mentioned in 22 of 24 reports, ranged from 2.5% to 10%, with median and mode (n = 14) of 5%. In 14 of 17 reports for which this information was available, one-sided type I error rates were used. On the other hand, a one-sided type I error rate of 2.5% or a two-sided type I error rate of 5% were reportedly used in only seven of 17 trials with complete information. Power was reported in 21 papers, and ranged from 59% to 90%, with median and mode (n = 15) of 80%.

In terms of patient accrual, 15 of 18 trials reporting the target number of patients to be accrued (and not exclusively the number of events to be observed) actually included at least that number of patients. For the other three trials, final accrual ranged from 58% to 78% of target accrual. All 24 trials described the population used for the primary analysis of non-inferiority. In 20 cases, the intent-to-treat (ITT) population was explicitly mentioned or the primary analysis was reportedly conducted with all randomized patients. In one case, the primary analysis was conducted using both ITT and per-protocol (PP) populations, and, the PP or response-evaluable population was used in the remaining three trials.

Positivity of non-inferiority and superiority trials

We could ascertain whether the primary objective of demonstrating non-inferiority of the experimental regimen could be achieved in 22 of the 24 trials (Table II). For one trial with RR as the primary endpoint, the 95% CI for the difference in RRs was not reported, nor were claims of non-inferiority made. In a second trial with the aim of demonstrating equivalence in one-year OS between groups, results were only presented for a superiority test, which yielded a non-significant result. Seventeen of

Table II. Selected features of the 24 non-inferiority trials analyzed.

First author	Tumor type	Interventions	Primary end point	Margin justification	Primary results
Smith	BC	C: 24-hour paclitaxel E: 3-hour paclitaxel	RR	No	C superior to E
Kaufmann	BC	C: megestrol acetate E: exemestane	RR	No	E non-inferior to C
Bonneterre	BC	C: tamoxifen E: anastrozole	PFS*	No	E non-inferior to C
Nabholtz	BC	C: tamoxifen E: anastrozole	PFS*	No	E non-inferior to C in RR and superior in PFS
Batist	BC	C: Conventional doxorubicin and cyclophosphamide E: Liposome-encapsulated doxorubicin and cyclophosphamide	RR	No	E non-inferior to C
Namer	BC	C: 5-fluorouracil, cyclophosphamide, and doxorubicin or epirubicin E: mitoxantrone and vinorelbine	RR	No	E non-inferior to C
Harris	BC	C: Conventional doxorubicin E: Liposome-encapsulated doxorubicin	RR	No	Identical RR in both groups, but 95% CI for difference not reported
Buzdar	BC	C: tamoxifen E: droloxifene	PFS	No	C superior to E
Rosell	NSCLC	C: paclitaxel and cisplatin E: paclitaxel and carboplatin	RR	No	E non-inferior to C for RR, but C superior for secondary end point of OS
Fossella	NSCLC	C: vinorelbine and cisplatin E 1: docetaxel and cisplatin E 2: docetaxel and carboplatin	OS	Yes, using ERM	E 1 non-inferior and superior to C E 2 non-inferior to C
O'Brien	BC	C: conventional doxorubicin E: pegylated liposomal doxorubicin	PFS	No	E non-inferior to C
Hanna	NSCLC	C: docetaxel E: pemetrexed	OS	Yes, using ERM	E non-inferior to C
Conte	BC	C: concurrent epirubicin and paclitaxel E: sequential epirubicin and paclitaxel	RR	Yes, on clinical grounds	E non-inferior to C
Chan	BC	C: epirubicin and cyclophosphamide E: liposomal doxorubicin and cyclophosphamide	RR	No	E non-inferior to C
Gradishar	BC	C: conventional paclitaxel E: nanoparticle albumin-bound paclitaxel	RR	No	E non-inferior and superior to C
Schuette	NSCLC	C: docetaxel every 3 weeks E: weekly docetaxel	OS	No	Equivalence sought, but results not presented as such; no significant differences between groups in OS
Ramlau	NSCLC	C: docetaxel E: oral topotecan	1-year OS	No	E non-inferior to C
Ohe	NSCLC	C: cisplatin and irinotecan E 1: carboplatin and paclitaxel E 2: cisplatin and gemcitabine E 3: cisplatin and vinorelbine	1-year OS	No	Non-inferiority not demonstrated, but there were no significant differences between C and E 1, E 2, or E 3
Novello	NSCLC	C: cisplatin and gemcitabine × 5 E: cisplatin and gemcitabine × 2 followed by gemcitabine × 3	OS	Yes, on clinical grounds	Non-inferiority not demonstrated, but there was no significant difference between C and E

