Eur J Cancer Clin Oncol, Vol. 22, No. 8, pp. 945–952, 1986.
Printed in Great Britain.

0277-5379/86$3.00+0.00
Pergamon Journals Ltd.

Standardization of Steroid Receptor Assays in Human Breast Cancer—IV. Long-Term Within- and Between-Laboratory Variation of Estrogen and Progesterone Receptor Assays

A. KOENDERS* and S. M. THORPE† (On behalf of the EORTC Receptor Group‡)

**Department of Experimental and Chemical Endocrinology, St. Radboud Hospital, Nijmegen, The Netherlands and †The Finsen Institute, Department of Clinical Physiology, Copenhagen, Denmark.*

Abstract—*One batch of lyophilized calf uterine cytosol was analyzed for estrogen and progesterone receptor (ER and PgR, respectively) content by* 12 *members of the EORTC Receptor Group on three different occasions over a total study period of* 1 yr:

(1) *One vial was included with each of* 20 *consecutive batches of routine tumor analyses between December* 1983 *and May* 1984.

(2) *Two vials were simultaneously assayed between July and August* 1984 *(within-run variation).*

(3) *One vial was analyzed at the end of* 1984 *(November–December).*

The overall mean ER and PgR values did not change systematically over the total study period of 1 yr. *Within the various laboratories, the between-run variations of both ER and PgR assays were considerable and differed from one institution to another* (7–26%). *For both ER and PgR measurements the average within-run* ($n = 2$) *and between-run* ($n = 20$) *coefficients of variation were similar* (8–9% *and* 16–17%, *respectively). Comparison of the results from multiple sequential assays* (c.v. = 12.9%) *with those from single assays* (c.v. = 21.2%) *showed that about* 60% *of the between-laboratory variance in ER could be explained on the basis of the between-run variance. With regard to PgR analysis, however, the between-laboratory variance decreased only* 25%.

Standardized use of one type of protein assay (Coomassie brilliant blue) and a standard protein solution (human serum albumin) has decreased the between-laboratory variation of the protein analysis results to less than 15%.

Accepted 7 January 1986.

‡*List of participating institutions:* Antoni van Leeuwenhoekhuis, Amsterdam, The Netherlands (W. Nooyen); Cancerologie Experimentale Laboratoire, des Recepteurs Hormonaux, Faculté de Médecine de Marseille, Marseille, France (P. Martin); Centro Medico Oncologico, Ospedale Riuniti di Parma, Parma, Italy (C. Bozzetti); Department of Biochemistry, University of Glasgow, Glasgow, U.K. (R. Leake); Finsen Institute/Fibiger Institute, Copenhagen, Denmark (C. Rose and S.M. Thorpe); Institut Jules Bordet, Brussels, Belgium (G. Leclercq); Istituto di Radiologia, University of Ferrara, Ferrara, Italy (D. Pelizzola, G. Giovannini and A. Piffanelli); Istituto di Ricerche Biomediche, 'Antoine Marxer' (RBM), Ivrea, Italy (A. Orlando and S. Fumero); Universitäts-Frauenklinik Hamburg, Isotopenlabor, Hamburg, F.R.G. (W. Jonat); Laboratoire d'endocrinologie experimentale, Centre Oscar Lambret, Lille, France (J. P. Peyrat); Radboudziekenhuis, Afd. Experimentele en Chemische Endocrinologie, Nijmegen, The Netherlands, (Th. Benraad and A. Koenders); Rotterdamsch Radio-Therapeutisch Instituut, Rotterdam, The Netherlands (M. A. Blankenstein and J. A. Foekens).

Requests for reprints should be addressed to A. Koenders.

INTRODUCTION

THE steroid receptor content in tumor tissue is of clinical significance with regard to prognosis and survival of patients with primary breast cancer and in selecting patients with advanced disease for endocrine treatment [1–6]. Estrogen and progesterone receptor (ER and PgR, respectively) assays are routinely performed in laboratories on several continents. The internal and external quality control of these complex and time-consuming assays is of utmost importance.

