PERGAMON

European Journal of Cancer 36 (2000) 615–620

European Journal of Cancer

www.ejconline.com

Clinical impact of dosimetry quality assurance programmes assessed by radiobiological modelling of data from the thermoluminescent dosimetry study of the European Organization for Research and Treatment of Cancer

S.M. Bentzen [a,*], J. Bernier [b], J.B. Davis [c], J.C. Horiot [d], G. Garavaglia [e], J. Chavaudra [f], K.A. Johansson [g], M. Bolla [h]

[a] *Gray Laboratory Cancer Research Trust and the Cancer Centre, PO Box 100, Mount Vernon Hospital, Northwood, Middlesex HA6 2JR, UK*
[b] *Department of Radio-Oncology, San Giovanni Hospital, Bellinzona, Switzerland*
[c] *Department of Radiation Physics, Universitätsspital, Zurich, Switzerland*
[d] *Department of Radiation Oncology, Centre G.F. Leclerc, Dijon, France*
[e] *Department of Radiation Oncology, The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands*
[f] *Department of Radiation Physics, Institut Gustave Roussy, Villejuif, France*
[g] *Department of Radiation Physics, Sahlgren University Hospital, Gothenburg, Sweden*
[h] *Department of Radio-Oncology, Hospital Universitaire, Grenoble, France*

Received 3 August 1999; received in revised form 2 November 1999; accepted 8 December 1999

Abstract

The European Organization for Research and Treatment of Cancer (EORTC) Radiotherapy Group initiated its mailed thermoluminescence dosimetry (TLD) programme in 1986. The aim of the present study was to evaluate the clinical relevance of variations in beam output detected in the period 1993 to 1996. A total of 140 beam outputs were checked (26 for cobalt-60 units and 114 for linear accelerators) in 35 centres. Clinical dose–response data for tumour control and normal tissue morbidity were used to assess the variation in clinical outcome resulting from variability in beam output. For 75 checked beams with nominal accelerating potentials (n.a.p.) of 6 MV or less the mean ratio, ± standard deviation (S.D.) of measured to stated output was 1.004 ± 0.020. For 65 beams with n.a.p. of 8 MV or more, the ratio was 1.009 ± 0.021. Even with this relatively high level of precision, broad distributions of estimated tumour control or normal tissue morbidity were found. In the 10% of the beams with the most pronounced underdosage, the loss in tumour control probability was estimated at 7–8 percentage points. Likewise, in the 10% of the beams with the most pronounced overdosage, the increase in mild/moderate morbidity was 19–22 percentage points. For severe morbidity the same beams raised the estimated incidence of severe complications from 5% to 9–10%. An estimation of the loss of uncomplicated cure probability was about 1% for both high and low energy beams. Sequential mailings considerably improved the uniformity of clinical outcome. We conclude that small deviations in beam output may lead to clinically important variations in outcome. Substantial reductions in the variation between measured and stated output can be achieved by sequential mailings. Mailed TLD checks should be an integral part of a continuously ongoing quality assurance activity in radiotherapy.

Keywords: Radiotherapy; Quality assurance; Dosimetry; Tumour control; Normal tissue morbidity

1. Introduction

Quality assurance (QA) is the systematic effort to monitor performance, to compare its quality to a defined standard, and to implement corrective measures if the performance does not meet the requirements within a certain margin. There is a growing interest in QA in medicine in general and one of the fields where this has been pioneered is radiation oncology. Several QA programmes conducted within single centres or by cooperative clinical research groups have demonstrated their value at the operational level: deviations from performance specifications have been detected, documented and corrected. Yet, the clinical effect of QA

* Corresponding author. Tel.: +44-1923-828-611; fax: +44-1923-835-210.
E-mail address: bentzen@graylab.ac.uk (S.M. Bentzen).

0959-8049/00/$ - see front matter
PII: S0959-8049(99)00336-6

programmes on treatment outcome is difficult to assess directly. The reason is, that many instances of substandard performance cause an expected small, although maybe clinically relevant, reduction in treatment efficacy and/or increase in toxicity. In theory, such differences could be studied in randomised controlled trials; in reality, this is not an option because once a substandard performance is identified it is unethical not to take corrective action.

