High-resolution characterisation of short-term temporal variability in the taxonomic and resistome composition of wastewater influent

Chau KK\textsuperscript{a}, Goodall T\textsuperscript{b}, Bowes M\textsuperscript{b}, Easterbrook K\textsuperscript{c}, Brett H\textsuperscript{c}, Hughes J\textsuperscript{c}, Crook DW\textsuperscript{a,d}, Read D\textsuperscript{b}, Walker AS\textsuperscript{a,e}, Stoesser N\textsuperscript{a,d}.

Affiliations

\textsuperscript{a}Nuffield Department of Medicine, University of Oxford, John Radcliffe Hospital, Headley Way, Headington, Oxford, OX3 9DU, UK

\textsuperscript{b}UK Centre for Ecology & Hydrology, MacLean Bldg, Benson Ln, Crowmarsh Gifford, Wallingford, OX10 8BB, UK

\textsuperscript{c}Thames Water, Clearwater Court, Vastern Road, Reading, RG1 8DB, UK

\textsuperscript{d}Department of Microbiology/infectious diseases, Oxford University Hospitals NHS Foundation Trust, John Radcliffe Hospital, Headley Way, Headington, Oxford, OX3 9DU, UK

\textsuperscript{e}NIHR Oxford Biomedical Research Centre, The Joint Research Office, Second Floor, OUH Cowley, Unipart House Business Centre, Garsington Road, Oxford, OX4 2PG, UK.

Corresponding author:

Kevin K. Chau

e-mail: kevin.chau@ndm.ox.ac.uk
Abstract

Wastewater-based epidemiology (WBE) for population-level surveillance of antimicrobial resistance (AMR) is gaining significant traction, but the impact of wastewater sampling methods on results is unclear. In this study we characterised taxonomic and resistome differences between single-timepoint-grab and 24H-composites of wastewater influent from a large UK-based wastewater treatment work (WWTW [population equivalent:223,435]).

We autosampled hourly influent grab samples (n=72) over three consecutive weekdays, and prepared additional 24H-composites (n=3) from respective grabs. For taxonomic profiling, metagenomic DNA was extracted from all samples and 16S-rRNA gene sequenced. One composite and six grabs from day one underwent metagenomic sequencing for metagenomic dissimilarity estimation and resistome profiling.

Taxonomic abundances of phyla varied significantly across hourly grab samples but followed a repeating diurnal pattern for all three days. Hierarchical clustering grouped grab samples into four time periods dissimilar in both 16S rRNA gene-based profiles and metagenomic distances. 24H-composites resembled mean daily phyla abundances and showed low variability of taxonomic profiles. Of the 122 AMR gene families (AGFs) identified across all day one samples, single grab samples identified a median of 6 (IQR:5-8) AGFs not seen in the composite. However, 36/36 of these hits were at lateral coverage <0.5 (median:0.19; IQR:0.16-0.22) and potential false positives. Conversely, the 24H-composite identified three AGFs not seen in any grab with higher lateral coverage (0.82; 0.55-0.84). Additionally, several clinically significant human AGFs (blaVIM, blaIMP, blaKPC) were intermittently or completely missed by grab sampling but captured by the 24H-composite.
Wastewater influent undergoes significant taxonomic and resistome changes on short timescales potentially affecting interpretation of results based on sampling strategy. Grab samples are more convenient and potentially capture low-prevalence/transient targets but are less comprehensive and temporally variable. Therefore, we recommend 24H-composite sampling where feasible. Further validation and optimisation of WBE methods is vital for its development into a robust AMR surveillance approach.

Keywords

Wastewater-based epidemiology, antimicrobial resistance, wastewater sampling, 16S rRNA, resistome, metagenomics

Highlights

- Influent undergoes significant taxonomic/resistome changes over short timescales.
- Taxonomic abundances fluctuate diurnally but repeat for the 3 weekdays sampled.
- Detection of less prevalent AMR determinants is time-dependent for grab sampling.
- Single timepoint grab samples may produce temporally variable metagenomic profiles.
- 24H-composites reflect mean daily taxa and more reliably captured AMR determinants.
1. Introduction

