Journal of Cancer Clinical Trials

Keikes et al., J Cancer Clin Trials 2:1

Commentary Open Access

Clinical Trial Performance in Metastatic Colorectal Cancer: An Evaluation of Participating Centers in the CAIRO studies of the Dutch Colorectal Cancer Group.

Lotte Keikes[1*], Linda Mol[2*], Martijn G.H. van Oijen[1], Miriam Koopman[3], Cornelis J.A. Punt[1], Petronella B. Ottevanger[4]

[1]*Department of Medical Oncology, Academic Medical Centre, Amsterdam, The Netherlands*

[2]*Clinical research department, Netherlands Comprehensive Cancer Organisation, The Netherlands*

[3]*Department of Medical Oncology, University Medical Centre, Utrecht, The Netherlands*

[4]*Department of Medical Oncology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands*

***Corresponding author:** Lotte Keikes, Department of Medical Oncology, Academic Medical Centre, Amsterdam, The Netherlands Tel :+31205665955; Fax: +31206919743; E-mail: l.keikes@amc.uva.nl

Received date: February 03, 2017, **Accepted date:** February 23, 2017, **Published date:** February 28, 2017

Abstract

Objective: High quality clinical trials are essential for further improvement of treatment strategies for prolonged survival and reliable evidence-based outcomes. However, there are no defined standards for the quality of clinical trial performance. The aim of this study is to examine and compare clinical trial performance with a composite score between (different types of) hospitals, to identify potentially predicting factors for a high trial performance and examine a learning curve in composite performance scores between early compared to subsequent included patients.

Methods: We evaluated trial performance in three large phase 3 randomized clinical trials in metastatic colorectal cancer (CAIRO studies of the Dutch Colorectal Cancer Group, total n=2131) with a newly introduced composite score, consisting of stratification errors, major protocol violations, number of included ineligible patients, and reporting of serious adverse events (SAE) on hospital and patient level. These data were supplemented with a hospital survey containing questions about number of beds, oncologists and research nurses. A logistic regression was performed to identify factors associated with better trial performance (3-4 points).

Results: We observed variation in trial performance between 84 participating hospitals. However, no differences in performance between hospital categories (university, teaching, regional hospitals) were identified and none of the examined variables could be linked to a high composite performance score. In top 10 ranking hospitals with highest inclusion rates, trial performance on patient level was significantly lower in the first three inclusions compared to subsequent patients.

Conclusions: Trial performance was comparable between different types of hospitals and no factors were able to predict a high composite trial performance score. In the highest including hospitals we identified a learning curve for trial performance. We therefore recommend increased support during the first patient inclusions in participating centres in order to improve trial performance. Our composite score could be used as a quality metric for trial performance for individually based hospital evaluation.

Keywords Colorectal cancer, quality of care, quality assurance, trial performance, phase 3 randomized clinical trial, protocol violations, quality metric

Introduction

Survival in metastatic colorectal cancer has improved substantially over time, which is for a large part due to the availability of more effective systemic therapies [1,2]. Clinical trials are essential in this process, and will be necessary for further improvement of treatment. Maintaining high quality within these trials is a requisite for reliable evidence-based outcomes. However, even in the presence of guidelines for the standard of care, there are no defined standards for the quality of clinical trial performance.

Quality assurance of clinical trials is a complex issue. Several factors contributing to a good performance can be identified [3], such as the requirement for adequate methodology and protocols in order to maintain reliability and validity of the obtained results. However, data on the level of protocol adherence are scarce. Multi-centre trials, especially if they are complex, may lead to variability in treatment and data collection. To prevent biased results and to maintain integrity of the data, trial protocols need to be followed as closely as possible. Quality assurance is therefore a prerequisite.

