Wastewater influent and population-level surveillance of antimicrobial resistance in Oxfordshire using metagenomic sequencing



Kevin Kaibond Chau

Hertford College, University of Oxford

Nuffield Department of Medicine

A thesis submitted for the degree of Doctor of Philosophy

Trinity 2022

Table of Contents

Abstract 6

0. Preface 8

0.1 Acknowledgements 8

0.2 Funding 9

0.3 Index of Figures and Tables 9

0.4 Papers arising directly from this work 13

Peer reviewed 13

Preprints currently under review 13

0.5 Other papers arising in the course of this work 13

Peer reviewed 13

Preprints currently under review 19

Under review without preprint 20

0.6 Declaration of contribution to thesis 20

0.6.1 Chapter 2 20

0.6.2 Chapter 3 21

0.6.3 Chapter 4 21

0.6.4 Chapter 5 21

0.6.5 Chapters 3, 4 and 5 22

0.7 Abbreviations 22

1. Introduction 24

1.1 Antimicrobial resistance 24

1.2 Direction of this thesis 25

1.3 Epidemiology, public health and syndromic surveillance 26

1.4 Asymptomatic carriage and colonisation 28

1.5 Microbial shedding and wastewater 30

1.6 Thesis outline 32

2. Systematic review of wastewater surveillance of antimicrobial resistance in human population 35

2.1 Introduction 35

2.2 Materials and Methods 39

2.2.1 Literature search 40

2.2.2 Eligibility criteria 40

2.2.3 Study selection and data extraction 43

2.2.4 Risk of bias and Certainty assessment 43

2.2.5 Data synthesis and analysis 45

2.3 Results 47

2.3.1 Study selection, risk of bias and characteristics 47

2.3.1.1 Literature screen 47

2.3.1.2 Risk of bias and certainty assessment 48

2.3.1.3 Summary of general study characteristics 49

2.3.1.4 Summary of AMR detection methods in included studies 51

2.3.2 Wastewater-human AMR concordance 51

2.3.2.1 Phenotypic wastewater-human AMR concordance 51

2.3.2.2 Genotypic wastewater-human AMR concordance 55

2.3.3 Synthesis of study features 57

2.3.3.1 Logistic regression of study features and wastewater-human AMR agreement 57

2.3.3.2 Study features descriptive synthesis 58

2.3.3.3 Studies without extractable data 59

2.4 Discussion 63

2.4.1 Effect of AMR detection method on wastewater-human AMR concordance 63

2.4.2 Effect of study features on wastewater-human AMR concordance 65

2.4.3 Recommendations regarding risk of bias 67

2.4.4 Limitations 69

2.5 Supplementary Material 71

3. Evaluation of metagenomic, 16S rRNA gene and ultra-plexed PCR-based sequencing approaches for profiling antimicrobial resistance gene and bacterial taxonomic composition of polymicrobial samples 109

3.1 Background 109

3.2 Methods 112

3.2.1 DNA mock community preparation and wastewater samples 112

3.2.2 Sequencing 113

3.2.3 Software 115

3.2.4 Taxonomic profile analysis 117
3.2.5 Sensitivity and specificity calculations for AMR gene detection 117
3.3 Results 120
3.3.1 Taxonomic profiling 120
3.3.2 AMR gene profiling 122
3.3.3 Real-world wastewater samples 126
3.3.3.1 Taxonomic profiles 126
3.3.3.2 AMR profiles 128
3.4 Discussion 129
3.4.1 Mock community evaluation 130
3.4.2 “Real-world” wastewater samples 132
3.4.3 Limitations 134
4. High-resolution characterisation of short-term temporal variability in the taxonomic and resistome composition of wastewater influent 137
4.1 Introduction 137
4.2 Methods 139
4.2.1 Wastewater sampling 139
4.2.2 Sequencing 140
4.2.3 Computational methods 142
4.3 Results 143
4.3.1 Taxonomic variation 143
4.3.2 Resistome variation 149
4.4 Discussion 153
4.4.1 Temporal taxonomic changes 153
4.4.2 Resistome profiling 155
4.4.3 Limitations 158
5. Fine-scale spatial and longitudinal survey of AMR prevalence, antibiotic prescribing and antibiotic residues across Oxfordshire WWTWs 160
5.1 Introduction 160
5.2 Methods 163
5.2.1 Site-level feature analyses and site selection 163
5.2.1.1 Data sources 163
5.2.1.2 Data processing 164

5.2.1.3 Practice-level antibiotic prescribing 166
5.2.1.4 Presence of care homes, agriculture and abattoirs 168
5.2.1.5 Relative Deprivation 168
5.2.1.6 Site selection summary 168
5.2.2 Wastewater sampling and processing 169
5.2.3 Sequencing 172
5.2.4 Contemporaneous primary care antibiotic prescribing data 173
5.2.5 Metagenomic analysis 173
5.2.5.1 AMR profiling 173
5.2.5.2 Taxonomic profiling 174
5.2.6 Statistical analysis 175
5.2.6.1 Differential abundance testing 175
5.2.6.2 Generalised linear modelling 176
5.3 Results 178
5.3.1 Distribution of taxa 178
5.3.2 Distribution of AMR genes 182
5.3.3 Generalised linear modelling 185
5.3.3.1 Association between wastewater AMR abundance, local primary care prescribing and wastewater antibiotic residues 186
5.3.3.2 Antibiotic residue concentrations significantly vary across WWTWs 194
5.3.4 Antibiotic residue concentrations reaching selective levels 199
5.4 Discussion 202
5.4.1 Significant differences in taxonomy and resistomes across sites and timepoints 202
5.4.2 Generalised linear modelling 205
5.4.2.1 Do city-level WWTW wastewater resistomes reflect regional WWTWs? 205
5.4.2.2 Does higher primary care prescribing associate to higher wastewater AMR prevalence? 206
5.4.2.3 Wastewater antibiotic residues and selection of AMR 207
5.4.3 Limitations 208
5.4.4 Conclusion 209
6. Conclusions and future work 211
6.1 Future work 218
6.2 Concluding remarks 220
Bibliography 222

Abstract

Antimicrobial resistance (AMR) represents a significant challenge to global health comprising substantial complexity driving its emergence and spread. Thus, surveillance efforts are essential to monitor trends, identify emergence and develop interventions, and of particular importance in curtailing rapid global dissemination facilitated by plasmids and other mobile genetic elements. Accordingly, the development and deployment of novel surveillance approaches is prioritised in the World Health Organisation's global AMR action plan. 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. Whilst WBE has found success in the surveillance of polioviruses and SARS-CoV-2, its application to monitoring population-level AMR prevalence remains limited to date. Part of the challenge lies in the paucity of clearly reported studies and validation of methodology which hinders effective iterative development.

In this thesis, I employed metagenomic sequencing methods to investigate several fundamental questions of methodology and study design regarding the application of WBE to population-level AMR surveillance. By initially conducting a comprehensive synthesis of shortcomings and knowledge gaps in the existing literature, I tailored my experimental chapters to specifically address gaps in knowledge whilst employing best practices gleaned from the review. I benchmarked the performance of multiple sequencing and bioinformatic approaches in the context of WBE AMR surveillance. I also investigated the impact of initial wastewater sampling methods on profiling taxonomic and resistome composition, with parallel characterisation of short-term temporal

fluctuation and associated drivers. Lastly, I conducted a multi-site longitudinal survey to explore factors associated with wastewater composition such as antibiotic prescribing and residue concentrations.

A common recurring theme throughout this thesis is the importance of validating individual methods to systematically identify sources of variation when developing workflows for WBE. The work I have conducted highlights the complexity involved in the analysis and deconvolution of wastewater for AMR surveillance, and demonstrates the need for clearer reporting of high-quality metadata to facilitate Iterative development of optimal practice. Additionally, interdisciplinary collaboration is vital for this field where insight from colleagues in areas of chemical and viral surveillance as well as the often-overlooked expertise of individuals working outside of research but within the wastewater industry is invaluable. Future work underscored by collaboration and the sharing of granular metadata will significantly accelerate the development of wastewater-based epidemiology into a robust surveillance tool for the mitigation of antimicrobial resistance.

0. Preface

0.1 Acknowledgements

I am immensely grateful for the incredible support, guidance, resources and encouragement given to me by my outstanding supervisors: Derrick Crook, Daniel Read, Sarah Walker and Nicole Stoesser. Their expertise, enthusiasm and mentorship have been invaluable during this period and were indisputably pivotal to the completion of this thesis. In particular, I thank Nicole for consistently going above and beyond in her role as my primary supervisor. I have undertaken many amazing opportunities I was plainly unaware of or did not believe myself capable of thanks to Nicole's encouragement and motivating presence. Her attentive support in challenging periods both academic and personal has my sincere gratitude. The work described herein would not have been possible without the substantial efforts of my supervisors and I consider myself extremely fortunate to have been their student.

I also extend my thanks to the whole Modernising Medical Microbiology group for their contributions, support, and friendship, particularly Ali Vaughan, Carla Wright, Gillian Rodger, Karina-Doris Vihta and Emma Pritchard. I thank my childhood friends too for boosting morale in times of need, particularly Anthony Peskett and Ryan Arding. I am also indebted to my wonderful partner Leanne Barker for her constant support and patience throughout the past four years, and who has picked me up (literally off the floor) many times when I stumbled throughout the DPhil – all the while as a medical student no less.

My family has also been a critical lifeline, especially my auntie Linh Sang who has been an inspiring example of persevering in the face of great adversity. Finally, to my parents Mui

and Hatu Chau who never had the chance to pursue education past the equivalent of primary school and who endured untold hardship growing up during the Vietnam War, I thank you wholeheartedly for your sacrifices and untiring efforts in providing me a better life and opportunities than afforded to you. I strive to embody their indomitable spirit and continue to apply myself with fervour in their honour.

0.2 Funding

This work was supported by the Medical Research Foundation National PhD Training Programme in Antimicrobial Resistance Research (MRF-145-0004-TPG-AVISO). Additional funding was provided by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Healthcare Associated Infections and Antimicrobial Resistance at University of Oxford in partnership with the UK Health Security Agency (UKHSA), and the NIHR Oxford Biomedical Research Centre (NIHR200915).

0.3 Index of Figures and Tables

All figures have been additionally deposited in higher resolution at (https://figshare.com/s/213d6678488e0f08fc7f)

Table 1 - Methodological features potentially contributing to variability in outcomes _____38
Figure 1 - Screening flowchart _____42
Figure 2 - PRISMA flowchart of search strategy and study inclusion/exclusions _____47
Figure 3 - Geographic distribution of wastewater sampling and test approach of included studies _____50
Figure 4 - AMR in wastewater isolates and human isolates for phenotypic (A) and genotypic (B) comparisons. _____53
Table S1 - Modifications of Cochrane risk of bias tool used for risk of bias assessment _____71
Table S2 - Adapted GRADE domains used for certainty assessment_____73

Table S3 - Certainty assessment results and rationale for upgrades and downgrades _____74
Figure S1 - Global density of wastewater sampling of included studies _____75
Table S4 – Descriptive table 1 of included studies _____76
Table S5 – Descriptive table 2 of included studies _____87
Figure S2 - AMR prevalence in wastewater and human isolates from low bias studies only _____92
Figure S3 - Phenotypic aminoglycoside resistance prevalence estimates in human and wastewater isolates for aminoglycosides by study _____93
Figure S4 - Phenotypic resistance prevalence estimates in human and wastewater isolates for beta-lactams by study _____94
Figure S5 - Phenotypic resistance prevalence estimates in human and wastewater isolates for beta-lactam/beta-lactamase inhibitor combinations by study _____94
Figure S6 - Phenotypic resistance prevalence estimates in human and wastewater isolates for cephalosporins combinations by study _____95
Figure S7 - Phenotypic resistance prevalence estimates in human and wastewater isolates for carbapenems by study _____95
Figure S8 - Phenotypic resistance prevalence estimates in human and wastewater isolates for fluoroquinolones/quinolones by study _____96
Figure S9 - Phenotypic resistance prevalence estimates in human and wastewater isolates for glycopeptides by study _____96
Figure S10 - Phenotypic resistance prevalence estimates in human and wastewater isolates for macrolides by study _____97
Figure S11 - Phenotypic resistance prevalence estimates in human and wastewater isolates for ESBLs and MRSA by study _____97
Figure S12 - Genotypic prevalence estimates in human and wastewater isolates for aac variants by study _____98
Figure S13 - Genotypic prevalence estimates in human and wastewater isolates for arr variants by study_____99
Figure S14 - Genotypic prevalence estimates in human and wastewater isolates for ctx variants by study_____100
Figure S15 - Genotypic prevalence estimates in human and wastewater isolates for erm variants by study _____101
Figure S16 - Genotypic prevalence estimates in human and wastewater isolates for fos variants by study_____102

Figure S17 - Genotypic prevalence estimates in human and wastewater isolates for oxa variants by study _____103
Figure S18 - Genotypic prevalence estimates in human and wastewater isolates for van variants by study _____104
Figure S19 - Genotypic prevalence estimates in human and wastewater isolates for qnr variants by study _____104
Table S6 - Univariable logistic regression of study features associated with high-agreement. __105
Table S7 - Study features excluded from logistic regression due to missingness _____107
Table S8 - Tally of features and agreement _____108
Figure 5 - Schematic overview of evaluated sequencing approaches and respective bioinformatic pipelines tested for taxonomic and resistome profiling of the DNA mock. _____113
Figure 6 - Bioinformatic pipeline versions and associated reference databases _____116
Figure 7 - AmpliSeek scoring thresholds _____118
Figure 8 - (A) Theoretical taxonomic composition of DNA mock as compared to actual generated pipeline outputs faceted by sequencing approach. _____121
Figure 9 - (A) Detection sensitivity of known AMR genes present in the mock at 100% identity. 123
Figure 10 - Heatmap of AmpliSeq pipeline calls labelled by reference marker ID and sequence identity. _____124
Figure 11 - Replicate results of AmpliSeq profiling._____125
Figure 12 - Relative abundance of phyla for wastewater samples (1-7) faceted by sequencing approach and pipeline combination._____127
Figure 13 - Distribution of diversity metrics across all wastewater samples coloured by sequencing approach and pipeline combination._____128
Figure 14 - Schematic overview of study methods and analyses _____141
Figure 15 - Temporal taxonomic abundance fluctuations. _____144
Figure 16 - Log 2-fold changes in abundance for ASVs undergoing significant abundance changes across grab samples. _____146
Figure 17 - Taxonomic and metagenomic differences between samples. _____148
Figure 18 - Heatmap of AMR gene family lateral coverage across samples from the same sampling day._____151
Figure 19 - Heatmap of selected AMR gene family lateral coverage across samples from the same sampling day. _____152
Figure 20 - Site selection summary _____166
Figure 21 - Historic practice-level prescribing of shortlisted sites_____167

Table 2 - Table of targeted chemicals for residue measurement – modified from Holton et al.[152] _____170
Figure 22 - Heatmap of significantly differentially abundant genera identified by consensus-based DESeq2 and ALDEx2 testing. _____180
Figure 23 - Principal coordinate analysis (left) with hierarchical clusters (ellipses) and SRS-normalised relative abundance (right) of taxonomic composition _____181
Figure 24 - Principal coordinate analysis of resistome composition (left) with hierarchical clusters (ellipses) and aggregated FPKM of AMR genes conferring resistance to antimicrobial classes (right).____183
Figure 25 - Heatmap of AMR gene FPKM aggregated by antimicrobial class _____185
Figure 26 - Generalised linear model assessing AMR against timepoint and WWTW. _____187
Figure 27 - Generalised linear models assessing class-level AMR against timepoint and WWTW 188
Figure 28 - Generalised linear model assessing macrolide AMR against timepoint, WWTW, macrolide prescribing and residue levels. _____190
Figure 29 - Generalised linear model assessing tetracycline AMR against timepoint, WWTW and tetracycline prescribing.____191
Figure 30 - Generalised linear model assessing nitroimidazole AMR against timepoint, WWTW and nitroimidazole residues.____192
Table 3 - Table of coefficients for univariable models assessing associations between prescribing and residues with class-level AMR abundance _____193
Figure 31 - Antibiotic residue concentrations across sites and timepoint. _____195
Figure 32 - Generalised linear model assessing class-level antibiotic concentrations against timepoint and WWTW _____197
Table 4 - Table of coefficients for univariable models assessing associations between class-level prescribing and residue concentrations. _____198
Figure 33 - Antibiotic residue concentration measurements for 10 antibiotics detected directly as their original chemical form. _____199
Figure 34 - Antibiotic residue concentration measurements for four antibiotics exceeding respective PNECs _____201

0.4 Papers arising directly from this work

Peer reviewed

Chau, K., Barker, L., Budgell, E., Vihta, K., Sims, N., Kasprzyk-Hordern, B., Harriss, E., Crook, D., Read, D., Walker, A. and Stoesser, N., 2022. Systematic review of wastewater surveillance of antimicrobial resistance in human populations. *Environment International*, 162, p.107171.

Preprints currently under review

Chau, K., Matlock, W., Constantinides, B., Lipworth, S., Newbold, L., Tipper, H., Goodall, T., Brett, H., Hughes, J., Crook, D., Eyre, D., Read, D., Walker, A. and Stoesser, N. 2022. Evaluation of metagenomic, 16S rRNA gene and ultra-plexed PCR-based sequencing approaches for profiling antimicrobial resistance gene and bacterial taxonomic composition of polymicrobial samples. *bioRxiv preprint*.

Chau, K., Goodall, T., Bowes, M., Easterbrook, K., Brett, H., Hughes, J., Crook, D., Read, D., Walker, A. and Stoesser, N. 2022. High-resolution characterisation of short-term temporal variability in the taxonomic and resistome composition of wastewater influent. *bioRxiv preprint*.

0.5 Other papers arising in the course of this work

Peer reviewed

Constantinides, B*., **Chau, K.K***., Phuong Quan, T., Rodger, G., Andersson, M.I., Jeffery, K., Lipworth, S., Gweon, H.S., Peniket, A., Pike, G., Millo, J., Byukusenge, M., Holdaway, M.,

Gibbons, C., Mathers, A.J., Crook, D.W., Peto, T.E.A., Sarah Walker, A., Stoesser, N., 2020. Genomic surveillance of Escherichia coli and Klebsiella spp. in hospital sink drains and patients. *Microbial Genomics* 6, 4–16.

* Joint first author

Matlock, W., **Chau, K.K**., AbuOun, M., Stubberfield, E., Barker, L., Kavanagh, J., Pickford, H., Gilson, D., Smith, R.P., Gweon, H.S., Hoosdally, S.J., Swann, J., Sebra, R., Bailey, M.J., Peto, T.E.A., Crook, D.W., Anjum, M.F., Read, D.S., Walker, A.S., Stoesser, N., Shaw, L.P & REHAB consortium., 2021. Genomic network analysis of environmental and livestock F-type plasmid populations. *ISME J*. 15, 2322–2335.

Shaw, L.P., **Chau, K.K**., Kavanagh, J., AbuOun, M., Stubberfield, E., Gweon, H.S., Barker, L., Rodger, G., Bowes, M.J., Hubbard, A.T.M., Pickford, H., Swann, J., Gilson, D., Smith, R.P., Hoosdally, S.J., Sebra, R., Brett, H., Peto, T.E.A., Bailey, M.J., Crook, D.W., Read, D.S., Anjum, M.F., Walker, A.S., Stoesser, N., 2021. Niche and local geography shape the pangenome of wastewater-and livestock-associated Enterobacteriaceae. *Scientific advances*. 7, 1–12. https://doi.org/10.1126/SCIADV.ABE3868

Lipworth, S., Vihta, K.D., **Chau, K**., Barker, L., George, S., Kavanagh, J., Davies, T., Vaughan, A., Andersson, M., Jeffery, K., Oakley, S., Morgan, M., Hopkins, S., Peto, T.E.A., Crook, D.W., Walker, A.S., Stoesser, N., 2021. Ten-year longitudinal molecular epidemiology study of Escherichia coli and Klebsiella species bloodstream infections in Oxfordshire, UK. *Genome Medicine*. 13, 1–13.

Lipworth, S., Vihta, K.D., **Chau, K**., Kavanagh, J., Davies, T., George, S., Barker, L., Vaughan, A., Andersson, M., Jeffery, K., Oakley, S., Morgan, M., Peto, T., Crook, D., Walker, A.,

Stoesser, N., 2021. Ten years of population-level genomic Escherichia coli and Klebsiella pneumoniae serotype surveillance informs vaccine development for invasive infections. *Clinical Infectious Diseases*, ciab006.

Lumley, S.F., Rodger, G., Constantinides, B., Sanderson, N., **Chau, K.K**., et al., 2021. An Observational Cohort Study on the Incidence of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection and B.1.1.7 Variant Infection in Healthcare Workers by Antibody and Vaccination Status. *Clinical Infectious Diseases*. 2, 1–12.

Lipworth S, Vihta KD, Davies T, Wright S, Tabirao M, Chau K, Vaughan A, Kavanagh J, Barker L, George S, Segal S, Paulus S, Barrett L, Oakley S, Jeffery K, Butcher L, Peto T, Crook D, Walker S, Kadambari S, Stoesser N. Molecular epidemiology and antimicrobial resistance phenotype of paediatric bloodstream infections caused by Gram-negative bacteria. *Communications Medicine*. 11;2:101.

AbuOun, M., Jones, H., Stubberfield, E., Gilson, D., Shaw, L.P., Hubbard, A.T.M., **Chau, K.K**., Sebra, R., Peto, T.E.A., Crook, D.W., Read, D.S., Gweon, H.S., Walker, A.S., Stoesser, N., Smith, R.P., Anjum, M.F., 2021. A genomic epidemiological study shows that prevalence of antimicrobial resistance in Enterobacterales is associated with the livestock host, as well as antimicrobial usage. *Microbial Genomics* 7.

Lumley, S.F., Constantinides, B., Sanderson, N., Rodger, G., Street, T.L., Swann, J., **Chau, K.K**., O'Donnell, D., Warren, F., Hoosdally, S., O'Donnell, A.M., Walker, T.M., Stoesser, N.E., Butcher, L., Peto, T.E., Crook, D.W., Jeffery, K., Matthews, P.C., Eyre, D.W., 2021. Epidemiological data and genome sequencing reveals that nosocomial transmission of SARS-CoV-2 is underestimated and mostly mediated by a small number of highly infectious individuals. *Journal of Infection*. 83, 473–482.

Auguet, O.T., Niehus, R., Gweon, H.S., Berkley, J.A., Waichungo, J., Njim, T., Edgeworth, J.D., Batra, R., **Chau, K**., Swann, J., Walker, S.A., Peto, T.E.A., Crook, D.W., Lamble, S., Turner, P., Cooper, B.S., Stoesser, N., 2021. Population-level faecal metagenomic profiling as a tool to predict antimicrobial resistance in Enterobacterales isolates causing invasive infections: An exploratory study across Cambodia, Kenya, and the UK. *EClinicalMedicine* 36, 100910.

Lumley, S.F., Eyre, D.W., McNaughton, A.L., Howarth, A., Hoosdally, S., Hatch, S.B., Kavanagh, J., **Chau, K.K**., Downs, L.O., Cox, S., Dunn, L., Justice, A., Wareing, S., Dingle, K., Rudkin, J., Auckland, K., Fyfe, A., Bolton, J., Paton, R., Mentzer, A.J., Jeffery, K., Andersson, M.I., James, T., Peto, T.E.A., Marsden, B.D., Screaton, G., Cornall, R.J., Klenerman, P., Ebner, D., Stuart, D.I., Crook, D.W., Stoesser, N., Kennedy, S.H., Thompson, C., Gupta, S., Matthews, P.C., 2020. SARS-CoV-2 antibody prevalence, titres and neutralising activity in an antenatal cohort, United Kingdom, 14 April to 15 June 2020. *Eurosurveillance* 25, 1–6.

Pouwels, K.B., House, T., Pritchard, E., Robotham, J. V., Birrell, P.J., Gelman, A., Vihta, K.D., Bowers, N., Boreham, I., Thomas, H., Lewis, J., Bell, I., Bell, J.I., Newton, J.N., Farrar, J., Diamond, I., Benton, P., Walker, Ann Sarah, Pouwels, K.B., Walker, A. Sarah, Crook, D., Matthews, P.C., Peto, T., Stoesser, N., Howarth, A., Doherty, G., Kavanagh, J., **Chau, K.K**., et al., 2021. Community prevalence of SARS-CoV-2 in England from April to November, 2020: results from the ONS Coronavirus Infection Survey. *Lancet Public Health*. 6, e30–e38.

Wei, J., Stoesser, N., Matthews, P.C., Ayoubkhani, D., Studley, R., Bell, I., Bell, J.I., Newton, J.N., Farrar, J., Diamond, I., Rourke, E., Howarth, A., Marsden, B.D., Hoosdally, S., Jones, E.Y., Stuart, D.I., Crook, D.W., Peto, T.E.A., Pouwels, K.B., Eyre, D.W., Walker, A.S.,

Lambert, A., Thomas, T., Black, R., Felton, A., Crees, M., Jones, J., Lloyd, L., Sutherland, E., Pritchard, E., Vihta, K.D., Doherty, G., Kavanagh, J., **Chau, K.K**., et al., 2021. Antibody responses to SARS-CoV-2 vaccines in 45,965 adults from the general population of the United Kingdom. *Nature Microbiology*. 6, 1140–1149.

Wei, J., Matthews, P.C., Stoesser, N., Maddox, T., Lorenzi, L., Studley, R., Bell, J.I., Newton, J.N., Farrar, J., Diamond, I., Rourke, E., Howarth, A., Marsden, B.D., Hoosdally, S., Jones, E.Y., Stuart, D.I., Crook, D.W., Peto, T.E.A., Pouwels, K.B., Walker, A.S., Eyre, D.W., Thomas, T., Cook, D., Ayoubkhani, D., Black, R., Felton, A., Crees, M., Jones, J., Lloyd, L., Sutherland, E., Pritchard, E., Vihta, K.-D., Doherty, G., Kavanagh, J., **Chau, K.K**., et al., 2021. Anti-spike antibody response to natural SARS-CoV-2 infection in the general population. *Nature Communications*. 12, 1–12.

Peto, T., Affron, D., Afrough, B., Agasu, A., Ainsworth, M., Allanson, A., Allen, K., Allen, C., Archer, L., Ashbridge, N., Aurfan, I., Avery, M., Badenoch, E., Bagga, P., Balaji, R., Baldwin, E., Barraclough, S., Beane, C., Bell, J., Benford, T., Bird, S., Bishop, M., Bloss, A., Body, R., Boulton, R., Bown, A., Bratten, C., Bridgeman, C., Britton, D., Brooks, T., Broughton-Smith, M., Brown, P., Buck, B., Butcher, E., Byrne, W., Calderon, G., Campbell, S., Carr, O., Carter, P., Carter, D., Cathrall, M., Catton, M., Chadwick, J., Chapman, D., **Chau, K.K**., et al., 2021. COVID-19: Rapid antigen detection for SARS-CoV-2 by lateral flow assay: A national systematic evaluation of sensitivity and specificity for mass-testing. *EClinicalMedicine* 36, 0–6.

Lumley, S.F., O'Donnell, D., Stoesser, N.E., Matthews, P.C., Howarth, A., Hatch, S.B., Marsden, B.D., Cox, S., James, T., Warren, F., Peck, L.J., Ritter, T.G., de Toledo, Z., Warren, L., Axten, D., Cornall, R.J., Jones, E.Y., Stuart, D.I., Screaton, G., Ebner, D., Hoosdally, S.,

Chand, M., Crook, D.W., O'Donnell, A.-M., Conlon, C.P., Pouwels, K.B., Walker, A.S., Peto, T.E.A., Hopkins, S., Walker, T.M., Jeffery, K., Eyre, D.W., **Oxford University Hospitals Staff Testing Group**., 2021. Antibody Status and Incidence of SARS-CoV-2 Infection in Health Care Workers. *New England Journal of Medicine*. 384, 533–540.

Kotanidis, C.P., Xie, C., Alexander, D., Rodrigues, J.C.L., Burnham, K., Mentzer, A., O'Connor, D., Knight, J., Siddique, M., Lockstone, H., Thomas, S., Kotronias, R., Oikonomou, E.K., Badi, I., Lyasheva, M., Shirodaria, C., Lumley, S.F., Constantinides, B., Sanderson, N., Rodger, G., Chau, K.K., et al., Constructing custom-made radiotranscriptomic signatures of vascular inflammation from routine CT angiograms: a prospective outcomes validation study in COVID-19. *The Lancet Digital Health*.

