Report

Ellis, T.; Scott, A.P.; Waring, C.; Margiotta-Casaluci, L.; **Pottinger, T.G.**; Morris, S.; Sumpter, J.P.. 2011 *Accuracy of methods of sex steroid determination.* Defra, 28pp. (CEH Project Number: C03801)

General enquiries on this form should be made to:
Defra, Procurements and Contracts Division (Science R&D Team)

Telephone No. 0207 238 5734
E-mail: research.competitions@defra.gsi.gov.uk

SID 5 Research Project Final Report

- **Note**
  In line with the Freedom of Information Act 2000, Defra aims to place the results of its completed research projects in the public domain wherever possible. The SID 5 (Research Project Final Report) is designed to capture the information on the results and outputs of Defra-funded research in a format that is easily publishable through the Defra website. A SID 5 must be completed for all projects.
- This form is in Word format and the boxes may be expanded or reduced, as appropriate.
- **ACCESS TO INFORMATION**
  The information collected on this form will be stored electronically and may be sent to any part of Defra, or to individual researchers or organisations outside Defra for the purposes of reviewing the project. Defra may also disclose the information to any outside organisation acting as an agent authorised by Defra to process final research reports on its behalf. Defra intends to publish this form on its website, unless there are strong reasons not to, which fully comply with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.
  Defra may be required to release information, including personal data and commercial information, on request under the Environmental Information Regulations or the Freedom of Information Act 2000. However, Defra will not permit any unwarranted breach of confidentiality or act in contravention of its obligations under the Data Protection Act 1998. Defra or its appointed agents may use the name, address or other details on your form to contact you in connection with occasional customer research aimed at improving the processes through which Defra works with its contractors.

Project identification

1. Defra Project code: CB0427

2. Project title:
   Accuracy of methods of sex steroid determination

3. Contractor organisation(s):
   Cefas Weymouth Laboratory
   Barrack Road
   The Nothe
   Weymouth
   Dorset
   DT4 8UB

4. Total Defra project costs (agreed fixed price): £

5. Project: start date ..... 1 September 2009
   end date ..... 31 March 2011

6. It is Defra's intention to publish this form.
Please confirm your agreement to do so.....**YES** ✓ **NO** ☐

(a) When preparing SID 5s contractors should bear in mind that Defra intends that they be made public. They should be written in a clear and concise manner and represent a full account of the research project which someone not closely associated with the project can follow.
Defra recognises that in a small minority of cases there may be information, such as intellectual property or commercially confidential data, used in or generated by the research project, which should not be disclosed. In these cases, such information should be detailed in a separate annex (not to be published) so that the SID 5 can be placed in the public domain. Where it is impossible to complete the Final Report without including references to any sensitive or confidential data, the information should be included and section (b) completed. NB: only in exceptional circumstances will Defra expect contractors to give a "No" answer.
In all cases, reasons for withholding information must be fully in line with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.

(b) If you have answered NO, please explain why the Final report should not be released into public domain

Executive Summary

7. The executive summary must not exceed 2 sides in total of A4 and should be understandable to the intelligent non-scientist. It should cover the main objectives, methods and findings of the research, together with any other significant events and options for new work.

Reproductive hormones (estrogenic and androgenic steroids) enter natural water bodies from various sources, e.g. human sewage, farmed livestock and vertebrate wildlife. A previous Defra research study (*SF0241 Impacts of intensive in-river aquaculture on wild salmonids*) reported very high concentrations of such steroids in two UK rivers (the River Test and R. Avon in Hampshire/Wiltshire) in the vicinity of trout farms. The mean reported concentrations of the natural steroids 11-ketotestosterone (11-KT), testosterone (T) and oestradiol (E2) ranged from 4 to 79 ng/L. The maximum level observed in an individual sample was 145 ng/L for 11-KT. Furthermore, the subject brown trout and rainbow trout farms were implicated as a source of the steroids, as mean levels reported "downstream" of the trout farms were 1.3 to 6 fold higher than those "upstream". Nevertheless, the steroid concentrations reported both upstream and downstream of the trout farms would be of great concern for wild fish in these rivers due to potential for endocrine disruption, i.e. when exogenous substances negatively impact the endogenous hormone systems and reproductive function of the organism.

The synthetic sex steroid ethinyl-estradiol (EE2) was also measured in the SF0241 study. EE2 originates from human prescription medicine (the contraceptive pill) via sewage treatment works, and has been studied in rivers due to its endocrine disruptive activity. Levels of EE2 reported in SF0241 were one to two orders of magnitude lower than the natural steroids. Mean values were in the range of 0.13 – 0.30 ng/L, and were not found to be elevated downstream of the trout farms.

The very high levels of the natural sex steroids that were reported were not only of concern because of potential adverse effects on resident fish, but were also surprisingly high. In the SF0241 study, all steroid levels in acquired water samples had been measured using commercial Enzyme-ImmunoAssay (EIA) kits (also referred to as Enzyme-Linked Immuno-Sorbent Assay, ELISA kits). However, the kits used to measure the natural steroids used a different enzyme component in the assay to the EE2 kits. Environmental water samples are typically concentrated before assay, and it was speculated that other compounds in the river water may have interfered with the enzyme stage in the kits for the natural steroids. It was therefore hypothesised that the very high concentrations of the natural steroids reported in SF0241 were false positives due to interference in the EIAs.

Due to concern over the possible inaccuracy of the EIA kits (and trout farms as a source of endocrine disrupting compounds), Defra's Chemicals and Nanotechnology Division funded this project (*CB0427*) to examine the "*Accuracy of methods of sex steroid determination*". The aim of the project was to test the above hypothesis by repeating the sampling and sample processing, and then assaying for the steroids

using the EIA kits and an additional method – radioimmunoassay (RIA) - expected to be less susceptible to interference.

Water samples were collected from the two trout farm sites, one on the R. Test and one on the R. Avon, between January and June 2010 and extracted (C18 solid phase extraction (SPE) followed by extract clean-up with aminopropyl SPE). Additional "spiked" and "blank" water samples were also prepared and processed.

Four independent laboratories conducted EIAs and RIAs (3 laboratories per assay technique) for the four steroids (11-KT, T, E2, EE2) using replicate aliquots of the same 44 samples, each aliquot representing ca 1 L of river water. Participating laboratories (other than the lead laboratory) conducted the assays blind (i.e. were unaware of the sample details) and returned the calculated steroid concentrations to the project leaders for collation. A few anomalous results were questioned and, upon investigation, were found to be due to human errors and were subsequently corrected.

There was broad agreement between the EIA and RIA measurements for all four steroids showing that the EIA kits did not generate erroneously high values. The base hypothesis for the project, i.e. that some EIA kits generate erroneously high values, was therefore rejected. The two assay methods were comparable for accuracy and precision.

Recovery of steroids using the SPE methodology was examined using water samples spiked with known amounts of steroids. It was found that the recovery efficiency varied between samples and steroids. Higher recoveries were evident for the oestrogens (mean recovery 69% and 67% for E2 and EE2 respectively) than for the androgenic compounds (mean recovery 29% and 46% for 11KT and T respectively).

Measured steroid concentrations in river water samples were all <0.6 ng/L. The measured concentrations cannot be considered definitive, as they are uncorrected for recovery efficiency. Nevertheless, they are considered low, unlikely to be of concern for endocrine disruption, and demonstrate that the water steroid levels reported in SF0241 were not typical of river steroid concentrations in 2010. The levels of 11KT, T, E2 and EE2 in the rivers were respectively 1300, 70, 90 and 3 times lower than reported in SF0241. In addition to the major difference in measured levels between the two studies:

- the relative concentrations also differed, with 11KT being the lowest rather than highest as reported in SF0241.
- river steroid levels were higher in Spring rather than the Winter period, being opposite to the seasonal effect reported in SF0241
- river EE2 levels in 2010 were too low to be detectable by EIA (equivalent to <0.05 ng/L).
- the clear elevation in river water concentrations of the natural steroids associated with the fish farms reported in SF0241 was not evident in the present study.

It was not possible to provide a definitive explanation for the differences found between the present results and those reported within SF0241. Eleven alternate hypotheses are discussed that could have contributed to the divergent results. The available evidence points towards miscalculation and assaying errors within SF0241 as the probable cause of differences. However, this hypothesis could not be definitively accepted because the raw assay data and the calculations from which the SF0241 results derived had not been retained.

Published information on steroid output from fish farms is presented. It is suggested that there is no urgent requirement to further examine the steroid output from UK fish farms. Although the current research did indicate a possible 0.13 ng/L increase in testosterone in the immediate outflow of one farm, this was questionable and, if real, is below concentrations that cause endocrine disruption and would be further diluted in the receiving channel and then main river.

During this project, a number of potential sources of error associated with the measurement of steroids from samples were identified and are discussed. It is suggested that guidance for quality control could be developed. Future research into river water steroids could also include:

- comparisons of the recovery efficiencies of different steroids from water samples, and a full optimisation and validation of the most appropriate extraction methodology for the different steroids
- a comparison of 'modelled' versus 'measured' river water steroid levels to examine whether the exclusion of other sources of steroids, e.g. agricultural livestock, is an important omission.

8. As a guide this report should be no longer than 20 sides of A4. This report is to provide Defra with details of the outputs of the research project for internal purposes; to meet the terms of the contract; and to allow Defra to publish details of the outputs to meet Environmental Information Regulation or Freedom of Information obligations. This short report to Defra does not preclude contractors from also seeking to publish a full, formal scientific report/paper in an appropriate scientific or other journal/publication. Indeed, Defra actively encourages such publications as part of the contract terms. The report to Defra should include:
   - the scientific objectives as set out in the contract;
   - the extent to which the objectives set out in the contract have been met;
   - details of methods used and the results obtained, including statistical analysis (if appropriate);
   - a discussion of the results and their reliability;
   - the main implications of the findings;
   - possible future work; and
   - any action resulting from the research (e.g. IP, Knowledge Transfer).

BACKGROUND

Reproductive hormones (estrogenic and androgenic steroids) enter natural water bodies from various sources, e.g. human sewage, farmed livestock and vertebrate wildlife (e.g. Kolodziej et al. 2004; Barel-Cohen 2006; Zhao et al. 2010). The previous Defra research study *SF0241 Impacts of intensive in-river aquaculture on wild salmonids* reported very high concentrations of sex steroid hormones in two UK rivers (the R. Test and R. Avon – Wiltshire/Hampshire) in the proximity of trout farms. The mean concentrations of the natural steroids 11-ketotestosterone (11-KT), testosterone (T) and oestradiol (E2; oestradiol-17β), ranged between 4 and 79 ng/L**.** The maximum level observed in an individual sample was 145 ng/L of 11-KT [individual sample data provided by SF0241 Principal Investigator (PI)]. Furthermore, the trout farms were implicated as a source of the steroids with mean concentrations reported "downstream" of the trout farms being 1.3 to 6 times higher than "upstream" levels.

Steroids, whether from humans, cows or fish, are chemically identical, and their physiological roles have generally been conserved in the vertebrate lineage. Steroids in water, no matter what their origin, can therefore potentially affect the physiology and behaviour of wild fish and amphibians. Steroids in water bodies are typically at far lower concentrations than within aquatic animals themselves, and it is only when exogenous water steroid concentrations exceed concentrations of a few ng/L that they are abundant enough to affect internal hormone systems and reproductive function, a process commonly known as endocrine disruption (Balaam et al. 2010; Zhao et al. 2010). The concentrations reported by SF0241 (both "upstream" and "downstream") would therefore be of concern for wild fish in the rivers due to the potential for endocrine disruption to occur.

The synthetic sex steroid ethinyl-estradiol (EE2; 17α-ethinylestradiol) was also measured within SF0241. EE2 derives from human prescription medicine (the contraceptive pill) and is known to enter receiving waters via sewage treatment works (Balaam et al. 2010). Its presence in rivers and concerns relating to endocrine disruption in fish are well documented (see Balaam et al. 2010). Mean levels of EE2 reported in the SF0241 study were 0.13 – 0.30 ng/L, lower than the natural steroids (11-KT, E2 and T), and not found to be elevated downstream of the trout farms.

The very high levels of the natural sex steroids reported in SF0241 were not only of concern from the perspective of endocrine disruption, but also surprisingly high. The steroids had been measured using commercial Enzyme-ImmunoAssay (EIA) kits (frequently referred to as Enzyme-Linked Immuno-Sorbent Assay, ELISA kits). Environmental water samples are typically extracted and concentrated before assay. It has been suggested previously that humic substances may interfere with steroid EIA measurements, the extent of which depends upon the actual bio-chemical processes involved in the assay (Hanselman et al. 2004). The EIA kits used to measure the natural steroids used a different enzyme component in the assay to the EE2 kits. It was therefore hypothesised that the very high concentrations of the natural steroids reported in SF0241 were false positives caused by interference within the EIAs from other compounds, co-extracted from the river water with the steroids.

EIA kits are widely used to measure water (and blood) concentrations of steroids in both laboratory and field research into endocrine disruption. EIA kits have previously been found to generate different values for blood plasma steroid (cortisol) concentrations when compared with a different assay method - radioimmunoassay (RIA) (Sink et al. 2008). Principally due to the concern over the possible inaccuracy of EIA kits (and also the possibility of trout farms acting as a source of endocrine disruption compounds), Defra's Chemicals and Nanotechnology Division funded the current project (*CB0427*) to examine the "*Accuracy of methods of sex steroid determination*". The aim of the project was to test the above hypothesis by repeating the SF0241 sampling and sample processing, and then assaying for the steroids using the EIA kits and RIA - expected to be less prone to interference. A comparison was made involving four independent laboratories conducting EIAs and RIAs (three laboratories per method) for the four steroids (11-KT, T, E2, EE2) using replicate aliquots derived from the same samples.