In 1979, the EORTC Receptor Group started a quality assurance program for laboratories assaying tumor samples of patients entering clinical trials of the EORTC Breast Cancer Cooperative Group and for representatives of existing national organizations [7]. Since 1984 this program, which utilizes lyophilized samples that can be mailed without temperature control and that tolerate pro-

longed storage at 4°C [8], has been open to all other individual laboratories for validation of receptor methodologies and results. Over 110 laboratories now co-operate within this European organization. Various external quality assurance organizations have demonstrated large between-laboratory variations in quantitative receptor values [9–13]. Such quality assurance programs generally rely on analysis of a single control sample by the various participants. The observed variations may, therefore, be ascribed to both between-laboratory variation that results from differences in methodology as well as within-laboratory variation that is related to the within- and between-run precision of receptor assays.

In the present investigation, the EORTC Receptor Group has paid special attention to the long-term between-run and between-laboratory variation of receptor results. Moreover, the contribution of between-run variation in the individual laboratories of routine assays to the overall between-laboratory variation has been evaluated.

MATERIALS AND METHODS

Preparation and distribution of control material

A control cytosol with an ER value of about 50 fmol/mg protein was prepared by 7-fold dilution of calf uterine cytosol with receptor-negative, normal calf muscle cytosol. Both cytosols were prepared by stirring pulverized tissue powder with phosphate buffer (0.02 M Na_2HPO_4/NaH_2PO_4, 1.5 mM EDTA, 3.0 mM NaN_3, 10 mM $NaMoO_4$, pH 7.5; w/v = 1 : 6) for 30 min at 4°C and centrifuging at 32,000 ***g*** for 30 min at 2°C. After thorough mixing of both cytosol solutions, 5 ml aliquots in glass vials were lyophilized in one batch by the National Institute of Public Health of the Netherlands. Vials were sealed under vacuum and stored at 4°C until mailing.

Twenty identical vials were distributed to all 14 members of the EORTC Receptor Group by mail at ambient temperature in December, 1983. All vials were stored at 4°C until use. Each participant received precise instructions with regard to the reconstitution of the lyophilized material. Thereafter, one vial was analyzed concurrently with each series of routine human breast tumor receptor assays. With few exceptions, all analyses were completed before the end of May.

In conjunction with two trials of the European Quality assurance program organized in 1984 by the EORTC Receptor Group, the same samples were analyzed two additional times in the course of the following year. For the first analysis (Trial 1) two of the four samples were from this particular batch of cytosols and they were assayed simultaneously (within-run variation). For the second trial (Trial 2) one of the four samples stemmed from this batch of cytosols. During both trials, participants were unaware of the identity of the samples. Thus, the same control sample was analyzed by participants on three different occasions and the between-laboratory variation of the receptor results could be evaluated over a period of 1 yr.

Receptor and protein assay conditions

With the following exceptions, the DCC receptor assay was performed in accord with previously accepted guidelines [7,14]: Laboratories A and G performed single dose saturation assays rather than multipoint Scatchard analysis. All participants used ^{3}H-Org 2058 to measure progesterone receptors except laboratory K who analyzed the first five vials using ^{3}H-R 5020 and the remaining 15 vials with ^{3}H-Org 2058. The results of the first five assays were, therefore, excluded from the study. Laboratories K and L used the Lowry method and BSA standard for protein analysis while the remaining laboratories all performed the Bradford technique (Coomassie brilliant blue) using human serum albumin as a standard as previously agreed. Finally, laboratory E only reported protein results after treatment of the cytosol with DCC.

Statistical methods

Within-run variation. Variation of estimates simultaneously obtained for the same sample within one series of measurements.

Between-run variation. Variation of estimates obtained for the same sample during different series of measurements.

Between-laboratory variation. Variation of estimates for one sample based upon the measurements performed by different laboratories. This estimate may be the result from one single analysis (Trial 2, November–December 1984), from two simultaneously performed measurements (Trial 1, July–August 1984), or from many, sequentially performed receptor measurements (long-term study, December 1983–May 1984).