Walker, A.S., Pritchard, E., House, T., Robotham, J. V., Birrell, P.J., Bell, I., Bell, J.I., Newton, J.N., Farrar, J., Diamond, I., Studley, R., Hay, J., Vihta, K.D., Peto, T.E.A., Stoesser, N., Matthews, P.C., Eyre, D.W., Pouwels, K.B., **COVID-19 Infection Survey team**., 2021. CT threshold values, a proxy for viral load in community sars-cov-2 cases, demonstrate wide variation across populations and over time. *Elife* 10.

Lumley, S.F., Wei, J., O'Donnell, D., Stoesser, N.E., Matthews, P.C., Howarth, A., Hatch, S.B., Marsden, B.D., Cox, S., James, T., Peck, L.J., Ritter, T.G., de Toledo, Z., Cornall, R.J., Jones, E.Y., Stuart, D.I., Screaton, G., Ebner, D., Hoosdally, S., Crook, D.W., Conlon, C.P., Pouwels, K.B., Walker, A.S., Peto, T.E.A., Walker, T.M., Jeffery, K., Eyre, D.W., **Oxford University Hospitals Staff Testing Group**., 2021. The Duration, Dynamics, and Determinants of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Antibody Responses in Individual Healthcare Workers. *Clinical Infectious Disease*. 73, e699–e709.

Lipworth, S., Pickford, H., Sanderson, N., **Chau, K.K**., Kavanagh, J., Barker, L., Vaughan, A., Swann, J., Andersson, M., Jeffery, K., Morgan, M., Peto, T.E.A., Crook, D.W., Stoesser, N., Walker, A.S., 2020. Optimized use of oxford nanopore flowcells for hybrid assemblies. *Microbial Genomics* 6, 1–12.

Eyre, D.W., Lumley, S.F., O'donnell, D., Campbell, M., Sims, E., Lawson, E., Warren, F., James, T., Cox, S., Howarth, A., Doherty, G., Hatch, S.B., Kavanagh, J., **Chau, K.K**., et al., 2020. Differential occupational risks to healthcare workers from SARS-CoV-2 observed during a prospective observational study. *Elife*.

Ainsworth, M., Andersson, M., Auckland, K., Baillie, J.K., Barnes, E., Beer, S., Beveridge, A., Bibi, S., Blackwell, L., Borak, M., Bown, A., Brooks, T., Burgess-Brown, N.A., Camara, S., Catton, M., **Chau, K.K**., et al., 2020. Performance characteristics of five immunoassays for SARS-CoV-2: a head-to-head benchmark comparison. *Lancet Infectious Disease*. 20, 1390–1400.

Preprints currently under review

Walker, A.S., Vihta, K.-D., Gethings, O., Pritchard, E., Jones, J., House, T., Bell, I., Bell, J.I., Newton, J.N., Farrar, J., Diamond, I., Studley, R., Rourke, E., Hay, J., Hopkins, S., Crook, D., Peto, T., Matthews, P.C., Eyre, D.W., Stoesser, N., Pouwels, K.B., **COVID-19 Infection Survey team**., 2021. Increased infections, but not viral burden, with a new SARS-CoV-2 variant. *medRxiv preprint*.

Lumley, S.F., O'Donnell, D., Stoesser, N.E., Matthews, P.C., Howarth, A., Hatch, S.B., Marsden, B.D., Cox, S., James, T., Warren, F., Peck, L.J., Ritter, T.G., de Toledo, Z., Warren,

L., Axten, D., Cornall, R.J., Jones, E.Y., Stuart, D.I., Screaton, G., Ebner, D., Hoosdally, S., Chand, M., **Oxford University Hospitals Staff Testing Group**., et al., 2020. Antibodies to SARS-CoV-2 are associated with protection against reinfection. *medRxiv preprint*.

Under review without preprint

Stanton, I.C., Tipper, H.J., Chau, K., Klumper, U. Subirats, J., Murray, A.K., 2022. Does environmental exposure to PPCP residues result in the selection of antimicrobial-resistant microorganisms, and is this important in terms of human health outcomes?. *Environmental Toxicology and Chemistry [Manuscript submitted for publication]*.

0.6 Declaration of contribution to thesis

I, Kevin Kaibond Chau, performed all fieldwork, DNA extraction and designed and conducted all of the analyses presented in this thesis with appropriate support from my supervisors and colleagues. I hereby declare and gratefully acknowledge the assistance I have received from others in relation to this work below.

0.6.1 Chapter 2

I am grateful to Elinor Harriss from the University of Oxford Bodleian Libraries who aided in the design and search strategy, as well as providing general advice regarding the systematic review. I also acknowledge the assistance of Leanne Barker and Natalie Sims for their contributions to the duplicate processing of records, as well as Karina-Doris Vihta and Eric Budgell for providing general statistical advice.

0.6.2 Chapter 3

Samuel Lipworth provided the previous sequencing data for the clinical isolates spiked into the mock community sample. Holly Tipper and Lindsay Newbold performed the 16S rRNA gene and shotgun sequencing for the mock community sample. Tim Goodall taught me how to perform 16S rRNA gene library prep and aided in the sequencing of seven wastewater samples. For assistance in the development and release of the AmpliSeek pipeline, I acknowledge David Eyre who advised on determining mapping thresholds, and Bede Constantinides and William Matlock who provided general programming advice.

0.6.3 Chapter 4

Mike Bowes performed nutrient analyses for 18 wastewater samples and Tim Goodall aided me in the sequencing of 75 wastewater samples. I also acknowledge the work of the team at the Wellcome Trust Centre for Human Genetics who undertook library preparation and metagenomic sequencing on my behalf for seven samples. In particular, I am grateful to Sarah Lamble who organised sequencing efforts and partially mitigated Covid-19 delays.

0.6.4 Chapter 5

I am grateful to Karina-Doris Vihta who provided me with her R code to help clean the data used in initial site selection as well as providing general statistical advice throughout. I am also grateful to Emma Pritchard who taught me aspects of statistical modelling. Chemical analyses were possible with the collaboration of Natalie Sims and Barbara Kaspryzk-Hordern from the University of Bath. In particular, I am grateful for the advice

and on-site solid phase extraction training. Natalie Sims also conducted UPLC-MS/MS on my behalf due to Covid-19 restrictions. I also acknowledge the work of the team at the Wellcome Trust Centre for Human Genetics who undertook library preparation and metagenomic sequencing on my behalf for 25 samples. The Modernising Medical Microbiology informatics team including Jeremy Swann and Geoffrey Ma provided invaluable infrastructure support. I also thank Soon Gweon for partially processing metagenomic data my behalf using the latest version ResPipe he is currently developing.

0.6.5 Chapters 3, 4 and 5

All wastewater sampling was facilitated by members of Thames Water who helped organise site access, required documentation and training. In particular, I acknowledge the assistance of Howard Brett, Jennifer Hughes, Karen Easterbrook, Robert Pritchard, Marta Gorska and Richard Whinton. I am especially grateful to Karen Easterbrook and Richard Whinton who explained processes in detail and consulted on sampling methodology.

0.7 Abbreviations

AMR – Antimicrobial resistance

AST - Antimicrobial susceptibility testing

ASV - Amplicon sequence variant

CARD - The Comprehensive Antibiotic Resistance Database

CCC – Concordance correlation coefficient

CI – Confidence interval

FPKM - Fragments Per Kilobase Million

GP – General practice

IQR – Interquartile range

LMIC - Low-middle-income country

LSOA - Lower layer super output areas

MAPE - Mean absolute proportion error

MIC - Minimum inhibitory concentrations

NPV - Negative predictive value

PCR – Polymerase chain reaction

PE – Paired-end

PNEC - Predicted no-effect concentration

PPV – Positive predictive value

qPCR - Quantitative polymerase chain reaction

UTI – Urinary tract infection

WBE – Wastewater-based epidemiology

WGS – Whole genome sequencing

1. Introduction

1.1 Antimicrobial resistance

In 1945, following on from his discovery of the world's first broadly-effective antimicrobial, Alexander Fleming concluded his Nobel lecture by sounding a note of warning – "It is not difficult to make microbes resistant to penicillin"[1]. Only seventeen years prior, Fleming made his fortuitous discovery which would save countless lives, inspire drug-discovery and herald the start of the antibiotic age. But in this short period, *Staphylococcus aureus* had already been experimentally rendered resistant to the action of penicillin[2], and unfortunately, despite his prescience, Fleming's warnings proved prophetic and antimicrobial resistance (AMR) has developed into a highly complex all-pervasive threat to global health[3]. Presently, AMR has evolved beyond the simple hypothetical scenarios proposed by Fleming, with highly diverse genetic underpinnings and interconnected transmission networks affected by globalisation, healthcare, agriculture, industrial pollution and waste management[4]. The antibiotic age has revolutionised modern medicine with antimicrobial chemotherapy critical to the management and prevention of infections across specialties, but consequently, all aspects of modern medicine are increasingly jeopardized by AMR. A potential post-antibiotic era has been touted as apocalyptic and although dramatic, annual deaths attributable to AMR are already estimated to exceed 1.27 million currently[5], and predicted to exceed 10 million by 2050 with corresponding economic costs of $100 trillion[6]. Reversing the tide is a difficult proposition but concerted global efforts have been established such as the WHO Global action plan on antimicrobial resistance which prioritises key objectives to address AMR in a multi-pronged strategic

manner[7]. One such objective is to "Strengthen the knowledge and evidence base through surveillance and research" which represents a primary focus of my thesis.

1.2 Direction of this thesis

Whilst the aim of the WHO Global action plan to enhance AMR surveillance encompasses several varied disciplines, I have focussed on a surveillance approach known as wastewater-based epidemiology (WBE) due to its significant potential to address multiple important gaps in knowledge highlighted by the WHO. These include:

- The development of novel surveillance approaches capable of providing AMR data in a timely manner to inform local, national and regional actions; and to monitor effectiveness of interventions.
- Understanding how AMR circulates between humans, animals, water and the environment to inform policies and regulations.
- The ability to rapidly identify and characterise emerging resistance mechanisms to facilitate corresponding rapid responses, and to characterise the wider "resistome" (i.e. the entire complement of AMR mechanisms present in a given sample) in an agnostic manner

To this end, I aimed to assess the basic and translational science behind WBE as applied to AMR for human population-level surveillance. In the following sections of this Introduction, I will briefly review the relevant history, disciplines, strategies and evidence underpinning WBE where not specifically addressed in my subsequent thesis chapters.

1.3 Epidemiology, public health and syndromic surveillance

At a high-level, epidemiology is the study of the who and why of disease occurrence to further evaluate disease dynamics[8]. Early examples include the work of hobbyist statistician John Graunt who was dissatisfied with the reporting of London plague deaths in 1634 and performed his own appraisal, analysis and collation of data to publish possibly the first large epidemiological report – "Natural and Political Observations Made Upon the Bills of Mortality"[9]. Whilst this report was pioneering, it was not until 1854 that epidemiology and public health made revolutionary, joint, progress where in perhaps the most famous application, John Snow halted an outbreak of cholera by attributing clusters of cases to a contaminated water pump, facilitating targeted intervention[10]. Whilst modern epidemiology is now enriched by large-scale datasets such as electronic healthcare records and DNA sequencing data, identifying common features amongst diseased or affected cases remains a fundamental principle.

When applied continuously to specific noteworthy diseases, epidemiology underpins public health surveillance and intervention - "the ongoing systematic collection, analysis, and interpretation of data, closely integrated with the timely dissemination of these data to those responsible for preventing and controlling disease"[11]. Tracking of affected cases in (or close to) real-time facilitates identification of trends or clusters at an early stage and allows timely response to maximise disease prevention. For the UK, rapid detection of potential outbreaks and epidemics has utilised the notifications of infection diseases system since 1968 which relies on statutory reporting upon suspicion of a notifiable disease or organism by registered medical practitioners[12]. However, this system and most passive

surveillance approaches are by and large extensions of the traditional clinical approach and thus mostly dependent on the sampling of individuals seeking healthcare. The lag phase between emerging disease in a community and its observation in healthcare represents a target of syndromic surveillance systems which aim to discern population health indicators before presentation in secondary healthcare settings (i.e. hospitalisation)[13].

Syndromic surveillance entails the collation of health-related data to infer possible future trends and provide early warning before cases become prevalent. In the UK, the UKHSA employs the number of individuals visiting general practices and emergency departments, the number of ambulance callouts, and the number of calls to NHS 111[14]. By tracking the number of consultations and the reasons for those consultations, patterns can be discerned such as the seasonality of consultations for respiratory syndromes without laboratory results to identify specific pathogens. Alternative digital data sources outside of healthcare are also worth noting such as the aggregation of online searches made by individuals. Much interest was garnered by the now defunct Google Flu Trends in 2008 which aimed to predict influenza circulation using Google Search queries of flu-like symptoms. Initial reports of 97% correlation to laboratory-based data with a two-week lead were extremely promising[15], but subsequent analyses found modelling deficiencies and data noise led to periods of significant inaccuracy[16]. However, lessons were learnt and recent models accurately forecasting SARS-CoV-2 prevalence with search aggregation have done so by addressing biases in search numbers attributable to public interest/news media coverage[17]. Interestingly, Twitter represents potentially the most useful digital data source due to its relatively open data policy and geolocation capability which facilitate modelling

of prevalence as a function of tweet frequency as well as disease spread as a function of geolocated movement[18]. Whilst these are excellent examples of currently relatively under-used data sources, such approaches are limited by reliance on an internet-savvy subpopulation and potentially inaccurate self-diagnoses of symptoms, and omit asymptomatic individuals entirely.

1.4 Asymptomatic carriage and colonisation

Syndromic surveillance and healthcare-related activity generally targets change in human behaviour due to symptoms (i.e. those experienced by an individual which characterise a health issue). However, in the absence of symptoms, individuals do not generally seek healthcare or undertake internet-related activity searching for or tweeting about symptoms. There is no significant change in behaviour and thus these surveillance approaches will be blind to asymptomatic individuals that may be contributing to dissemination of potential pathogens and/or AMR mechanisms. This blind spot to asymptomatic carriers represents a key limitation of traditional surveillance[19], highlighted historically by the case of "Typhoid Mary". Mary Mallon was the first identified asymptomatic carrier of *Salmonella enterica* Typhi and responsible for infecting up to 122 individuals in New York City with typhoid from 1900-1915[20]. She experienced no symptoms but infected almost all individuals she cooked for or lived with during this period. Epidemiological efforts identified Mallon as the common link between outbreaks and whilst her subsequent treatment by health authorities remains controversial, Mary Mallon represents a well-known example of how asymptomatic carriage can silently further the spread of infectious diseases[21].

In addition to spread in the community, asymptomatic carriage poses a risk to the carrier themselves by frequently representing a risk factor for infection[22]. For example, this may arise due to immune disruption, indwelling devices, injury, surgery or stochastically. Given that surveillance is dependent on accurately estimating the rate of disease, the lack of insight into the prevalence of asymptomatic carriers of potential pathogens may result in inadequate public health responses, as previously demonstrated for *Clostridium difficile*[23] and influenza[24]. Monitoring AMR reflects a key issue with strategies focused on syndromic surveillance only as most AMR genes are carried and exchanged asymptomatically within human or other niches as part of bacterial interactions in the microbiome[25]. This exchange is facilitated through multiple modes of horizontal gene transfer via mobile genetic elements, such as plasmids[26]. Prior asymptomatic colonisation by AMR bacteria has been associated with increased risk of subsequent resistant infections[27], as well as representing a major route for wider dissemination. As such, precautions against some silently circulating AMR-associated pathogens (such as methicillin-resistant Staphylococcus aureus [MRSA] or carbapenemase-producing Enterobacterales [CPE]) are routinely in place across UK hospitals to prevent nosocomial spread. For example, national and trust-level policies generally require MRSA screening for high-risk patients, patients being admitted to high-risk units (e.g. ICUs) and inter-hospital transfers[28]. Upon positive result, patients undergo decolonisation/suppression treatment before admission or treatments with high infection risk such as surgery. Whilst this approach curtails a subset of nosocomial infections, it is currently only conducted for a handful of known target organisms.

Therefore, disease-, syndromic- or organism-focused surveillance systems are largely inadequate for properly evaluating the true burden of AMR by failing to capture asymptomatic carriage of AMR in healthcare- and particularly non-healthcare-associated populations. Direct screening of community-associated populations may provide useful data similar to NHS population screening programmes which target specific conditions/pathogens such as the HPV virus in specific subpopulations[29]. However, individual-level screening of a large enough and representative sample presents many logistical and ethical barriers.

1.5 Microbial shedding and wastewater

Humans constantly shed their microflora into the immediate environment. For instance, dead skin cells slough off with commensal microbes still attached and oral microflora escape in fomite droplets to hands and surfaces. Chambered bioaerosol studies have estimated approximately one million biological particles are emitted by an individual per hour during seated office-work, increasing up to 6-fold with movement[30]. This of course facilitates the onward transmission of many pathogens but simultaneously represents a potential surveillance target. The largest and most relevant human microbiome with respect to AMR is represented by organisms in the gastrointestinal tract, which harbours a vast milieu of microbial species, including both benign/beneficial commensals and potential pathogens[31]. This microbiome is readily reflected in faecal samples, which can be analysed to monitor infection/colonisation status of an individual. In fact, gastrointestinal pathogens such as enteroviruses heavily rely on shedding to facilitate faecal-oral transmission, and this has been exploited for surveillance purposes.

Poliovirus[32] and most recently SARS-CoV-2[33-36] have been accurately monitored through analysis of wastewater whereby excreted virions/viral RNA can be detected in an approximate dose-dependent fashion based on the number of individuals infected/shedding the viral material in the wastewater catchment area. This approach is known as wastewater-based epidemiology (WBE) and relies on the analysis of wastewater to generate population-scale information on the catchment area targeted[37].

The use of wastewater potentially bypasses several aforementioned limitations including the capability to provide data on the traditionally inaccessible asymptomatic population[38,39]. However, this approach does require that the surveillance target is readily shed into wastewater streams such as for enteroviruses. AMR is a potential suitable candidate for wastewater-based surveillance since gut colonisation by AMR bacteria usually precedes infection and disease, with additional flora such as from the skin, oral and upper respiratory tract also contributing to waste streams (e.g. toilets, showers, sinks). As such, WBE for AMR surveillance is likely mechanistically similar to monitoring for enterovirus or SARS-CoV-2 with regard to shedding, although the extent to which microbes in wastewater systems establish their own, distinct biomes remains unclear. The application of WBE to monitor population-level AMR prevalence is at a more developmental stage than established syndromic-/organism-focused surveillance approaches, and the separate use case of WBE for population-level chemical analyses, such as for recreational drug use[40]. Part of the aim of this thesis (Chapter 2) was therefore to evaluate previous research using WBE for AMR surveillance to identify the gaps in knowledge which need addressing. By tailoring the rest of my thesis to these

fundamental questions of study design and methodology, I aimed to contribute to the development of a robust methodology to underpin future efforts to use WBE to effectively monitor population-level AMR.

1.6 Thesis outline

The aim of this thesis is to identify and address fundamental questions regarding the application of wastewater-based epidemiology (WBE) to population-level AMR surveillance, with a focus on sampling and data processing methodologies. In Chapter 2, I conducted a comprehensive systematic review of previous studies employing wastewater to monitor human AMR prevalence, with the aim of characterising the strength of the evidence supporting the approach and assessing potential best practices and remaining knowledge gaps. Given that no systematic review on this topic existed to date, this chapter was purposefully conducted to provide insight on the state of the developing field as a whole through evidence-based synthesis. Using the insights gained from the systematic review in Chapter 2, I tailored my remaining thesis chapters to specifically addressing methodological aspects of employing WBE for population-level AMR surveillance (Chapter 3 and Chapter 4), and to evaluating whether WBE could be used to reflect primary-care level antimicrobial prescribing within a given locality (Chapter 5).

In Chapter 3, I benchmarked the performance of several culture-independent sequencing and bioinformatic approaches for profiling the taxonomic and resistome composition of a well-defined mock microbial community. Furthermore, given the relative simplicity of the

mock community, I also investigated the variation between profiles generated by the different methods when applied to real-world wastewater samples.

Given the heterogeneity in wastewater sampling methods observed across studies in Chapter 2, in Chapter 4 I explored the impact of the two most commonly utilised sampling methods (24-hourly composite and grab sampling) on the interpretation of wastewater composition. I also characterised the short-term temporal variability of wastewater composition, investigated potential explanatory factors, and made recommendations regarding the potential use case and limitations of each method.

Lastly, in Chapter 5, I conducted a fine-scale longitudinal survey of wastewater across Oxfordshire using metagenomic sequencing and measurement of antibiotic residue levels. I investigated the spatiotemporal variation in taxonomic and resistome composition as well as modelled the contributions of local primary care prescribing volume and antibiotic residue levels on wastewater taxonomic and AMR gene composition.

I conclude the thesis with a discussion of implications for future studies utilising WBE for AMR surveillance as well as outlining both immediate future work arising from my thesis findings and more broadly from other recent research studies within the field.

Uniquely, this thesis represents several “start-to-finish” projects encompassing considerable multi-disciplinarity which has facilitated my development of a diverse range

of skills. In addition to further developing my molecular skillset regarding laboratory aspects of DNA sequencing, I have acquired skills in establishing a multi-site longitudinal study and conducted extensive fieldwork to collect all wastewater samples myself. Furthermore, I arranged an interdisciplinary collaboration enabling chemical analyses of antibiotic residues and my training in chemistry workflows. Lastly, I developed computational skills for the bioinformatic, statistical and modelling analyses of generated data.

2. Systematic review of wastewater surveillance of antimicrobial resistance in human population

Note: This chapter has previously been published[41]. I authored the text and figures below (which are reproduced here largely unaltered from the published work) with input from supervisors commensurate with the amount of input/advice that would be considered appropriate for a DPhil thesis. All authors have agreed to the inclusion of this published work in my thesis.

2.1 Introduction

Antimicrobial resistance (AMR) is a significant threat to global health[6] and a multi-faceted problem compounded by diverse drivers facilitating its emergence and spread. AMR surveillance is critical to understanding trends, monitoring interventions and developing empiric treatment guidelines, as prioritised in the World Health Organisation's global AMR action plan[7]. Large networks sharing AMR data have been established to meet this need, including the European Antimicrobial Resistance Surveillance Network (EARS-Net) and the Global Antimicrobial Resistance Surveillance System (GLASS). However, current surveillance can be limited by the reliance on individual-level sampling, which is often affected by selection bias towards healthcare-associated settings[42]. For example, both EARS-Net and GLASS target AMR in clinical specimens from hospitalised patients; this however does not reliably capture AMR prevalence in commensal organisms, thought to silently constitute most of the true AMR burden[19,38,43]. Additionally, data collection is often limited to a subset of culturable species, and on susceptibility phenotypes rather

than AMR genotypes. This lack of genotyping hampers the surveillance of high-risk AMR-associated clones and specific AMR-associated genetic determinants[44]. In many settings, particularly low- and middle-income countries (LMICs) where the burden of AMR is largest, the laboratory infrastructure to support individual, patient-level surveillance is lacking.

Wastewater-based epidemiology (WBE) is an epidemiological approach based on the analyses of wastewater (e.g. sewage) to generate information on human populations on a community scale[37]. WBE has the potential to overcome some of the aforementioned challenges by simultaneously sampling both healthcare- and community-associated populations at scale[45]. The approach has already been successful in illicit drug monitoring[40] and pathogen surveillance[32,46], including SARS-CoV-2[36], and its application to AMR surveillance is gaining traction[38]. Difficulties in standardising AMR detection methods and targets across surveillance networks[44] could potentially be circumvented by using metagenomics to agnostically probe wastewater resistomes[39,47]. Recent wastewater-based studies have investigated seasonal/geographic AMR distributions[48], quantified global AMR gene abundance[43] and identified associations between AMR in wastewater and clinical samples[49,50]. However, heterogeneous study designs and methods likely contribute to differences in outcomes/interpretations. The impact of methodological approaches such as grab sampling (i.e. taking single samples at a single timepoint)[51], snapshot versus longitudinal study design, sampling in the presence of unrepresentative and “contaminating” AMR-associated point sources, and/or

characterising AMR based on phenotypic testing of isolates versus genotypic profiling remains poorly understood (Table 1).

Despite the increasing use of WBE for AMR surveillance/evaluation purposes, there has been no attempt to review the available data, synthesise the evidence, and assess remaining knowledge gaps. I therefore systematically reviewed studies using wastewater (i.e. the "wastewater compartment") for AMR surveillance in human populations (i.e. the "human compartment"), seeking firstly to characterise the strength of the AMR prevalence associations observed between wastewater and human compartments (i.e. reflecting whether this appears to be a promising surveillance approach) and secondly to identify methodological factors that might optimise these associations (i.e. in support of a standardised approach for wastewater-based AMR surveillance going forwards). I specifically focussed on study design and methodological approaches, including AMR detection methods, highlighting limitations and recommendations for future work.

Table 1 - Methodological features potentially contributing to variability in outcomes

Methodological features	Examples of methodological feature	Aspects potentially introducing variability in outcomes	References
Wastewater sampling point type	Wastewater treatment works (WWTW) sampling point e.g. influent versus effluent	Treatment processes can transform microbial and AMR composition resulting in differences between e.g. influent and effluent samples	Tong et al., 2019[52]; Zhang et al., 2020[53]
	• Hospital effluent • Domestic sewers/manholes	Focussed sampling may only represent specific sub-populations*	Jakobsen et al., 2008[54]; Larson et al., 2020[34]*
	Informal sewer systems	Informal sewer systems (often with low flow) may be susceptible to homogeneity	Fahrenfeld and Bisceglia, 2016[38]
Wastewater sampling method	Grab (single sample)	• Single grab samples can be flooded by homogenous solids • Wastewater composition can vary significantly over short time periods	Reinthaler et al., 2013[51]; Guo et al., 2019[55]
	• Composite sampling (combining grabs) • Proportional sampling (flow/time/volume)	Composite and proportional samples capture average composition but may be unable to discriminate peak values during sampling period*	Michael-kordatou et al., 2020[33]*
WWTW sewershed inputs	• Hospitals • Other healthcare facilities (e.g. care homes) • Industry • Abattoirs	• Effluent from AMR-associated sources may obscure detection of true population-level trends (e.g. elevated levels of unique AMR, co-selection of plasmids, non-human associated AMR)* • Unique substrate properties of industry effluent can influence WWTW communities	Jakobsen et al., 2008[54]; Shchegolkova et al., 2016[56]; Larson et al., 2020[34]*
	• Agriculture • Rural communities	Subsistence farming and inadequate waste management can result in frequent contact between human and environmental reservoirs	Pehrsson et al., 2016[57]
WWTW properties and sampling conditions	Size of WWTW sewershed and infrastructure	• Long conveyance times from population to sampling point may impact composition due to transformation in unique sewer environment (anaerobic, temperature, biofilms) • Presence of pre-treatment infrastructure (e.g. pumping stations, balancing tanks) may also play a role in transforming wastewater	Fahrenfeld and Bisceglia, 2016[38]
	Treatment methods	When sampling treated wastewater, differing levels of treatment can selectively transform AMR and microbial composition	Tong et al., 2019[52]; Zhang et al., 2020[53]
	Geography and weather (seasons, rainfall, temperature, latitude)	• Heavy rainfall dilutes wastewater in combined sewer systems via rainwater runoff and by infiltration of groundwater (dislodged biofilms, freshwater taxa) • Local ambient environment and climate can influence both human-associated microorganisms entering the system and resident sewer microbiota	Shanks et al., 2013[58]

Methodological features	Examples of methodological feature	Aspects potentially introducing variability in outcomes	References
	Flow rate	• Combined sewer overflows impact composition of post-treatment samples collected during these events • Flow rate is also associated to hydraulic retention time and the level of WWTW treatment	Honda et al., 2020[59]
Sample processing methods	• Filtration • Storage conditions • Sodium thiosulfate (chlorine neutraliser)	Different sample processing methods may selectively affect recovery yields of specific species*	Ahmed et al., 2020[36]*
	Freeze-thaw cycles	Multiple freeze-thaw cycles shown to select for Firmicutes, Actinobacteria, and eukaryotic microorganisms	Poulsen et al., 2021[60]
	DNA extraction methods	Use of different metagenomic DNA extraction kits and procedures has been shown to modulate inferred microbial composition*	Knudsen et al., 2016[61]; Michael-kordatou et al., 2020[33]*
AMR detection methods	• Culture-based and phenotypic (selective media, disk-diffusion, microbroth dilution) • Culture-based and genotypic (WGS, PCR, qPCR) • Direct-from-sample genotypic (qPCR, metagenomics)	Methods based on culturing isolates may only capture a fraction of the diversity present even with detailed sampling	Shaw et al., 2021[62]
		Culture-based methods may be subject to variations from phenotyping method and breakpoint selection	Davies, 2019[63]; Davies et al., 2020[64]
		Bioinformatic deconvolution can be subject to variation depending on which tools/databases/references are used.	Lal Gupta et al., 2020[65]

*Also recently studied as sources of variability in outcomes for SARS-CoV-2 wastewater-based epidemiology and likely relevant for AMR surveillance.

