Eur J Cancer Clin Oncol, Vol. 25, No. 8, pp. 1241–1247, 1989.
Printed in Great Britain

0277–5379/89$3.00 + 0.00

Quality Control of Validity of Data Collected in Clinical Trials

K. VANTONGELEN,* N. ROTMENSZ† and E. VAN DER SCHUEREN‡
for the EORTC Study Group on Data Management (SGDM)

**EORTC Data Quality Control, UZ Leuven, Radiotherapy Department, Kapucijnenvoer* 33, 3000 *Leuven, Belgium*, †*SGDM, EORTC Data Center, Boulevard de Waterloo* 125, 1000 *Brussels, Belgium and* ‡*EORTC Quality Control Committee, UZ Leuven, Radiotherapy Department, Kapucijnenvoer* 33, 3000 *Leuven, Belgium*

Abstract—*In a study initiated by the EORTC Study Group on Data Management,* 15 *site visits to main participating centers in ongoing cancer clinical trials have been carried out over a* 1 *year period. The aim was to evaluate the quality level of EORTC clinical trial data, to find out the order of magnitude of possible problems encountered and to test a technique to objectively assess the quality of data. The process of data collection and the quality of data transfer from hospital charts to EORTC case report forms (CRF) were checked. The data quality was scored and the causes of incorrectness were evaluated. Percentages of correct data ranged from* 78% *up to* 98%; 11/15 *centers had* >90% *correct data. The median rate of error encountered in key data was* 2.8% (*range* 0.5–7%). *The main source of error was incorrect transfer of the information recorded in the patient chart to the CRF.*

Equally good overall results have been observed in the centers where data managers fill in the forms (DM) and those centers without an administrative trial structure (PH). The mean percentage of correct data for both types of centers is 91.4%. *The wider range in percentage for incorrect data (DM mean value* 3.0%, *range* 0.5–7%; *PH mean value* 2.3%, *range* 1.4–3.1) *suggests the important impact of the knowledge and experience of the people involved in data management. The data quality evaluation was hampered by the impossibility of checking part of the data present on the CRF,* 0.4–14.5%.

Besides knowledge and experience, the main aspects influencing good data quality appeared to be the efficacy of the internal organization and good local data monitoring. The importance of the design of CRFs was also highlighted. As this study was run for on-going protocols, the site visiting team had the opportunity to point out and report to the trial coordinator all shortcomings and controversial points that could thus be corrected during the course of the trial.

INTRODUCTION

THE 17 EORTC (European Organization for Research and Treatment of Cancer) Co-operative Groups are conducting around 200 clinical trials involving a large number of institutions spread all over Europe. Over the past years Co-operative Groups have gradually developed a number of quality control checks covering mainly diagnostic techniques, treatment administration and tumor response assessment [1–5].

There is no doubt that participation in cancer clinical trials entails an increased workload at the hospital and tends to complicate the already complex activities that arise when diverse disciplines are involved [6]. Moreover, in the central data center, the handling of data coming from many different institutions necessitates a rigid system of quality control in order to guarantee the uniformity, the completeness and correctness of the clinical trial data [7–11].

A few studies have analyzed this problem, some of them focusing on the quality of institutional participation, with, however, rather controversial results [12, 13]. Other studies concentrate on the quality of data management systems with emphasis on the need for specific training of administrative clinical research personnel and the benefits of a well functioning local data collection system [14–16].

In order to introduce a formalized procedure for assessing the accuracy of data collection and the efficacy of data transfer, the EORTC Study Group on Data Management has piloted a method for Data Quality Control in clinical trials. The program consists of site visits to the main contributors of on-going EORTC clinical trials during which the data sent to the Data Center are compared with the

Accepted 27 April 1989.

information available in the medical charts. At the local hospital level the process of data collection is observed, the quality of data transfer is evaluated and the causes for incorrectness or deviations are noted. In order to differentiate possible causes for poor data, a careful check on completeness and clearness of the selected protocols and corresponding forms is included in the program as well.

This article will focus on the procedure and the results achieved in the 15 centers visited.

METHOD

The study started on the basis of a written protocol submitted and accepted by an expert committee. Three large ongoing studies were selected: two Phase III breast cancer trials, one for adjuvant therapy and one for advanced disease, and one randomized Phase III study in advanced soft tissue sarcoma.

