DORE
SUEZ ENVIRONNEMENT

Round Robin Report

Effluent and influent of waste water treatment Plant surface water and groundwater sludge

Réf : 0512 MEPr 118

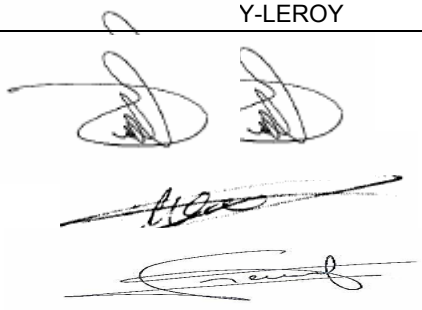
VEOLIA ENVIRONNEMENT

Global Water Research Coalition

SUEZ ENVIRONNEMENT

Pôle Expertise Analytique

REPORT
Effluent and influent of waste water treatment plant surface water and groundwater sludge

Rédacteurs	Vérificateur	Approbateur
Mar ESPERANZA /Claudine GRANET/ Gaëla Y-LEROY	Mar ESPERANZA	Zdravka DO-QUANG
		

Date : 02/17/2006

N/Réf. : 0512MEPr118_9Round Robin

Global Water Research Coalition

Global cooperation for the generation of water knowledge

GWRC is a non-profit organization that serves as a collaborative mechanism for water research. The benefits that the GWRC offers its members are water research information and knowledge. The Coalition focuses on water supply and wastewater issues and renewable water resources: the urban water cycle.

The members of the GWRC are: the Awwa Research Foundation (US), CRC Water Quality and Treatment (Australia), EAWAG (Switzerland), Kiwa (Netherlands), Suez Environment- CIRSEE (France), Stowa - Foundation for Applied Water Research (Netherlands), DVGW – TZW Water Technology Center (Germany), UK Water Industry Research (UK), Veolia- Anjou Recherché (France), Water Environment Research Foundation (US), Water Research Commission (South Africa), Water Reuse Foundation (US), and the Water Services Association of Australia.

These organizations have national research programs addressing different parts of the water cycle. They provide the impetus, credibility, and funding for the GWRC. Each member brings a unique set of skills and knowledge to the Coalition. Through its member organizations GWRC represents the interests and needs of 500 million consumers.

GWRC was officially formed in April 2002 with the signing of a partnership agreement at the International Water Association 3rd World Water Congress in Melbourne. A partnership agreement was signed with the U.S. Environmental Protection Agency in July 2003. GWRC is affiliated with the International Water Association (IWA).

DISCLAIMER

This study was jointly funded by GWRC members. GWRC and its members assume no responsibility for the content of the research study reported in this publication or for the opinion or statements of fact expressed in the report. The mention of trade names for commercial products does not represent or imply the approval or endorsement of GWRC and its members. This report is presented solely for informational purposes.

SUMMARY

PREFACE

Endocrine Disrupting Compounds (EDC) is one of the priority issues of the research agenda of the Global Water Research Coalition.
Analytical methods to determine the occurrence and fate of EDC in the water cycle are of vital importance to study the impact of EDC on public health and the aquatic environment. Over the years many analytical methods have been developed by a number of organisations but an international comparison of these methods has not been performed to date.

The GWRC likes to acknowledge Suez Environnement and Veolia Environnement for their joint leadership to organise this first international interlaboratory exercise and the participating organisations for the support.

EXECUTIVE SUMMARY

14 leading laboratories from 7 different countries took part in the first international interlaboratory exercise on hormones. Following Global Water Research Coalition's initiative, the CIRSEE and Anjou Recherche were in charge of the organization and logistics of the test.
The compounds targeted in this test were estradiol (alpha and beta), estrone and ethinylestradiol. The analytical methods to be compared had been developed by the experts of each participating laboratory and their performance had been already tested in several environmental matrixes.
The challenge consisted on determining the levels of hormones in 6 environmental samples (5 different matrixes): raw wastewater influent, treated wastewater, spiked treated wastewater, influenced surface water, spiked ground water and digested sludge from a wastewater treatment plant.

Based on the results of this first international interlaboratory exercise the following overall conclusions can be made:

- The majority of the analytical methods employed by the laboratories reported acceptable results in surface, ground waters and effluent wastewaters. Concentrations reported were well distributed around the average, and acceptable reproducibility and repeatability deviations were found.
- The limited number of laboratories (3) that analysed the sludge samples prohibited to draw firm conclusions.
- No clear tendencies were observed when comparing results obtained by the different analytical instruments. This suggests that acceptability of the results is not only a function of the final analytical technique but of the combination of all the steps in the procedure. Nevertheless, LC-coulometry and GC/MS detection methods non including a previous clean up step seem not to be adapted to the analysis of estrogenic hormones in environmental samples. Certain problems were detected as well when the concentrations of hormones were low and/or the complexity of the samples increased.
- Considering all the data and the statistical results, the optimum conditions for the analysis of hormones in complex matrixes seem the combination of a powerful purification stage followed by a highly selective and sensitive quantification technique.

INTRODUCTION

A considerably important number of analytical methods have been developed in the last years to determine the occurrence of estrogenic hormones in the environment, their behaviour in wastewater treatment process and their fate in aquifers, rivers and drinking water treatment plants. When considering the published effect concentrations – from 0.1 to 1 ng/L depending of the source – and the complexity of the targeted matrixes, the determination of hormones in environmental samples appears as a true analytical challenge. The procedures include an extraction step followed by one or more clean up procedures of divers complexity, separation and detection. The variety of techniques is wide, specially regarding the clean up and detection. Solid phase extraction, gel permeation, HPLC fractionation, size exclusion or a combination of these techniques is used to eliminate interferences. Regarding separation and detection, gas and liquid chromatography followed by mass spectrometry detectors (MS, MS-MS, TOF, etc..) are the most commonly used techniques.

Considering the volume of data published and the heterogeneity of the analytical procedures employed, it is becoming necessary to compare the results obtained by different methods and techniques.

The GWRC challenge the laboratories of its partners as well as other world leading laboratories, to test the performance of their procedures. The Round Robin test allowed to compare the performance of the existing analytical methods for environmental matrixes of different complexity. The test was proposed with the main objective of harmonizing the results within the GWRC laboratories and creating guidelines for future work on hormones within the Coalition.

This report includes details about the logistics of the Round Robin test, raw and statistical treated results and the conclusions withdrawn from them.

The results of this international interlaboratory exercise have been discussed by the participants at a meeting in le Pecq (France) on 12 October 2005. The report have been circulated for review by all participating organisations.

1 PREPARATION OF THE ROUND ROBIN

1.1 Analysed parameters

Three natural hormones and one synthetic hormone were analysed during this round robin :
- natural hormones : 17β-Estradiol, 17α-Estradiol, Estrone
- synthetic hormones : Ethinylestradiol

1.2 Tested samples

Six different samples were tested :

- Sample A : Effluent of waste water treatment plant
- Sample B : Spiked effluent of waste water treatment plant (+ 5 ng/L)
- Sample C : Influent of waste water treatment plant
- Sample D : Surface Water
- Sample E : Spiked groundwater (+ 5 ng/L)
- Sample F : Sludge

1.3 Samples preparation and shipping

For the first shipment, samples were prepared on April 26, 2005 and sent on April 27, 2005 by the CAE-Veolia Environnement laboratory (FRANCE)

Preparation :

- Sample A : A glass tank was filled with 40 liters of an effluent of waste water treatment plant collected on April 26, 2005. A 5 liters flask was used to fill this glass tank. The batch was mixed during 2 hours and the 2.5 L bottles were filled with approximately 2.2 L in three times. The 2.5 L bottles were labelled : Sample A – Effluent of WWTP.

- Sample B : First of all , the glass tank was filled with 10 liters of an effluent of waste water treatment plant collected on April 26, 2005. A 5 liters flask was used to fill this glass tank. For the spiking, a 1 L flask was filled with the waste water and spiked with 200 µL of a solution at 1 ng/µL of hormones. This flask was transferred to the tank and rinsed 4 times with the waste water. The 4 liters of rinsing were added to the tank and it was completed with 25 L of effluent of waste water treatment plant. The final concentration of the spiking was 5 ng/L. The batch was mixed during 2 hours and the 2.5 L bottles were filled with approximately 2.2 L in three times. The 2.5 L bottles were labelled : Sample B – Spiked effluent of WWTP.

- Sample C : A glass tank was filled with 40 liters of an effluent of waste water treatment plant collected on April 26, 2005. A 5 liters flask was used to fill this glass tank. The batch was mixed during 2 hours and the 2.5 L bottles were filled with approximately 2.2 L in three times. The 2.5 L bottles were labelled : Sample C – Influent of WWTP.

All the samples were stored at -80°C during the night.

Shipment :

Frozen samples were sent in an isothermal box. Together with the samples, each laboratory received a reception form.

- Sample D: Batch D consisted on surface water. The sample was taken from the Aubergenville drinking water production plant at the Seine inlet on May 17, 2005. The Seine water was packaged in 20 L jerrycans for transport to CIRSEE.

  Upon arrival at CIRSEE, the contents of 4 jerrycans, about 80 L, were added to a 150 L tank. The tank was homogenized with a submersible pump at half the height of the tank with recirculation just below the surface of the water. A second pump of the same type was placed at the bottom of the tank to fill the bottles.

  Filling was as follows:
  - ✓ 2 L of sample for the control of batch homogeneity
  - ✓ 8 bottles of 2 L
  - ✓ 2 L of sample for the control of batch stability
  - ✓ 8 bottles of 2 L
  - ✓ 2 L of sample for the control of batch stability
  - ✓ 8 bottles of 2 L
  - ✓ 2 L of sample for the control of batch stability
  - ✓ 6 bottles of 2 L
  - ✓ 2 L of sample for the control of batch homogeneity
  - ✓ 2 bottles of 2 L (reserve)

  The bottles destined for participating laboratories and stability tests were placed in a -20°C freezer.

- Sample E : Batch E was spiked ground water. The ground water was sampled at the Croissy drinking water production plant following activated charcoal treatment on May 17, 2005. It was packaged in 20 L jerrycans for transport to CIRSEE. Spiking was carried out in the laboratory. The contents of the jerrycans were added to a 150 L tank up to the 80 L mark. The tank was homogenized with a submersible pump at half the height of the tank with recirculation just below the surface of the water. To 1 L of water taken from the surface of the tank, 400 µL of a solution was added, that contained 10 mg/L of each hormone to essay. After manually homogenizing, this solution was added back to the tank at the level of recirculation. The flask that previously contained the spiked water was conscientiously rinsed with water from the tank to assure complete transfer of the spike. A second pump of the same type was placed at the bottom of the tank with recirculation at the middle of the tank. Homogenization by the two pumps lasted 4 hours. Bottles were then filled using the same procedure as for batch D.

- Sample F was a digested sludge from the wastewater treatment station. The sludge was sampled at Asnières-sur-Oise on April 29, 2005. It was freeze dried and grinded before packaging in 200 mL bottles.

Shipment:
Frozen samples, 2 times 2 L of batch D, 2 times 2 L of batch E and 1 bottle of batch F (for participants in this test) were packed in insulated containers and spaces were filled with vermiculite. The containers were placed in cartons filled with packing material.
Packages were shipped to all the participants by DHL. They were picked up on May 18, 2005.

1.4 Round Robin Scheme

Samples had to be analysed as soon as possible after reception. All water samples had to be filtered through glass fibre filter (1.2µm) before extraction and only free steroid hormones were analysed in filtrated samples.
It was asked that all the samples have to be extracted and analysed in duplicate.

1.5 Participants

14 Laboratories took part in this round robin. The list of the participants was the following :

- Al control Laboratories (UK)
- Anglian Water (UK)
- Berliner Wasser Betrieb (Germany)
- CAE – Veolia Environnement (France)
- CRC WQT (Australia)
- CIRSEE – Suez Environnement (France)
- EAWAG (Switzerland)
- Institut for Environmental Studies (Netherlands)
- Kiwa Water Research (Netherlands)
- National Laboratory Service (UK)
- TZW (Germany)
- USEPA (USA)
- USGS (USA)
- WRC (South Africa)

2 HOMOGENEITY AND STABILITY TESTS

The homogeneity and stability tests were performed in order to verify that there were no differences between all the samples sent to the laboratories.

Results for homogeneity and stability tests are presented in the followings tables.
Regarding to the general means, the standard deviations and the relative standard deviations, all the samples tested (samples A to E) were homogenous and stable.

SUMMARY TABLE OF THE BATCHES HOMOGENEITY AND STABILITY TESTS

	Sample A				Sample B			
	Estrone ng/L	Beta Estradiol ng/L	Alpha Estradiol ng/L	Ethinylestradiol ng/L	Estrone ng/L	Beta Estradiol ng/L	Alpha Estradiol ng/L	Ethinylestradiol ng/L
HOMOGENEITY TEST								
Date :	*April 27, 2005*				*April 27, 2005*			
Number of values	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*
General mean	**1.08**	**0.45**	**0.17**	**< 0.4**	**5.56**	**2.20**	**3.49**	**3.55**
Standard deviation	0.03	0.02	0.02		0.15	0.05	0.14	0.17
Variation coefficient %	0.03	0.04	0.13		0.0	0.0	0.0	0.0
STABILITY TEST								
Date :	*May 3, 2005*				*May 3, 2005*			
Temperature of storage	*minus 20°C*				*minus 20°C*			
Number of values	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*
General mean	**1.13**	**0.50**	**0.18**	**<0.4**	**5.85**	**1.96**	**3.51**	**4.32**
Standard deviation	0.00	0.02	0.05		0.06	0.00	0.29	0.42
Variation coefficient %	0.2	4.4	25.61		1.1	0.2	8.2	9.8
Variation percentage	*4.1*	*9.8*	*7.6*		*5.2*	*-11.2*	*0.7*	*21.7*
STABILITY TEST								
Date :	*May 10, 2005*				*May 10, 2005*			
Temperature of storage	*minus 20°C*				*minus 20°C*			
Number of values	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*
General mean	**1.20**	**0.58**	**0.15**	**<0.4**	**5.76**	**2.00**	**3.51**	**4.60**
Standard deviation	0.08	0.00	0.01		0.10	0.03	0.32	0.27
Variation coefficient %	6.3	0.2	5.81		1.7	1.4	9.1	5.8
Variation percentage	*11.3*	*27.1*	*-12.4*		*3.5*	*-9.2*	*0.5*	*29.7*
STABILITY TEST								
Date :	*May 24, 2005*				*May 24, 2005*			
Temperature of storage	*minus 20°C*				*minus 20°C*			
Number of values	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*
General mean	**1.25**	**0.44**	**0.19**	**<0.4**	**6.40**	**2.95**	**4.52**	**4.85**
Standard deviation	0.02	0.02	0.00		0.04	0.06	0.06	0.03
Variation coefficient %	1.9	5.6	1.84		0.6	1.9	1.4	0.6
Variation percentage	*15.2*				*15.1*	*34.0*	*29.6*	*36.8*
General mean	***1.16***	***0.49***	***0.17***	***<0.4***	***5.89***	***2.28***	***3.76***	***4.33***
Standard deviation	***0.07***	***0.06***	***0.02***		***0.36***	***0.46***	***0.51***	***0.57***
Conclusion	**Homogeneous Stable**	**Homogeneous Stable**	**Homogeneous Stable**	**- -**	**Homogeneous Stable**	**Homogeneous Stable**	**Homogeneous Stable**	**Homogeneous Stable**

SUMMARY TABLE OF THE BATCHES HOMOGENEITY AND STABILITY TESTS

	Sample C			
	Estrone ng/L	**Beta Estradiol ng/L**	**Alpha Estradiol ng/L**	**Ethinylestradiol ng/L**
HOMOGENEITY TEST				
Date :	*April 27, 2005*			
Number of values	*4*	*4*	*4*	*4*
General mean	**45.48**	**24.56**	**3.53**	**0.81**
Standard deviation	0.74	0.86	0.05	0.02
Variation coefficient %	3.9	3.9	2.5	1.9
STABILITY TEST				
Date :	*May 3, 2005*			
Temperature of storage	*minus 20°C*			
Number of values	*2*	*2*	*2*	*2*
General mean	**47.82**	**24.70**	**3.69**	**1.19**
Standard deviation	1.86	0.62	0.07	0.25
Variation coefficient %	3.9	2.5	1.9	20.9
Variation percentage	*5.1*	*0.6*	*4.7*	*46.4*
STABILITY TEST				
Date :	*May 10, 2005*			
Temperature of storage	*minus 20°C*			
Number of values	*2*	*2*	*2*	*2*
General mean	**50.00**	**27.22**	**3.55**	**1.46**
Standard deviation	1.45	0.32	0.11	0.04
Variation coefficient %	2.9	1.2	3.0	2.8
Variation percentage	*9.9*	*10.8*	*0.8*	*80.5*
STABILITY TEST				
Date :	*May 24, 2005*			
Temperature of storage	*minus 20°C*			
Number of values	*2*	*2*	*2*	*2*
General mean	**48.73**	**25.85**	**3.67**	**1.23**
Standard deviation	0.05	0.03	0.03	0.02
Variation coefficient %	0.1	0.1	0.9	1.6
Variation percentage	*7.1*	*5.3*	*4.0*	*51.5*
General mean	***48.01***	***25.58***	***3.61***	***1.17***
Standard deviation	***1.91***	***1.23***	***0.08***	***0.27***
Conclusion	**Homogeneous Stable**	**Homogeneous Stable**	**Homogeneous Stable**	**Homogeneous Stable**

SUMMARY TABLE OF THE BATCHES HOMOGENEITY AND STABILITY TESTS

	Sample D				Sample E			
	Estrone ng/L	**Beta Estradiol ng/L**	**Alpha Estradiol ng/L**	**Ethinylestradiol ng/L**	**Estrone ng/L**	**Beta Estradiol ng/L**	**Alpha Estradiol ng/L**	**Ethinylestradiol ng/L**
HOMOGENEITY TEST								
Date :	*May 17, 2005*				*May 17, 2005*			
Number of values	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*
General mean	**5.30**	**< 1**	**< 1**	**< 1**	**3.50**	**5.40**	**3.70**	**5.00**
Standard deviation	0.10				0.10	0.20	0.10	0.30
Variation coefficient %	2.50				2.9	3.7	2.7	6.0
STABILITY TEST								
Date :	*May 19, 2005*				*May 19, 2005*			
Temperature of storage	*minus 20°C*				*minus 20°C*			
Number of values	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*
General mean	**4.77**	**< 1**	**< 1**	**< 1**	**3.49**	**4.56**	**3.13**	**4.49**
Standard deviation	0.08				0.06	0.12	0.03	0.47
Variation coefficient %	1.7				1.7	2.6	1.0	10.5
Variation percentage	*-10.0*				*-0.3*	*-15.6*	*-15.4*	*-10.2*
STABILITY TEST								
Date :	*May 20, 2005*				*May 20, 2005*			
Temperature of storage	*minus 20°C*				*minus 20°C*			
Number of values	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*
General mean	**6.06**	**< 1**	**< 1**	**< 1**	**3.25**	**5.01**	**3.92**	**4.39**
Standard deviation	0.84				0.10	0.04	0.59	0.16
Variation coefficient %	13.9				3.1	0.8	15.1	3.6
Variation percentage	*14.3*				*-7.1*	*-7.2*	*5.9*	*-12.2*
STABILITY TEST								
Date :	*June 6, 2005*				*June 6, 2005*			
Temperature of storage	*minus 20°C*				*minus 20°C*			
Number of values	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*
General mean	**5.28**	**< 1**	**< 1**	**< 1**	**3.19**	**5.21**	**3.77**	**4.66**
Standard deviation	0.07				0.10	0.21	0.16	0.75
Variation coefficient %	1.3				3.1	4.0	4.2	16.1
Variation percentage	*-0.4*				*-8.9*	*-3.5*	*1.9*	*-6.8*
General mean	***5.4***	***<1***			***3.36***	***5.05***	***3.63***	***4.64***
Standard deviation	***0.5***				***0.16***	***0.36***	***0.35***	***0.27***
Conclusion	**Homogeneous**	**-**	**-**	**-**	**Homogeneous**	**Homogeneous**	**Homogeneous**	**Homogeneous**
	Stable	**-**	**-**	**-**	**Stable**	**Stable**	**Stable**	**Stable**

3 PRESENTATION OF THE REPORT

3.1 Method information

Since the objective of the exercise was to compare the performance of the different methods developed within the coalition, no analytical procedure or analytical technique was imposed. As a consequence, and because of each participant conducted the analysis following their own procedure, the exercise was virtually performed using a different method per laboratory. The most important steps of the procedures carried out by each laboratory are shown in Table 3.1.1.

3.1.1. Scheme of the analytical procedures used by each participant

Lab Code	Extraction	Clean up	Clean up 2	Derivatization	Analytical Technique	Ionization mode
1	SPE-PolyStyrene-divinylbenzene	LC-Agilent PLGel	SPE-Aminopropyl	Dansyl Chloride	LC-MS-MS	APPI +
2	SPE-C-18	SPE-Florisil	-	-	LC-MS-MS	ESI -
3	SPE-Oasis-HLB	-	-	-	LC-MS-MS	ESI -
4	SPE-PolyStyrene-divinylbenzene	LC-C-18	-	SIGMA SIL-A	GC-MS-MS	EI
5	SPE-Oasis-HLB	-	-	-	LC-MS	ESI -
6	SPE-C-18	SPE-C18	-	-	LC-Coulometry	-
7	SPE-C-18	SPE-Florisil	-	-	LC-MS-MS	ESI -
8	SPE-C-18	-	-	MSTFA	GC-MS	EI
9	SPE-Baker SDB1	LC-Zorbax Cyano	LC-GPC PlGel50	-	LC-TOF	APPI -
10	SPE-LichrolutEN-C-18	SPE-Silica Gel	-	-	LC-MS-MS	ESI -
11	SPE-C-18	SPE-Silica Gel	-	-	LC-MS-MS	ESI -
12	SPE-Oasis-HLB	-	-	BSTFA	GC-MS	EI

3.2 Generalities on the statistical tests

The statistical methods used by VEOLIA and SUEZ-ENVIRONNEMENT to process the data were very similar. Both laboratories have exploited separately all the data and the conclusions were identical.
For each parameter, the lab results were alternately subjected to two statistical tests (Cochran and Grubbs). It should be emphasized that such statistical tests are intended only as tools for an evaluation of the consistency of the lab results among themselves.

3.2.1 Normality test

Two different methods for testing the normality were used by CAE and CIRSEE but the results and the conclusions were similar.

CAE Exploitation

The normality was verified using a normal probability plot (**Henry straight line**). The normality probability plot is a graphical technique for assessing whether or not a data set is approximately normally distributed. The data were plotted against a theoretical normal distribution in such way that the points should form an approximately straight line. Differences from this straight line indicate deviations from normality.

CIRSEE exploitation

Application of the above tests presumes a normal or at least mono-modal distribution of the results. Yet, this was not always the case (particularly when the labs used several methods). Normality in results distribution was then tested according to the **Shapiro-Wilk test** based on standard NF X 06-050 both before and after deletion of the outliers. This test enables either acceptance or rejection of the assumption of distribution normality, but fails to provide any confirmation that the distribution is indeed normal. Rejection of the assumption is associated to a risk α (1st order) that the true assumption may be rejected.