2.2 Materials and Methods

For this systematic review, I sought firstly to evaluate concordance between wastewater and human AMR prevalence estimates for each study, stratified by the AMR detection method used (i.e. phenotypic versus genotypic). Secondly, I adapted the PECOTS (Population, Exposure/Intervention, Comparator, Outcome, Target Condition, Study Design) systematic review framework and formulated the following statement to assess association between study methods and outcomes: Among studies jointly evaluating

AMR prevalence in wastewater and humans, what is the effect of methodological approaches (e.g. wastewater sampling methods, AMR detection methods) on the concordance between these metrics?

A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist is available at 10.6084/m9.figshare.20873110.v1, and the complete PROSPERO (International prospective register of systematic reviews) protocol is available at: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019134946.

2.2.1 Literature search

The search string was developed through iterative preliminary searches in consultation with a librarian experienced with systematic reviews. Full search strings adapted for each database are available at 10.6084/m9.figshare.20873434.v1. Searches were conducted on 01/02/2019 in: MEDLINE (National Library of Medicine), EMBASE (Excerpta Medica dataBASE) Global Health, CAB Abstracts, Scopus and Web of Science Core Collection. Searches were updated on 09/01/2021 using identical search strings. Results were limited to the English language and de-duplicated.

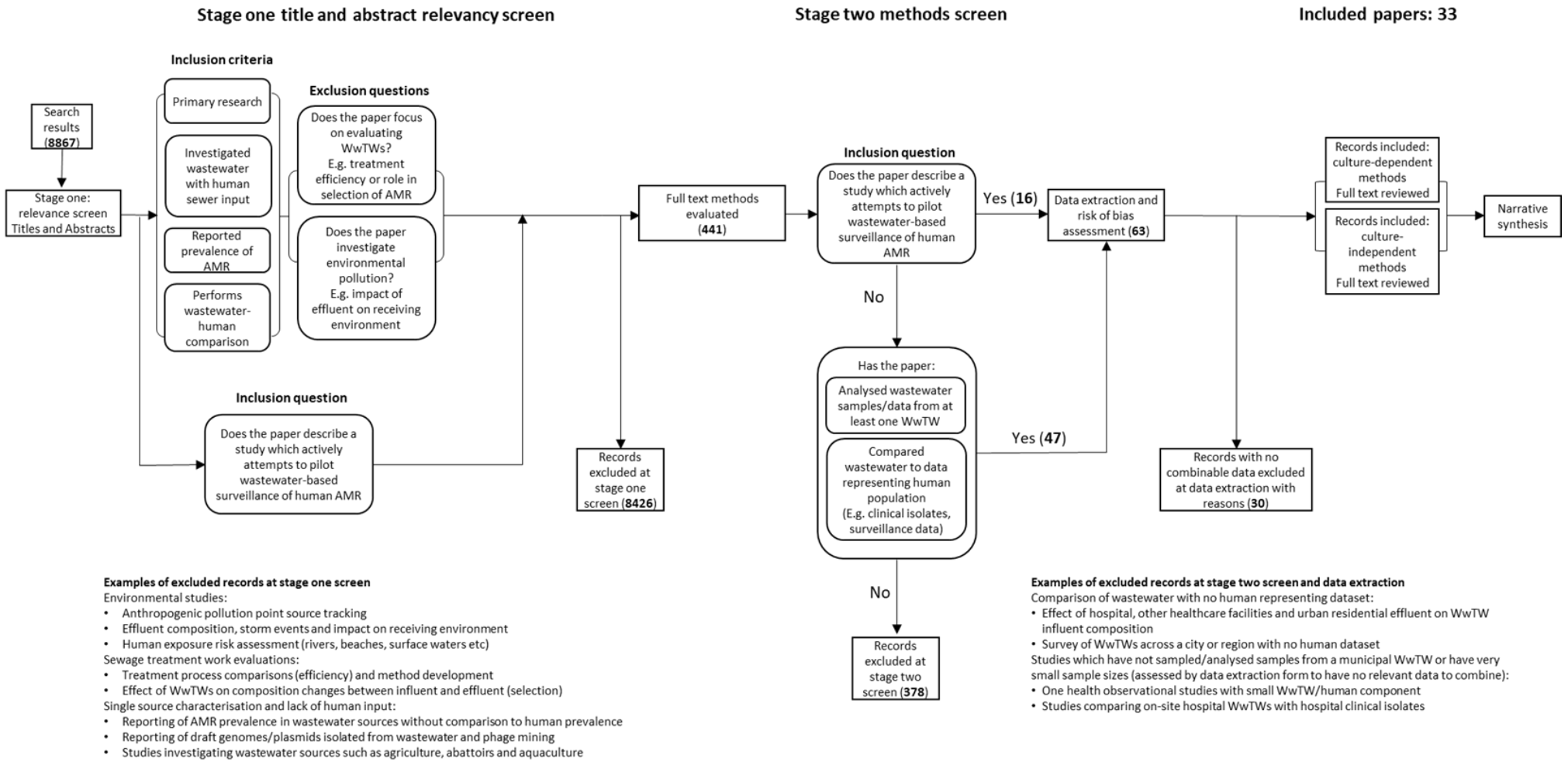
2.2.2 Eligibility criteria

Records were assessed through a two-stage screen (Figure 1) to capture both studies piloting wastewater-based AMR surveillance and studies conducting relevant wastewater-human AMR comparisons. Briefly, stage one screened titles/abstracts to

determine if the study was: (i) primary research conducted by the author/s, (ii) investigated wastewater which at least in part was constituted of human waste, (iii) reported AMR prevalence as a result of the work/analyses undertaken, and (iv) performed comparative analyses of the investigated wastewater to a separate non-wastewater compartment which potentially represents AMR prevalence in a human population. If it was unclear whether a study met criteria based on title and abstract alone, the study was passed onto the next stage. Stage two reviewed full-text methods and studies included if: (i) the wastewater analysed originated from at least one conventional WWTW where multiple waste streams converge, and (ii) the compartment being used as a comparator to wastewater AMR prevalence must directly represent a human population such as in clinical isolates or resistance network data. A universal inclusion question was used to include studies explicitly performing wastewater-based surveillance of human AMR irrespective of meeting stage one or two criteria.

Figure 1 - Screening flowchart

Screening criteria described for all stages with numbers of passing/excluded records. Examples of excluded records are described at each stage.



2.2.3 Study selection and data extraction

Records were independently screened in duplicate by a collaborating co-author and myself, and data was extracted from included records using a pre-tested data extraction form piloted on five random included records. This form consisted of both pre-determined fields using data validation and non-restricted write in fields to record data relating to (non-exhaustive): study design (wastewater sampling strategy, WWTW metadata, human sample type, sample sizes), methods (wastewater sampling, AMR detection, statistical methods) and outcomes (reported wastewater-human comparison results). The form can be found at 10.6084/m9.figshare.20873581.v1. If available, raw resistance prevalence data (total and resistant isolate counts regarding individual antibiotics or resistance genes) were extracted for antibiotics on the WHO critically important antimicrobials list[66]. These counts were used to calculate point estimates (±95% confidence intervals) representing the proportion of AMR isolates amongst all isolates tested for either wastewater or human compartments; this metric is referred to as either the wastewater or human AMR prevalence from here on.

2.2.4 Risk of bias and Certainty assessment

Risk of bias was assessed independently by a collaborating co-author and myself using a qualitative approach based on the Cochrane risk of bias tool[67]; my modified tool focused on systematic differences at the study level as outcomes reported were highly diverse. Modified interpretations of five bias domains are detailed in Table S1; a brief summary of my interpretation of each bias domain is provided here. Attrition bias referred to differences introduced by missing data (e.g. missing sampling timepoints in longitudinal

studies). Performance bias referred to differences introduced by use of different methods between or within sampling compartments (i.e. AMR detection methods). Reporting bias referred to selective reporting where outcomes were measured but not reported or disproportionately reported in the text. Selection bias referred to differences introduced across the wastewater compartment by use of different wastewater sampling or processing methods (e.g. selecting more colonies for susceptibility testing from one WWTW compared to another). Other bias referred to a lack of acknowledgement of AMR-influencing wastewater inputs (i.e. unreported sewer inputs from healthcare, abattoir or agricultural sources), which may have modulated AMR profiles in sampled wastewater. If there was insufficient information present to assess the risk of bias this was denoted as "unclear". The rationale for risk of bias assessments was recorded, including examples where applicable. Discrepancies in risk of bias assessments were resolved by discussion. An overall qualitative measure (high, low and unclear) was assigned to each study as per the Cochrane risk of bias tool approach to summary assessment[67].

Certainty assessment (assessment of overall confidence in the evidence included) was conducted using an adaptation[68] of the GRADE (Grading of Recommendations, Assessment, Development and Evaluations) system designed for clinical studies[69], where evidence is given an initial rating ("high", "moderate", "low", "very low") then upgraded or downgraded based on study characteristics. The outcome being evaluated was whether or not high concordance between AMR prevalence estimates in wastewater and human compartments was observed. Conventionally, GRADE assigns "high" and "low"

initial ratings to randomised trials and observational studies respectively. A randomised controlled trial design is of limited applicability to the studies being evaluated as part of this review, and therefore using the adapted version of GRADE[68], I assigned initial "moderate" ratings to my body of evidence. Upgrading or downgrading was based on a subset of adapted GRADE domains (Table S2). Briefly, these included my risk of bias summary assessment, inconsistency, indirectness, imprecision, publication bias and concordance (all defined in Table S2). In my adaptations I omitted two of the original GRADE criteria ("large effects", "residual confounding"), as these could not be readily adapted to my context.

2.2.5 Data synthesis and analysis

For extracted resistance prevalence data, I used Lin's concordance correlation coefficient (CCC – R package DescTools) with 95% confidence intervals (CIs) to quantify the concordance between the proportion of resistant wastewater isolates (i.e. wastewater AMR prevalence) and the proportion of resistant human isolates (i.e. human AMR prevalence), with the latter representing the reference standard. As perfect concordance is unrealistic, I arbitrarily defined "high concordance" to represent a ±10% difference in AMR prevalence between wastewater and human compartments. Concordance was considered separately for comparisons derived from phenotypic versus genotypic approaches. In addition to evaluating concordance across all studies, I also evaluated concordance stratified by bacterial species, resistance to specific antibiotic classes and for AMR gene families. Since Lin's CCC does not reflect error in AMR prevalence estimates, I

also compared Clopper-Pearson 95% confidence intervals (CIs) stratified by study, antibiotic class or AMR gene.

As I aimed to identify approaches that could optimise wastewater-based AMR surveillance, I classified studies based on wastewater-human AMR concordance where a "high agreement" study was defined as a study with >70% of its wastewater and human AMR prevalence estimate comparisons being within ±10% of each other (i.e. highly concordant comparisons, see above). I then used logistic regression to identify if any study features were associated with this high agreement classification in STATA/IC v.16.1 (StataCorp, College Station, USA). As study feature reporting was highly inconsistent, I only tested features that were reported by at least 75% of studies.

In addition, given the heterogeneity of study features, their inconsistent reporting across studies, and the small number of studies limiting power to detect associations, I descriptively synthesised features potentially associated with "high agreement" studies (i.e. where >70% of wastewater and human AMR prevalence estimate comparisons were within ±10% of each other). For this, in addition to the high agreement category, I further classified studies as moderate agreement (30-70% of wastewater and human AMR prevalence estimate comparisons being within ±10% of each other) and low agreement studies (<30% of wastewater and human AMR prevalence estimate comparisons being within ±10% of each other). Agreement classifications were considered separately for comparisons derived from phenotypic versus genotypic approaches.

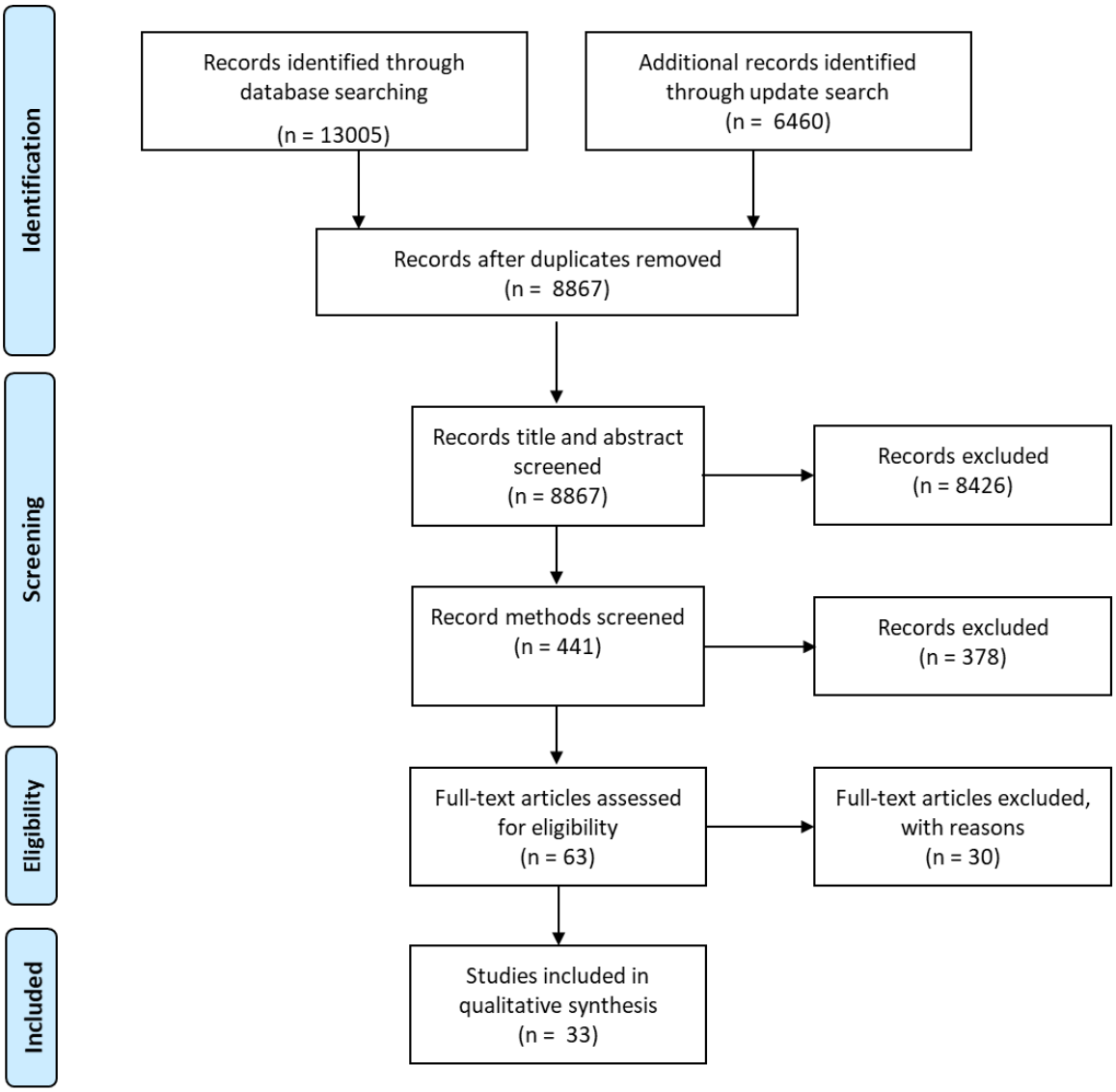
2.3 Results

2.3.1 Study selection, risk of bias and characteristics

2.3.1.1 Literature screen

Of 8,867 de-duplicated studies identified using my search strategy, full-text methods for 441 relevant studies were reviewed, and based on pre-specified inclusion criteria (see Methods section 2.2.2), 33 studies were included in the review (Figure 2).

Figure 2 - PRISMA flowchart of search strategy and study inclusion/exclusions



2.3.1.2 Risk of bias and certainty assessment

Based on my modified bias domains (Table S1), 19/33 studies were judged to have overall high-risk of bias, 7/33 with an unclear-risk and 7/33 with low-risk – full results available at 10.6084/m9.figshare.20874265.v1. To avoid splitting my analyses and losing statistical power, I present a synthesis of all studies and provide a summary of the risk of bias across studies below (as recommended by Higgins et al., 2011[67] when the study pool is small and stratified analyses according to risk categories is not feasible).

Selection bias was present in eight studies[43,49,51,54,70-73] where 7/8 sampled multiple wastewater sources using different sampling methods. Performance bias was not present in any studies, with laboratory methods consistent within studies. Attrition bias occurred in three studies as missing pre-specified sampling timepoints due to logistical factors (e.g. WWTW closure, problems with sampling permissions[50], or missing datapoints[74,75]. Reporting bias was present in three studies where sample sizes or antibiotic phenotyping data were incompletely reported for one or more group[73,76,77]. Other bias was evident in six studies[57,70,78-81] which did not report any information regarding sewer inputs to sampled wastewater.

Based on my certainty assessment, I rated the overall quality of included bodies of evidence as “low to moderate” regarding the outcome of identifying concordance between wastewater and human AMR prevalence estimates (Table S3). Initial “moderate” ratings were downgraded due to high overall risk of bias and the indirectness of study aims in the context of my review questions. Upgrades were warranted by

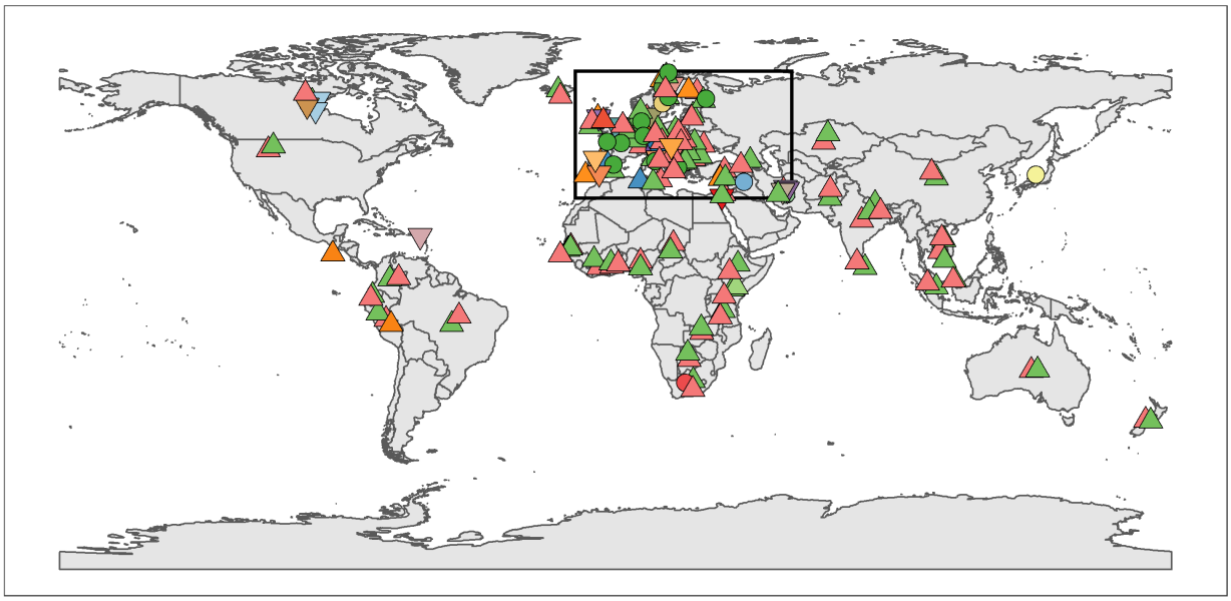
statistically significant concordance (adapted GRADE “dose-response” assessment) across most studies resulting in the final “low to moderate” rating of the evidence base on which the following data synthesis and recommendations are made.

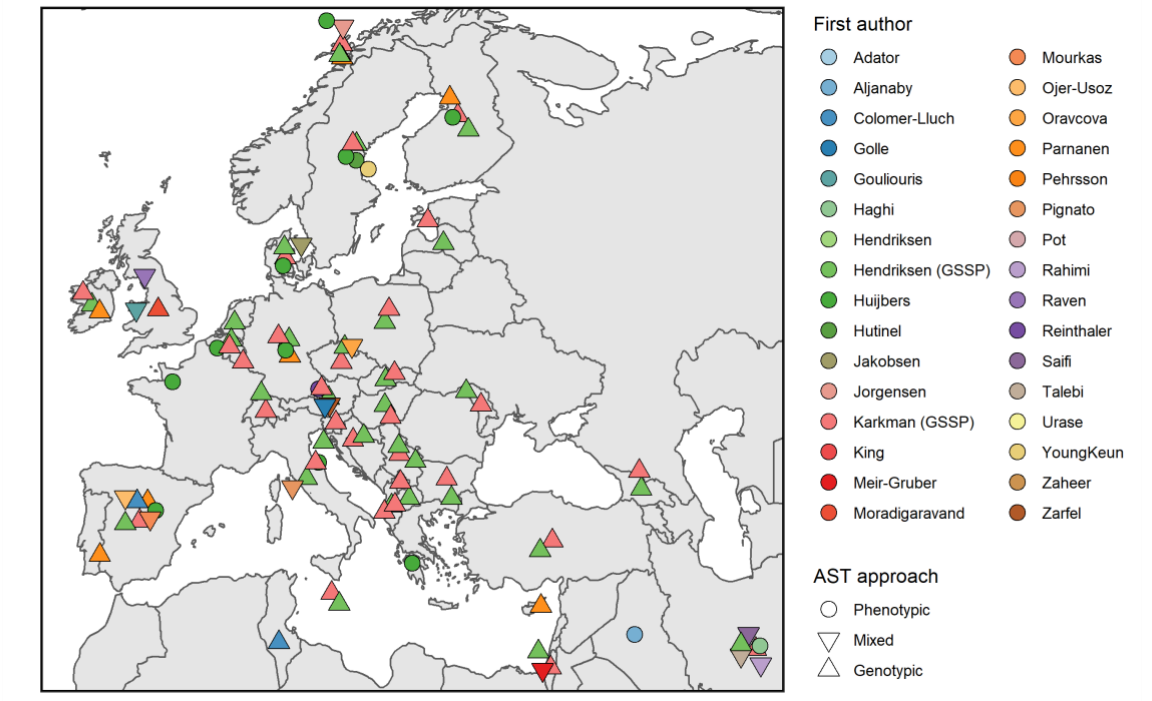
2.3.1.3 Summary of general study characteristics

Studies explicitly using wastewater for human population-level AMR surveillance made up 12/33 included studies. The remaining 21 studies included relevant comparisons between wastewater and human AMR, but were not directly set up as wastewater-based AMR surveillance studies (e.g. studies focussed on One Health). Amongst the 33 included studies, 73 unique countries were sampled, although most (48/73) were represented as part of a single global study[43] (Figure 3). World Bank regions covered by the studies were as follows: East Asia and Pacific (n=3 studies), Europe and Central Asia (n=19), Latin America and the Caribbean (n=4), Middle East and North Africa (n=9), North America (n=5), South Asia (n=2), Sub-Saharan Africa (n=4). World Bank income classifications showed a sampling skew towards high-income countries (high income [n=24 studies]; middle income [n=11] and low income [n=3]) (Figure S1). Three studies covered multiple regions and income classifications. Publication dates ranged from 2007-2020, with most published in the last three years (n=24) Full study descriptions are presented in Table S4 and Table S5

Figure 3 - Geographic distribution of wastewater sampling and test approach of included studies

Centroids of countries sampled by included studies are plotted with colours and shapes according to citation and antimicrobial susceptibility test (AST) approach respectively. Centroid are plotted with jitter to avoid overplotting and do not represent exact sampling locations within countries.





2.3.1.4 Summary of AMR detection methods in included studies

Evaluations of AMR were undertaken using genotypic-only methods (n=7), phenotypic-only methods (n=8), or a mixed approach combining both (n=18). Genotypic-only studies employed metagenomics (n=4), qPCR (n=2) and single isolate whole genome sequencing (WGS) (n=1). Phenotypic-only studies employed disk-diffusion (n=4), microbroth dilution (n=3), or both (n=1). Mixed approach studies combined disk-diffusion/microbroth dilution with qPCR (n=1), PCR (n=9) or single isolate WGS (n=8). For data synthesis and analysis across studies, relevant phenotypic data was extracted from 22 studies and genotypic data from 12 studies (24 studies in total as both data types could be extracted from 10 studies). These 24 studies conducted phenotypic-only (n=7), genotypic-only (n=1) or combined (n=16) AMR detection. All extracted genotypic data consisted of isolate-level qPCR, PCR or WGS; no metagenomic data at the sample-level was synthesised. For nine studies there were no relevant data that could be extracted for inclusion in a combined summary; these were either where isolate counts were not reported, or only raw sequencing data was available which was beyond the scope of my analysis.

2.3.2 Wastewater-human AMR concordance

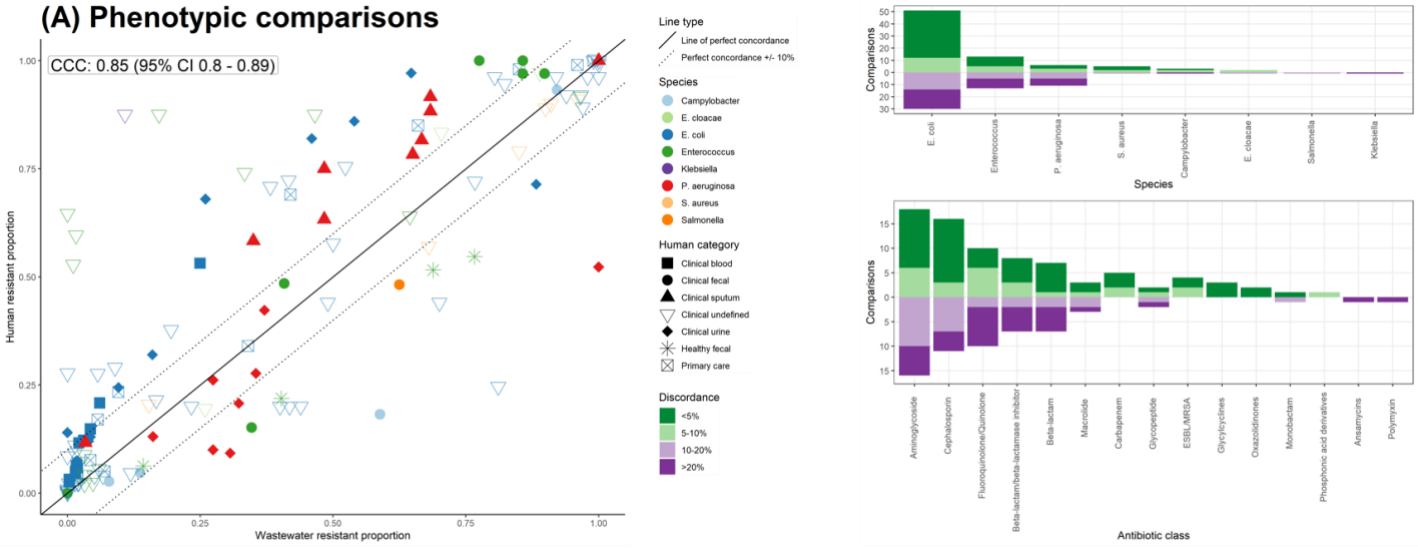
2.3.2.1 Phenotypic wastewater-human AMR concordance

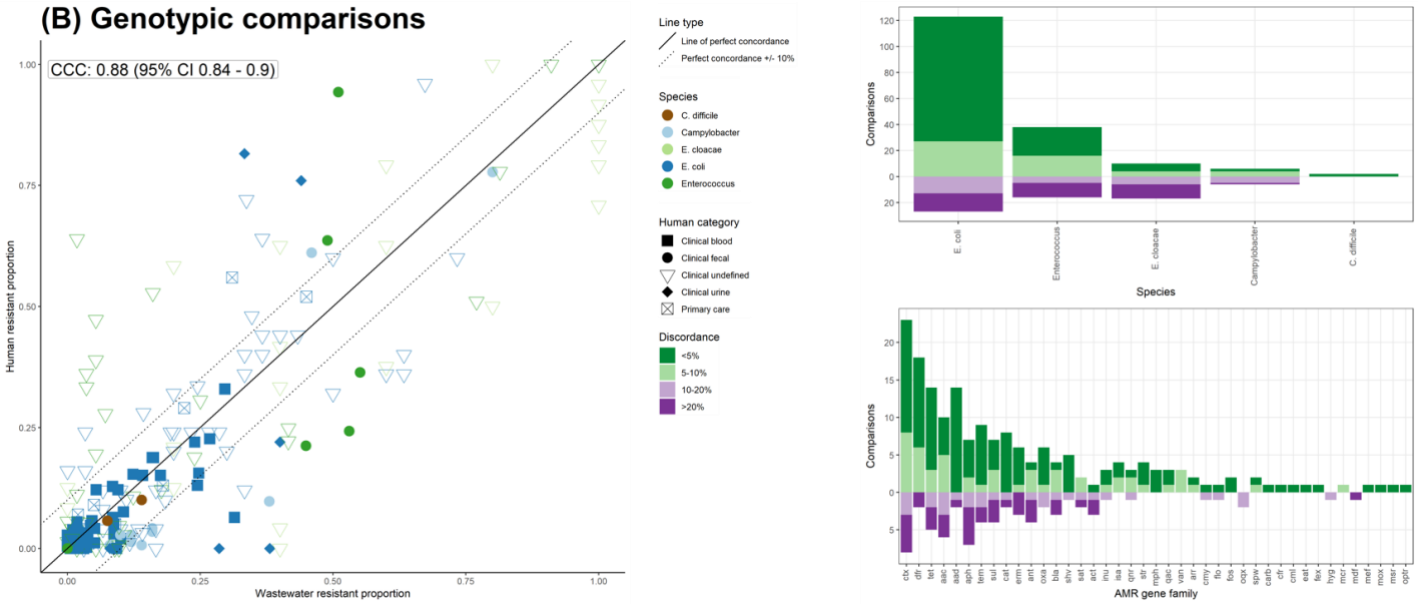
Phenotypic data from 22 studies covered fifteen WHO Critically Important Antimicrobial classes and comprised 139 comparisons for which overall concordance between resistance prevalence estimates in wastewater and human compartments was reasonably high (CCC=0.85 [95% CI 0.8-0.89]) (Figure 4A). The median number of comparisons (i.e. AMR prevalence for a specific species-drug across both wastewater and

human compartments) per study was 6 (IQR: 3-9). For any comparison, the median number of isolates analysed in humans was 130 (IQR: 50-232), and in wastewater 98 (IQR: 58-186). AMR prevalence estimates were highly concordant (defined as wastewater AMR prevalence within ±10% of human AMR prevalence) for 80/139 (58%) comparisons.