Computerized check lists were designed, according to each study, listing all items present on the CRFs. A set of CRFs is composed of an on-study form, treatment form(s) including data on treatment and evaluation of side-effects, a follow-up form and a summary form used for the patients leaving the trial. It was indeed the aim to verify, item by item, all the information collected on the original CRFs and sent to the Central Data Center for analysis.

Site visits were run in the five major participating institutions of the selected clinical trials. 'Major participants' were defined as institutions entering at least 20 patients in the trial. The files to be reviewed were randomly selected, 30 files per center in the adjuvant study and 20 files per study and per center, for the two advanced disease trials.

The visiting panel was composed of the trial study coordinator, a data manager of the Data Center and a data manager external to the visited center. At the local level, the presence of the physician(s) in charge of the study and the local data manager, if applicable, was requested. In 10/15 centers there was administrative help, functioning in different local data management structures. The five other centers had no administrative help available and the physicians were filling in the study forms themselves.

Prior to the site visit, the protocol under study was evaluated on the basis of a checklist, looking for completeness and clearness (Fig. 1).

The quality of data observed was doubly coded: Code A related to the type and the relevance of the errors:

1 = correct: data on the form fully corresponds with the information in the medical file
2 = deviation: minor discrepancy in more general data without direct impact on the trial outcome
3 = incorrect: discrepancy in key data with direct influence on the trial outcome
4 = missing: data are missing on the form
5 = data on the form, not in the file: impossibility of checking correctness of the data.

Code A5 was added after the first two site visits as it became clear that in some centers the information coded on the forms was not always available in the medical file.

In the majority of the centers, as far as side-effects documentation was concerned, we had to accept that the absence of comments in the files was to be considered as 'no toxicity observed', while very often it was quoted on the CRFs on the basis of nurse observation, orally transmitted and thus impossible to verify. Consequently, any coded toxicity on the CRF for which we could not find information in the patient file was coded 'data on form, not in file'. The results of this code have been counted against the correct percentage in order to preserve clean percentages of 'checked' correct and incorrect data. On the other hand, side-effects described in the patient file but not transferred on the CRF, are coded 'incorrect'.

Code B related to the causes for incorrect data:

1 = administrative mistake ('slip of the pen')
2 = information in file needing interpretation for transfer
3 = incorrect information in file
4 = information is missing in file
5 = incorrect transfer of information
9 = mistake related to ambiguous question on the CRF.

After each site visit a confidential report was sent to the trial study coordinator and to the local responsible physician. This report summarized the impression of the site visitors, including a general appreciation of the local organization. It included also the total number of items checked, the results detailed by type of form, and the overall results for all patients charts reviewed.

The visited centers were assigned a number in order to guarantee anonymous reporting.

RESULTS

Eleven out of the 15 centers visited have >90% correct data (Table 1). Although this percentage of correct data varies from 78% to 98%, it is to be noted that for some centers these results are mainly influenced by the impossibility of verifying data quality as we could not find in the charts the information reported on the CRFs. Code A5: 'data on form, not in file', range: 0.4 up to 14.5%. The evaluation of the results in this category is rather complex as it is partly caused by inaccuracy, partly by internal organizational aspects. Example: in one hospital a trial patient was treated at different

PROTOCOL CONTENTS CHECK

1. Protocol Number

	Clearly defined	not clear
3. Objectives	_____	_____
4. Eligibility criteria	_____	_____
5. Trial design	_____	_____
THERAPEUTIC REGIMEN		
6. Surgery	_____	_____
7. Radiotherapy	_____	_____
8. Chemotherapy	_____	_____
9. -dose description	_____	_____
10. -dose modification	_____	_____
11. -drug administration	_____	_____
12. -drug availability	_____	_____
- for new drugs,		
13. mode of drug conservation	_____	_____
14. TREATMENT TOXICITIES - are all toxic effects WHO defined ? YES - NO		
15. - if NO, reporting is	_____	_____
TREATMENT EVALUATION		
16. Laboratory tests	_____	_____
17. Radiological evaluation	_____	_____
18. Follow up	_____	_____
19. Response evaluation	_____	_____
ADMINISTRATIVE MATTERS		
20. Registration/randomization	_____	_____
21. Forms submission	_____	_____
22. Contact persons	_____	_____
23. Statistical considerations	_____	_____
24. Commitment for patients accrual	_____	_____

Comments :

Date : Signature :

Fig. 1. Protocol contents check.

departments without a central file system and insufficient exchange of information between the different physicians involved. We were told that part of the missing information was partly collected by telephone, partly by oral communication in order to enable form filling. The correctness of these data was impossible to verify during the site visit.