(Continued)

Table II. (Continued).

First author	Tumor type	Interventions	Primary end point	Margin justification	Primary results
Park	NSCLC	C: 6 cycles of platinum-based CT E: 4 cycles of platinum-based CT	1-year OS	No	E non-inferior to C
Scagliotti	NSCLC	C: cisplatin and gemcitabine E: cisplatin and pemetrexed	OS	No	E non-inferior to C
Kim	NSCLC	C: docetaxel E: gefitinib	OS**	Yes, using ERM	E non-inferior to C
Mok	NSCLC	C: carboplatin and paclitaxel E: gefitinib	PFS	No	E non-inferior and superior to C
Maruyama	NSCLC	C: docetaxel E: gefitinib	OS	No	Non-inferiority not demonstrated, but there was no significant difference between C and E

BC, breast cancer; C, control arm; CT, chemotherapy; E, experimental arm; ERM, effect retention method; NSCLC, non-small cell lung cancer; OS, overall survival; PFS, progression-free survival, time to tumor progression, or time to treatment failure; RR, response rate.
*RR was a co-primary end point.
**Superiority in OS in the subgroup with high epidermal growth factor receptor-gene-copy number was co-primary objective.

the 22 NI trials for which positivity could be ascertained achieved their primary objective (nine in BC and eight in NSCLC). For two of the five non-inferiority trials that were negative for their primary objective, the absence of significant treatment differences between groups was mentioned in the abstract of the paper.

We could ascertain trial positivity for the primary endpoint in all superiority trials ($n = 171$), 60 of which were positive. Thus, non-inferiority trials were more likely than superiority trials to yield positive results ($p < 0.001$). Tumor type was also a significant predictor of trial positivity (48.9% in BC vs. 26.4% in NSCLC; $p = 0.02$), whereas number of patients per arm and study period were not associated with trial positivity (data not shown).

Discussion

In the current study, we found that nearly one in eight recent phase III trials in advanced BC and advanced NSCLC had a non-inferiority primary hypothesis, which was demonstrated in nearly 80% of cases. There appears to be a perceived notion that non-inferiority cancer trials are increasing in frequency, but if we found evidence of such trend in NSCLC, we observed the opposite trend in BC. More importantly, our work demonstrates that most trials use non-inferiority margins with no reported justification for their choice. The frequent similarities in the margins used both across trials with time-to-event endpoints (typically, a lower limit of the 95% CI for the hazard ratio of 0.8) and those with categorical endpoints (a lower limit of the 95% CI for the difference between groups ranging from 10% to 25%) suggest that such margins are based on tradition rather than on statistical calculations specific to each situation. Of note, the current study suggests that only a small fraction of non-inferiority trials in oncology use the frequently recommended one-sided type I error rate of 2.5% [1,3]. Few trials use the PP population as the primary analysis population, which arguably leads to more conservative inference about non-inferiority, though this is a matter of debate [10,11]. Even though the median number of patients was significantly higher in non-inferiority trials than in superiority trials, still higher sample sizes would be needed if most trials used one-sided type I error rates of 2.5% (rather than 5%). Finally, our study demonstrates that non-inferiority trials are more likely to yield positive results than superiority trials. This finding, per se, is of no major concern; however, we believe that the influence of non-inferiority margin width on trial results should be further investigated.