Figure 4 - AMR in wastewater isolates and human isolates for phenotypic (A) and genotypic (B) comparisons.

Left: Concordance plot of AMR prevalence in wastewater and human isolates stratified by AMR detection approach (i.e. phenotypic (A) versus genotypic (B) approaches). Each point represents a single wastewater-human comparison conducted where columns corresponds to bacterial species tested and shape corresponds to human sample type used. Lin's concordance correlation coefficient (CCC) is labelled with 95% confidence intervals. Unbroken line of *y=x* is plotted as perfect concordance between wastewater and human resistance. Dashed lines of *y=x*+0.1 and *y*=x-0.1 represent high concordance, i.e. ±10% from perfect concordance respectively. Right: Individual wastewater-human comparisons tallied by level of discordance (<5% and 5-10% coloured in green, 15-20% and >20% coloured in purple) between compared wastewater and human AMR prevalence estimates, and plotted to show number of comparisons at each level of discordance, stratified by the target species and antibiotic class (A-right) or AMR gene family (B-right).





Hutinel et al., 2019[82] and Ojer-Usoz et al., 2017[80] contributed the most highly concordant comparisons (13/16 [81%] and 14/22 [64%] within-study comparisons respectively); however only the Hutinel study was therefore defined as a "high agreement" study. Most phenotypic comparisons utilised isolates cultured from human samples originating from healthcare settings (135/139 comparisons), with higher AMR prevalence than corresponding wastewater estimates (in 86/135 (64%) comparisons). The remaining four comparisons (all from Haghi et al. 2019[70]) were unique in comparing faecal isolates from healthy volunteers with no recent history of antibiotic use to wastewater; all four showed reduced AMR prevalence in the human isolates. Sensitivity analysis using data from low bias studies only (n=5; 31 comparisons) showed a slight decrease in overall concordance (CCC=0.81 [95% CI 0.65-0.9]; 95% CI overlaps with that of the full phenotypic dataset described above) (Figure S2).

The most common species and antibiotic class investigated using phenotypic susceptibilities were *E. coli* and aminoglycosides respectively. When considering discordance between wastewater and human AMR prevalence estimates across the top two represented species, 38/81 (47%) comparisons for *E. coli* showed <5% discordance and 51/81 (63%) of comparisons ≤10% discordance, whereas 8/26 (31%) of comparisons across compartments in *Enterococcus* spp. showed <5% discordance and 13/26 (50%) ≤10% discordance (Figure 4A, top right panel). Differences between the top antibiotic classes evaluated were similarly seen; those investigating aminoglycoside resistance showed <5% discordance for 11/34 (32%) and ≤10% discordance for 18/34 (53%) of

comparisons, while cephalosporin resistance comparisons showed <5% discordance for 13/27 (48%) and ≤10% discordance for 16/27 (59%) (Figure 4A, bottom right panel).

By individual WHO critically important antimicrobial, AMR prevalence estimates (i.e. point estimate [[95% confidence intervals]) in human and wastewater isolates overlapped for: (i) aminoglycosides (26/34 comparisons; Figure S3 (ii) beta-lactams (10/14 comparisons; Figure S4); (iii) beta-lactam/beta-lactamase inhibitor combinations (8/15 comparisons; Figure S5); (iv) cephalosporins (18/26 comparisons; Figure S6); (v) carbapenems (5/5 comparison; Figure S7); (vi) fluoroquinolones/quinolones (7/20 comparisons; Figure S8); (vii) glycopeptides (3/4 comparisons; Figure S9); and (viii) erythromycin (5/6 comparisons; Figure S10). Overlap was also seen specifically for extended-spectrum beta-lactamase (ESBL)-producing isolates (2/3 comparisons) and methicillin-resistant *Staphylococcus aureus* (MRSA) (1/1 comparison; both Figure S11). Non-overlapping estimates were mostly associated with five studies[51,81,83-85].

2.3.2.2 Genotypic wastewater-human AMR concordance

Extracted genotypic data (single isolate WGS and PCR) from 12 studies comprised 245 comparisons between AMR gene prevalence estimates between wastewater and human compartments. Correlation between compartments was slightly higher than for phenotypic data (CCC=0.88 (95% CI 0.84-0.9)) and overall spread away from perfect concordance was reduced (Figure 4B), with high concordance (wastewater AMR prevalence within ±10% of human AMR prevalence) in 179/245 (73%) comparisons. The median number of comparisons (i.e. AMR prevalence for a specific species-AMR gene

combination across both wastewater and human compartments) per study was 11 (IQR: 5-23). For any comparison, the median number of isolates analysed in humans was 94 (IQR: 25-437), and in wastewater 56 (IQR: 30-388).

Using genotypic analyses, most high concordance comparisons were observed in Raven et al., 2019[86] (71/73), followed by Adator et al., 2020[87] (33/45) and Zaheer et al., 2020[84] (19/22). All genotypic comparisons evaluated isolates cultured from human samples originating from healthcare settings, with higher AMR prevalence than the corresponding wastewater estimate in 102/245 (42%). Sensitivity analysis of wastewater-human genotypic concordance using data from low bias studies only (n=4; 154 comparisons) showed a slight increase (CCC=0.9 [95% CI 0.87-0.93]; confidence interval overlapping the CCC estimate [95% CI] derived from the full genotypic dataset) (Figure S2).

For genotypic comparisons, the most common species and AMR gene family investigated were *E. coli* and CTX-M β-lactamases respectively. When considering discordance across species, 95/150 (63%) comparisons for *E. coli* showed <5% discordance and 123/145 (85%) of comparisons ≤10% discordance, while comparisons in *Enterococcus* spp. showed 32/54 (59%) <5% discordance and 38/54 (70%) ≤10% discordance. *Enterobacter cloacae and Campylobacter* spp. comparisons shared more even distributions of discordance levels (Figure 4B top right panel). For AMR gene family, comparisons of most gene families across compartments were in high concordance (i.e. ≤10% discordance) such as for CTX-M β-lactamases 23/31 (74%), *dfr* 18/20 (90%) and *aad* 14/16 (88%) (Figure 4B bottom right panel).

For a subset of eight genes conferring resistance to WHO critically important antimicrobial and investigated by multiple studies, 95% CIs around prevalence estimates in both compartments overlapped as follows: (i) *aac* (13/16 (81%) comparisons; Figure S12 [displayed for individual allelic variants]); (ii) *arr* (1/2 comparisons; Figure S13); (iii) CTX-M 25/29 comparisons; Figure S14); (iv) *erm* (6/9 comparisons; Figure S15); (v) *fos* (2/2 comparisons; Figure S16); (vi) *oxa* (8/8 comparisons; Figure S17); (vii) *van* (3/3 comparisons; Figure S18) and (viii) *qnr* (3/4 comparisons; Figure S19).

2.3.3 Synthesis of study features

2.3.3.1 Logistic regression of study features and wastewater-human AMR agreement

If more than 70% of wastewater-human AMR prevalence estimate comparisons conducted by a study were highly concordant (i.e. prevalence estimates within ±10% of one another), studies were classed as high agreement overall; this was the case for 6/22 studies (27%) with phenotypic data and 5/12 (42%) studies with genotypic data. Of the ten studies with both phenotypic and genotypic data available only 1/10 showed high agreement for both approaches[87]; the remaining 9/10 either showed high agreement for only one approach (3/10) or neither approach (6/10). No statistically significant associations between study features and higher wastewater-human AMR concordance were identified (Table S6). The limited number of eligible studies, the substantial heterogeneity of combinations of approaches deployed across studies, and missingness of data meant power to detect independent associations was low (Table S7).

2.3.3.2 Study features descriptive synthesis

I therefore synthesised study features descriptively, additionally assigning studies to moderate (30-70% of wastewater and human AMR prevalence estimate comparisons being within ±10% of each other) and low agreement categories (<30% of wastewater and human AMR prevalence estimate comparisons being within ±10% of each other), and treating phenotypic- and genotypic-approaches as separate study subsets (Table S8). Sampling of influent, either alone or in conjunction with effluent, appeared most consistently associated with moderate-/high-agreement (15/28; 54% where reported) in estimates of AMR prevalence between wastewater and human compartments. Longitudinal sampling was the most common study design (29/32; 91% where reported) with moderate-/high-agreement in most (22/29; 76%). The four studies undertaking snapshot (i.e. single-timepoint) sampling all had moderate-/high-agreement whereas the single study conducting a mixed sampling design was associated with low agreement. For longitudinal studies, the timeframe of sampling was potentially relevant: Of the eight low-agreement longitudinal studies, 7/8 sampled for ≤12 months and only 1/8 for >12 months. Conversely, 14/20 medium/high-agreement longitudinal studies sampled for >12 months, and only 2/20 for <6 months. Studies deployed several wastewater sampling methods, with grab and flow-proportional sampling studies equally distributed across agreement categories (8/12 [67%] and 4/6 [67%] moderate-/high-agreement respectively) but interestingly, composite sampling was more associated with moderate-/high-agreement (8/9; 89%). Of note, sampling point or method was not reported by six and five studies respectively. Most studies performed comparisons on wastewater at least in part derived from the human population sampled (i.e. direct comparisons, 18/34; 53%), while eight conducted indirect comparisons, one conducted both and seven were

unclear/unreported. Most moderate-/high-agreement studies conducted direct comparisons (14/21; 67% where reported). Lastly, studies investigating 1-2 WWTWs made up the majority (22/34) but were similarly associated to moderate-/high-agreement (17/22; 77%) as those investigating ≥3 sites (7/10; 70%).

2.3.3.3 Studies without extractable data

Targeted qPCR-based studies

Two studies performed direct AMR gene detection using qPCR of either 229[50] or eight AMR genes[76]; both reported a relationship between wastewater AMR and national AMR data. Pärnänen et al. conducted a longitudinal survey of 13 European WWTWs across seven countries compared with contemporaneous surveillance data from EARS-Net. The relative abundance of influent AMR genes clustered significantly into two distinct country groups coinciding with high and low antibiotic consumption, with AMR gene distributions mirroring EARS-Net (i.e. human AMR surveillance)-described north-to-south and west-to-east geographic gradients. A higher relative abundance of most AMR gene classes was observed from high antibiotic consumption countries except for tetracycline and macrolide-lincosamide-streptogramin B AMR genes. The study also performed counts of antibiotic-resistant culturable bacteria on samples but showed no significant correlation between counts and AMR gene quantification.

Colomer-Lluch et al. compared wastewater samples in Tunisia and Spain, and their respective national AMR surveillance data. Intra-country comparisons showed the most

commonly recovered beta-lactamase gene (*bla*$_{TEM}$) from wastewater was also most commonly reported in clinical surveillance. This was also true of the quinolone resistance genes *qnr*S and *qnr*A but not for CTX-M-9 group ESBL genes, where a higher prevalence was identified in wastewater. Inter-country comparisons also highlighted higher AMR gene detection in samples from Spain coinciding with higher antibiotic use, and high *mec*A prevalence thought to be potentially consistent with high-pig farming densities.

Metagenomic-based studies

Four studies employed metagenomics to identify potential wastewater-human AMR associations. Two of these studies appeared to demonstrate an association while the other two were inconclusive. Hendriksen et al. conducted a large-scale cross-sectional survey of global wastewater and compared resistomes to multiple epidemiological variables[43]. This study showed total AMR gene abundances from a given site/country correlated with sanitation and general health metrics (with up to 89% of observed resistome variation significantly explained regardless of within-sample AMR gene diversity). A high human development index (i.e. generally high-income countries) was linked to significantly lower abundance of AMR genes in wastewater metagenomes.

Karkman et al. compared wastewater resistomes (generated by Hendriksen et al.) to local clinical surveillance data extracted from surveillance networks (EARS-Net, CAESAR and ResistanceMap)[49]. A beta regression model showed relative abundance of AMR genes could partially explain clinical resistance patterns but focussing on individual classes of

resistance only improved model fit for aminopenicillin resistance. The best models were produced by incorporating country-level socioeconomic factor data (GDP, basic sanitation and proportion of urban population), which improved model fit for all resistance classes except aminoglycoside resistance.

The first inconclusive metagenomics-based study conducted longitudinal surveillance comparing resistomes generated from drainage ditch wastewater and contemporaneous local surveillance data consisting of household morbidity and healthcare usage[88]. Significant increases/decreases in resistome read abundance were not associated with contemporaneous increases in reported diarrhoea or clinical faecal isolates. However, authors highlighted non-significant increases in pathogen read abundances appeared to coincide with increased reported illness/clinic visits, possibly reflecting a lack of sensitivity to detect significant changes in read abundances.

The second inconclusive study was a metagenomic and functional metagenomic longitudinal evaluation of wastewater and human faecal samples collected from individuals within the WWTW catchment[57]. Characterisation of resistomes, including the use of AMR gene networks, showed significantly higher phylogenetic species diversity and abundance of AMR proteins in wastewater (WWTW influent, street-access wastewater) compared to faecal samples. Faecal samples were significantly enriched for drug efflux mechanisms whereas wastewater samples contained more aminoglycoside acetyltransferase, class D β-lactamase, and dihydrofolate reductase genes. However,

extensive sharing of sulphonamide-resistance conferring AMR genes was detected between sample sets indicating possible resistance-specific relationship.

Mixed-approach and phenotypic-only studies

Two studies used mixed approaches combining phenotypic AST with qPCR (2-targets)[77] and single isolate WGS[89]; one study used a phenotypic approach only[73]. All three studies appeared to show a wastewater-human AMR association. Meir-Gruber et al. compared prevalence of bla_{KPC}/bla_{NDM-1}-carrying CREs, MRSA and VRE in WWTW influent and domestic sewer systems collected nationally over one year compared to contemporaneous clinical samples. The study found a limited observational association where wastewater and clinical isolates shared the same most commonly isolated bla_{KPC}-carrying species (*K. pneumoniae*) in addition to the other top three isolated bla_{KPC} species.

The second mixed approach study[89] compared wastewater *E. faecium* isolates from 20 WWTWs and isolates from clinical infection preceding and during study sourced from hospitals in the same region. They found wastewater and clinical isolates were phylogenetically interspersed, indicating mixing, with divergence events implying recent emergence and dissemination in the geographic region. Each WWTW sampled also contained genetically diverse populations, but these remained comparable to diversity in national bloodstream infection isolates. Network analysis showed geographic clustering of WWTWs and bloodstream isolates, and linked bloodstream isolates to 9/20 WWTWs (three without hospital sewer input).

Lastly the phenotypic-only study[73] compared prevalence of resistant *E. coli* isolated from WWTW influent, effluent and hospital effluent to national resistance prevalence of clinical blood isolates with a non-overlapping sampling period. The authors found a lower prevalence of resistance in WWTW isolates when compared to hospital effluent. An observational wastewater-human association was reported in the increasing resistance prevalence of WWTW isolates which may reflect increases in the clinical dataset.

2.4 Discussion

From my review and synthesis of the available data, I found characterisation of AMR in wastewater shows promise in reflecting AMR in human populations, irrespective of diverse target species, target resistances and study locations, although associations may be stronger for some species and AMR mechanisms than others, and may vary by setting and over time. The strength of this relationship varied across studies and was likely influenced by study features (e.g. design, setting, spatiotemporal sampling strategies) and AMR detection method (i.e. genotypic/phenotypic); the heterogeneity of methodological approaches and lack of clear reporting of key study features made any quantitative synthesis very difficult.

2.4.1 Effect of AMR detection method on wastewater-human AMR concordance

My estimates of concordance (Figure 4A, 4B) supporting a wastewater-human AMR correlation are in line with estimates in individual studies[49,82,83]. In particular, Huijbers et

al., 2020[83] reported coefficients of determination as 0.62-0.72 for individual antibiotics and 0.85 when data was combined for four antibiotic classes – similar to my findings of 0.85 and 0.88 for class-unrestricted phenotypic and genotypic data respectively. Although data was too limited to robustly estimate Lin's CCC for individual species and AMR, variability in the level of discordance of wastewater-human comparisons (Figure 4A, 4B – right panels) and overlap in 95% CI around point estimates (Figure S3-19) is likely attributable in part to specific species and AMR mechanism that are optimally suited to wastewater-based AMR surveillance. This phenomenon was also reported in several studies without extractable prevalence data where specific AMR classes/genes exhibited notably higher/lower wastewater-human concordance.

Genotypic AMR detection methods showed a slight performance increase (non-significant) for both Lin's CCC and in the reduced proportion of strongly discordant comparisons (Figure 4B right panels) which may reflect their relatively species- and mechanism-agnostic nature (mostly WGS in extracted data) over phenotypic methods which may be more susceptible to variations from differing growth media/conditions and interpretation of resistance breakpoints. Problems with accurately characterising AMR prevalence when only small numbers of isolates are analysed (median of 94-130 for human and 91-98 for wastewater compartments across pheno-/genotypic comparisons here) is also a concern highlighted by previous researchers, particularly when few resistant isolates are available[83].

Although not a focus in my review, genotypic profiling potentially affords some additional advantages over phenotypic analyses, and is relevant to confirming that genetic mechanisms underpinning phenotypes are also similar. Genomic approaches, such as sequencing of isolates or whole sample metagenomics, enable a more agnostic approach to be adopted than for qPCR, in that analyses do not need to be restricted to a subset of predefined genes/gene variants. Genomic data and profiles can also be more readily shared with the wider community to allow for cross-study comparisons and data synthesis; as demonstrated by Karkman et al[49]. Genomics approaches also allow for the evaluation of genetic relatedness and quantitation of either isolates or microbial populations across compartments (e.g. through phylogenetics, taxonomic/strain-level profiling and strain-based comparisons using metagenomes). Genomic approaches may be higher-resolution and more flexible, but at a higher resource cost; sensitivity for the detection of AMR genes is also dependent on sequencing depth, and accurately associating specific AMR gene markers with strains or species in short-read based metagenomes remains difficult[90].

2.4.2 Effect of study features on wastewater-human AMR concordance

My review highlights the need for clear guidance on performing and reporting these studies in a more standardised way, with a view to consolidating best-practice approaches in a workflow whilst enabling some flexibility to account for differences in any given setting. Although logistic regression found no significant associations, descriptive synthesis identified study features potentially associated with higher

wastewater-human AMR concordance and these are discussed in the context of existing literature here.

WWTW influent is likely the most population-representative wastewater sample for AMR surveillance using either phenotypic or genotypic approaches. This is not unexpected, as previous studies have described transformation of microbial and AMR gene composition during treatment[52,53]. Transformed samples may remain useful for wastewater-based AMR surveillance, potentially dependent on treatment process; however differing levels of treatment have been shown to select for different species/AMR determinants[52]. Additionally, a temporospatial overlap of wastewater sampled and the target surveillance population is likely helpful; a feature of the majority of moderate-/high-agreement studies as well as sampling fewer WWTWs which may also relate to the closeness of the populations compared.

Composite sampling also seems sensible as wastewater composition changes significantly over short timescales[55] and individual grab samples may be "flooded" by homogenous solid material[51]. However, grab sampling is convenient and avoids significant autosampler-associated workload and capital costs, and was the most common sampling method used in the studies analysed. Further research is needed into characterising how effectively single timepoint grab samples versus composite/proportional samples reflect temporal AMR changes.

Longitudinal sampling with timeframes over 12 months was most common in both phenotypic and genotypic high agreement studies, consistent with data from two studies which could not be directly synthesised[72,88]. In these studies, both used two-weekly sampling intervals over 12- and 3-month timeframes respectively. The former observed an association between ampicillin-resistant wastewater and contemporaneous clinical isolates, whereas the latter found no relationship between contemporaneous public health surveillance and wastewater metagenomic read abundances.

2.4.3 Recommendations regarding risk of bias

Only seven included studies were judged as low risk based on my risk of bias assessment; however a sensitivity analysis of wastewater-human AMR prevalence concordance in these data (Figure S2) showed similar results to the main analysis of all studies. Although high risk of bias does not necessarily mean that the data cannot be considered in a data synthesis, and my concordance sensitivity analysis appears to show this for included studies, minimising the risk of bias is key to producing robust data for future evaluations of wastewater-based AMR surveillance.

Several potential biases may be challenging to manage in the context of wastewater sampling, for example those that arise from logistical reasons concerning inaccessibility of sampling sites or sampling equipment. While unexpected wastewater site closures cannot be anticipated[50] and achieving exactly the same sampling methods across highly variable sampling sites is challenging in large-scale collaborations[43], future studies should attempt to minimise any differences across wastewater sampling sites where feasible.

For example, Huijbers et al.[83] consistently sampled wastewater sites during the mid-week to avoid potential weekend effects known to affect chemical wastewater-based epidemiology, and Urase et al.[75] purposefully avoided sampling combined sewer wastewater sites during rain events to avoid potential dilution and flow variation. Additional metadata such as sample storage conditions or freeze/thaw cycles are also pertinent to interpretation as investigated by Poulsen et al.[60] who similarly suggested detailed reporting of these features.

Future studies should clearly report sewer inputs, including any unique AMR-associated inputs (e.g. hospitals, agricultural sources) that may confound AMR prevalence estimates[38]. The importance of specific AMR-associated inputs is likely linked to whether the AMR mechanisms under evaluation are uniquely associated with that specific source, or already widely disseminated in the community. For example, one study[86] sampled WWTWs with and without hospital input, and found the most clinically-prevalent *E. coli* ESBL gene was ubiquitous in all WWTWs, indicating prior widespread dissemination in the community. Another study[54] focussing on *E. coli* gentamicin resistance in hospital effluent, receiving WWTW influent and domestic-only wastewater, found significantly lower prevalence in domestic-only wastewater compared to hospital effluent and WWTW influent which shared similar prevalence, indicating that the presence of any hospital-associated wastewater in influent was not representative of community-based estimates.

2.4.4 Limitations

My study has several limitations. Although I conducted a comprehensive risk of bias assessment, the study pool was too small to feasibly conduct risk-stratified synthesis and meta-regression without substantial loss of power[67]. To mitigate this, a modified version of the GRADE system was used to incorporate summary assessments of the quality of the evidence into the interpretation of results[67]. Certainty of the evidence base was rated as "low to moderate" which potentially indicates reduced confidence in my conclusions/recommendations, however, my certainty assessment is a conservative estimate omitting two upgrade domains, and true certainty may be higher if data was available to assess these domains. Since most human datasets available in the included studies were clinical in origin, human compartment AMR prevalence estimates were potentially susceptible to biases outlined in my introduction (i.e. overestimation of the "true" population-level AMR burden). However, as seen in existing literature and in the results of this review, clinical and wastewater AMR prevalence estimates do appear to mirror each other. I excluded non-English publications, potentially missing some relevant studies. Studies were highly diverse in reported features, design and outcomes, making a comprehensive synthesis difficult. In particular, many features were poorly characterised and could not be explored in my analyses. For my study feature analysis I focused *a priori* on features that optimised the identification of an association between wastewater and human AMR prevalence, however it may be that in some circumstances there is genuinely no such association. In many settings globally, established wastewater infrastructures are not available, and an analysis of, for example, WWTW influent may not be feasible; open sewerage systems may represent an alternative sampling point in these contexts.

In conclusion my review suggests that overall, wastewater-based surveillance has significant potential for monitoring population-level AMR, particularly for some species, despite high diversity in study design, methods and metadata. I found that no specific study feature or AMR detection method conferred a clear increase in the ability of a study to detect an association between wastewater and human AMR prevalence. However, based on limited available data, I would recommend that where feasible, genotypic AMR detection, composite sampling of influent with longitudinal timeframe >12 months, and contemporaneous sampling of wastewater and human compartments that are directly associated (i.e. the human population sampled contributes to the wastewater sampled) are used to generate more robust data to better evaluate the strengths and limitations of this approach for surveillance purposes. Clear reporting of study methods and features are essential, and this will facilitate the iterative development of optimal practice guidelines for this emerging surveillance tool.

