ELSEVIER

European Journal of Cancer 40 (2004) 1244–1249

European Journal of Cancer

www.ejconline.com

Five-year change in statistical designs of phase II trials published in leading cancer journals

S. Thezenas[a,*], J. Duffour[b], S. Culine[b], A. Kramar[a]

[a]*Department of Biostatistics, Val d'Aurelle Regional Cancer Center, Parc Euromédecine, 208, rue des Apothicaires, 34298 Montpellier cedex 05, France*
[b]*Department of Medical Oncology, Val d'Aurelle Regional Cancer Center, Montpellier, France*

Received 20 October 2003; received in revised form 30 December 2003; accepted 12 January 2004

Abstract

This study compares the evolution in statistical design reporting for phase II cancer clinical trials published in the six following leading journals: *American Journal of Clinical Oncology*, *Annals of Oncology*, *British Journal of Cancer*, *Cancer*, *European Journal of Cancer* and *Journal of Clinical Oncology*. Only articles where tumour response was considered as the primary endpoint were selected. A total of 393 phase II trials published in 1995 ($n=185$) and 2000 ($n=208$) were reviewed. Neither sample size nor design parameters were specified in 157 (85%) and 113 (46%) papers in 1995 and 2000, respectively. 28 (15%) and 95 (46%) papers included at least some information on the statistical designs used: Gehan (4.3% and 3.3%), Fleming (2.2% and 4.3%), and Simon (2.7% and 11.0%). Ad hoc, non-referenced methods were used in 5.9% and 27.3% articles in 1995 and 2000, respectively. Although there is an increase in the mention of at least some statistical design parameters in phase II cancer clinical trials over a 5-year period in these selected cancer journals, the use of referenced methods is still short or often inadequate.

Keywords: Statistical design reporting: Phase II clinical trials; Leading cancer journals

1. Introduction

In oncology, the importance of phase II clinical trials as part of the drug-development process is clearly accepted, since, following analysis, the molecule may be definitively abandoned if it does not show sufficient therapeutic promise. The determination of the success or failure of a therapeutic treatment after a phase II trial depends to some extent on the quality of the statistical design put in place from the outset. This design generally allows for the continuation or abandon of the trial early on according to the initial conditions. The literature contains numerous statistical methods for the planning and analysis of multistage phase II trials, but their use is far from universal.

Gehan [1] from the beginning of the 1960s developed one of the best-known statistical designs. This design was established in order to reject a molecule if no responses were observed in a first cohort of patients. Twenty years later, in 1982, Fleming [2] developed more efficient analytical designs allowing for the trials to be ended prematurely in case of sufficient efficacy in terms of response. Since then, there has been an increase in the number of articles devoted to the development of statistical methods in this field [3–5]. Each of these recent methods allows the establishment of designs that ensure an ethical dimension by minimising the number of subjects exposed to an insufficiently active drug.

In a recent review of the literature on the statistical quality of 308 articles involving phase II cancer trials undertaken between 1990 and 1996 and published in 1997, Mariani [6] showed that the overall quality of the statistical section was poor. In fact, only 20% of the trials correctly described the statistical methods employed. Previous to this study, a similar study of 83 trials published in five journals in 1995 was carried out [7]. The number of articles was later increased to a total of 156. Our study compares the evolution of the most commonly used statistical designs between 1995 and 2000 in a review of the literature of phase II cancer trials published in six leading cancer journals indexed in Medline.

* Corresponding author. Tel.: +33-4-67-61-30-35; fax: +33-4-67-61-37-18.
E-mail address: sthezenas@valdorel.fnclcc.fr (S. Thezenas).

0959-8049/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2004.01.008

2. Patients and methods

2.1. Identification and selection

An electronic bibliographic search in Medline allowed us to identify the necessary articles for this literature review (Table 1). We limited our search to phase II clinical trials published in 1995 and 2000 in the following journals: *American Journal of Clinical Oncology* (AJCO), *Annals of Oncology* (AO), *British Journal of Cancer* (BJC), *Cancer* (CAN), *European Journal of Cancer* (EJC) and *Journal of Clinical Oncology* (JCO). We did not take into account reports concerning pilot studies, or phase II trials where either the quality of life (two cases) or the PSA rate (three cases) was considered as the primary measure of outcome. However, we kept the phase I/II or II/III trials that listed tumour response as the main criterion for evaluation.