With regard to the latter study, the between-run variation and the between-laboratory variation were estimated by one-way analysis of variance [15]. The relationship between variations, i.e. ER and PgR values observed within the same laboratories, was tested for significance with Spearman rank correlation coefficients (r_s).

RESULTS

Estradiol receptor (fmol/ml *cytosol*)

Twelve of the 14 laboratories in the EORTC Receptor Group participated in this long-term investigation. Figure 1 shows all ER results ex-

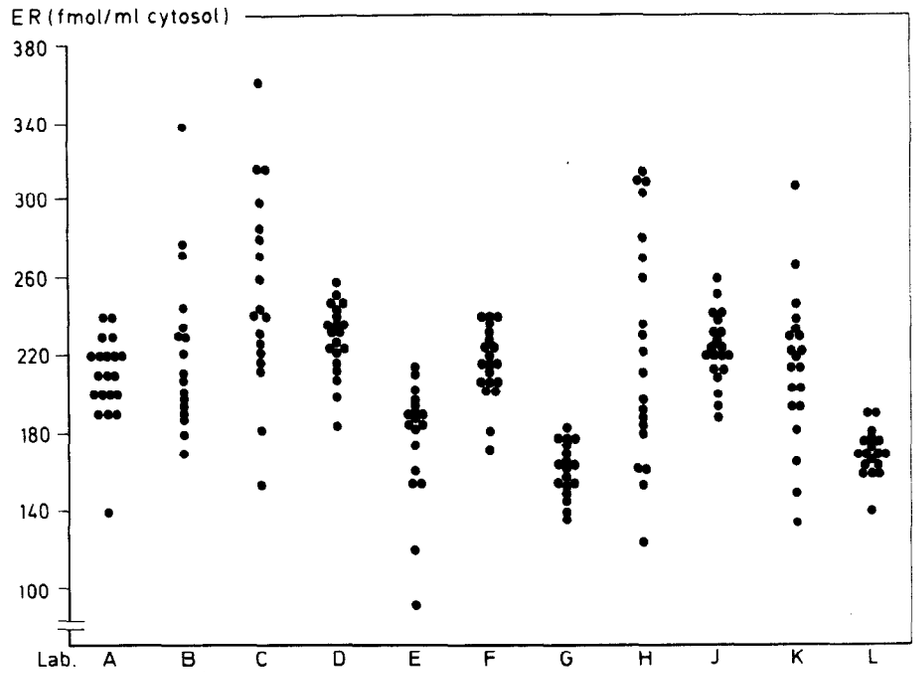

Fig. 1. Long-term variation of routinely performed ER assays determined by 11 members of the EORTC Receptor Group (excluding laboratory I).

pressed as fmol/ml cytosol reported from the 11 of the 12 participants. In some laboratories the total number of reported results is under 20; several vials were broken during shipping, and occasional single results — never more than one per laboratory — were missing due to technical errors in the assay. All participants except laboratory E observed receptor values that were normally distributed about the mean. Laboratory E's results were skewed toward low values. The two laboratories using single dose saturation assays (A and G) reported receptor values that were within the range observed by laboratories performing Scatchard assays. After 13 consistent assays (202–276 fmol/ml), the estrogen receptor values of laboratory I suddenly decreased by about 50% and demonstrated poor precision (Fig. 2). This coincided with the use of a new batch of ^{3}H-estradiol. When this laboratory employed a third batch of ^{3}H-estradiol, receptor values coincided very well with assays 1–13. These results were excluded from the analysis of variance of between-laboratory results. The PgR determinations for all 19 analyses in this particular laboratory were normally distributed.

The between-run variation of receptor results ranged between 7.4% (laboratory L) and 26.3% (laboratory H). The average coefficient of variation computed from these results was 16.0% (Table 1). Despite these large differences in day-to-day precision of assays, the mean values of all participants agreed reasonably well (range 161–254 fmol/ml). On the basis of the mean values, the between-laboratory coefficient of variation was 12.9% with an overall mean of 209 fmol/ml cytosol.