Antimicrobial resistance (AMR) represents a significant challenge to global health (O’Neil, 2016). Surveillance efforts are essential to monitor trends, identify emergence and develop interventions (WHO, 2019). Wastewater-based epidemiology (WBE) is of increasing interest as a convenient surveillance approach, leveraging the pooling of human excreta to generate information on human populations at scale (Choi et al., 2018). Recent AMR-focussed WBE studies have surveyed global AMR gene distributions (Hendriksen et al., 2019), identified associations between AMR genes in clinical and wastewater samples (Karkman et al., 2020; Pärnänen et al., 2019) and characterised the relationship between wastewater AMR abundance, antimicrobial usage and clinical isolate phenotype (Perry et al., 2021). In addition to being more convenient than individual sampling for population-level surveillance, WBE can circumvent selection biases such as the typical healthcare-associated focus of most current surveillance efforts (WHO, 2018), by simultaneously capturing healthcare- and community-associated populations, and reflecting carried organisms thought to silently comprise most of the true AMR burden (Newton et al., 2015). Genotypic/sequencing-based approaches to WBE may also provide useful information on specific AMR-associated genetic determinants and high-risk clones (Tacconelli et al., 2018) whilst presenting a means to standardise AMR detection methods such as through metagenomic-based profiling of wastewater resistomes (Aarestrup and Woolhouse, 2020; Wright, 2007).

However, WBE for AMR surveillance is relatively new in terms of method development and validation. AMR-focussed WBE studies have used highly heterogenous methods, likely contributing to systematic differences in outcomes and interpretations (Chau et al., 2022). Several studies have aimed to validate specific methodology for AMR-focussed WBE such as
DNA extraction (Knudsen et al., 2016), or quantify the effect of freeze-thawing (Poulsen et al., 2021) or healthcare effluents (Perry et al., 2021), but many knowledge gaps remain.

A critical component of WBE is the approach used to collect wastewater for analyses. Previous work has suggested sampling of untreated wastewater influent as optimal due to transformation of microbial and AMR gene composition during treatment (Chau et al., 2022; Tong et al., 2019; Zhang et al., 2020). However, the composition of influent may vary significantly over short periods due to human behaviour (Guo et al., 2019) (e.g. increased flows entering the wastewater system in the morning). To account for this, methods such as flow-proportional and composite sampling combine multiple samples automatically collected throughout a defined period. These methods require autosampling equipment and multiple site visits (i.e. setup and collection), representing a significant workload. Alternatively, grab sampling entails the collection of a single volume sample at one timepoint without the need for specialised equipment or multiple site visits. However, a single sample may not be representative of daily fluctuations and could under- or over-estimate specific taxa and AMR determinants, as well as be flooded by homogenous solids (Reinthaler et al., 2013). Despite this, grab sampling remains the most common wastewater sampling method used by AMR-focussed WBE studies (Chau et al., 2022).

The trade-offs between the convenience of grab sampling and temporal coverage of composite sampling has not been previously explored for WBE-based AMR surveillance. We therefore conducted a high-resolution temporal study of wastewater influent comparing taxonomic and resistome profiles generated from contemporaneous hourly grab and 24H composite sampling. Additionally, we characterised changes in influent composition in relation to metadata such as
time-of-day, flow rate and nutrient concentrations. We aimed to highlight limitations of each method and recommend potential use cases.

2. Methods

2.1 Wastewater sampling

Wastewater influent was sampled from the crude inlet (post-mechanical screening out of large solids) of a large UK-based wastewater treatment work (WWTW) with population equivalent 223,435 and consented flow of 50,985 m$^3$/day. The WWTW received urban wastewater and healthcare effluent from multiple sites including two large hospitals (>1000 beds). Hourly influent grab samples (n=72) were collected via ice-filled autosampler (Hach AS950, Colorado, USA) over three consecutive dry-weather weekdays (14/05/19-16/05/19 [Tues-Thurs]) (Fig.1). Grab samples were collected every 24 hours, stored on ice and processed within three hours of collection. Autosampler ice was replaced twice daily (morning/evening) to keep samples cool before collection. A subset of hourly samples (n=18, 4H-interval) underwent nutrient analyses (UKCEH, Wallingford, UK) as previously described (Bowes et al., 2018). Daily 24H composites (n=3; 9am-8am) were prepared by combining 100 ml from each respective grab after thorough mixing, and all samples were pelleted (5,300RPM, 10min, 4°C) and stored at -80°C before metagenomic extraction using the DNeasy PowerSoil kit following manufacturer specifications (QIAGEN, Hilden, Germany). All sampling and sample processing was conducted following cold chain principles, with storage on ice or refrigeration during sampling, transport and processing.
2.2 Sequencing

