Impact of in-sewer degradation of pharmaceutical and personal care products (PPCPs) population markers on a population model

Jake William O'Brien, Andrew Phillip William Banks, Andrew Joseph Novic, Jochen F. Mueller, Guangming Jiang, Christoph Ort, Geoff Eaglesham, Zhiguo Yuan, and Phong Thai

Environ. Sci. Technol., Just Accepted Manuscript • DOI: 10.1021/acs.est.6b02755 • Publication Date (Web): 28 Feb 2017

Downloaded from http://pubs.acs.org on March 7, 2017

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.
Title: Impact of in-sewer degradation of pharmaceutical and personal care products (PPCPs) population markers on a population model


Queensland Alliance for Environmental Health Science (QAEHS), The University of Queensland, 39 Kessels Road, Coopers Plains, QLD 4108, Australia
Advanced Water Management Centre (AWMC), The University of Queensland, St. Lucia, QLD 4072, Australia
Eawag, Swiss Federal Institute of Aquatic Science and Technology, CH 8600 Dübendorf, Switzerland
International Laboratory for Air Quality and Health, Queensland University of Technology, Brisbane, QLD 4001, Australia

*Tel. +61 (0)7 3274 9009, Fax +61 (0)7 3274 9003
†Queensland Alliance for Environmental Health Sciences incorporates the former National Research Centre for Environmental Toxicology (Entox).

Abstract
A key uncertainty of wastewater-based epidemiology is the size of the population which contributed to a given wastewater sample. We previously developed and validated a Bayesian inference model to estimate population size based on 14 population markers which: (1) are easily measured and (2) have mass loads which correlate with population size. However, the potential uncertainty of the model prediction due to in-sewer degradation of these markers was not evaluated. In this study we addressed this gap by testing their stability under sewer conditions and assessed whether degradation impacts the model estimates. Five markers which formed the core of our model, were stable in the sewers while the others were not. Our evaluation showed that the presence of unstable...
population markers in the model did not decrease the precision of the population estimates providing that stable markers such as acesulfame remained in the model. However, to achieve the minimum uncertainty in population estimates, we propose that the core markers to be included in population models for other sites should meet two additional criteria: (3) negligible degradation in wastewater to ensure the stability of chemicals during collection; and (4) <10% in-sewer degradation could occur during the mean residence time of the sewer network.

1. Introduction:

Wastewater-based epidemiology (WBE) has become a useful approach for monitoring of drug use at the population level. Challenges remain to improve the accuracy of the approach and a key uncertainty associated with WBE relates to the population that has contributed to a given wastewater sample which is essential for the normalisation of the estimates. Normalised data is important to ensure that WBE is comparable across cities and even countries. A number of proxies have previously been proposed to address the population uncertainty. These include nutrients such as N and P, biological oxygen demand (BOD) and chemical oxygen demand (COD), creatinine, caffeine and nicotine, as well as selected pharmaceuticals and personal care products (PPCPs) and an artificial sweetener. But none of the proposed population markers were calibrated against accurate population counts and thus were not validated yet. In our previous paper, a Bayesian inference model was developed to estimate de facto population size using 14 anthropogenic markers (Table 1) using samples from 10 wastewater treatment plants (WWTPs) collected during the 2011 Australian Census. We subsequently applied this model successfully to support a study on daily drug use monitoring in a catchment in South East Queensland, Australia. However, we recognised that some chemicals may degrade in the sewer. Therefore, it is important to investigate whether degradation during sewer passage contributes to the uncertainty of the population estimation model.
In-sewer degradation is a result of both hydrochemical and biotransformation processes in the sewer and as such is dependent on the chemistry (e.g. metals, micropollutants, nutrients, pH) of the wastewater, the bioactivity in the sewer and residence time of the wastewater in the sewer\textsuperscript{15, 16}. In real sewers the chemistry, bioactivity and residence times are dependent on the flow and composition of water entering from the catchment at the time\textsuperscript{16}. Additionally, sewers are complex systems comprising of a multitude of pipes ranging from small, pressurised pipes called rising mains (RM) to large partially filled gravity fed pipes called gravity sewers (GS). As such the surface area, which biofilm can grow on is, dependent on both the size of the pipe and how full the pipe remains. The ratio of area of biofilm versus volume of wastewater is thus termed the A/V ratio. Sewer networks comprise of sections ranging from low A/V to high A/V with differing residence times in each. A major limitation of in-sewer chemical degradation monitoring is that the observed degradation rates need to be statistically higher than the uncertainty of the chemical analysis and thus high A/V and residence times may be required. Using laboratory-scale sewer reactors overcomes this limitation as both the A/V and residence time can be controlled\textsuperscript{14, 17}.