The sample used for the long-term evaluation of reproducibility in ER and PgR analyses was analyzed on two additional occasions during 1984. Since duplicate samples were analyzed in the first of these two trials, the within-run variability of receptor and protein determinations can be analyzed for the 12 laboratories. The average coefficient of variation of duplicate ER determinations is 7.9% for 11 of the 12 laboratories (Table 2). The result

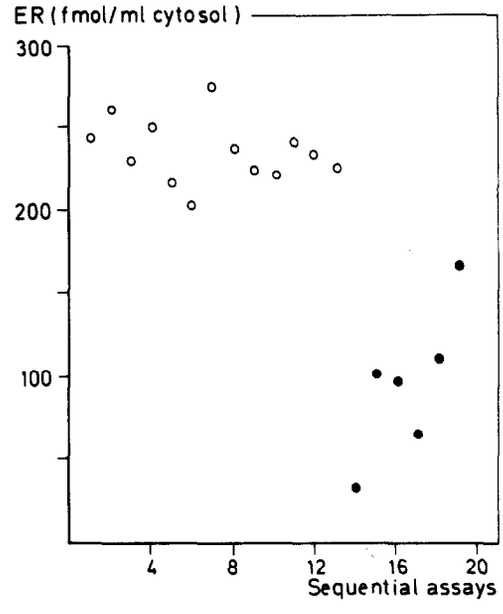

Fig. 2. Long-term variation of 20 sequential ER assays performed by laboratory I of the EORTC Receptor Group.

Table 1. Between-run and between-laboratory variation of long-term receptor and protein analysis

	Estrogen Receptor (fmol/ml cytosol)	Progesterone Receptor (fmol/ml cytosol)	Protein (mg/ml cytosol)
Range of mean values	161–254	347–665	3.45–4.38
Overall mean	209	519	3.78
Average between-run S.D.	34	86	0.29
Between-run C.V. (%)	16.0	16.6	7.7
Between-lab S.D.	27	98	0.30
Between-lab C.V. (%)	12.9	18.9	7.9

Table 2. Within-run variability of receptor and protein results assayed in duplicate

	Estrogen Receptor (fmol/ml cytosol)	Progesterone Receptor (fmol/ml cytosol)	Protein (mg/ml cytosol)
Mean value	214	462	4.00
Average S.D. (C.V.)	16.8 (7.9%)	41.8 (9.0%)	0.1 (2.5%)
No. of participants	11	11	12

Table 3. Between-laboratory variation of estrogen receptor, progesterone receptor and protein assays performed by 12 laboratories over a period of 1 yr

	Long-term study* December 1983–May 1984	Trial 1† July–August 1984	Trial 2‡ November–December 1984
ER : mean (fmol/ml cytosol)	209	214	197
: S.D. (C.V.%)	27 (12.9)	43 (20.3)	42 (21.2)
PgR : mean (fmol/ml cytosol)	519	462	471
: S.D. (C.V.%)	98 (18.9)	98 (21.2)	103 (21.9)
mg protein/ml cytosol			
: mean	3.78	4.00	3.88
: S.D. (C.V.%)	0.30 (7.9)	0.59 (14.9)	0.55 (14.0)

*20 sequential assays.
†Duplicate sample analyzed in one assay.
‡Single sample analyzed.

from laboratory H was excluded since the duplicate determination was extremely poor (256 and 130 fmol/ml cytosol).

The mean values of all 12 laboratories on the three occasions during the course of a year are very similar (Table 3). However, the between-laboratory variation ranged from 12.9% (long-term study of 20 sequential assays) to 21.2% (one single assay), which represents a 63% reduction in inter-laboratory variance (S.D.^{2}).

Progesterone receptor (fmol/ml *cytosol*)

The distribution of results for all laboratories except laboratory H appeared to be normal (Fig. 3). The results from laboratory F suddenly decreased at one point in time from about 520 to 260 fmol/ml cytosol, while precision was maintained (c.v. 9–10%) and consistency of ER values was observed. Several months later (Trial 2, 1984), this participant obtained a receptor value of 660 fmol/ml cytosol. There is no obvious explanation for the difference found, and this laboratory's PgR results have been excluded from further statistical calculations.