All samples (n=75) underwent 16S rRNA gene sequencing using 515F-806R primers on the MiSeq (Illumina, San Diego, USA) as previously described (Seaton et al., 2021) generating 250bp paired-end reads. Two grab samples (2/72; timepoints 4:00 and 14:00) were excluded from analyses due to low sequencing quality. A subset of samples from day 1 (n=7; six grab samples, one 24H composite sample) also underwent metagenomic sequencing on the NovaSeq6000 at the Wellcome Trust Centre for Human Genetics (Oxford, UK), generating 150 bp paired-end reads. This was conducted to a depth of ~75 million reads (~23Gb) per sample based on previous deep sequencing and rarefaction analyses demonstrating this as the minimum depth required to capture most AMR gene diversity in samples from the same WWTW (Gweon et al., 2019).
Figure 1: Schematic overview of study methods and analyses
2.3 Computational methods

Taxonomic processing of 16S rRNA gene sequence data was conducted using DADA2 v1.16 as previously described (Seaton et al., 2021). Briefly, Illumina demultiplexed 16S rRNA gene sequences underwent DADA2 workflows to quality filter, merge, denoise and assign taxonomy using SILVA v138.1 as the training database. The resulting amplicon sequence variant (ASV) tables were processed using PhyloSeq v1.38.0 (McMurdie and Holmes, 2013) and DESeq2 v1.34.0 (Love et al., 2014) in R v4.1.2. DESeq2 was used to address heteroscedasticity of count data via local model fitting (after visual inspection of dispersion estimates to determine best fitting model) and the alternate estimation of size factor method which excludes zero counts, as per recommendations for zero-inflated metagenomic count data (McMurdie and Holmes, 2014). Principal component analyses (PCA) were conducted on variance stabilised counts using R packages stats v4.1.2, factoextra v1.0.7 (Kassambara, 2020), FactoMineR v2.4 (Lê et al., 2008) and vegan v2.6.2 (Oksanen, 2022) for ordination, hierarchical clustering and univariable fitting of environmental variables. Nutrient concentrations were normalised to flow rate before fitting. Significant abundance differences were assessed using DESeq2 likelihood ratio test with a reduced design formula focused on sampling time.

Metagenomic sequencing data was processed through ResPipe (Gweon et al., 2019) v1.4.0 for AMR genotyping, and Mash v2.3 for comparing overall similarity via metagenome distance estimation (Ondov et al., 2016) [21 mers, sketch size=10,000, minimum abundance threshold=10 k-mers]. ResPipe outputs lateral coverage as the proportion of an AMR gene covered by at least a single base of mapped reads. A higher lateral coverage reflects more complete mapping of a gene and increased confidence in true presence, with a score of 1 representing complete capture. Low lateral coverage either reflects low abundance (i.e. few
reads to map to the reference gene) or false-positives where conserved regions from different AMR determinants are erroneously mapped. AMR gene family (AGF) lateral coverage underwent column- and row-wise hierarchical clustering using pheatmap v1.0.12 (cutree_rows=3) into three groups representing predominantly high or variable lateral coverage, and those with intermittently zero lateral coverage across samples.