3.2.2 Cochran Test

Test designed to check whether the highest relative standard deviation is statistically aberrant as compared with the full set of relative standard deviations. If so, the maximum relative standard deviation is neglected and the test is then applied to the standard deviation immediately below. The procedure is repeated as many times as needed. On the second graphs, the labs Cochran-tested as statistically aberrant are indicated with « * » shown after their code number. Whenever the relative standard deviation of a lab was regarded as aberrant, the code number is followed by « ° » (case of labs shown in brackets or labs who supplied results showing overall inconsistency).

3.2.3 *Grubbs Test*

Test designed to check if the highest or lowest (or two highest or lowest) mean(s) is(are) statistically aberrant as compared with the full set of mean values. If so, the extrema value(s) is(are) neglected and the test is then applied to the remaining mean values. The procedure is repeated as many times as needed. On the second graphs, the labs Grubbs-tested as statistically aberrant are indicated with « ** » shown after their code number. Whenever the lab's mean value was regarded as aberrant, the code number is followed by « °° » (case of labs shown in brackets or whose results are inconsistent as compared to other labs).

3.3 Results presentation for each parameters

For each parameter, statistical results are presented as follows on different pages:

- Summary of the statistical calculation giving the following information :
  - **M** = General mean of all labs
  - The total number of participant laboratories
  - The number of labs included in statistical tests
  - The number of labs used for the calculation
  - The number of the laboratory excluded because of the normality test
  - The number of the laboratory excluded because of statistical treatment (Grubbs or Cochran)
  - **Sr** =Repeatability standard deviation
  - **SR** = Reproductibility standard deviation
  - **CVr** = (Repeatability variation coefficient
  - **CVR** = Reproductibility variation coefficient
  - **CVR/CVr** = Predominance of inter-lab error

- 2 graphs : illustrating visually the performances of each lab (mean value and standard deviation) as compared with other labs and with a comparison value whenever available ; Results are expressed as follows:

Graph 1 :

◆ values of lab i

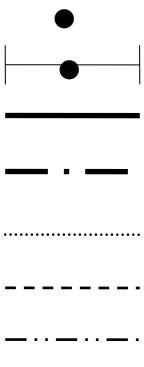
_____ General mean m of all labs

_____ Limits of interval m ± 2 * SR

_____ Limits of interval m ± 3 * SR

Each lab is represented by a code and information about the methods and limits of quantification are shown on the graph

Graph 2 :



Mean value of lab i (pattern corresponding to analysis method used)

Standard deviation related to mean mi

General mean m of all labs

Comparison value Comp = value of spiking (bold alternate dotted line)

Limits of interval m ± 1 * SR

Limits of interval m ± 2 * SR

Limits of interval m ± 3 * SR

Each lab is represented by a code. The annotations shown alongside this code have the following meaning:

* = lab whose relative standard deviation is regarded as aberrant by Cochran test

° = lab whose relative standard deviation is regarded as aberrant by statistician

** = lab whose mean is regarded as aberrant by Grubbs test

°° = lab whose mean is regarded as aberrant by statistician

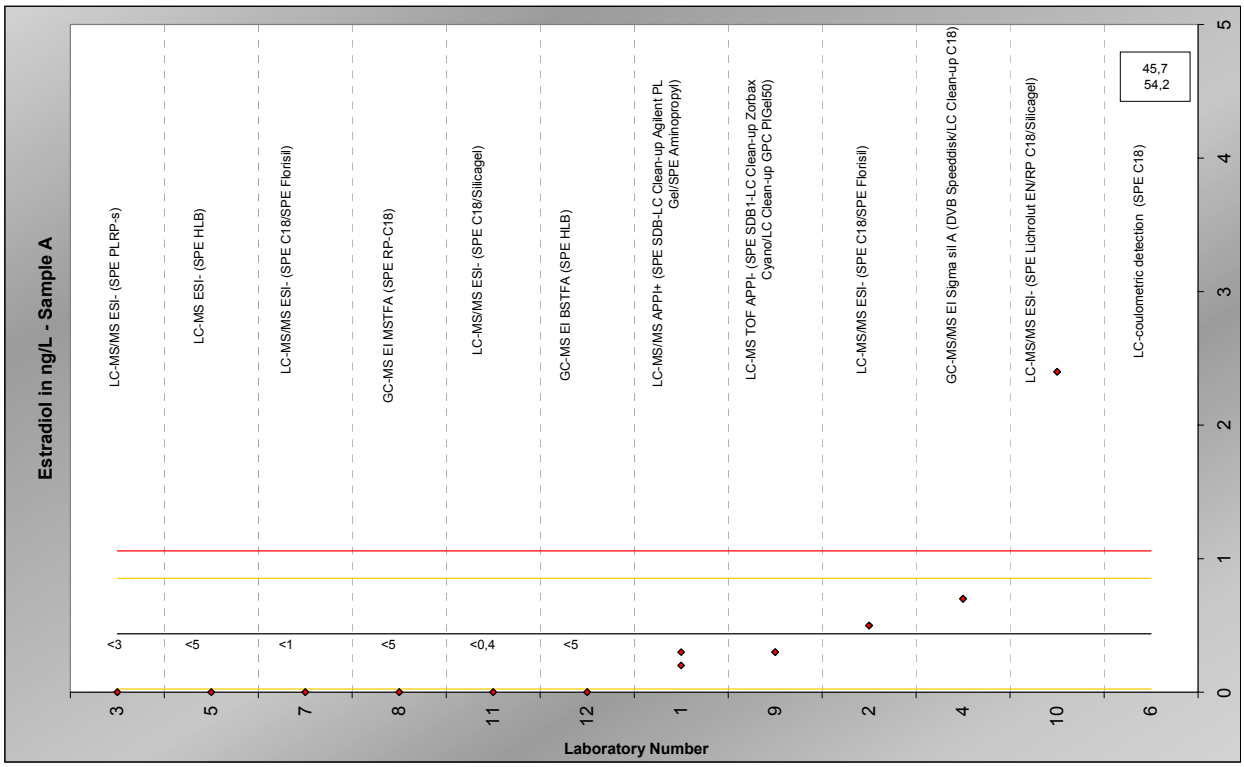
4 RESULTS

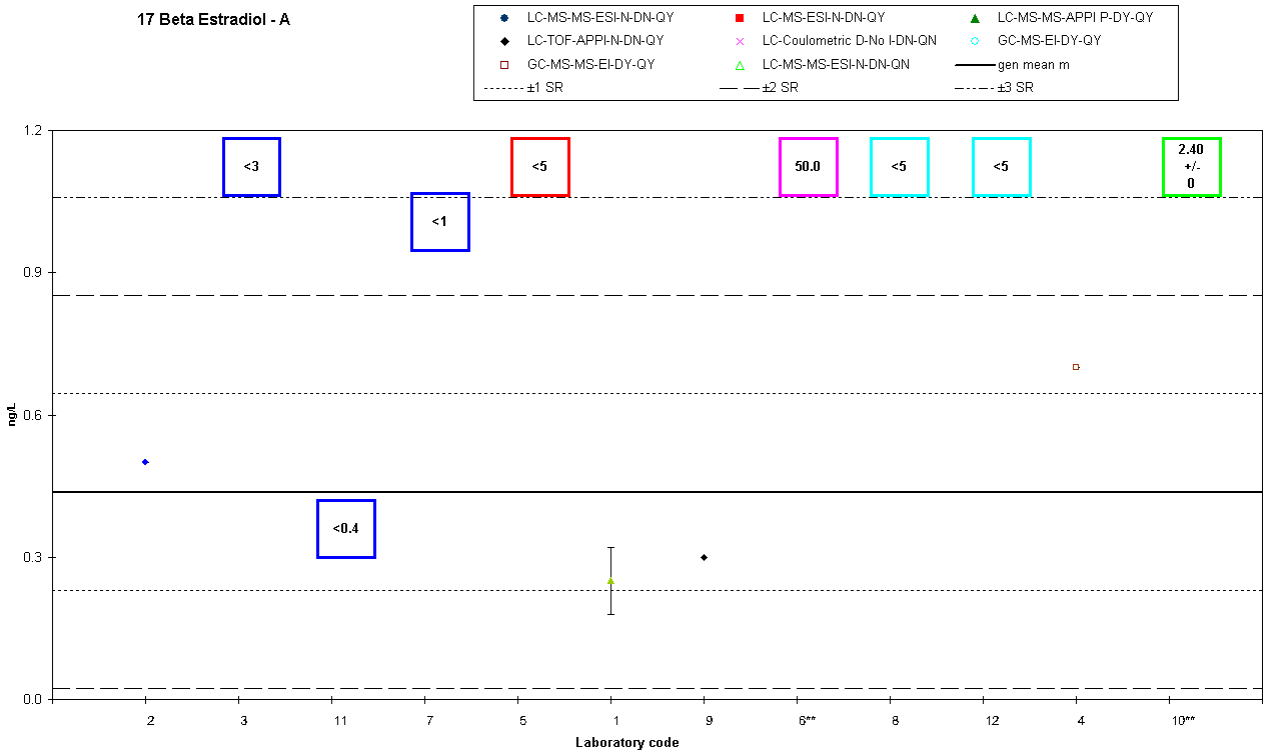
4.1 17β-Estradiol

4.1.1 Statistical Summary

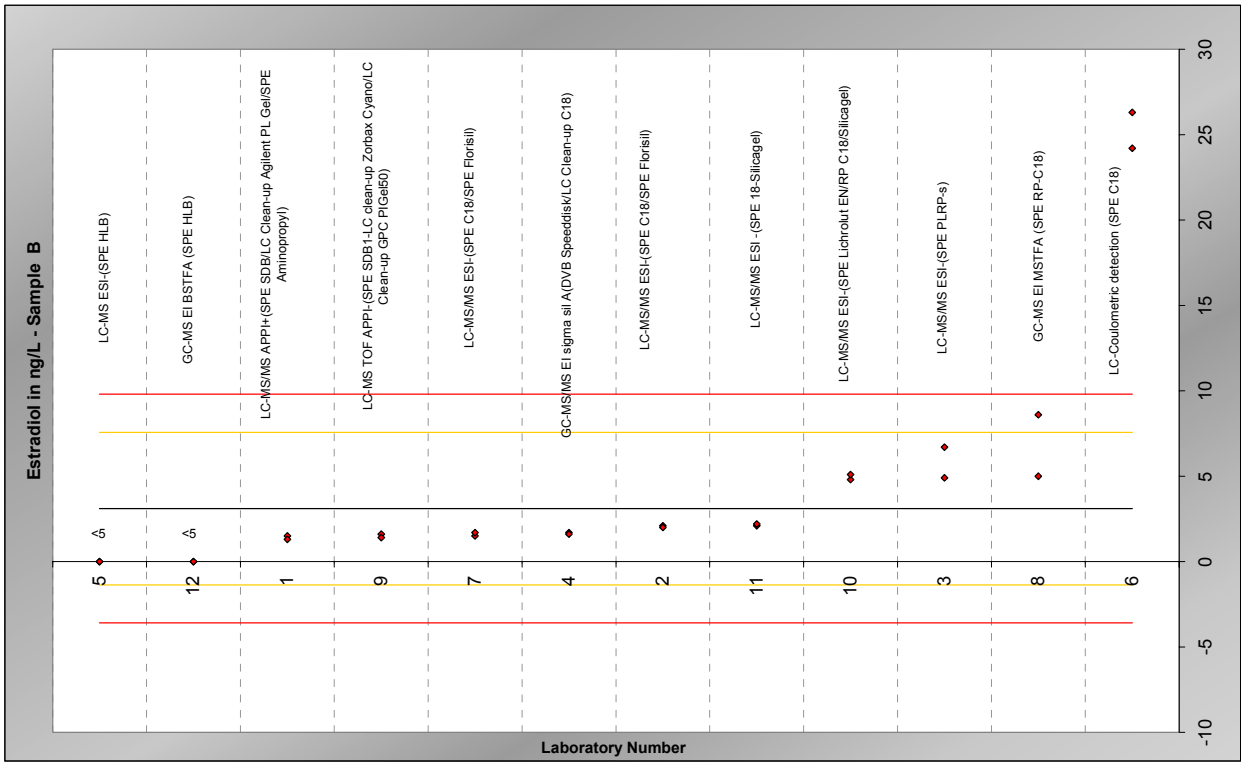
	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F
Mean value M (ng/L ou ng/g)	0,44	3,10	27,70	1,17	5,97	5,51
Total number of participing laboratories	12	12	9	12	12	6
Number of labs included in statistical tests	6	10	7	7	11	3
Number of labs used for the calculation	4	9	7	5	11	3
Exclusion because of the normality test	Labs 6 and 10	Lab 6	/	Labs 6 and 10	/	/
Exclusion because of statistical treatment (Grubbs or Cochran)	Lab 1 aberrant for Cochran test but used for the calculation	Lab 8 aberrant for Cochran test but used for the calculation	/	/	Lab 6 doubtful for Grubbs test but used for the calculation	/
Repeatability standard deviation Sr (ng/L or ng/g)	0,04	0,96	0,90	0,05	0,99	0,7
Reproductibility standard deviation SR (ng/L or ng/g)	0,21	2,23	3,63	0,22	1,46	0,7
Repeatability variation coefficient CVr(%)	8,1	30,8	3,3	4,7	16,5	12,8
Reproductibility variation coefficient CVR (%)	47,8	72	13,1	19,1	24,4	10
Predominance of inter-lab error	5,86	2,34	4,03	4,07	1,48	1