Since 1987, the Radiotherapy Group of the European Organization for Research and Treatment of Cancer (EORTC) has conducted a mailed thermoluminescent (TL) dosimetry (TLD) programme checking external beam therapy machine output on the beam axis. The ratio between the measured absorbed dose to water and that stated by the radiotherapy department was used as a measure of agreement. The deviations, ΔD, between actual and stated dose, were classified as acceptable ($\Delta D \leqslant 4\%$), minor ($5\% < \Delta D \leq 7\%$) or major ($\Delta D > 7\%$).

In the EORTC institutions, the first independent checks of therapy machine output were conducted in the late 1980s [1]. An early report [2], published in 1993, covered the first 6-year period (1987–1992): 55 institutions were monitored, 127 mailings were sent to radiotherapy departments and a total of 357 beams were measured. This first report showed that the level of deviation between the measured and stated dose was acceptable (within $\pm 4\%$) for 90% of the measured beams. The 714 measurements carried out on 357 beams yielded a mean ratio of 1.007 and a standard deviation (S.D.) of 4%. Only 9 beams out of 357 (2.5%) exceeded a 7% difference between measured and stated dose. Subsequent site visits and extensive correspondence were activated and maintained for several months between the EORTC experts and each of the institutions where major deviations had been found. In a second mailing, a deviation larger than 7% was observed in only one of these centres.

Results of this programme have been presented in terms of the distribution of deviations seen. There is, however, a growing body of clinical data that allows realistic estimates of dose–response relationships to be made [3,4]. The aim of the present study was to assess the clinical impact of the measured dose deviations on local tumour control and late normal tissue reactions, and to quantify the therapeutic gain that may result from corrective actions taken in the context of long-term QA programmes.

2. Patients and methods

2.1. TL dosimetry

The methodology and feasibility of TLD programmes have been previously described [5–7]. TLD is suited for periodical monitoring of absorbed doses at specified beam reference points. The precision of the TL dosimeters is sufficient to identify errors in absorbed dose above 3% for photon beams of Cobalt-60 machines or linear accelerators with nominal accelerating potentials (n.a.p.) up to 30 MV.

The EORTC mailed TLD programme was organised as follows. After the centres had provided the EORTC panel with relevant information on the beam qualities of their radiotherapy equipment, they were sent a set of dosimeters, containers for in-water irradiation and instructions for irradiation. After irradiation, all the material was returned to the reference centre for read-out. From 1986 until 1992 the reference centre was in the radiation physics department of the Sahlgren University Hospital in Göteborg, Sweden; thereafter it was in the radiation physics department of the Gustave Roussy Institute, in Villejuif, France.

It was decided to base the current study on the TLD results in the period 1993 to 1996 in order to make them more representative for what can be achieved after a QA programme has been in place for several years.

2.2. Radiobiological modelling

Evaluation of the clinical consequences of a given deviation between stated and actual dose is based on published clinical data on the steepness of dose–response curves for human tumours and normal tissues. Steepness was quantified by the normalised dose–response gradient, γ, which may be thought of as the change in response (in percentage points) for a 1% relative change in dose [3,8]. Thus, as an example, for $\gamma = 3$, a 4% underdosage would lead to an estimated 12% decrease in the incidence of a specific endpoint. This simple multiplicative conversion from relative change in dose into resulting change in outcome, is only a good approximation in a relatively narrow interval around a baseline dose. The reason is the S-shape of the dose–response curve, which means that the local value of γ is not constant as a function of dose. More precise estimates of incidence were obtained using the exact parameterisations of dose–response models discussed by Bentzen and Tucker [9].

Table 1 shows the dose–response parameters used. The assumed level of normal-tissue complication prob-

Table 1
Incidence and steepness parameters assumed in the radiobiological modelling

Endpoint	Incidence (%)	γ_N
Tumour control	37	2.3
Mild/moderate reactions[a]	40	5.2
Severe complications[a]	5	0.9

[a] $\gamma_N = 5.5$ at normal tissue complication probability (NTCP) = 0.5.