3. Results

3.1 Taxonomic variation

Normalised read abundances of specific phyla varied significantly across the three 24H periods but followed a diurnal pattern with peaks at 12:00, 23:00-3:00 and 7:00 (Fig.2 and Fig.S1), most notably for Firmicutes and Proteobacteria. Remaining phyla were less abundant and more consistent across the 24H periods but also appeared to follow a similar pattern. One grab sample (Day 2, 12:00) did not detect Actinobacteria, which were otherwise ubiquitous. 24H composites were similar for each day and appeared to approximately reflect the mean of phyla abundances across the hourly grab samples, capturing all phyla represented (Fig.S1). Assessing model fit via likelihood-ratio test confirmed significant abundance differences attributable to grab sampling time for 77 classifiable ASVs (adjusted p-value <0.05, 55/77 <0.01) (Fig.S2). A batch effect was also identified between all grab samples collected on day 1 versus day 3 although this was relatively minimal as only 3/77 ASVs underwent log 2-fold changes between these timepoints. Most ASVs belonged to Proteobacteria (24/77), Firmicutes (22/77) and Bacteroidota (17/77), with the rest classified as Actinobacteriota (6/77), Campylobacteriota (3/77), Fusobacteriota (2/77), Acidobacteriota (2/77) and Chloroflexi (1/77). Accordingly, PCA of 16S rRNA gene-based profiles demonstrated association between time of grab
sampling and taxonomic variation (Fig.3A). Hierarchical clustering on principal components identified four main clusters with most separation for grab samples taken between 04:00-08:00 and 17:00-22:00. Consistent with the regular daily pattern across the three sampling days (Fig.2 and Fig.S1), samples taken at the same time-of-day clustered together, regardless of the day they were collected. Remaining sampling times were represented by two partially overlapping clusters. Notably, 24H composite samples demonstrated comparably low taxonomic variability and nested centrally, surrounded by the hourly grab samples.

**Figure 2: Temporal taxonomic abundance fluctuations.** Each point represents the mean DESeq2-normalised abundance of Phyla (coloured) across the three sampling days for hourly grab samples (left) and 24H composites (right). Shaded areas (grab) or intervals (composite) convey 95% confidence intervals.
Mash distance of full metagenomic content for a subset of seven samples from sampling day one showed the same clustering pattern by PCA on 16S rRNA gene data alone (Fig.3B). Grab samples taken at 21:00 and 05:00 were most dissimilar, corresponding to the 16S-rRNA gene-based clustering around 17:00-22:00 and 04:00-08:00 respectively. The remaining grab samples were less dissimilar and congruent with the two partially overlapping clusters. As for PCA, the 24H composite was nested between grab samples although did not exactly replicate the magnitude of variance from 16S rRNA gene data alone (e.g. composite clustered closest to 01:00 based on Mash distance but 09:00 for 16S rRNA gene data). This likely reflects the higher resolution afforded by metagenomic Mash distance over targeted 16S rRNA gene sequencing methods.
Figure 3: Taxonomic and metagenomic differences between samples. (A) Principal component analysis of taxonomic profiles for single timepoint grab samples (time of sampling labelled/coloured) and composite samples (red). Sampling day and hierarchical clusters are represented by point shape and ellipses respectively. Points highlighted by squares underwent additional shotgun sequencing (B) Heatmap of Mash distance between a subset of samples from the same day (A: squares). Higher scores (darkness) represent increased dissimilarity in pairwise comparison.
Considering environmental variables, flow rate (p=0.001), dissolved fluoride (p=0.002) and dissolved ammonium (p=0.028) were independently significantly associated with 16S-gene based taxonomy, with a weaker association for dissolved nitrite (p=0.068) (Fig.S3).

3.2 Resistome variation

From the seven samples undergoing metagenomic sequencing, in total 122 AMR gene families (AGFs) were identified at various levels of lateral coverage (i.e. mapping completeness). Of the 122 AGFs, single grab samples identified between 90/122 (74%) and 101/122 (83%) AGFs, while the 24H composite identified 101/122 (83%) AGFs. When comparing all AGFs identified in all grab samples to all AGFs identified in the 24H composite, grab samples identified a median six (IQR:5-8) AGFs not seen in the composite (total: 36 AGFs across grab samples of which 21 were uniquely identified), and the composite identified three AGFs not seen in any grab sample. However, 36/36 of grab sampling hits missed by composite had lateral coverage <0.5 (median:0.19; IQR:0.16-0.22) whereas composite hits missed by all grab samples had significantly higher lateral coverage (0.82; 0.55-0.84). Therefore, when comparing single timepoint grab sampling (i.e. considering only one grab) to 24H composite, grab sampling likely misses more truly present AGFs and identifies more false-positives than 24H composite sampling.