Between-run variations for PgR analysis were considerable and differed from institution to institution. The coefficients of variation within the various laboratories ranged from 7.8% (laboratory G) to 25.5% (laboratory H) with an average value of 16.5% (Table 1). One-way analysis of variance

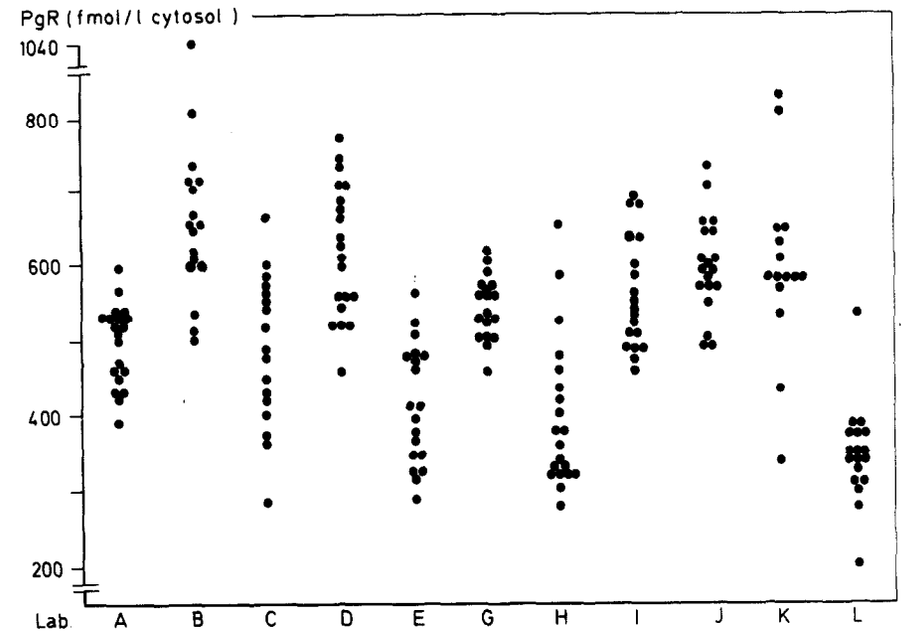

Fig. 3. Long-term variation of routinely performed PgR assays determined by 11 members of the EORTC Receptor Group (excluding laboratory F).

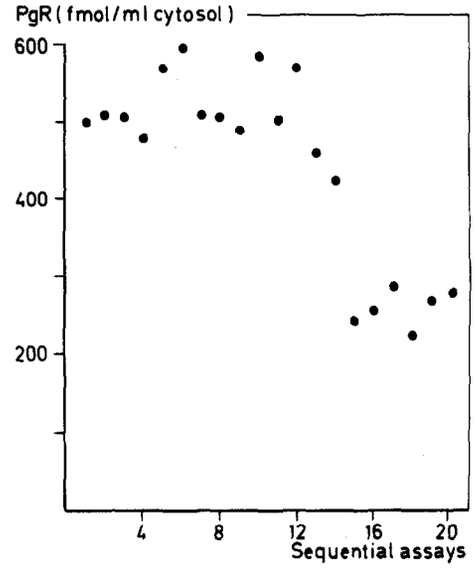

Fig. 4. Long-term variation of 20 sequential PgR assays performed by laboratory F of the EORTC Receptor Group.

resulted in a mean value of 510 fmol/ml cytosol (135 fmol/mg protein) and there was a between-laboratory c.v. of 18.9%.

The within-run variability of duplicate PgR analyses was found to be 9.0% (Table 2). The result from laboratory G is excluded because of poor concordance between the duplicate determinations (362 and 624 fmol/ml cytosol).