The aim of this study is to 1) assess the stability of the 14 population markers proposed in O’Brien et al.\textsuperscript{13} under sewer conditions and 2) assess whether potential degradation of these population markers in the sewer impacts the population model. From this we propose a protocol to assess whether a chemical can be used as a population marker.

2. Materials and Methods:

2.1 Chemicals and reagents

Analytical grade acetic acid was purchased from Sigma Aldrich (Castle Hill, Australia). Analytical grade hydrochloric acid 32\% was purchased from Univar (Ingleburn, Australia). Water was purified through a MilliQ system (Millipore, 0.22 μm filtered, 18.2 mΩ cm\textsuperscript{-1}). High purity PPCP native and
labelled analytical standards were purchased from various suppliers as outlined in the SI. Calibration standards were prepared in MilliQ water. Liquid chromatography grade methanol was purchased from Merck (Darmstadt, Germany). Mobile phases were filtered using Sartorius Stedim 0.45 µm RC filters (Goettingen, Germany).

2.2 Laboratory-scale sewer reactors

Laboratory-scale sewer reactors, with high A/V (70.9 m²/m³) which is similar to A/V ratios of small pipes, were used to investigate the biotransformation and degradation of the population markers. These reactors have been described elsewhere. Briefly, the system comprised of laboratory-scale RM, GS and a control reactor (CR). Earlier studies using these reactors have found them suitable for studies regarding in-sewer processes, the control of sewer biofilm activities and the in-sewer degradation of chemicals. The sewer reactors are fed with domestic wastewater collected on a weekly basis from a local pumping station (Indooroopilly, Brisbane) which is typically at pH 7.5 with sulphide concentrations <3 mg S/L, sulfate between 10 and 25 mg S/L, total COD between 450 and 600 mg/L, soluble COD between 260 and 450 mg/L which contains volatile fatty acids between 50 to 120 mg COD/L. The collected wastewater is stored at 4 °C and warmed to ~20 °C prior to feeding into the reactors. Feeding events occur at 6 hour intervals to mimic the typical hydraulic retention time of a real sewer and last approximately 2 minutes during which time one reactor volume (0.75 L) of wastewater is delivered to the reactor. Magnetic stirrers set to 250 rpm (Heidolph MR3000) homogenise the wastewater in the reactors. Further description of the reactors can be found in Thai et al.

