SUPPORTING INFORMATION

Manuscript Title: Parameterization of pharmaceutical emissions and removal rates for use in UK predicted exposure models: steroid estrogens as a case study

Authors: J. D. Heffley[1,3], S. D. W. Comber[2], B. W. Wheeler[1], C. H. Redshaw[1]

[1]European Centre for Environment and Human Health, University of Exeter Medical School, Truro, Cornwall, UK.

[2]School of Geography, Earth and Environmental Sciences, University of Plymouth, Plymouth, Devon, UK.

[3]University of Miami, University of Miami Miller School of Medicine, Department of Public Health Sciences, Miami, Florida, United States.

Contents:

Predicting Synthetic Estrogen Usage: A Case Study

Understanding the dataset: British National Formulary (BNF) codes

The BNF is split into chapters (i.e. endocrine system), and further divided into multiple sections (i.e. sex hormones, female sex hormones, oestrogens and hormone replacement therapy), which are then further classified as codes for specific pharmaceuticals and combinations of pharmaceuticals (i.e. Combined Ethinylestradiol 30mcg"). These codes are often comprised of a number of different formulations, which vary by brand, dose and/or route of administration; no specific coding or labelling is given for individual formulations. For example, in the UK there are 16 different formulations within the BNF code for "Combined Ethinylestradiol 30mcg" (BNF code 070301F0). The mass of active ingredient in each formulation may differ over a range of brands, routes of administration, or combinations with other drugs (see S1 for case study example).

	Average prescriptions per month (R_t)			
Pharmaceutical name & BNF Code	GP Surgery 1	GP Surgery 2	GP Surgery 3	Total for selected catchment
Estradiol (systemic) 0604011G0	48.6	22.8	25.4	96.8
Estradiol Valerate 0604011K0	2.6	7.2	3.4	13.2
Estradiol with Progestogen 0604011L0	28.2	15	21	64.2
Estradiol (topical) 0702010Ga0	35.2	14	28.2	77.4
Estradiol Val & Estradiol Val + Dienogest 0703010R0	0	0.2	0	0.2
Oestrogens Conjugated 0604011P0	4.6	6.8	6.4	17.8
Oestrogens Conjugated with Progestogen 0604011Q0	0.8	5	5.2	11
Combined Ethinylestradiol 20mcg 703000	1.6	2.2	11.4	15.2
Ethinylestradiol 0604011D0	2.2	0	1.4	3.6
Combined Ethinylestradiol 30mcg 070301F0	59.8	39.8	70.8	170.4
Combined Ethinylestradiol 35mcg 070301G0	19.6	7.6	7.4	34.6
Phased Formulations Of Ethinylestradiol 070301P0	0.8	1.8	5	7.6
Co-Cyprindiol (Cyprote Acet/Ethinlestr) 1306020C0	2	1	3.2	6.2
Etonogestrel/Ethinylestradiol 070301A0	0	0	0.2	0.2

Table S1: Estrogen Prescription Data for the Selected River Catchment

Mean number of prescriptions calculated across 01.09.11 to 31.01.12. Source: National Health Service, Prescribing by GP Practice, The Information Centre for Health and Social Care, 2011 [1].

Conversion of number of prescriptions to prescribed mass

In the UK there are 16 different formulations within the BNF code for "Combined Ethinylestradiol 30mcg" (BNF code 070301F0; table S1). The mass of active ingredient in each formulation may differ over a range of brands, routes of administration, or combinations with other drugs, but in this particular case all formulations contained 30 micrograms of EE2. NHS data [2] detailed for the UK as a whole the number of prescriptions written for each individual formulation (I_f) and the quantity of drug dispensed (Q_f). The units for Q_f depend upon how each individual formulation is dispensed, and can be a "pill", "pack", "millilitre", "gram", "capsule", etc. To illustrate this point, data for "Loestrin 30_Tab" (which is a formulation within BNF code 070301F0) is discussed (see table S2). There were 24,528 prescriptions written and 2,380,457 pills of Loestrin 30_Tab dispensed in the UK between October and December of 2011. Therefore each prescription of Loestrin 30_Tab contained on average 97 pills, which equates to 2.91 mg of active ingredient (Table S2). The average mass of active ingredient per prescription for each formulation of EE2 ($M_{i,f}$), across the whole of the UK was calculated in this manner (Eq S1; Table S2); however this did not account for regional (or catchment level) prescribed mass.