It must be appreciated that neither EIA nor RIA measure amounts of steroid directly. Both assay methods rely on quantifying the amount of steroid in samples from the proportion of added antibody that it binds to. The proportion of bound antibody is then measured, but again by an indirect technique involving either enzymatic activity (EIA) or radioactively-labelled steroid (RIA) (Fig 1).

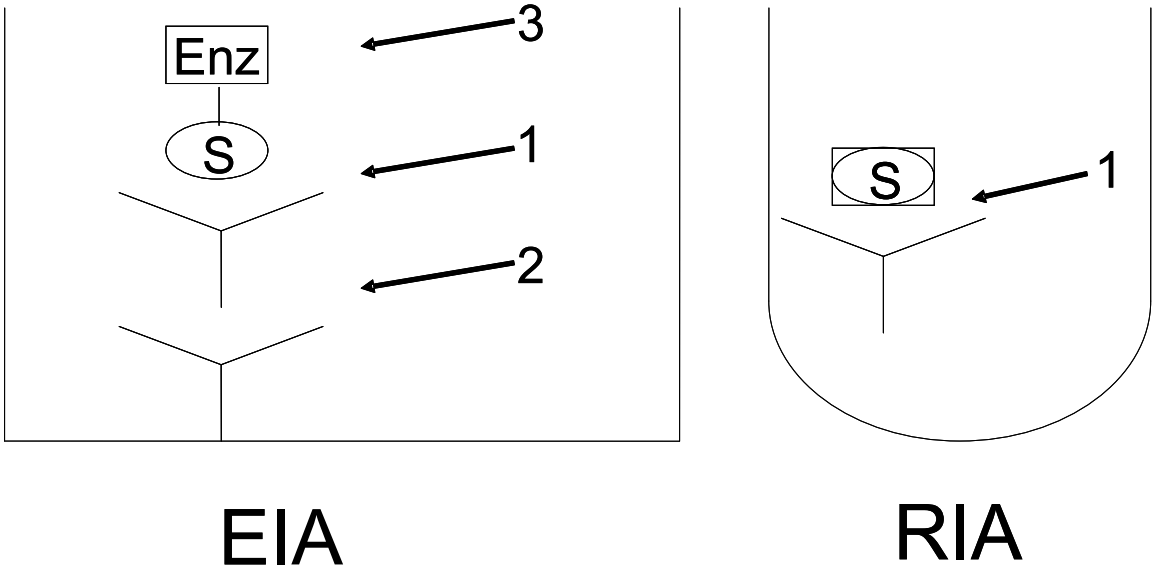


FIGURE1: The potential sources of non-specific interference in steroid EIA and RIA: 1) disruption of binding of labelled steroid to primary antibody; 2) disruption of binding of primary antibody to secondary antibody; and 3) disruption of enzyme activity.

Both EIA and RIA rely on the generation of a predictive standard curve from standards included in the assay:

- for EIA: Measurements of optical density/absorbance (representing enzymatic activity proportional to the amount of bound steroid) versus steroid concentration in standards are made, whereas
- for RIA: Measurements of radioactivity (measured via fluorescence in scintillant fluids and representing undisplaced labelled steroid) versus steroid concentrations in standards are undertaken.

Values for optical density or radioactivity of the unknown samples are converted to (estimated) steroid concentrations using equations derived from the standard curve. EIAs are generally conducted in off-the-shelf 96-well plates, standardised for use with optical density readers. In contrast RIAs use a series of individual test-tubes and scintillation vials.

This project was led by Cefas. The University of Portsmouth was a key project participant – they had conducted the original SF0241 study (under sub-contract to Cefas), and their participation ensured the SF0241 sampling, extraction and EIA assaying was replicated as far as possible. Additional laboratories were recruited to provide independent assay measurements using both EIA and RIA techniques (three laboratories per assay method) and these were Brunel University and the Centre for Ecology & Hydrology (CEH), Lancaster. It should be noted that the current study was not designed as an inter-laboratory study to evaluate the performance of the two assay techniques (e.g. MacMaster et al. 2001; Heath et al., 2010); it was designed primarily to assess the validity of the EIA kits used to generate the SF0241 reported water steroid levels, and secondarily to assess the steroid output from trout farms.

The specific objectives were:

1. Collection and extraction of samples from trout farms
2. Distribution of samples to participants
3. Completion of EIA and RIA assays of all samples by participants
4. HPLC analysis of immunoactive compounds in water extracts
5. Collation and analysis of results
6. Submission of final report

All objectives were achieved, apart from #4, as described below.

MATERIALS & METHODS

Sampling sites

Initial visits were made to the R. Test (Hampshire) and R. Avon (Wiltshire) sites to view sampling stations, discuss the project and sampling with the fish farmers and land owners, and gain agreement for access and provision of stock details for the SF0241 and CB0427 sampling dates. All parties were very helpful. Details of the farms, management practices and trout stocks held during the SF0241 and CB0427 sampling periods are provided in Appendix 1.

For brevity, the SF0241 report simply indicated that, at both sites, samples were taken 50 m upstream and 50 m downstream of where the trout farm effluent entered the river. These locations represented control and impacted stations respectively. However, the SF0241 upstream station was not 100 m above the downstream station at either site but had been located on nearby parallel channels for various reasons (proximity of the stations; private ownership of land ensuring security; staff health and safety; historical use of the sites in previous projects; operation away from fish farm inflows for biosecurity). The water courses around both trout farms are complex with division of the rivers into several channels, and none of the SF0241 stations were on the main rivers (Fig 2).

At the **R. Test** site, the channels supplying the upstream and downstream stations diverged 900 m upstream of the fish farm (Fig 2):

- The channel supplying the upstream station flows through agricultural land and no pollution incidents were recorded on the Environment Agency (EA) "What's in your backyard" website
- The downstream channel flows past industrial and residential areas with nearby roads, before receiving the fish farm effluent. It also receives water from two additional channels, the input from which is understood to be variable, depending upon river levels and sluice control. These additional channels flow through different agricultural areas, and pass other residential and industrial areas before joining the downstream channel above the fish farm. Two pollution incidents are reported on the EA "What's in your backyard" website for the channel network supplying the SF0241 downstream station:
    - 10 Jan 2003 – sewage materials - “significant” impact to water
    - 1 Feb 2004 – oils and fuel – “major” impact to water

Both these incidents were within the SF0241 study period (December 2002 – June 2006), but did not coincide with sampling. Whether they are indicative of background anthropogenic inputs is unknown.

At the **R. Avon** site, the channel supplying the “SF0241 upstream” station diverges from the main river 200 m downstream of where the fish farm supply channel diverges (Fig 2). Both channels are controlled by sluices.

- The upstream channel flows across a field for 250 m before the “upstream” station.
- The fish farm supply flows 350 m around the hedged perimeter of a field, before splitting into the fish farm supply and bypass channel.
- The bypass channel flows along the hedged perimeter of the field for 100 m before joining the upstream channel, just downstream of the SF0241 upstream station.
- There are 3 separate outflow points from the fish farm into the combined upstream/farm bypass channel. The “SF0241 downstream” station was downstream of the third farm outflow point.
- In addition to the river water, both the upstream channel and farm supply channel receive additional input from separate springs. However, the contribution from these springs was considered (by the farm manager) to be negligible (< 5%) and this impression was supported by personal observation.
- No pollution incidents were recorded on the EA "What's in your backyard" website for the vicinity. The field(s) through which the channels flow are used for livestock: livestock are visible in satellite images (Google Earth) and cattle faeces were noted in the fields. There is therefore a possibility of differential localised inputs from livestock and other natural sources between the SF0241 upstream and downstream stations. The SF0241 downstream station is also adjacent to a road.

The initial CB0427 proposal was to collect water samples from 3 stations per site: “SF0241 upstream”, “SF0241 downstream” and an additional fish farm outflow (prior to dilution by receiving channel flow). The outflow station at the R. Avon farm site (with 3 separate outflows) was from the pond containing the highest biomass of fish on-site, which were also larger and more reproductively mature. As the SF0241 upstream stations could not be considered to represent farm inflow (due to the possibility of other agricultural, industrial, road and residential inputs) additional fish farm inflow samples were taken to enable assessment of the role of the fish farms (Fig 2).

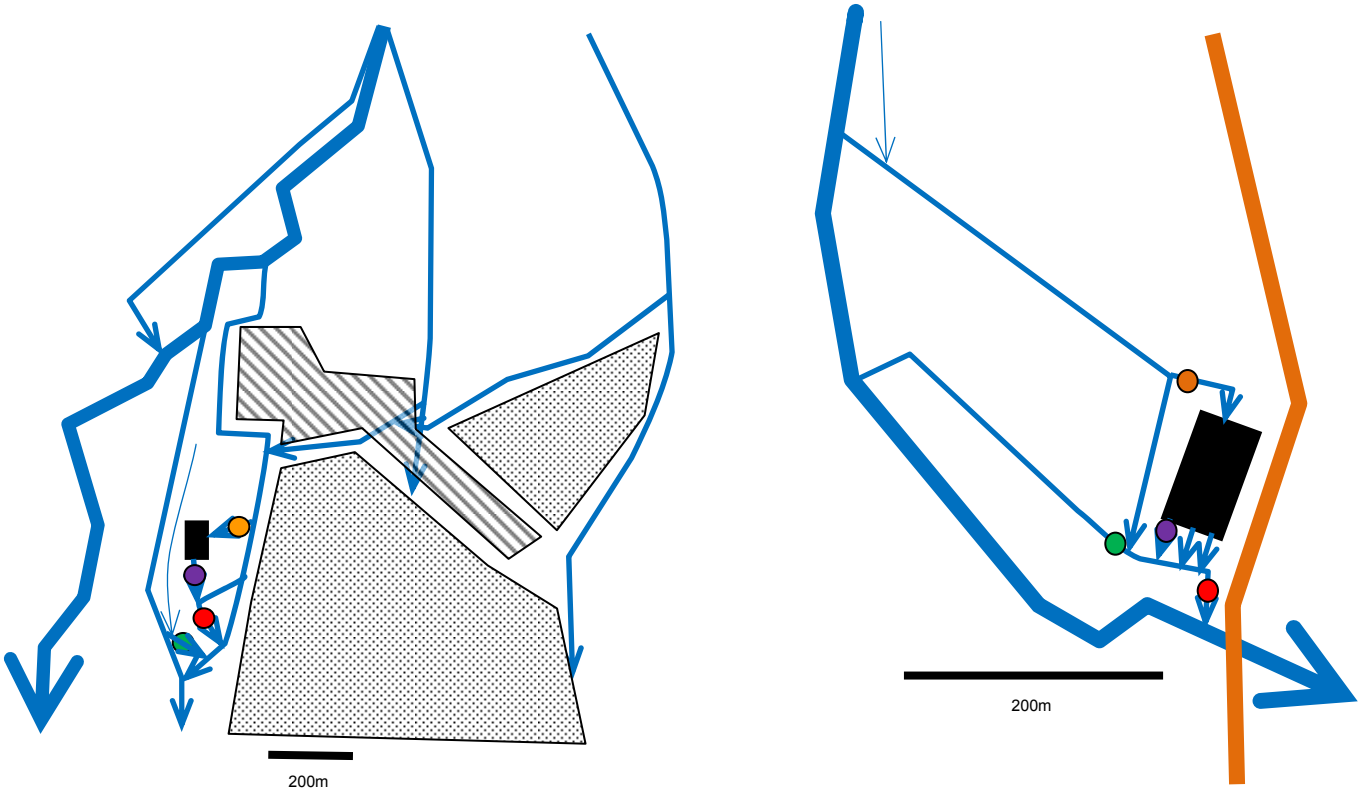


FIGURE 2: Diagrammatic representation of locations of sampling stations in relation to river channels and fish farms on the R Test (left) and R Avon (right). Key below.

fish farm		SF0241 upstream sampling station	
main river (+ flow direction)		SF0241 downstream sampling station	
channel (+ flow direction)		CB0427 fish farm inflow sampling station	
ditch (+ flow direction)		CB0427 Fish farm outflow sampling station	
residential area		industrial area	
road		200 m scale bar	

Sampling

Samples were collected from both sites on two separate occasions in the Winter period (Jan/Feb 2010), and on three occasions in the Spring season (Apr-June 2010). For each sample, river water was collected in 25 L HDPE drums for subsequent processing. Sufficient water was collected per sample to provide:

- six 1 L sub-sample equivalents for the assays (3 for EIA, 3 for RIA)
- one 1 L sub-sample equivalent for HPLC analyses
- spare 1 L sub-sample equivalents in case of mishap (≤ 5, varying between samples).

Additional validation samples were collected or prepared to assess the performance of the extraction and assay procedures.

- Three replicate samples were collected from one station on each river site on a single occasion (21/04/10; 11/05/10). This was to assess the precision (variability) associated with water steroid measurements.
- Two deionised water samples – to assess assay values for "blank" samples.
- Two "spiked" river water samples (one from each river) – to assess the recovery efficiency of the extraction methodology
- Two "spiked" deionised water samples – to assess if other compounds in the river water samples interfered with recovery and the performance of the assays

The spiking of river and deionised water samples was conducted by a staff member at Cefas who was not involved in the sample collection, extractions or assays. A stock solution was prepared containing a mixture of 11KT (0.498 ng/µL), T (0.501 ng/µL), E2 (0.503 ng/µL) and EE2 (0.024 ng/µL). From this, exactly 1 mL was added to 10 L of each water sample to provide equivalent concentrations of 49.8 ng/L of 11KT, 50.1 ng/L of T, 50.3 ng/L of E2, and 2.4 ng/L of EE2. These concentrations were similar to those reported in SF0241 for the natural steroids, and a factor of 10 higher for the synthetic steroid (EE2). The stock solution containing the mixture of steroids was retained, and included in the assays as an unknown sample. Unspiked methanol was also retained for inclusion in the assays.