The same situation is created in cancer centers, where patients continue (part of) the protocol treatment at their local hospitals.

On the other hand, the category A5 also includes situations in which the physician codes the information directly on the form, in due time, but without making notes in the patient file. Although the percentage of reliability of this information is high, we applied the same rules for the 15 centers involved: data which could not be checked in the patient file are coded 'data on the form, not in the file'.

The percentage of missing data on the CRFs is rather low (0.3% up to 2.9%) and is partly caused by missing information in the medical chart, partly by incorrect transfer. Deviations were not often encountered (0.2–1.2%).

Special attention was paid to the percentage of incorrect data. The median rate of incorrect data is 2.8% (range 0.5–7%). The main cause of incorrect data, as reflected in Table 2, is the incorrect transfer of information from the medical chart to the CRF. Errors related to the ambiguity of the question on the CRF (code B9) were only observed in two centers for one specific trial, where a few questions on the forms were subject to different interpretations. The other causes, administrative mistake

Table 1. *Overall results*

Centers	101	102	103	104	105	201	202	203	204	205	301	302	303	304	305
DM structure	DM	DM	DM	PH	DM	DM	DM	DM	DM	PH	PH	DM	DM	PH	PH
Total number of data checked	1146	1202	745	1132	1448	2151	2107	1887	3255	2365	3511	4546	3040	3569	4929
Correct(%)	97.2	97.8	78.2	92.6	94	87	94.4	90	98	82.6	96.4	91.4	86	92.7	93
Deviations	0.7	0.5	1.2	1	1	1	1.2	0.7	0.8	0.4	0.4	0.2	0.5	0.8	—
Incorrect(%)	1.5	1.2	3.3	2.5	3	4.2	2.5	7	0.5	3.1	1.4	3.7	3.5	2.6	2.3
Missing(%)	0.5	0.3	2.8	0.3	1.5	0.6	1.1	1.2	0.3	2.9	0.2	1	2	1.2	0.2
Not in file(%)	—	—	14.5	3.6	0.5	7.2	0.8	1.1	0.4	11	1.6	3.7	8	2.7	4.5

Table 2. *Causes of 'incorrect data'*

Centers	101	102	103	104	105	201	202	203	204	205	301	302	303	304	305
No. of incorrect data/ total number of data checked	17/1146	15/1202	25/745	28/1132	43/1448	91/2151	54/2107	132/1887	16/3255	75/2365	48/3511	170/4546	104/3040	94/3569	112/4929
Administrative mistake	1	1	1	1	4	5	—	2	—	2	—	3	4	8	1
Information needing interpretation	2	1	1	—	5	1	5	2	—	2	5	3	4	2	1
Incorrect information in file	—	—	1	—	—	—	—	—	—	—	1	—	1	—	—
Information missing in file	1	1	—	1	—	4	1	—	—	9	—	—	—	8	—
Incorrect transfer	13	12	22	8	28	81	48	128	16	62	42	164	95	76	110
Form related	—	—	—	18	6	—	—	—	—	—	—	—	—	—	—

and information needing interpretation, were responsible for very few errors. The cause 'incorrect information in medical chart' was difficult to evaluate in general. The few cases reported were evident mistakes, admitted by the local physician.

We analyzed separately the results obtained in the two types of centers included in the site visits: centers with some kind of data management structure available (DM) (101, 102, 103, 105, 201, 202, 203, 204, 302, 303) and centers without administrative support (PH) (104, 205, 301, 304, 305). Exactly the same median rate of correct data (91.4%) was observed in both groups. In (DM) centers the difference ranged from 78.2 up to 98%, in the (PH) centers from 82.6 up to 96.4%. The percentages of 'incorrect' data reflected almost the same balance: mean values 3% (DM) and 2.3% (PH) ranging from 0.5 up to 7% (DM) and 1.4 up to 3.1% (PH). The evaluation of 'data on form, not in file' also resulted in very small differences: median percentage 4.5 (DM), range 0.4–14.5; against 4.6 (PH), range 1.6–11%. The extreme values obtained in both groups can only be explained by further analysis and discussion of the nominal aspects encountered.