2.3 Batch tests for the degradation of PPCPs in sewer reactors

To investigate the degradation of the PPCPs under different sewer conditions, batch tests were conducted in triplicate. To meet OECD guideline No. 314, wastewater from the sampling pump station was warmed to 20 °C however no pH adjustment was necessary. These conditions are comparable with other studies on the stability of chemicals under sewer conditions.
Preliminary analysis of the wastewater was conducted to determine the spiking amount. Only furosemide, iopromide and norfloxacin required spiking for degradation to be measured accurately to give a spiked concentration of 5-10 μg.L$^{-1}$ above concentrations already present in the wastewater but within the range usually found in previous wastewater studies. For this purpose, a working solution of PPCP standards from 5 - 10 μg.mL$^{-1}$ was prepared in 50:50 methanol/deionised water and spiked in to the wastewater. Other chemicals can be tested with the residues already available in the wastewater matrix. The wastewater was then pumped using a peristaltic pump (Masterflex 7520-47) at a flow rate of 0.375 L/min into the RM and GS reactors ensuring that the wastewater in each reactor was entirely replaced with fresh wastewater. Mixing occurred continuously for the duration of the batch tests using magnetic stirrers set to 250 rpm (Heidolph MR3000). This not only improved homogeneity of the wastewater but also enhanced surface aeration (dissolved oxygen of ~0.5 mg/L) producing both aerobic and anaerobic conditions in the GS reactor. The CR was basically the same as the GS except that the walls are regularly cleaned hence preventing the growth of biofilms, however, mixing still occurred which enhanced the aerobic conditions for the CR (DO = 0.5 mg/L). Wastewater samples were taken at 0, 0.25, 0.5, 1, 2, 3, 6, 9 and 12 hours after the feeding event. Samples were immediately filtered using 0.20 μm PTFE syringe filters (Phenomenex, Australia) into 2 mL vials and 15 μL of 2 M HCl was added to each sample to acidify the sample to ~pH 2 to reduce the biological activity$^{19}$ and then samples were frozen at -20 °C prior to analysis.

**2.4 Chemical Analysis**

A slightly modified version of the chemical analysis outlined in O’Brien et al$^{13}$ was performed using a Sciex QTrap 6500 (Sciex, Concord, Ontario, Canada) with electrospray ionization (ESI) interfaces coupled to Shimadzu Nexera HPLC systems (Shimadzu Corp., Kyoto, Japan). Separation was achieved on a Kinetix Biphenyl column (2.6 μm, 100 Å, LC Column 50 mm x 2.1 mm, Phenomenex) using a mobile phase gradient of 1 to 95 % methanol with 0.1% acetic acid. The QTrap 6500 was operated in scheduled multiple reaction monitoring (sMRM) mode with optimised parameters (SI Table 1).
acquisition was performed using the Sciex software package Analyst Software 1.6 and quantification was performed using MultiQuant 3.0. Quantification was done using relative response factors of mass labelled internal standards.

For dissolved sulphide, samples were analysed within 24 h of sampling using an ion chromatograph with an UV and conductivity detector (Dionex ICS-2000). For methane analysis, BD vacutainer tubes were allowed to reach gas/liquid equilibrium overnight. Methane in the gas phase was measured by gas chromatography (Shimadzu GC-9A) equipped with a flame ionization detector. Concentrations of methane in sewage were calculated using mass balance and Henry’s law.

3 Results and Discussion:

3.1 Biological activities in sewer reactors

The biological activity indicated by methane and sulphide gas production of the RM and GS reactors used in this experiment has previously been found to represent the biological activity and associated degradation of chemicals within actual sewers. The control reactor showed no significant methane or sulphide gas production as it doesn’t contain the sewer biofilms compartment as seen before. The RM reactor produced sulphide and methane indicating that it was under anaerobic conditions. Activities of sulphate-reducing bacteria and methanogenic archaea in the RM reactor were measured at 7.21 ± 0.74 mg S/L-h and 29.73 ± 0.63 mg COD/L-h respectively which is similar to previously reported values for both real and laboratory-scale sewers. Dissolved oxygen in the GS reactor was measured below 0.33 mg/L despite continuous aeration which indicates aerobic activity consuming oxygen. It is also expected that anaerobic conditions may be present in the bottom of the reactor where oxygen could not reach. This is supported by the low reduction of sulphate (4.21 ± 0.77 mg S/L-h) and the low production of methane (14.8 ± 2.31 mg COD/L-h).

3.2 Degradation of PPCPs in the sewer reactors
Figure 1 shows the concentration profile of chemicals of interest in the system over time normalised to the concentration at the start of the experiment expressed in percent (i.e. $C_t/C_0 \times 100$). To understand the degradation kinetics of the PPCP population markers, both linear regression (zero order) and pseudo first order regression were fitted to the data obtained from the batch tests (Table 1). For the pseudo first order, the regression intercept was set through 100% at time 0. The kinetic model chosen for the degradation of each compound was based on the fit to the model, i.e. the model with the higher $R^2$ value. For compounds where degradation observed was within the uncertainty of measurements, we concluded that no degradation had occurred over the studied period (12 hours) and thus these compounds are stable. The term “degradation” in this study refers to either biological or chemical transformation as sorption to organic matter of the studied chemicals was considered negligible as sorption is only considered significant for log octanol-water partitioning coefficients ($\log K_{OW}$) higher than 4.0 ($\log K_{OW}$) (see SI Table 2 for $\log K_{OW}$ values).