Sample Processing

Sample processing involved various steps to extract the steroids from the water samples (several litres) and concentrate into small volumes (a few mLs) of methanol solvent. Further steps were required to prepare buffered solutions appropriate for the assays. The procedure was conducted to replicate the SF0241 procedure (based upon Rubio et al., 2004) as far as possible, although modifications were made due to the greater sample volumes processed and suitability / availability of equipment. The steps involved are summarised in Appendix 2. Departures from the SF0241 procedure were discussed between Cefas and Portsmouth beforehand, and considered to have no likely impact on the recovery of steroids, although this would be reviewed in the light of the assay results. The sample processing resulted in replicate 1 mL aliquots of 44 samples.

Assaying

The assaying involved a total of 44 samples:

- 19 samples from the R. Test (collected on five separate occasions)
- 15 samples from the R. Avon (collected on five separate occasions)
- One steroid-spiked R. Test water sample
- One R. Test water sample spike control (1 mL unspiked methanol added before extraction)
- One steroid-spiked R. Avon water sample
- One R. Avon water sample spike control (1 mL unspiked methanol added before extraction)
- Two unspiked deionised water samples
- Two steroid-spiked deionised water samples
- One sample of the steroid spike methanol stock solution (0.1 mL spiking methanol + 900 uL methanol; equating to spike in 1 L of water)
- One sample of unspiked methanol stock

Sub-sample aliquots (1 mL) were coded prior to distribution to ensure that assays were conducted blind (by all participants apart from Cefas).

EIA kits were purchased by Portsmouth University and distributed to the participating laboratories (Cefas, Brunel, and Portsmouth). For participants applying RIAs (Cefas, Brunel, CEH Lancaster), CEH Lancaster were provided with radiolabels and antibodies for 11K-T, T, E2 and EE2 assays from Cefas stocks, but the RIA standards used by CEH, and all the materials for the Brunel RIAs, had been sourced independently. The ranges of steroid concentration in the standard curves are provided (Table 1).

TABLE 1: Details of standard curves in the EIAs and RIAs.

Assay	Steroid	Manufacturer/Participant	No of standards	Minimum ng/mL	Maximum ng/mL
EIA	11-KT	Cayman Chemical Company	8	0.00078	0.1
	T		8	0.0039	0.5
	E2		8	0.0066	4.0
	EE2	Tokiwa Chemical Industries	4	0.05	3.0
RIA	11-KT	Cefas	9	0.01	2.5
	T		9	0.02	5.0
	E2		9	0.005	1.25
	EE2		9	0.02	5.0
	11-KT	Brunel University	7	0.02	1.25
	T		7	0.02	1.25
	E2		8	0.01	1.25
	EE2		7	0.02	1.25
	11-KT	CEH, Lancaster	8	0.06	8.0
	T		8	0.06	8.0
	E2		8	0.06	8.0
	EE2		8	0.06	8.0

As lead contractor, Cefas undertook assays prior to distribution to identify any potential issues with the limited amounts of material provided (1 mL extract, equivalent to 1 L water sample, being provided for assay of 4

different steroids). Guidance was subsequently provided to participants on generic amounts of samples used by Cefas in their assays, in addition to storage of EIA kits and handling the extracts to reduce the risk of cross-contamination. However, the participants were free to proceed as they preferred. Assay participants were provided with a spreadsheet with the sample code number/letter and water volume the sample derived from (where relevant) and asked to provide the amount of each of the four steroids in each 1 mL extract (ng/mL), and the water concentration (ng/L) that this equated to. Results were provided to the Cefas project leaders for collation. The involvement of an independent quality control officer, for coding of samples and aliquots and collation of results, did not prove possible due to constraints of time and staff availability.

RESULTS & DISCUSSION

Steroid concentrations measured in distributed sample extract aliquots

The measured aliquot concentrations, as reported by each participant, are presented in Table 2. Aliquots coded A-E relate to the methanol-based steroid stock solution used for spiking and the spiked river and deionised water samples. With the exception of sample numbers 25 (blank –methanol), 26 and 29 (unspiked deionised water blanks), other sample numbers relate to river water samples. Some cells for EE2 contain 2 values (highlighted red) where the initially reported values (first quoted value) clearly differed from those from the other participants. Upon investigation, these were revealed to be errors and recalculated values are also provided (second quoted value). Cells shaded in yellow indicate values provided outside of the standard curve range.

TABLE 2A: Concentrations (in ng/mL) of 11-keto Testosterone (11KT) and Testosterone (T) measured in the 44 sample aliquots by EIA and RIA by the project participants, 3 participants per assay methodology. A value below the EIA detection limit is denoted as <DL.

Aliquot No	11-KT						T					
	EIA			RIA			EIA			RIA		
	1	2	3	4	5	6	1	2	3	4	5	6
1	0.01	0.01	0.01	0.08	0.03	0.03	0.05	0.11	0.16	0.17	0.22	0.25
2	0.01	0.01	0.01	0.09	0.03	0.00	0.07	0.10	0.12	0.17	0.18	0.21
3	0.01	0.01	0.01	0.09	0.03	0.00	0.08	0.12	0.12	0.20	0.26	0.23
4	0.01	0.01	0.02	0.11	0.03	0.00	0.11	0.20	0.28	0.30	0.38	0.32
5	0.01	0.01	0.01	0.09	0.04	0.00	0.09	0.10	0.16	0.16	0.16	0.17
6	0.01	0.01	0.01	0.09	0.03	0.00	0.07	0.12	0.13	0.21	0.18	0.22
7	0.02	0.02	0.01	0.11	0.03	0.00	0.11	0.21	0.28	0.26	0.30	0.20
8	0.01	0.01	0.01	0.08	0.04	0.00	0.09	0.13	0.34	0.18	0.20	0.26
9	0.01	0.01	0.01	0.07	0.03	0.00	0.06	0.11	0.11	0.16	0.20	0.21
10	0.01	0.01	0.01	0.10	0.03	0.00	0.07	0.11	0.11	0.16	0.18	0.20
11	0.01	0.01	0.01	0.11	0.03	0.00	0.09	0.20	0.11	0.22	0.32	0.29
12	0.01	0.01	0.03	0.10	0.04	0.00	0.11	0.14	0.19	0.24	0.32	0.27
13	0.02	0.02	0.02	0.12	0.05	0.00	0.10	0.18	0.22	0.26	0.36	0.24
14	0.03	0.04	0.02	0.15	0.07	0.00	0.10	0.16	0.21	0.29	0.35	0.34
15	0.02	0.02	0.04	0.10	0.05	0.00	0.11	0.20	0.29	0.26	0.36	0.26
16	0.03	0.03	0.03	0.12	0.04	0.02	0.13	0.21	0.30	0.26	0.31	0.24
17	0.02	0.02	0.03	0.11	0.04	0.00	0.11	0.18	0.22	0.25	0.28	0.30
18	0.02	0.02	0.01	0.11	0.05	0.00	0.14	0.20	0.15	0.28	0.36	0.35
19	0.01	0.01	0.01	0.10	0.05	0.00	0.12	0.21	0.14	0.27	0.37	0.24
20	0.01	0.01	0.02	0.09	0.03	0.00	0.17	0.26	0.22	0.31	0.44	0.32
21	0.02	0.02	0.02	0.10	0.04	0.00	0.19	0.29	0.27	0.39	0.47	0.37
22	0.02	0.02	0.02	0.12	0.05	0.00	0.23	0.29	0.31	0.45	0.65	0.48
23	0.02	0.02	0.03	0.12	0.04	0.00	0.27	0.38	0.59	0.48	0.67	0.46
24	0.02	0.02	0.03	0.11	0.04	0.00	0.25	0.33	0.87	0.46	0.62	0.45
25	0.02	0.03	0.03	0.05	0.04	0.00	0.05	0.09	0.06	0.03	0.06	0.05
26	0.00	0.00	0.00	0.04	0.03	0.00	0.01	0.01	0.02	0.02	0.03	0.00
27	0.03	0.03	0.02	0.13	0.04	0.00	0.33	0.46	0.28	0.63	0.81	0.55
28	0.03	0.02	0.04	0.13	0.04	0.01	0.32	0.44	0.56	0.60	0.78	0.58
29	0.00	<DL	0.00	0.02	0.03	0.00	0.01	0.01	0.02	0.02	0.02	0.03
30	0.02	0.02	0.02	0.12	0.04	0.00	0.20	0.36	0.27	0.48	0.58	0.41
31	0.03	0.03	0.09	0.13	0.05	0.01	0.25	0.37	0.84	0.54	0.60	0.51
32	0.02	0.02	0.03	0.11	0.06	0.00	0.25	0.28	1.08	0.49	0.65	0.47
33	0.03	0.02	0.03	0.11	0.06	0.00	0.26	0.47	0.70	0.56	0.63	0.50
34	0.03	0.02	0.02	0.11	0.04	0.02	0.30	0.39	0.33	0.51	0.62	0.48
35	0.04	0.02	0.02	0.12	0.04	0.02	0.32	0.51	0.32	0.54	0.74	0.52
36	0.03	0.02	0.02	0.12	0.04	0.00	0.21	0.31	0.34	0.41	0.52	0.40
37	0.02	0.02	0.02	0.09	0.06	0.00	0.23	0.33	0.33	0.45	0.50	0.44
38	0.02	0.02	0.02	0.12	0.05	0.03	0.28	0.42	0.32	0.63	0.77	0.59
39	0.02	0.02	0.03	0.12	0.04	0.02	0.36	0.54	0.61	0.70	0.79	0.60
A	45	39	27	37	50	38	40	42	32	47	51	40
B	13	7	6	13	12	11	18	21	13	24	21	23
C	11	7	6	10	11	11	17	14	12	24	21	23
D	10	5	7	11	13	10	15	16	12	18	19	18
E	17	11	10	16	13	16	20	18	20	31	27	26

TABLE 2B: Concentrations (in ng/mL) of Estradiol (E2) and Ethynylestradiol (EE2) measured in the 44 sample aliquots by EIA and RIA by the project participants, 3 participants per assay methodology. Cells shaded red indicate initial erroneous values that were subsequently re-calculated. Cells shaded yellow indicate values provided outside of standard curve range.

Aliquot No	E2						EE2					
	EIA			RIA			EIA			RIA		
	1	2	3	4	5	6	1	2	3	4	5	6
1	0.22	0.23	0.17	0.18	0.16	0.14	<0.05	<0.05	nd	0.17	0.03	0.08
2	0.17	0.22	0.12	0.15	0.17	0.14	<0.05	<0.05	nd	0.14	0.03	0.09
3	0.15	0.20	0.09	0.14	0.09	0.14	<0.05	<0.05	nd	0.14	0.03	0.05
4	0.20	0.22	0.16	0.15	0.16	0.16	<0.05	<0.05	nd	0.15	0.03	0.05
5	0.19	0.17	0.11	0.15	0.17	0.14	<0.05	<0.05	nd	0.15	0.03	0.03
6	0.20	0.20	0.12	0.14	0.18	0.14	<0.05	<0.05	nd	0.18	0.02	0.04
7	0.23	0.26	0.12	0.15	0.20	0.16	<0.05	0.03	nd	0.14	0.02	0.06
8	0.30	0.20	0.17	0.15	0.17	0.13	<0.05	0.02	nd	0.17	0.02	0.07
9	0.23	0.26	0.18	0.16	0.24	0.17	<0.05	<0.05	nd	0.14	0.02	0.06
10	0.19	0.22	0.11	0.15	0.18	0.15	<0.05	<0.05	nd	0.16	0.02	0.03
11	0.22	0.32	0.12	0.16	0.19	0.14	<0.05	<0.05	nd	0.15	0.03	0.03
12	0.19	0.21	0.14	0.14	0.16	0.14	<0.05	<0.05	nd	0.15	0.03	0.03
13	0.17	0.17	0.11	0.14	0.11	0.14	0.05	<0.05	nd	0.15	0.03	0.08
14	0.20	0.21	0.12	0.15	0.15	0.16	<0.05	0.02	nd	0.17	0.02	0.09
15	0.19	0.23	0.15	0.15	0.16	0.16	<0.05	0.02	nd	0.16	0.02	0.06
16	0.20	0.19	0.17	0.13	0.18	0.15	<0.05	<0.05	nd	0.16	0.03	0.10
17	0.17	0.17	0.16	0.14	0.12	0.14	<0.05	<0.05	nd	0.15	0.03	0.04
18	0.20	0.23	0.12	0.14	0.17	0.17	<0.05	<0.05	nd	0.14	0.03	0.08
19	0.17	0.20	0.10	0.13	0.14	0.14	<0.05	<0.05	nd	0.13	0.03	0.02
20	0.21	0.25	0.15	0.16	0.18	0.15	<0.05	<0.05	nd	0.17	0.03	0.07
21	0.21	0.22	0.15	0.15	0.17	0.16	<0.05	<0.05	nd	0.16	0.02	0.06
22	0.21	0.23	0.14	0.14	0.18	0.15	<0.05	<0.05	nd	0.18	0.02	0.07
23	0.22	0.29	0.18	0.14	0.21	0.17	<0.05	<0.05	nd	0.20	0.13	0.06
24	0.18	0.21	0.19	0.14	0.16	0.16	<0.05	<0.05	nd	0.16	0.02	0.03
25	0.81	0.94	0.49	0.59	1.01	0.45	0.14	0.16	0.03 / 0.10	0.27	0.03	0.13
26	0.04	0.04	0.03	0.03	0.05	0.04	<0.05	<0.05	nd	nd	0.03	0.02
27	0.37	0.39	0.17	0.24	0.35	0.22	<0.05	<0.05	nd	0.20	0.03	0.04
28	0.32	0.37	0.25	0.24	0.36	0.21	<0.05	<0.05	nd	0.19	0.03	0.08
29	0.10	0.07	0.04	0.05	0.09	0.06	<0.05	0.03	nd	0.03	0.04	0.00
30	0.40	0.40	0.24	0.28	0.48	0.25	<0.05	<0.05	nd	0.24	0.04	0.17
31	0.33	0.43	0.34	0.24	0.25	0.22	<0.05	<0.05	nd	0.25	0.04	0.11
32	0.26	0.34	0.34	0.26	0.29	0.22	<0.05	<0.05	nd	0.27	0.03	0.09
33	0.33	0.38	0.27	0.27	0.19	0.22	<0.05	<0.05	nd	0.28	0.04	0.12
34	0.34	0.35	0.19	0.25	0.30	0.22	<0.05	<0.05	nd	0.29	0.03	0.11
35	0.41	0.48	0.19	0.27	0.28	0.24	<0.05	<0.05	nd	0.29	0.04	0.09
36	0.28	0.27	0.20	0.21	0.24	0.18	<0.05	<0.05	nd	0.21	0.04	0.06
37	0.23	0.30	0.18	0.23	0.34	0.20	<0.05	<0.05	nd	0.22	0.03	0.07
38	0.22	0.30	0.19	0.25	0.32	0.20	<0.05	<0.05	nd	0.22	0.04	0.06
39	0.24	0.38	0.30	0.24	0.21	0.21	<0.05	<0.05	nd	0.23	0.25	0.03
A	65	55	21	46	63	38	4.4	5.0	0.2 / 2.1	6.0	97 / 3.9	8.0
B	47	44	18	42	50	31	3.7	4.0	0.2 / 2.0	5.6	83 / 3.3	7.7
C	52	35	22	38	34	34	3.2	3.0	0.2 / 2.1	5.7	60 / 2.4	5.7
D	26	31	13	26	34	26	2.7	2.5	1.0 / 0.9	4.8	85 / 3.4	5.6
E	56	43	22	42	48	36	3.6	3.3	0.2 / 1.9	5.8	14 / 1.5	6.6