ability (NTCP) and tumour control probability (TCP) were chosen arbitrarily, but they are representative for standard curative radiotherapy schedules. The value for mild/moderate reactions was derived from a study of moderate subcutaneous (s.c.) fibrosis after post-mastectomy radiotherapy [10]. The γ value for severe reactions is the resulting value at NTCP equal to 5% for a dose–response curve with the same steepness as for mild/moderate reactions. These calculations were performed assuming a logistic dose–response relationship. For tumour control, the γ value was derived from the review by Bentzen [3] who found that 2.0 was a representative value for dose–response curves generated with a fixed dose per fraction for various tumour sites in the head and neck. Assuming that the α/β-ratio of the linear-quadratic model is 10 Gy, this corresponds to a γ value of 2.3 for a dose–response curve generated with a fixed *number* of fractions. In view of the results presented below, it should be noted that the values used here for the steepness of the dose–response curves are not extreme. Some normal tissue endpoints have steeper dose–response curves than s.c. fibrosis and vocal cord tumours have consistently been found to have steeper dose–response curves than other sites in the head and neck.

Following the 1936 paper by Holthusen [12], the probability of uncomplicated cure, $P+$, was used as a composite measure of the trade off between tumour control and normal tissue reactions in over- or under-dosing beams. To this end, normal tissue complications and tumour recurrence were assumed to be statistically independent, i.e. the relevant probabilities were multiplied to yield the probability of a joint event. Recent analyses of clinical data provide support for this assumption [13,14]. The γ values at the steepest point of the dose–response curves were still as in Table 1 but the dose for NTCP = 50% was chosen at 75 Gy in order to simulate an optimisation of $P+$ for more severe complications.

3. Results

3.1. TL-dosimetry results

The data set used for the present analysis were collected in the mailed dosimetry programme during the years 1993 to 1996. A total of 140 beam outputs, were checked (26 for cobalt-60 units and 114 for linear accelerators) in 35 centres. For the purpose of this analysis the beams were grouped according to n.a.p., the low-energy beams being defined as having an n.a.p. of 6 MV or less and the high energy beams defined as n.a.p. of 8 MV or more. Table 2 shows some descriptive statistics on the ratio of EORTC measured and institution-stated doses. The average was close to one and the S.D. was around 2%. The 90 percentile of the distribution of dose ratios was used to characterise the extreme tail of the distribution (i.e. the 10% of the beams with the largest overdosage) and this was around 3%. These fairly good results were also reflected using the classification of previous reports: the number of beams falling within the acceptable levels of variation (0.960–1.040) was 130 (93%), whilst minor deviations (5–7%) were found for 10 beams only (7%). No major deviation, defined as $\Delta D > 7\%$, was found.

Table 2
Results of the EORTC mailed TLD programme 1993–1996

Beam energies	No. of beams	Average ratio	Standard deviation	90 percentile
6 MV or less	75	1.004	0.020	1.025
8 MV or more	65	1.009	0.021	1.034

3.2. Clinical consequences

Figs. 1 and 2 show the cumulative distribution of changes in TCP and NTCP due to unintended beam output variations for low and high energy megavoltage beams, respectively. For tumour control, approximately 35% of the beams were found to underdose the tumours and, therefore, were expected to yield an inferior tumour

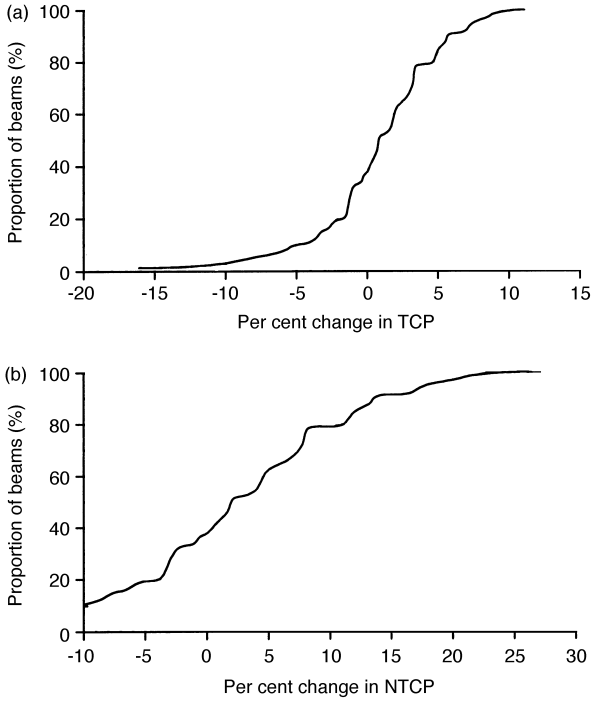

Fig. 1. Cumulative distribution of the proportion of beams with dosimetric deviations corresponding to various estimated changes in (a) tumour control probability (TCP) and (b) normal tissue complication probability (NTCP) for beams of 6 MV or less.