The most important limitation of the current study is publication bias, as we have only analyzed papers on two tumor types published within a limited period of time in selected medical journals. By using this study design, we have not been able to ascertain whether non-inferiority trials on other tumor types, or if those that that remain unpublished or that have been published in other journals have different features from the sample of trials analyzed herein. However, BC and NSCLC are the two tumor types with the largest number of randomized trials, and the journals analyzed publish most of the trials that are likely to influence practice. Lange and Freitag have conducted a systematic review of 332 non-inferiority and equivalence trials, 29 of which in oncology [12]. Notwithstanding the fact that no details particular to the oncology trials were presented in their broad review, the general findings of their study suggested substantial variability on the choice of non-inferiority/equivalence margins. Moreover, ideal requirements concerning the choice of such margins were fulfilled in only about 8% of all trials.

In the present study, 100% of the trials reported the non-inferiority margin, but with some justification in only 20.8% of cases. Similarly, the non-inferiority or equivalence margin was specified in 96.3% of 162 reports retrieved by French investigators through MEDLINE and the Cochrane Central Register of Controlled Trials and published in a two-year period, but a justification for the margin was provided in only 20.4% of reports [13]. Likewise, the rate of positive results that we found is nearly the same as the one reported by Dutch investigators in a recent meta-analysis of non-inferiority trials conducted in various medical fields. The 170 trials analyzed had been published in 121 core clinical journals, and the experimental treatment was considered to be non-inferior in 74% of cases [14]. Given the similarities between the main results of our study in two important fields in medical oncology and the results from broader assessments [12–14], we believe our results allow for an overview of contemporary statistical practice regarding non-inferiority trials in medical oncology.

The choice of the non-inferiority margin involves statistical and clinical reasoning, especially in oncology. For medical fields and clinical settings for which a placebo control is appropriate, the choice of the non-inferiority margin aims at retaining some of the effect of active control over placebo. As a result, such choice may be done following a few possible rules, one of which is to define the non-inferiority margin as 50% of the treatment effect when the active control was compared with placebo in a previous trial [1,3]. In fact, three of the five non-inferiority trials with margin justification reportedly used such method for determining the non-inferiority margin [15–17]. In addition to this 'effect retention method', other techniques may be used for selecting the non-inferiority margin. However, there is no consensus regarding which technique is superior. What is most important is that physicians participating in a trial, the larger community of physicians and patients, and regulatory agencies have a clear understanding of what should be a clinically relevant difference in a given clinical setting [18].

A critical issue regarding non-inferiority trials is the importance of seeing them through to planned completion. The key premise of these trials is that the difference in efficacy between the treatments under comparison is sufficiently small to permit use of the one that provides other benefits. The estimated efficacy difference between treatments becomes more precise as information accumulates, and thus early stopping is rarely warranted or desirable in non-inferiority trials [18]. In the current study, we have found that most of the non-inferiority trials analyzed achieved their target accrual, although such feature could not be ascertained for the six trials in which the planned number of patients was not reported.

For superiority trials, the ITT population is considered the primary analysis population because it tends to avoid the over-optimistic estimates of efficacy that may result from a PP population [19]. The withdrawal of patients after inclusion in the study may introduce bias towards the type I error, i.e. an exacerbation of differences between groups [20]. Thus, it may be of interest to conduct the primary analysis of non-inferiority trials in the PP population, as some authors have argued that this is a more conservative approach in these trials [2,10]. However, there is no consensus among statisticians and researchers about the best population for primary analysis in non-inferiority trials, and it has been argued that both ITT and PP analyses should be conducted and reported [11,13,19,21]. Interestingly, only four trials analyzed herein reported primary analyses involving the PP population, whereas the vast majority gave precedence to the ITT population.

In summary, the current study suggests that non-inferiority margins for cancer trials are often chosen on historical grounds. Given the fact that most non-inferiority trials achieve their primary objective, the extent to which the choice of margins has influence on such results remains to be determined.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

[1] Rothmann M, Li N, Chen G, Chi GY, Temple R, Tsou HH. Design and analysis of non-inferiority mortality trials in oncology. Stat Med 2003;22:239–64.

[2] Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of noninferiority and equivalence randomized trials: An extension of the consort statement. JAMA 2006;295:1152–60.