Do EIA kits generate higher steroid values than the RIA method?
No. Generally, there was broad agreement between the EIA and RIA measurements (Table 2). Ranges of concentration from the two assay techniques were similar for the unspiked river water samples (#1-39; <1 ng/mL) and spiked samples (#A-E; >2.5 ng/L). Both EIA and RIA returned steroid values for deionised (blank) water samples (#26 & 29), although the amounts were negligible.

A statistical comparison of the measurements by the two methods (EIA v RIA) of steroid concentrations in the sample aliquots was made for each steroid (Minitab General Linear Model ANOVA of transformed (log_{10} [concentration +1]) data with 3 factors: assay method, sample, participant {nested within assay method}). Separate comparisons were made for river water samples (aliquots 1-39 excluding 25 (blank - methanol) and 26 and 29 (blank – extracted deionised water)) and spiked samples (A-E) due to the 10 to 100-fold difference in levels (Table 3).

TABLE 3: Statistical comparisons of measured steroid concentrations in river water extracts by EIA and RIA: Least Squares Mean (LSM, untransformed) and probability values provided.

	River water samples					Spiked samples			
	EIA LSM (ng/mL)	RIA LSM (ng/mL)	Ratio: EIA LSM ÷ RIA LSM	p		EIA LSM (ng/mL)	RIA LSM (ng/mL)	Ratio: EIA LSM ÷ RIA LSM	p
11KT	0.02	0.05	0.4	0.000		11.7	15.6	0.7	0.000
T	0.24	0.38	0.6	0.000		19.2	26.1	0.7	0.000
E2	0.22	0.19	1.2	0.000		33.2	38.1	0.9	0.007
EE2		Comparison not possible				2.8	4.7	0.6	0.000

These statistical comparisons illustrated that in all cases there was a significant difference in values generated by the two methods, RIA typically giving higher values than EIA (Table 3; Fig 3). The exception was E2 for river water samples where EIA gave slightly higher values than RIA. These findings are contrary to the base hypothesis that EIA would give far higher values than RIA. The differences between the two methods were, however, not large: EIA values were on average 75% of the RIA values (mean of ratios in Table 3).

In addition to the assay method being significant in the analyses summarised in Table 3, sample and participant were also significant ($p<0.02$) in all analyses.

A brief summary of an alternative, more complex comparison of the two assay methods undertaken is detailed in Appendix 3.

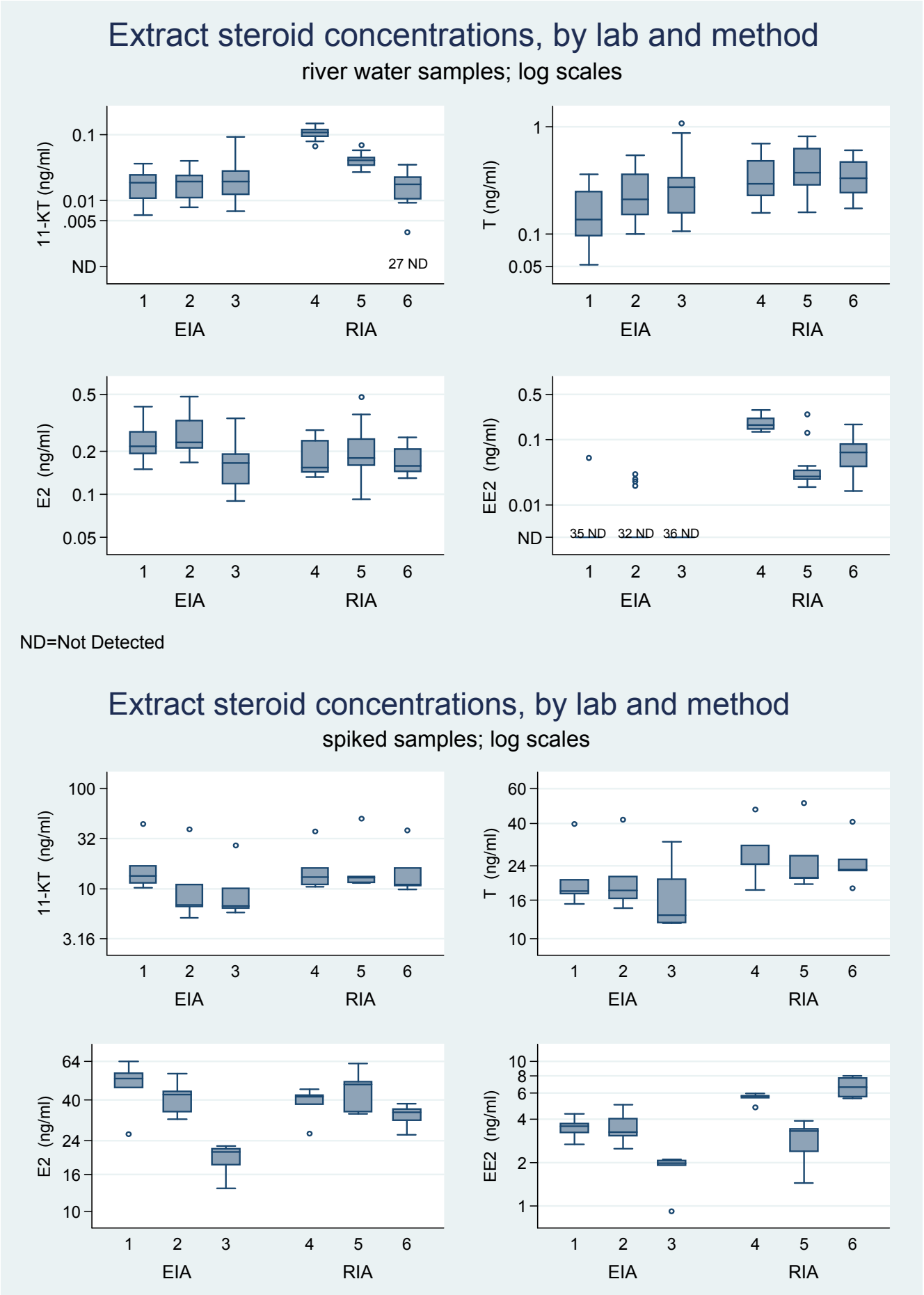


FIGURE 3: Box and whisker plots showing distributions of measured concentrations of steroid (11-KT, T, E2 and EE2) in replicate aliquots by EIA and RIA by participant laboratories (denoted by 1-3 for EIA, 4-6 for RIA). Separate plots for river water samples (top) and spiked samples (bottom) on log scales. Plots show quartiles as box, median as horizontal line within box, whiskers to expected range “adjacent values”, and outliers outside the whiskers.

Are EIA kit results inherently more variable than RIA?
No. The variability of each assay for the determination of each steroid was calculated as the Coefficient of Variation (CV = standard deviation / mean, expressed as percentage; also known as Relative Standard Deviation) for the three measurements by EIA and RIA for each sample (Table 4). Visual comparisons of the CV ranges indicates that RIA provided less variable concentration data for T and E2, whereas the EIA technique provided less variable concentration data for 11KT and EE2 (although the EE2 comparison is compromised by the paucity of EIA values). The higher variability between RIA participants for 11KT and EE2 is likely to be due to the very low amounts in the samples, with the RIAs of one participant consistently producing higher values for these two steroids.

TABLE 4: Basic descriptors (mean, minimum and maximum) of CV, in the amounts measured by the three different participants, in the 44 samples, of the four different steroids, by EIA and RIA.

Steroid	EIA		RIA	
	n	Mean CV (min-max)	n	Mean CV (min-max)
11KT	43	22% (1-76%)	44	92% (4-126%)
T	44	32% (7-88%)	44	16% (3-71%)
E2	44	27% (3-49%)	44	15% (2-43%)
EE2	6	33% (22-47%)	44	81% (24-115%)

The variability (or degree of precision) associated with each technique represents a combination of inter-participant and inter-assay variation as participants were not required to conduct more than one assay. It is therefore unknown whether there are consistent differences between participants, or the variation reflects inter-assay variation that would be revealed between assays repeated by a single participant. Furthermore, samples were not distributed to examine within-assay repeatability (i.e. the degree to which repeated measurements under unchanged conditions show the same results).

Are EIA kits less accurate than RIA?
No. However, it should be noted that our ability to evaluate this performance characteristic of the assays is limited with the current samples and data.

Accuracy is the degree of closeness of measurements of a quantity to its actual (true) value. We only have measurements for a single sample (A – methanol spike) for which an expected value is known. (Samples B-E will be affected by recovery efficiency of the extraction procedure which is discussed below). Expected and reported concentrations for each steroid in sample A are provided in Table 5. It should be noted that the methanol spiked solution that was initially prepared was diluted prior to distribution and again prior to assay. By undertaking dilutions, extra steps are introduced into the chain of procedures and this becomes a potential source of error in this method for assessing accuracy.

TABLE 5: Expected and measured steroid concentrations for Sample A (methanol spike) and percentage accuracy for median measured concentration.

	Expected steroid concentration after dilution (ng/mL)	EIA measured		RIA measured	
		Steroid concentrations (ng/mL)	Accuracy for median conc. (%)	Steroid concentrations .(ng/mL)	Accuracy for median conc. (%)
11KT	49.8	45, 39, 27	78%	37, 50, 38	76%
T	50.1	40, 42, 32	80%	47, 51, 40	94%
E2	50.3	65, 55, 21	109%	46, 63, 38	91%
EE2	2.4	4.4, 5.0, 2.1	183%	6.0, 3.9, 8.0	250%

Accuracies of both EIA and RIA for the measurement of 11-KT, T and E2 were similar and ranged from 76% to 109%. However, both EIA and RIA returned higher than expected concentrations for EE2 (accuracies of 183% and 250% respectively). It is difficult to explain why this was apparent. A plausible explanation is that a weighing error occurred with the EE2 during the initial preparation of the stock solution containing the four steroid compounds from which sample A was derived. The mass of the EE2 weighed out was only 5% of that of the other steroids. In hindsight, rather than combine it with the other steroids to make a solution containing a mixture of compounds, a more appropriate method would have been to prepare the EE2 as an individual solution at a higher initial concentration, which was then serially diluted to the appropriate concentration.

Please note that the design of this study (e.g. 3 participants per assay, one EIA per steroid, inter-laboratory comparisons of sub-divided extracts rather than original samples, no examination of intra-assay variation, limited spiked samples) prevents direct comparisons to previous studies examining the reliability of assay techniques

(e.g. McMaster et al., 2001; Hanselman et al. 2004; Sink et al. 2008; Heath et al. 2010) in which the analysis and presentation of the results also differ. Nevertheless, the indices of accuracy and variability obtained in the present study are in line with such studies (e.g. McMaster et al. 2001: Inter-laboratory CVs of 60% and 70% for measurements of plasma T and E2).

Water steroid levels.

The above comparisons established that EIAs and RIAs give comparable results for water sample extracts, albeit with some statistical differences. We have no grounds for assuming one assay technique gives more reliable or accurate results than the other, or one participant's results were more reliable than another's. Therefore, to examine water steroid levels, the median extract concentration from the 6 values (≥3 values for EE2 due to many values below the EIA limit of detection) was converted to a water steroid concentration (Table 6). The median, rather than the mean, was used intentionally to reduce the effect of possible outliers. It should be noted that within CB0427 all participants correctly converted their measured extract steroid concentration to a water steroid concentration using the equivalent water volume.