AGFs identified hierarchically clustered into three main groups based on lateral coverage across samples: predominantly high, variable and intermittently identified. 53/122 (43%) AGFs had predominantly high lateral coverage (median: 0.90; IQR: 0.74-1), 44/122 (36%) intermittent (0; 0-0.20) and 25/122 (20%) variable (0.54; 0.44-0.62) coverage. Several beta-lactamase AGFs of significant human clinical importance had lateral coverage following these
groupings (Fig.S4), such as $bla_{OXA}$ and $bla_{CTX-M}$ families with consistently high and variable lateral coverage respectively. Notably, the carbapenemase gene families $bla_{VIM}$, $bla_{KPC}$ and $bla_{IMP}$ were completely missed by several grab samples but identified in the composite. Conversely, the carbapenemase gene family $bla_{NDM}$ was only identified in a single grab sample, and not seen in the composite.

Figure 5: Heatmap of AMR gene family (AGF) lateral coverage across samples from the same sampling day. Row-wise hierarchical clustering divides AGFs into consistently high lateral coverage (top), intermittently 0 lateral coverage (middle) and variable lateral coverage (bottom).
4. Discussion

In this study, we characterised the short-term temporal changes in wastewater taxonomy using 16S rRNA gene sequencing on an hourly scale over three days, profiled taxonomic distributions and resistome profiles using metagenomics four-hourly throughout a single day, and concurrently compared the profiling performance of single timepoint grab sampling versus 24H composite sampling. We find results are significantly influenced by both the sampling method used and the time of sampling for grab samples, representing an important consideration for future studies analysing wastewater.

4.1 Temporal taxonomic changes

Taxonomic abundances fluctuated on an hourly basis and reflected a diurnal pattern which repeated for all three sampling days. Firmicutes and Proteobacteria were the most abundant phyla overall, consistent with their classifications as dominant human gut microbiota (Rinninella et al., 2019). They also constituted most ASVs (47/77; 60%) undergoing significant abundance changes. Three other human gut associated phyla also underwent significant fluctuation (Bacteroidota, Actinobacteriota, Fusobacteriota), further reflecting human faecal content in wastewater. However, whilst wastewater does reflect human microbiomes (Newton et al., 2015), its composition is also influenced by the sewer environment during conveyance to WWTWs (Fahrenfeld and Bisceglia, 2016), and by contributions from other wider environmental and anthropogenic sources. Accordingly, our abundance estimates for wastewater phyla are not completely consistent with estimates of Firmicutes and Bacteroidota in direct human metagenomic studies (typically ~90% abundance (Rinninella et al., 2019)), but are similar to other wastewater estimates (Azli et al., 2022; Numberger et al., 2019; Polo et al.,...
2020). Campylobacteriota, Acidobacteriota and Chloroflexi also significantly fluctuated, and have been previously reported in humans, soil and wastewater respectively (Speirs et al., 2019). Interestingly, Chloroflexi significantly increased in abundance during periods of lower flow rate, possibly indicating increased sewer community contributions (Santo Domingo et al., 2011).

Whilst hourly taxonomic composition fluctuated significantly, overall daily variation was relatively consistent across the three consecutive sampling days, and day of sampling did not appear to drive taxonomic variation. This was also true for 24H composite samples which consistently approximated mean abundance of phyla from each contributing set of 24-hourly grab samples. Principal components of 24H composites across three days projected closely, indicating low variation between composites, and nested centrally amongst the grab samples, suggesting representativeness. We purposefully avoided sampling on weekends and wet weather days to prevent previously described factors such as weekend household behaviour (Polo et al., 2020) and surface water runoff/infiltration (Shanks et al., 2013) from impacting influent composition. Without external factors such as these, we find that sampling one weekday in our setting is representative of others in the same week, although extended longitudinal sampling would be needed to assess this in further detail.