2.2. Data extraction

Using individual forms we extracted the following information from each article: the journal (volume, number, pages), the authors, the country, the cancer site, the study structure, the statistical method used, the number of stages, the *a priori* success probabilities ($p0$, $p1$), the number of planned patients (n), the type I and II error rates (α, β), the number of patients actually included, the number of assessable patients, the type and number of tumour responses: complete response, partial response, stable disease or progression. This information was computerized into an ACCESS database. The data were then transferred by *Stat/Transfer* to *Stata 7.0* for analysis.

2.3. Statistical methods

Categorical variables were reported by means of contingency tables according to year of publication. Continuous variables, such as the number of assessable patients, are presented with summary statistics. The number of assessable patients was also categorized into four groups ($\leqslant 20$, 21–40, 41–60 and 60+). Certain items were grouped together: America (USA, Canada, Mexico and South America); Europe; and 'rest of world' (Japan, China, Taiwan and Australia). In the case of multicentre studies the 'country of origin' value was the country of the first centre mentioned. The journals were also classified according to their impact factor in 2000: the three highest (high) versus low.

To investigate the association between the frequency of statistical design reporting and trial features, univariate statistical analyses were performed using Pearson's χ^2-test, or Fisher's exact test if applicable for categorical variables and using either the non-parametric Kruskal–Wallis test or Student t-test for continuous variables.

In the multivariate analyses we adopted a logistical regression (overall and stepwise) model that included as predictor variables: continent, study structure, tumour type etc. Associations were assessed using Wald's test. Other trial features were not considered in the latter analyses due to the lack of any apparent association at univariate analysis. All P-values reported are two-sided, and differences were considered as significant at the 5% level for all statistical tests.

3. Results

3.1. Sample description

A total of 393 published phase II trials were selected from the following journals (1995 and 2000): AJCO (32 and 25), AO (29 and 56), BJC (17 and 21), CAN (33 and 41), EJC (24 and 12) and JCO (50 and 53). The authors were essentially of American (36%) or European (49%) origin. The proportion of articles published in these countries hardly changed between 1995 and 2000: America (39% and 32%), Europe (50% and 49%) and other (11% and 19%) (Table 2). Similarly, the representation rate for phase II trials in the different journals that we studied remained the same between 1995 and 2000. Among the selected journals the JCO published the most phase II trials, with 27% and 25% in 1995 and 2000, respectively. The most frequent cancer sites cited in the 1995 and 2000 publications were: lung (20%) including lung/NOS, small and non-small cell lung cancer; gastrointestinal cancer (18%) including colon, pancreatic, oesophagus and liver cancer; gynaecological cancer (20%) including breast, ovary and cervix cancer; genitourinary cancer (11%) including kidney, bladder, prostate and testicular cancer.

The study structure, mainly single centre, evolved little between 1995 and 2000, with 52% and 57%, respectively.

Table 1
Search strategy adopted to identify published phase II cancer studies (Medline)

01	Entry date	Year 2000, year 1995
02	Text word	Phase II
03	Publication type	Clinical trials
04	Journal	*Am J Clin Oncol*, *Ann Oncol*, *Br J Can*, *Cancer*, *Eur J Can*, *J Clin Oncol*
05	Language	English or French

3.2. Statistical designs mentioned

In 1995 and 2000, 157 (85%) and 113 (54%) articles, respectively, cited none of the parameters of the statistical method used (α, β, $p0$, $p1$, or sample size) (Table 3). The percentage of articles that mentioned a statistical design increased significantly from 15% to 46% between 1995 and 2000 ($P<0.001$). However, the exact formulations of the designs used were rarely mentioned in the articles examined: Simon (7.1%), Fleming (3.3%), Gehan (3.8%) and other non-referenced (17%). The median number of assessable patients was statistically significant according to whether or not a statistical design was mentioned, 36 (range 8–255) and 42 (range 6–216), respectively ($P=0.039$). The frequency of statistical methods mentioned was similar in articles published by American teams as compared to European ones (37% versus 28%; $P=0.176$). Multicentre trials (39% versus 27%; $P=0.015$), studies with a higher number of assessable patients ($P<0.031$) and journals with high impact factor ($P=0.002$) were significantly related to the mention of a statistical design. Except in one case, all journals greatly increased their percentage of statistical design reporting between 1995 and 2000 (Table 4).