2.5 Supplementary Material

Table S1 - Modifications of Cochrane risk of bias tool used for risk of bias assessment

Bias domain	**Cochrane tool source of bias**	**Cochrane tool bias description**	**Modified tool source of bias**	**Modified tool bias description**
Selection bias	Random sequence generation Allocation concealment	Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence Selection bias (biased allocation to interventions) due to inadequate concealment of allocations before assignment	Selective sampling of wastewater sites	Selection bias introduced by use of significantly different wastewater sampling method/strategy for specific sampling locations (e.g. WWTWs receiving hospital waste vs domestic waste only) or points (e.g. influent vs effluent)
Performance bias	Blinding of participants and personnel	Performance bias due to knowledge of the allocated interventions by participants and personnel during the study	Use of non-uniform performing methods during outcome assessment	Performance bias introduced between and within sample groups by the use of different processing/AMR detection approaches
Attrition bias	Incomplete outcome data	Attrition bias due to amount, nature, or handling of incomplete outcome data	Incomplete outcome data	Attrition bias introduced by incomplete outcome data where data fields may be missing across comparison groups (e.g. longitudinal studies in particular may be missing timepoints due to logistical reasons)
Reporting bias	Selective reporting	Reporting bias due to selective outcome reporting	Selective reporting	Reporting bias due to selective outcome reporting where outcomes may be measured but not reported or disproportionately reported compared to other outcomes
Other bias	Anything else, ideally prespecified	Bias due to problems not covered elsewhere	Uncertainty of sewer inputs	Confounding bias referring to the uncertain presence of AMR/AMR-influencing sources which could have impacted study outcomes

Bias domain	Cochrane tool source of bias	Cochrane tool bias description	Modified tool source of bias	Modified tool bias description
Detection bias	Blinding of outcome assessment	Detection bias due to knowledge of the allocated interventions by outcome assessment	Not applicable as knowledge of allocated groups is essential for specific sample processing	NA

Table S2 - Adapted GRADE domains used for certainty assessment

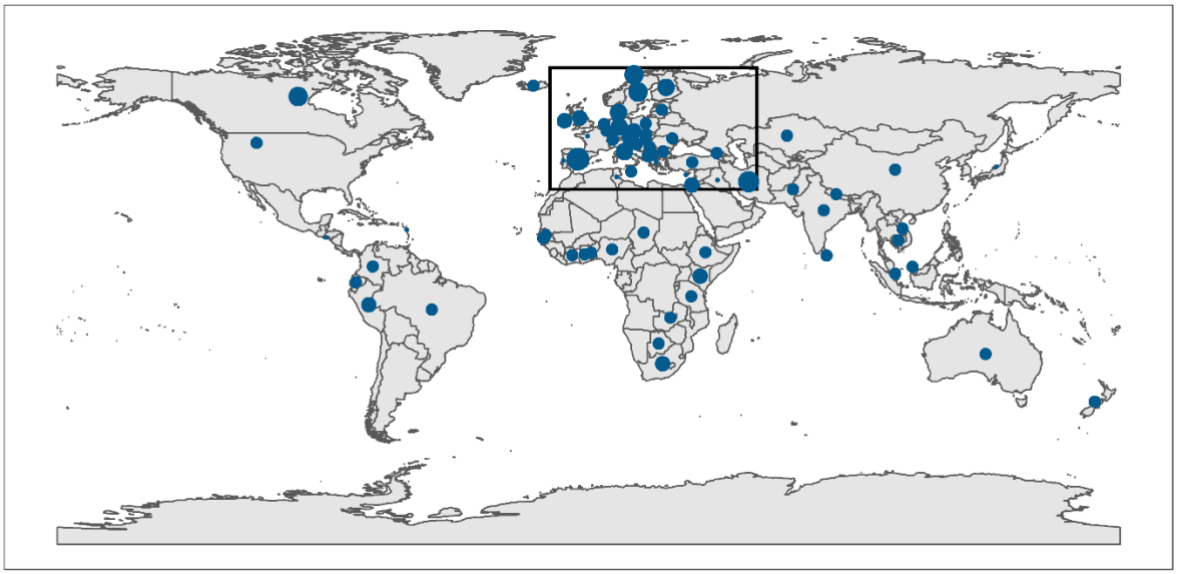
	GRADE domain	Summary of criteria	Example for further clarification where required
Domains considered for downgrading	Risk of bias	Summary assessment of study limitations consisting of attrition-, performance-, reporting-, selection- and other-bias.	Examples explained in the main text
	Inconsistency	High variability in the direction of effect.	The body of evidence contains studies reporting both positive and negative correlation between wastewater and human AMR prevalence
	Indirectness	Evidence is not directly related to the systematic review question.	Studies not specifically set up to pilot population-level surveillance using wastewater
	Imprecision	Relating to the overall number of isolates tested in a body of evidence.	
	Publication bias	Incomplete body of evidence resulting from inadequate search strategy.	
Domains considered for upgrading	Dose response (adapted for this study to "concordance")	Effect size mirrored across compartments	AMR prevalence estimates in the human compartment is associated with AMR prevalence in the wastewater compartment (i.e. High AMR prevalence estimate in wastewater concords with high AMR prevalence estimate in human compartment).
	Large effect	Not applicable to our systematic review	
	Confounding minimizes effect	Not applicable to our systematic review	

Table S3 - Certainty assessment results and rationale for upgrades and downgrades

Identification of association between AMR prevalence in humans and wastewater									
Initial confidence (no. of records)	**Domains decreasing confidence** **"---" if not present; "↓" if enough evidence to downgrade**					**Domains increasing confidence** **"---" if not present; "↑" if enough evidence to upgrade**			**Final confidence rating**
	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication bias	"Concordance"	Large effect	Residual confounding	
Surveillance studies									
Initial moderate (12 records)	↓	---	---	---	---	↑	NA	NA	Moderate
Relevant comparison studies									
Initial moderate (21 records)	↓	---	↓	---	---	↑	NA	NA	Low
Rationale									
	Majority of records judged to have overall high risk of bias across both bodies of evidence (8/12, 11/21).	Judging magnitude of effect consistency is not feasible due to heterogenous outcomes but direction of effects is mostly consistent (i.e. a wastewater-human association is observed).	Research evidence for surveillance studies is not dissimilar to our review questions but majority relevant comparison studies aim to answer dissimilar questions.	High total number of participants (i.e. isolates tested) across all studies in both bodies of evidence.	It was not feasible to assess publication bias via funnel plot in our bodies of evidence, but bias is less likely when a comprehensive search strategy like ours is used.	Majority of records conducting statistical testing show a statistically significant association between human and wastewater AMR prevalence across both bodies of evidence.	Not applicable to our bodies of evidence or review question so not assessed.	Not applicable to our bodies of evidence or review question so not assessed.	

Figure S1 - Global density of wastewater sampling of included studies

Count of studies sampling countries plotted by points on country centroids with size reflecting number of studies.



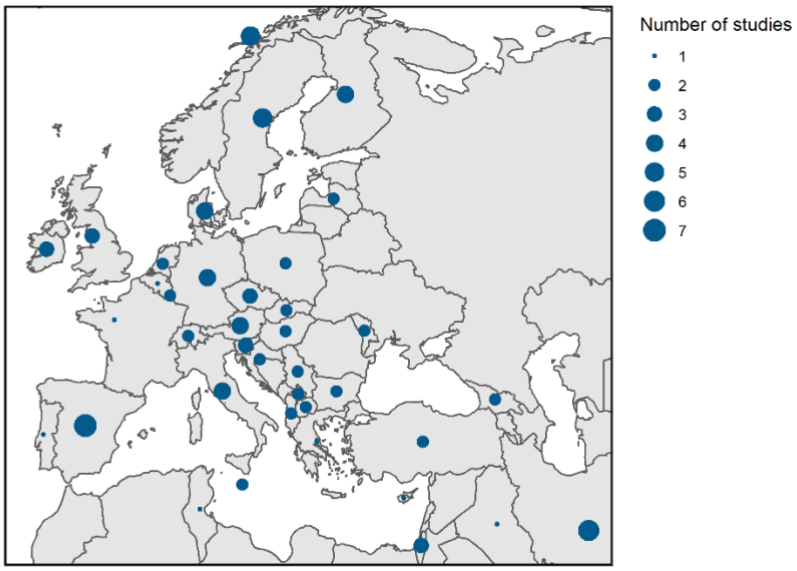


Table S4 – Descriptive table 1 of included studies

Citation	Country	Relevant comparison	Sampling method	WWTW details	AMR detection method	Human dataset	Wastewater sample sizes	Data analyses methods	Relevant outcomes	Overall risk of bias
Adator et al. 2020a	Canada	Comparison of *E. coli* antibiograms from wastewater isolates (selective and non-selective media) and human clinical isolates.	Composite of influent and effluent (1L).	Influent and effluent (post-grit tank) from 2 WWTWs – no details reported.	Selective and non-selective culture - AST via disk diffusion and PCR of 8 AMR targets.	Clinical isolates from blood, urine and abdominal samples (n=25).	12 wastewater samples and 194 wastewater isolates.	To compare wastewater and human group effect on AMR prevalence, univariate multinomial logistic regression models were used.	Wastewater and human components showed more similar phenotypic and genotypic resistance prevalence than when compared to other one health samples. Human isolates were not exclusively higher or lower in MDR or genotypic AMR testing than wastewater isolates.	Unclear
Adator et al. 2020b	Canada	Comparison of ESBL *E. coli* between wastewater and human clinical isolates.	Composite of influent and effluent (1L).	Influent and effluent (post-grit tank) from 2 WWTWs – with agriculture input.	Disk-diffusion and WGS.	Human clinical isolates randomly selected from the same period as sampling (subset for WGS = 25).	Unclear number of wastewater samples, WGS of wastewater isolates = 30.	Analyses conducted on generated pan genomes. Separate binary logistic regression used to example wastewater-human ARG prevalence differences. Maximum likelihood phylogenetic tree also generated.	Phylogenetic AST showed human clinical samples and wastewater isolates had similar resistance prevalence. Both groups had higher prevalence for certain antibiotics than the other. Genotypic analyses showed a similar outcome with genotype-phenotype agreement ranging from 93.2-100% depending on the antibiotic.	Unclear
Aljanaby et al. 2018	Iraq	Antibiotic susceptibility of *P. aeruginosa* to 10 antibiotics compared for 60 isolates from patients with pneumonia and 60 wastewater isolates from the same city.	Single grab samples of "sewage water" from undefined sampling point.	"Purified water station" located in the same city as where clinical isolates were collected.	Disk diffusion for 10 antibiotics and ESBL kit tests.	60 clinical isolates from sputum samples from patients with pneumonia.	60 wastewater samples and 60 wastewater isolates analysed.	Chi-squared test.	Significantly higher prevalence of MDR in clinical isolates compared to wastewater isolates. Overall higher prevalence of individual antibiotic resistance in clinical isolates also. Non-significant difference in prevalence of ESBLs between the two compartments.	Unclear
Colomer-Lluch et al. 2014	Tunisia, Spain	Abundance of several AMR genes in geographically-distant wastewater sources with differing characteristics and historical resistance backgrounds. Country-level comparisons of	Grab samples collected in sterile conditions.	Influent from four WWTWs with tourist facility and abattoir sewer inputs (sheep, cattle).	Samples filtered and phases extracted by QIAamp DNA Blood Mini Kit (retained bacteria) and phenol:chloroform in-house method (phage containing supernatant). qPCR of genes: blaTEM, blaCTX-M-1, blaCTX-M-9, mecA,	Country-specific historical resistance rates and specific resistance prevalence.	Unclear total wastewater samples: 26 wastewater samples underwent DNA extraction and 28 underwent viral phage fraction analyses.	None conducted.	Most common ESBL isolated reflects most reported clinical ESBL. Higher ARG density from high-antibiotic use country.	High

		resistance rates - indirect.			armA, qnrA, qnrS, sul1. Eight target genes.					
Golle et al. 2017	Slovenia	Assessing the prevalence/variability of carbapenem-resistant *P. aeruginosa* (CRPA) strains isolated from clinical isolates and WWTW effluent in the same region using PFGE and WGS.	Not reported	Effluent from two with distinct community and hospital inputs (1300 or 260 beds). Additional six healthcare facilities including geriatric primary care unit and psychiatric long-term care facility). Median flow m3/d 22207.5 (IQR 7928.5).	Disk diffusion (10 antibiotics). DNA extraction of subset (QIAamp DNA Mini Kit), MiSeq sequencing and identification of acquired carbapenemase genes using Resfinder.	CRPA cultured during the same sampling period from all respiratory and urine samples from hospitals, other health care facilities and GPs (n=130).	83 CRPA isolated from effluent samples. Totals including clinical: 208 PFGE types and 112 sequenced.	None conducted	No pattern or uniformity of PFGE types observed over the year with low overlap (4 isolates) between clinical and effluent samples regardless of region. Low overlap also seen between hospitals and other clinical settings (2). Slightly higher overlap between WWTWs (6). AST-based resistance patterns: clinical isolates have higher diversity in resistance patterns but resistance to individual antibiotics is more prevalent in effluent for 6/8 reported.	Low
Gouliouris et al. 2019	UK	Genomic comparison of wastewater *E. faecium* isolates and isolates from clinical infection preceding and during study.	Composite (two subsamples) of consecutive 500 ml samples. Hospital grab samples of 1 L.	Influent and effluent from 20 WWTWs where 10 received hospital waste and 10 did not. 15 urban and 5 rural sites. Median PE 47957.5 (IQR 127521) and flow m3/d 27043.2 (IQR 48621.6).	Selective media for AMR counts and VITEK microbroth dilution for AST. Hiseq for WGS. Phylogenetic WGS analyses with ResFinder (ARIBA) and plasmidSPAdes/BLASTN.	187 *E. faecium* associated with bloodstream infection in 187 patients. 140/187 patients in hospitals served by sampled WWTWs preceding wastewater sampling period and 23/187 collected during sampling period.	40 WWTW wastewater samples from which 388 *E. faecium* isolates sequenced. Five hospital wastewater samples from which 40 isolates sequenced. Total of 423 in data analysis after dropping 5 low quality sequences.	Maximum-likelihood phylogeny (no recombination/MGE).	ST typing show high diversity in influent and much lower in clinical samples – consistent with specific ST's association with drug resistance. Comparing relatedness between sample groups showed minimum SNP distance between any bloodstream to wastewater isolates did not differ significantly for plants with hospital input or not.	Low
Haghi et al. 2019	Iran	Prevalence and resistance patterns of Enterococci isolated from healthy humans and wastewater from the same region during the same period.	Grab samples collected in sterile bottles.	Six sites of "urban sewage waters".	Disk diffusion with seven antibiotics for ASTs. MIC by agar dilution method.	100 healthy faecal samples with no history of diarrhoea and antibiotic therapy for at least one month.	100 wastewater samples and 100 faecal specimens - total of 141 Enterococci isolated. 77/141 from wastewater samples and 64/141 from faeces.	None conducted	Top three single-resistances shared. Wastewater had higher prevalence of resistance to three other antibiotics, multi-drug resistance and VRE.	High
Hendriksen et al. 2019a	Kenya	Resistomes of wastewater compared to surveillance data and laboratory-confirmed cases of	Grab samples of 500 ml.	Untreated waste from two confluence points of drainage ditch networks with	DNA extracted (QIAamp Fast DNA Stool mini kit on pellet) and Viral RNA/DNA extracted (Nucleospin RNA XS kit	PBIDS local surveillance system comprised of household morbidity and	42 wastewater samples sequenced.	Confidence intervals calculated on mean read abundance. Weekly number of reads for specific	Increases/decreases in read abundances for bacteria/AMR genes were not reflected in the contemporaneous PBIDS data. Non-significant observed increase	Unclear

		wastewater source populations.		the high surface flow. Each point drains household latrine waste from a separate village: population (population density): 1845 (8.4/m2) and 3727 (8.8/m2).	on supernatant). Hiseq and Miseq (PE, 2 x 250bp) used to sequence pellet and supernatant extracts respectively. MGmapper used to map against genomic/viral databases and Resfinder for AMR genes.	health care usage data contemporaneously collected through home visits. Key measures used in study: total cases of diarrhoea, fever, acute febrile illness and number of clinic visits. Stool culture also conducted.		bacteria, viruses or AMR genes surpassing upper limit defined as significant.	in read abundances of pathogens appeared to coincide with reported illness/visits to clinic.	
Hendriksen et al. 2019b	60 countries sampled	Resistomes of sewage influent collected on a global scale compared to source regional antimicrobial use, and socio-economic, health, and environmental factors.	24H flow proportional sampling or composites (3 sub samples) over a short time period (minimum 15 min).	Influent from 53 WWTWs. 44/53 received hospital input, 34/53 received industrial input, 17/53 received abattoir input and 14/53 received agricultural input.	DNA extracted (QIAmp Fast DNA Stool Mini Kit with twice input material and bead beating changes), libraries prepped (NEXTflex PCR-free Library Prep Kit) and sequenced (HiSeq3000 2 x 150bp paired-end). MGmapper used to align against reference sequence databases and AMR genes annotated using ResFinder.	World Bank Health, Nutrition, Population, Development indicator data sets collected between the years 2000 and 2016 for 259 countries and territories.	Samples sequenced for 80 sites (WWTW and others) - results from 79 reported.	Multilevel Poisson model including observation-level random effects and fixed effects.	Total AMR gene abundances detected across sites corresponded to existing data describing their country/continent of origin. Country-specific variables made up from World Bank's Health, Nutrition and Population, Development indicator data sets (mostly sanitation and general health) explained up to 89% of observed resistome variation between samples regardless of diversity of AMR genes. Countries with higher HDI have significantly lower abundance of AMR genes and the number of passengers flying into a country has no effect on abundance. No significant association between temperature at collection with AMR genes abundance. No significant effect was found when investigating total usage of all antimicrobials on the abundance of AMR genes on a class-level. No significant association was found between the abundance of AMR genes on a class-level and the antimicrobial residue levels measured.	High
Huijbers et al. 2020	Denmark, Finland, Norway, Sweden, Belgium, France, Germany,	Wastewater-isolated *E. coli* AST compared against AMR surveillance network data for the country samples at a similar time period.	Time-proportional, flow-proportional, volume-proportional.	Influent from 10 large WWTWs (>500,000 PE each).	Non-selective chromogenic plating and isolation of suspected *E. coli* (confirmed by MALDI). Microbroth dilution AST against six antibiotics.	EARS-net data of invasive *E. coli* isolated from blood and CSF clinical samples extracted for countries sampled for 2016 (each country has	10 wastewater samples, 2507 *E. coli* isolated and underwent AST.	Resistance prevalence and 95% CIs calculated based on the binomial probability function. Chi-squared test used to assess relationship between isolate	Resistance prevalence lower in wastewater *E. coli* compared to clinical *E. coli* but followed geographic trends based on the ten countries studied. Wastewater-human comparisons showed significant correlation for 2/4	Low

	Greece, Italy, Spain					contributed >1000 isolates to EARS-net previously).		prevalence and national surveillance data. Simple linear regression used for wastewater-human relationships.	antibiotic classes with higher sample sizes.	
Hutinel et al. 2019	Sweden	Resistance rates of *E. coli* isolates from hospital/WWTW wastewater compared to *E. coli* isolated from hospital ward and primary care clinical samples during the same period.	24H WWTW flow proportional sampling comprised of a minimum of 224 subsamples. 24H hospital effluent composite comprised of subsamples taken every 9 min (160 subsamples).	Influent from WWTW with PE 746,882.	Microbroth dilution of stocked isolates against 11 antibiotics. ESBL screening by multiple disk diffusion.	*E. coli* urine/blood isolates from patients on hospital wards connected to hospital wastewater sampling point or primary care patients from municipalities connected to the sampled WWTW (n=6270). Contemporaneous and before wastewater sampling period.	14 wastewater samples from which 1252 *E. coli* were isolated.	Resistance across clinical/sewage modelled with linear regressions.	Resistance prevalence of hospital effluent, clinical samples and primary care urine isolates correlated with annual mean resistance rates. ESBLs higher in hospital effluent than WWTW influent. Clinical samples had double resistance rates of influent isolates except for one antibiotic.	Unclear
Jakobsen et al. 2008	Denmark	Gen-R resistance *E. coli* isolates from sewage with/out hospital sewer input and from contemporaneous hospital patient UTI samples.	24H flow proportional WWTW sampling. Composites (flow-respective mix) of hospital effluent. Composites of residential area outlet comprising 5 samples taken with 10 min interval (5 subsamples).	Influent from WWTW with 230,000 PE receiving industrial hospital (794 bed) input.	Microboth dilution for ASTs to five antibiotics. PCR detection of four genes: aac(3)-II, aac(3)-IV, ant(2")-I, and armA. Additional PCR detection of aac(3)-I and aac(3)-III in subset.	38 Gen-R *E. coli* isolates from hospital patients with UTIs during wastewater sampling period with addition month either side.	33 wastewater samples from which 55 GEN-R *E. coli* isolates cultured.	None conducted	Prevalence of Gen-R *E. coli* similar between hospital effluent and WWTW influent - lowest in residential outlet. No sharing of PFGE types between any patient isolates and wastewater isolates.	High
Jorgensen et al. 2017	Norway	Genomics of ESBL-*E. coli* isolated from influent compared to those isolated from UTI patients living in the WWTW catchment	24H flow proportional sampling.	Influent from one WWTW.	Vitek2 microbroth AST and E-test. PCR for blaCTX-M and DNA microarray for AMR 11 genes for isolates negative for blaCTX-M, Hiseq (150 PE) WGS and	94 patients with community acquired UTI caused by ESBL-*E. coli* living within catchment area of the WWTW	Five wastewater samples from which 91 ESBL *E. coli* isolates cultured.	Univariate and multivariate analyses conducted by binary logistic regression.	Clinical isolate single and multi-drug resistance rates higher than wastewater isolates. Most prevalent ESBL gene was shared between groups and prevalence for other ESBL genes also similar except M-14 higher in wastewater.	High

		area with temporal overlap.			analysis using ResFinder and PlasmidFinder.	collected before wastewater sampling.			Prevalence of acquired resistance genes similar between groups but higher in urine isolates. No difference in plasmid replicon typing between groups.	
Karkman et al. 2020	60 countries sampled (GSSP)	Resistomes of sewage influent collected on a global scale compared to local clinical surveillance data and socio-economic, health, and environmental factors.	Mixed (24H flow proportional and composites)	As described for Hendriksen (GSSP).	DNA extracted (QIAmp Fast DNA Stool Mini Kit with twice input material and bead beating changes), libraries prepped (NEXTflex PCR-free Library Prep Kit) and sequenced (HiSeq3000 2 x 150bp paired-end). MGmapper used to align against reference sequence databases and AMR genes annotated using ResFinder.	Clinical resistance data for invasive *E. coli* (blood, CSF) against four classes of antibiotics extracted from networks EARS-Net, CAESAR and ResistanceMap.	As described for GSSP.	Prevalence of clinical resistance was modelled using beta regression. Further analyses with Principal component analysis, correlation plots and heatmaps.	Sewage resistomes correlated to surveillance data of invasive *E. coli* in terms of prevalence on a global scale. Resistome ARGs were mostly insufficient to explain variability of clinical resistance. Similarly, resistance to antibiotic classes could not be associated with corresponding ARG resistance classes except for beta-lactam ARGs and aminopenicillin resistance. Combining sewage resistome data and socioeconomic data from respective countries produced the best performing models.	High
King et al. 2020	South Africa	Comparison of AMR profiles of wastewater-isolated Klebsiella and clinical Klebsiella isolates	Grab sampling.	Influent and effluent from 2 WWTWs (no details reported)	Disk diffusion and AmpC/ESBL detection kits.	Undefined clinical isolates (n=72).	4 wastewater samples and 37 Klebsiella isolated.	None conducted.	Common resistance profiles were detected between clinical and wastewater isolates. Wastewater isolates had a lower resistance prevalence compared to clinical but much fewer wastewater isolates were collected. Carbapenem resistance was detected in clinical sampled but not in WWTW samples (was detected in hospital effluent though).	Unclear
Meir-Gruber et al. 2016	Israel	Prevalence of blaKPC/blaNDM-1 carrying CREs, MRSA and VRE in WWTW influent and domestic sewer systems collected nationally during 2012-2013 compared to in clinical samples from a major hospital during 2013. Distribution of CREs, MRSA and VRE from sampling sites with and without hospital sewage input.	24H WWTW and sewer composites via autosampler (24 or 48 subsamples).	Influent from six WWTWs serving large urban areas across Israel (population high = 462670, low = 56943). 4/6 WWTWs received healthcare input (hospitals and nursing homes).	Microbroth dilution (Phoenix) susceptibility testing (antibiotics not provided). RT-PCR for detection of blaKPC and blaNDM-1.	All blood cultures, general cultures and rectal swabs received by one major hospital clinical microbiology laboratory in 2013 (semi-contemporaneous) Sample size not reported.	Six WWTW influent samples and 10 sewer system samples. Total number of isolates cultured or tested unclear: approximately 560 tested and 112 CRE included in analyses.	None conducted	WWTW samples and clinical isolates shared the same most commonly isolated blaKPC-carrying species (*K. pneumoniae*). They also shared the three other most commonly isolated blaKPC species but at different rates. Higher prevalence of MRSA and VRE in hospital wastewater and the receiving WWTWs than in domestic wastewater.	High

Moradigaravand et al. 2018	UK	Comparison of resistance gene prevalence in *C. difficile* isolated from hospital patients and wastewater treatment works both downstream and unconnected to hospitals.	Composites of two consecutively collected 0.5L grabs from both influent and effluent.	Influent and effluent from 20 conventional WWTWs where 10 received hospital effluent and 10 did not.	Single isolate WGS.	70 clinical isolates from patients with *C. difficile*-associated diarrhoea.	40 wastewater samples and 186 wastewater isolates analysed.	None conducted	Low overall prevalence of AMR genes, most common being ermB (13%) and tetM (7%).	High
Mourkas et al. 2019	Spain	Comparison of wastewater Campylobacter isolates and clinical isolates.	Grab samples" collected in sterile containers".	Effluent from one site of "urban effluents".	Microbroth dilution, WGS of subset.	Campylobacter isolates associated with campylobacteriosis cases from sewershed hospitals (n=152).	50 wastewater samples and 53 wastewater isolates.	Sum of core genes used to create pangenome. Neighbour joining clustering method used to generate phylogenetic trees.	Phenotypic resistance prevalence of individual antibiotics tested were similar between human and sewage isolates – both groups showed high prevalence than the other in specific antibiotic comparisons. Genotypic comparisons showed human isolates generally had higher prevalence of the AMR genes tested.	High
Ojer-Usoz et al. 2017	Spain	Comparison of resistance to individual antibiotics and presence of resistance genes in ESBL *E. coli* isolated from wastewater and hospital patients.	Not reported.	No details reported.	Microbroth dilution for 35 antibiotics and PCR for 5 AMR genes.	130 clinical isolates of mixed origin provided by hospital routine diagnostics.	279 wastewater samples and 132 wastewater isolated analysed.	R × C contingency test or Fisher's exact test.	Similar resistance prevalence to individual antibiotics across compartments for ~half of antibiotics tested, differences with some significant for the rest - both higher and lower prevalence seen in wastewater compared to human compartment. Very similar presence of bla genes detected across compartments human (112/130) and wastewater (117/132) but varying specific gene distribution - CTX-M-15 being notably higher in human isolates (39.2% vs 11.1%).	High
Oravcova et al. 2017	Czech Republic	Comparison of resistance to individual antibiotics and presence of resistance genes in VRE isolated from wastewater and hospital patients on	Moore swab method exposing a cotton swab to wastewater effluents for 7 days.	Effluent from a large urban WWTW receiving effluent from multiple hospitals and long-term care facilities.	Disk diffusion for 14 antibiotics and PCR for 14 AMR genes.	33 isolates from rectal swabs collected from patients with current or previous oral antibiotic therapy.	37 wastewater samples and 49 wastewater isolates analysed.	None conducted.	Presence of identical isolates from the WWTW and hospitalized patients, with presence of vanA-enterococci in WWTW outflows. All isolates cross compartments possessed vanA, but none harboured vanB. Similar phenotypic resistances to individual antibiotics seen across WWTW and clinical	Low

		oral antibiotic treatments.							isolates but varying prevalence of different genotypic AMR determinants.	
Pärnänen et al. 2019	Portugal, Spain, Ireland, Cyprus, Germany, Finland, and Norway	The relative abundance of targeted AMR genes in WWTW influent/effluent collected across multiple European countries compared with AMR surveillance data from the EARS-Net database and national human antibiotic consumption amounts.	24H composite samples (24 hourly sub samples totalling 1-2L) collected on three consecutive days (Tuesday, Wednesday and Thursday).	Influent and effluent from 13 WWTWs from seven countries. Median PE 173839 (IQR: 289906) and flow 33032.5 m3/d (IQR: 100529). Hospital sewer input described for 7/13 WWTWs.	DNA extraction (PowerWater Kit) and qPCR array with 384 primer sets targeting AMR genes and mobile genetic elements.	EARS-Net clinical AMR surveillance data from 30 European countries, specifically the 2017 report for 2013-2016 describing "a north-to-south and a west-to-east gradient is evident in Europe".	168 wastewater samples. 142/168 underwent qPCR (92 effluent and 50 influent).	PCoA using the Bray-Curtis dissimilarity index.	Wastewater ARG distribution consistent with country-level gradient from EARS-Net report. Relative abundance of ARG classes reflects country-level antibiotics consumption but more insertion sequences in low-use countries. Effluent antibiotics concentrations and primary care antibiotics consumption not associated to ARG abundance.	High
Pehrsson et al. 2016	Peru, El Salvador	Resistomes and functional metagenomics of human faecal samples compared to wastewater collected from WWTW serving the human community sampled.	Samples collected in sterile containers.	Influent and effluent from one WWTW serving population of 700,00 including two shanty towns of low-income families.	Metagenomic DNA extracted from pellet via in-house phenol-chloroform bead beating method. Small-insert shotgun expression libraries created (pZE21 in *E. coli* DH10B), colony libraries selectively screened and pooled for DNA extraction. Metagenomic inserts amplified (PCR) and sequenced (HiSeq 2000). Assembly with PARTFuMS and annotation with Resfarms. Metagenomic DNA sheared (300–400 bp) and sequenced (HiSeq/NextSeq 2 × 150 PE). ShortBRED24 used to quantify abundance of AMR genes in the metagenomes based on generated functional metagenomics, CARD, human faecal metagenome	Human faecal samples from shanty town serving WWTW sampled - collected semi-contemporaneously with wastewater. 5.7 ± 3.0 individuals per house sampled. Up to 44 faecal samples analysed.	24 wastewater samples (12 influent, 12 effluent). Metagenomics on 22 wastewater samples after dropping low-yield extracts. Functional libraries created from 31 human samples from four households. Metagenomic libraries created from 44 human faecal samples.	Non-parametric Student's t-tests with Bonferroni correction.	Street-access wastewater and WWTW influent resistomes were equally similar to human-associated resistomes. Street-access wastewater was more similar to influent than human faeces in terms of bacterial composition. All wastewaters had higher phylogenetic diversity and more AMR proteins per sample than the human faeces. Drug efflux antibiotic resistance mechanisms were higher overall in human faecal vs wastewater resistomes. Investigating genetic context of cosmopolitan AMR proteins showed network of gene sharing between all sample metagenomes.	High