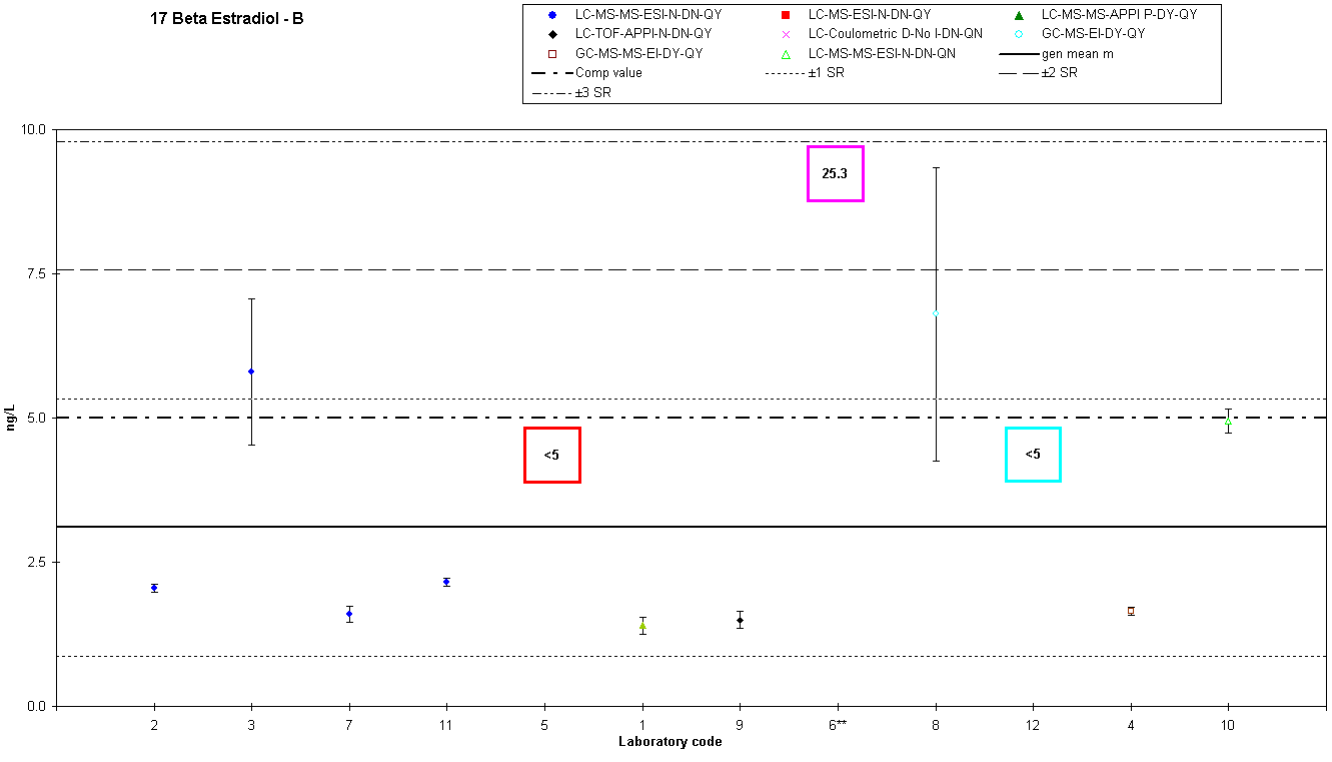
Sample A - Graphs



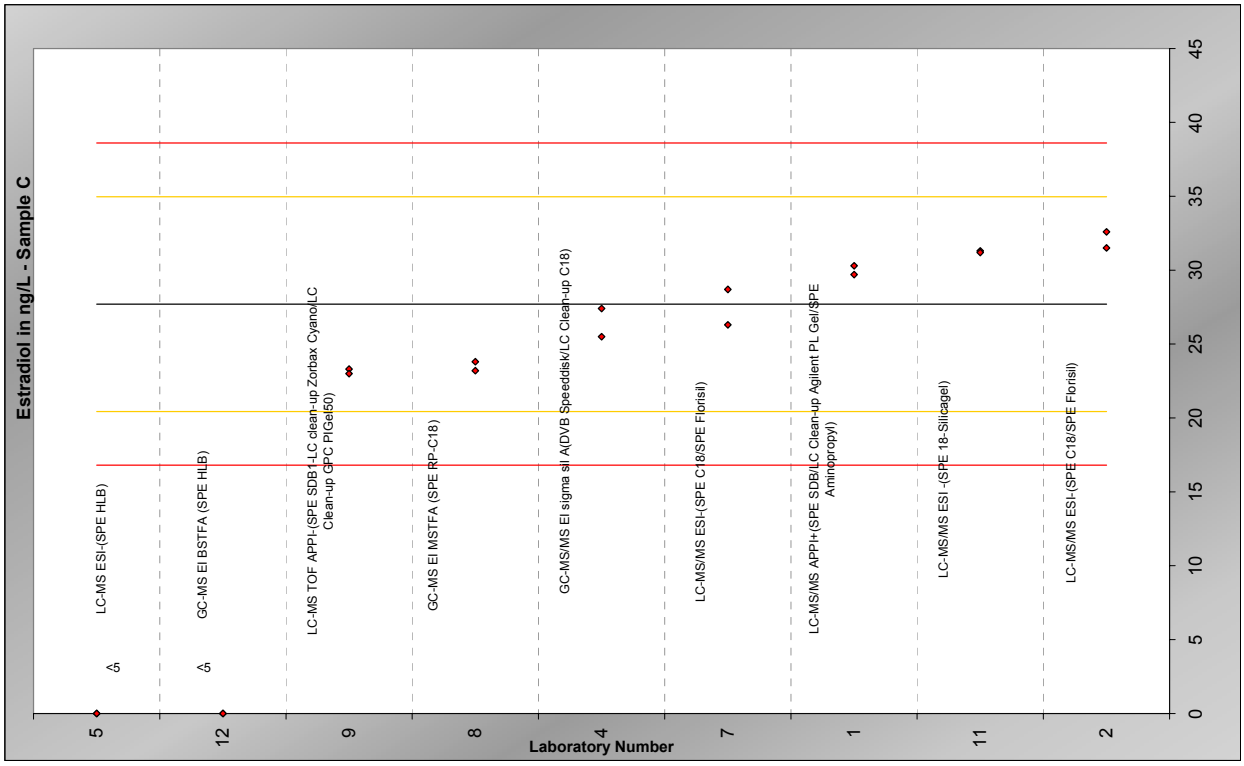


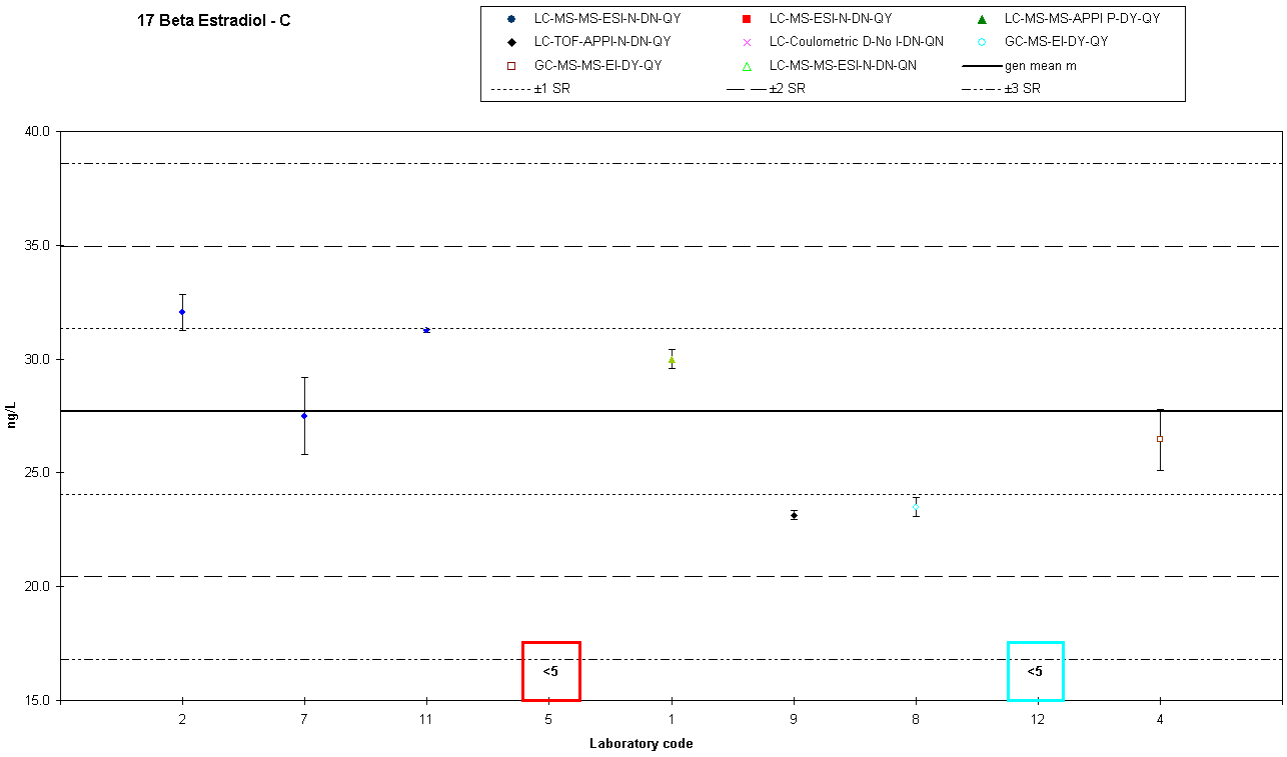
Sample B - Graphs



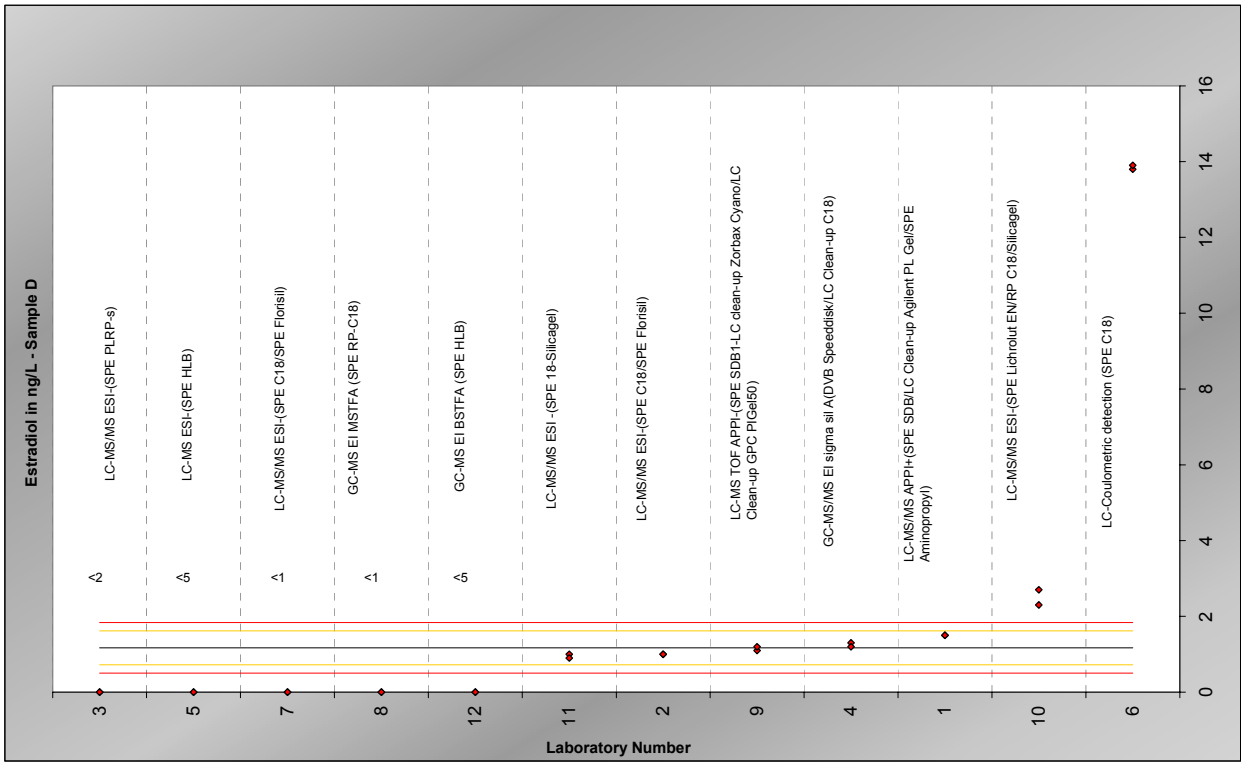


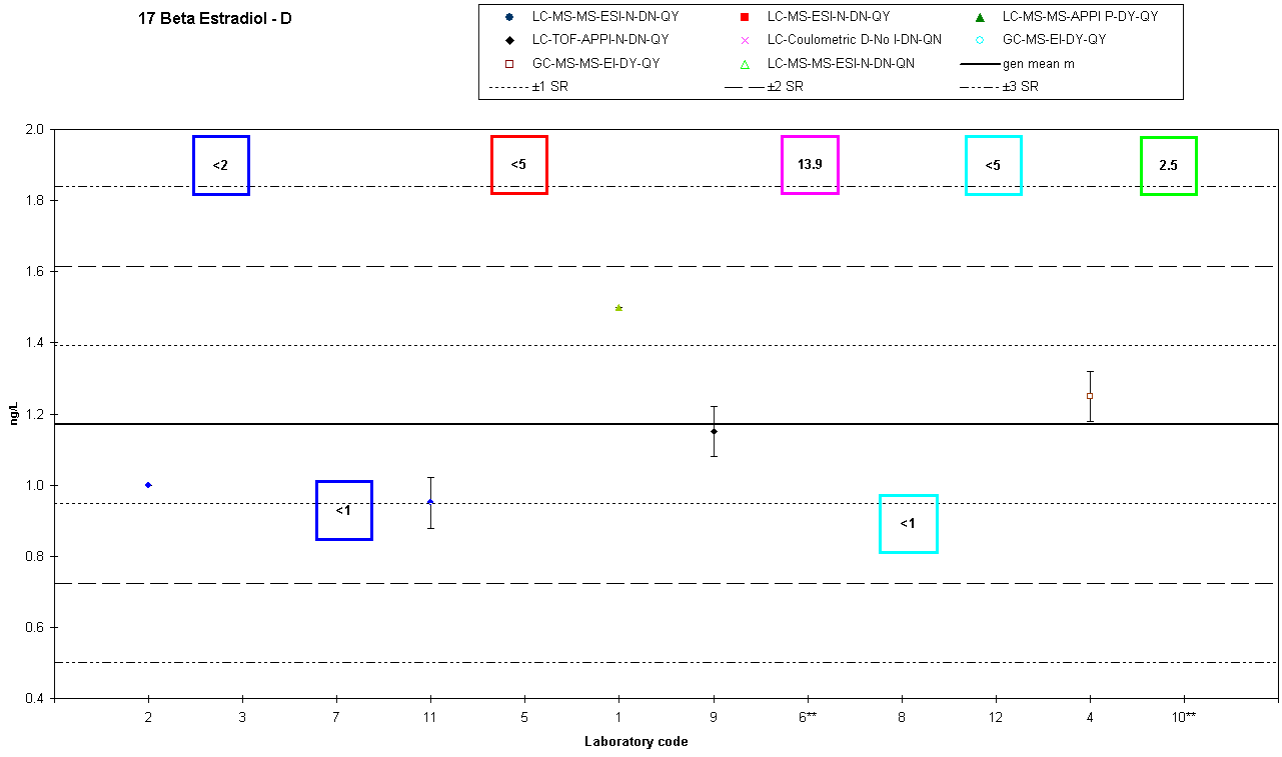
Sample C - Graphs



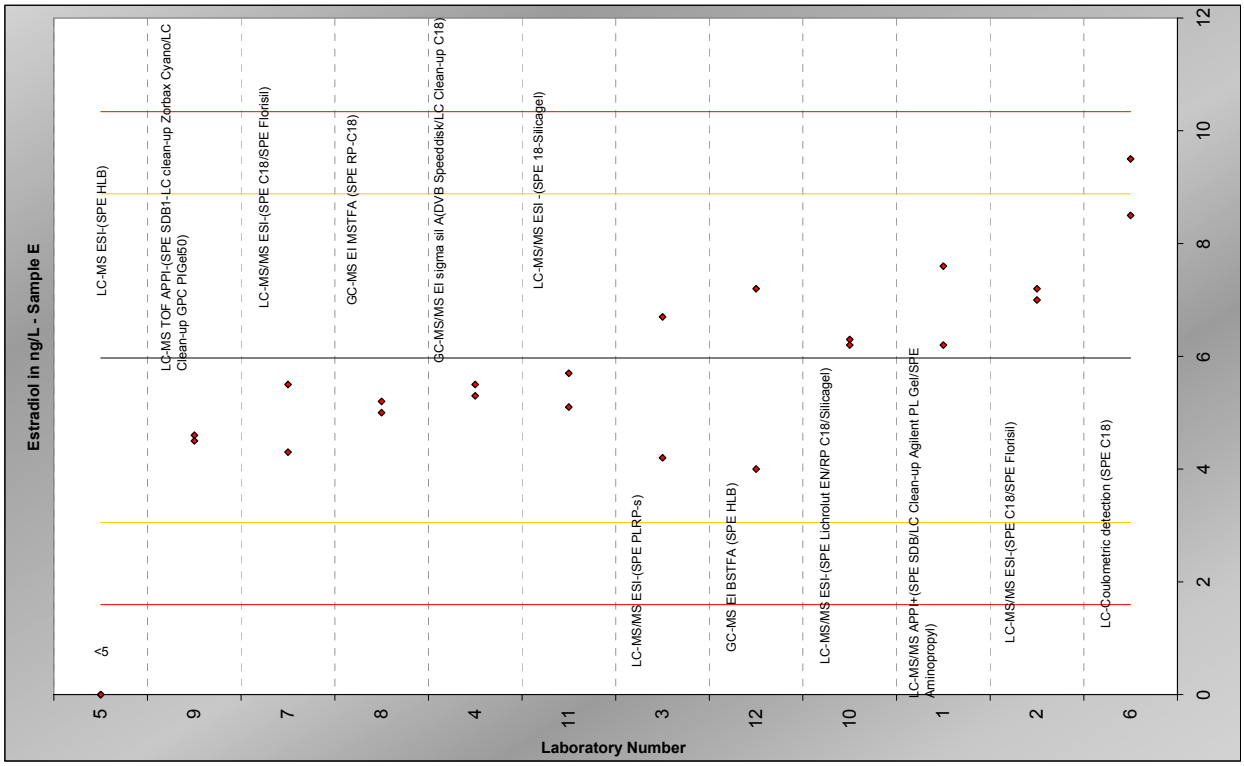


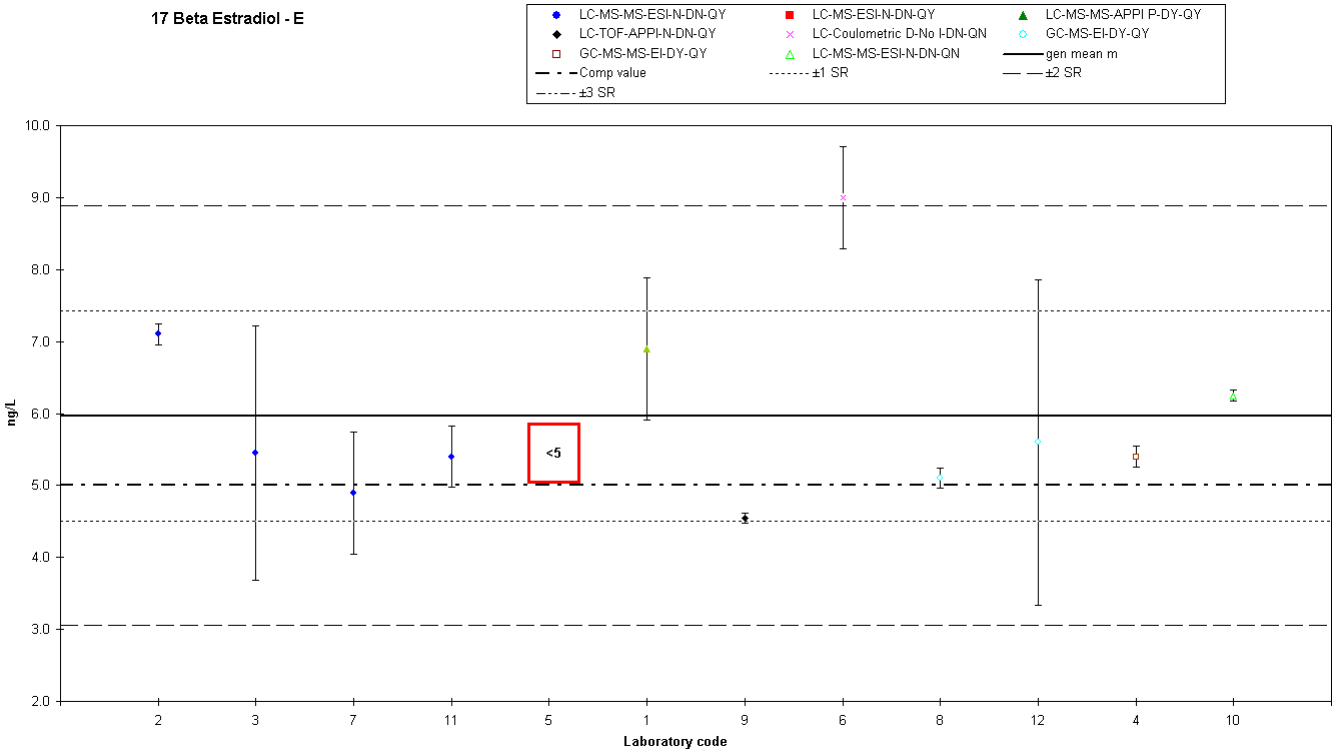
Sample D - Graphs



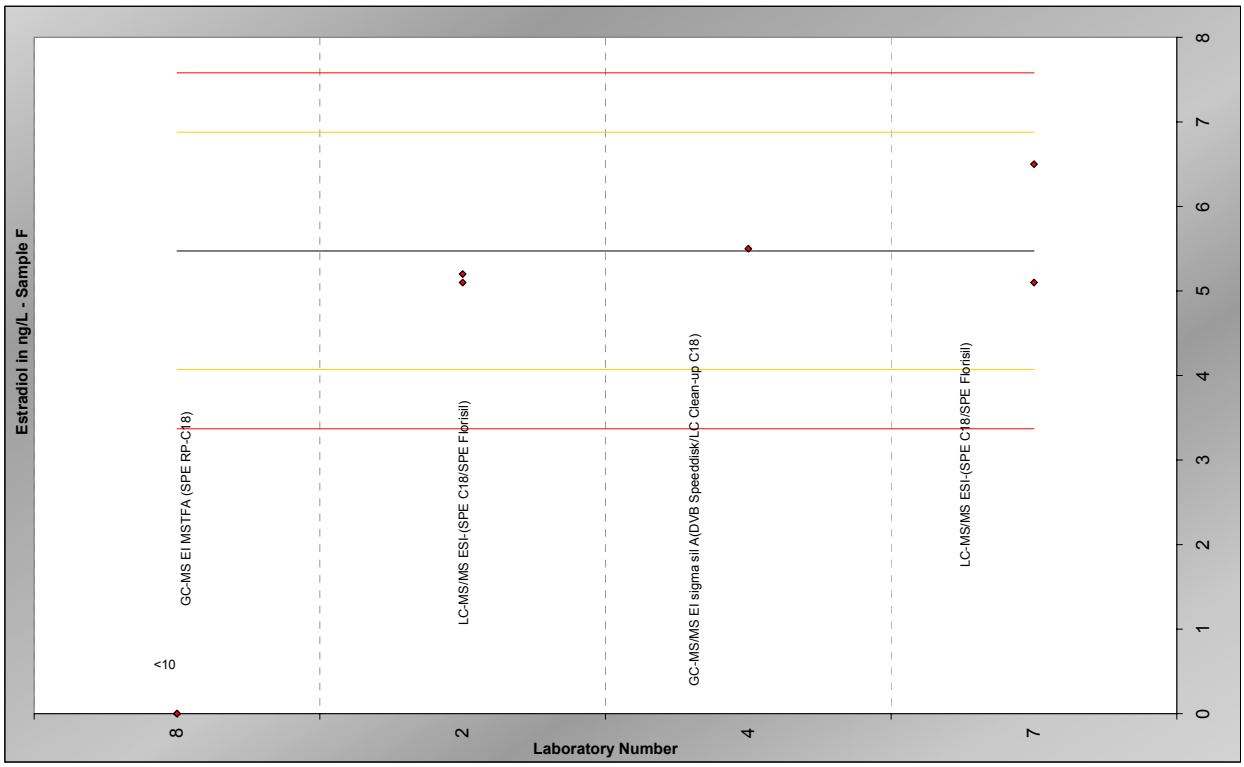


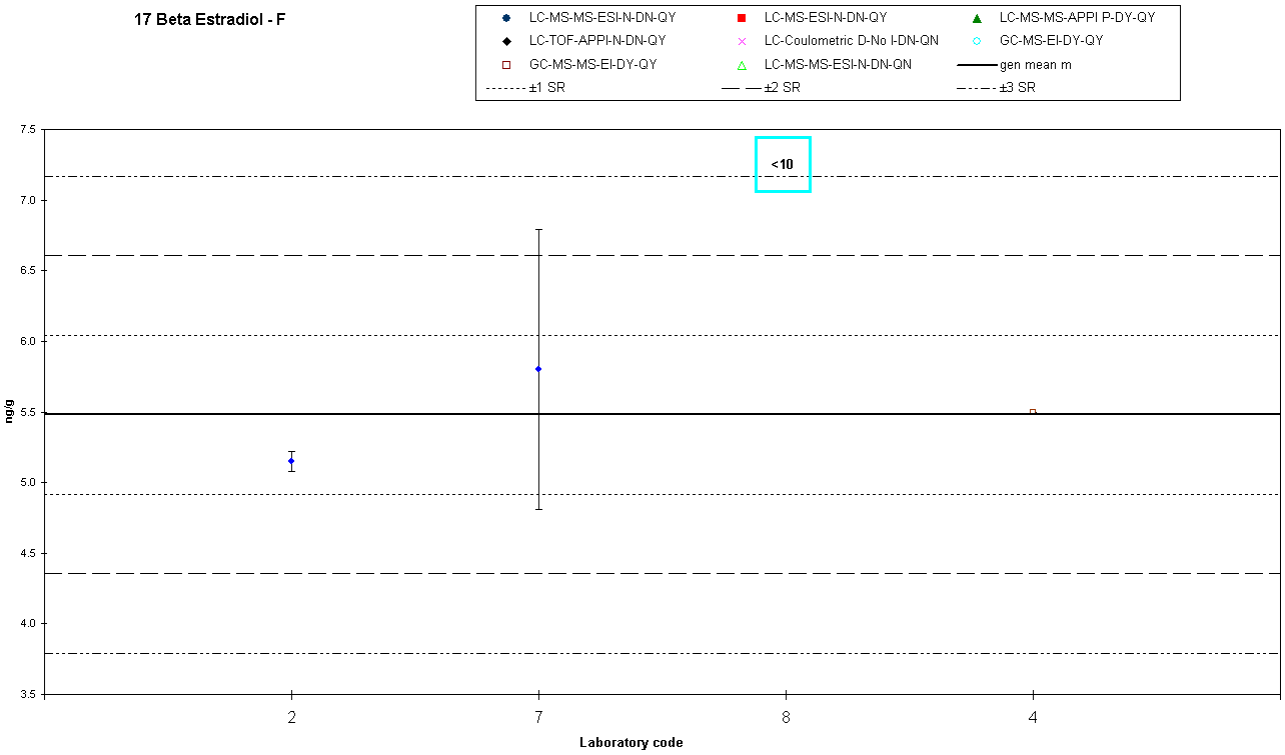
Sample E - Graphs





Sample F - Graphs



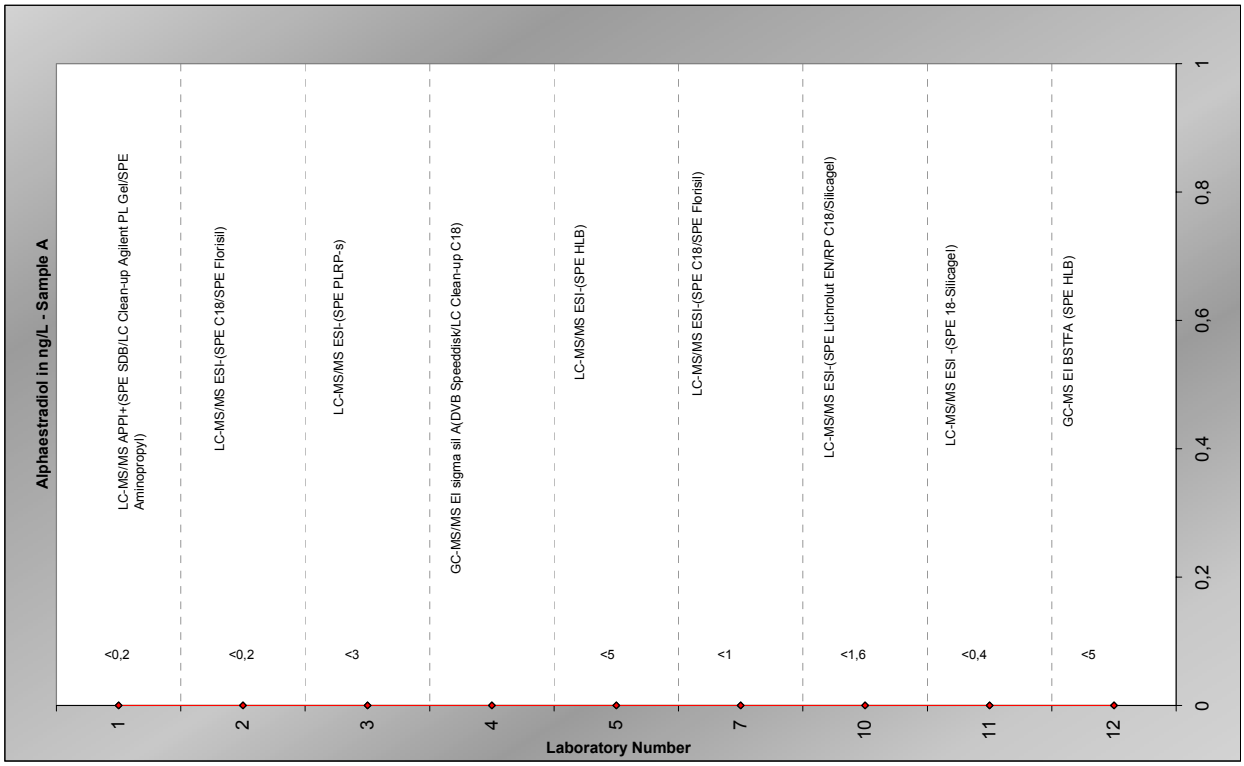


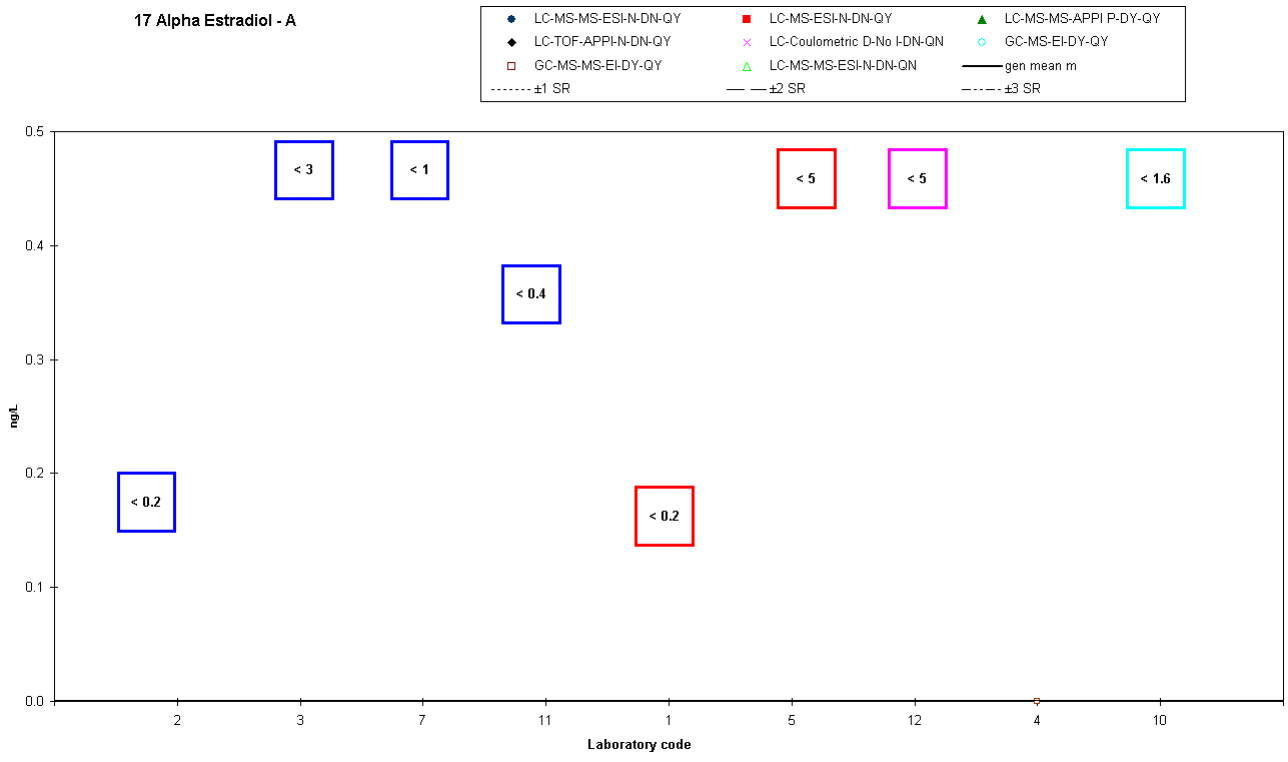
4.2 17α-Estradiol

4.2.1 *Statistical Summary*

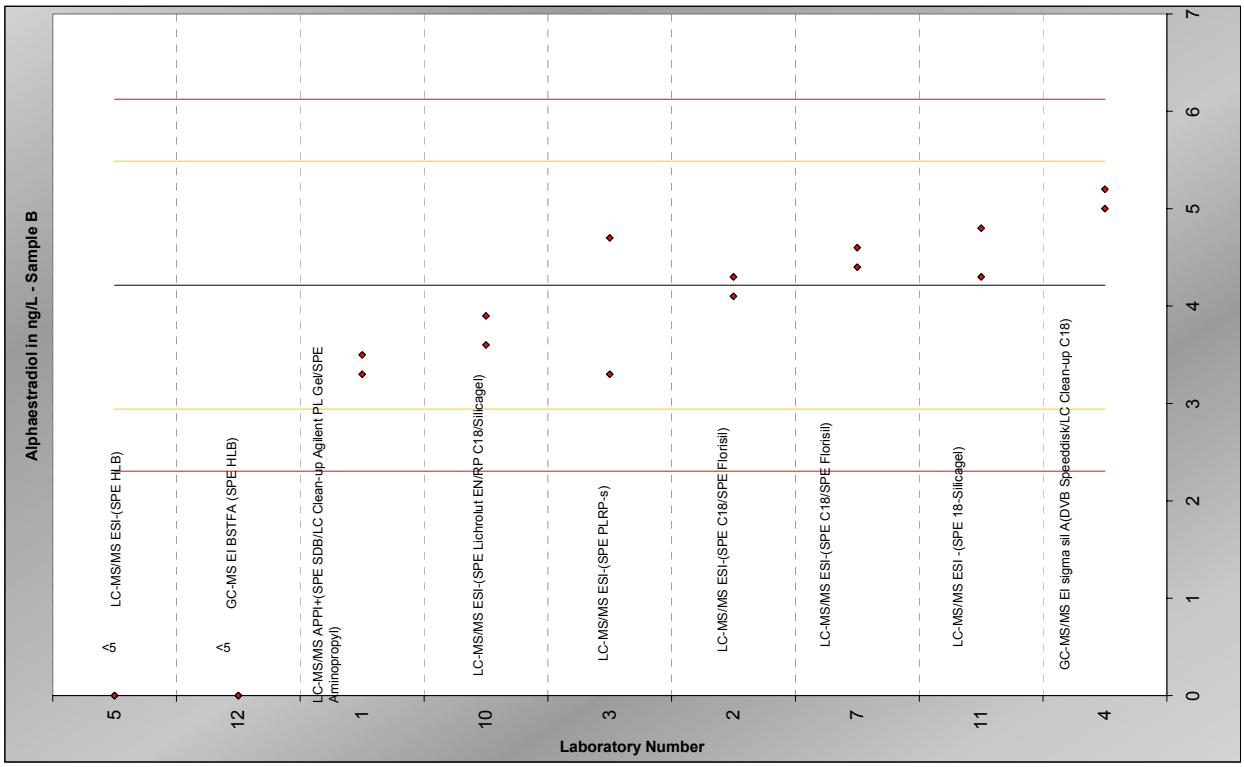
	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F
Mean value M (ng/L ou ng/g)	<LQ	4,21	5,33	0,37	4,51	7,86
Total number of participing laboratories	9	9	7	9	9	5
Number of labs included in statistical tests	0	7	5	3	9	3
Number of labs used for the calculation	0	7	5	3	7	3
Exclusion because of the normality test	/	/	/	/	Lab 2	/
Exclusion because of statistical treatment (Grubbs or Cochran)	/	Lab 3 doubtful for Cochran test but used for calculation	/	/	/	/
Repeatability standard deviation Sr (ng/L or ng/g)	/	0,42	0,46	0	0,67	0,65
Reproductibility standard deviation SR (ng/L or ng/g)	/	0,64	1,67	0,12	1,36	3,39
Repeatability variation coefficient CVr(%)	/	9,9	8,6	0	14,8	8,3
Reproductibility variation coefficient CVR (%)	/	15,1	31,2	31,5	30,2	43
Predominance of inter-lab error	/	1,52	3,64	/	2,04	5,2