Although no validated indicators for clinical trial performance are available, a potentially useful criterion for clinical trial adherence may be the number of protocol deviations or protocol violations. Protocol deviations are not caused or preventable by the investigator, in contrast to protocol violations. Therefore, the number of protocol violations may be used to assess and compare (investigator) trial performance

[4]. Major protocol violations are defined as deviations, which may result in harm to the patient and may impact the integrity of data. These violations may have major impact on data interpretation and may result in the assumption of wrong recommendations [4,5]. Protocol violations are usually underreported and differ widely among studies [5]. The number of study participants included per participating hospital in multicentre trials is earlier suggested as a potential indicator of trial performance [6]. However, this finding prompted several comments that agreed [7,8] and disagreed with this indicator [9,10]. Also conflicting results regarding trial performance between different types of hospitals have been reported [6,11,12]. In case of any reported difference in performance, there are no data on the underlying contributing causes. Therefore, the aims of our study are: 1) to examine clinical trial performance of hospitals that participated to national phase 3 studies in metastatic colorectal cancer using a scoring system based on stratification errors, major protocol violations, serious adverse event (SAE) reporting and the number of ineligible patients that was included; 2) to identify factors that may explain differences, if any, in clinical performance between different hospital categories; 3) to compare clinical trial performance between hospitals with low and with high accrual rates; and 4) to examine whether a learning curve can be identified per hospital between early included patients compared to subsequent included patients in the trial.

Methods

We merged data from all hospitals participating in three nationwide large phase 3 randomized clinical trials in metastatic colorectal cancer for this study: CAIRO [NCT00312000] [13], CAIRO2 [NCT00208546] [14], and CAIRO3 [NCT00442637] [15]. All Dutch hospitals were allowed to participate in the CAIRO studies. These studies investigated highly relevant clinical research questions for patients with metastatic colorectal cancer, to which the great majority of Dutch hospitals participated.

The CAIRO study identified comparable overall survival for combination versus sequential use of cytotoxic drugs; the CAIRO 2 study investigated the addition of cetuximab to fluoropyrimidine-based chemotherapy and bevacizumab which resulted in shorter progression-free survival and inferior quality of life and the CAIRO 3 study examined and concluded survival benefit for maintenance treatment with capecitabine and bevacizumab after induction treatment with capecitabine, oxaliplatin and bevacizumab. These studies are published in leading medical journals and include accurate on-site monitoring which is essential to answer our research questions.

The Dutch Colorectal Cancer Group (DCCG) was the sponsor of all studies, and a total of 84 hospitals participated. This is a representative amount since there are 88 hospitals in The Netherlands. Of the participating hospitals 43 (51%) were regional hospitals, 32 (38%) teaching hospitals, 8 (10%) university hospitals, and 1 cancer institute (1%).

We have added the latter to the university hospitals group in all analyses. For all 3 studies, regional initiation meetings were organized, and accrual and participation in the study was only allowed in hospitals of which relevant staff had been present at these meetings.

The primary outcome of our study was trial performance, consisting of a composite of 4 dichotomous items: 1) number of errors during stratification (cut-off 10%); 2) number of major protocol violations (cut-off 10%); 3) number of included ineligibles (cut-off 5%), and 4) reported serious adverse events (SAE) within 7 days (cut-off 75%). For all 4 items a participating hospital could gain 1 point: the composite score ranges between 0 and 4 per participating hospital per study. A higher score indicates a better trial performance.

The cut-off scores are based on relevance for clinical practice: therefore the cut-off value for ineligibles is for example lower than the cut-off value for stratification errors, as ineligibility is expected to be more harmful than a stratification error. All data regarding the primary outcome were collected prospectively during the trials with extensive on-site monitoring and verification of the study variables in each individual hospital. To compare hospitals with low and high accrual rates, scores were compared for sites with low (less than 5 patients) and high (5 or more patients) accrual.

To identify variations in trial performance between early and later patient inclusions per hospital (learning curve), we labelled the first 3 included patients in the top-10 hospitals with highest inclusion rates as 'early included' and subsequent patients as 'later included'. An adjusted composite score of 3 dichotomous items (stratification errors, major protocol violations and ineligibility) was calculated at patient level, SAE reporting was excluded since SAE's may occur at any time during the course of the study and therefore may not be a valid measurement to identify a learning curve.

A questionnaire was sent to local investigators of participating hospitals to the CAIRO3 study during its conduct (March 2008), which consisted of the following items, which were to be scored per hospital: the number of beds, the full-time equivalence (FTE) of medical oncologists, the number of hours per week of a research nurse and the number of newly diagnosed patients with colorectal cancer per year.