					assemblies, Lahey b-lactamase database.					
Pignato et al. 2010	Italy	Prevalence of ampicillin resistance between *E. coli* and Salmonella isolates cultured from wastewater compared to Salmonella isolates cultured from symptomatic hospitalized patients.	Composite sampling (10 subsamples).	Influent and effluent from two WWTWs serving separate towns with hospital input for one site. Median PE 21500 (IQR 16000) and flow m3/d 3801.6 (IQR 3110.4)	Selective media at cut-off points to indicate resistance for eight antibiotics. Disk diffusion for subset of isolates (66) against 32 antibiotics to determine resistance patterns. PCR of four bla genes and class 1 integron gene cassettes.	274 Salmonella isolates cultured from faeces of hospitalized patients with gastroenteritis at hospital served sampled WWTW during wastewater sampling period.	108 total wastewater samples from which 64 Salmonella isolates cultured. 273 previously cultured *E. coli* included.	None conducted	Resistance prevalence in Salmonella isolates was of the same order in isolates from wastewater and clinical samples. Salmonella resistance profiles were similar between sample sources according to serovar. Specific six-drug pattern observed with a similar frequency in in Salmonella serovar Typhimurium isolates from wastewater and clinical specimens. Persistent isolation of suspected (not phage-typed confirmed) Salmonella DT104 endemic clone accounting for 40% of human salmonella isolates in the study country. Suspected DT104 represented 50% of wastewater isolates but only 22.2% of clinical isolates. Another potential clone with different restricted resistance pattern also appeared to be endemic in the same population.	High
Pot et al. 2020	Guadeloupe	Comparison of ESBL *Enterobacter cloacae* from wastewater and sewershed clinical infections.	Grab samples (500ml)	Influent and effluent from 1 large WWTW treating hospital and urban wastewater - activated sludge process.	Selective media, disk diffusion, subset underwent WGS.	ESBL-producing *Enterobacter cloacae* associated with clinical infection (n=36).	20 wastewater samples and 63 wastewater isolates.	Maximum likelihood phylogenetic tree. Temporal analysis via linear regression between root-to-tip and isolation year.	Focussing on wastewater-human component of the one health study: ESBL *E. coli* strains isolated from wastewater mirrored the prevalence in human clinical isolates. Regional ESBL-producing *E. coli* ST114 divided into two lineages both including wastewater and clinical isolates.	Low
Rahimi et al. 2014	Iran	MRSA isolates collected during the same sampling period from sewage and human clinical infections compared using biochemical fingerprinting, SCCmec typing and ASTs.	Three grab samples collected in sterile 250 ml bottles.	Urban WWTW west of Tehran.	MIC by E-test for two antibiotics. AST by disk diffusion for 16 antibiotics. PCR detection of mecA, nucA, SCCmec (eight types/subtypes) and ccr (four types).	489 clinical samples (wound, urine, sputum, blood, CSF, nose and abscess) from hospitalized patients collected during the same year as wastewater sampling.	Three wastewater samples from which 653 S. aureus cultured.	None conducted	Higher rates of multi-drug resistance in wastewater isolates versus clinical. Wastewater isolates also had higher single resistance rates to five other antibiotics. Homogeneity in biochemical typing consistent with contemporaneous outbreaks and dissemination of clones.	High
Raven et al. 2019	United Kingdom	Genomic comparison of WWTW *E. coli* isolates (with and without hospital sewage input) to *E. coli*	Composite (two subsamples) of consecutive	Influent and effluent from 20 WWTWs where 10 received hospital waste	Microbroth dilution (Vitek2) AST testing/phenotypic carbapenem test. DNA extracted (Qiagen	Genomes of 437 *E. coli* isolates from hospital patients blood cultures (13/437 overlap	40 wastewater samples from which 394 *E. coli* isolates sequenced. Total	Phylogenetic analysis of ST131 based on core genome SNPs.	Top three most common ST types from clinical blood cultures present in most WWTWs sampled (both receiving and not receiving hospital waste). Most prevalent ESBL gene	Low

		blood isolates from hospitalised patients in the same region preceding and during WWTW sampling period.	500 ml samples.	and 10 did not. 15 urban and 5 rural sites. Median PE 47957.5 (IQR 127521) and flow m3/d 27043.2 (IQR 48621.6).	QIAxtractor) and sequenced (Hiseq2000). AMR genes identified with ARIBA and ResFInder. Mobile genetic elements encodings five bla genes identified using in silico PCR approach.	with wastewater sampling, 424/437 are prior during 2006-2013).	of 388 in data analysis after dropping 6 low quality sequences.		from clinical samples was detected in all WWTWs. Prevalence of most common ESBL genes approximately reflected between wastewater and clinical samples although higher in wastewater. Majority of AMR gene variants identified were shared between groups. No detection of carbapenem or colistin resistance, consistent with very low resistance to these antibiotics in study area.	
Reinthaler et al. 2012	Austria	Prevalence of antimicrobial resistant *E. coli* and resistance patterns of patient isolates collected in 2000 and 2009 compared to sewage sludge isolates collected during the same period.	Grab samples using sterile wide-mouth bottles.	Activated sludge from five WWTWs with no sewer input from hospitals.	Microbroth dilution (Vitek2) for identification and susceptibility testing to 15 antibiotics.	Random subset (500) of *E. coli* collection from patient samples (urine, sputum, stool, wound, skin and respiratory tract) taken by general medical practitioners during wastewater sampling periods.	55 total wastewater samples from which 356 *E. coli* isolates cultured.	None conducted	Most commonly isolated resistances shared between patient and sludge isolates but lower overall prevalence in sludge. Increase in individual resistances over time seen in both groups.	High
Saifi et al 2009	Iran	Assessing clonality and resistance prevalence of gentamicin-resistance *Enterococcus faecium* isolates isolated from clinical and WWTW samples. Both WWTWs and clinics in the same city but overlap unclear but associated in time.	Not reported	Influent from three WWTWs located in different parts of Tehran.	Disk-diffusion and microbroth dilution for antimicrobial susceptibility testing (10 antibiotics). PCR for five genes (aac (6 ')-Ie-aph (2 ")-Ia, aph (2 ")-Ib , aph (2 ")-Ic, aph (2 ")-Id and ant (4 ')-Ia).	320 enterococcal isolates cultured from urine and wound samples from six outpatient clinics during sampling period.	Unclear total wastewater samples and isolates: total of 106 high-level gentamicin-resistant isolates (clinical = 48, WWTW = 58)	None conducted	Clinical isolate resistance rates higher than wastewater isolates. Individual resistances match national prescribing levels. Low PFGE type overlap between groups.	High
Talebi et al. 2008	Iran	Prevalence of resistance to single/combinations of antibiotics in clinical VRE and wastewater VRE isolates. PFGE and biochemical fingerprinting patterns in clinical VRE compared to wastewater VRE isolates.	Grab samples using 250 ml sterile wide-mouth bottles.	Three urban WWTWs located in different areas of Tehran.	VRE isolated by classical culture-based methods. PCR for detection of vanA, vanB. AST disk diffusion and E-test for eight antibiotics.	Enterococcus species isolated from clinical samples (450) (predominately urine) from three major hospitals.	Six wastewater samples from which 593 enterococcal isolates cultured.	None conducted	MICs higher in human isolates compared to wastewater isolates. Multi-drug resistance detected at the same prevalence between sample groups. Human isolates showed higher PFGE type diversity consistent with no reported contemporaneous outbreaks compared to lower diversity in wastewater isolates.	High

Urase et al. 2020	Japan	Comparison between wastewater-isolated CTX-R *E. coli* and MPEM-R Enterobacteriaceae prevalence compared to JANIS national surveillance data.	Not reported	Effluent "after chlorination" from 12 WWTWs in Tokyo and suburbs. Details: 9 receive hospital and use activated sludge treatment, 2 use oxidation ditches in rural areas, 1 is a community plant with no medical facilities.	Selective chromogenic media incubated at both 37 and 42C - isolates identified by 16S rRNA sequencing, disk-diffusion AST.	Resistance prevalence from Japan Nosocomial Infections Surveillance (JANIS 2019) - both inpatients and outpatients.	17 wastewater samples, 605 *E. coli* isolated tested against CTX, 976 Enterobacteriaceae tested against MEPM.	None conducted.	Lower resistance prevalence in effluent wastewater isolates (5.7%) compared to clinical surveillance data of both in-(27.5%) and outpatients (17%).	High
YoungKeun et al. 2015	Sweden	Resistance rates of *E. coli* isolated from WWTW influent, effluent and hospital effluent in 2013-2014 compared to national resistance rates of human blood isolates from 2007-2012.	24H WWTW "continuous" sampling via autosampler. 4H hospital composites comprising 50 ml grab samples every 10-15 minutes (16 subsamples).	Influent from WWTW with PE >750,000 and flow 250,000 m3/d. Receives hospital input (800 bed) but no significant agricultural sewer input.	Microbroth dilution (PhenePlate) with 10 antibiotics. Microarray targeting six ESBL and six plasmid-mediated AmpC genes.	Resistance data from EARS-Net European database on human blood isolates limited to studied country for the years 2007-2012 (preceding wastewater sampling).	17 influent samples from which 1326 *E. coli* isolates cultured. Three effluent samples from which 117 *E. coli* isolates cultured. Six hospital effluent samples from which 451 *E. coli* cultured.	None conducted	Increasing resistance prevalence of WWTW isolates reflect increase in resistance of human blood isolates. Lower prevalence of resistance in WWTW isolates compared to hospital effluent.	High
Zaheer et al. 2020	Canada	Comparison of wastewater-isolated Enterococcus spp. and clinical isolates.	Composite of influent and effluent (1L).	Influent and effluent (post-grit tank) from 2 WWTWs – with agriculture input.	Selective culture, disk diffusion and WGS of subset.	Clinical isolates - mostly non-sterile sites but blood, other tissue and rectal swabs (subset for WGS = 299).	43 wastewater samples and subset of 169 isolates underwent WGS.	Phylogenetic analysis via single nucleotide variate phylogenomics of core genes using unassembled sequence data.	MDR prevalence in clinical isolate was much higher than wastewater isolates. VRE was only detected in clinical isolates. The majority of wastewater and human isolates were closely related when compared to other one health samples, and harboured similar resistance genes. Phylogenetic clustering showed clinical and wastewater isolates grouped together separately compared to other one health samples indicating surveillance potential of human carriage.	Low
Zarfel et al. 2013	Austria	AST, biochemical phenotypes, plasmid replicon types and bla gene presence of ESBL	Grab samples collected using sterile	Activated sludge from five WWTWs "PE range <10,000 to	AST for 11 antibiotics using microbroth dilution (Vitek). Three additional antibiotics	50 ESBL *E. coli* isolates from hospitalized patients with UTIs	30 sludge samples from which 50	None conducted	Most prevalent ESBL genes detected were shared across wastewater sludge and UTI isolates. TEM-1 gene detected at similar	Unclear

		E. coli isolates from WWTW sludge compared to isolates from hospital patients within WWTW catchment and collected in the same sampling period.	wide-mouth bottles.	>100,000 and flow rate range 100–1200 L/min".	via disk diffusion. CLSI confirmation of ESBL. PCR for five b-lactamase gene families.	collected contemporaneously with wastewater.	ESBL *E. coli* isolated.		rates between groups. Both groups shared the same level of resistance pattern diversity. UTI isolates showed reduced phenotypic variation but both groups clustered together with no distinction. Higher resistance rates against clinical antibiotics seen in UTI isolates but not for non-clinical antibiotics.	

Table S5 – Descriptive table 2 of included studies

Citation	Study aims	Wastewater sampling strategy	Targeted organism and resistances	Wastewater handling and storage	Sampling season and weather reporting
Adator et al. 2020a	To investigate ESBL *E. coli* from a one health continuum linked to beef production.	Longitudinal sampling over 24 months – frequency of visits not reported.	ESBL *E. coli*.	Not reported.	All seasons - weather conditions not reported.
Adator et al. 2020b	To examine similarities between AMR patterns/prevalence of *E. coli* along a one health continuum.	Longitudinal over 24 months with a two-monthly interval.	ESBL *E. coli*.	Not reported.	All seasons - weather conditions not reported.
Aljanaby et al. 2018	To investigate the relationship between virulence genes and antibiotic resistance in pathogenic bacteria isolated from clinical and environmental sources.	Longitudinal sampling over 10 months - interval not reported.	*P. aeruginosa.*	Not reported.	Not reported.
Colomer-Lluch et al. 2014	To evaluate the abundance of several ARGs in bacteria and bacteriophage DNA derived from wastewater collected from distant sources with differing socio-economic and cultural characteristics, climate and geographic background.	Sampling between 2011 and 2014 with unclear interval.	Metagenomics targeting blaTEM, blaCTX-M-1, blaCTX-M-9, mecA, armA, qnrA, qnrS, sul1.	Frozen at -80°C and stored in dry ice before filtration (0.45µm) and processing.	Not reported.
Golle et al. 2017	To evaluate possible overlap between clinically relevant and environmental strains and determine transfer from wastewater to WWTW.	Longitudinal sampling between Jan-Dec 2014 with one month interval - 12 timepoints.	Carbapenem-resistant *P. aeruginosa* targeting 1411 carbapenemase genes.	Refrigerated for up to 6H before filtration (0.45µm).	All seasons - weather conditions not reported.
Gouliouris et al. 2019	To generate indirect evidence for the extent to which healthcare-associated *E. faecium* is disseminated in the community.	Snapshot sampling between June 2014 and January 2015.	Vancomycin-resistance *E. faecium*. Non-limited genotyping.	Sodium thiosulphate-based preservation.	Summer, autumn, winter - weather conditions not reported.
Haghi et al. 2019	To investigate potential transmission of VRE from wastewater to humans by sampling VRE in wastewater and faecal samples of healthy carriers.	Longitudinal samples "collected randomly" from June to October 2017. Unclear timepoints.	Vancomycin-resistant *Enterococcus* spp.	Kept at 4°C until processing - no other details reported.	Summer - weather conditions not reported.
Hendriksen et al. 2019a	To monitor circulating pathogens and AMR genes via metagenomics of urban wastewater with comparison	Longitudinal sampling every Monday and Wednesday between June 16 and August 26 2014 - 42 timepoints.	Metagenomics – non-limited genotyping.	Samples whole frozen at -80°C within 2 hours of collection before	Summer - weather conditions not reported.

Citation	Study aims	Wastewater sampling strategy	Targeted organism and resistances	Wastewater handling and storage	Sampling season and weather reporting
	to other concurrent disease surveillance.			shipment without coolant to authors. Samples were thawed for 48H at 4°C prior to processing.	
Hendriksen et al. 2019b	To characterise bacterial resistomes and explain AMR variation in untreated wastewater from sites sampled on a global scale.	Snapshot sampling between 25 Jan and 5 Feb 2016.	Non-limited	Samples frozen at -80°C (48H) before shipping to authors without cold chain. Samples then thawed at 20 for 12H before pelleting via centrifugation and storage at -20 or -80°C before processing.	Winter, spring summer - weather conditions reported as air temperature at time of sampling and brief description (sunny/cloudy/rainy/foggy/snowy).
Huijbers et al. 2020	To evaluate relationship between AMR prevalence in *E. coli* from wastewater and clinical samples with traditional clinical surveillance comparison.	Snapshot sampling, one timepoint per WWTW between Dec 2016 and Feb 2017.	*E. coli* resistant to ampicillin, ciprofloxacin, cefotaxime, ceftazidime, gentamicin, meropenem.	Cooled (4°C) until processing 12H after collection.	All seasons – based on reference flow provided during dry conditions, majority of sampling was during wet conditions.
Hutinel et al. 2019	To investigate the relationship between *E. coli* resistance rates in WWTW and hospital wastewater compared to clinical samples.	Longitudinal samples collected in 2016 with 1-2-month interval. Up to 12 timepoints.	*E. coli* - no specific resistance targeted.	Kept at 4°C and processed within three hours of collection.	All seasons - weather conditions not reported.
Jakobsen et al. 2008	To investigate the potential spread of gentamicin resistant (GEN-R) *E. coli* isolates or GEN-R determinants from a Danish university hospital to wastewater.	Longitudinal monthly sampling between October 2002 - August 2003 - 11 timepoints.	Gentamicin-resistant *E. coli* (aac(3)-II, aac(3)-IV, ant(2")-I, armA, aac(3)-I, aac(3)-III).	Not reported	All seasons - weather conditions not reported.
Jorgensen et al. 2017	To investigate possible transmission routes of ESBL-*E. coli* by comparing isolates	Longitudinal sampling (May 2010, two visits in June, Aug, and Sep) - 5 timepoints.	Third generation cephalosporin-resistant *E. coli* (blaCTX-M,	Filtered - no other details reported.	Summer - weather conditions

Citation	Study aims	Wastewater sampling strategy	Targeted organism and resistances	Wastewater handling and storage	Sampling season and weather reporting
	from various niches (freshwater lakes, saltwater basins, WWTW and UTIs).		blaTEM, blaSHV, blaNDM, blaKPC).		not reported.
Karkman et al. 2020	To characterise bacterial resistomes and explain AMR variation in untreated sewage from sites sampled on a global scale.	As described for GSSP.	Non-limited metagenomics exploration – phenotypic clinical surveillance data limited to four antibiotic classes: aminopenicillins, fluoroquinolones, third generation cephalosporins and aminoglycosides	As described for GSSP.	As described for GSSP.
King et al. 2020	To investigate the AMR profiles of *Klebsiella* spp. from hospitals, hospital effluents and WWTWs.	Longitudinal: two timepoints over three months.	*Klebsiella* spp. with focus on ESBLS and carbapenemase resistance.	Transferred on ice and processed within 3H of collection.	Winter and Spring - weather conditions not reported.
Meir-Gruber et al. 2016	To evaluate the population carriage of pan-resistant bacteria and to establish a convenient and fast method for wastewater sampling.	Snapshot sampling between Apr 2012 and Nov 2013.	Carbapenem, methicillin and vancomycin resistance in Enterobacteriaceae, *S. aureus* and *Enterococci* spp (blaKPC, blaNDM-1).	Not reported	Not reported
Moradigaravand et al. 2018	To investigate *C. difficile* in wastewater treatment plants and compare to clinical isolates from the same geographical region to identify if wastewater treatment plants are potential reservoirs.	Snapshot sampling of individual sites over 8 months - interval between sites not reported.	*C. difficile*, non-limited genotypic detection using srst2 package database.	Mixing of wastewater samples with sodium thiosulphate, filtration and alcohol shock for selected samples.	Summer, Autumn, Winter - weather not reported.
Mourkas et al. 2019	To compare genomes of AMR campylobacter from multiple niches.	Longitudinal sampling over 36 months – frequency of visits not reported.	Campylobacter with high resistance to ciprofloxacin, tetracycline, erythromycin, streptomycin and gentamicin.	Stored at 4°C – no other details reported.	Not reported
Ojer-Usoz et al. 2017	To characterise the clonal diversity of ESBL-*E. coli* from food, humans and environmental niches.	Longitudinal sampling - length and interval not reported.	ESBL *E. coli* and presence of blaCTX-M-14, blaCTX-M-1, blaCTX-M-15, blaTEM-42, blaSHV-12.	Not reported.	Not reported.

Citation	Study aims	Wastewater sampling strategy	Targeted organism and resistances	Wastewater handling and storage	Sampling season and weather reporting
Oravcova et al. 2017	To investigate if VRE is discharged by the sampled WWTW, if gulls are colonised by effluent VRE if present, and compare to admitted hospital patients.	Longitudinal sampling over 10 months - interval not reported.	VRE - Presence of vanA, vanB, vanC, tet(K), tet(L), tet(M), tet(O), erm(A), erm(B), mef(A), ant(4')-Ia, aac(6')-Ie-aph(2")-Ia, and aph(3')-IIIa.	Processed within 30m of collection.	All seasons - weather not reported.
Pärnänen et al. 2019	Trans-European antibiotic resistance surveillance in WWTWs - an initial survey for basing future regular surveillance protocols.	Longitudinal sampling with three timepoints over one year (autumn 2015, spring and autumn 2016). Four sites missing timepoints.	Targeting 229 resistance genes (not organism specific).	Filtered (0.22µm) and stored at -80°C before processing.	Autumn, spring - weather conditions reported as average annual air temperature based on "timeanddate.com".
Pehrsson et al. 2016	To map resistance gene dissemination between humans and their environments with focus on low-income populations.	Longitudinal sampling between May 2012 and Jan 2013 with 1-2-month interval. Up to nine timepoints.	Non-limited.	Frozen (−20 to −80°C) until shipment to authors where samples were stored at −80°C before filtration (0.22 µm).	All seasons - weather conditions not reported.
Pignato et al. 2010	To investigate whether wastewater isolates can be a useful/easy for epidemiological monitoring of AMR prevalence in Salmonella clones circulating in the human population.	Longitudinal sampling between Jan-Dec 2004 with 0.5-1 month interval. Up to 24 timepoints.	Ampicillin-resistant *Salmonella* spp. and *E. coli* (blaTEM, blaSHV, blaPSE-1, blaOXA-1 group).	Refrigerated (4°C) before filtration and processing.	All seasons - weather conditions not reported.
Pot et al. 2020	One health study to assess ESBL *Enterobacter cloacae* between WWTWs, humans and animals.	Longitudinal: Ten timepoints between April 2018 and March 2019.	ESBL *Enterobacter*.	Stored at 4°C and processed within 4H of collection.	All seasons - weather conditions not reported.
Rahimi et al. 2014	To determine epidemiological relatedness between MRSA strains isolated from wastewater and human infections.	Snapshot samples collected in 2010 - no other details.	MRSA (mecA).	Cold chain transport before filtration (0.45µm) and processing.	Not reported.
Raven et al. 2019	To evaluate whether the sequencing of *E. coli* isolated from WWTWs could support surveillance of pathogenic *E.*	Snapshot sampling between June 2014 and January 2015.	Third generation cephalosporin-resistant *E. coli* with non-limited genotype targeting.	Sodium thiosulphate-based preservation	Summer, autumn, winter - weather conditions

Citation	Study aims	Wastewater sampling strategy	Targeted organism and resistances	Wastewater handling and storage	Sampling season and weather reporting
	coli lineages and AMR genes in a specific region.			before filtration.	not reported.
Reinthaler et al. 2012	To assess transmission of antibiotic resistances from human sources into the environment.	Two sampling periods of six months each divided by ~nine years (five days between Apr - Sep 2000) and (six days between Apr - Sep 2009).	*E. coli* - no specific resistance targeted.	Kept refrigerated (4-8°C) for up to 24H before processing.	Spring, summer, autumn - weather conditions not reported.
Saifi et al 2009	To analyse and compare clonal diversity of gentamicin-resistant *E. faecium* in clinical and WWTW samples to assess release into the environment.	Longitudinal sampling with unclear interval between September 2005 - September 2006. Unclear timepoints.	Gentamicin-resistant *E. faecium* (6 ')-Ie-aph (2 '')-Ia, aph (2 '')-Ib , aph (2 '')-Ic, aph (2 '')-Id, ant (4 ')-Ia).	Filtration (0.45µm) - no other details reported.	Not reported.
Talebi et al. 2008	This study was undertaken to determine the genetic relatedness between the VRE isolates of wastewater and human urinary tract infection isolates.	Snapshot sampling during 2005 - no other details reported.	Vancomycin-resistant *E. faecium* (vanA and vanB).	Kept refrigerated before filtration (0.45 µm) and processing.	Not reported.
Urase et al. 2020	Comparison of ESBL *E. coli* and CRE in treated wastewater to national surveillance data on nosocomial infections.	Snapshot and longitudinal: One timepoint for 9 WWTWs, four timepoints for 1 WWTW and two timepoints for 1 WWTW - taken 2018-2019.	*E. coli* possessing ESBLs, and CREs.	Sampled collected in the morning and processed in the afternoon – no cool chain mentioned.	All seasons – samples taken to avoid rain events.
YoungKeun et al. 2015	To provide evidence for AMR wastewater surveillance for monitoring resistance trends and as an early warning system.	Longitudinal sampling with five timepoints over one year: Jan, Feb, July 2013, Jan 2014, Feb.	*E. coli* (blaCTX-M, blaTEM, blaSHV, blaNDM, blaKPC).	Kept at 4°C and analysed within 4H.	Winter, summer - weather conditions not reported.
Zaheer et al. 2020	To assess contributions of beef cattle to AMR in environmental niches and transmission chains to humans.	Longitudinal sampling over 24 months – frequency of visits not reported.	Enterococcus resistant to critically and highly important antibiotics.	Not reported.	All seasons - weather conditions not reported.
Zarfel et al. 2013	To investigate wastewater environmental samples to understand transmission of AMR to humans.	Longitudinal monthly sampling between Feb and July 2009. Six timepoints.	Third generation cephalosporin-resistant *E. coli* (blaTEM, blaSHV, blaCTX-M, blaVEB and blaGES).	Kept at 4–8°C for up to 24H before processing.	Winter, spring, summer - weather conditions not reported.

Figure S2 - AMR prevalence in wastewater and human isolates from low bias studies only

Concordance plot of AMR prevalence in wastewater and human isolates utilising data from low bias studies only. Each point represents a single wastewater-human comparison conducted where colour corresponds to bacterial species tested and shape corresponds to human sample type used. Lin's concordance correlation coefficient (CCC) is labelled with 95% confidence intervals. Unbroken line of $y=x$ is plotted as perfect concordance between wastewater and human resistance. Dashed lines of $y=x+0.1$ and $y=x-0.1$ represent high agreement, i.e. ±10% from perfect concordance respectively.