To our knowledge only Jelic et al. have investigated the fate of selected PPCPs under sewer conditions. That study however was conducted only in a pressurised sewer which is similar to that of the RM in our study. It is important to note though that the sewer reactors used for our experiment had fixed area to volume (A/V) ratios of 70.9 m$^2$/m$^3$ which is similar to A/V ratios of small pipes but higher than that of average sewers which typically use larger diameter pipes. Therefore the chemicals investigated in this study may have had more contact with biofilms and hence more degradation than would be expected in a typical sewer. This may explain why some discrepancy was observed between our results and that of Jelic et al. who had an A/V ratio of 8 m$^2$/m$^3$. As no other studies have looked at the degradation within a GS, we conducted a comparison between our results with available data on the removal efficiency of WWTPs for those PPCPs (Table 1). Such a comparison helps support our findings although we acknowledge that fate and transport processes during wastewater treatment are different from those during sewer passage ($S1$ SI).
It was observed that there are two groups of chemicals, those which had no measurable degradation (acesulfame, atenolol, carbamazepine, gabapentin and ibuprofen) and those which degraded over the studied period with half-life estimates ranging from 0.6 to 10 h in the GS and 0.6 to 11h in the RM. Detailed discussion for the degradation of each chemical can be found in the SI.
Figure 1 Degradation profiles of 14 population markers under different sewer conditions. Error bars represent the standard deviation of 3 replicate samples. The x axis is the time after the feeding event. The y axis is the concentration relative to the starting concentration (time zero) expressed as percent. Concentrations under control conditions are indicated by ×. Concentrations under gravity sewer (GS) conditions are indicated by ■. Concentrations under rising main (RM) conditions are indicated by ▲.
Table 1 Selection of kinetics models for degradation of the 14 population markers under sewer conditions. The model which fitted with the higher $R^2$ value was selected except where linear regressions indicated no significant degradation (i.e. less than the uncertainty of the measurement). The selected model for each compound under the particular sewer condition is indicated by bolded text. Time before 10% loss under each sewer condition was calculated using the selected model for each compound.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Correlation</th>
<th>Linear regression (Zero order)</th>
<th>First-order Kinetics</th>
<th>Time before 10% loss (h)</th>
<th>WWTP</th>
</tr>
</thead>
</table>