Table 2
Distribution of the 393 identified studies according to context-related factors

	1995 ($n=185$)	2000 ($n=208$)	Total ($n=393$)
Continent			
American	73 (39.5%)	67 (32.2%)	140 (35.6%)
European	92 (49.7%)	102 (49.1%)	194 (49.4%)
Other	20 (10.8%)	39 (18.7%)	59 (15.0%)
Journal			
Am J Clin Oncol	32 (17.3%)	25 (12.0%)	57 (14.5%)
Ann Oncol	29 (15.7%)	56 (26.9%)	85 (21.7%)
Br J Cancer	17 (9.2%)	21 (10.1%)	38 (9.7%)
Cancer	33 (17.8%)	41 (19.7%)	74 (18.8%)
Eur J Cancer	24 (13.0%)	12 (5.8%)	36 (9.2%)
J Clin Oncol	50 (27.0%)	53 (25.5%)	103 (26.2%)
Study structure			
Single centre	86 (52.1%)	119 (57.2%)	205 (55.0%)
Multicentre	79 (47.9%)	89 (42.8%)	168 (45.0%)
Sites			
Lung			
Lung/NOS	2 (1.0%)	4 (1.9%)	6 (1.5%)
SCLC	12 (6.4%)	10 (4.7%)	22 (5.5%)
NSCLC	18 (9.7%)	32 (15.1%)	50 (12.7%)
Gastrointestinal			
Colon	17 (9.1%)	15 (7.1%)	32 (8.1%)
Pancreas	6 (3.2%)	13 (6.2%)	19 (4.8%)
Oesophagus	5 (2.7%)	9 (4.3%)	14 (3.5%)
Liver	2 (1.0%)	2 (0.9%)	4 (1.0%)
Other GI	1 (0.5%)	1 (0.4%)	2 (0.5%)
Gynaecological			
Breast	25 (13.1%)	26 (12.4%)	51 (12.9%)
Ovary	12 (6.4%)	12 (5.7%)	24 (6.0%)
Cervix	4 (2.1%)	1 (0.4%)	5 (1.2%)
Genitourinary			
Kidney	7 (3.7%)	11 (5.2%)	18 (4.5%)
Bladder	4 (2.1%)	6 (2.8%)	10 (2.5%)
Prostate	5 (2.7%)	5 (2.3%)	10 (2.5%)
Testicular	3 (1.6%)	2 (0.9%)	5 (1.2%)
Others			
Brain	2 (1.0%)	0 (0.0%)	2 (0.5%)
Head/neck	7 (3.7%)	8 (3.8%)	15 (3.8%)
Leukaemia	7 (3.7%)	3 (1.4%)	10 (2.5%)
Lymphoma	15 (8.1%)	14 (6.7%)	29 (7.3%)
Nervous system	6 (3.2%)	12 (5.7%)	18 (4.5%)
Peritoneum	0 (0.0%)	1 (0.4%)	1 (0.2%)
Sarcoma	8 (4.3%)	3 (1.4%)	11 (2.7%)
Skin	11 (5.9%)	7 (3.3%)	18 (4.5%)
Other	6 (3.2%)	12 (5.7%)	18 (4.5%)

SCLC, small cell lung cancer; NSCLC, non-small cell lung cancer.

3.3. Design parameters

Among the 123 articles mentioning a statistical design, two-stage designs were the most often used (65%), and were reported in 64% and 83% in 1995 and

Table 3
Distribution of the 393 identified studies according to design parameters

	1995 ($n=185$)		2000 ($n=208$)		Total ($n=393$)	
Method						
Not indicated	157	(84.8%)	113	(54.1%)	270	(68.5%)
Gehan	8	(4.3%)	7	(3.3%)	15	(3.8%)
Fleming	4	(2.2%)	9	(4.3%)	13	(3.3%)
Simon	5	(2.7%)	23	(11.0%)	28	(7.1%)
Other	11	(5.9%)	57	(27.3%)	68	(17.3%)
Number of included patients						
Mean (S.D.)	45.8	(37.6)	55.9	(47.6)	50.7	(43.3)
Median	38		44		41	
Range	6–295		9–455		6–55	
Assessable patients						
NP	1	(0.5%)	18	(8.6%)	19	(4.8%)
6–20	43	(23.2%)	27	(12.9%)	70	(17.8%)
21–40	71	(38.4%)	67	(32.1%)	138	(35.0%)
41–60	44	(23.8%)	58	(27.7%)	102	(25.9%)
>60	26	(14.1%)	38	(18.7%)	64	(16.5%)

Table 4
Percent evolution of statistical design reporting in 1995 and 2000

Journal*	% Statistical design report	
	1995	2000
AJCO	12.5	36.0
Br J Cancer	29.4	28.6
Cancer	15.1	56.1
Eur J Cancer	8.3	33.3
JCO	20.0	59.3
AO	6.9	39.3

*Full names of journals appear in the text.