Data from the questionnaires were entered into an electronic database and merged with the data of the primary outcome per participating hospital per study. Characteristics of participating hospitals were compared between the three categories of hospitals. Categorical variables were analysed using chi-squared testing, or fisher's exact test if appropriate. Continuous variables were compared between groups using Kruskal-Wallis analysis. The primary outcome and its individual components were tested against the three hospital categories using a chi-squared test. A uinvariable en multivariable logistic regression analysis was performed to identify factors associated with a higher composite performance score and additionally univariable logistic regression was performed on the individual components of the composite score. All analyses were performed for all 3 trials combined as well as for each individual trial. All tests were two-sided and a p-value of <0.05 was considered to be statistically significant.

Results

Overall, 84 Dutch hospitals participated in one or more CAIRO studies. A total of 66 hospitals participated in the CAIRO study (total included patients; n=820), 73 hospitals in the CAIRO2 study (n=755), and 61 hospitals in the CAIRO3 study (n=556).

This concerned 75%, 83% and 69% of all Dutch hospitals, respectively. The response rate to the questionnaire was 45% (38/84).

The characteristics of participating hospitals are shown in Table 1.

Variable		University hospitals	Teaching hospitals	Regional hospitals	p-value
Number of accruing hospitals	CAIRO	7	26	33	
	CAIRO2	9	31	33	
	CAIRO3	7	27	27	
Accrual per hospital					
CAIRO	Median (IQR)	14 (8-19)	18 (9-24)	7 (5-12)	<0.05
CAIRO2	Median (IQR)	12 (8-16)	13 (8-16)	7 (4-10)	<0.05
CAIRO3	Median (IQR)	7 (3-12)	7 (4-11)	8 (2-12)	0.91
Number of beds per hospital	Median (IQR)	882 (715-1200)	666 (480-930)	384 (314-486)	<0.05
Number of hours research nurse per week per hospital	Median (IQR)	45 (7-80)	16 (0-43)	3 (0-16)	0.18
Number of full-time oncologists per hospital	Median (IQR)	9 (6-16)	3 (2-4)	2 (2-2)	<0.05

Table 1: Characteristics of participating hospitals.

The median inclusion of patients was significantly different between categories of hospitals in the total dataset with a median inclusion of 12 patients in university hospitals, 11 in teaching hospitals and 7 in regional hospitals (p <0.01). Furthermore, in each study the median inclusion in university, teaching and regional hospitals was 14, 18 and 7 patients (p<0.05) in CAIRO, 12, 13 and 7 patients (p< 0.05) in CAIRO2, and 7, 7 and 8 patients in CAIRO3 (p=0.91), respectively (Table 1).

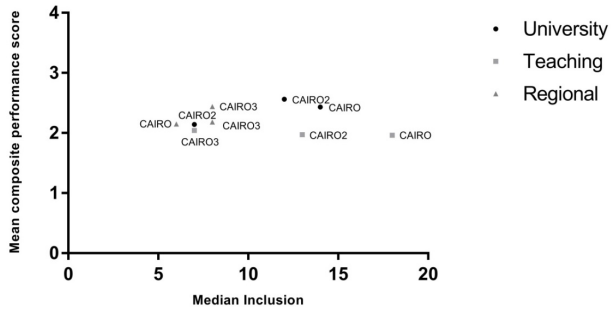

Figure 1: Mean composite performance score

The median overall composite performance score was 2 (IQR 2-3) which was not significantly different between categories of hospitals: university 2 (IQR 2-3), teaching 2 (IQR 1-3) and regional 2 (IQR 2-3). The median overall composite performance score was 3 (IQR 2-3) for hospitals including less than 5 patients, and 2 (IQR 1-3) for hospitals including 5 patients or more (p <0.01).

In Figure 1, the mean composite performance score is plotted against the median inclusion rates for the different types of hospitals for CAIRO, CAIRO2 and CAIRO3. In CAIRO and CAIRO2, the university hospitals had the highest score, but in the CAIRO3 study, regional hospitals scored highest.

Twenty-one hospitals that participated in at least 2 CAIRO studies, had a persisting low composite performance score (0-2 points). On the other hand, 6 hospitals (4 regional, 1 teaching, 1 university) had a persisting high performance score (3-4 points).