TABLE 6: Calculated water steroid concentrations (ng/L) from median extract steroid concentrations (ng/mL) and water volume equivalent. Please note that EE2 values for river water samples were derived almost exclusively from RIA (thought to generate higher values than EIA) as samples were almost universally below the limit of detection of EIA (equivalent to 0.05 ng/L).

No	Water volume equivalent (L)	Sample date	River	Sampling station	Sample treatment	Median extract concentration (ng/mL)				Calculated water concentration (ng/L)			
						11KT	T	E2	EE2	11KT	T	E2	EE2
1	0.986	21/1/10	Test	Upstream	None	0.02	0.17	0.17	0.08	0.02	0.17	0.17	0.08
2	1.000	21/1/10	Test	FF Inflow	None	0.01	0.15	0.16	0.09	0.01	0.15	0.16	0.09
3	0.868	21/1/10	Test	Downstream	None	0.01	0.16	0.14	0.05	0.01	0.19	0.16	0.06
4	0.986	21/1/10	Test	FF outflow	None	0.01	0.29	0.16	0.05	0.01	0.29	0.16	0.05
5	0.994	28/1/10	Avon	Upstream	None	0.01	0.16	0.16	0.03	0.01	0.16	0.17	0.03
6	1.041	28/1/10	Avon	Downstream	None	0.01	0.15	0.16	0.04	0.01	0.15	0.15	0.04
7	0.956	4/2/10	Avon	Upstream	None	0.02	0.23	0.18	0.05	0.02	0.24	0.19	0.05
8	0.992	4/2/10	Avon	Downstream	None	0.01	0.19	0.17	0.05	0.01	0.19	0.17	0.05
9	1.019	11/2/10	Test	Upstream	None	0.01	0.14	0.20	0.06	0.01	0.13	0.20	0.05
10	1.029	11/2/10	Test	FF Inflow	None	0.01	0.14	0.17	0.03	0.01	0.14	0.16	0.03
11	1.023	11/2/10	Test	Downstream	None	0.01	0.21	0.18	0.03	0.01	0.20	0.17	0.03
12	1.028	11/2/10	Test	FF outflow	None	0.02	0.22	0.15	0.03	0.02	0.21	0.14	0.03
13	1.019	21/4/10	Avon	Upstream	None	0.02	0.23	0.14	0.06	0.02	0.22	0.14	0.06
14	0.975	21/4/10	Avon	FF Inflow	None	0.03	0.25	0.15	0.06	0.03	0.25	0.16	0.06
15	1.035	21/4/10	Avon	Downstream	None	0.03	0.26	0.16	0.04	0.03	0.25	0.15	0.04
16	1.007	21/4/10	Avon	Downstream	Replicate 2	0.03	0.25	0.18	0.10	0.03	0.24	0.18	0.10
17	0.980	21/4/10	Avon	Downstream	Replicate 3	0.03	0.23	0.15	0.04	0.03	0.24	0.16	0.04
18	1.040	21/4/10	Avon	FF outflow	None	0.02	0.24	0.17	0.08	0.02	0.23	0.16	0.08
19	1.033	11/5/10	Test	Upstream	None	0.01	0.23	0.14	0.03	0.01	0.22	0.14	0.03
20	1.037	11/5/10	Test	FF Inflow	None	0.01	0.28	0.17	0.07	0.01	0.27	0.16	0.07
21	1.012	11/5/10	Test	Downstream	None	0.02	0.33	0.16	0.06	0.02	0.32	0.16	0.06
22	1.005	11/5/10	Test	FF outflow	None	0.02	0.38	0.16	0.07	0.02	0.38	0.16	0.07
23	1.035	11/5/10	Test	FF outflow	Replicate 2	0.03	0.47	0.19	0.13	0.03	0.45	0.19	0.12
24	0.990	11/5/10	Test	FF outflow	Replicate 3	0.03	0.46	0.17	0.03	0.03	0.46	0.17	0.03
25		25/5/10	Methanol control			0.03	0.05	0.70	0.14				
26	1.097	25/5/10		Deioinsed water		0.00	0.02	0.04	0.02	0.00	0.01	0.04	0.01
27	1.072	25/5/10	Test	Downstream	control	0.03	0.50	0.30	0.04	0.03	0.47	0.28	0.04
28	1.053	25/5/10	Test	Downstream	solvent control	0.03	0.57	0.29	0.08	0.03	0.54	0.27	0.08
29	1.066	27/5/10		Deioinsed water		0.00	0.02	0.07	0.03	0.00	0.02	0.06	0.03
30	1.129	27/5/10	Avon	Downstream	control	0.02	0.38	0.34	0.17	0.02	0.34	0.30	0.15
31	1.061	27/5/10	Avon	Downstream	solvent control	0.04	0.52	0.29	0.11	0.04	0.49	0.27	0.10
32	1.056	8/6/10	Avon	Upstream	None	0.03	0.48	0.28	0.09	0.03	0.45	0.26	0.08
33	1.054	8/6/10	Avon	FF Inflow	None	0.03	0.53	0.27	0.12	0.03	0.50	0.25	0.11
34	1.097	8/6/10	Avon	Downstream	None	0.03	0.43	0.27	0.11	0.03	0.40	0.25	0.10
35	1.040	8/6/10	Avon	FF outflow	None	0.03	0.51	0.27	0.09	0.03	0.49	0.26	0.09
36	1.086	15/6/10	Test	Upstream	None	0.02	0.37	0.22	0.06	0.02	0.34	0.20	0.06
37	1.045	15/6/10	Test	FF Inflow	None	0.02	0.39	0.23	0.07	0.02	0.37	0.22	0.07
38	1.076	15/6/10	Test	Downstream	None	0.02	0.51	0.23	0.06	0.02	0.47	0.22	0.06
39	1.043	15/6/10	Test	FF outflow	None	0.02	0.61	0.24	0.23	0.02	0.58	0.23	0.22
A		25/5/10	Methanol spike			39	41	51	4.7				
B	1.073	25/5/10		Deioinsed water	spiked	11	21	43	3.9	11	20	40	3.6
C	1.041	25/5/10	Test	Downsteam	spiked	11	19	34	3.2	10	18	33	3.0
D	1.045	27/5/10		Deioinsed water	spiked	10	17	26	3.1	10	16	25	2.9
E	1.029	27/5/10	Avon	Downsteam	spiked	15	23	42	3.5	14	22	41	3.3

Is the extraction methodology 100% efficient?
No. The recovery of the steroids from the water samples can be assessed by comparing the (median) measured steroid concentration in the spike methanol (A) to the concentrations in the extracts of the spiked water samples (after correcting for the amount in unspiked water samples) (Table 7).

TABLE 7: Calculation of recovery efficiency of the methodology to extract steroids from water samples into aliquots for assay.

Steroid	Spike added (A) (ng/L)	Spiked / corresponding unspiked water sample number	Measured in spiked water sample (ng/L)	Measured in equivalent unspiked water sample (ng/L)	Difference (ng/L)	Recovery	Mean recovery
11KT	39	B/26	11	0.00	11	27%	29%
		C/28	10	0.03	10	26%	
		D/29	10	0.00	10	25%	
		E/31	14	0.04	14	37%	
T	41	B/26	20	0.02	20	48%	46%
		C/28	18	0.57	18	43%	
		D/29	16	0.02	16	40%	
		E/31	22	0.52	22	54%	
E2	50	B/26	40	0.04	40	79%	69%
		C/28	33	0.29	33	65%	
		D/29	25	0.07	25	49%	
		E/31	41	0.29	41	81%	
EE2	4.7	B/26	3.6	0.02	3.6	77%	67%
		C/28	3.0	0.08	3.0	62%	
		D/29	2.9	0.03	2.9	61%	
		E/31	3.3	0.11	3.3	68%	

These limited results (and compromised due to accidental transfer of some of sample D to E during rotary evaporation) indicate that the efficiency of the solid phase extraction method (and storage) varies between steroids. Recovery was higher for the oestrogens (49-81% for E2 & EE2) than the androgens (25-54% for 11KT & T). This may be due to the extraction methodology (Rubio et al. 2004) being suggested for oestrogens. RIA by Cefas of spare subsamples of 1-8 retained after C18 extraction, but not exposed to aminopropyl SPE clean-up indicates that this latter stage does cause loss of steroid: 60% loss for T but only 15% loss for E2. However, whether the differences in recovery efficiency between the steroids, and between individual water samples, is an artefact of the present data, or does represent real differences would require further and extensive work.

Unless recovery is established during method validation and shown to be within the range that is generally accepted (i.e. 80 to 110%) and reproducible (i.e. CV≤10%), then measurements should be corrected for recovery efficiency before reporting as water steroid concentrations. However, correction for recovery efficiency has not been applied here because

- definitive correction factors for the methodology have not been derived from this work
- the recovery information from the spiked river water samples (based upon concentrations reported within SF0241) may not be applicable at the much lower, observed river steroid concentrations
- comparison to the SF0241 reported steroid concentrations is required, which were not corrected for recovery efficiency.

The reported concentrations should, therefore, not be considered absolute, due to uncertainties with the efficiency of the extraction procedure.

How variable is the extraction and assay methodology?
The reproducibility of the extraction (and assay) methodology can be partially examined (Table 8) from the two sets of three replicate river water samples (#15-17; #22-24). Reproducibility is acceptable (CV ≤11%) for 11KT, T, and E2. [NB: The CV values are derived from the data in Table 6, representing medians of the ≤6 values from RIA and EIA]. However, variation for EE2 is noticeably higher for both sets of samples (CV ≥58%), indicating less reliable values for EE2. This will reflect the variation in RIA values for EE2 at low concentrations.

TABLE 8: Variation (expressed as CV) in water steroid concentrations from replicate samples (n=3).

Steroid	R Test Samples 22-24	R Avon Samples 15-17
11KT	7%	6%
T	11%	3%
E2	7%	8%
EE2	67%	58%

Do the CB0427 results correspond with the SF0241 steroid concentrations in the R Test and R Avon?
No. Mean steroid concentrations reported in SF0241 (samples collected 9/12/02 – 1/6/06) and this project - CB0427 (samples collected 21/1/10 - 16/6/10) - are collated in Table 9. There are major differences (2-3 orders of magnitude) for 11KT, T and E2, and a 3 fold difference for EE2. Furthermore, SF0241 reported values indicated that 11KT>T>E2>EE2, whereas the current results indicate that T>E2>EE2>11KT. There is therefore not only disagreement in the absolute levels of the steroids, but also a marked difference in the relative amounts of the individual steroids.

TABLE 9: River water steroid concentrations (ng/L) found within the present study (CB0427) and reported in SF0241 (NB: Winter and Spring values combined). Values are means ± standard error. Please note that EE2 concentrations were reported as pg/L in SF0241.

Steroid	R Test					R Avon				
	CB0427		SF0241		Fold difference	CB0427		SF0241		Fold difference
	n	Mean ± SE (ng/L)	n	Mean ± SE (ng/L)		n	Mean ± SE (ng/L)	n	Mean ± SE (ng/L)	
11KT	19	0.02 ± 0.00	42	20 ± 3	1000	15	0.02 ± 0.00	42	32 ± 5	1600
T	19	0.31 ± 0.03	42	19 ± 3	59	15	0.29 ± 0.03	42	25 ± 3	83
E2	19	0.18 ± 0.01	42	17 ± 4	89	15	0.20 ±.0.01	42	18 ± 2	90
EE2	19	0.07 ± 0.01	42	0.24 ± 0.02	3.4	15	0.07 ± 0.01	42	0.17 ± 0.01	2.4

How do the SF0241 and CB0427 results compare to model predictions?
Richard Williams and Andrew Johnson (CEH Wallingford) provided predicted river concentrations of E2, EE2 and Estrone (E1, not examined in either SF0241 or CB0427) for the two locations. These concentrations were derived from a model (LF200-WQX) that calculates a distribution of concentrations that are likely to occur from loadings to sewage treatment works (STW) upstream and river flows. The model takes into account the variability in inflow load to STW, the effectiveness of removal by the STWs, and variations in river flows. Two values were provided – the mean and the 90th percentile, that latter being the concentration that is expected to be exceeded only 10% of the time (Table 10).

TABLE 10: Mean and 90th percentile values predicted by a model for three steroids in the R Test and R Avon at the study sites. See text for more detailed explanation.

	R Test		R Avon	
	Mean (ng/L)	90th %tile (ng/L)	Mean (ng/L)	90th %tile (ng/L)
E2	0.031	0.067	0.049	0.093
EE2	0.009	0.019	0.014	0.026
E1	0.26	0.67	0.45	0.78

The model predictions for both E2 (<0.1 ng/L) and EE2 (<0.03 ng/L) are considerably lower than values reported in SF0241 (cf values in Table 9). The model predictions are also lower than CB0427 measured values (uncorrected for recovery) (cf values in Table 9). Whether this latter difference represents error in either the CB0427 measurements, the data inputs to the model, assumptions made for the model, or omissions from the model (e.g. non-inclusion of other sources of steroids such as farmed livestock or manure run-off (Kolodziej et al., 2004; Shore et al. 2004; Zhao et al. 2010) may merit future research. Balaam et al. (2010) similarly found that model predictions for E2 were an order of magnitude lower than measured concentrations.