2000 ($P=0.032$). 26% did not mention a theoretical-success probability nor the planned sample size (23%) (Table 5). However, the success percentage reported was similar between 1995 and 2000 (71% versus 75%, not significant).

The value of the probabilities of maximal inefficacy ($p0$) and minimal efficacy ($p1$) were rarely mentioned. The median values for $p0$ (20%) and $p1$ (30%) were similar between 1995 and 2000 (range 10–80 and 5–85).

The median difference, δ ($p1$–$p0$), calculated on only 44 articles, was also constant at 20% (range 1995: 15–25; range 2000: 5–25.5).

The median type I α-error values were 0.05 (range 0.04–0.10) and 0.05 (range 0.03–0.16) in 1995 and 2000, respectively. The median type II β-errors were 0.10 (range 0.05–0.10) and 0.12 (range 0.05–0.20) for 1995 and 2000, respectively.

3.4. Number of patients

The median number of included patients increased significantly (38 versus 44; $P=0.004$) between 1995 and 2000. The number of planned subjects, mentioned in only 10% and 37% of articles in 1995 and 2000, varied between 14 and 250 (median 40.5). Protocol adherence was good, as the median difference in the number of assessable patients and the number of planned patients was zero (range −88 to 89). The number of patients actually included was stated in 388 trials (99%). The number of patients assessable for response was recorded in 374 (96%) trials, with a median value of 38 patients (range 6–255). 64 articles (16%) included more than 60 assessable patients. In 24 out of 393 (6%) trials the number of assessable patients was greater than 100. The higher the number of assessable patients, and higher was the percent of statistical design reported ($P=0.031$). Finally, 14 trials, all in 2000, specified nothing other than the number of patients included.

Table 5
Description of the features in articles with a statistical design reported

	1995 ($n=28$)	2000 ($n=95$)	Total ($n=123$)
Stages			
Number of stages indicated			
No	10 (35.7%)	16 (16.8%)	26 (21.1%)
Yes	18 (64.3%)	79 (83.2%)	97 (78.9%)
1	1 (3.6%)	13 (13.7%)	14 (11.4%)
2	17 (60.7%)	63 (66.3%)	80 (65.0%)
3	0 (0.0%)	3 (3.2%)	3 (2.4%)
% Success			
No	8 (28.6%)	24 (25.3%)	32 (26.0%)
Yes	20 (71.4%)	71 (74.7%)	91 (74.0%)
δ ($p1$–$p0$)[a]			
0.05	0 (0%)	2 (5.9%)	2 (4.6%)
0.10	0 (0%)	4 (11.8%)	4 (9.1%)
0.15	1 (10%)	5 (14.7%)	6 (13.6%)
0.16	0 (0%)	1 (2.9%)	1 (2.3%)
0.20	8 (80%)	18 (52.9%)	26 (59.1%)
⩾0.25	1 (10%)	4 (11.7%)	5 (11.3%)
Expected patients			
No	9 (32.1%)	19 (20.0%)	28 (22.8%)
Yes	19 (67.9%)	76 (80.0%)	95 (77.2%)
Assessable patients			
Mean	46.0	50.5	49.4
S.D.	29.2	34.2	33.1
Median	42	42	42
Range	6–172	12–216	6–216

[a] A total of 44 trials.

3.5. Number of responses

The number of objective responses was mentioned in 337 (86%) trials, with a median percentage of objective responses of 37% (95% confidence interval (CI): 34–40%). This percentage did not change significantly between 1995 and 2000, with 38% (95% CI: 34–42%) and 37% (95% CI: 33–41%), respectively. Conversely, 32 trials (8.1%) with an average of 21 patients (range 6–69) declared a 0% objective response rate. Among these trials, 20 out of 32 trials (62%) did not mention the statistical method used. We observed, in 2000, a lower rate of objective responses in trials that clearly specified the statistical method used (30% versus 43%; $P=0.0001$); in contrast to 1995 (36% versus 38%; $P=0.697$). Besides, four trials declared a 100% objective-response rate, but without clearly specifying the statistical design used.