As localised prescription data is not detailed down to the individual formulation level (only BNF codes), it was necessary to ascertain the relative proportion of each formulation prescribed within each BNF code to obtain typical UK wide prescribing practices, which could then be applied to localised catchment level data. For example, 3% of the 934,016 "Combined Ethinylestradiol 30mcg" (BNF code 070301F0) prescriptions written in the UK were of the "Loestrin 30_Tab" formulation (Eq S2; Table S2). Assuming that the prescription habits of GPs are uniform throughout the UK, 3% of the 170.4 monthly prescriptions for "Combined Ethinylestradiol 30mcg" in the selected river catchment were the "Loestrin 30_Tab" formulation (Eq S3). This output was then used in conjunction with average mass per prescription (Eq S1) of a specific drug formula to calculate the total mass of Loestrin 30_Tab prescribed in this catchment per month and per day (Eq S4). A

simplified summary formula covering all stages of calculation is presented in Eq S5 summary. This process was repeated for all formulations contained within each BNF code of interest and were then summed to give the total mass prescribed for each chemical of interest (E2, EE2, CE) (Eq S5).

Formulation Name	**EE2 dosage (mg)** (D)	**Items (1000's)** (I_f)	**Quantity (1000's)** (Q)	**Mean mass per prescription (mg)** (M_I)	**Proportion prescribed** (P)	**Number of prescriptions / month for catchment** (R_f)	**Mass prescribed per day for catchment (mg)** (M_D)
Elevin_Tab 150mcg/30mcg	0.03	2.33	222.48	2.87	0.00	0.42	0.04
Femodene ED_Tab	0.023	1.20	135.42	2.54	0.00	0.22	0.02
Femodene_Tab	0.03	36.5	3651.70	3.00	0.04	6.66	0.66
Gedarel_Tab 30/150mcg	0.03	18.7	1800.22	2.88	0.02	3.42	0.32
Katya 30/75_Tab	0.03	0.20	17.87	2.65	0.00	0.04	0.00
Levest 150/30_Tab	0.03	12.10	1194.52	2.96	0.01	2.21	0.21
Levest 150/30_Tab (Actavis)	0.03	1.18	114.67	2.92	0.00	0.22	0.02
Loestrin 30_Tab	0.03	24.53	2380.46	2.91	0.03	4.47	0.43
Marvelon_Tab	0.03	47.99	4718.38	2.95	0.05	8.76	0.85
Microgynon 30 ED_Tab	0.0225	26.94	3171.78	2.65	0.03	4.91	0.43
Microgynon 30_Tab	0.03	492.78	48562.78	2.96	0.53	89.90	8.74
Millinette_Tab 30/75mcg	0.03	2.41	231.32	2.89	0.00	0.44	0.04
Minulet_Tab	0.03	0.00	0.25	3.78	0.00	0.00	0.00
Ovranette_Tab 150mcg/30mcg	0.03	46.79	4488.25	2.88	0.05	8.54	0.81
Rigevidon_Tab	0.03	39.97	3953.00	2.97	0.04	7.29	0.71
Yasmin_Tab	0.03	180.39	17653.28	2.94	0.19	32.91	3.18
Totals		**934.02 (I_T)**				**170.40 (R_T)**	

Table S2: Formulation level data for BNF code 070301F0

Formulae for a generalised method to predict prescription usage

A general method for determining daily use of prescription medications in a targeted catchment is presented below.

Eq S1. The average mass per prescription for a specific drug formulation, $M_{i,f}$, is given by

$$M_{i,f} = \frac{D_f Q_f}{I_f}$$

where D_f is the mass of a single dose of a given formulation, Q_f is the quantity of doses dispensed of this formation in the UK within a given time frame (e.g. number of pills dispensed from October-December 2011), and I_f is the number of prescriptions written of a given formulation UK wide (within the same time frame as Q_f).

Eq S2. The relative proportion prescribed of each formulation on a UK wide basis within a given BNF code, P_f, is given by

$$P_f = \frac{I_f}{I_t}$$

Where I_t is the sum of all I_f within a given BNF code.

Eq S3. The number of prescriptions written of a given formulation in a localised catchment per month, R_f, can then be estimated by

$$R_f = P_f R_t$$

where R_t is the average number of prescriptions written per month in the catchment for a given BNF code (from Table S1).