Do the CB0427 results reproduce the SF0241 effects of fish farms and season?
No. SF0241 reported that levels of 11KT, T and E2 were elevated downstream of fish farms, in both rivers, in both seasons by 1.3 to 5.9 fold (based on average values). Furthermore, SF0241 reported that "*on the whole, water hormone concentrations were lower in the Spring compared to the Winter*". The CB0427 water steroid concentration data (in ng/L, $\log_{10}$ transformed) were therefore analysed (ANOVA GLM in Minitab) to examine effects of River (Test v Avon), season (Winter = Jan/Feb; Spring = April-June) and sampling station ("Upstream", "Downstream", Fish farm inflow, Fish farm outflow) for each of the 4 steroids. (N.B. Due to limited degrees of freedom for the analyses, sampling occasion was not included and neither was the 3-way interaction; 2-way interactions were tested in sequential analyses and dropped when non-significant).

The only effect that was significant (Table 11) was season, with concentrations consistently being greater in the Spring than Winter. This is likely to be explained, at least in part, by river flows now recognised as a major factor affecting river steroid levels (Johnson, 2010). River flow rates at the times of the Winter sampling (Jan/Feb 2010) were greater than at the times of the Spring sampling (April-June 2010) (GLM ANOVA p=0.003; Least Squares Mean 11.1 v 6.9 m^3/s; River flow data in Appendix 4). Higher flows would increase dilution and thereby decrease water steroid concentrations. Interestingly SF0241 found the opposite effect, i.e. water steroids levels were higher in Winter. River flows were similarly higher in the Winter than in the Spring for the SF0241 sampling dates (GLM ANOVA of flow data in Appendix 4 for SF0241 sampling dates; p=0.041; Least Squares Mean 8.1 v 6.2 m^3/s), although local water flows due to sluice operation may have differed.

TABLE 11: Results of ANOVA assessing the effects of river, season and sampling station on river water steroid level data collected in CB0427 (n=34; spiked, de-ionised, and blank-methanol data excluded). LSM refers to Least Squares Mean values, and p indicates significance of factor.

	River			Season			Sampling station				
	R Test LSM (ng/L)	R Avon LSM (ng/L)	p	Winter LSM (ng/L)	Spring LSM (ng/L)	p	Up-stream (ng/L)	Fish farm inflow (ng/L)	Down-stream (ng/L)	Fish farm outflow (ng/L)	p
11KT	0.02	0.02	0.135	0.01	0.02	0.000	0.02	0.02	0.02	0.02	0.912
T	0.26	0.24	0.494	0.18	0.35	0.000	0.23	0.27	0.24	0.27	0.525
E2	0.18	0.19	0.742	0.17	0.20	0.071	0.18	0.20	0.18	0.18	0.884
EE2	0.06	0.07	0.434	0.05	0.08	0.051	0.05	0.10	0.05	0.06	0.308

It should be noted that although sampling station was not significant in any of the analyses, for Testosterone the interaction term river*sampling station did approach significance (p=0.073).

Do trout farms release steroids?

The location of water sample collection (sampling station) was not significant in any of the four analyses above, contrasting with the SF0241 finding of elevated levels of the natural steroids downstream of fish farms. (NB: The SF0241 upstream and downstream sites are on different river channels, so any difference cannot be assumed to be solely due to the fish farms). However, the statistical analysis above (Table 11) does lack sensitivity for assessing whether fish farms do increase river steroid levels - if effects are masked by differences between rivers or sampling occasions. Paired t-tests of the steroid levels in fish farm inflow and outflow samples, on each sampling occasion, for each river (Table 12) allow a more direct comparison of the water flowing into and out of the fish farms, although the power of the tests has to be recognised as severely limited by the sample size (n=4 pairs for R Test, n=2 pairs for R Avon). This analysis, and visual inspection of the CB0427 data, largely supports the null hypothesis that fish farms do not elevate river steroid levels.

TABLE 12: Comparisons of water steroid levels in fish farm inflow and outflow samples. Mean values given with probabilities from paired t-tests.

Steroid		R Test			R Avon			
	N	Mean inflow (ng/L)	Mean outflow (ng/L)	p	n	Mean inflow (ng/L)	Mean outflow (ng/L)	p
11KT	4	0.013	0.018	0.182	2	0.030	0.025	0.500
T	4	0.23	0.36	0.021	2	0.38	0.36	0.205
E2	4	0.18	0.17	0.718	2	0.21	0.21	0.500
EE2	4	0.07	0.09	0.559	2	0.09	0.09	1.000

In one analysis – testosterone in R. Test – was there a statistically significant increase. It must be remembered that statistical significance is a guide, rather than a definitive assessment, as to whether factors have a real effect on the measured parameter. The significant effect for testosterone must be viewed in relation to the multiple testing of the data set (8 separate analyses) which will increase the risk of Type I errors (rejecting the null hypothesis when it is true) (Petrie & Sabin, 2000). If a probability level of 0.05 is used for rejecting the null hypothesis and accepting an effect, then there is a 1 in 20 likelihood that any single result could have occurred simply by chance without there being a true difference; this likelihood increases to 8 in 20 in the current set of analyses. If the Bonferroni approach for correcting for multiple testing is applied, then a significance level of 0.006 (0.05 ÷ 8) would be required (Petrie & Sabin, 2000) and the increase in testosterone observed is not significant. Furthermore, the apparent significant increase (0.13 ng/L) due to the farm may be an artefact – a product of the

degree of variability (precision) associated with the determination of testosterone using the assays: the variation between replicate samples from the R. Test (#22-24) was 0.08 ng/L.

However, on balance, it must also be recognised that the apparent increase in testosterone at the R. Test site is quite possibly a real effect:

- the Bonferroni approach is recognised as conservative (Petrie & Sabin, 2000) and can lead to Type II errors (i.e. not rejecting the null hypothesis when it is false).
- the power of the statistical test is severely limited by the small sample size (n=4)
- the difference between inflow and outflow is in the direction expected if the fish farm were releasing testosterone
- visual examination of the separate "downstream" station levels for the R Test (Table 6) supports a slight elevation in testosterone from the trout farm
- although the farm holds only female rainbow trout (Appendix 1), female rainbow trout do produce noteworthy amounts of testosterone (Scott et al. 1980; Sumpter et al. 1984)

Only additional, more extensive sampling could clarify whether this apparent increase in testosterone from the R. Test fish farm is a real effect or an artefact. However we do not consider this worthwhile because if the apparent 0.13 ng/L increase is real, then it is already well below levels thought to precipitate biological effects, and will be diluted when the outflow mixes with the river water in the receiving channel, and further when joining other channels.

It should be noted that the sampling and assays were directed at the biologically active (free = non-conjugated) form of the 3 natural steroids. The free form typically comprises only a small component of excreted steroids, the majority being released as conjugates (i.e. sulphate or glucuronide) via the urine and bile within faeces (e.g. Vermeirssen & Scott 1996). After release into the river, natural microbiological processes might deconjugate and hence reactivate excreted conjugated steroids (Heath et al. 2010), thereby contributing additional free steroid to levels in the river.

In addition to the SF0241 and CB0427 data for steroids emanating from trout farms:

- Kolodziej et al. (2004) used gas chromotagraphy-tandem mass spectrometry (GC/MS/MS) to examine water steroids (estrone, testosterone and androstenedione) in a limited number of water samples taken in the inflow, outflow and downstream of US rainbow/steelhead trout farms. Steroids were detected in some effluent samples at levels up to 0.7 ng/L, but were undetectable in others; it was implied that this was due to differences in fish size / maturity. River steroid levels immediately downstream of caged adult Chinook salmon approached 1.7 ng/L, and steroids were detected at around 0.5 ng/L in a wild salmonid spawning ground. It should also be noted that levels recorded around dairy cattle farms were 2-3 orders of magnitude higher.
- Shore et al. (2004) found that Israeli fish ponds had 1.0-5.8 ng/L estrogen (E1+E2), as measured by RIA, but provided no detail on the sites (e.g. fish species farmed, farm management practices). The possibility that the steroids derived from cattle pasture runoff (shown to be a contributing factor to river levels) into the inflow, or manure applied directly to the pond, cannot be discounted.
- Barel-Cohen et al. (2006) found that Israeli fish pond effluent contained 5-10 ng/L of both testosterone and estrogen (E1+E2), as measured by RIA. Again no detail is provided on the sites and the possibility that the steroids derived from other sources than the fish themselves cannot be discounted.
- Pelissero et al. (1989) found that commercial fish diets incorporating fish meal contained notable amounts of androgens (11KT+T: ≤11 ng/g), E2 (9.4 ng/g) and E1 (6.2 ng/g), as measured by RIA. However, we are not aware that this work has been followed up in subsequent publications or independently substantiated, so it must be considered a single unconfirmed study.

Based on the evidence available, it is judged that there is no urgent requirement to examine the steroid output from UK fish farms, due to their probable minor contribution relative to STW and farmed livestock. However, the possibility of a small, localised contribution to river steroid levels cannot be dismissed.

HPLC analysis of immunoactive compounds in water extracts

This objective was dropped. High Performance Liquid Chromatography (HPLC) provides a means of separating the various compounds within a sample extract into different fractions. The fractions that particular steroids occur in are well documented and can be confirmed by passage of radiolabelled steroids. Had the EIA/RIA comparison supported the original hypothesis of false positive measurements from EIAs, then EIA of separate HPLC fractions of the water sample extracts would have provided a means of further examination:

- if the bulk of the observed EIA immuno-activity had occurred in the appropriate fraction(s), then it is likely to represent the target steroid
- if the bulk of the observed EIA immuno-activity had occurred in the other fractions, then the assay is definitely not detecting the target steroid.

However, contrary to the original hypothesis, the EIA results did not markedly exceed the RIA results and EIA results were far lower than those reported in SF0241. This objective was therefore deemed irrelevant. Furthermore, the amounts of the immuno-reactive compounds in the samples were insufficient for examination of individual HPLC fractions.

Can the disparities between the SF0241 and CB0427 results be explained?
No explanation could be firmly accepted that would explain the differences between the SF0241 and CB0427 results, i.e.:

- the major differences between studies in absolute levels
- the change in relative levels (i.e. 11KT being lowest in CB0427 rather than highest in SF0241)
- the loss of effect of trout farms (apart from possibly T at the R. Test site)
- a reversal of the effect of season
- the lost ability to detect river EE2 levels by EIA.

The samples were collected several years apart (SF0241: 9/12/02 – 1/6/06; CB0427: 21/1/10 - 16/6/10) and the assays were conducted by different personnel. Hypotheses that have been put forward to account for the differences are discussed below.

Hypothesis 1: SF0241 results invalid due to miscalculation and/or assaying errors
Cefas personnel met with the SF0241 PI to review records of the assay procedures and associated data for the SF0241 assays. However, none of the raw assay data or calculation spreadsheets had been retained. It was therefore impossible to establish, or disprove, that miscalculation or assaying errors occurred. The SF0241 PI acknowledged personal unfamiliarity with EIAs and that the SF0241 EIAs were conducted by other staff. Upon examination of extant spreadsheets of related, but unreported samples, the SF0241 PI recognised and accepted two probable sources of error in the SF0241 results:

- A 1000 fold error - due to the spreadsheet column containing the calculated values for the 1 mL extracts being labelled as “ng/L” which was then misinterpreted for the concentration in the original 1 litre water sample. This is recognised as the probable primary cause of the very high SF0241 water steroid concentrations.
- Erroneous values – a spreadsheet used to calculate EE2 concentration from EIA absorbance generated erroneous values. This would compound the 1000-fold error.

Further anomalies were apparent in the spreadsheets, but as these were for preliminary analyses of early samples (when SF0241 staff were gaining experience), they may or may not have been perpetuated throughout the study.

Hypothesis 2: SF0241 results invalid due to differing sampling procedures to CB0427
During SF0241 water sampling, concurrent caging and sampling of fish at the sampling stations disturbed the substratum. The SF0241 PI has suggested that the SF0241 water samples may have become contaminated with substratum-derived sediment which could have then contributed to the measured water steroid levels. All water samples (both upstream and downstream) would have to have been contaminated with variable amounts of sediment. No attempt was made during CB0427 to explore this hypothesis due to the difficulty in replicating variable substratum disturbance and the considerable additional effort required to assess sediment loads, sediment steroid content, and the leaching rate of steroids from sediment particles into water samples during transport, overnight storage and processing.

Hypothesis 3: SF0241 results invalid due to unreliable EIA kits which have since improved
The same suppliers were used and, as far as we can tell given a lack of documentation from SF0241, the methodology and standards have remained the same.

Hypothesis 4: SF0241 results invalid due to external sources of contaminants in the water samples interfering with the EIAs which have since decreased
There is no information available. The aminopropyl step in the extraction procedure is intended to clean-up the samples, and was used within both SF0241 and CB0427.

Hypothesis 5: SF0241 results invalid due to samples becoming contaminated during collection, extraction or assaying.
Samples can be contaminated with steroids excreted from researchers themselves, or from standards in the laboratory. All SF0241 samples, collected and processed across 3½ years, would have to have been accidentally contaminated with all 4 steroids to varying degrees. 11-KT, reported to be the dominant steroid in the SF0241 results, could not be derived from human contamination as it is only produced by male fish. Other compounds may have contaminated the samples (during collection, processing and assaying) and interfered with the SF0241 EIAs, but this possibility is also considered remote.

Hypothesis 6: SF0241 results valid, but fish farm practices have since changed or were not representative at the times of the CB0427 sampling

Major changes in farm practices would have to have occurred between the two study periods to explain the 2-3 order of magnitude differences in natural steroid concentrations. The farm managers supplied information about their sites against which this hypothesis can be judged (Appendix 1). This information does not indicate any fundamental changes in farm practices. Environment Agency water quality information (R. Test site only) supplied by the farm manager further indicates that the physiological loading of fish within this farm had not changed between the two periods. During CB0427 the fish farm managers were aware of the general timing of the sampling, but during SF0241 the trout farm managers were largely unaware.