[3] U.S. Department of Health and Human Services. Food and drug administration. Center for drug evaluation and research (cder). Center for biologics evaluation and research (cber). Guidance for industry. Non-inferiority clinical trials. Accessed 2011 Sep 4. Available from: http://www.Fda.Gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm202140.Pdf

[4] Temple R, Ellenberg SS. Placebo-controlled trials and active-control trials in the evaluation of new treatments. Part 1: Ethical and scientific issues. Ann Intern Med 2000;133:455–63.

[5] Hung HM, Wang SJ, O'Neill R. A regulatory perspective on choice of margin and statistical inference issue in non-inferiority trials. Biom J 2005;47:28–36; discussion 99–107.

[6] Carroll KJ. Active-controlled, non-inferiority trials in oncology: Arbitrary limits, infeasible sample sizes and uninformative data analysis. Is there another way? Pharm Stat 2006;5:283–93.

[7] European Medicines Agency. Committee for medicinal products for human use (chmp). Guideline on the choice of the non-inferiority margin. (July 2005) Accessed 2011 Sep 3. Available from: http://www.Ema.Europa.Eu/docs/en_gb/document_library/scientific_guideline/2009/09/wc500003636.Pdf

[8] Saad ED, Katz A, Buyse M. Overall survival and post-progression survival in advanced breast cancer: A review of recent randomized clinical trials. J Clin Oncol 2010;28:1958–62.

[9] Saad ED, Katz A. Progression-free survival and time to progression as primary end points in advanced breast cancer: Often used, sometimes loosely defined. Ann Oncol 2009;20:460–4.

[10] International conference on harmonisation. Ich topic e 9. Statistical principles for clinical trials. Accessed on 2011 Aug 15. Available from: http://www.Ich.Org/fileadmin/public_web_site/ich_products/guidelines/efficacy/e9/step4/e9_guideline.Pdf

[11] Garrett AD. Therapeutic equivalence: Fallacies and falsification. Stat Med 2003;22:741–62.

[12] Lange S, Freitag G. Choice of delta: Requirements and reality – results of a systematic review. Biom J 2005;47:12–27.

[13] Le Henanff A, Giraudeau B, Baron G, Ravaud P. Quality of reporting of noninferiority and equivalence randomized trials. JAMA 2006;295:1147–51.

[14] Soonawala D, Middelburg RA, Egger M, Vandenbroucke JP, Dekkers OM. Efficacy of experimental treatments compared with standard treatments in non-inferiority trials: A meta-analysis of randomized controlled trials. Int J Epidemiol 2010;39:1567–81.

[15] Fossella F, Pereira JR, von Pawel J, Pluzanska A, Gorbounova V, Kaukel E, et al. Randomized, multinational, phase III study of docetaxel plus platinum combinations versus vinorelbine plus cisplatin for advanced non-small-cell lung cancer: The tax 326 study group. J Clin Oncol 2003;21:3016–24.

[16] Kim ES, Hirsh V, Mok T, Socinski MA, Gervais R, Wu YL, et al. Gefitinib versus docetaxel in previously treated non-small-cell lung cancer (interest): A randomised phase III trial. Lancet 2008;372:1809–18.

[17] Hanna N, Shepherd FA, Fossella FV, Pereira JR, De Marinis F, von Pawel J, et al. Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy. J Clin Oncol 2004;22:1589–97.

[18] Dignam JJ. Early viewing of noninferiority trials in progress. J Clin Oncol 2005;23:5461–3.

[19] Matilde Sanchez M, Chen X. Choosing the analysis population in non-inferiority studies: Per protocol or intent-to-treat. Stat Med 2006;25:1169–81.

[20] Lachin JM. Statistical considerations in the intent-to-treat principle. Control Clin Trials 2000;21:167–89.

[21] Brittain E, Lin D. A comparison of intent-to-treat and per-protocol results in antibiotic non-inferiority trials. Stat Med 2005;24:1–10.

Supplementary material available online

Supplementary reference list available online at http://informahealthcare.com/doi/abs/10.3109/0284186X.2012.702924.