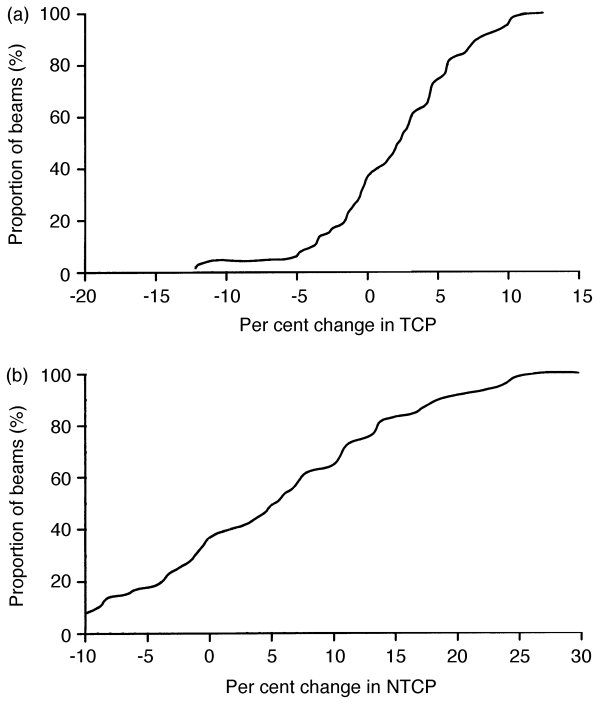

Fig. 2. Cumulative distribution of the proportion of beams with dosimetric deviations corresponding to various estimated changes in (a) tumour control probability (TCP) and (b) normal tissue complication probability (NTCP) for beams of 8 MV or more.

control probability. In the 10% of the beams with the most pronounced underdosage, the average loss in tumour control probability was estimated at 8 or 7% for the low and high energy beams, respectively.

Around two thirds of the beams gave a ratio of more than one between the TL and the stated dose (Figs. 1b and 2b). This resulted in an increase in the average NTCP for all beams of 2 and 5 percentage points for the low and the high energy beams, respectively. However, this average resulted from a widespread range of actual NTCPs. The 10% worst beams in terms of overdosage gave on average an estimated 19 or 22 percentage points increase in the incidence of mild or moderate side-effects for the low or high energy beams, respectively. The same beams were estimated to raise the probability of severe complications from 5% to an average of 9% or 10%. Thus, the frequency of patients with severe side-effects would roughly be doubled for patients treated on those machines.

Optimisation of the probability of uncomplicated cure, $P+$, was done numerically. For the values assumed here (see Materials and Methods section) the optimal prescribed dose was 68.1 Gy. At this dose the TCP was 77.6% and the NTCP was 11.7%. The local values of γ were 1.8 and 2.1 for the local control and the morbidity dose–response curves, respectively. The average change in $P+$ was estimated at 0.2% loss for the low energy beams and 0.3% loss for the high energy beams. The average loss for the worst 10% of the beams was 0.9% and 1.2% for the low and high energy beams, respectively.

3.3. The effect of repeat mailings

Table 3 summarises the distribution of the ratios between measured and stated doses, observed since the activation of the EORTC QA programme in radiation physics, both for on site visits (1982–1986) and in the framework of the mailed TLD programme (1987–1992 and 1993–1996). To elucidate the clinical impact of these repeated dosimetry checks, we defined considerable overdosage as a beam giving a more than 15 percentage points increase in NTCP for mild or moderate reactions. Using this definition, the proportion of beams with considerable overdosage dropped from 40–50% to less than 10% during the three checks.