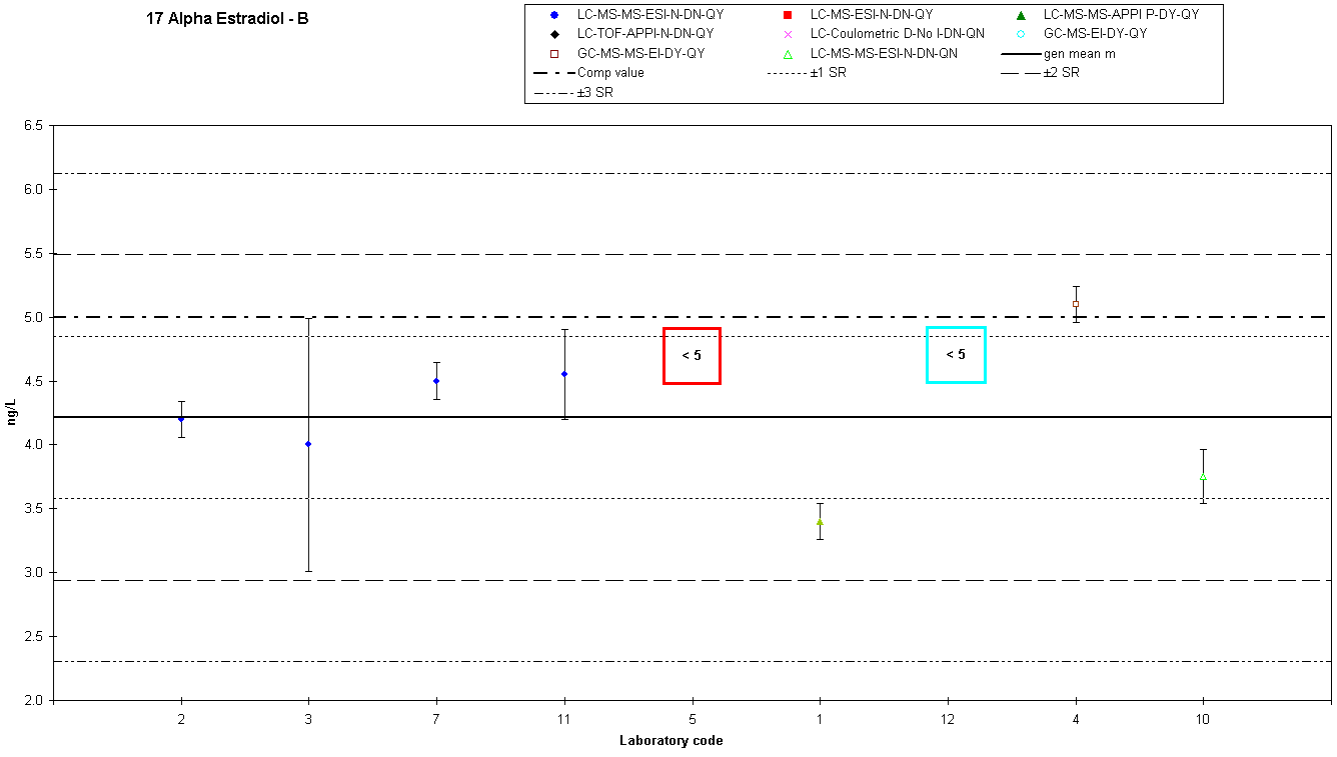
Sample A - Graphs



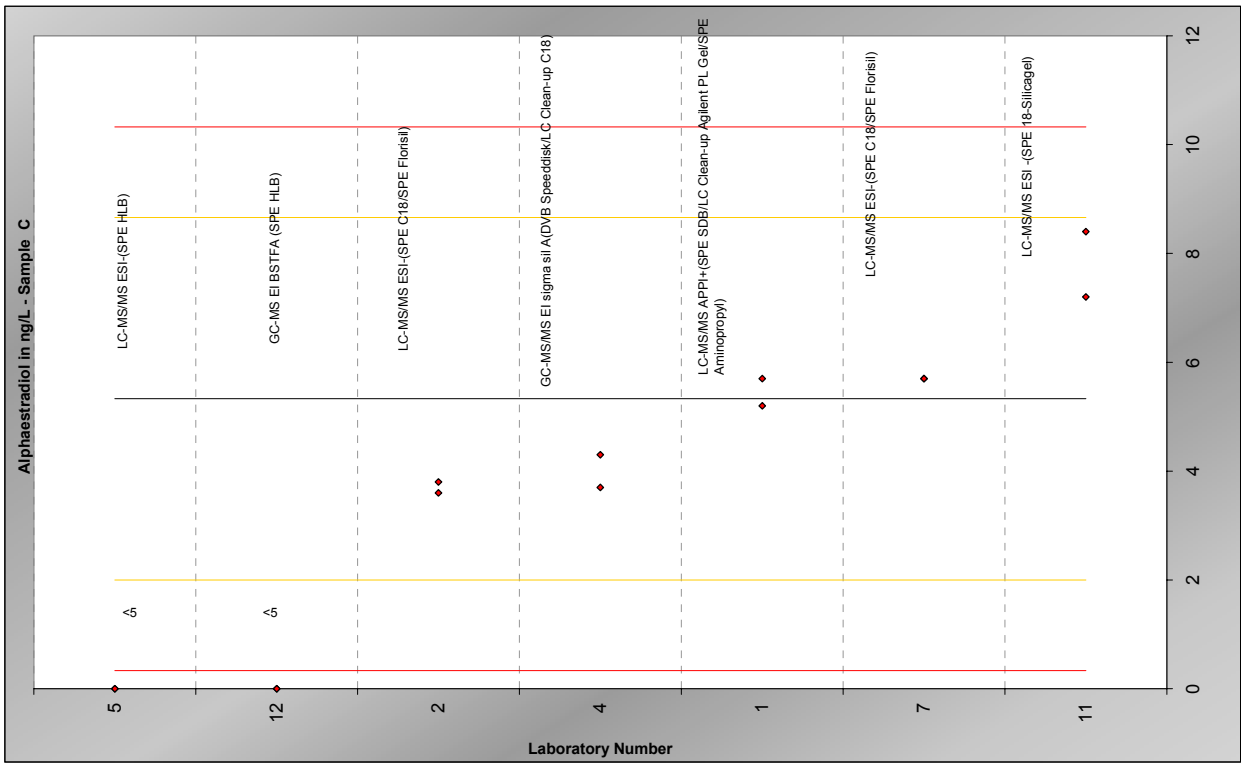


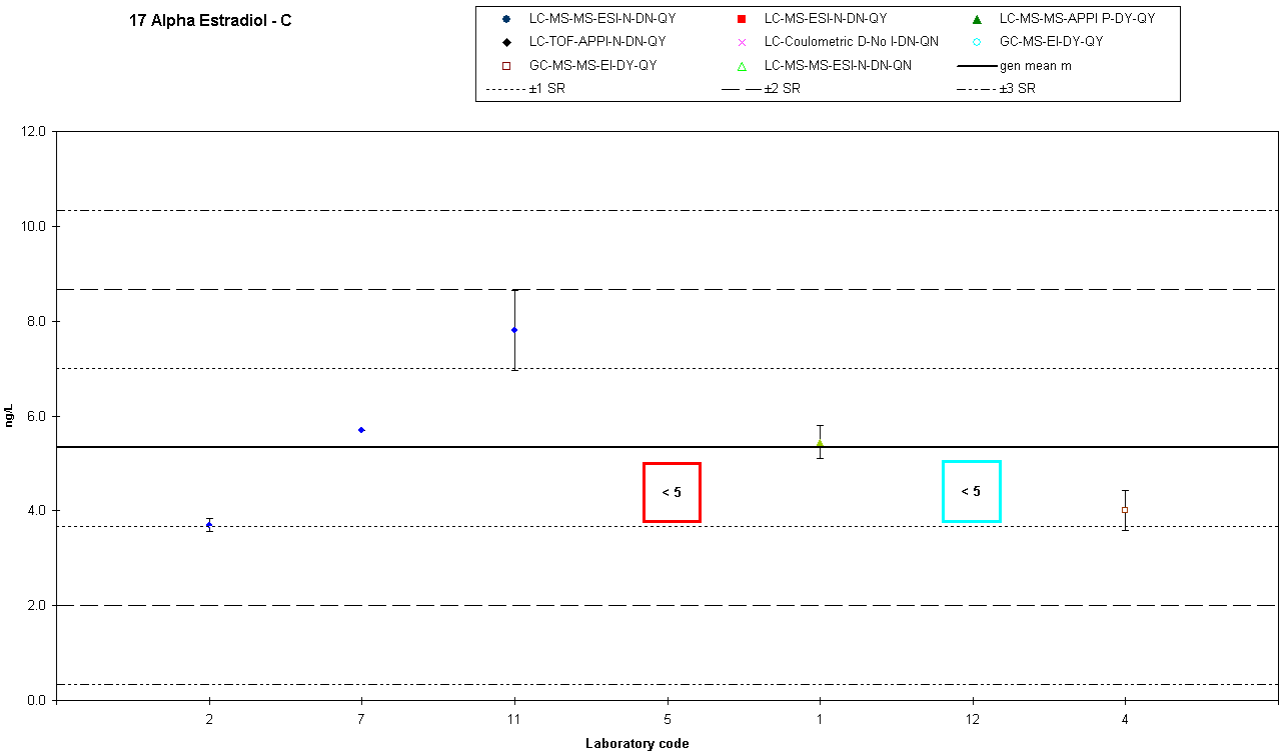
Sample B - Graphs



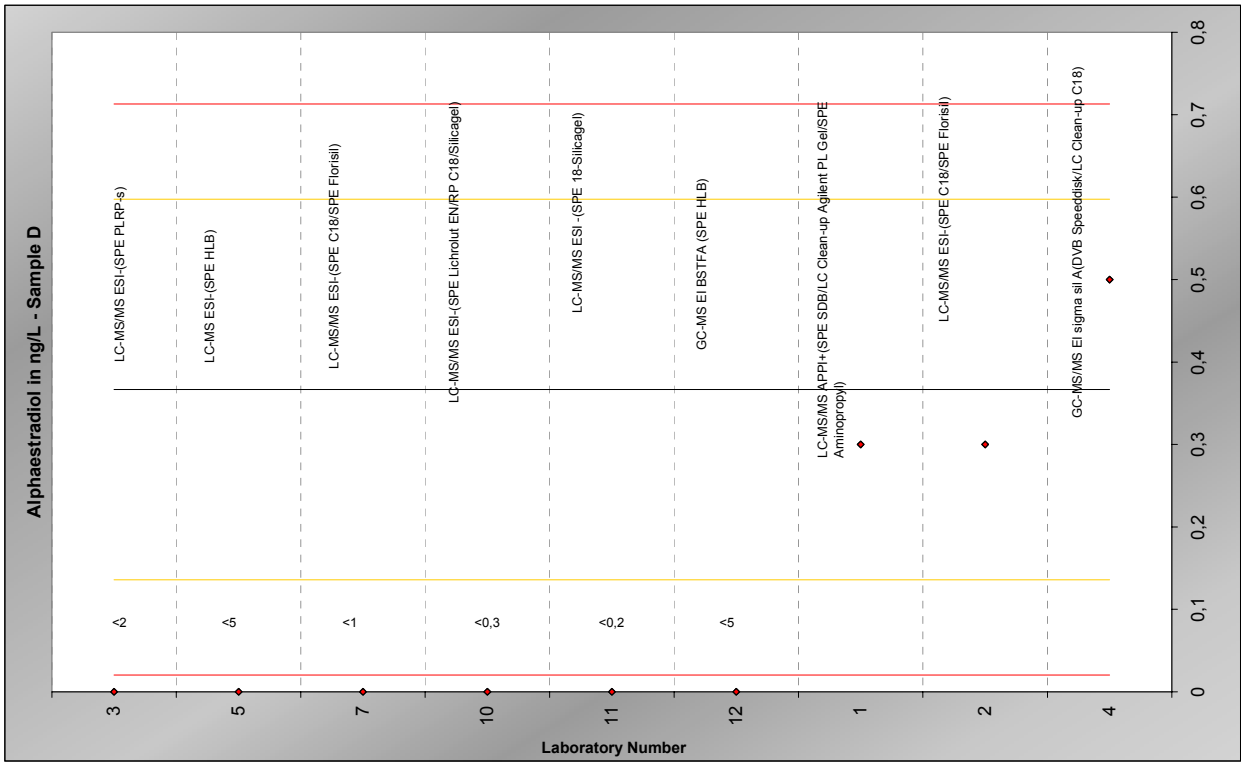


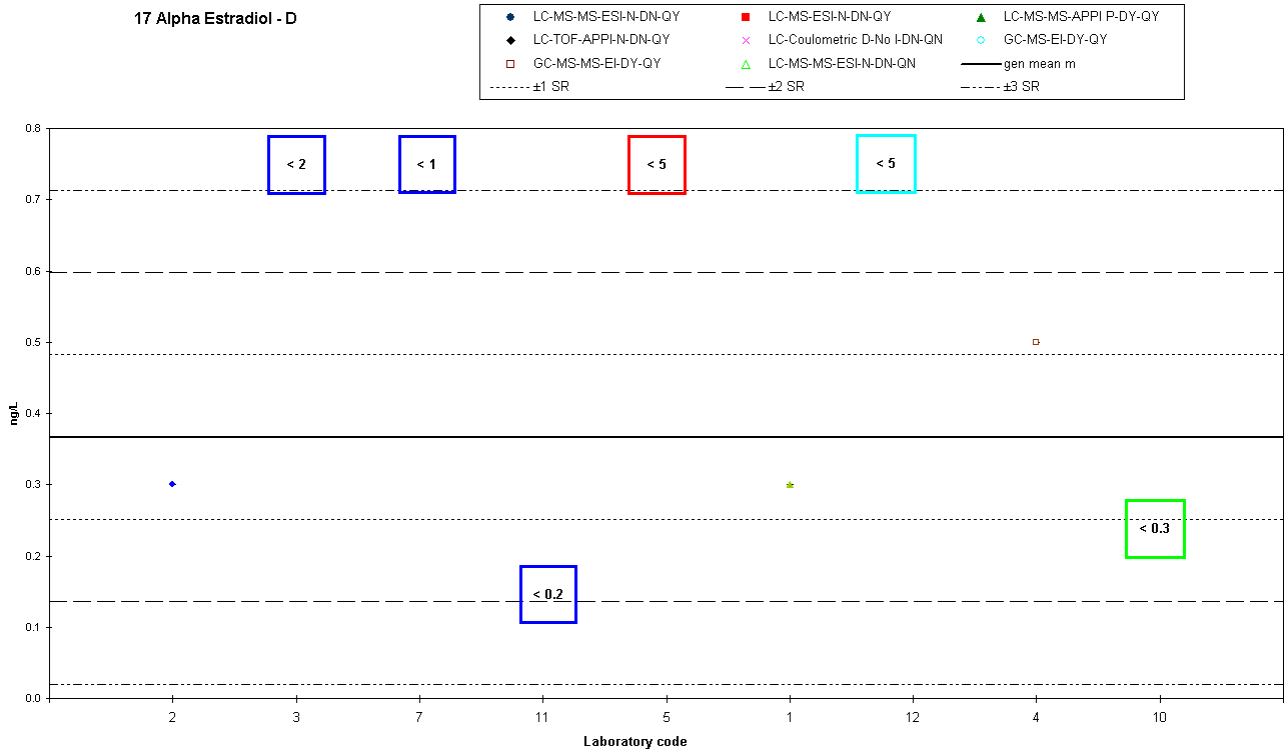
Sample C - Graphs



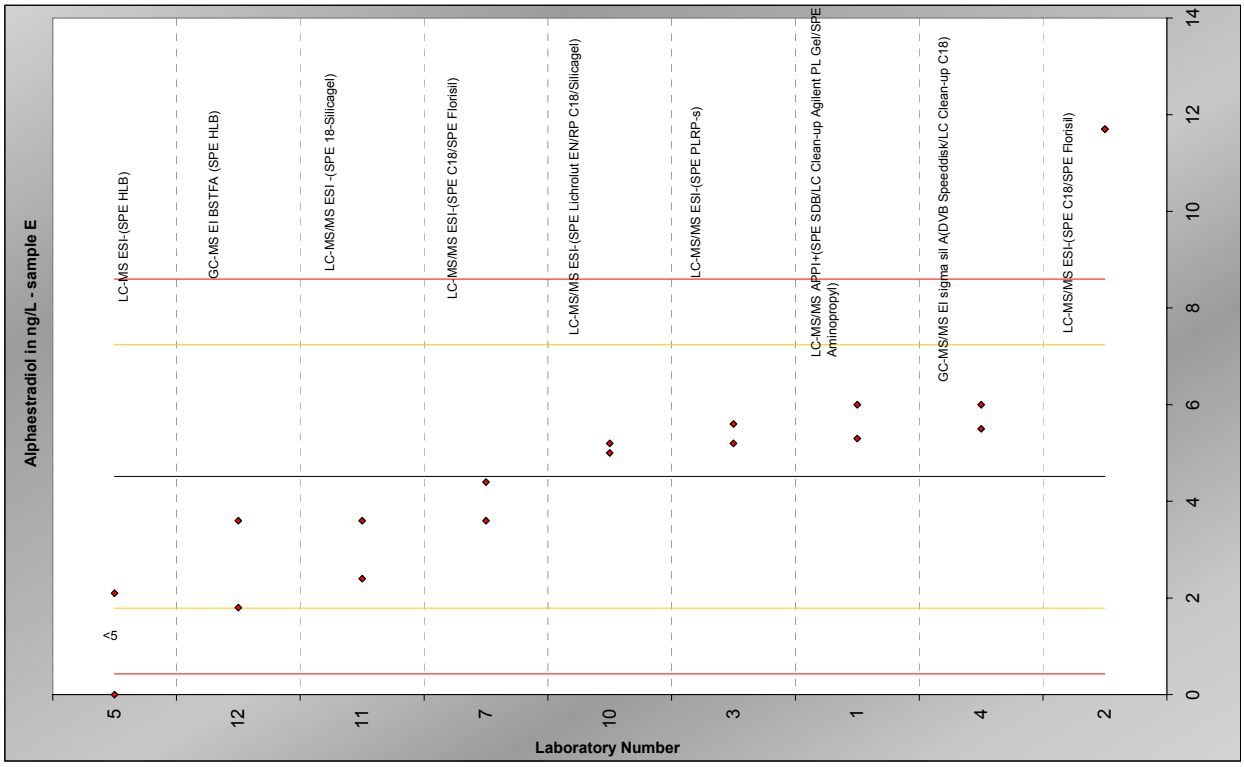


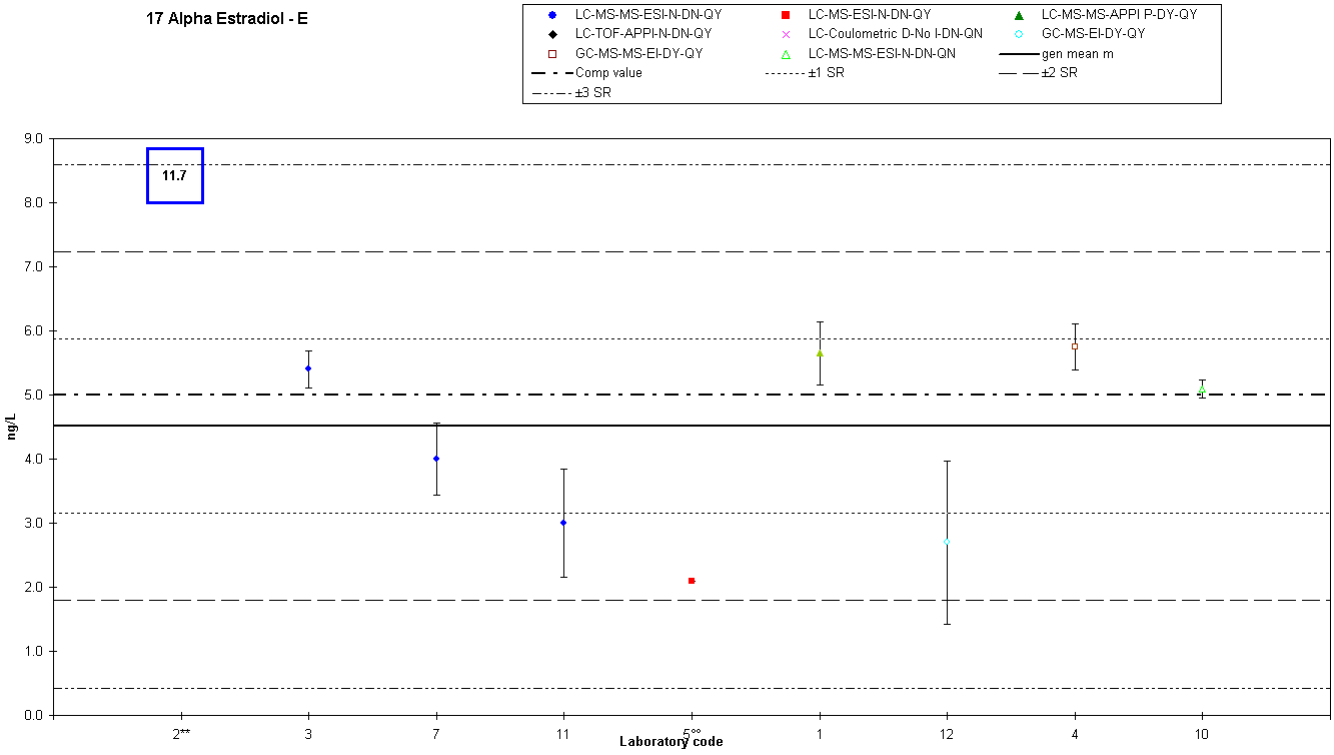
Sample D - Graphs



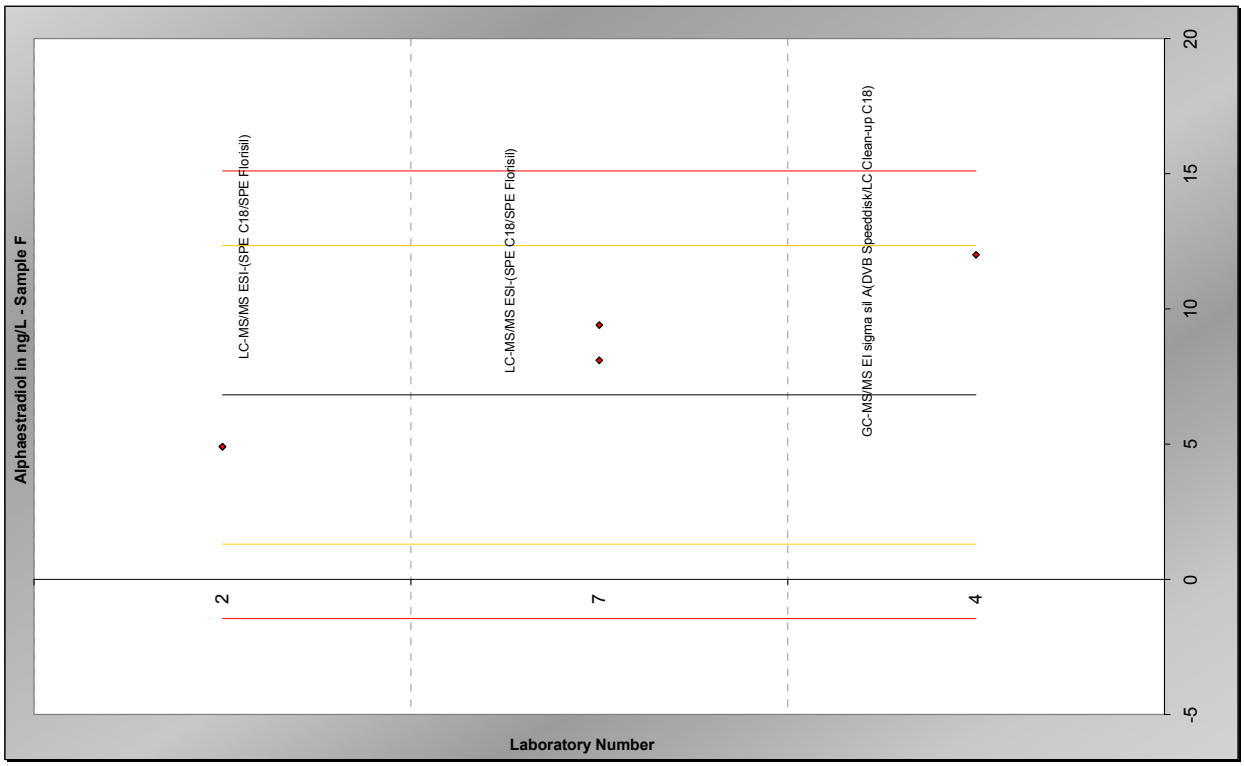


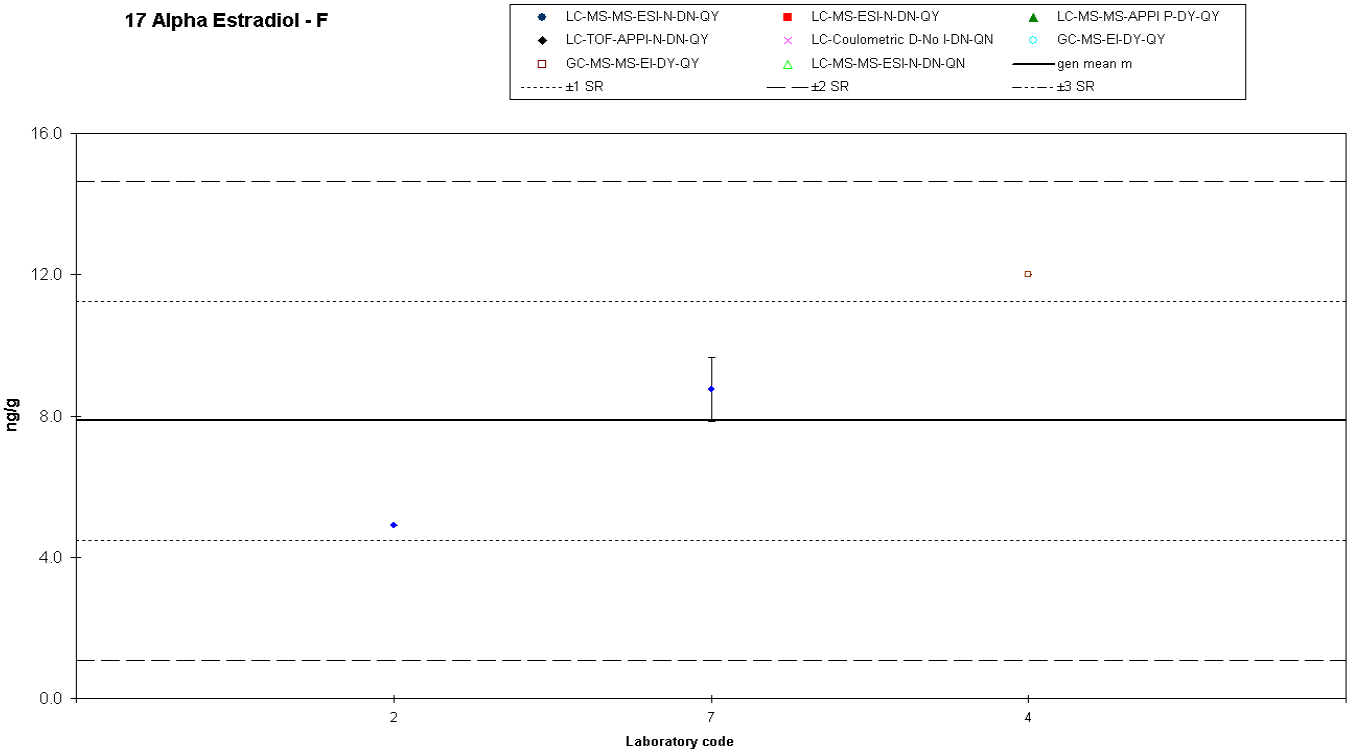
Sample E - Graphs





Sample F - Graphs



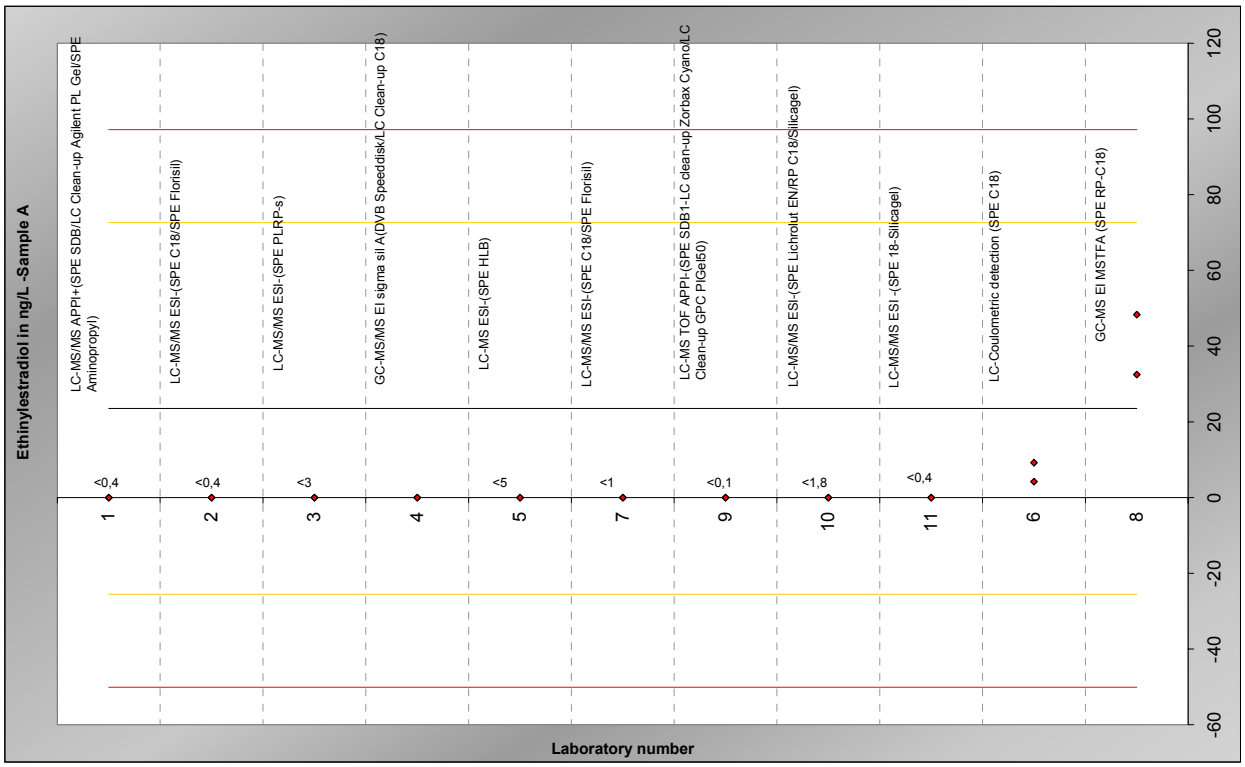


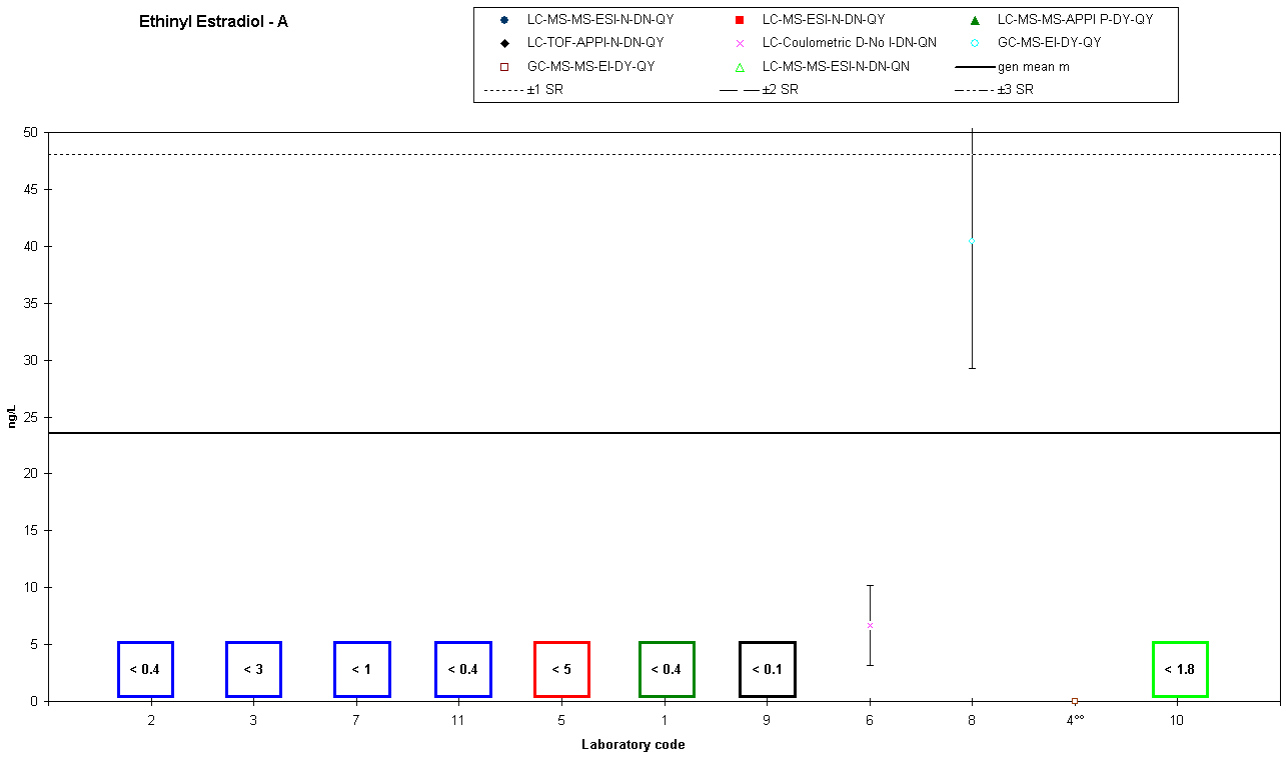
4.3 Ethinylestradiol

4.3.1 Statistical Summary

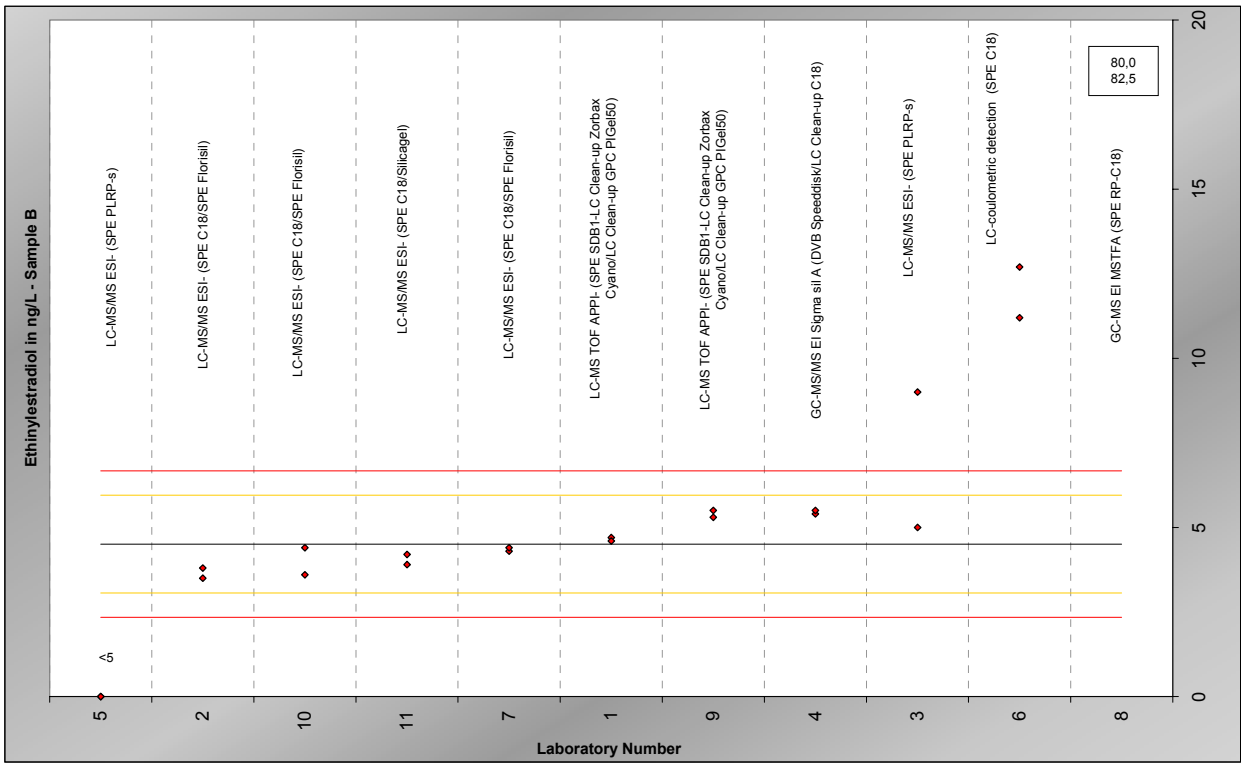
	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F
Mean value M (ng/L ou ng/g)	23,55	4,51	1,38	0,23	4,94	0,73
Total number of participing laboratories	11	11	8	11	11	5
Number of labs included in statistical tests	3	10	7	6	10	2
Number of labs used for the calculation	2	7	4	3	9	2
Exclusion because of the normality test	/	Labs 6 and 8	Lab 8	/	Lab 6	/
Exclusion because of statistical treatment (Grubbs or Cochran)	/	Lab 3 aberrant for Cochran test	/	Lab 8 aberrant for Grubbs test	/	/
Repeatability standard deviation Sr (ng/L or ng/g)	8,29	0,25	0,10	0	0,28	0,14
Reproductibility standard deviation SR (ng/L or ng/g)	24,54	0,72	0,23	0,06	0,79	0,78
Repeatability variation coefficient CVr(%)	35,2	5,6	8,2	0	5,7	19
Reproductibility variation coefficient CVR (%)	104,2	16	19	24,7	15,9	106
Predominance of inter-lab error	2,96	2,86	2	/	2,81	5,5