Table 3 compares the mean PgR values computed from the results of the same laboratories observed on the three different occasions throughout the year. The mean values agreed reasonably well, and comparison of the between-laboratory variations obtained on these two occasions shows that only 25% of the inter-laboratory variance is eliminated by means of multiple sequential assays.

A statistically significant direct correlation ($r_s = 0.67$, $P < 0.05$) was observed between the coefficients of variation observed by the participants for both the ER and the PgR assays. In contrast, similar consistency was not observed with regard to the mean values for ER and PgR ($r_s = 0.41$, $P > 0.1$).

Protein measurement

Figure 5 shows the protein results of all 12 participants. Three participants (laboratories E, K and L) were excluded from further analysis for the reasons mentioned in *Materials and Methods*. The between-run precision of the protein assays within the other laboratories ranged from 4.4% (laboratory J) to 11.1% (laboratory H) with an average value of 7.7% (Table 1). The mean values of these nine participants were between 3.45 and 4.38 mg/ml cytosol with a between-laboratory coefficient of variation of 7.9%.

With one exception, the protein values reported during the two external trials conducted in July and November, 1984 remained within the range of values obtained during the long-term laboratory study (December 1983–May 1984). During Trial 1, laboratory A found protein values of 5.3 mg/ml cytosol while during the long-term study as well as

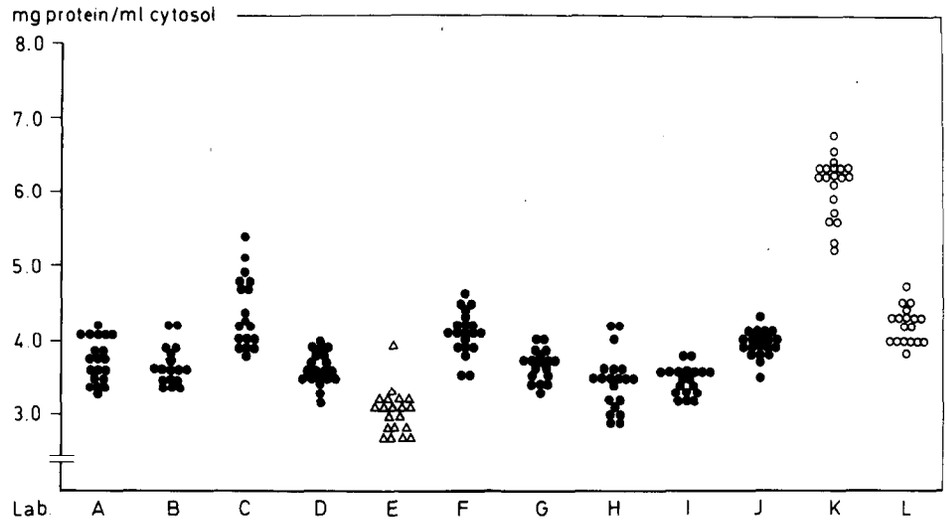

Fig. 5. Long-term variation of protein measurements. Nine participants (●) used the recommended Bradford technique (Coomassie brilliant blue) and human serum albumin standard, two (○) used the Lowry method with BSA standard, and one participant (△) reported Bradford protein results only after treatment of cytosol with DCC.

during trial 2 the values ranged between 3.3 and 4.2 (3.76 ± 0.29 mg/ml, $n = 20$). As a result of these deviating values found in laboratory A, the mean value of all participants was higher during this trial compared to the mean values from the other two investigations (Table 3). Excluding the results of this one participant, the calculated mean value of 3.84 ± 0.39 (c.v. = 10.1%) was again very similar. Comparison of the between-laboratory variance of Trial 2, i.e. results from one single assay, with those from the long-term study, i.e. a mean result computed from about 20 sequential assays, revealed a 68% difference in variance.

Estrogen and progesterone receptor (fmol/mg *cytosol protein*)

With regard to the long-term study, the between-laboratory variation of receptor results expressed as fmol/ml cytosol or fmol/mg protein was similar, i.e. ER = 57 ± 6 fmol/mg cytosol protein (c.v. = 11.7) and PgR = 146 ± 25 fmol/mg cytosol protein (c.v. = 17.3%) (compare with results of fmol/ml cytosol shown in Table 1).