3.6. Association between statistical design and parameters

The frequencies of the reporting of statistical designs according to the different parameters and years are presented in Table 6. Significant results were observed for the type of study ($P=0.015$), the number of assessable patients ($P=0.031$), the journal ($P=0.05$) and the year of publication ($P=0.0001$). The continent was not statistically significant ($P=0.176$).

Using a stepwise logistic-regression model we were able to identify the most significant variables associated with the reporting of statistical designs in the published articles (Table 7): multicentre trials (odds ratio = 2.0; $P=0.011$), articles published in 2000 (odds ratio = 6.7, $P=0.0001$), and journals with the higher impact factor (odds ratio = 2.1; $P=0.007$).

4. Discussion

The importance of using a well-planned statistical design during the protocol writing of a phase II trial is indisputable. Funding agencies are more and more

Table 6
Number of studies with statistical designs mentioned

	D*/T	%	*P*
Continent			
American	52/140	37.1	
European	54/194	27.8	
Other	17/59	28.8	0.176
Year			
1995	28/185	15.1	
2000	95/208	45.7	0.000
Structure			
Single centre	55/205	26.8	
Multicentre	65/168	38.7	0.015
Assessable patients			
0–20	14/70	20.0	
21–40	44/138	31.9	
41–60	35/102	34.3	
>60	28/64	43.7	0.031
Journal[a]			
AJCO	13/57	22.8	
Ann Oncol	24/85	28.2	
Br J Cancer	11/38	28.9	
Cancer	28/74	37.8	
Eur J Cancer	6/36	16.7	
JCO	41/103	39.8	0.050

*D, number of studies with a statistical method specified; T, total number of studies.
[a] Full names of journals appear in the text.

Table 7
Results from a multivariate logistic regression of significant variables on the mention of a statistical design

	OR	SE	*P*	95% Confidence interval
Overall model				
Year of publication				
1995	1*			
2000	6.67	1.9	<0.001	3.83–11.62
Continent				
America	1*			
European	0.56	0.2	0.05	0.32–0.99
Others	0.55	0.2	0.13	0.25–1.18
Number of evaluable patients				
⩽20	1*			
21–40	1.87	0.8	0.12	0.84–4.15
41–60	1.53	0.8	0.32	0.66–3.56
>60	2.00	1.2	0.14	0.80–4.99
Study structure				
Single centre	1*			
Multicentre	1.99	0.5	0.01	1.17–3.37
Journal impact factor (JIF)				
Low	1*			
High	2.11	0.6	<0.01	1.22–3.64
Stepwise model				
Year 2000	6.4	1.8	<0.001	3.77–11.16
Multicentre	2.1	0.5	<0.01	1.28–3.45
High JIF	2.5	0.6	<0.001	1.48–4.13

*Reference categories.

reluctant about their involvement in poorly designed studies, and statistical sections are being scrutinised more carefully. The statistical literature for the design of multistage phase II cancer trials has been increasing over the past years but their use in actual studies is still inadequate, due partly to the transfer of this new technology to the production floor. It is nevertheless very important to quickly bring these methods into use, since they are based on ethical principles that minimise the number of patients exposed to an inferior treatment. It is no longer possible to keep giving patients treatments that clearly are not working, especially in cancer. The application of the statistical design to the trial results guarantees trials of high quality and sufficient validity.

Lee [8], Kramar [9] and Mariani [6] have already drawn our attention to these issues with other similar literature reviews involving phase II clinical trials. Our study, like that of Mariani, is certainly biased by the selection of journals and the choice of a single search engine (Medline). However, scientists are in agreement over the superior quality of the articles referenced in Medline. Also, the journals were selected specifically in order to allow for a comparison with a preliminary study performed in 1995 [7,9]. In 2000, a little over half of the articles (54%) did not detail the statistical method employed during the planning stage of phase II trials. However, this number is a quantitative improvement on the figures from 1995, when only 15% reported the use of a statistical design. The articles with methods classified as 'other' concerned those articles that actually presented a statistical method without clearly identifying it: the two-stage designs could have been Gehan, Fleming or Simon. All of the journals selected except one had a significant rise in the publication of articles mentioning the statistical design. At the same time, this improvement should not hide certain observations.

In agreement with Lee [8] and Mariani [6], in the majority of cases the statistical design is almost never suitably named and is too often badly clarified: it does not record the necessary information indispensable to its construction: α (type I error), β (power), $p0$ (maximum probability of inefficacy), $p1$ (minimum probability of efficacy), δ ($p1$–$p0$), and the number of subjects required.