Page 10 of 21
<table>
<thead>
<tr>
<th>Substance</th>
<th>R²</th>
<th>Ranking*</th>
<th>Control</th>
<th>Gravity Sewer</th>
<th>Rising Main</th>
<th>Control</th>
<th>Gravity Sewer</th>
<th>Rising Main</th>
<th>Control</th>
<th>Gravity Sewer</th>
<th>Rising Main</th>
<th>Removal Efficiency %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acesulfame</td>
<td>0.995</td>
<td>1</td>
<td>n.s.</td>
<td>n.s.</td>
<td>&gt;500</td>
<td>0.005</td>
<td>&gt;500</td>
<td>0.000</td>
<td>100</td>
<td>0.075</td>
<td>&gt;24</td>
<td>&gt;24</td>
</tr>
<tr>
<td>Atenolol</td>
<td>0.823</td>
<td>12</td>
<td>n.s.</td>
<td>n.s.</td>
<td>94</td>
<td>0.070</td>
<td>96</td>
<td>0.062</td>
<td>43</td>
<td>0.108</td>
<td>&gt;24</td>
<td>&gt;24</td>
</tr>
<tr>
<td>Caffeine</td>
<td>0.869</td>
<td>9</td>
<td>n.s.</td>
<td>-8.08 ± 0.37</td>
<td>0.952</td>
<td>-8.27 ± 0.8</td>
<td>0.809</td>
<td>200</td>
<td>0.024</td>
<td>4.0</td>
<td>0.967</td>
<td>1.8</td>
</tr>
<tr>
<td>Carbamazepine</td>
<td>0.849</td>
<td>10</td>
<td>n.s.</td>
<td>n.s.</td>
<td>150</td>
<td>0.024</td>
<td>170</td>
<td>0.011</td>
<td>&gt;500</td>
<td>0.000</td>
<td>&gt;24</td>
<td>&gt;24</td>
</tr>
<tr>
<td>Codeine</td>
<td>0.908</td>
<td>8</td>
<td>1.02 ± 0.49</td>
<td>0.117</td>
<td>-7.13 ± 0.65</td>
<td>0.784</td>
<td>-7.77 ± 0.73</td>
<td>0.820</td>
<td>&gt;500</td>
<td>3.8</td>
<td>0.867</td>
<td>2.1</td>
</tr>
<tr>
<td>Furosemide</td>
<td>0.839</td>
<td>11</td>
<td>-6.08 ± 0.77</td>
<td>0.715</td>
<td>-5.9 ± 0.81</td>
<td>0.680</td>
<td>-4.57 ± 0.91</td>
<td>0.501</td>
<td>5.7</td>
<td>0.764</td>
<td>5.8</td>
<td>0.731</td>
</tr>
<tr>
<td>Gabapentin</td>
<td>0.968</td>
<td>2</td>
<td>n.s.</td>
<td>n.s.</td>
<td>460</td>
<td>0.001</td>
<td>90</td>
<td>0.062</td>
<td>110</td>
<td>0.029</td>
<td>&gt;24</td>
<td>&gt;24</td>
</tr>
<tr>
<td>Hydrochlorothiazide</td>
<td>0.944</td>
<td>3</td>
<td>1.39 ± 0.42</td>
<td>0.252</td>
<td>n.s.</td>
<td>-1.11 ± 0.48</td>
<td>0.178</td>
<td>&gt;500</td>
<td>0.000</td>
<td>300</td>
<td>0.009</td>
<td>59</td>
</tr>
<tr>
<td>Ibuprofen</td>
<td>0.919</td>
<td>6</td>
<td>n.s.</td>
<td>n.s.</td>
<td>320</td>
<td>0.002</td>
<td>71</td>
<td>0.029</td>
<td>270</td>
<td>0.004</td>
<td>&gt;24</td>
<td>&gt;24</td>
</tr>
<tr>
<td>Iopromide</td>
<td>0.377</td>
<td>14</td>
<td>-2.26 ± 0.86</td>
<td>0.215</td>
<td>-3.55 ± 0.63</td>
<td>0.557</td>
<td>-2.99 ± 0.9</td>
<td>0.307</td>
<td>17</td>
<td>0.234</td>
<td>10</td>
<td>0.609</td>
</tr>
<tr>
<td>Naproxen</td>
<td>0.912</td>
<td>7</td>
<td>n.s.</td>
<td>-3.15 ± 1.1</td>
<td>0.200</td>
<td>-3.26 ± 0.94</td>
<td>0.323</td>
<td>48</td>
<td>0.038</td>
<td>23</td>
<td>0.193</td>
<td>22</td>
</tr>
<tr>
<td>Norfloxacin</td>
<td>0.929</td>
<td>4</td>
<td>-4.46 ± 1.1</td>
<td>0.396</td>
<td>-4.89 ± 1.11</td>
<td>0.439</td>
<td>&gt;500</td>
<td>0.001</td>
<td>0.6</td>
<td>0.819</td>
<td>0.6</td>
<td>0.864</td>
</tr>
<tr>
<td>Paracetamol</td>
<td>0.811</td>
<td>13</td>
<td>n.s.</td>
<td>-8.31 ± 0.96</td>
<td>0.693</td>
<td>-6.74 ± 1.2</td>
<td>0.560</td>
<td>120</td>
<td>0.070</td>
<td>1.5</td>
<td>0.949</td>
<td>0.8</td>
</tr>
<tr>
<td>Salicylic acid</td>
<td>0.922</td>
<td>5</td>
<td>-8.26 ± 0.55</td>
<td>0.899</td>
<td>0.850</td>
<td>-7.89 ± 1.14</td>
<td>0.635</td>
<td>5.2</td>
<td>0.869</td>
<td>2.6</td>
<td>0.953</td>
<td>1.3</td>
</tr>
</tbody>
</table>