Hypothesis 7: SF0241 results valid, but fish farm diets have changed
The fishmeal content of fish diets (a putative source of steroids; Pelissero et al., 1989) may have decreased between the SF0241 and CB0427 sampling periods. It is considered unlikely that diet changes could account for the magnitude and number of differences observed.

Hypothesis 8: SF0241 results valid, but river flows were greater for CB0427
No information is available on water flow rates within the local river channels. (NB: The flow data reported within SF0241 has units of m/s, and therefore represents current speed rather than water flow). The ranges of river flow recorded at official monitoring stations downstream on the sampling days within CB0427 were within the ranges observed for the sampling days in SF0241 (Appendix 4: R Avon SF0241: 1.7 – 8.6; CB0427: 2.8 to 7.3 m^3/s; R Test SF0241: 6.7 – 18.6; CB0427: 8.7 – 16.1 m^3/s). On average flows were 24% higher on the CB0427 sampling dates than the SF0241 dates (Least Squares Mean 8.8 v 7.1 m^3/s; ANOVA, p= 0.075), but this would not account for the magnitude of the differences.

Hypothesis 9: SF0241 results valid, but timing of sampling differed between SF0241 and CB0427
The winter samples for CB0427 were collected later in the season than for SF0241, but the Spring season sampling showed a considerable overlap. All river water samples were collected at a similar time of day (morning within the working day) in SF0241 and CB0427.

Hypothesis 10: SF0241 results valid, but recovery from CB0427 samples was poor
The SF0241 extraction and assay procedure was followed as closely as possible for the CB0427 samples. SF0241 did not include any recovery (or blank sample) data, and recovery can therefore be assumed to be similar to CB0427 (or possibly lower due to the longer periods of storage; Appendix 2). Even if 100% recovery had been achieved in SF0241, recoveries of 25-81% observed in CB0427 would not explain a 2-3 order of magnitude difference to SF0241.

Hypothesis 11: SF0241 results valid, but external sources of steroids have changed
Changes in external sources, e.g. wild fish, livestock, wild birds, sewage treatment plants, have been suggested to account for the differences between the SF0241 and CB0427 data. Information is lacking to support or refute this, although data on sewage treatment plants and wild fish could be sought from the water companies and the Environment Agency if considered worthwhile pursuing.

In conclusion the available evidence points towards *Hypothesis 1: SF0241 results invalid due to miscalculation and/or assaying errors* as the most probable explanation of the differences between the SF0241 and CB0247 results. However, this cannot be confirmed as records relating to the SF0241 assay procedures, results and ensuing calculations were not available. The possibility that other alternate hypotheses are true or contributed therefore cannot be discounted. However, it is not considered worthwhile pursuing these unless additional evidence is forthcoming which allows Hypothesis 1 (as the most probable) to be rejected. Nevertheless, the current project has demonstrated that the SF0241 reported water steroid levels are not typical of river steroid concentrations in 2010.

Are steroid measurements always accurate and reliable?
No. Steroid concentration data are not simple direct measurements like length or temperature data. Both EIA and RIA techniques provide indirect measurements of steroid concentrations which require a large number of procedures including:

- extracting (and cleaning) a known volume of water (or plasma)
- drying down the extract
- making up the extract in a known volume of assay buffer
- making a known dilution of the extract so that, when it is added to the assay tube/well, its reading will fall within the range of the standard curve
- making up a series of standards at the correct concentrations
- using appropriate mathematical tools to calculate the concentration of samples in the assay tubes
- rejecting values that fall too high or too low on the curve
- being familiar with units in which the standards are expressed (which may variously be ng/ml, pg/ml, ng/L, µg/ 100 ml and pg/tube)
- multiplying the calculated concentrations by the dilution ratio of the extract
- correcting for the volume of assay buffer that was added to the extract

- correcting for the volume of water that was originally extracted
- correcting, ideally, for extraction efficiency.

Both assay techniques also require much manual manipulation of samples and standards with frequent dilutions and placement within an assay template in a known order. As a result of this complexity, there are many opportunities for human error.

The present project itself has illustrated that such errors are easily made (Table 2B), even by researchers experienced in the techniques and aware that their results will be scrutinised, e.g.

- miscalculation – entering the wrong values for calibration standards in software calculation packages generates erroneous data for samples
- miscalculation – entering the wrong dilution factor for a sample
- miscalculation – using values from a different assay to predict steroid concentrations
- manipulation errors - mis-ordering of samples within an assay and thereby attributing erroneous values to samples
- reporting of data extrapolated beyond the extremes of the standard curve (samples should be re-assayed after further dilution or concentration).
- Contamination - solvents can become contaminated when preparing spikes (or standards). The blank methanol (sample 25) was clearly contaminated with E2 and EE2 (but not 11KT or T), although the source of the contamination could not be identified.
- Cross contamination of samples - some of sample D was accidentally transferred into sample E during rotary evaporation. (It should be noted that this was recognized and recorded at the time).
- Miscalculation/weighing errors can occur during preparation of spiking solutions or calibration standards – a possible explanation for the ca. 200% accuracy value for EE2 (Table 5).
- Readings from blank samples falling on the standard curve, mis-indicating a low concentration of steroids within a quality control sample.

Other gross errors can occur from measurement equipment, antibody or enzyme malfunction. Additional minor errors are also inevitably associated with

- imprecision in manual pipetting of sample extracts and standards and making dilutions
- the estimation of concentrations from calibration curves
- the "distance" (several times removed) between the actual measurements (radioactivity or optical density) and the value of interest, i.e. the steroid concentration in the extract

It is these smaller errors that affect the accuracy and precision of the techniques.

It has to be recognised that the process of immunoassay is not a simple task and experience does count. There is a danger with EIA kits in that people follow the instructions without appreciating the principles of the process. Staff need to be trained and supervised until they are at ease making up dilutions, handing regression curves, using mathematical transformations to generate measured values, etc.

What "lessons can be learnt" for future steroid measurement research?

This project has illustrated the potential for errors to occur in the generation of steroid data (Table 2B). What is needed are safeguards to ensure that erroneous data is spotted and corrected. This was possible within CB0427 due to independent assay of the same samples. Safeguards for future could include:

- typical assay validation work, e.g. examination of cross-reactivity, accuracy, inter and intra-variation, parallelism.
- documentation of extraction recovery efficiency and reproducibility, e.g. through spiking, blanks, replicates
- "blind" assaying to eliminate the possibility of bias in selecting reported data
- guidance for internal scrutiny of assay data with tolerance limits, e.g. for duplicate variation, standard curve fits, etc
- comparisons to previously reported values for similar water bodies and/or values predicted by models.
- independent scrutiny of data by a recognised expert, or more extremely, independent extraction and assay of replicate samples
- retention of raw assay data and associated calculations used to derive final steroid concentration values in case of queries.

Requirements for safeguards will inevitably increase the cost of generating water steroid data. It seems that guidance on best practice might be valuable indicating the level of such requirements, which will be relative to the policy importance of the data generated.

Acknowledgements

Thanks to: the farmers and land-owners for their cooperation and discussions; Richard Williams and Andrew Johnson (CEH) for supplying 2010 river flow data and model predictions for water steroid levels in the two rivers; John Elliott (EA) for provision of maps of the R Test river system.

Appendix 1: **Details of subject trout farms and husbandry factors**.

This information has been collated from discussions between the CB0427 project manager and the fish farm managers. It has been examined by Cefas Fish Health Inspectors familiar with the sites.

R Test site

This long-standing site is owned by a company producing rainbow trout for the table market. It is comprised of concrete raceways and is considered an intensive farm. The site holds only female (diploid) rainbow trout (*Oncorhynchus mykiss*). Fish are brought onto site from hatcheries and on-grown to harvest. The approximate size range of fish on site is therefore 5 to 800 g. Fish are harvested before becoming reproductively mature, although occasional fish are noted with ovarian development. Supplemental oxygen is supplied, and swirl concentrators reduce the suspended solids load in the outflow. The local sewage treatment works (STW) discharges into the main river 0.4 Km downstream of the fish farm. There is another rainbow trout farm in the vicinity (1 km upstream), but its discharge joins the main river downstream of where the channels supplying the sampling stations diverge. The farm manager supplied details of stocks on site and feeding rate for all the water steroid sampling dates within SF0241 and CB0427, in addition to water quality data from EA inspections closest to the sampling dates of the two projects.

- Water flow (abstraction) through the farm has not changed between 2002-2006 and 2010
- The infrastructure has not changed: the raceways are original, the swirl concentrators were installed in the 1980s, and oxygenation was installed in the mid 1990s.
- The management personnel of the site have not changed.
- The species, gender and ploidy of the fish stocks have not changed.
- It was not assessed whether the origins of the stocks have changed. The origins depend upon what is available from commercial hatcheries which will have sourced eggs from UK, USA, Denmark and South Africa. The information could be tracked, albeit with some effort, if required.
- The total fish biomass on site has not changed (n=21 and 5, ANOVA p=0.7).
- There are differences in the mean size and numbers of fish on site. On average, on the SF0241 sampling dates there were 17% more (n=21 and 5, ANOVA p=0.05) but 24% smaller (n=21 and 5, ANOVA p=0.02) fish on site than on the CB0427 sampling dates.
- The amount of feed fed per day has not changed (n=21 and 5, ANOVA p=0.96).
- The supplier of the feed has not changed. (It should be noted that commercial feed suppliers do vary the composition of diets over the short and long terms, due to availability and cost of ingredients).
- The EA recorded outflow oxygen saturation, inflow ammonia, outflow ammonia, and ammonia increase from the farm have not changed (n=21 and 5, ANOVA p=0.90, 0.97, 0.36, 0.24 respectively).

R Avon site

This long-standing site is owned by an angling club and is comprised of three earth ponds; additional small tanks are occasionally used for temporary rearing of small fry. It is considered a low intensity farm. The site holds only normal (diploid) mixed sex brown trout (*Salmo trutta*): broodstock and their progeny being on-grown to restock local angling waters. Reproductively mature fish can therefore be expected. Fish on site range between approximately 1 g and 1 kg. Eggs (from the broodstock) are hatched and initially reared at a separate off-site hatchery, brought back to the site at 1-2 g and on-grown for approximately 2 years before stocking into the river. The farm manager supplied (from farm movement records) very approximate numbers (to nearest hundred or thousand) and individual sizes (either in inches or pounds) of stocks on site for all the water steroid sampling dates within SF0241 and CB0427. No flow rate data is collected; flows through the farm do vary over time, being largely dependent upon the amount of debris collecting on the inflow screen. There is no water treatment on site, and water quality data is not collected. Feed data is not retained; however, feeding is consistent over time, and approximate amounts were provided from memory. There is a nearby lake fishery stocked with rainbow trout, but the outflow joins the R Avon downstream of "SF0241 downstream" station, so can be excluded from considerations of this site. The farm manager indicated that during 2002-2006, grayling were common in the river and spawned in the spring. Pike and roach were also considered common in the vicinity.

- The infrastructure of the site has not changed between 2002-2006 and 2010.
- The species, gender and ploidy of the stocks has not changed.
- The fish biomass on site has not changed (n=21 and 5 respectively. ANOVA, p=0.8).
- The mean individual fish size on site has not changed (n=21 and 5 respectively. ANOVA, p=0.2).
- The origins of the stocks have changed. The stocks of fish held in 2002-2006 came from various sources across the UK; in 2010 stock originated solely from local rivers.
- The timing of maturation has changed: in 2002-2006 broodstock would show signs of maturing from late September and would be stripped for gametes from mid November to early December. In 2010, the broodstock was stripped in late January / early February, at the time the local wild brown trout spawn.
- The amount of feed fed per day has not changed
- The supplier of the feed has not changed. (It should be note that commercial feed suppliers do vary the composition of diets over the short and long terms, due to availability and cost of ingredients).