DISCUSSION

In addition to the figures shown in Tables 1 and 2, we found that other causes, which had not been expected beforehand, were influencing the rate of errors.

1. *Structures of local data management*

The visited centers are all major participants in EORTC clinical trials and the majority have administrative support in a more or less organized structure. In 10/15 centers the CRFs were partly (centers 302, 303) or completely filled in by data managers (centers 101, 102, 103, 105, 201, 202, 203, 204). In the five remaining centers the treating physician was also the administrator of the study data (centers 104, 205, 301, 304, 305). The data managers had different types of paramedical background (nurses, medical secretaries, physical therapists) and different levels of experience. From Table 1 we cannot conclude that the quality of data is *a priori* better in one or another specific structure, although the imbalance (10 against 5) hampers a statistically significant evaluation. The best results, however, were observed in centers 204, 102 and 101, with well trained and experienced data managers involved (>10, 3 and 4 years). In these centers, the good quality of data was influenced by the knowledge and training of the local data manager but also by the kind of structure in which the data manager was functioning. In these centers, the organization of data collection and trial monitoring in general, was efficiently organized.

Prior to the first site visits (101 and 102), it was indeed not realized that poor data collection in hospital files would hamper the evaluation possibilities and influence the complexity of the results. Based on the site visit findings in the third center (103) we decided to include a code for data which we could not verify (code A5).

In connection with our visit to this center, one very specific structure of data management will be highlighted, the situation of the so-called 'gipsy' data manager [17]. These are data managers travelling from the Comprehensive Cancer Centers (CCC) where they are based, to the different centers which can apply for their services. Gipsy data managers are usually involved in many different trials with few patients in every trial and their work is divided between many and usually small hospitals. Multiple consequences result from this situation. For instance in one CCC, one data manager is involved in 43 clinical studies with 160 patients/year in total. It is almost impossible to know all the pitfalls, backgrounds and protocol interpretations necessary for the good understanding of the specific study questions. Moreover, the administrative help, in this kind of structure, is mainly concentrated on the completion of CRFs. Gipsy data managers are filling in the CRFs in the different hospitals but have no direct impact on the system of local trial organization.

The results obtained in this center should be discussed in view of these circumstances. First it is to be noted that only 745 data have been checked, which is a markedly reduced number, compared with centers 101, 102, 104, 105. This was due to missing files (10/30 indicated on the list) referring to 'unknown' trial patients, for which no CRFs had been sent to the Data Center. The high percentage of data 'on form, not in file' (14.5%) was influenced by the poor local data collection and the complete lack of data monitoring. At the time the data manager visited the center for completion of the CRFs, she was confronted with poor documentation in the patient charts and limited possibilities for reconstruction. Amongst the many other practical problems related to the system, a very important one is the restricted possibility of communication between the data manager and the physicians in so many different hospitals.

Results obtained in centers without administrative support (104, 205, 201, 304 and 305) are mainly influenced by the quality of the patient file and the time at which the physician completes the forms. Overdue forms, up to 12 months may have elapsed since the protocol treatment was given, filled in 'from memory' by the physician, did have a negative influence on the overall quality of data obtained (205).

Whatever the local data management structure available, the link between the administrative help and the clinical trial activities is a priority. Moreover, the data manager should be guided by the physician and be trained to the level of complete reliability. Good communication possibilities between physicians and data managers are the best guarantee to achieve this goal [18].

2. *Data collection and completeness of the hospital file*

Besides the survival and time to progression, the main endpoints in Phase II and III cancer clinical trials are tumor response and toxicity. In EORTC trials, most of the time, protocols and forms refer to the WHO standard grading systems to guarantee a uniform collection of these data [19]. In view of the importance of these evaluations, the information should be 'primary-source information', avoiding interpretation of descriptive notes in the patient file by the data manager, for transfer on the trial forms. In general, only few centers have complete information on treatment toxicity in the patient file, including evaluation of side-effects which are present and absent.