* = population biomarker rankings are from O'Brien et al. (2014) with 1 being the best marker, 14 the worst
n.s. = not significantly different from zero; bolded values = kinetic model chosen
3.3 Implications of in-sewer degradation for PPCPs on a population model and recommendations for future population models

The sewer degradation data obtained from this study represents a worst case scenario in terms of degradation processes as the A/V ratio was 70.9 m²/m³ which is about 2-10 times higher than that of average sewers. It can then be suggested that uncertainty from degradation is not an issue for acesulfame, gabapentin, atenolol, carbamazepine and ibuprofen as these compounds showed no measurable degradation under any sewer conditions over the 12 hour study period. Their stability probably help acesulfame and gabapentin achieve the first and second ranking in our list of potential population markers with high correlation between mass load and population size. The linear regressions also take into account uncertainties such as degradation/transformation in the sewer, variability from the different sampling methods, chemical analysis, and any differences in consumption and excretion in the different regions. This is because the samples collected in our earlier study were from sewer catchments of both small and large population sizes and from different localities and jurisdictions within Australia – using different sampling modes and frequencies. Therefore, it is likely that the variability observed in the linear regressions for the mass loads of these population markers versus population size presented in O’Brien et al. 13 reflects the variance associated with other factors such as differences in consumption and excretion as well as uncertainty associated with sampling and analysis.

To understand if there is a link between degradation and the correlations between mass loads and population, we correlated the R² values for the correlation between mass load and population size from the previous paper 13 and the time taken for 10% loss under each of the sewer conditions in this study and found that there was no relationship (Spearman’s rank correlation, p < 0.05) (SI Table 3). For example, norfloxacin was considered unstable under sewer conditions but had better correlation (R² = 0.912) between population size and mass load than carbamazepine (R² = 0.849) which was stable under all sewer conditions. One explanation for why degradation did not affect the
correlations between mass loads and population is that for degradation to affect a correlation there
must be a crossover between residence time and degradation (e.g. the observed degradation in
some catchments would be higher than others because of longer residence times). We speculate
that other factors may be responsible for the higher uncertainty but this requires further
investigation.

To better understand the impact of in-sewer degradation on the uncertainty of our population
estimate, we compared the coefficient of variation (CV) of the population estimate for each
catchment based on all chemicals (s14 in the paper); the stable chemicals identified in this study and
the unstable chemicals identified in this study (Figure 2). Individual substance distributions for the
three largest WWTPs are shown in SI Figure 2. Comparing stable chemicals with all chemicals (stable
and unstable chemicals combined), for all WWTP catchment population estimates made using the
model, the CV was lower (higher precision) by using all 14 chemicals in the model. However,
comparing stable chemicals with unstable chemicals a few observations can be made. Firstly, the CV
decreases (precision increases) as population increases regardless of the chemicals used in the
model. For unstable chemicals however, the uncertainty appears to plateau at ~20%. Secondly, we
observe across all catchment sizes, stable chemicals produce more precise estimates than unstable
chemicals and that the difference in the CV increases as catchment size increases. The model used a
constant error term which means the magnitude in the deviation in mass loads is constant across
population sizes (i.e. mass/day/person uncertainty stays constant across populations). It would then
be expected that smaller differences would be observed between unstable and stable chemicals
with smaller population sizes, but as the population increases, the difference in precision also
increases. Despite this, the third observation is that the inclusion of stable and unstable chemicals in
the model resulted in the most precise estimates. In Bayesian inference, the posterior distribution is
the product of the likelihood function and prior probability. In our model, the chemicals with the
highest correlation between mass load and population size (as indicated by their R2 value) implicitly
have the most weight on the likelihood that a given estimate will provide accurate population
estimates. Because stable chemicals resulted in the highest $R^2$ values (acesulfame and gabapentin), it meant that the less stable chemicals had less impact on the posterior distribution.