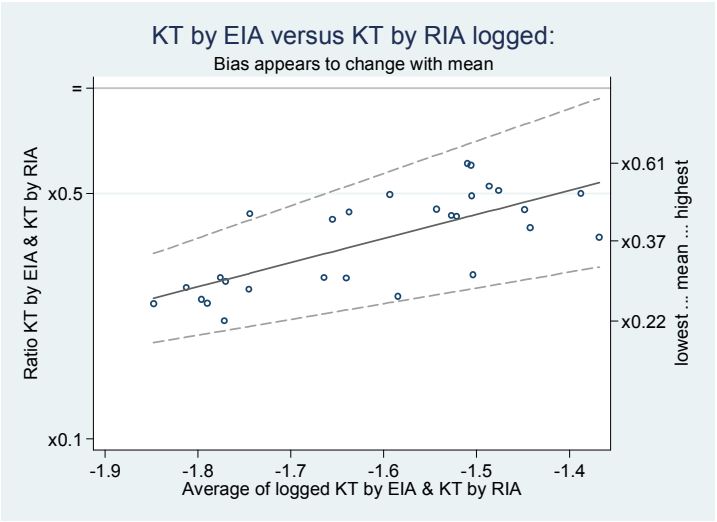
Appendix 2: Details of sample extraction methodology employed in CB0427 and SF0241

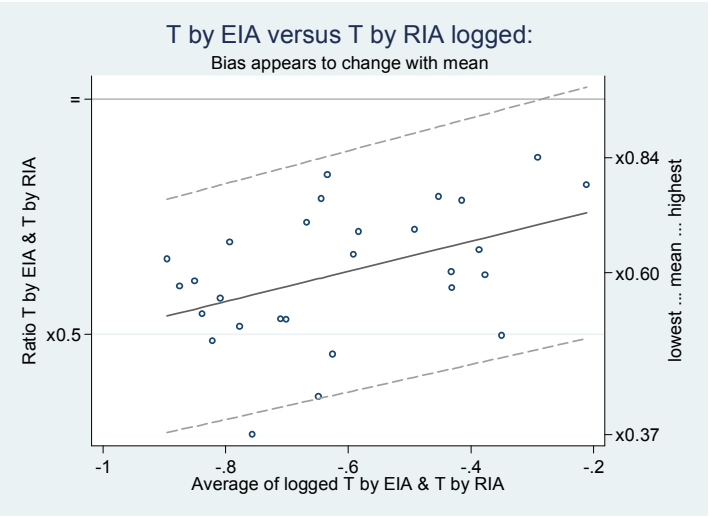
	CB0427	SF0241
Sample collection	Water samples collected in 25 L HDPE drums	Water samples collected in 1 or 2 litre glass bottles
Transportation to lab	Transport to Cefas Weymouth or Portsmouth University	Transport to Portsmouth University
Acidification	(Post filtration)	2 ml of HCl.
Filtration	Vacuum filtration through GFC filter (150 mm) to remove particulates	Vacuum filter through GFC filter (50 mm) to remove particulates
Acidification	2 ml/L HCl (36%)	(Pre-filtration)
Overnight storage	Overnight storage in dark at 4°C	Overnight storage at 4°C
Precondition C18 solid phase extraction columns (SPEC)	Varian Bond Elut (1 g, 6 mL): 5 ml methanol followed by 10 ml of distilled water.	Varian Bond Elut (1 g, 6 mL): 5 ml methanol followed by 10 ml of distilled water.
Sample extraction	Pump 1 L of acidified, filtered water per sub-sample at <15 mL/min through SPEC. Typically 9 sub-samples per sample.	Pass filtered water sample (1 or 2 L) through SPEC (max flow rate 10-15 ml/min) under vacuum
Wash SPEC	Wash SPEC with 5 ml distilled/deioinsed water, allow to dry under air / under vacuum for 1 min, wash SPEC with 5 ml hexane	Wash SPEC with 5 ml distilled water, allow to dry under air for 1 min, wash SPEC with 5 ml hexane
Elution of SPEC	Elute with 5 ml dichloromethane into glass stoppered tubes	Elute with 5 ml dichloromethane into glass stoppered tubes
Dichloromethane evaporation	Evaporate under nitrogen at 37°C	Evaporate under nitrogen at 37°C
Redissolving	Add 1 ml methanol and vortex	Add 1 ml methanol and vortex
Freeze storage	-20°C for 1-6 months	-20°C for various periods up to 3 years
Sub-sample combination	8 sub-samples combined into single sample	Not applicable
Precondition aminopropyl SPEC	Precondition aminopropyl SPEC (Varian Mega Bond Elut, 10 g, 60 mL) with 50 ml methanol at < 20 ml/min (Rubio et al., 2004).	Precondition aminopropyl SPEC (Varian Bond Elut 1 g, 6 mL) with 5 ml methanol
Methanol extract clean-up using aminopropyl SPEC	Pass 8 mL sample methanol (+ 2 mL sample tube rinse) through aminopropyl SPEC (under gravity) at <3 mL/min (adjusted via tap). Collect in glass bottle.	Wash through the 1 ml methanol sample and keep the methanol filtrate in small glass bottles
Rinse aminopropyl SPEC	Pass 40 ml methanol rinse through the aminopropyl cartridges at <3 mL/min and collect.	Add 5 ml methanol to the aminopropyl SPEC and add the methanol eluate to the 1 ml filtrate
Storage	Store in at -20°C for 2 days	Stored at -20°C for various periods
Sample methanol volume reduction	Reduce methanol volume to ca 6 mL under partial vacuum at 43°C using rotary evaporator. Transfer methanol to graduated stoppered tube and top up to 8 mL with rinsings/ fresh methanol.	
Sample aliquoting	Dispense 1 mL methanol aliquots into 8 labelled screw-cap glass sample vials using Eppendorf multipipettor. Parafilm vials and store at -20°C.	
Storage & distribution	Storage at -20°C and transport to participants in cool boxes.	
Preparation for assay	Dependent upon participant. Cefas evaporated the methanol to dryness (at 45°C under nitrogen) and reconstituted in 10 µL methanol + 1000 µL assay buffer.	Methanol dried under a stream of nitrogen at 37°C and 1 mL of EIA buffer added
Inclusion of spiked and blank samples	Included as checks for method validity	Not included

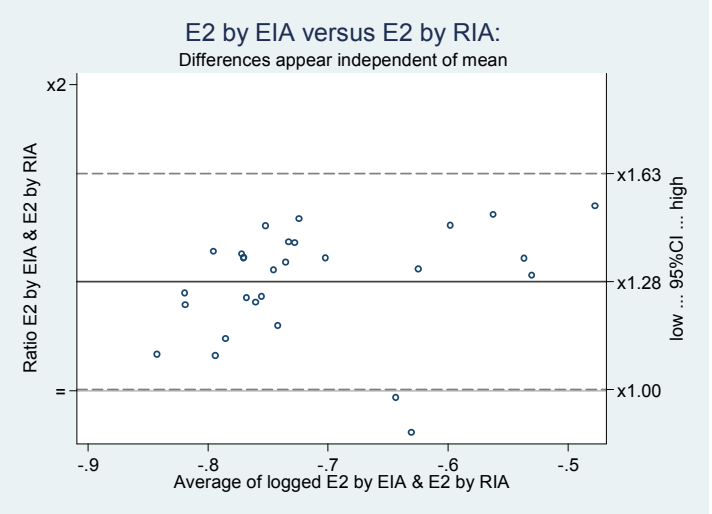
Appendix 3: Comparison of EIA and RIA measurements of river water sample aliquots

A recommended method for comparing two alternative measurement methods is that of Bland-Altman[1]:

- The median of the three replicates for each method was selected and log transformed
- The difference (EIA value – RIA value) was plotted against the average of the two medians.
- As the measurements are logged, this difference equates to the ratio of the two measurements.
- The plots are presented below showing the individual data points, the mean (solid regression line), and "limits of agreement" (dashed lines i.e. ± 2SD from the mean).
- If there were no difference between measurements, all point would lie on the line of equality (=) where the ratio is 1.







These plots illustrate that:
For 11KT and T:

- RIA values were higher than EIA values (Ratio EIA:RIA < 1)
- The difference between the methods decreased with increasing steroid concentration

For E2:

- EIA values were higher than RIA values (Ratio EIA:RIA > 1)
- The difference between the two methods did not change with increasing steroid concentration

The Bland-Altman method therefore confirmed the differences highlighted by ANOVA and indicates further subtle differences between the EIA and RIA measurements.

However, as the differences within the measurement range were judged not to be biologically relevant, we are justified in using the approach of taking a median of the three EIA and three RIA values to derive a water steroid concentration for further examination.

[1] Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet, 1986; i: 307-310 **(http://www-users.york.ac.uk/~mb55/meas/ba.htm)**

Appendix 4: Flow data

Mean Gauged Daily Flow data downloaded from the Centre for Ecology and Hydrology's National River Flow Archive (NRFA http://www.ceh.ac.uk/data/nrfa/data/retrievals.html) for the SF0241 sampling dates (2002 -2006). Data for the CB0427 sampling dates (2010) were obtained from Richard J Williams (CEH). Data was obtained for the nearest gauging stations (No 43005 on the R Avon at Amesbury and No 42004 on the R Test at Broadlands). The Gauged Daily Flow is the mean river flow in cubic metres per second in a water-day, (09.00 to 09.00 GMT). Measuring authorities typically calculate river flows on the basis of measurements at 15-minute intervals. This high resolution data is used to calculate the mean Gauged Daily Flow. These mean daily flow data are submitted to the NRFA. Data submitted to the NRFA are systematically submitted to a sequence of quality control checks. However, the ability of individual gauging stations to provide accurate and representative flows on a sensibly continuous basis varies greatly from station to station and, often, through time at the same site.

River	Date	Gauged Daily Flow m^3/s		River	Date	Gauged Daily Flow m^3/s
R Avon	10/12/2002	8.6		R Test	09/12/2002	18.6
R Avon	08/04/2003	4.4		R Test	08/04/2003	15.0
R Avon	13/05/2003	3.0		R Test	06/05/2003	12.0
R Avon	06/12/2003	2.2		R Test	05/12/2003	11.2
R Avon	10/12/2003	1.9		R Test	09/12/2003	10.3
R Avon	17/12/2003	2.3		R Test	16/12/2003	10.7
R Avon	10/01/2004	4.2		R Test	09/01/2004	17.4
R Avon	14/01/2004	5.8		R Test	13/01/2004	20.9
R Avon	10/03/2004	3.9		R Test	09/03/2004	11.7
R Avon	04/04/2004	3.7		R Test	03/04/2004	10.9
R Avon	14/12/2004	1.7		R Test	13/12/2004	8.7
R Avon	18/12/2004	2.6		R Test	17/12/2004	9.2
R Avon	03/01/2005	2.6		R Test	03/01/2005	9.8
R Avon	07/01/2005	2.7		R Test	07/01/2005	9.8
R Avon	10/05/2005	2.6		R Test	09/05/2005	7.2
R Avon	13/05/2005	2.5		R Test	12/05/2005	7.1
R Avon	14/03/2006	3.4		R Test	14/03/2006	7.9
R Avon	21/04/2006	2.8		R Test	21/04/2006	7.6
R Avon	28/04/2006	2.6		R Test	28/04/2006	7.3
R Avon	24/05/2006	3.0		R Test	24/05/2006	7.9
R Avon	01/06/2006	2.4		R Test	01/06/2006	6.7
R Avon	28/01/2010	6.9		R Test	21/01/2010	16.1
R Avon	04/02/2010	7.3		R Test	11/02/2010	14.2
R Avon	21/04/2010	4.5		R Test	11/05/2010	12.4
R Avon	27/05/2010	2.8		R Test	25/05/2010	10.3
R Avon	08/06/2010	2.8		R Test	15/06/2010	8.7

References to published material

9. This section should be used to record links (hypertext links where possible) or references to other published material generated by, or relating to this project.

Balaam JL, Grover D, Johnson AC, Jürgens M, Readman J, Smith AJ, White S, Williams R, Zhou JL (2010). The use of modelling data to predict levels of estrogens in a river catchment: How does modelled data compare with chemical analysis and in vitro yeast assay results? Science of the Total Environment 408, 4826-4832.

Barel-Cohen K, Shore LS, Shemesh M, Wenzel A, Mueller J, Kronfeld-Schor N (2006). Monitoring of natural and synthetic hormones in a polluted river. Journal of Environmental Management 78, 16-23

EA "What's in your backyard; accessed 22/02/11: http://maps.environment-agency.gov.uk/wiyby/wiybyController?ep=maptopics&lang=_e

Hanselman TA, Graetz DA, Wilkie AC (2004). Comparison of three enzyme immunoassays for measuring 17β-Estradiol in flushed dairy manure wastewater. J Environ Qual 33: 1919-1923.

Johnson AC (2010). Natural variations in flow are critical in determining concentrations of point source contaminants in rivers: an estrogen example. Environ. Sci. Technol. 44, 7865–7870.

Kolodziej EP, Harter T, Sedlak DL (2004). Dairy wastewater, aquaculture, and spawning fish as sources of steroids hormones in the aquatic environment. Environ Sci Technol 38, 6377-6384.

Pelissero C, Cuisset B, Menn F Le (1989). The influence of sex steroids in commercial fish meals and fish diets on plasma concentration of estrogens and vitellogenin in cultured Siberian sturgeon *Acipenser baeri*. Aquatic Living Resources 2, 161-168.

Petrie A, Sabin, C (2000). Medical statistics at a glance. Blackwell, Oxford, 138 p.

Rubio F, Ramsey K, Stackelberg P, Meyer M (2004) Monitoring of selected herbicides, antibiotics, steroids, and industrial chemicals in water by ELISA. Presentation at 2004 National Monitoring Conference Presentations and Results, Chattanooga, TN - May 17 – 20, http://acwi.gov/monitoring/conference/2004/conference_agenda_links/power_points_etc/03_ConcurrentSessionA/24_Rm13_RubioFernando.pdf

Scott A P, Bye VJ, Baynes SM (1980). Seasonal variations in sex steroids of female rainbow trout (*Salmo gairdneri* R.). J Fish Biol 17, 587-592.

Shore LS, Reichmann O, Shemesh M, Wenzel A, Litaor MI (2004). Washout of testosterone in a watershed. Science of the Total Environment 332, 193-202.

Sink TD, Lochmann RT, Fecteau KA (2008). Validation, use and disadvantages of enzyme-linked immunosorbent assay kits for detection of cortisol in channel catfish, largemouth bass, red pacu, and golden shiners. Fish Physiol Biochem 34: 95-101

Sumpter JP, Scott AP, Baynes SM, Witthames PR (1984). Early stages of the reproductive cycle in virgin female rainbow trout (*Salmo gairdneri* R.). Aquaculture 43, 235-242.

Vermeirssen ELM, Scott AP (1996). Excretion of free and conjugated steroids in rainbow trout (*Oncorhynchus mykiss*): evidence for branchial excretion of the maturation-inducing steroid, 17,20b-dihydroxy-4-pregnen-3-one. General and Comparative Endocrinology 101, 180–194.

Waring CP, Moore A, Best JH, Crooks N, Crooks LE (In press). Do trout farm effluents affect Atlantic salmon smolts? Preliminary studies using caged salmon smolts. Aquaculture.

Zhao S, Zhang P, Melcer ME, Molina JF (2010). Estrogens in streams associated with a concentrated animal feeding operation in upstate New York, USA. Chemosphere 79, 420-425.