Eq S4. The mass prescribed for each formulation in a given BNF code per month in the targeted catchment, M_m, can now be calculated by

$$M_m = M_{i,f} R_f$$

which can be converted to the daily mass prescribed of each formulation in a given BNF code, M_d, which is given by

$$M_d = \frac{12 M_m}{365}$$

Eq S5. The series of equations above can be simplified to the following formula:

$$M_d = \frac{12 D_f Q_f R_t}{365 I_t}$$

Eq S6. To ascertain the total mass of a drug prescribed in a given catchment this series of calculations was repeated for each formulation within each BNF code of a given chemical, e.g. all 47 formulations of EE2 which are contained within all 7 BNF codes that include EE2. This method of determining the mass of pharmaceuticals used is contingent on the assumption that all prescribed medications are actually ingested by the patient. Typically, only 50% of prescriptions are taken [3], but data suggests that patient adherence to contraceptive and HRT regimens is much higher [4, 5]. Therefore, assuming all prescribed medications are consumed, the sum of all M_d across all BNF codes containing a given chemical is the total daily consumption for the given chemical in the target catchment, M_t, and can be given by

$$M_t = \sum_{f=1}^{n} M_{d,f}$$

where n is equal to the number of formulations of a given chemical and the subscript f denotes the fth formulation of a given chemical.

Modifying case study model inputs to generate predictions for England

As there was no measured crude sewage steroid data for the select catchment, the modelling approach developed was modified to apply English population values to the results of the case study excretion predictions, in order to allow comparison with measured sewerage influent data for England as a whole. For the values based on prescription data (HRT users and EE2 excretion), this modification is predicated on the assumption that prescribing practices for physicians in the case study catchment mimic those of physicians in England as a whole. Outputs from this modification and a description of the process are presented in Table 3.

			Total excretion (natural + synthetic) in mg day^{-1}					
			CASE STUDY CATCHMENT			**ENGLAND**		
Population Group	% of catchment Population	% of England's Population	E1	E2	EE2	E1	E2	EE2
A. Pregnant	0.78	0.99	95.15	67.99	0	267020	190798	0
B. Menstrual Females (not pregnant)	20.60	24.37	53.40	14.60	9.37	140121	38324	24587
C. HRT users	3.16	2.38	11.34	14.45	0	18881	24059	0
D. Menopausal Females (non-medicated)	20.80	15.64	8.31	4.62	0	13835	7686	0
E. Males	47.70	48.68	27.55	19.08	0	62194	43057	0
Average for total population of area (µg / day / capita)			8.82	5.44	0.42	10.22	6.19	0.50

Table S3: Comparison between demographics and modelled estrogen excretion rates for the select catchment and England as a whole

Population percentages are calculated from 2001 census data [6] by completing the following calculations, followed by conversion to a percentage of total population: A = population x birth rate x multiplier accounting for pregnancy duration; B = female population aged 13-49 minus the number of pregnant females; C = female population aged 50+ x HRT usage rate (13.2%); D = female population aged 50+ - number of HRT users; E = obtained from census data. Average per capita excretion data is presented in reference to total population (i.e. adults and children). Total natural excretion was calculated based upon figures detailed in the manuscript (Figure 1) and census data. Total synthetic excretion for England was based upon average catchment level per capita excretion of synthetic compounds (mass excreted in catchment ÷ number of individuals in corresponding catchment population group) and English population group specific demographics. Total natural and synthetic excretion were summed and divided by the total population size to find average for total population of area.

Predicting in-sewer transformations: a case study

The UK water industry has recently undertaken a £30 million research programme to determine priority chemicals entering STW, impacts of primary, secondary and tertiary treatment and effluent concentrations. A total of 25 STW were sampled over the course of a year and influent concentrations and removal rates calculated for chemicals including the steroid estrogens [7]. Table S4 provide a summary of the influent data reported in an earlier predictive study, the UK Chemical Investigation Programme and this modelling approach (using a range of E1 to E2 transformation rates). For the purpose of an accurate comparison, concentrations have been calculated as µg capita^{-1} day^{-1}.