Figure 2 Coefficient of variation expressed as percentage for posterior estimates for population size for each of the WWTPs from O’Brien et al. (2014) using all chemicals (s14, indicated by ▲), using only the chemicals identified as stable in the sewer by this study (indicated by ▼) and using only the chemicals identified as unstable in the sewer by this study (indicated by ■).

For future calibration of a population model, if the only chemicals identified are unstable in the sewer and/or where the stable chemicals have no relationship between mass load and population size, we propose to still use less stable markers but also include a metabolite/metabolites or transformation products which are produced from the degradation of the compound within the sewer. This will then provide insight into degradation within the sewer and to also counterbalance the impact of the degradation of the parent chemical on the model i.e. as the mass load of the parent chemical decreases from in-sewer activity, the mass load of the metabolites/transformation products increases. Additionally, as shown in the earlier study, the inclusion of more chemicals (even those with lower correlation) in the model lead to increased precision of the population estimate.

3.4 Importance of the quality of calibration data used in a population model

The samples from our previous study used to calibrate the model covered a broad range of catchments from across Australia including small catchments, large catchments, different climate and geographic zones (from cold temperate to subtropical), inland and coastal catchments and catchments with potential cultural and socioeconomic differences. These factors may all influence
the consumption, excretion and degradation of the population markers investigated and thus influence the calibration data. Despite these factors, the correlation between mass load and population size was still above $R^2 = 0.8$ for 13 of the 14 population markers.

In the model, an error term was included that accounts for differences in consumption between locations and daily variations of population size. It was implied that daily variations would include factors that would contribute to either the addition (increase in mass load) or removal (decrease in mass load) of a chemical in wastewater.

Degradation is therefore one variable in daily variations and its effect on the population model is the extent that the degradation of a chemical would affect the variance in the correlation between mass load and the population size. However, if the combined impact of A/V ratios and mean residence times are all similar among catchments, then the chemicals would degrade to a similar extent and we would get a systematic underestimation of total consumption. In such a case, there is still a good correlation between mass load and population size (assuming relatively homogenous consumption and excretion) and thus we expect that the population estimate would be unaffected by the degradation. The quality of the population estimates is therefore dependent on the quality (representativeness) of the calibration data and as such, calibration data for a different area or country should include catchments of both long and short residence times as well as a combination of GS and RM sewers. By including these in the calibration data, variability due to degradation is covered implicitly. Using in-situ calibration data, the $R^2$ term helps explain homogeneity in consumption of all chemicals across catchments. To highlight this, the two highest $R^2$ correlations between mass load and population size were acesulfame and gabapentin. Acesulfame is an artificial sweetener, gabapentin is a medication used to treat epilepsy, neuropathic pain and a range of other conditions. Despite the difference in intended consumption, in order for them to have such high $R^2$ correlations there needed to be homogenous consumption amongst all catchments in the calibration dataset. Assuming non-homogenous consumption, the $R^2$ values would be lower and their influence on the model would diminish empirically.
3.5 Applicability to other sewer catchments

Degradation within sewers is dependent on the residence time in the sewer. Using data from the most comprehensive WBE study which covered 25 WWTPs in 2012 and 47 WWTPs in 2013, we correlated mean residence time of each WWTP who provided a mean residence time (n = 50) against population size (Spearman’s rank correlation, \( p < 0.05 \)) and found there is no relationship between mean residence time and population size (SI Figure 1). Therefore we consider that residence time is independent of population size and as such degradation can also be considered independent of population size.