DISCUSSION

This study confirms that lyophilized calf uterine samples can be stored at 4°C for extensive periods of time and can, therefore, be used for routine purposes as within- and between-laboratory control preparations. The mean ER and PgR values calculated from the results of the 12 participating laboratories were very similar over a total study period of 1 yr (Table 3). The same could be concluded from the results of about 80 different laboratories that analyzed two lyophilized calf uterine control preparations with a time interval of about 6 months (unpublished observations).

The between-run variation of receptor assays routinely performed over a period of several months could be computed from the results of 11 of the 12 participants. Such an internal quality control checks routine measurements, traces persistent trends and systematic fluctuations, and can be used to identify and reject grossly inaccurate results. The results from laboratory I demonstrate a sudden decrease in observed levels of ER that was found to coincide with employment of a new batch of ^{3}H-estradiol. This batch of tracer was apparently the source of the error since assays performed with a third batch of ^{3}H-estradiol at a later date yielded the same, high receptor values originally observed. Another laboratory observed a sudden and systematic decline of 50% of PgR levels, but the reasons for this difference remain obscure.

In some laboratories a few receptor results were much lower than all other values (e.g. Fig. 1, laboratories A and E). None of these deviating values were excluded from calculations of the intra-laboratory variation. However, between-run reproducibility still needs to be improved in some of these experienced laboratories despite their all using well-defined, standardized methodologies.

Despite the appreciable between-run variations, the mean ER values agreed very well. Comparison of the results from multiple sequential assays with those from single assays showed that about 60% of the between-laboratory variance in ER could be explained on the basis of the between-run variance. While the average between-run precision of multiple sequential ER and PgR assays was similar (16.4 and 16.6%, respectively) (Table 1), between-laboratory variance decreased only 25% when the mean values of many sequential assays were compared with the results from one single assay for

PgR. Thus it appears that the ER assay shows potentially a better between-laboratory comparability of the two assays, and that the sizable inter-laboratory variation seen on the basis of single analyses may be due to true minor methodological differences among the laboratories and/or to poor between-run precision. Ryan *et al.* also observed that the between-laboratory variability of progesterone receptor assays was larger than the variability of estrogen receptor assays [16]. PgR assays were both inaccurate and imprecise.

Comparison of the results from the present report with those of previous studies ([14,17–18] and Table 3) demonstrated that the between-laboratory coefficients of variation for PgR analyses have gradually decreased with time from more than 50% in 1980 [18] to around 20% in 1984 (Table 3).

During the first trials of the EORTC Receptor Group, unacceptably large variations in protein values were noted. Thereafter, standardization of the type of protein assay and standard protein solution has resulted in a decline of the variability of protein results to between 10 and 15% ([14,17–18] and Table 3). The same decline has been observed in a quality assurance program conducted in the Netherlands [19]. Because of this improvement, the between-laboratory variation in receptor values is similar regardless of whether results are expressed as fmol/ml cytosol or fmol/mg cytosol protein.

Acknowledgements—The major burden of this work has been borne by the Department of Experimental and Clinical Endocrinology in Nijmegen and is most gratefully acknowledged. We would also like to thank the chairpersons of the three meetings, R. J. B. King (5th and 6th EORTC Workshop) and Th. J. Benraad (Information Exchange of Practical Aspects of ER and PgR analyses in breast cancer biopsies). The two former meetings were held in Hannover (hosts, P. W. Jungblut and H. O. Hoppen) and in Venice, Italy (host, A. Piffanelli), while the latter meeting was held in Nijmegen (host, Th. J. Benraad), The Netherlands. Financial support for the above-mentioned meetings was provided by the German, Italian and Dutch divisions of ICI/Pharma, the Max-Planck Institute, and Ciba-Geigy (Italian directions).