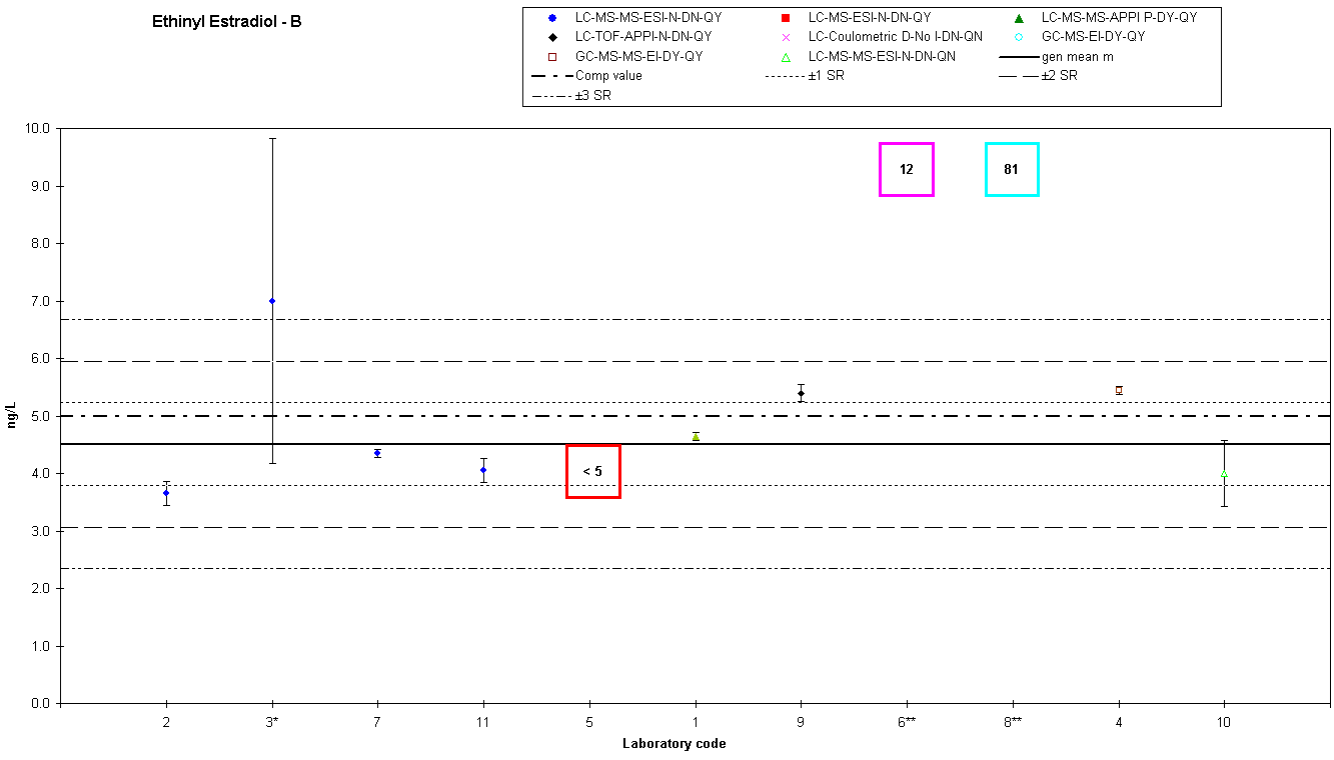
Sample A - Graphs



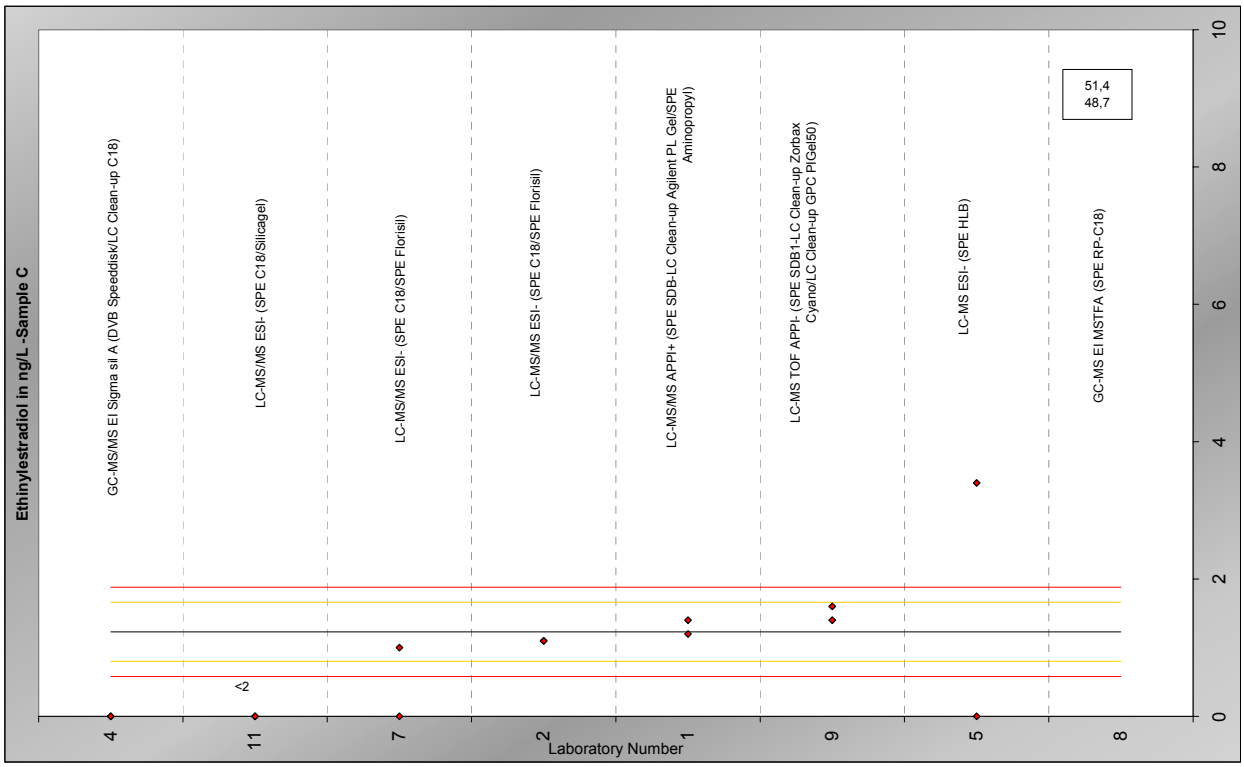


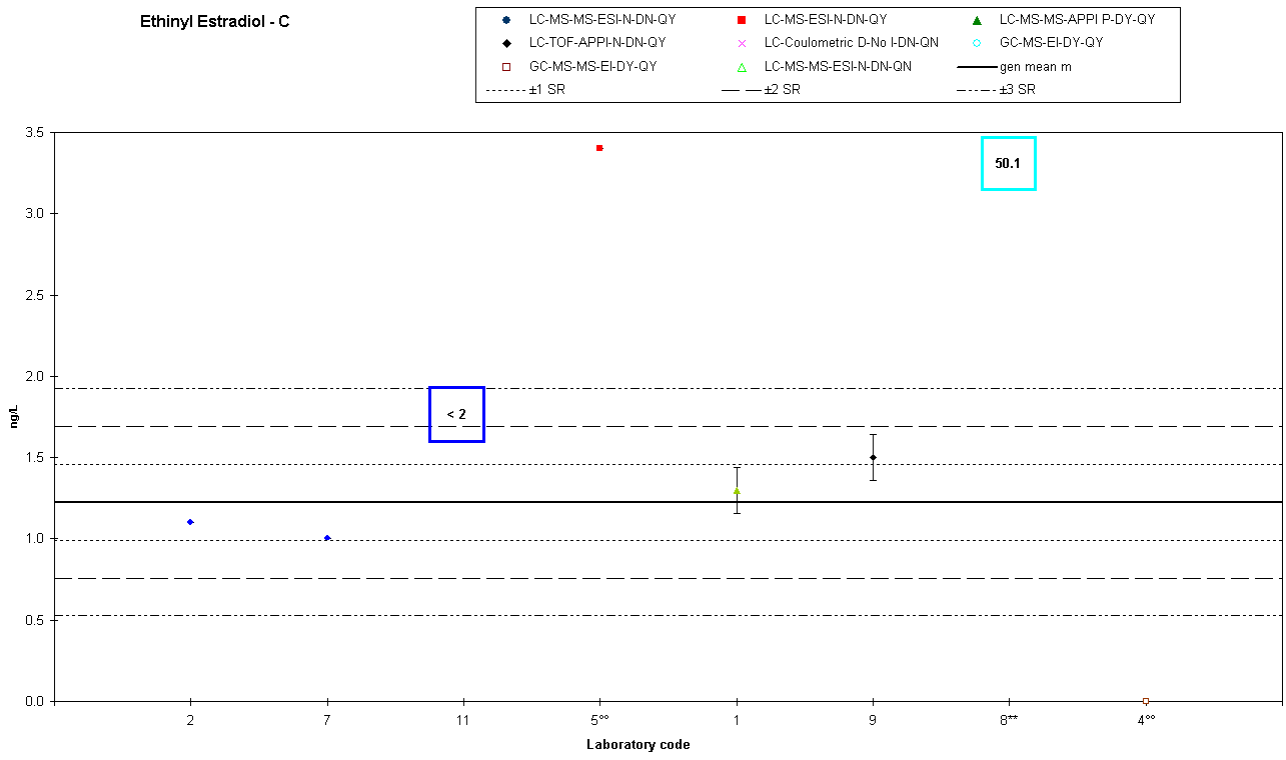
Sample B - Graphs



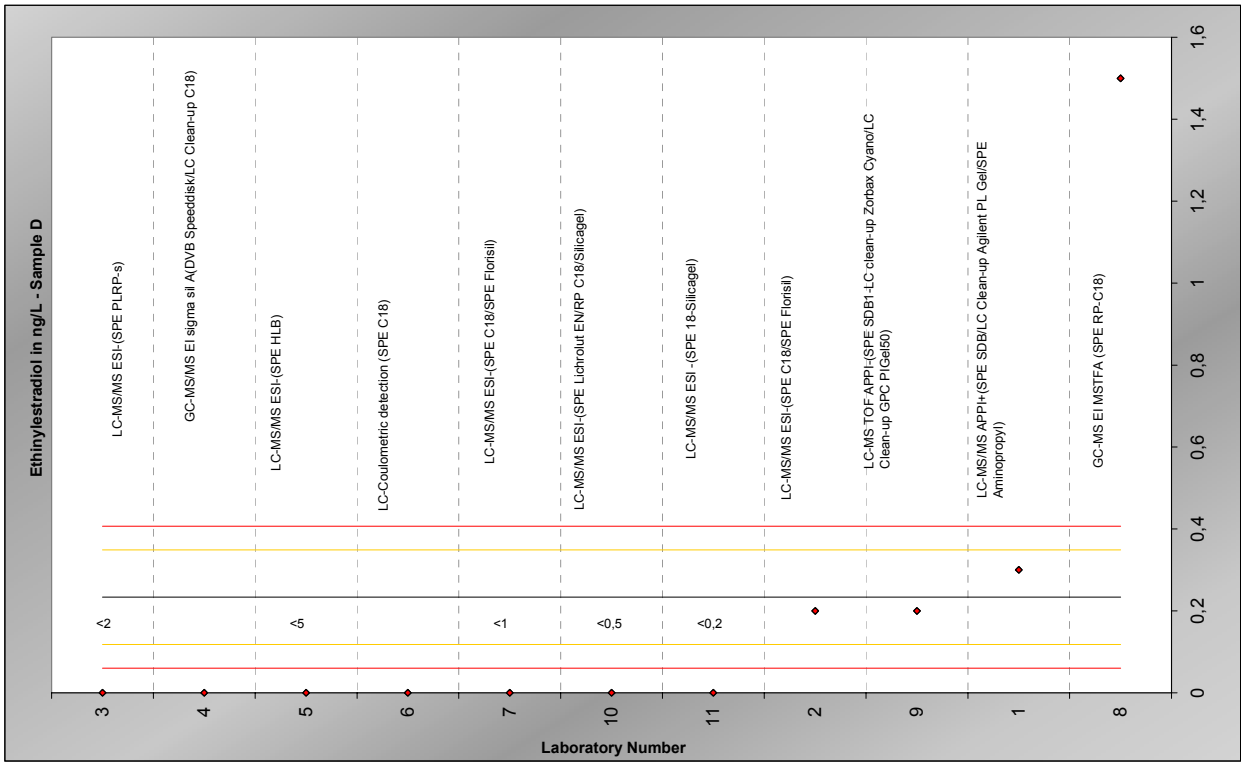


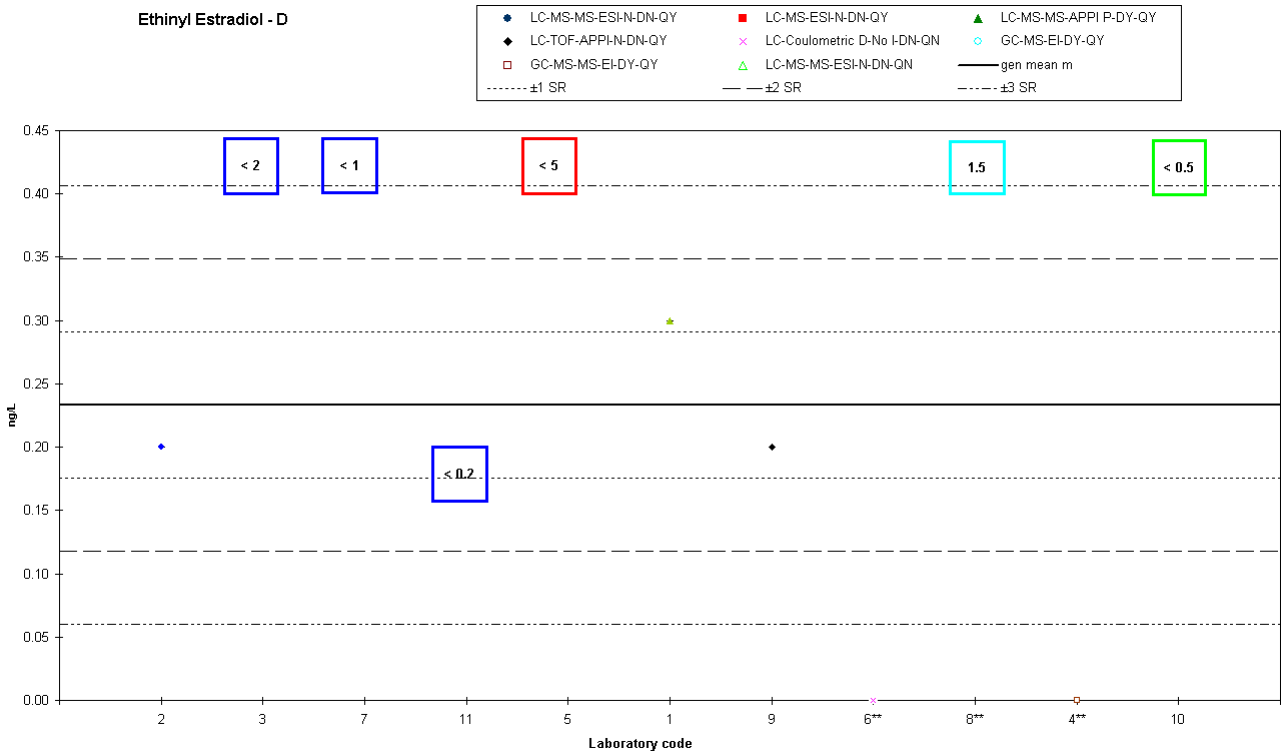
Sample C - Graphs



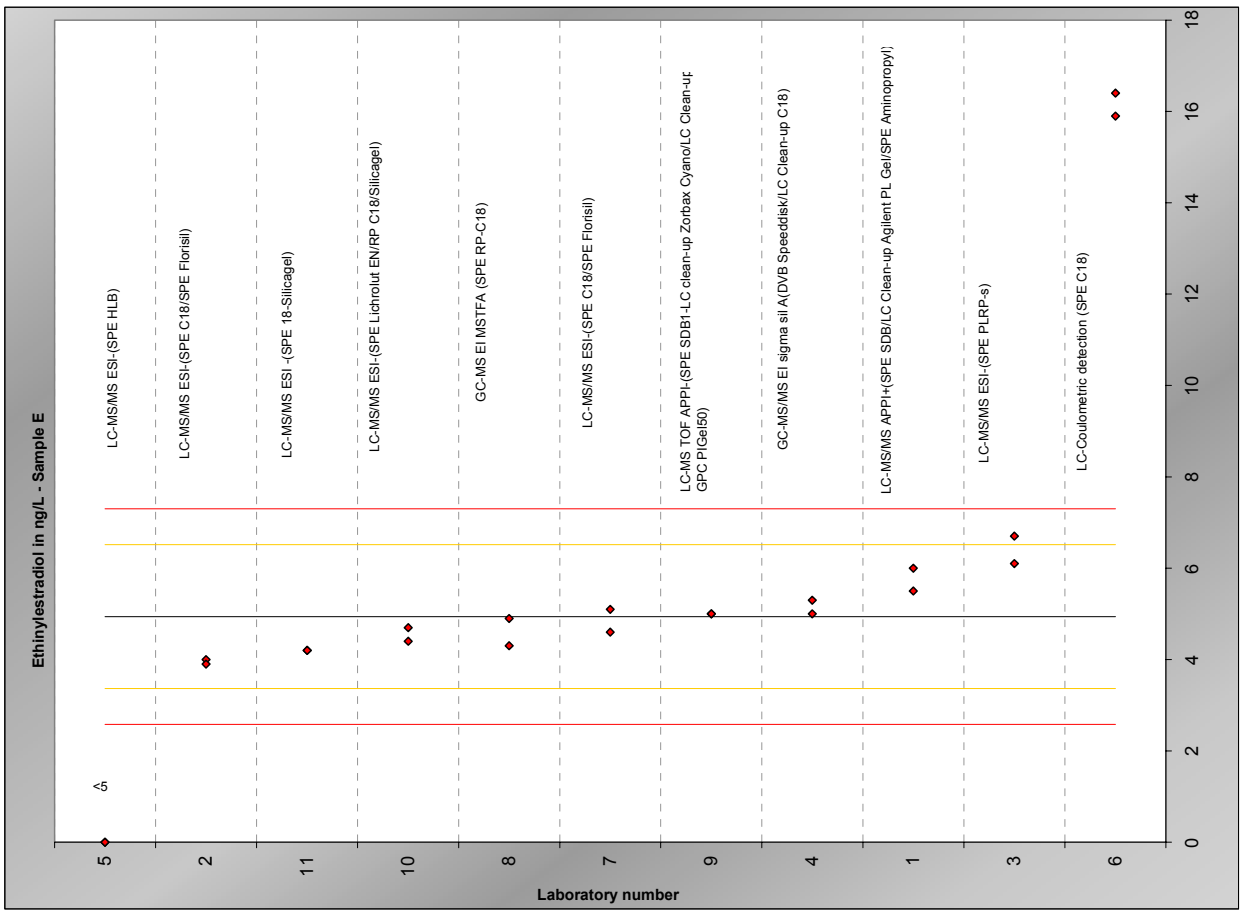


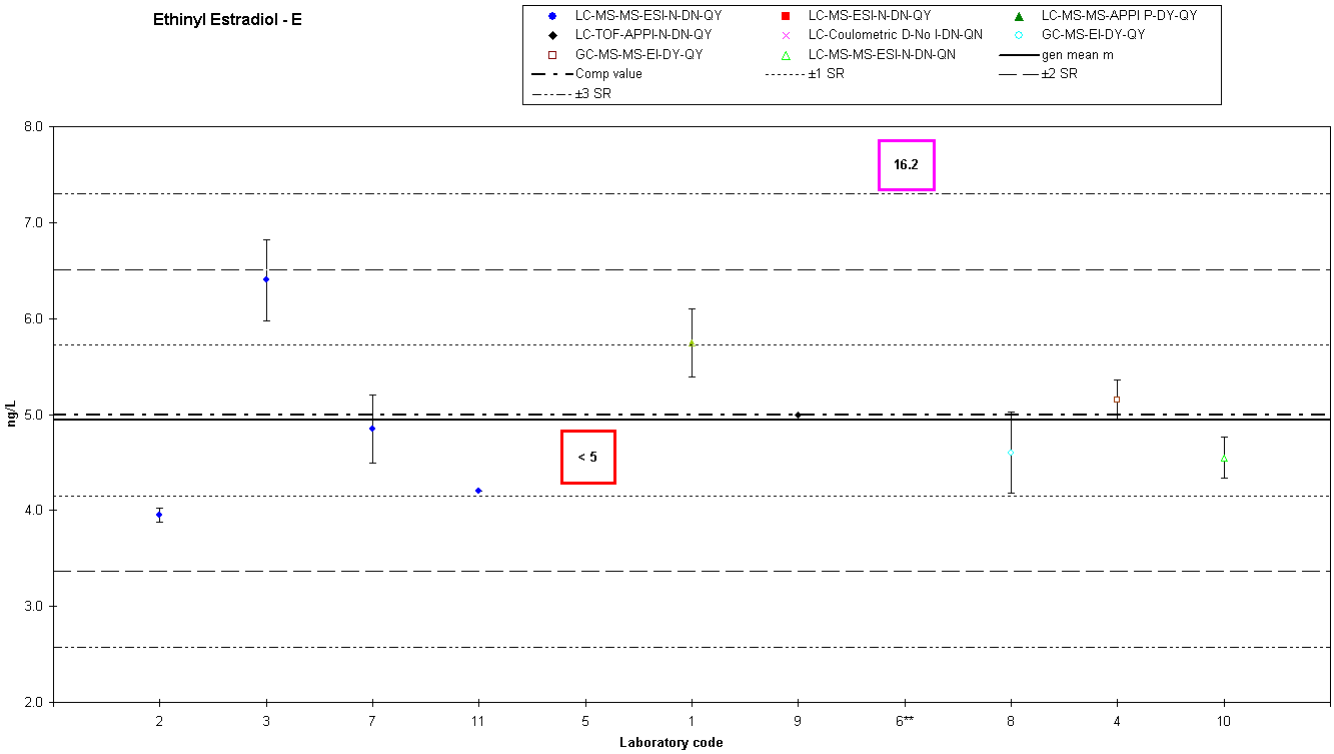
Sample D - Graphs



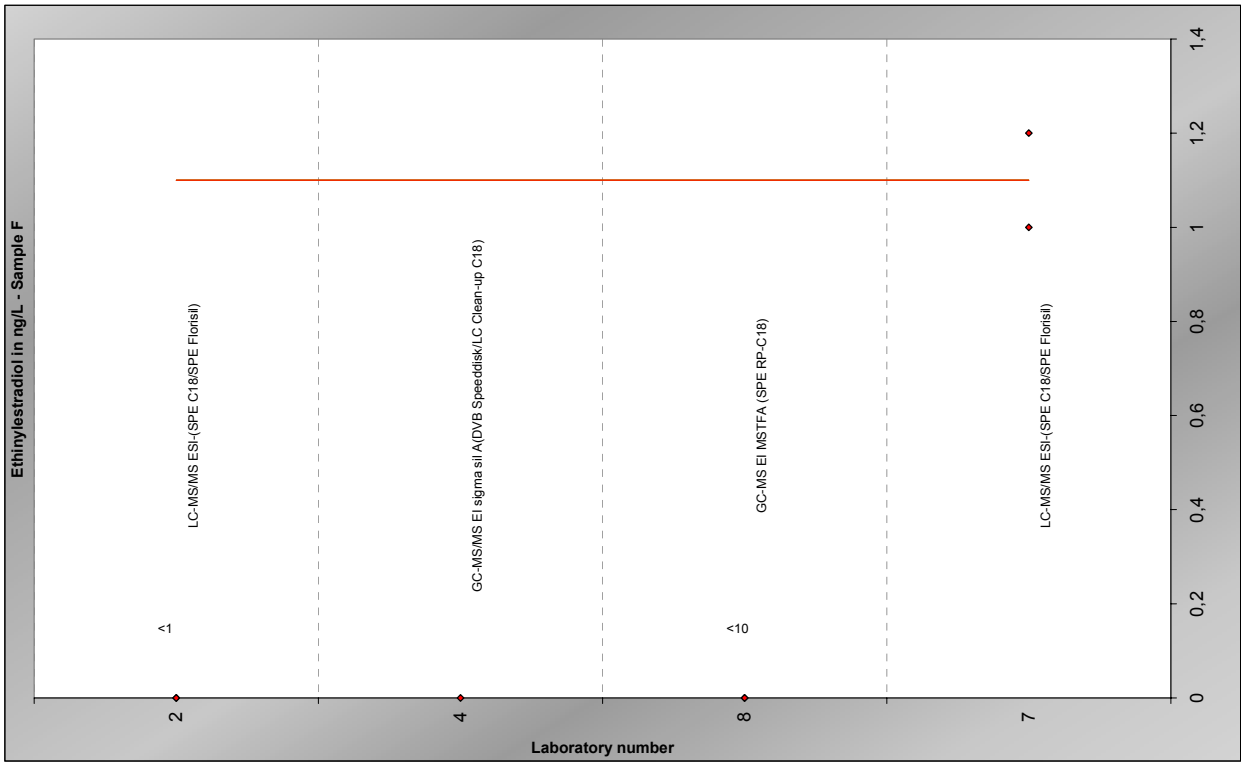


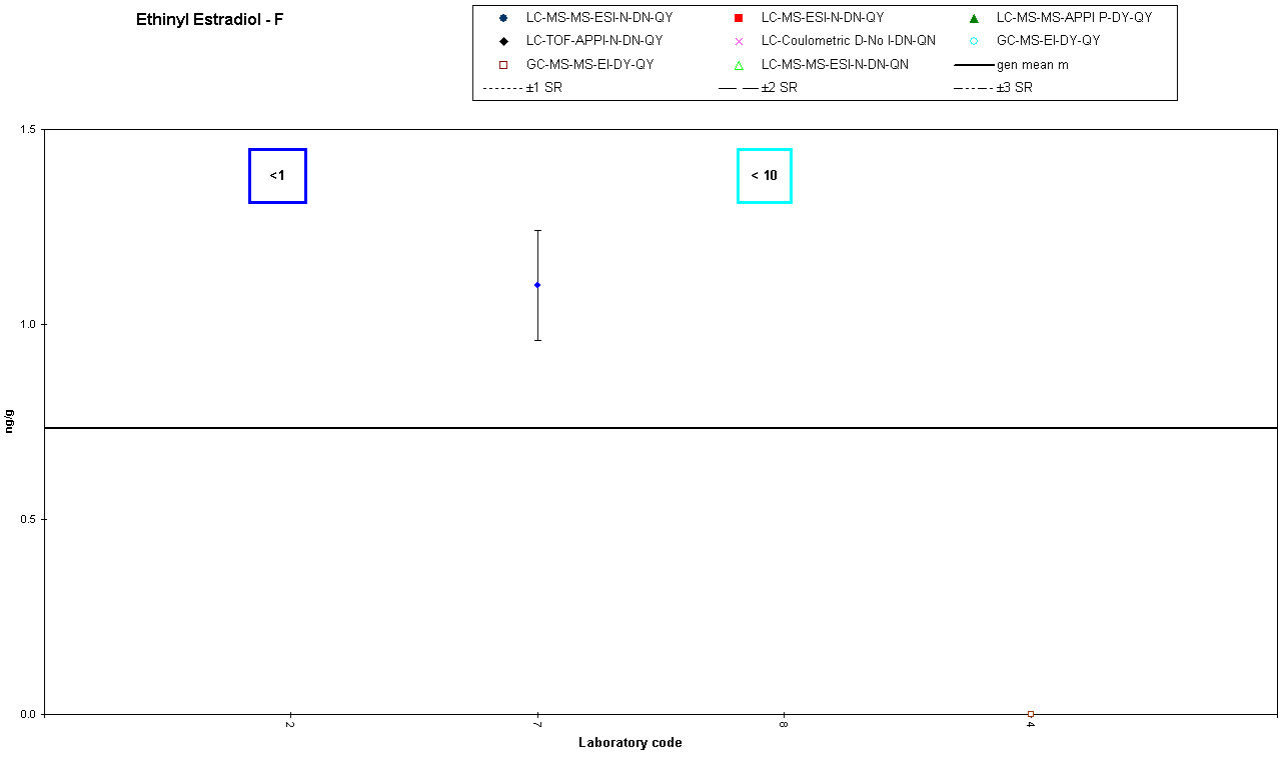
Sample E - Graphs





Sample F - Graphs



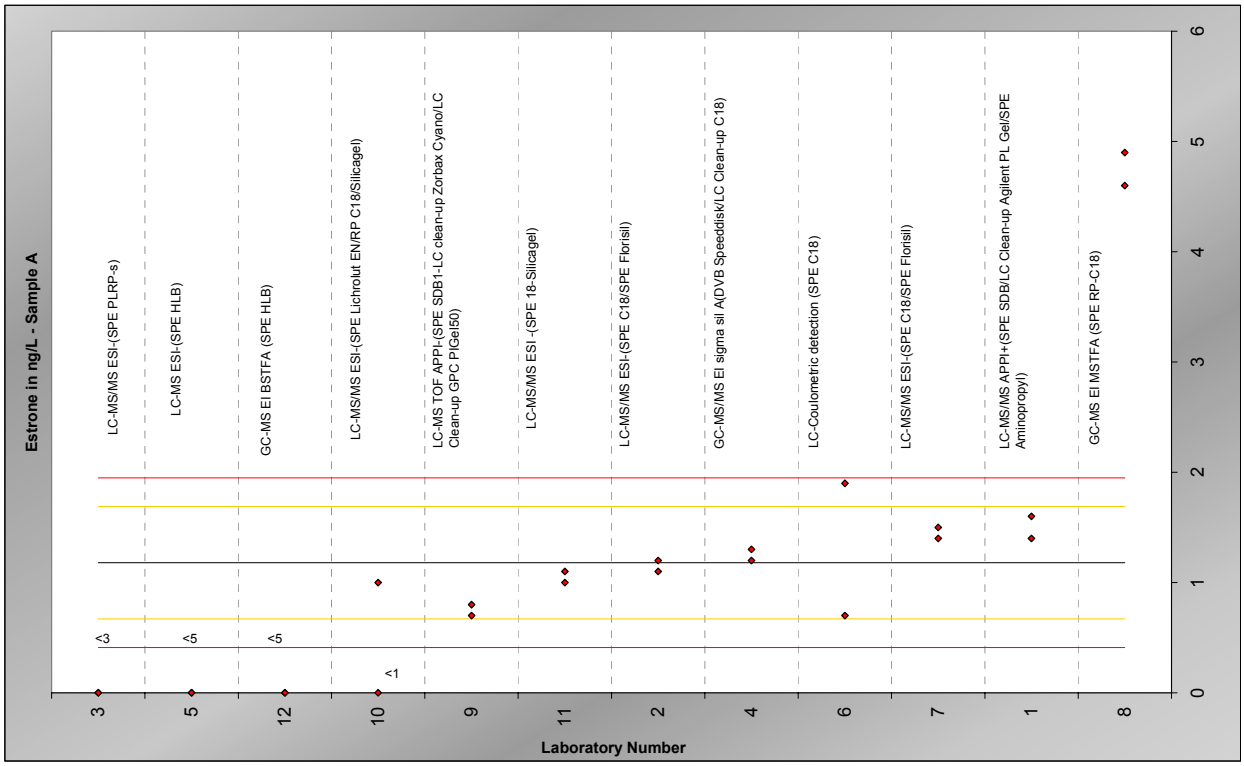


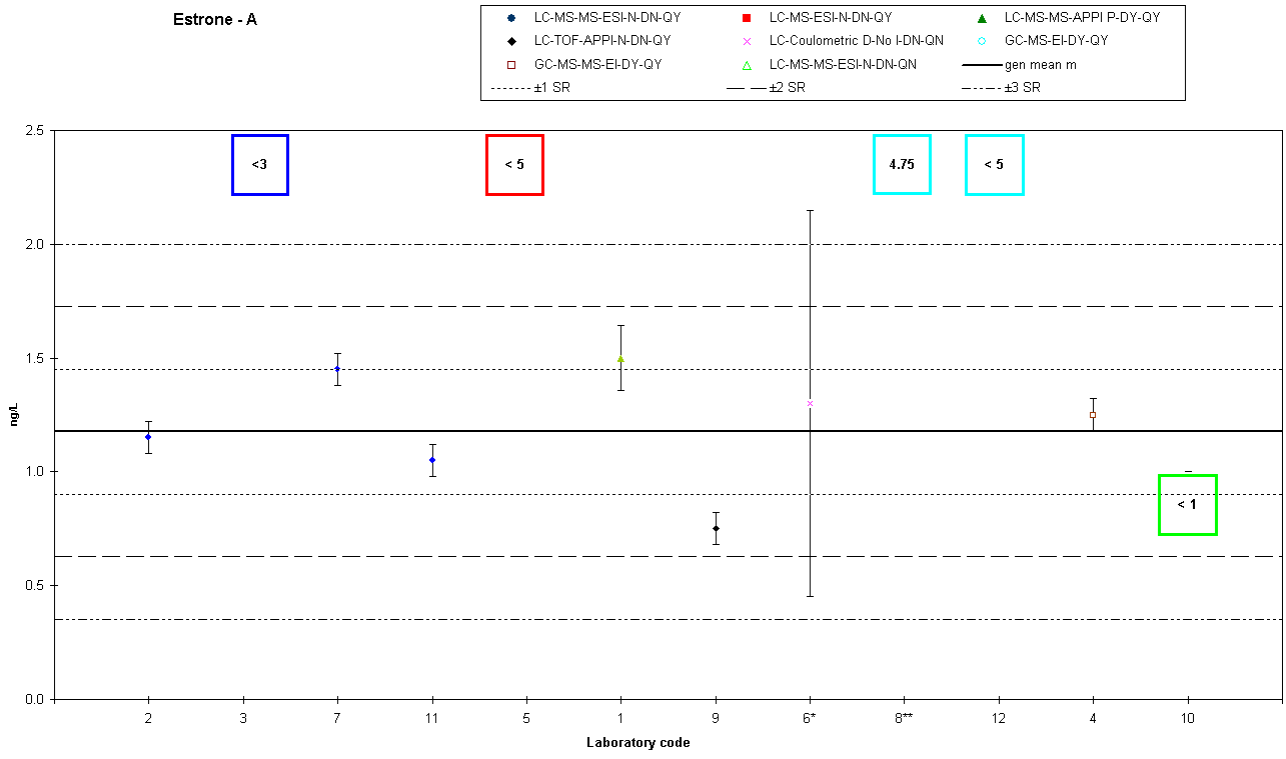
4.4 Estrone

4.4.1 Statistical Summary

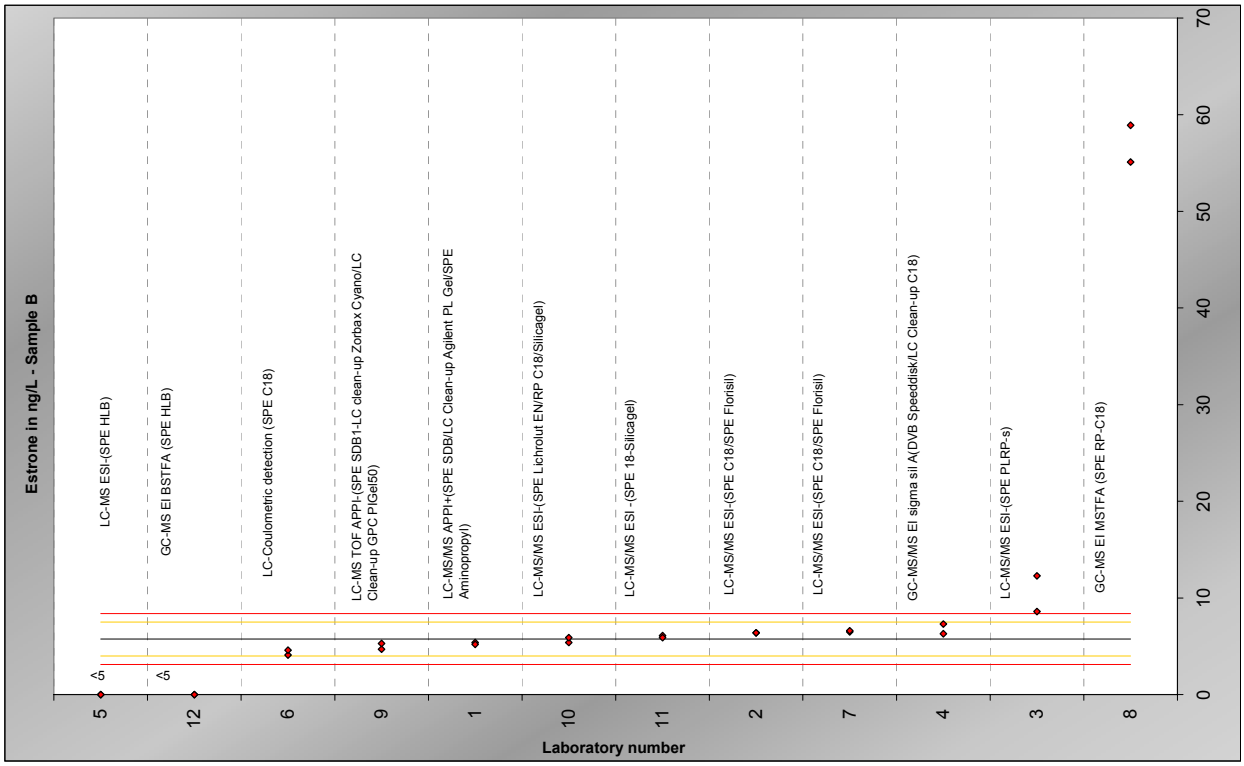
	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F
Mean value M (ng/L ou ng/g)	1,18	5,76	55,11	6,89	4,48	91,20
Total number of participing laboratories	12	12	9	12	12	6
Number of labs included in statistical tests	8	10	7	11	11	4
Number of labs used for the calculation	6	8	7	9	10	4
Exclusion because of the normality test	Lab 8	Lab 8	/	Lab 6	Lab 6	/
Exclusion because of statistical treatment (Grubbs or Cochran)	Lab 6 aberrant for Cochran test	Lab 3 aberrant for Cochran and grubbs tests	/	Lab 12 aberrant for Cochran test	/	/
Repeatability standard deviation Sr (ng/L or ng/g)	0,09	0,35	2,54	0,42	0,35	4,4
Reproductibility standard deviation SR (ng/L or ng/g)	0,27	0,88	6,01	1,05	1,3	48,4
Repeatability variation coefficient CVr(%)	7,4	6,1	4,6	6,1	7,9	4,8
Reproductibility variation coefficient CVR (%)	23,3	15,2	10,9	15,2	29	53
Predominance of inter-lab error	3	2,51	2,37	2,5	3,67	11