Hierarchical clustering of principal components produced distinct clusters for grab samples taken during 04:00-08:00 or 17:00-22:00. This may further reflect changes in flow and input as these periods mostly avoid apparent flow balancing events (see below). Flow rate was also identified as significantly associated with taxonomic profiling, likely contributing both directly as a result of discharged community effluent and indirectly via carriage of sewer communities (Fahrenfeld and Bisceglia, 2016). Our findings corroborate previously reported diurnal
compositional fluctuations associated with flow rate related to human behaviour/household discharge patterns (Guo et al., 2019). However, in addition to morning peaks reported by Guo et al., we observed additional abundance increases unrelated to household behaviour (i.e. occurring outside normal waking period) at ~1:00am. This pattern may have been missed by Guo et al. due to lower resolution sampling (4-hourly) or instead reflect WWTW infrastructure differences since our sampling site utilised multiple pumping stations, balancing tanks and hydro-brakes to limit maximum flow during peak use and divert wastewater for treatment during low flow periods (e.g. overnight). In addition to 16S rRNA gene-based taxonomic assessments, metagenomic evaluation of a subset of samples showed similar clustering by sampling time, indicating that this also impacts on the capacity to detect important functional features in addition to taxonomic variation, including antimicrobial resistance markers. WWTW and sewer infrastructure and approaches to wastewater management are therefore important factors to consider when interpreting results, but are often overlooked (Chau et al., 2022).

4.2 Resistome profiling

The 122 AMR gene families (AGFs) identified from the metagenomic subset of same-day samples (9:00, 13:00, 17:00, 21:00, 1:00, 5:00, 24H composite) divided into three groups of consistent, variable and intermittent detection. When identified with lateral coverage=1 (high confidence in true presence (Gweon et al., 2019)) in more than one sample, AGFs were generally also captured at 1 or high lateral coverage by most other samples (i.e. consistent detection), as seen for 53/122 (43%) AGFs. This suggests parts of the resistome remain consistently detectable throughout the day and can be identified with confidence at variable timepoints; these are likely highly prevalent AMR determinants circulating in the community.
and/or wastewater. Variable AGFs (25/122 (20%)) were predominantly detected with high confidence in at least one sample (e.g. MOX beta-lactamases, fosfomycin thiol transferases), and therefore are likely truly present but undergo temporal flux within a day. Thus, these AGFs likely represent circulating AMR of intermediate prevalence whose identification is somewhat temporally dependent but identifiable with reasonable confidence at varied timepoints. Lastly, detection of intermittent AGFs (44/122 (36%)) was highly time dependent, suggesting low prevalence and/or transient AMR determinants. Detecting these may rely on sampling during or temporally close to shedding events as hypothesised for SARS-CoV-2 (Michael-Kordatou et al., 2020; Polo et al., 2020). However, the low lateral coverage seen for most intermittent AGFs may alternatively represent mapping artifacts; in this case, the AGFs may not be truly present.

While grab sampling identified several more AGFs in total across all six individual timepoints, the 24H composite more consistently identified AGFs with high lateral coverage. Almost all AGFs identified by grab samples and missed by the composite had low lateral coverage and were potentially false-positives or transiently present. Conversely, composite sampling identified AGFs likely to be truly present but missed by all grab samples. It is important to note that the 24H composite was derived from all hourly grab samples over 24 hours and not just the six timepoints which underwent metagenomic shotgun sequencing. Therefore, when an AGF such as blaKPC is missed by most or all the six grab samples but captured with high lateral coverage in the composite, we can assume the AGF is present in other timepoints comprising the composite. If relying on grab sampling only (i.e. a single grab sample), the broader picture may be misrepresented by temporal fluctuation since a single timepoint grab only represents a narrow snapshot. For example, blaVIM, blaIMP and blaKPC beta-lactamases were all detected in the grab sample taken at 13:00, whereas all were absent in the grab sample at 21:00 and absent.
from at least one other timepoint. In this respect, the 24H composite is more representative, capturing the intermittently detectable VIM, IMP and KPC carbapenemase families. Therefore, 24H composites appear more reliable for detecting variable and intermittently detectable AGFs (moderate to low prevalence in community/wastewater) subject to temporal flux which may be missed by a single timepoint grab depending on sampling time.

However, sensitivity of detection also depends on sequencing depth and sample complexity (Gweon et al., 2019) which may explain why certain single grab samples identified AGFs missed in composite samples. For example, blaNDM was identified in one grab sample but completely missed in all other grabs and the 24H composite. This may represent extremely rare circulation (consistent with local clinical prevalence (UKHSA, 2021)) which is captured by chance sampling within the timeframe of a shedding event from colonised individuals. Since this single timepoint grab sample and others are effectively diluted in the composite, sensitivity to detect low prevalence, infrequently shed, AGFs is reduced in the composite.