In far too many articles, when present at all, the section 'statistical observations' or 'statistical methodology' is only a simple summary of the different univariate tests (χ^2, Fisher, Student) used for analysis. It would be appropriate if journal reviewers demanded more details of the statistical design used in each article submitted for publication. The number of assessable and the number of included patients were systematically recorded. This also applies to the number of stages. However, it would have been equally interesting to be able to know the number of subjects initially planned at the time of the protocol design. This would allow a ver-

ification of the compliance of the trials with their respective protocols.

Contrary to the results reported in the study by Lee [8], in our review of the literature, for the articles mentioning the elements that permitted such a review, compliance, the difference between the number of patients planned and the numbers of assessable patients were quite good. Contrary to preconceptions, some phase II trials (7% in our study) can be carried out on a large number of patients.

We have shown in this article that the phase II trials that reported the statistical method had a statistically smaller percentage of objective responses. There is probably no relation between cause and effect in this finding, yet it makes us aware of the fact that it is impossible to control the appropriateness and the exactness of the responses obtained in trials without a planned statistical design. This last result is counterbalanced by the fact that 62.5% of the trials that did not obtain any objective responses did not mention the statistical method.

Several factors seem to correlate with the presence of a statistical method in articles and our results are in accord with those of Mariani [6] on this point. The country of origin, journals with the higher impact factor, study structure, large sample size and the year of publication are all influential factors on the quality of the statistical composition of the articles. We are not questioning the quality of the articles, we are just drawing attention to a section that is often present but insufficiently detailed in the articles.

Many errors could easily be avoided if the biostatistical part of the article was written in a manner conforming to the guidelines published in the ICH9 recommendation [10,11] and/or checked by a biostatistician.

Taking into account the time for accrual, analysis, writing, acceptance and publication, the phase II trials that were published in 2000 were obviously started in earlier years. This explains why some of the more recent methods that incorporate toxicity [5] and early progression [12] are not yet represented in this literature review, as they were most certainly either still unknown or far too recent at the time the trials began.

In conclusion, the present study shows that the quality of the statistical section of the articles published involving phase II trials has improved considerably since 1995. However, one-half of articles are still published despite the fact that they do not, or hardly, mention the statistical method employed. Furthermore, in those that do mention it, some effort still needs to be put into making sure that it is done correctly, including all of the parameters allowing it to be properly identified. We plan to repeat this study in 2005 to monitor the evolution of statistical reporting in cancer journals.

Acknowledgements

We would like to thank N. Syz and S. Gourgou-Bourgade for their help in gathering the data, and G. Varey for her help in translation.

References

1. Gehan EA. The determination of the number of patients required in a preliminary and a follow-up trial of a new chemotherapeutic agent. *J Chronic Dis* 1961, **13**, 346–353.
2. Fleming TR. One-sample multiple testing procedures for phase II clinical trials. *Biometrics* 1982, **38**, 143–151.
3. Simon R. Optimal two-stage designs for phase II clinical trials. *Control Clin Trials* 1989, **10**, 1–10.
4. Ensign LG, Gehan EA, Kamen DS, *et al*. An optimal three-stage design for phase II clinical trials. *Stat Med* 1994, **15**, 1727–1736.
5. Bryant J, Day R. Incorporating toxicity considerations into the design of two stage phase II clinical trials. *Biometrics* 1995, **51**, 1372–1383.
6. Mariani L, Marubini E. Content and quality of currently published phase II cancer trials. *Journal of Clinical Oncology* 2000, **8**, 429–436.
7. Vabre F, Kramar A. Evaluation du nombre de sujets nécessaires dans la planification des essais de phase II en cancérologie- Revue de la littérature. Eurocancer: John Libbey Eurotext, Paris III.3.4; 1996; 408.
8. Lee YJ. Phase II trials in cancer—present status and analysis methods. *Drugs Exp Clin Res* 1986, **12**, 57–71.
9. Kramar A, Potvin D, Hill C. Multistage designs for phase II clinical trials: statistical issues in cancer research. *Br J Cancer* 1996, **74**, 1317–1320.
10. Simon R, Wittes RE. Methodology guidelines for phase II cancer clinical trials. *Cancer Treat Rep* 1985, **69**, 1–3.
11. ICH9. Statistical principals for clinical trials < CPMP/ICH/363/96.1998.
12. Dent S, Zee B, Dancey J, *et al*. Application of new multinomial phase II stopping rule using response and early progression. *J Clin Oncol* 2001, **19**, 785–791.