Variable		OR (95% CI)	OR (95% CI)
Hospital type		Unadjusted	Adjusted*
	University	reference	reference
	Teaching	0.55 (0.21-1.39)	0.70 (0.07-6.67)
	Regional	0.60 (0.24-1.51)	1.39 (0.10-19.4)

Number of beds		1.00 (1.00-1.00)	1.00 (1.00-1.00)
FTE oncologist		1.15 (1.01-1.31)	1.15 (0.85-1.56)
Hours of research nurse		1.01 (1.00-1.02)	1.00 (0.98-1.02)
Number of new metastatic colorectal patients per year	11-20 patients	reference	reference
	21-30 patients	0.73 (0.24-2.19)	0.62 (0.16-2.37)
	31-40 patients	1.84 (0.51-6.70)	2.23 (0.49-10.2)
	41-50 patients	2.11 (0.35-12.6)	2.07 (0.17-25.5)
	>50 patients	0.84 (0.21-3.44)	0.57 (0.06-5.23)
Number of included patients (<5 or ≥ 5 patients)		0.19 (0.09-0.40)	0.35 (0.08-1.55)

Table 2: Unadjusted and adjusted odds ratios (OR) with 95% Confidence Intervals (CI) for a higher (3-4 points) composite performance score (* adjusted for all variables listed).

Table 2 describes the univariable and multivariable analysis for factors that are potentially associated with a high composite performance score (3-4 points). The following were associated with the outcome: FTE oncologist (OR 1.15 per additional FTE; 95% CI 1.01-1.31), hours of research nurse (OR 1.01 per additional hour; 95% CI 1.00-1.02) and number of included patients (OR 0.19 for ≥ 5 patients included; 95% CI 0.09-0.40). However, none of these variables remained significant after adjusting for all covariates listed in Table 2.

None of the examined variables could be linked to the individual components of the composite score, except for low and high inclusion rates as expected, since hospitals with less than 5 inclusions had a higher mean overall composite score than hospitals including 5 patients or more. Odds ratios were 0.36 (95% CI 0.18-0.72) for SAE reporting, 0.43 (95% CI 0.19-0.99) for major protocol violations and 0.34 (0.16-0.69) for stratification errors, respectively.

The individualized patient composite scores in early included patients (first 3 inclusions) were significantly lower compared to subsequently included patients ($p<0.05$). This learning curve was observed in the top-10 ranking hospitals with highest inclusion rates.

Conclusions

We studied clinical trial performance in three large randomized clinical trials investigating treatment strategies in metastatic colorectal cancer. Based on a composite trial performance score including stratification errors, major protocol violations, SAE reporting and inclusion of ineligible patients, we identified a large variation in trial performance between hospitals. Inclusion rates were significantly higher in university and teaching hospitals compared to regional hospitals. However, we did not find a significant difference in trial performance between hospitals categories. There were no hospital-based factors identified (number of beds, FTE oncologist, and hours of research nurses available), that could explain differences between hospitals with high or lower trial performance. Interestingly, hospitals with a low inclusion rate had a higher mean composite score compared to hospitals with a high inclusion rate. In the top-10 hospitals with highest inclusion rates, we observed a learning curve for trial performance.

Our results are in line with earlier work by Begg et al.[11] They studied trial performance by rates of ineligibility, compliance with the protocol, and submission of data and concluded that the quality of participation of different types of hospitals was comparable. In agreement with their results, we observed no differences between types of hospitals using a composite endpoint that also included SAE reporting. However, we did identify a learning curve in hospitals with highest inclusion rates, an issue that was not addressed by Begg et al. We found a higher composite score for hospitals with a low accrual rate compared to hospitals with a high accrual rate. Although some studies showed lower performance scores for hospitals with low accrual rates [6], results of other studies did not show any correlation between accrual rates and clinical trial performance [11,16]. This latter observation implicates those hospitals with low accrual rates, which are often regional hospitals, should not be excluded and may be even encouraged to participate in clinical trials [12].