	E1	**E2**	**EE2**
Predicted influent (µg capita^{-1} day^{-1}) – this work, select catchment; 50% E2 to E1 conversion	11.5	2.72	0.42
Predicted influent (µg capita^{-1} day^{-1}) – this work, select catchment; 28.75% E2 to E1 conversion	10.4	3.88	0.42
Predicted influent (µg capita^{-1} day^{-1}) – this work, select catchment; 6% E2 to E1 conversion	9.1	5.11	0.42
Predicted influent (µg capita^{-1} day^{-1}) – this work, England as a whole; 50% E2 to E1 conversion	13.3	3.09	0.5
Predicted influent (µg capita^{-1} day^{-1}) – this work, England as a whole; 28.75% E2 to E1 conversion	12.0	4.41	0.5
Predicted influent (µg capita^{-1} day^{-1}) – this work, England as a whole; 6% E2 to E1 conversion	10.6	5.81	0.5
Predicted influent (µg capita^{-1} day^{-1}) – [8]	13.8	3.30	0.89
Mean measure influent (µg capita^{-1} day^{-1}) - [7]	16.1	5.90	0.31
Median influent (µg capita^{-1} day^{-1}) - [7]	16.7	5.90	0.23
Standard Deviation (µg capita^{-1} day^{-1}) - [7]	3.9	1.4	0.21
Range (µg capita^{-1} day^{-1}) - [7]	7.6-25.9	3.1-9.6	0.12-1.12

Table S4: Comparative findings for the case study catchment: influent loads

Description of Ratio-Approach used to ascertain In-Sewer Transformation rates

Based upon Johnson and Williams assumption of 50% degradation of E2 to E1, the modelling approach developed herein and modified with English demographics data, predicts E1 and E2 influent loads of 13.3 and 3.09 μg capita^{-1} day^{-1} respectively, which is in line with Johnson and Williams predictions (E1 = 13.8, E2 = 3.30 μg capita^{-1} day^{-1}), but is lower than measured means (E1 = 16.1, E2 = 5.9 μg capita^{-1} day^{-1}). Reduction of the degradation rate within the sewerage system will decrease E1 influent predictions, but increase E2 predictions. However, as predicted E1 + E2 equates to ~16 μg capita^{-1} day^{-1} in total, modification of the sewer removal rate will not result in achieving the higher values seen in measured data for both compounds simultaneously. Therefore sewer removal rates were optimised to achieve the same ratio of E1:E2 observed in the measured data. An identical E1:E2 ratio of 2.72 is achieved with a 28.75% removal rate; resulting in E1 and E2 predictions of 12.0 and 4.41 μg capita^{-1} day^{-1} respectively, with an 'optimised value' of 75% of the observed influent values, for the England prediction. Lowering the removal rate further, to the 6% indicated in the findings of M.E. Jarvie and D.W. Hand [9], changes the E1:E2 ratio to 1.82, resulting in substantially higher proportions of E2 relative to E1 than is found in observed data, with 98% and 66% of observed values for E2 and E1 being predicted, respectively. Based on this empirical data it may therefore be concluded that using a value of 28.75% loss of E2, with subsequent conversion to E1, within the sewer system, is the most accurate assumption for the purpose of this risk assessment for achieving realistic predictions of E1 and E2 simultaneously. Application of this 28.75% removal rate to the catchment scale model, results in E1 and E2 predictions of 10.4 and 3.88 μg capita^{-1} day^{-1} respectively.

Predicting losses during sewage treatment: a case study

The data (Figure S1) show that measured E1 and E2 removal rates are similar to those used in previous risk assessments and are within the 95% confidence intervals. For EE2, however, a very different result is observed, with previous estimates using a removal rate of 85%. The recent UK data however, generated widely varying removal rates, of between 0 and 92% with a mean of only 27% and median of 52% [7], greater removal rates were achieved with increased biological treatment, including nitrifying processes for ammonia removal (Figure S1). One explanation for this is analytical errors and measurements below the limit of detection which lead to ranges much more extreme than those calculated for E1 and E2, which are present at an order of magnitude or higher.