4. Discussion

Dose–response relationships in radiotherapy are in many curative indications quite steep and even small variations in dose will have a relevant influence on clinical outcome. Whilst a change in cure probability of a few per cent may seem modest, it is important to see this on a scale of other means to improve cancer treatment outcome. An example could be combined chemo-radiotherapy where a meta-analysis of head and neck

Table 3
Effect of repeated dosimetry checks

Mailing	Linear accelerators		^{60}Co	
	Ratio of measured to stated dose[a]	Beams with considerable overdosage[b] (%)	Ratio of measured to stated dose[a]	Beams with considerable overdosage[b] (%)
First	1.022 ± 0.024	41	1.025 ± 0.027	47
Second	1.013 ± 0.017	20	1.006 ± 0.014	6
Third	1.007 ± 0.013	6	0.994 ± 0.004	0

[a] Average ± 1 standard deviation. Data from Horiot and colleagues [18].
[b] Estimated percentage of beams giving a more than 15% increased probability of mild/moderate side-effects.

trials suggest a four percentage point improvement in survival. Similarly, an increase in the incidence of severe complications of just a few per cent may represent a major burden, not just to the individual patient but also to the National Health Service. What does complicate matters is the fact that a beam that overdoses patients will improve tumour control at the same time as giving an increased risk of side-effects. Yet, if this increased incidence of treatment-related morbidity is justifiable because of the improved tumour control probability, the only reasonable conclusion is that the prescribed total dose should be increased. Obviously, patients who get unintentionally high/low doses are receiving sub-standard radiotherapy.

The trade off between tumour control and treatment-related morbidity is a very complex issue. The probability of uncomplicated cure, $P+$, should be interpreted with some caution. An increased risk of even fairly severe complications may be acceptable in situations where the alternative is an increased risk of a possibly fatal locoregional recurrence. However, if $P+$ is actually at its optimum, as shown here, any variability in beam output will be associated with a loss of uncomplicated cure probability.

Direct comparison of clinical outcomes in patients treated on machines that over- or underdosage is hampered by a multitude of patient and treatment factors that vary amongst centres or even from one machine to the other. However, Ash and Bates [15] showed that tumour control was notably lower than expected for several tumour sites in 1045 patients who received 5–20% lower-than-prescribed doses of radiation because of an error in the application of a computerised dose-planning system. Also, studies applying individual patient dosimetry have found a significant relationship between small individual deviations from the prescribed dose and the resulting normal tissue reaction [16].

Obviously, some of the variability in the ratio of measured to stated doses is attributable to uncertainty in the TL-dosimetry. Hansson and Johansson [1] determined the reproducibility of their TL dose readings to be 0.5% (1 S.D.). They estimated the combined uncertainty of the determined absorbed dose to be approximately 2%. This figure included uncertainties arising from the exact irradiation conditions, energy dependence and the reading procedure, fading and supralinearity of the TL dosimeters. Actual data from the later rounds of repeat mailings in the TLD QA project suggest that the 2% may be an overestimate, the true uncertainty being closer to 1% (1 S.D.). In the latter case, most of the variation in readings will be due to true variation in machine output with a, say, 2% measured S.D. of the ratios being equivalent to a real S.D. of 1.73%. This again means that most of the spread analysed in this paper will actually affect clinical outcome as predicted in the modelling.

Several authors have considered the general issue of precision requirements in radiotherapy and this includes consideration of non-uniform dose distribution in individual patients and variation amongst patients due to a multitude of factors. The present study focuses on variation in beam output that will affect the whole population of patients treated on a specific machine. As shown here TL dosimetry programmes are effective in reducing the magnitude of such variations.

Most progress in the field of quality assurance has been achieved by cooperative clinical research groups [17]. Originally, the QA programme of the Radiotherapy Group of the EORTC was designed to make simple checks of both the compliance to protocol guidelines and the consistency of data recorded and reported by a large number of active centres. However, this was soon extended to cover the whole chain of events, ranging from the detection and score of any variation from performance specifications in the field of radiation physics, to the tracking of random and systematic errors in the clinic. As a consequence, the programme was, therefore, redirected towards the prospective implementation of measures to prevent, reduce or reverse the risk of deviations. In the TL programme, in most cases where deviations exceeded 7%, site visits and correspondence between EORTC and the monitored centres demonstrated that these large deviations were due either to errors in applying the dosimetry protocol or to the application of an erroneous calibration factor of the ionisation chambers.