In our previous study, chemicals for population estimation were chosen using two criteria: (1) the chemical must be measurable via direct injection on LC-MSMS in all of the collected samples; (2) the mass load of the chemical must show a correlation with population size (\( R^2 > 0.8 \)). With the use of Bayesian inference to combine the chemicals which met these criteria into a population model it was apparent that they were capable of estimating the population size with high accuracy and that uncertainty reduced with the inclusion of more markers. While these two criteria were observed as suitable criteria for our specific catchments, to extend the applicability of the population model to any other catchments based on the knowledge about chemical degradation obtained in this study we now propose two additional criteria for “Best Practice” population markers: (3) degradation should be negligible under CR conditions to ensure the stability of chemicals during wastewater collection; and (4) the mean residence time of the WWTP sewer network (RM and/or GS) should be shorter than the time for 10% degradation to occur for the compound under sewer conditions. Degradation of less than 10% is considered acceptable for our population markers as it is the acceptable “Best Practice” tolerance for degradation of sewage drug biomarkers. If no mean residence time is provided by WWTPs, we recommend using the 95th percentile of the mean residence time of all WWTPs (non-normally distributed) from the Ort et al. study which was 10

ACS Paragon Plus Environment
hours. A protocol for selecting “Best Practice” suitable population markers based on these criteria is shown in Figure 3.

Based on these criteria, for our calibration data and catchments, acesulfame, atenolol, carbamazepine, gabapentin and ibuprofen would all be considered “Best Practice” population markers particularly considering they had negligible degradation under “high in-sewer degradation processes” and hence should be considered as potential population markers for all sites regardless of knowledge of mean residence time. Their stability was also greater than the longest mean residence time of the sewer catchments in the Ort et al. study (15 hours). Hydrochlorothiazide can also be considered as a potential “Best Practice” population marker providing the sewer catchments do not have rising main sewers with mean residence times greater than 9 hours (Table 1).

Furosemide, iopromide and salicylic acid were identified as highly degradable as all compounds degraded in the CR and thus should not be used as “Best Practice” marker (failing criteria 3). The use of naproxen, caffeine, codeine, norfloxacin and paracetamol as “Best Practice” markers would require further investigation as all had greater than 10% loss in <3 hours under both gravity and rising main sewer conditions (Table 1).

Figure 3 Selection criteria of “Best Practice” population markers based on measurable levels (i.e. above LOQ), degradation in wastewater and within the sewer and correlation between mass load and population. Y = Yes. N = No.

LOQ = limit of quantification. MRT = mean residence time. * = if mean residence time is unknown, assume 10 h as it is the 95th percentile of the mean residence times from the largest wastewater-based epidemiology study.

**Supporting Information**

Comparison of sewer degradation of the population markers in this study against the literature. LC-MS/MS parameters. Population marker log $K_{OW}$ values. Pearson correlation analysis between the $R^2$ for the correlation between mass load and population size and time before 10% loss in the sewer.
reactors. Mean residence time versus population size. Individual substance distributions for estimating population size for all Large WWTPs.

Acknowledgements

The authors would like to thank Andreas Scheidegger from Eawag for his invaluable assistance in the initial model development and for the advice he provided for this study. The Queensland Alliance for Environmental Health Sciences, The University of Queensland, gratefully acknowledges the financial support of the Queensland Department of Health. This project was conducted as part of an ARC Discovery project (DP150100645). Jake O’Brien and Andrew Novic are both funded by APA PhD Scholarships. Phong K. Thai is funded by the QUT Vice Chancellor Research Fellowship. Jochen Mueller is funded through the ARC Future Fellowship (FT120100546). Guangming Jiang is funded by a Queensland State Government Early Career Accelerate Fellowship.

References


Determining "best practice" population markers to be included in a population size estimation model

- Is the chemical above LOQ? [Criteria 1] - N
- Is there a correlation between mass load & population? [Criteria 2] - N
- Is the time before 10% loss under CR conditions > 24 h? [Criteria 3] - N
- Is the time before 10% loss in the sewer > MRT*? [Criteria 4] - N

Unsuitable

Bayesian inference: all markers estimate

Population [1000, People]

STP Estimate

Actual Population

Posterior

Prior