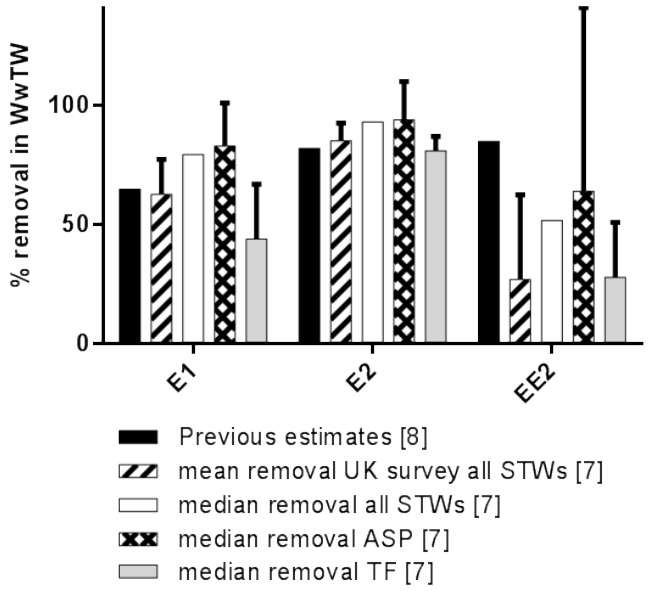

Figure S1: Reported STW removal rates for estrogens previously measured or estimated

Error bars represent 95% confidence intervals for overall mean removal and variations in median removal rates between activated sludge processes (ASP) and biological filter works (trickling filter: TF)

		E1	E2	EE2
Predicted influent (ng L^{-1})		48.0	17.6	2.0
Measured Effluent (ng L^{-1}) [& % observed removal]	Mean all STW	17.8 [63]	2.6 [85]	0.56 [27]
	Median all STW	10.9 [80]	1.3 [93]	0.36 [28]
	Range for all STW (95% conf)	10.9 – 24.7	1.3 – 3.9	0.33 – 0.78
	Median TF only	27.1 [44]	2.7 [81]	0.58 [53]
	Median ASP only	10.3 [83]	0.88 [94]	0.36 [64]
Required predicted: measured ratio		0.75	0.75	1.61
Predicted Effluent (ng L^{-1}; using optimised % removal)	Mean all STW	13.3	1.9	0.9
	Median all STW	8.1	1.0	0.6
	Median TF only	20.2	2.0	0.9
	Median ASP only	7.7	0.7	0.6
Optimised % removal	Mean all STW	72.3	89.0	54.9
	Median all STW	83.1	94.5	71.0
	Median TF only	57.9	88.6	53.3
	Median ASP only	84.0	96.3	71.0
	Range	58 – 84	89 – 96	53 - 71

Table S5: Optimised STW removal rates (based upon a ratio approach) for calculation of predicted effluents

Predicted data presented above are for the case study presented herein, for England's population demographics. Predicted influent concentrations are calculated on a basis of 28.8% sewer transformation of E2 to E1. Required predicted:measured ratio is calculated from influent data (Table S4).

Optimised removal rates for E1 and E2 correlate well with observed rates (i.e. mean & median across all STW; median TF and median ASP; Table S5). Of more interest are the optimised removal rates

generated for EE2. For specific STW types EE2 optimised removal rates (ASP 71%; TF 53%) and measured removal rates (ASP 64%; TF 53%) are alike. However, when the specific STW processes are not considered, and average or median values across different types of STW are used, optimised removal rates are found to be double those observed. These findings therefore suggest that use of predictive models would achieve more accurate effluent predictions if the specific treatment processes at a given works is taken into consideration. However, in the absence of process data for the works within the case study catchment, the optimised removal rates (calculated upon England's demographic data and presented above) were applied, resulting in effluent concentrations as summarised (Table S5).

Reference for SI

1. National Health Service, *Prescribing by GP practice*. 2011, The Health and Social Care Information Centre.
2. National Health Service, *Quarterly Prescription Cost Analysis*. 2011, The Information Centre for Health and Social Care.
3. R.B. Haynes, E. Ackloo, N. Sahota, H.P. McDonald, and X. Yao, *Cochrane Db. Syst. Rev.*, 2008. **2**(2).
4. M.Y. Hou, S. Hurwitz, E. Kavanagh, J. Fortin, and A.B. Goldberg, *Obstet. Gynecol.*, 2010. **116**(3), 633.
5. B.-S.S.M. Torgerson Dj, *JAMA, J. Am. Med. Assoc.*, 2001. **285**(22), 2891.
6. UK Office for National Statistics, UK Data Service Census Support: http://casweb.mimas.ac.uk.
7. M. Gardner, V. Jones, S. Comber, M. Scrimshaw, T. Coello-Garcia, E. Cartmell, J. Lester, and B. Ellor, *Sci. Total Environ.*, 2013. **456-457**, 359.
8. A.C. Johnson and R.J. Williams, *Environ. Sci. Technol.*, 2004. **38**(13), 3649.
9. M.E. Jarvie and D.W. Hand, *Water Environ. Res.*, 2009. **81**(2), 131.