In one center (204), where data managers transfer the information from the patient file onto the study forms, all possible side-effects that could occur were directly coded by the physician and noted in the patient file, thus enhancing the reliability of the data collected. In other centers, however, information on toxicity is partly or even scarcely documented, notes from nursing papers being the only reliable information available which in turn need interpretation by the data manager for transfer to the CRF (103, 201, 203, 205, 302). Mistakes and missing information are unavoidable in this situation.

The standard evaluation of the patient's performance status was perhaps the most poorly documented characteristic in the notes, and yet it is one that does have a major bearing on prognosis. During the site visit we discussed the need to train the physicians in using the referred grading systems as this is the only way to obtain complete and correct documentation in the patient files. The recording of all events, occurring during the protocol treatment, is essential and should be emphasized.

3. *Different interpretations of the WHO toxicity scale and general protocol terms*

Although the availability of toxicity data is a question of well organized local data collection and well trained physicians, there are major differences in toxicity grading between centers. One of the most important differences to emerge was the widely inconsistent toxicity grading on nausea and vomiting. This is illustrated in Table 3.

The first part of the table represents the WHO grading for nausea and vomiting. The second part summarizes the different interpretations by the first three hospitals included in the pilot study. All

patients were treated in the same clinical trial and received prophylactic antiemetics prior to the perioperative adjuvant CAF (cyclophosphamide, Adriamycin® and 5-FU), administered within 24 h after surgery. Each center had consistently interpreted the WHO grading but in a different way. More consistency in the use of grading scales would positively influence the reliability of the trial results. Setting up a standard policy on the grading of nausea and vomiting was suggested, differentiating the patients between those who received prophylactic antiemetics and the others.

Clarification is also required on different terms used in the protocols. It came out that well known items, such as date of diagnosis, dominant site of disease and response to treatment caused confusion in the different centers. Their interpretations were widely inconsistent and may necessitate a clarification manual that could be part of the protocol.

4. *Treatment protocol and corresponding forms*

Notwithstanding the overall high standard of the quality of protocols and the forms involved, some shortcomings have been observed. The lack of complete and clear documentation in the protocol and the presence of ambiguous questions on the forms have influenced the quality of some data obtained. For instance in a breast cancer protocol, the CRF collected information on the tumor staging:

0 = *tumor not palpable*
1 = <2 cm
2 = 2–5 cm
3 = >5 cm
4 = *any size + locally advanced*
9 = unknown.

There was no indication, either on the form or in the protocol, that the information was meant to be either clinical or pathological. The content of the codes was found confusing and incorrect, specially as it became clear that it was the intention to collect pathological staging information, which is clearly in contradiction with codes 0 and 4. These kinds of shortcomings, revealed during the site visits, have been reported to the trial coordinator and corrected during the course of the trial.

CONCLUSION

The question of data quality and reliability is recurrently raised when multicenter clinical trials are evaluated. It was therefore decided to test a method intended to highlight possible problems and to solve them during the course of the study. Rather than proper data management, our results are showing that 'knowledge' is the key for good data reports. It means that we have not succeeded in showing that data management support is a guarantee for correctness in data transfer. Actually, equally good findings have been observed in all types of structure, the main issue being good communication, internally as well as with the central office.

Good local organization with tight internal control will further enhance the reliability of the data coming from the different collaborating centers. If the hospital has a local data management structure available, the data managers should focus their attention on the organization of accurate data collection. Therefore they should give all necessary administrative support to the physicians and should not restrict themselves to form filling alone. Prospective reminders, protocol summaries and notes in the patient's charts, detailing what must be done, will markedly improve the quality of data collection. In order to function well, they need adequate training and full support by the physicians.

For the different people involved in the quality control exercise we may conclude that the occasion

Table 3. WHO grading system for nausea/vomiting

Grade 0	Grade 1	Grade 2	Grade 3	Grade 4
None	Nausea	Transient vomiting	Vomiting requiring therapy	Intractable vomiting
Institute 1		Institute 2		Institute 3
Grade 3 = all patients received premedication		Grade 1 = nausea + accept. vomiting Grade 2 = vomiting for several days requiring additional medication Grade 3 = never used		Grade 0 = premedication + some nausea

and incentive to sit together, discuss and evaluate the system of data collection and internal communication, was positively welcomed by all the groups involved.

Centers with poor quality will be revisited after one year. At that time they will have to prove the positive impact of their local re-organization.