4.3 Limitations

Resource limitations meant that we were only able to use 16S rRNA gene sequencing to characterise the 72 hourly grab samples collected; shotgun metagenomics enables a much higher-resolution and robust characterisation of genetic features (Langille et al., 2013; Manor and Borenstein, 2017). 16S rRNA gene sequencing may also be prone to amplification biases and normalisation issues. To mitigate these issues, we undertook metagenomic sequencing of a subset of samples which supported our 16S rRNA gene-based conclusions and enabled us to additionally evaluate the impact of sampling approach on resistome profiles. We also employed DESeq2 internal normalisation and variance stabilising transformation to avoid
heteroscedasticity introduced by relative abundance or rarefaction methods. The sensitivity of metagenomic sequencing to accurately profile species and AMR genes in samples is dependent on sequencing depth (Gweon et al., 2019); we sequenced at an optimised depth previously determined by ultradeep sequencing of wastewater from the same WWTW as our study. Our study was carried out in a single setting and represents sampling over three days of a single week only; findings may therefore not be fully generalisable, but they nevertheless clearly highlight the relevance of considering the impact of sampling approach on results and their interpretation.

5. Conclusion

The composition of wastewater influent undergoes significant taxonomic and resistome changes on short timescales which may affect interpretation of results based on sampling strategy. Grab sampling at a single timepoint generated significant differences in taxonomic profile which clustered based on sampling time, driven in part by human behaviour and flow rate. Although single grab sampling had greater sensitivity for potential rare/transient AMR determinants, resistome profiles were temporally variable and might not be generalisable beyond that timepoint. Additionally, many AMR determinants detected by grab sampling but missed by 24H composites were probable false-positives. 24H composites reflected overall daily taxonomic composition and more reliably captured AMR determinants present throughout the day or completely missed by grab sampling.

Consequently, we recommend 24H composite sampling over single timepoint grab samples where feasible. If grab sampling is undertaken for convenience, consideration should be given to the fact that it may misrepresent taxonomic and AGF abundances. Future work identifying
an optimal sampling time for grab sampling might be useful but challenging due to extensive
differences in relevant factors in different settings such as sampling site size, population and
infrastructure. Passive sampling methods may also represent an effective alternative to
traditional grab or composite sampling, and have shown potential in monitoring SARS-CoV-2
(Schang et al., 2021). Validation and optimisation of WBE methods for AMR surveillance is
vital for its development as a robust surveillance approach; context-specific validation for
methods selected could be useful to ensure that results and interpretation are robust.

**Funding**

This work was supported by the Medical Research Foundation National PhD Training
Programme in Antimicrobial Resistance Research, the National Institute for Health Research
(NIHR) Health Protection Research Unit in Healthcare Associated Infections and
Antimicrobial Resistance at University of Oxford (NIHR200915) in partnership with the UK
Health Security Agency (UKHSA), and the NIHR Oxford Biomedical Research Centre. The
report presents independent research. The views expressed in this publication are those of the
authors and not necessarily those of the NHS, NIHR, UKHSA, the Department of Health and
Social Care or Thames Water.

KKC is a Medical Research Foundation PhD student (ref. MRF-145-0004-TPG-AVISO). NS
is an Oxford Martin School fellow and an NIHR Oxford BRC Senior Research Fellow. ASW
is an NIHR Senior Investigator.

**Availability of data and materials**

Raw sequencing data are available under ENA project PRJEB52722
(https://www.ebi.ac.uk/ena/browser/view/PRJEB52722) and PRJEB54594
(https://www.ebi.ac.uk/ena/browser/view/PRJEB54594), with all metadata included in supplementary file 1. R scripts and markdown renders are available at https://github.com/KaibondChau/Chau_etal_Hires.

**Acknowledgements**

We thank Thames Water for facilitating access to the WWTW and flow data.

**Contributors**

A complete list of author contributions as per CRediT; Contributor Roles Taxonomy is given in the appendix.
References