In our study, we evaluated and compared trial performance between (different categories of) hospitals with a newly introduced composite score, because no scoring system for trial performance exists. However, the usability and validity of this score needs to be determined in follow-up studies. We were not able to explain the variation in our composite score between hospitals, even though we included data from three large clinical trials. This implies that, without proper validation, our composite score should be used with caution for comparison between hospitals. However, a possible implication for our composite score could be evaluation of trial performance per individual centre over time. This is supported by our finding of a learning curve for individual hospitals with a high inclusion rate. An explanation for the learning curve could be an improved comprehension of and experience with the study protocol after inclusion of several patients. An important implication is to optimize support of participating hospitals during trials to prevent protocol deviations, with specific focus on the first patients included.

Clinical trials are essential for improvement of treatment possibilities. High quality within these trials is a requisite for reliable evidence-based outcomes. There is a current need for an evidence-based instrument to evaluate trial performance between hospitals, possibly with an adjusted version of our composite score included.

Evaluation and validation should be performed in different studies and research areas, to make it more widely applicable.

In conclusion, trial performance, evaluated with a composite score including stratification errors, major protocol violations, SAE reporting and the inclusion of ineligible patients, was comparable between different types of hospitals. In the highest including hospitals we identified a learning curve for trial performance. Consequently, we recommend additional support during the first patient inclusions in every participating centre, to prevent stratification errors, ineligibility and major protocol violations. For individually-based hospital evaluation, our composite score could be used as a quality metric for trial performance.

Acknowledgements

There are no acknowledgements.

References

1. Brenner H, Bouvier AM, Foschi R, Hackl M, Larsen IK, et al. (2012) Progress in colorectal cancer survival in Europe from the late 1980s to the early 21st century: the EUROCARE study. Int J Cancer 131: 1649-1658.
2. Kopetz S, Chang G, Overman J, Eng C, Sargent J, et al. (2009) Improved survival in metastatic colorectal cancer is associated with adoption of hepatic resection and improved chemotherapy. J Clin Oncol 27: 3677-3683.
3. Ottevanger P, Therasse P, van de Velde C, Bernier J, van Krieken H, et al. (2003) Quality assurance in clinical trials. Crit Rev Oncol Hematol 47: 213-235.
4. Wolf G, Makuch R (1980) Editorial: a classification system for protocol deviations in clinical trials. Cancer Clin Trials 3: 101-103.
5. Sweetman E, Doig G (2011) Failure to report protocol violations in clinical trials: a threat to internal validity? Cancer Clin Trials 12: 214.
6. Sylvester RJ, Pinedo HM, De Pauw M, Staquet MJ, Buyse ME, et al. (1981) Quality of institutional participation in multicenter clinical trials. N Engl J Med 305: 852-855.
7. Thomas P (1982) Quality of institutional participation in multicenter clinical trials [letter]. In: Novak JW 814.
8. Palva I (1982) Quality of institutional participation in multicenter clinical trials [letter]. N Engl J Med. 813-814.
9. Begg C (1982) Quality of institutional participation in multicenter clinical trials [letter]. N Engl J Med 813.
10. De Dombal F (1982) Quality of institutional participation in multicenter clinical trials [letter]. N Engl J Med. 813.
11. Begg C, Carbone P, Elson P, Zelen M (1982) Participation of community hospitals in clinical trials: analysis of five years of experience in the Eastern Cooperative Oncology Group. N Engl J Med 306: 1076-1080.
12. Koretz MM, Jackson PM, Torti FM, Carter SK (1983) A comparison of the quality of participation of community affiliates and that of universities in the Northern California Oncology Group. J Clin Oncol 1: 640-644.
13. Koopman M, Antonini NF, Douma J, Wals J, Honkoop AH, et al. (2007) Sequential versus combination chemotherapy with capecitabine, irinotecan, and oxaliplatin in advanced colorectal cancer (CAIRO): a phase III randomised controlled trial. Lancet 370: 135-142.
14. Tol J1, Koopman M, Cats A, Rodenburg CJ, Creemers GJ, et al. (2009) Chemotherapy, bevacizumab, and cetuximab in metastatic colorectal cancer. N Engl J Med 360: 563-572.
15. Simkens L, van Tinteren H, May A, ten Tije, A, Creemers G, et al. (2015) Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): a phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. Lancet 385: 1843-1852.
16. Pollock B (1994) Quality assurance for interventions in clinical trials. Multicenter data monitoring, data management, and analysis. Cancer 74: 2647-2652.