Hansson and Johansson [1] showed that the results of the EORTC mailed TLD programme were somewhat better than those reported by the IAEA and similar to the data recorded in North America. Thus, the variations in clinical outcome estimated here will be fairly representative for these regions as well. In parallel with screening for cancer, the first screenings or quality checks would give a much larger expected improvement in outcome. The results from the sequential mailing programme illustrate the effectiveness of such programmes in improving the dosimetric accuracy in the participating institutions. In the present study, it was decided to use data from the late stage of a TLD QA programme. The estimated clinical consequences of beam output variations support the continued effort to do mailed TL dosimetry checks as part of a routine QA programme.

Acknowledgements

The TL dosimetry programme was supported in the frame of the 4th Medical and Health Research Programme, concerted action 1989–1992 of the DG XII of the European Union on the theme Quality Assurance in Clinical Cancer Research.

References

1. Hansson U, Johansson KA. Quality audit of radiotherapy with EORTC mailed in water TL-dosimetry. *Radiother Oncol* 1991, **20**, 191–196.
2. Hansson U, Johansson KA, Horiot JC, Bernier J. Mailed TL dosimetry programme for machine output check and clinical application in the EORTC radiotherapy group. *Radiother Oncol* 1993, **29**, 85–90.
3. Bentzen SM. Radiobiological considerations in the design of clinical trials. *Radiother Oncol* 1994, **32**, 1–11.
4. Bentzen SM. Quantitative clinical radiobiology. *Acta Oncol* 1993, **32**, 259–275.
5. Boyd AW, Eisenlohr HH. IAEA/WHO 60Co teletherapy dosimetry service using mailed LiF dosimeters. A survey of results obtained during 1975–1982. *Med Phys* 1983, **10**, 491–492.
6. Eisenlohr HH, Jayaraman S. IAEA-WHO cobalt-60 teletherapy dosimetry service using mailed LiF dosemeters. A survey of results obtaining during 1970–75. *Phys Med Biol* 1977, **22**, 18–28.
7. Pfalzner PM, Malo AS. Intercomparison of absorbed dose in cobalt 60 teletherapy using mailed LiF dosimeters. *Acta Radiol Ther Phys Biol* 1968, **7**, 379–388.
8. Brahme A. Dosimetric precision requirements in radiation therapy. *Acta Radiol Oncol* 1984, **23**, 379–391.
9. Bentzen SM, Tucker SL. Quantifying the steepness of dose–response curves. *Int J Radiat Biol* 1997, **71**, 531–542.
10. Bentzen SM, Thames HD, Overgaard M. Latent-time estimation for late cutaneous and subcutaneous radiation reactions in a single-follow-up clinical study. *Radiother Oncol* 1989, **15**, 267–274.
12. Holthusen H. Erfahrungen über die verträglichkeitsgrenze für röntgenstrahlen und deren nutzanwendung zur verhütung von schäden. *Strahlentherapie* 1936, **57**, 254–269.
13. Geara FB, Peters LJ, Ang KK, *et al*. Comparison between normal tissue reactions and local tumor control in head and neck cancer patients treated by definitive radiotherapy. *Int J Radiat Oncol Biol Phys* 1996, **35**, 455–462.
14. Bernier J, Thames HD, Smith CD, Horiot JC. Tumor response, mucosal reactions and late effects after conventional and hyperfractionated radiotherapy. *Radiother Oncol* 1998, **47**, 137–143.
15. Ash D, Bates T. Report on the clinical effects of inadvertent radiation underdosage in 1045 patients. *Clin Oncol* 1994, **6**, 214–226.
16. Turesson I, Nyman J, Holmberg E, Oden A. Prognostic factors for acute and late skin reactions in radiotherapy patients. *Int J Radiat Oncol Biol Phys* 1996, **36**, 1065–1075.
17. Horiot JC, van der Schueren E, Johansson KA, Bernier J, Bartelink H. The programme of quality assurance of the EORTC radiotherapy group. A historical overview. *Radiother Oncol* 1993, **29**, 81–84.
18. Horiot J-C, van der Schueren E, Johansson KA, Bernier J, Bartelink H. Quality assurance in radiotherapy. In Tobias JS, Thomas PRM, eds. *Current Radiation Oncology*. London, Edward Arnold, 1994, 365–381.