Acknowledgements—We are indebted to the following study coordinators, participants in the Quality Control Study whose contribution made this paper possible: L. Beex (Nijmegen), J.P. Julien (Rouen), M. Nooij (Leiden), E. Rankin (Glasgow), O. Repelaer van Driel (Leiden), A. Santoro (Milan), C.H. van de Velde (Leiden), J. Verweij (Rotterdam).

The Quality Control work is supported by the EORTC through a grant from the EEC.

REFERENCES

1. Suntharalingham N, Johansson KA. Quality assurance/physics/dosimetry. *Int J Radiat Oncol Biol Phys* 1988, **14**, S21–S24.
2. Horiot JC, Johansson KA, Gonzalez DG, van der Schueren E, van den Bogaert W, Notter G. Quality assurance control in the EORTC cooperative of radiotherapy. 1. Assessment of radiotherapy staff and equipment. *Radiother Oncol* 1986, **6**, 275–284.
3. Johansson KA, Hanson WF, Horiot JC. Workshop of the EORTC Radiotherapy Group on quality assurance in cooperative trials of radiotherapy: a recommendation for EORTC Cooperative Group. *Radiother Oncol* 1988, **11**, 201–203.
4. Johansson KA, Horiot JC, van der Schueren E. Quality assurance control in the EORTC cooperative group of radiotherapy. 3. Intercomparison in an anatomical phantom. *Radiother Oncol* 1987, **9**, 289–298.
5. Johansson KA, Horiot JC, van Dam J, Lepinoy D, Sentenac I, Sernbo G. Quality assurance control in the EORTC cooperative group of radiotherapy. 2. Dosimetric intercomparison. *Radiother Oncol* 1986, **7**, 269–279.
6. Dale Williams O. A framework for the quality assurance of clinical data. *Clin Pharmacol Ther* 1979, **25**, 700–702.
7. Buyse M. Quality control in multi-centre cancer clinical trials. In: Buyse ME, Staquet MJ, Sylvester RJ, eds. *Cancer Clinical Trials, Methods and Practice*. Oxford, Oxford University Press, 1984, 102–123.
8. Knatterud GL. Methods of quality control and of continuous audit procedures for controlled clinical trials. *Controlled Clin Trials* 1981, **1**, 327–332.
9. Canner PL, Krol WF, Forman SA. External quality control programs. *Controlled Clin Trials* 1983, **4**, 441–466.
10. Glicksman AS, Reinstein LE, Brotman R, McShan D. Quality assurance programs in clinical trials. *Cancer Treat Rep* 1980, **64**, 425–433.
11. Renard J, Van Glabbeke M. Quality control and data handling. In: Rotmensz N, Vantongelen K, Renard J, eds. *Data Management and Clinical Trials*. Amsterdam, Elsevier, 1989, 147–162.
12. Sylvester RJ, Pinedo HM, De Pauw M *et al.* Quality of institutional participation in multicenter clinical trials. *N Engl J Med* 1981, **305**, 852–855.
13. Eastern Cooperative Oncology Group. Participation of community hospitals in clinical trials: analysis of five years of experience in the Eastern Cooperative Oncology Group. *N Engl J Med* 1982, **306**, 1076–1080.
14. O'Fallon JR, Golenzer HJ, Taylor WF *et al.* Mayo Clinic. ASA 1980, Proceedings of the Statistical Computing Section, Washington.
15. vd Putten E, vd Velden JW, Siers A, Hamersma EAM for the Cooperative Study Group of Dutch Data Managers. *Controlled Clin Trials* 1987, **8**, 96–100.
16. Boyd NF, Pater JL, Ginsburg AD, Myers RE. Observer variation in the classification of information from medical record. *J Chron Dis* 1979, **32**, 327–332.
17. Konings M. Multicentric data management, the data gipsy. Proceedings of the Third Annual Meeting of the EORTC Study Group on Data Management. 1987.
18. Rotmensz N, Aaronson NK. The role of data managers in cancer clinical trials. Results of an EORTC survey. Proceedings of the Second Annual Meeting of the EORTC Study Group on Data Management, 1986.
19. *WHO Handbook for Reporting Results of Cancer Treatment*. WHO offset publication No. 48, Geneva, WHO, 1979.