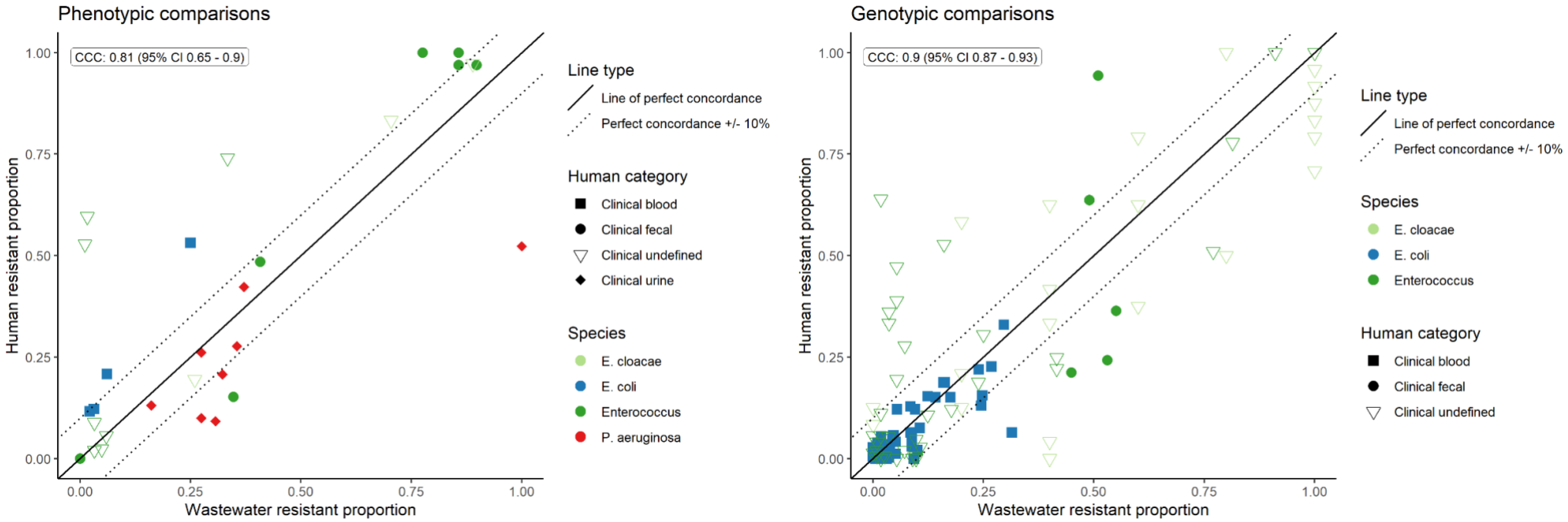


Figure S3 - Phenotypic aminoglycoside resistance prevalence estimates in human and wastewater isolates for aminoglycosides by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

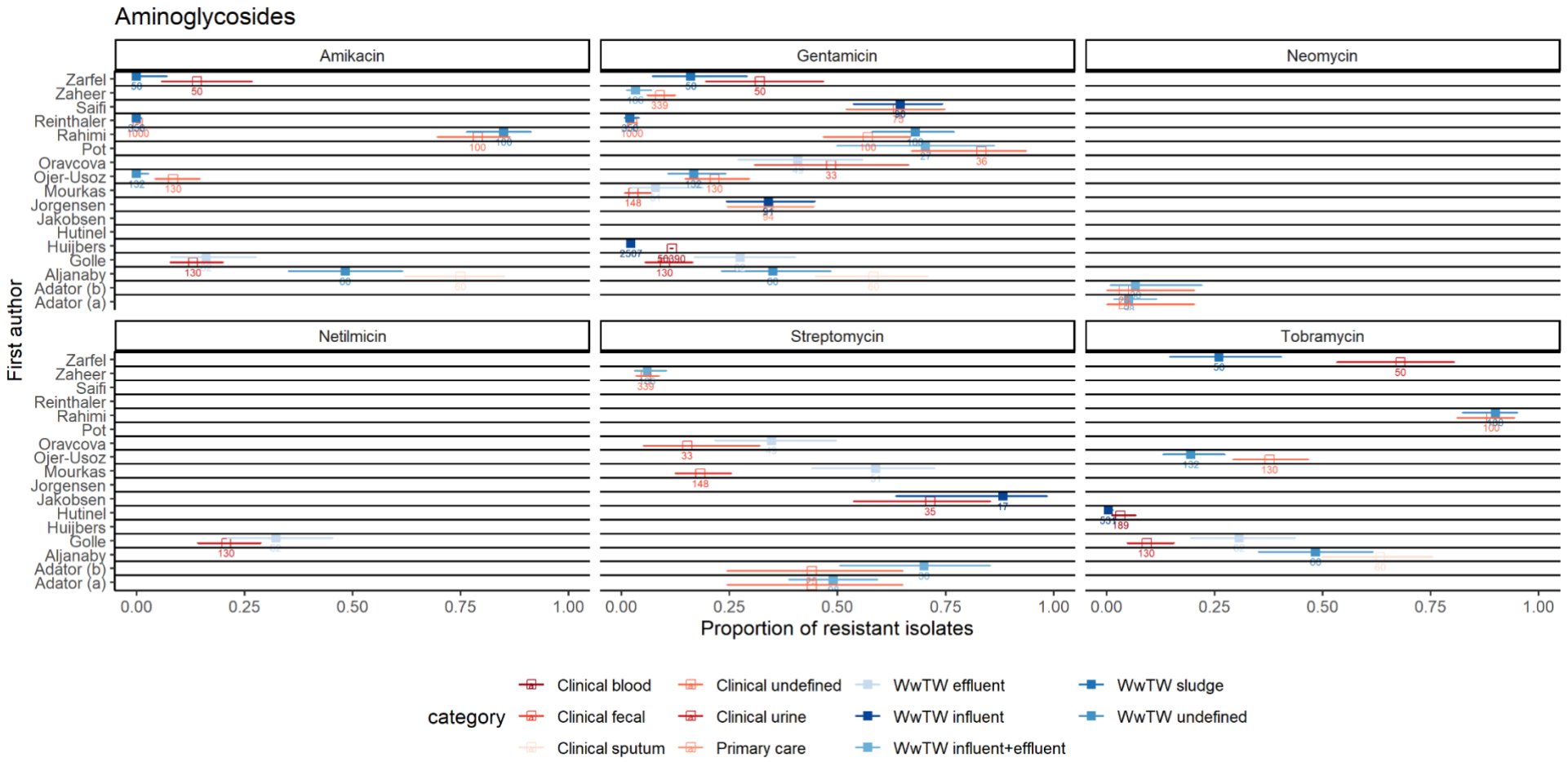


Figure S4 - Phenotypic resistance prevalence estimates in human and wastewater isolates for beta-lactams by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

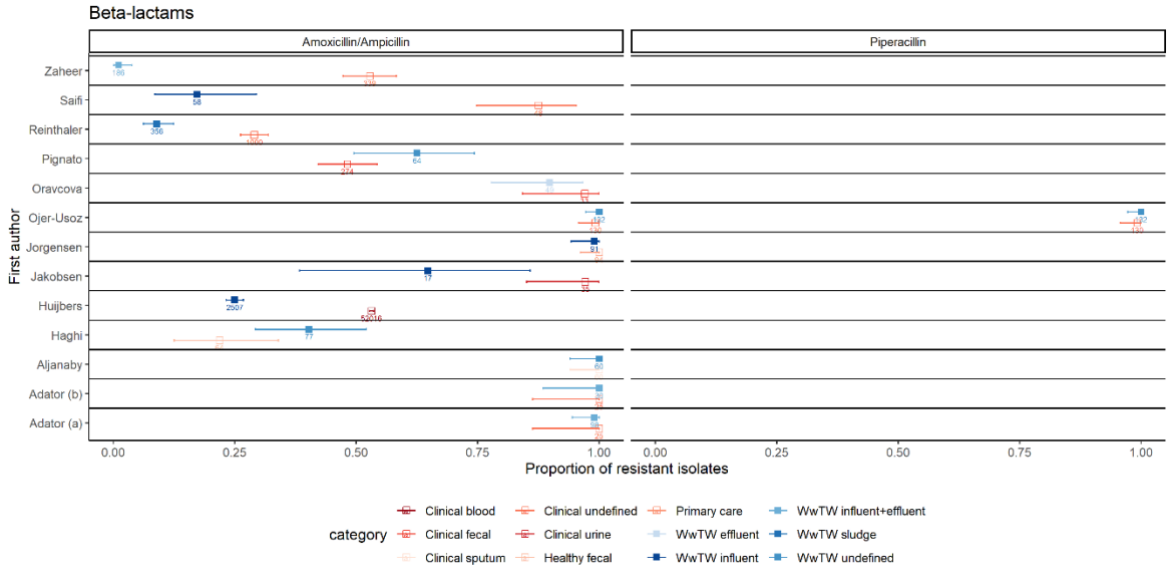


Figure S5 - Phenotypic resistance prevalence estimates in human and wastewater isolates for beta-lactam/beta-lactamase inhibitor combinations by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

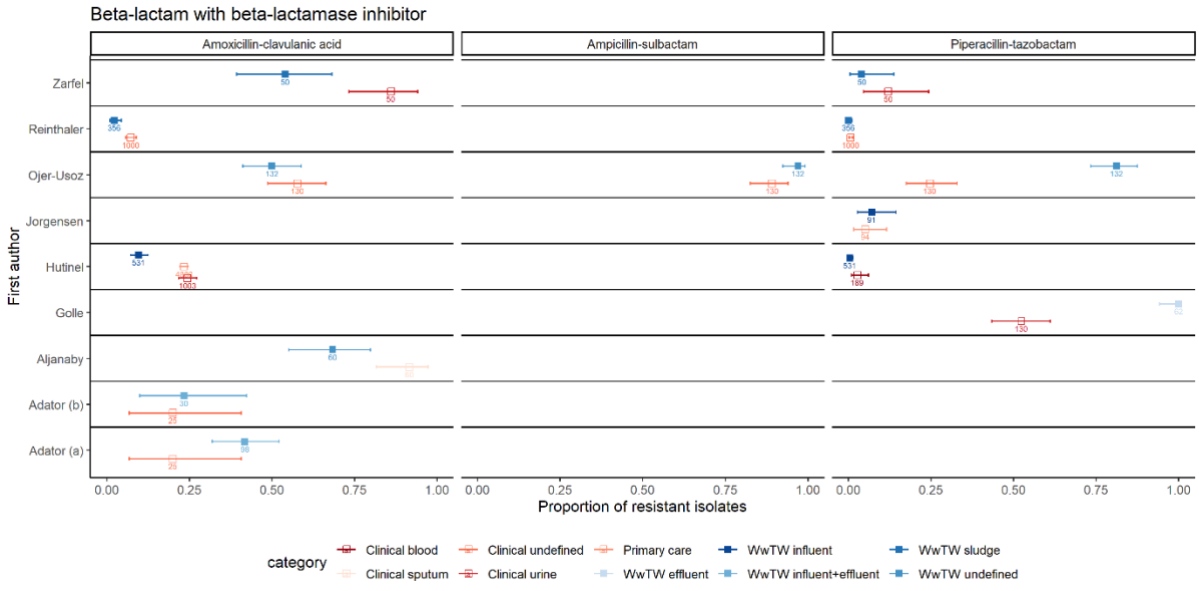


Figure S6 - Phenotypic resistance prevalence estimates in human and wastewater isolates for cephalosporins combinations by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

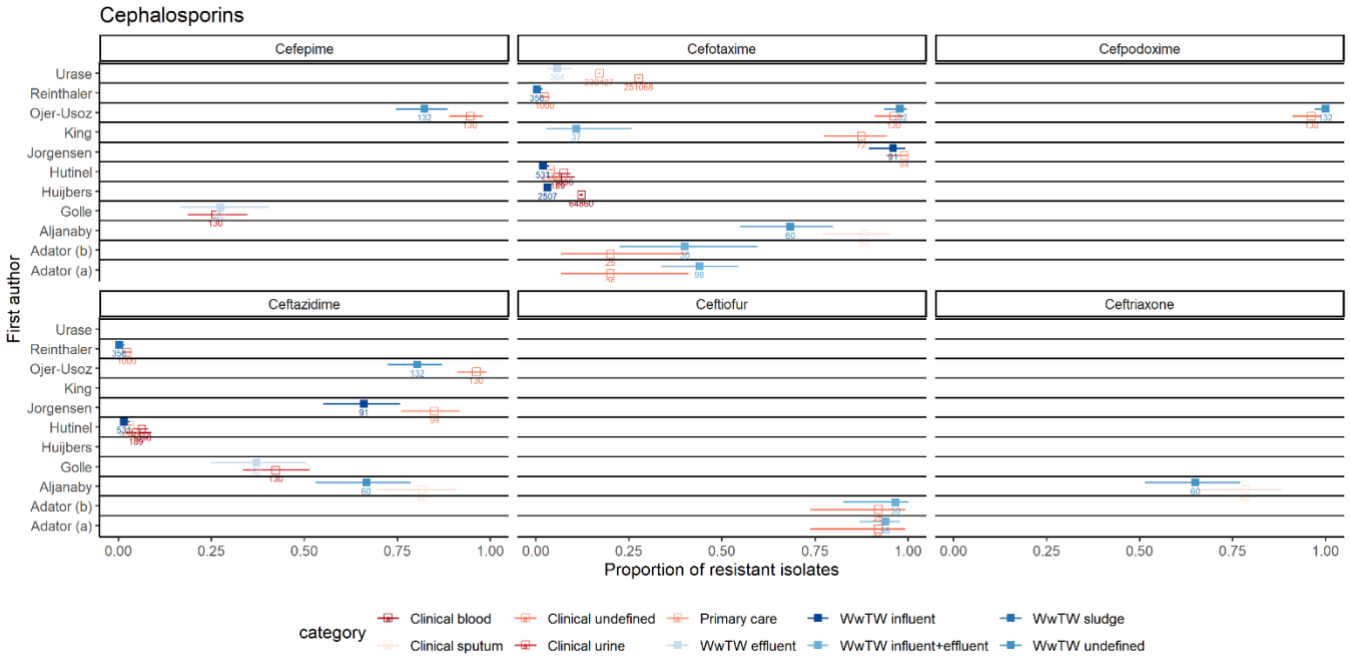


Figure S7 - Phenotypic resistance prevalence estimates in human and wastewater isolates for carbapenems by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

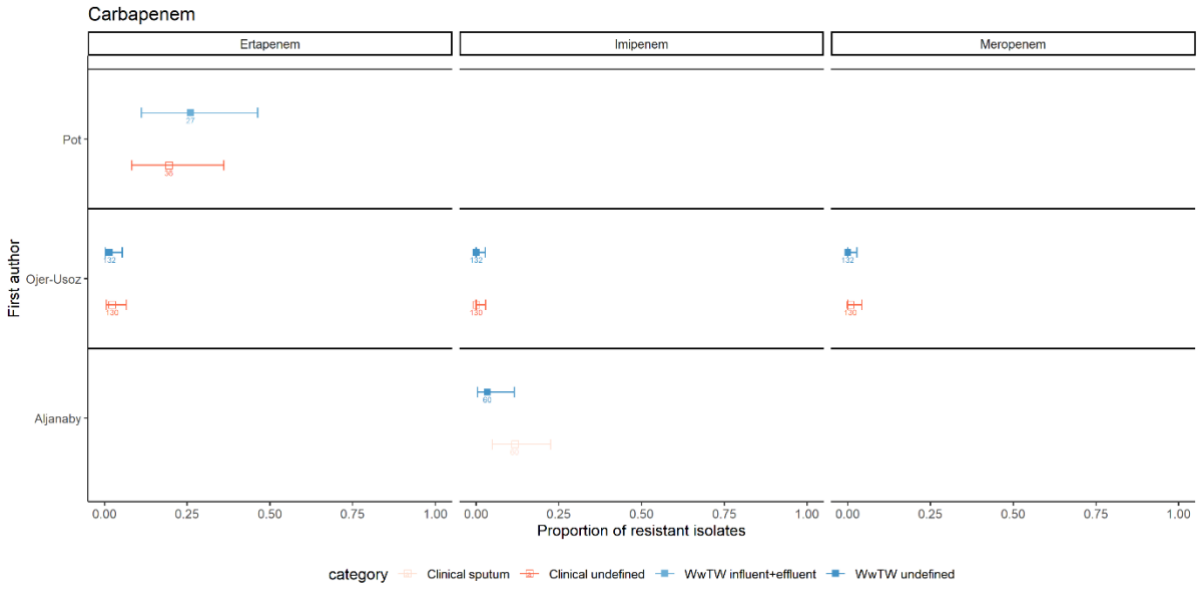


Figure S8 - Phenotypic resistance prevalence estimates in human and wastewater isolates for fluoroquinolones/quinolones by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

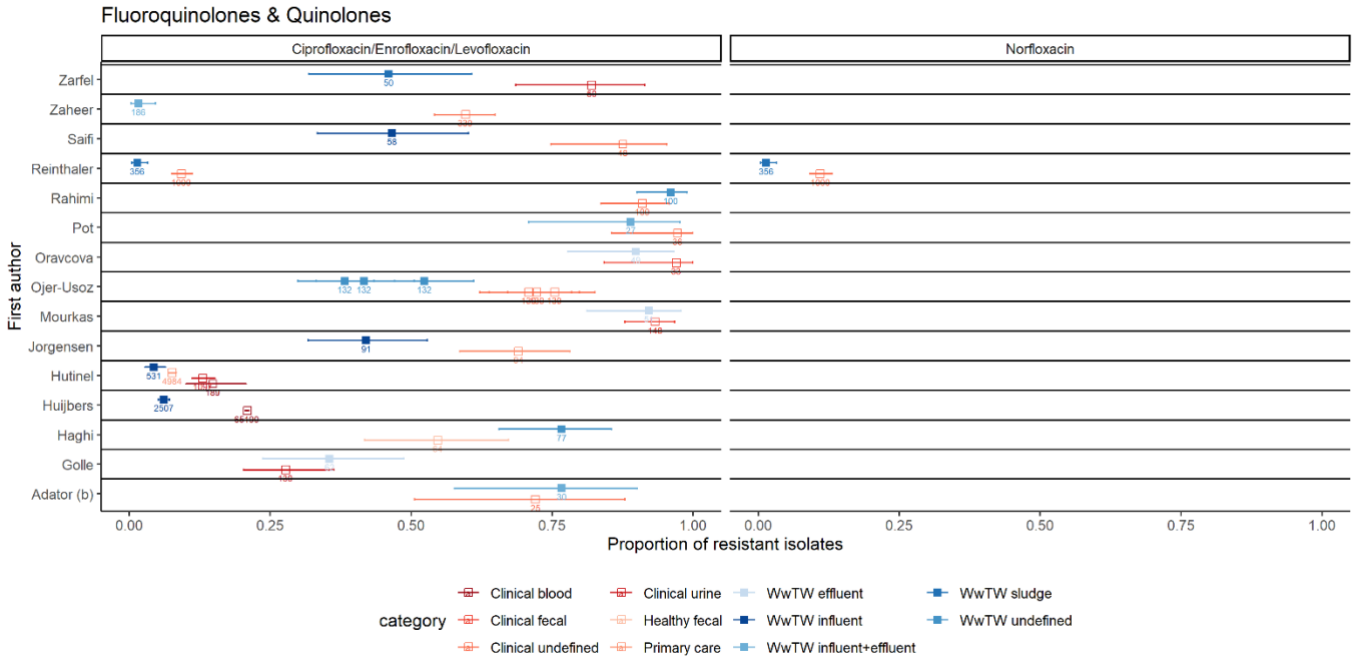


Figure S9 - Phenotypic resistance prevalence estimates in human and wastewater isolates for glycopeptides by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

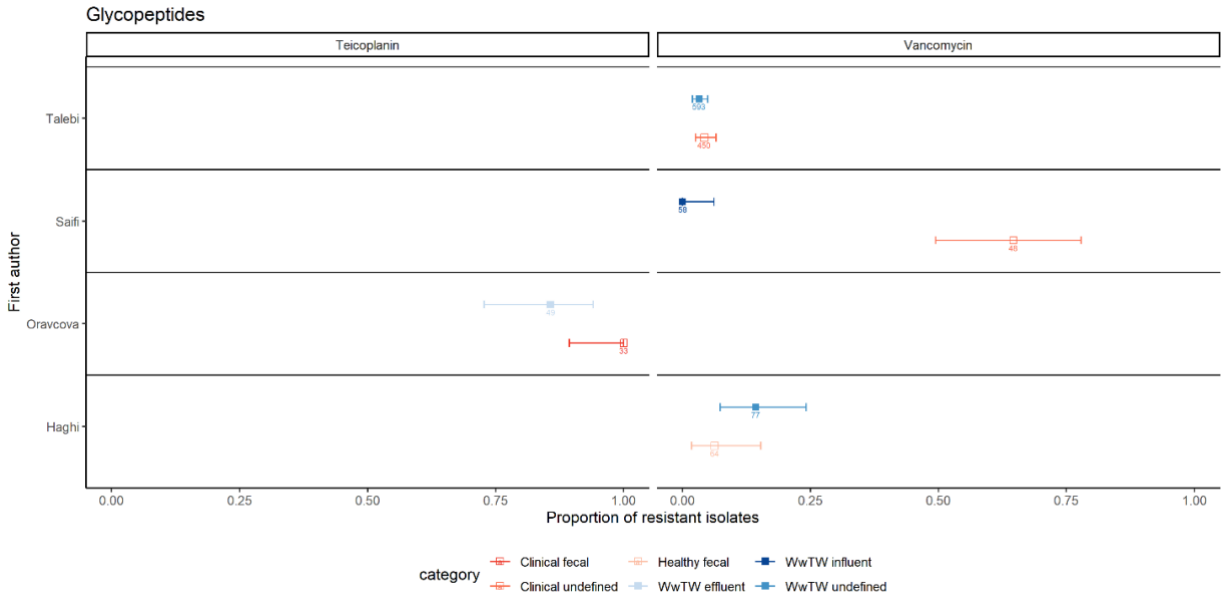


Figure S10 - Phenotypic resistance prevalence estimates in human and wastewater isolates for macrolides by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

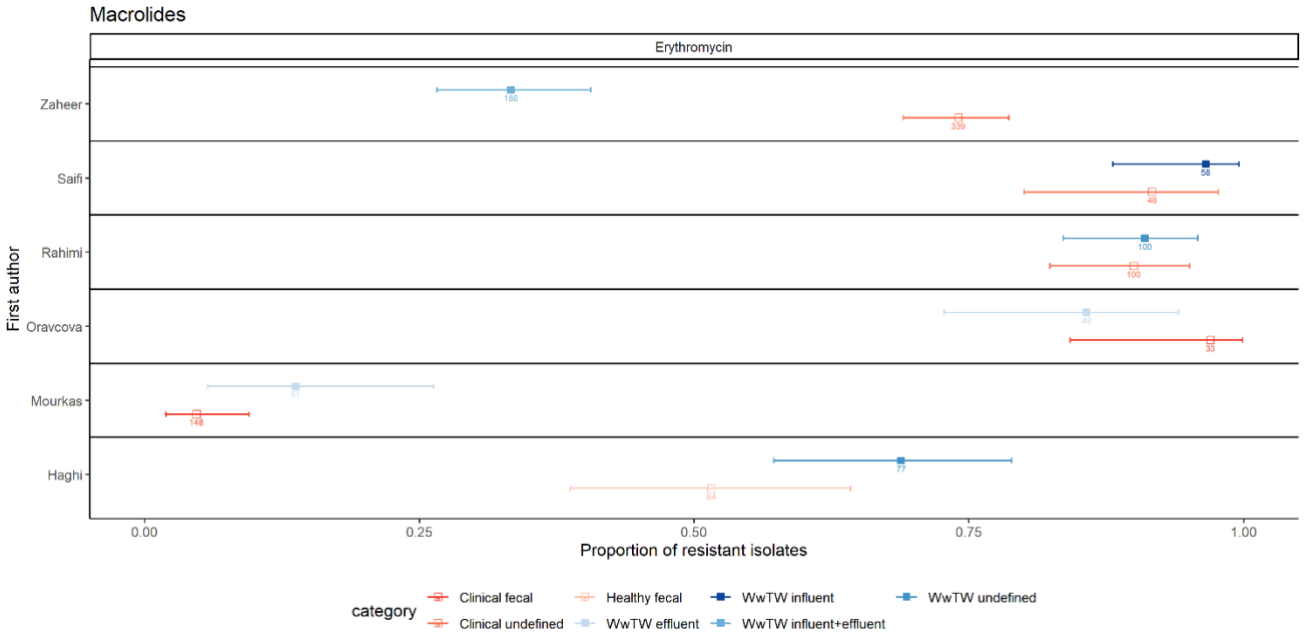


Figure S11 - Phenotypic resistance prevalence estimates in human and wastewater isolates for ESBLs and MRSA by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

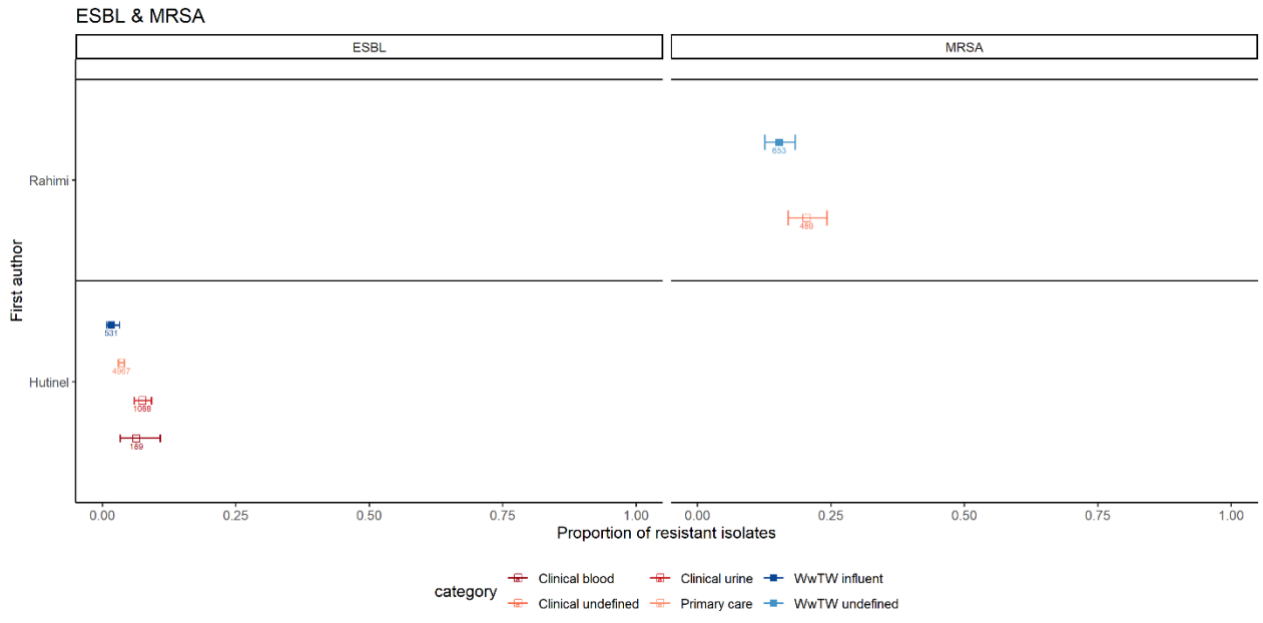


Figure S12 - Genotypic prevalence estimates in human and wastewater isolates for aac variants by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

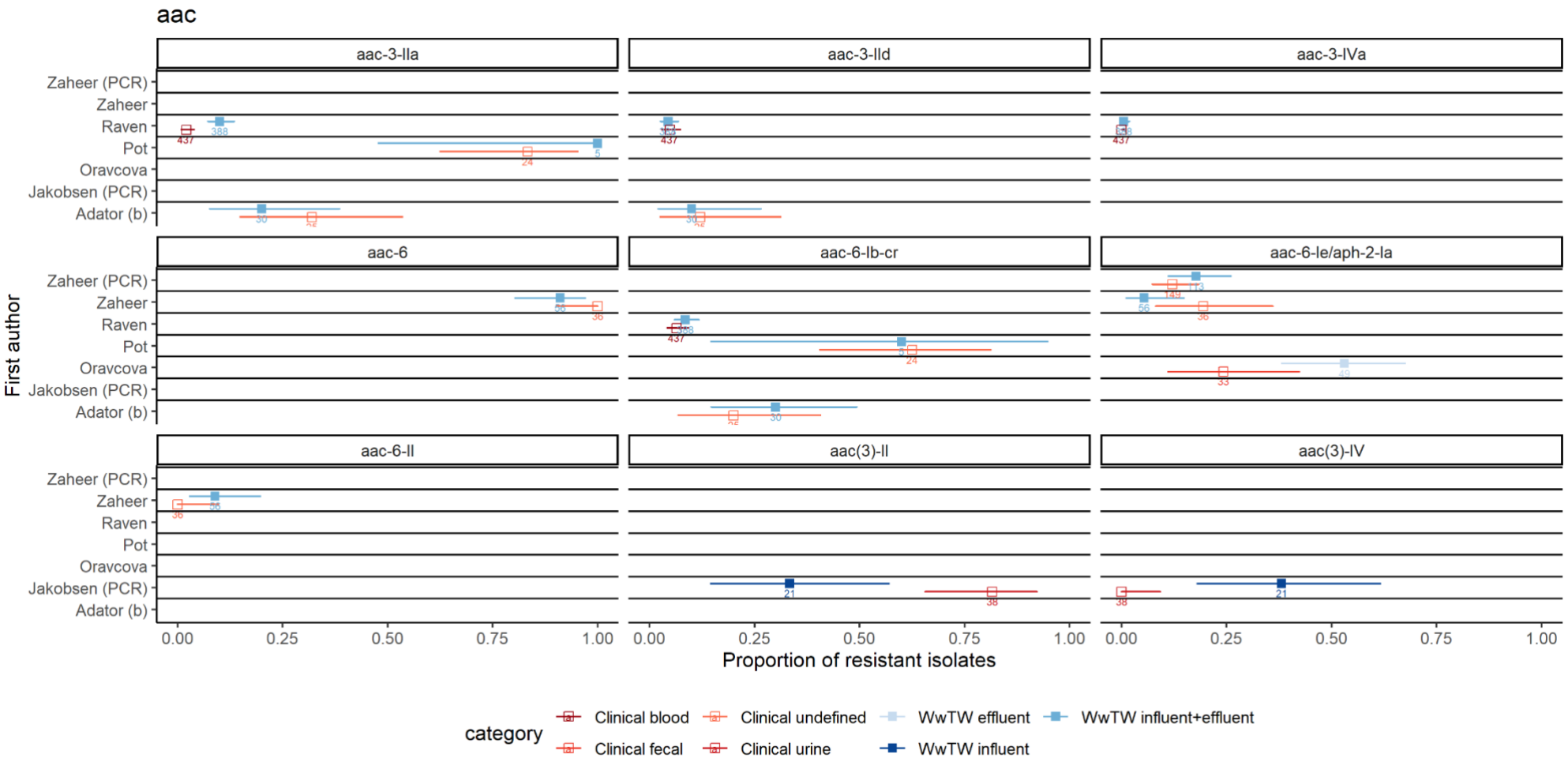


Figure S13 - Genotypic prevalence estimates in human and wastewater isolates for arr variants by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