REFERENCES

1. Jensen EV, Polley TZ, Smith S, Block GE, Ferguson DT, DeSombre ER. Prediction of hormone dependence in human breast cancer. In: McGuire WL, Carbone PP, Vollner ER, eds. *Estrogen Receptors in Human Breast Cancer*. New York, Raven Press, 1975, 37–56.
2. Cooke T, George D, Shields R, Maynard P, Griffiths K. Oestrogen receptors and prognosis in early breast cancer. *The Lancet* 1979, 995–97.
3. Osborne CK, Yochmowitz MG, Knight WA, McGuire WL. The value of estrogen and progesterone receptors in the treatment of breast cancer. *Cancer* 1980, **46**, 2884–2888.
4. Croton R, Cooke T, Holt S, George WD, Nicolson R, Griffiths K. Oestrogen receptors and survival in early breast cancer. *Br Med J* 1981, **283**, 1289–1291.
5. Stewart JF, Rubens RD, Millis RR, King RJB, Hayward JL. Steroid receptors and prognosis in operable (Stage I and II) breast cancer. *Eur J Cancer Clin Oncol* 1983, **19**, 1381–1387.
6. Rose C, Thorpe SM, Andersen KW, Pedersen BV, Mouridsen HT, Blichert-Toft M, Rasmussen BB. Beneficial effect of adjuvant tamoxifen therapy in primary breast cancer patients with high oestrogen receptor levels. *The Lancet* 1985, **i**, 16–19.
7. EORTC Breast Cancer Cooperative Group. Revision of the standards for the assessment of hormone receptors in human breast cancer. Report of the Second EORTC Workshop, held on 16–17 March 1979, in The Netherlands Cancer Institute. *Eur J Cancer* 1980, **16**, 1513–1515.
8. Benraad TH, Koenders A. Estradiol receptor activity in lyophilized calf uterus and human breast tumor tissue. *Cancer* 1980, **46**, 2762–2764.
9. Sarfaty GA, Nash AR, Keightly DD, eds. *Estrogen Receptor Assays in Breast Cancer. Laboratory Discrepancies and Quality Assurance*. New York, Masson, 1981.
10. King RJB. Quality control of estradiol receptor analysis: the United Kingdom experience. *Cancer* 1980, **12**, 2822–2824.
11. Fumero S, Piffanelli A. Results of the Italian interlaboratory quality control program for estradiol receptor assay. *Tumori* 1981, **67**, 301–306.
12. Zava DT, Guelpa C. A quality control study to assess the interlaboratory variability of routine estrogen and progesterone receptor assays. *Eur J Cancer Clin Oncol* 1982, **18**, 713–721.
13. Leclercq G, Toma S, Paridaens R, Heuson JC, eds. Clinical interest of steroid hormone receptors in breast cancer. *Recent Results Cancer Res* 1984, **91**, 89–138.
14. Koenders A, Thorpe SM (on behalf of the EORTC Receptor Group). Standardization of steroid receptor assays in human breast cancer—I. Reproducibility of estradiol and progesterone receptor assays. *Eur J Cancer Clin Oncol* 1983, **19**, 1221–1229.
15. Scheffé H. *The Analysis of Variance* New York, John Wiley, 1959.
16. Ryan ED, Clark AF, Mobbs BG, Ooi TC, Sutherland DJA, Tustanoff ER. Interlaboratory quality control of estrogen and progesterone receptor assays in breast cancer tissue using lyophilised cytosols. *Clin Biochem* 1985, **18**, 20–26.

17. Koenders A, Thorpe SM (on behalf of the EORTC Receptor Group). Standardization of steroid receptor assays in human breast cancer—II. Samples with low receptor contents. *Eur J Cancer Clin Oncol* 1983, **19**, 1467–1472.
18. Koenders A, Benraad Th J. Standardization of steroid receptor analysis in breast cancer biopsies: EORTC Receptor Group. *Recent Results Cancer Res* 1984, **91**, 89–138.
19. Koenders T, Benraad Th J. Quality control of estrogen receptor assays in The Netherlands. *Breast Cancer Research and Treatment* 1983, **3**, 255–266.