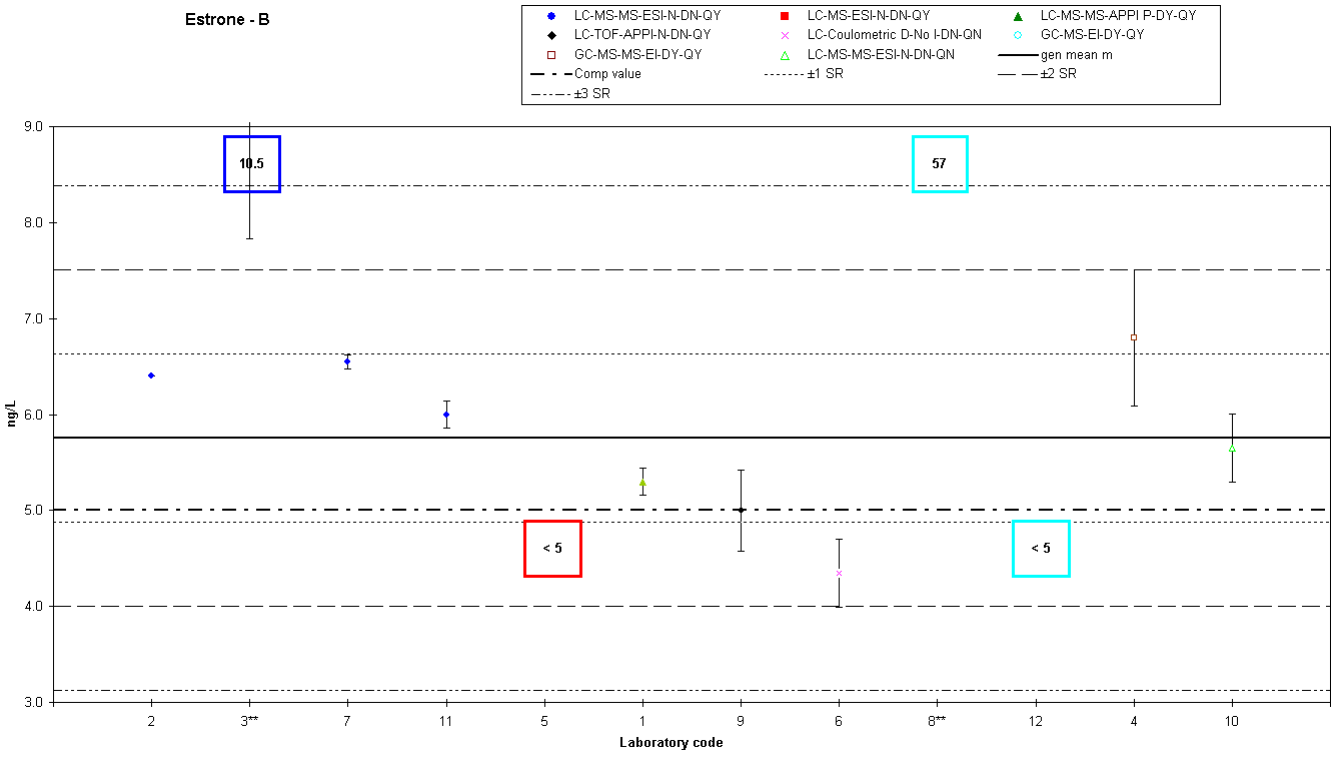
Sample A - Graphs



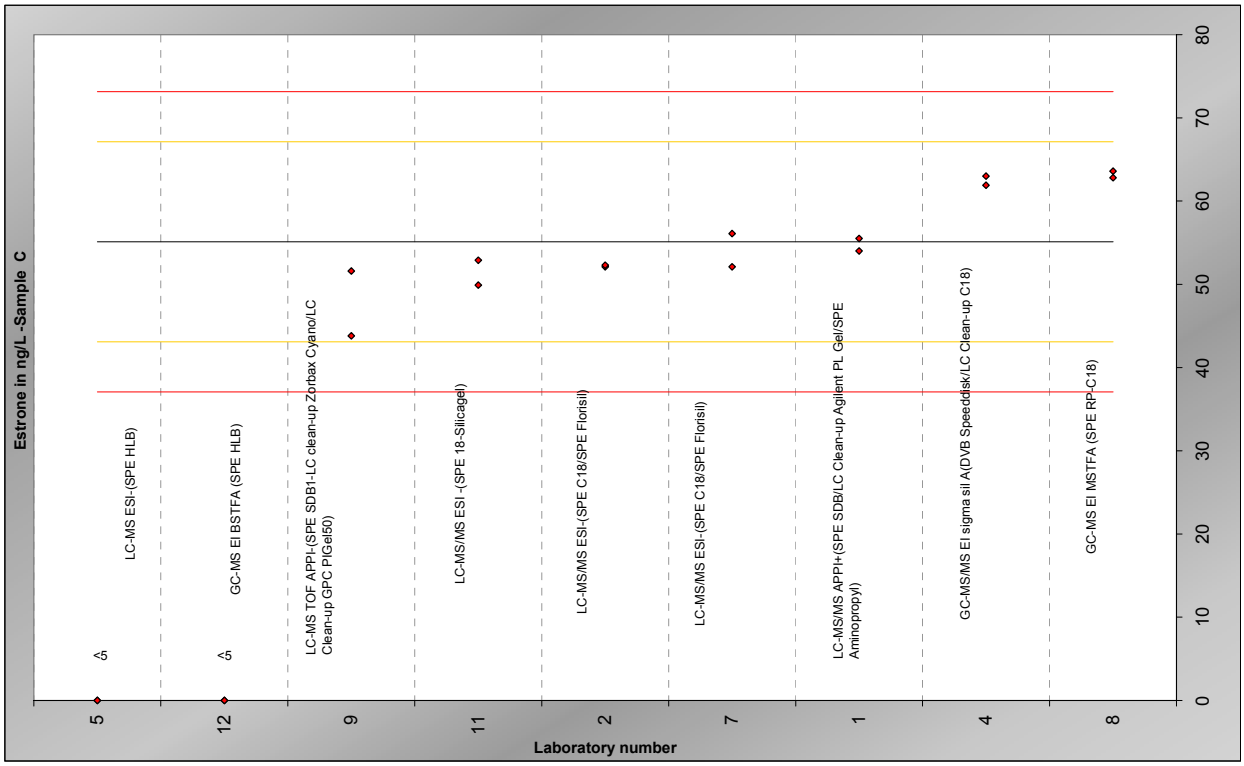


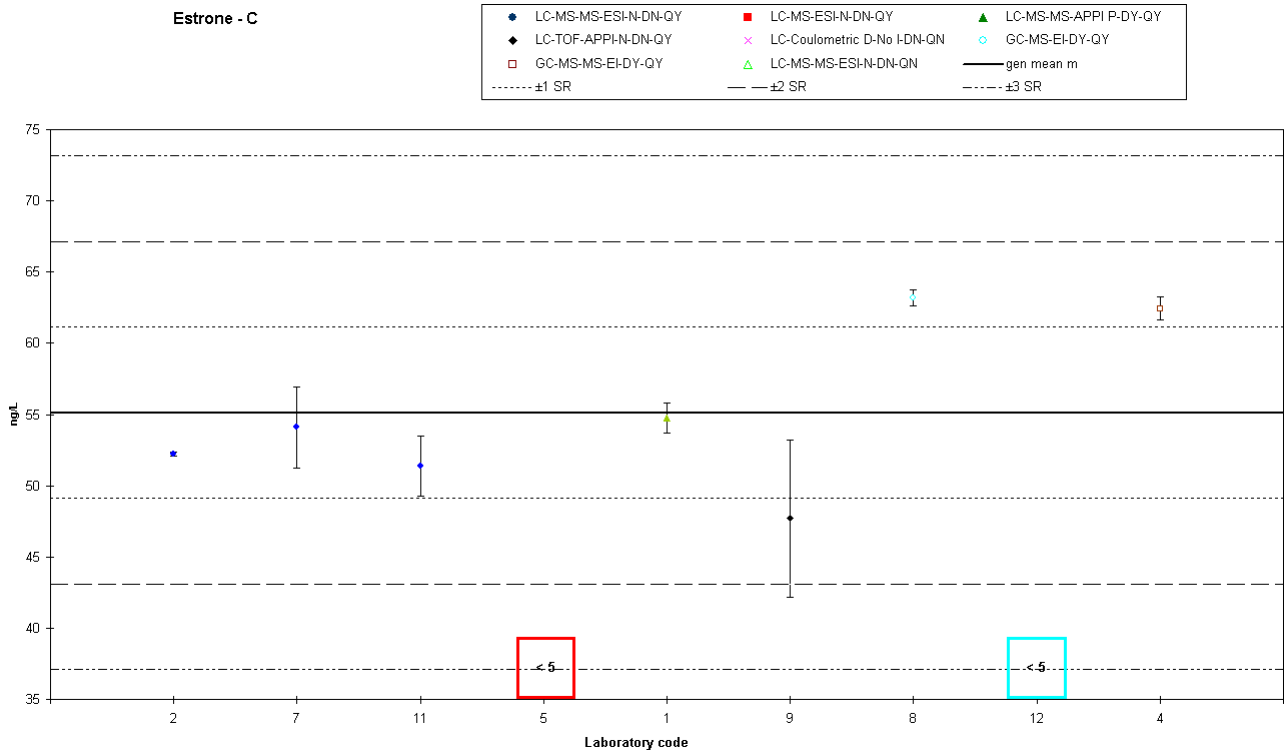
Sample B - Graphs



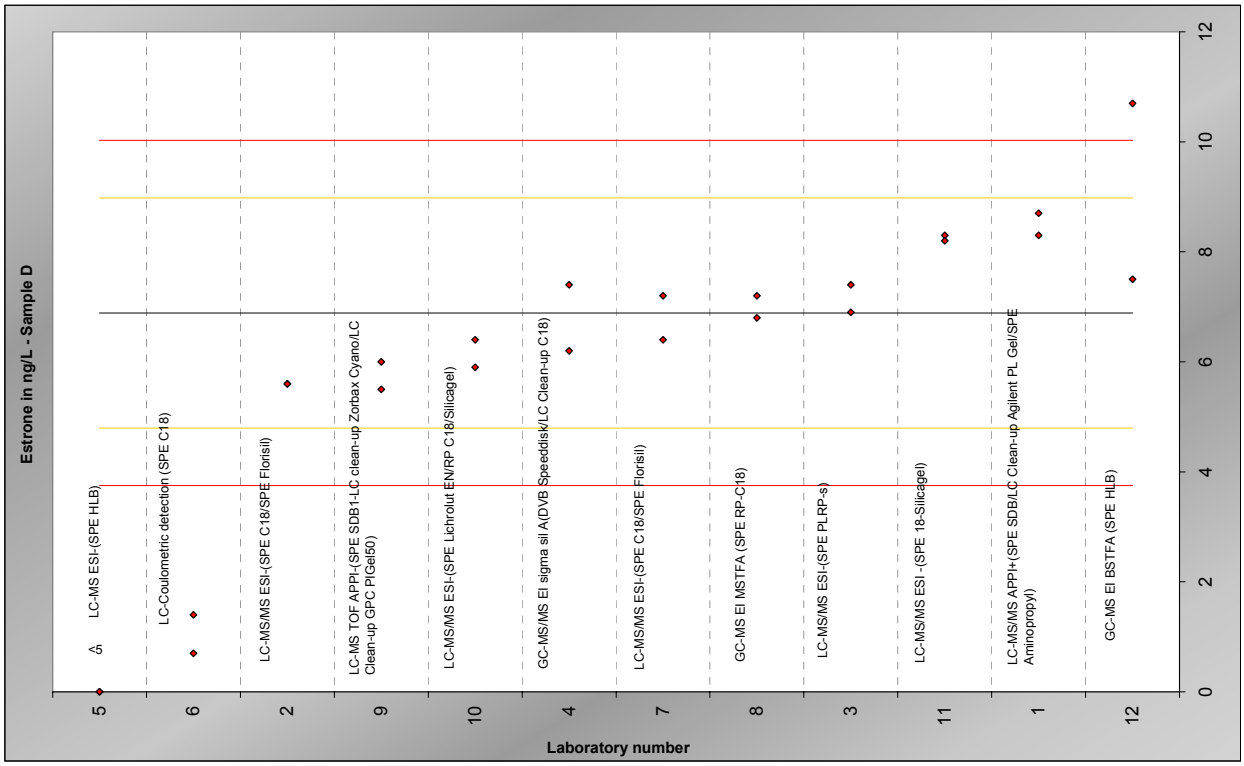


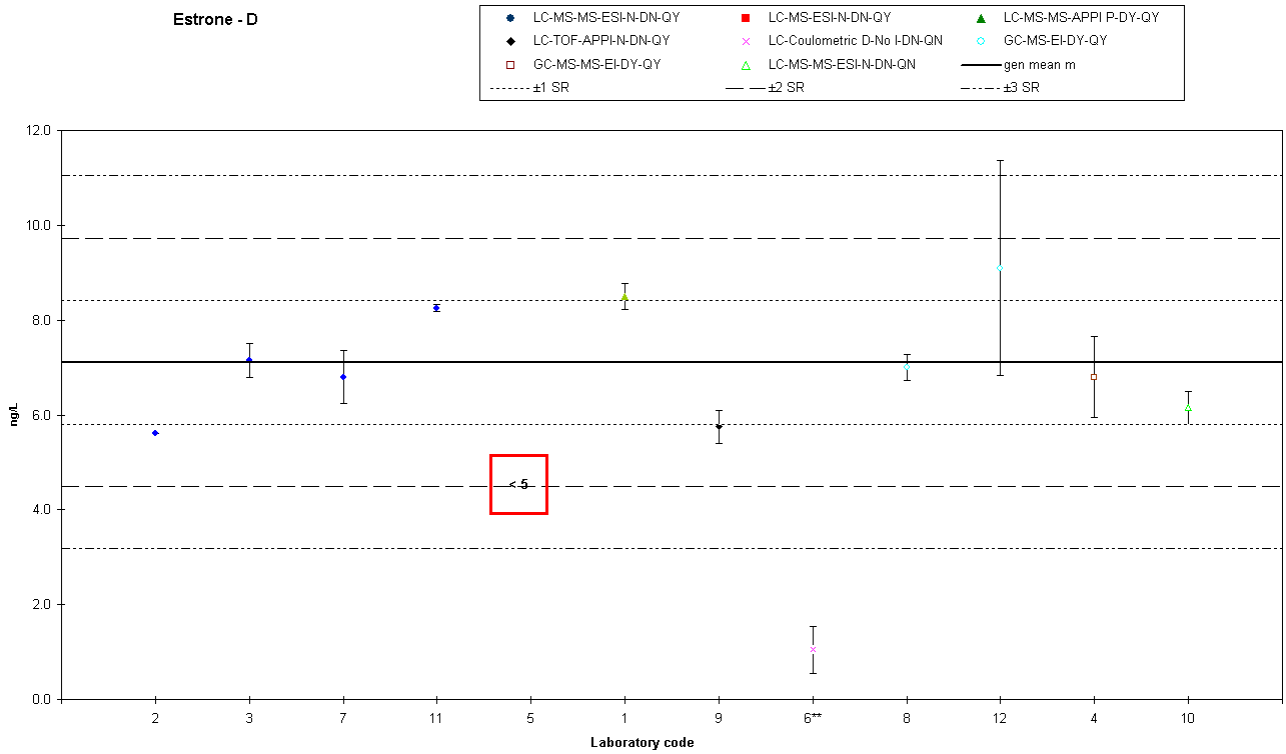
Sample C - Graphs



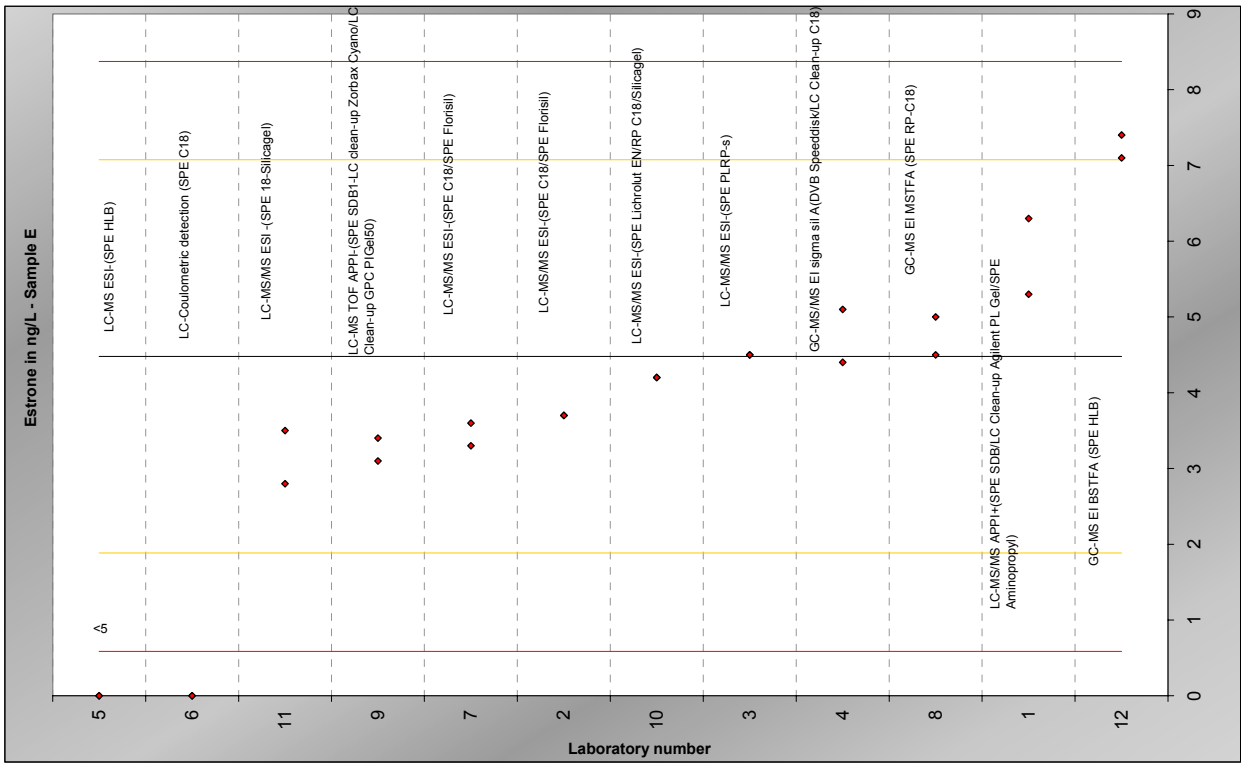


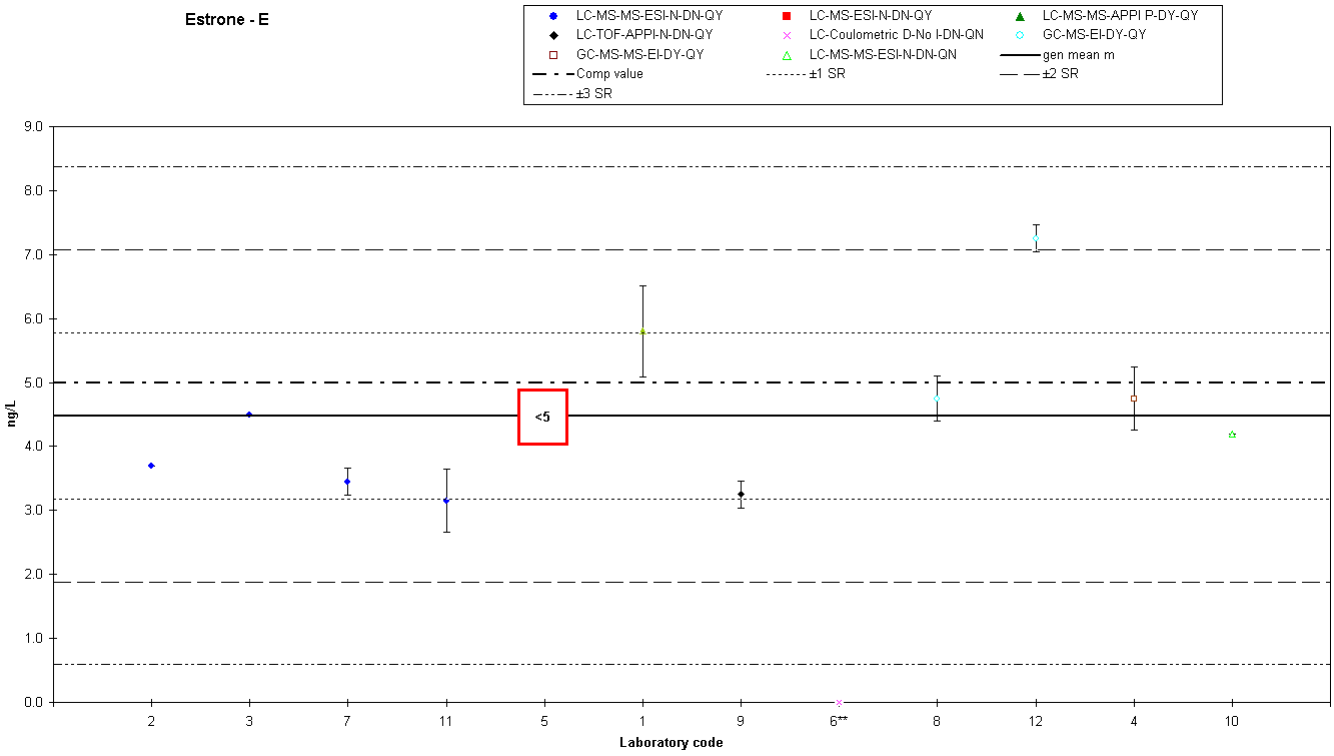
Sample D - Graphs



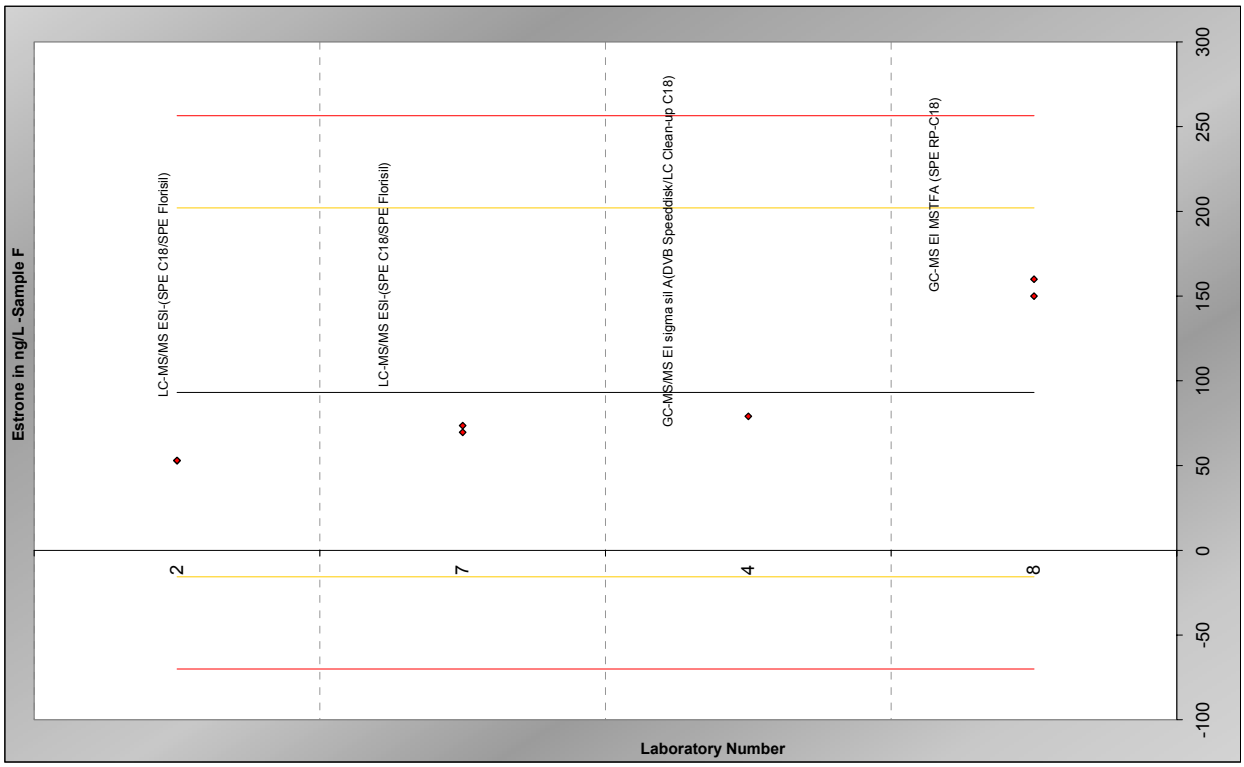


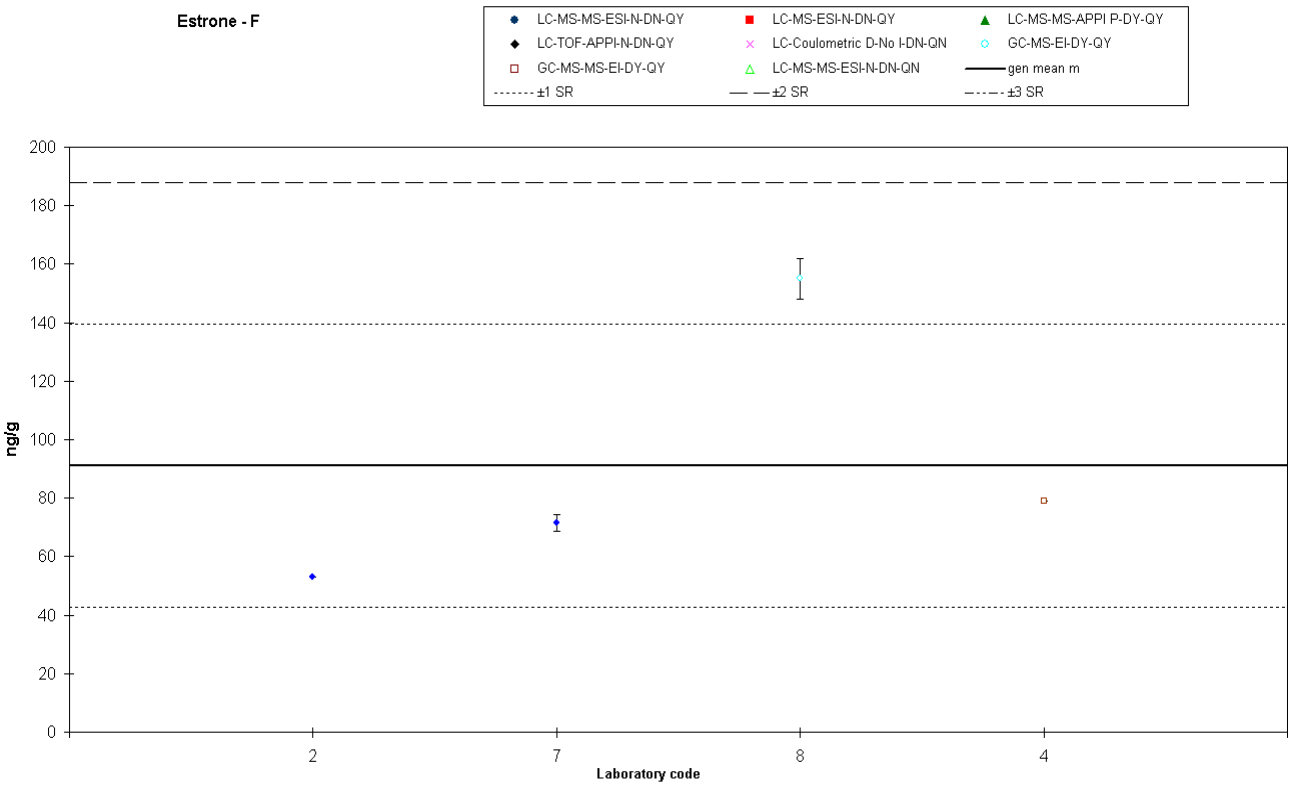
Sample E - Graphs





Sample F - Graphs





5 DISCUSSION OF THE RESULTS

In the first shipping, a standard solution containing an unknown concentration of the targeted compounds was sent to each participant laboratory. The content of the vial was to be diluted and quantified to report the concentration found. The standard was sent in order to detect possible calibration deviations among the laboratories which, if observed, could help to justify differences in the reported results. The values received have not been processed because of, what was considered, a misunderstanding in the instructions which generated a non-uniformity in the concentration format.

Concerning the data, the majority of the analytical methods employed by the laboratories reported acceptable results in surface, ground waters and effluent wastewaters. Concentrations reported were well distributed around the average, and acceptable reproducibility and repeatability deviations were found. As an example, there is a good agreement in the concentrations reported for estrone in all matrixes. No clear tendencies were observed when comparing results obtained by the different analytical instruments. This suggests that acceptability of the results is not only a function of the analytical technique but of the combination of steps in the procedure. Nevertheless, certain problems were detected when the concentrations of hormones were low and/or the analysis was performed on dirty matrixes.

Reiterative tendencies to over quantify were observed for certain methods suggesting that the quantification may be affected by interferences coming from co-extracted material. The graphs for ethinylestradiol in influent and spiked groundwater, and β-estradiol and estrone in effluent can be consulted regarding this matter.
Certain methods showed significantly higher standard deviations for their duplicates when compared to the other labs average repeatability standard deviations. The data for β-estradiol in effluent and spiked ground water and α-estradiol in the influent serve as a good example to illustrate this particular problem.
Elevated detection limits were reported by several labs when compared to the non-effect concentrations reported in the literature. In some cases, a factor of 25 separates two detection limits reported by two different laboratories for the same compound in a particular matrix. Clear examples are β-estradiol in influent and both α and β-estradiol in surface water.
All three problems aggravated as the complexity of the sample increased.

Considering the data and the statistical results, the optimum conditions for the analysis of hormones in complex matrixes seem the combination of a powerful purification stage followed by a highly selective and sensitive quantification technique. Instrumental limitations can be overthrown by the implementation of extensive clean up procedures which eliminates interferences and allows for robust and accurate results. In this case, clean up becomes essential when facing complex environmental matrixes such as raw water or sludge. On the other hand, when the quantification is conducted using a highly selective instrument, the need for an extremely clean extract, even if advisable, may not be as vital. Existing interferences are not accounted for because of the detector's specificity.
From the statistical process of the data and comparison of results we can conclude that analytical techniques such as HPLC-coulometry or those including GC-MS preceded of no clean up seem not to be adapted for the analysis of hormones in environmental matrixes. Results acquired using coulometry point to a lack of selectivity and sensitivity in the detection whereas important clean up of the extract is advisable when facing hormone quantification with a single quadrupole detector.

More development will be needed for the extension of the methods for the quantification of hormones in the sludge. since only a reduced number of laboratories were able to report results. It is important to point out the differences in the area of application of the participant laboratories. Certain methods were developed to detect estrogenic hormones in clean water matrixes such as surface, ground or drinking water. For those laboratories, the extension of their methods to the quantification of hormones in wastewater related matrixes is of no interest. In this case, this conclusion may be used to underline possible improvement lines in the case that a more challenging analysis is required.

No regulation exists in any of the laboratory's country of origin that limits the release of estrogenic hormones to the environment. Nevertheless, the United Kingdom is considering to set the no effect limit at only 0.1 ng/L which will translate into analytical detection limits of 0.03 ng/L. Limits between 1 and 3 ng/L for all receiving waters are being considered in the Netherlands in order to avoid the effect of estrogenic hormones on fish. Even though the creation of directives regulating the levels of estrogens in wastewater treatment plants discharges is not in sight, the increasing concern regarding endocrine disruption by different governments worldwide suggests that they may be established in the coming years.

6 CONCLUSIONS

Based on the results of this first international interlaboratory exercise the following overall conclusions can be made:

- The majority of the analytical methods employed by the laboratories reported acceptable results in surface, ground waters and effluent wastewaters. Concentrations reported were well distributed around the average, and acceptable reproducibility and repeatability deviations were found.
- The limited number of laboratories (3) that analysed the sludge samples prohibited to draw firm conclusions.
- No clear tendencies were observed when comparing results obtained by the different analytical instruments. This suggests that acceptability of the results is not only a function of the final analytical technique but of the combination of all the steps in the procedure. Nevertheless, LC-coulometry and GC/MS detection methods non including a previous clean up step seem not to be adapted to the analysis of estrogenic hormones in environmental samples. Certain problems were detected as well when the concentrations of hormones were low and/or the complexity of the samples increased.
- Considering all the data and the statistical results, the optimum conditions for the analysis of hormones in complex matrixes seem the combination of a powerful purification stage followed by a highly selective and sensitive quantification technique.