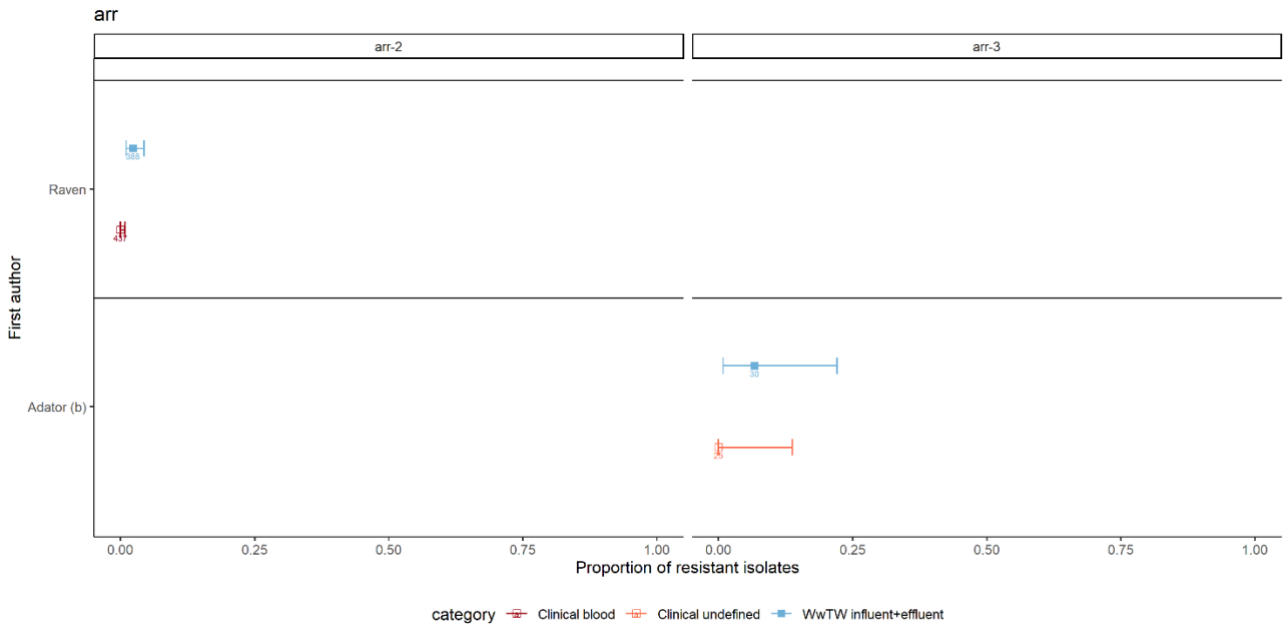


Figure S14 - Genotypic prevalence estimates in human and wastewater isolates for ctx variants by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

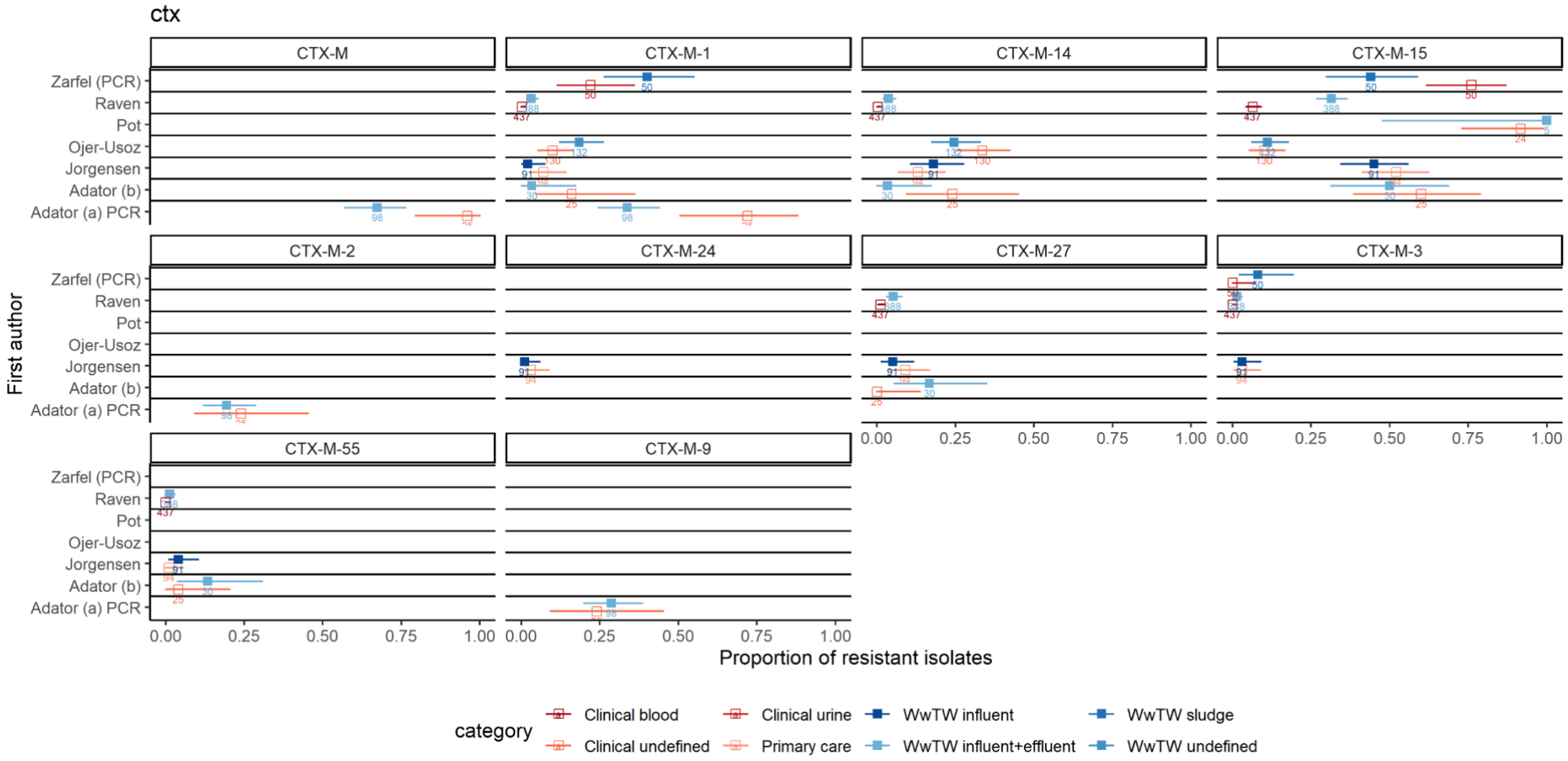


Figure S15 - Genotypic prevalence estimates in human and wastewater isolates for erm variants by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

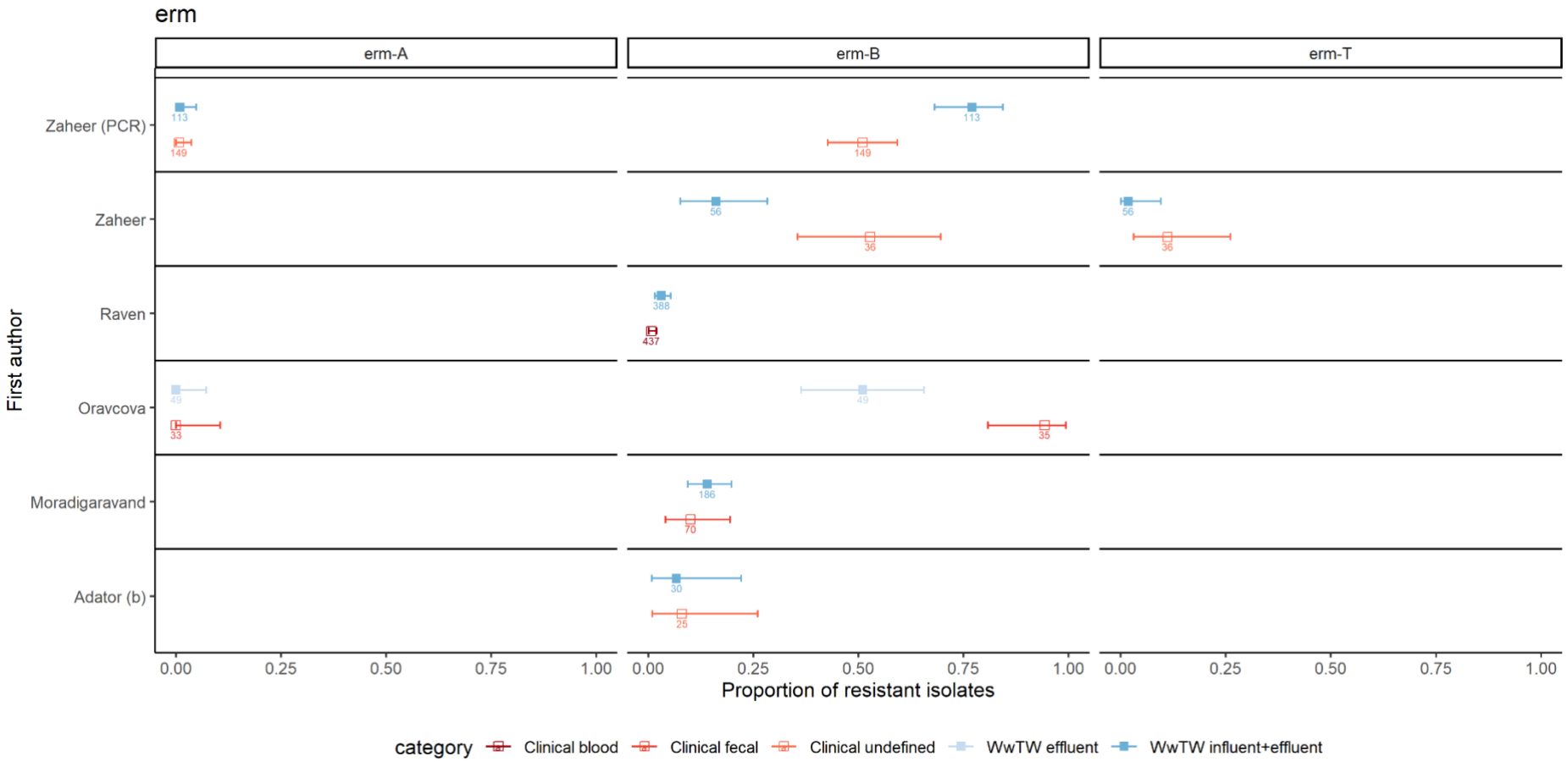


Figure S16 - Genotypic prevalence estimates in human and wastewater isolates for fos variants by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

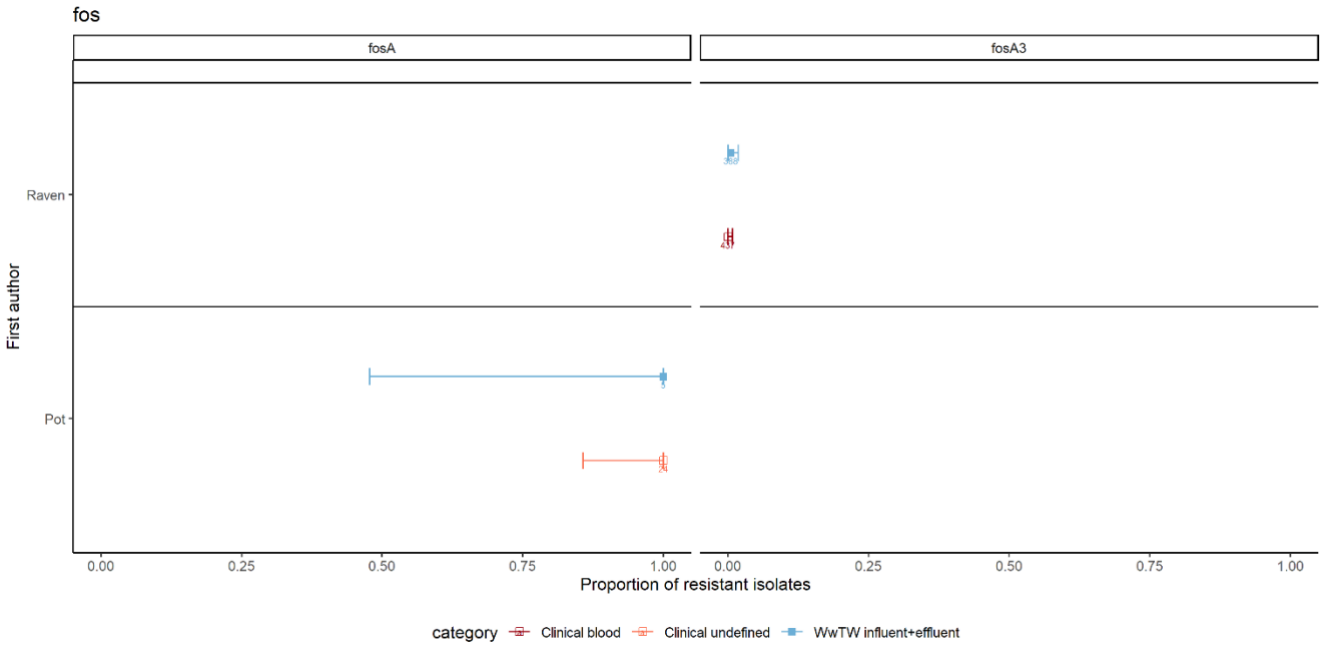


Figure S17 - Genotypic prevalence estimates in human and wastewater isolates for oxa variants by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

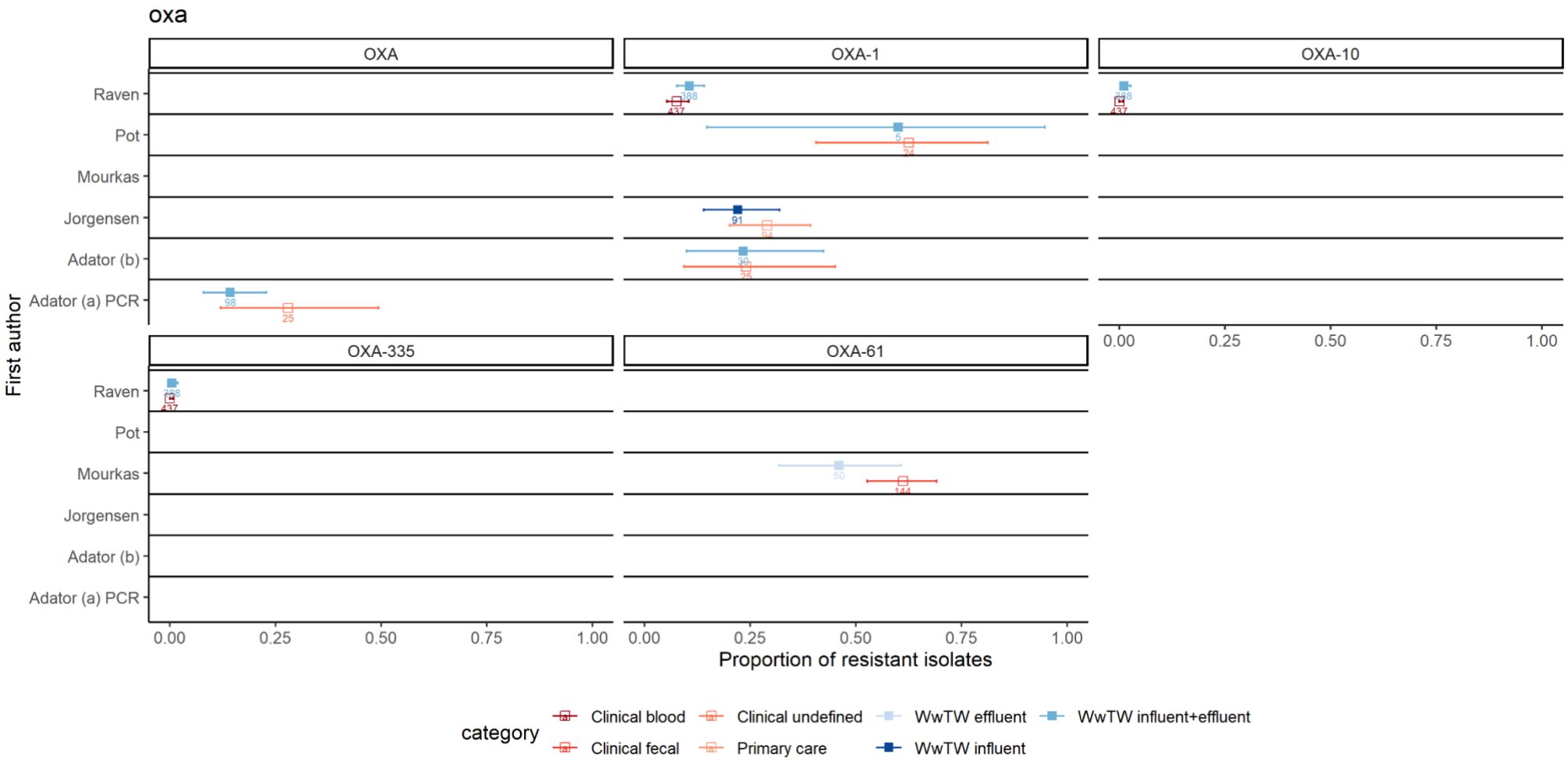


Figure S18 - Genotypic prevalence estimates in human and wastewater isolates for van variants by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

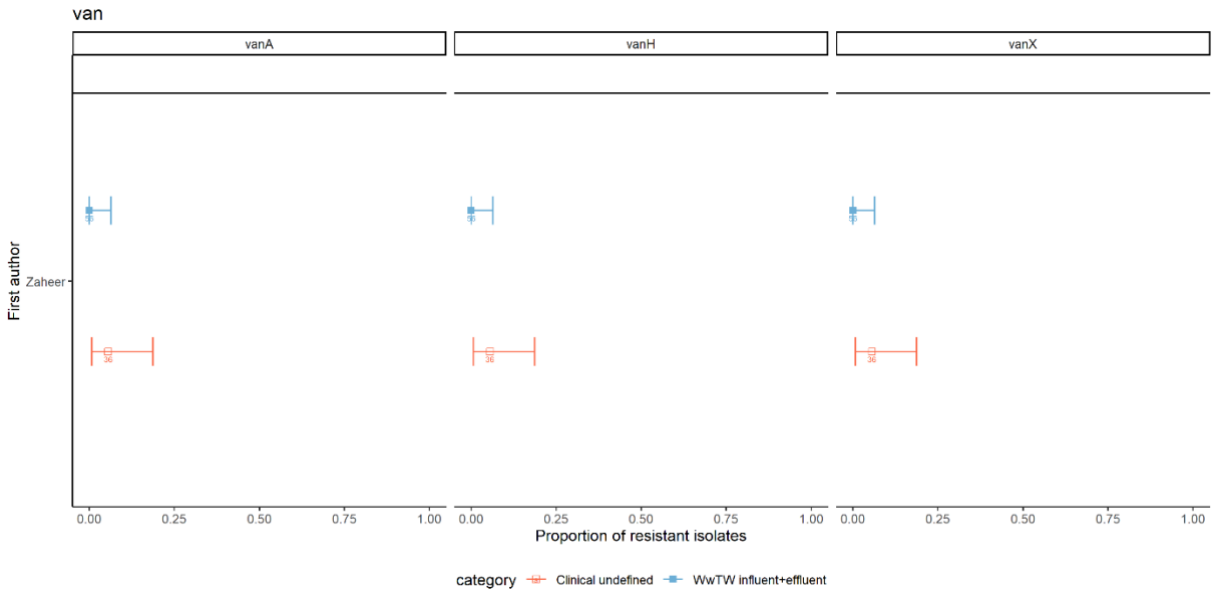


Figure S19 - Genotypic prevalence estimates in human and wastewater isolates for qnr variants by study

Point estimates with 95% confidence intervals plotted for human isolates (hollow squares coloured human sample category) and wastewater isolates (filled squares coloured by wastewater sample category). Total number of isolates analysed labelled under each estimate.

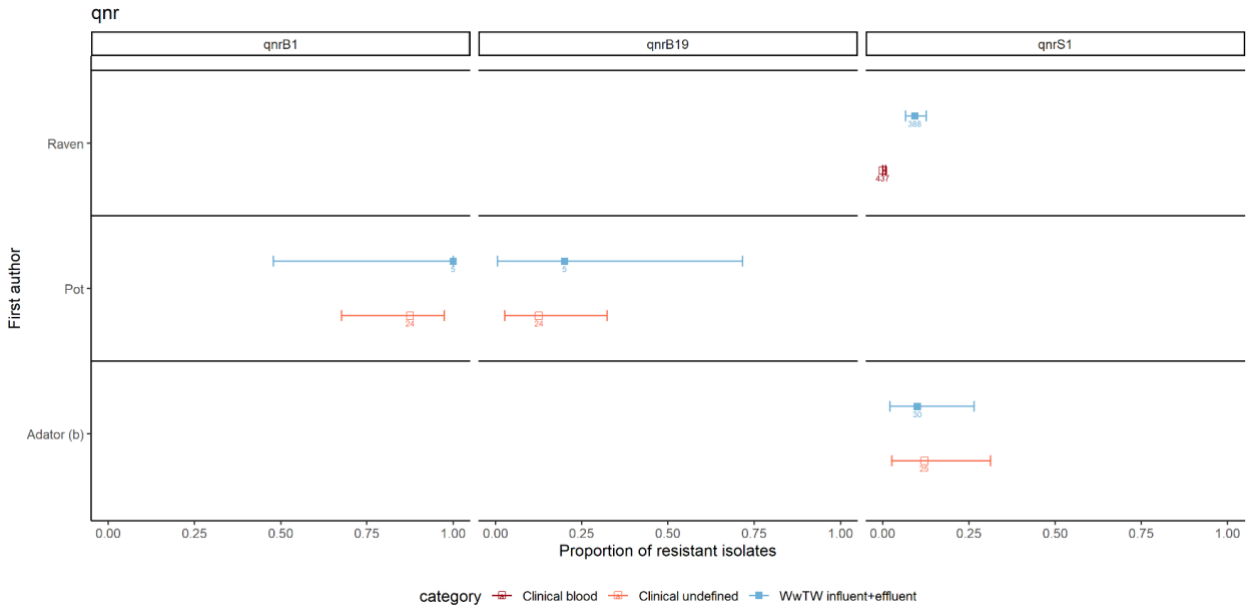


Table S6 - Univariable logistic regression of study features associated with high-agreement.

Study feature	n (% high) or IQR total (IQR high)	OR (95% CI)	p
Study design	24/24 reported	1.61 (0.34-7.69)	0.55
longitudinal	19/24 (32%)		
snapshot	4/24 (75%)		
both	1/24 (0%)		
Target species	24/24 reported	0.97 (0.69-1.38)	0.88
C. difficile	1/24 (100%)		
Campylobacter spp.	1/24 (0%)		
E. coli	11/24 (45%)		
E. cloacae	1/24 (0%)		
Enterococcus spp.	5/24 (40%		
Klebsiella spp.	1/24 (0%)		
P. aeruginosa	2/24 (0%)		
S. aureus	1/24 (100%)		
Salmonella spp.	1/24 (0%)		
AMR detection approach of extracted data	24/24 reported	2.59 (0.62-10.80)	0.19
genotypic	2/24 (100%)		
phenotypic	12/24 (33%)		
mixed	10/24 (30%)		
Human sample type[1]	24/24 reported	0.76 (0.43-1.35)	0.35
clinical blood	2/24 (50%)		
clinical faecal	4/24 (25%)		
clinical sputum	1/24 (0%)		
clinical urine	3/24 (0%)		
clinical unspecified	10/24 (50%)		
faecal carriage	1/24 (0%)		
primary care	3/24 (67%)		
Number of human isolates analysed	24/24 reported	0.99 (0.99-1.00)	0.56
IQRs	46.5-314.75 (70-450)		
Region sampled income classification	24/24 reported	0.53 (0.09-2.83)	0.45
high	17/24 (41%)		
medium	6/24 (33%)		
low	1/24 (0%)		
Number of wastewater isolates analysed	24/24 reported	1.00 (0.99-1.00)	0.92
IQRs	50.75-205.5 (91-388)		
Pre-defined surveillance study	24/24 reported	1.14 (0.15-8.59)	0.89
yes	5/24 (40%)		
no	19/24 (37%)		

Study feature	n (% high) or IQR total (IQR high)	OR (95% CI)	p
Number of WWTWs sampled	23/24 reported (1 missing)	1.09 (0.93-1.27)	0.29
IQRs	1-5 (1-5)		
Wastewater sampling method	20/24 reported (4 missing)	0.76 (0.34-1.68)	0.49
grab	9/24 (33%)		
composite	6/24 (67%)		
flow-proportional	3/24 (67%)		
time-flow-volume-proportional	1/24 (0%)		
moore swab	1/24 (0%)		
not reported	4/24 (0%)		
Wastewater sampled has a direct input from the human population surveyed	18/24 reported (6 missing)	0.44 (0.07-2.69)	0.38
yes	11/24 (45%)		
no	6/24 (33%)		
both	1/24 (0%)		
not reported	6/24 (33%)		
Wastewater sampling point	17/24 reported (5 missing)	0.93 (0.38-2.26)	0.87
influent	5/24 (40%)		
effluent	4/24 (0%)		
influent & effluent	8/24 (57%)		
sludge	2/24 (50%)		
not reported	5/24 (40%)		

[1]Primary care was removed from testing due to the uncertainty to which this category belongs in respect to other sample types. Sensitivity analysis including primary care showed no significant difference (p=0.363)

Table S7 - Study features excluded from logistic regression due to missingness

Study feature	n (% high) or IQR total (IQR high)
Wastewater samples undergone freeze-thawing?	14/24 reported (10 missing)
no	14/24 (36%)
not reported	10/24 (40%)
Presence of unique sewer inputs	12/24 reported (12 missing)
agriculture	2/24 (100%)
healthcare	7/24 (29%)
healthcare and industry	2/24 (0%)
no unique sewer inputs	1/24 (100%)
not reported	12/24 (33%)
Wastewater storage temperature	10/24 reported (14 missing)
4°C	10/24 (40%)
not reported	14/24 (36%)
Wastewater sample time-to-processing (hours)	8/24 reported (16 missing)
IQRs	3-15 (8.25-18.75)
not reported	16/24 (44%)
WWTW population equivalent	7/24 reported (17 missing)
IQRs	11722-186837.5 (10261-140867)
not reported	17/24 (41%)
WWTW sampling point flow rate (m^3 day)	6/24 reported (18 missing)
IQRs	1987.2-51861.6 (5464.8-54086.4)
not reported	18/24 (44%)
Weather conditions during wastewater sampling	2/24 reported (22 missing)
specifically avoiding rain events	1/24 (0%)
mixed but mostly during rain events	1/24 (0%)
not reported	22/24 (38%)

Table S8 - Tally of features and agreement

		Overall study agreement		
	Study features	<30%	30-70%	>70%
Sampling point	influent	2	3	2
	influent & effluent	2	5	5
	effluent	1	4	1
	sludge	1	1	1
	not reported	2	2	2
Study design	longitudinal	7	14	8
	snapshot	0	1	3
	both	1	0	0
Longitudinal timeframe	<6 months	2	1	1
	6-12 months	5	3	1
	>12 months	1	8	6
	not reported	0	2	0
	NA (snapshot studies)	0	1	3
Sampling method	grab	4	4	4
	composite	1	3	5
	flow-proportional	2	1	2
	flow-proportional+	0	1	0
	moore swab	0	2	0
	not reported	1	4	0
Comparison type	direct	4	8	6
	indirect	1	4	3
	both	1	0	0
	unclear	2	3	2
Sewer AMR input	agriculture	0	1	3
	hospital	2	5	2
	hospital & industry	2	0	1
	none	0	0	1
	not reported	4	8	5
Number of WWTWs	1	3	6	4
	2	2	4	3
	3-10	2	3	2
	>10	1	0	2
	not reported	0	2	0
Human category	clinical blood	0	1	1
	clinical urine	3	2	0
	clinical sputum	1	0	0
	clinical faecal	1	3	2
	clinical undefined	1	8	6
	primary care	1	1	2
	healthy faecal	1	0	0
Surveillance study	yes	2	1	2
	no	6	14	9

3. Evaluation of metagenomic, 16S rRNA gene and ultra-plexed PCR-based sequencing approaches for profiling antimicrobial resistance gene and bacterial taxonomic composition of polymicrobial samples

Note: This chapter has previously been published as a preprint[91]. I authored the text and figures below (which are reproduced here largely unaltered from the published work) with input from supervisors commensurate with the amount of input/advice that would be considered appropriate for a DPhil thesis. All authors have agreed to the inclusion of this published work in my thesis.

3.1 Background

Shotgun DNA metagenomic characterisation of polymicrobial samples is increasingly used in both clinical and environmental studies, facilitating agnostic sequencing of all DNA present in a sample and enabling flexible comparisons with reference databases to determine sample composition[92]. For example, the RefSeq[93] and CARD databases[94] can be used for taxonomic and antimicrobial resistance (AMR) gene (i.e. "resistome") profiling respectively. This approach has underpinned multiple studies characterising microbiome-disease associations[95], evaluating community diversity and anthropogenic impacts[62], and investigating AMR[96,97]. There is also growing interest in the use of shotgun metagenomics to profile wastewater for population-level surveillance of AMR (i.e. wastewater-based epidemiology [WBE])[39,41]. However, shotgun metagenomics can be expensive, and bioinformatic deconvolution of the data challenging, especially when

using short-read sequencing and trying to characterise the genetic context of specific loci, such as AMR genes[90]. Sequencing depth impacts sensitivity to detect rare targets of interest, and therefore, most studies involve a sequencing depth-sensitivity trade-off where uncommon sequences may lack coverage or be missed completely. Additionally, for many studies, only a fraction of the metagenome (e.g. bacterial sequences) will be of interest, and in this context, much of the wider metagenome represents wasted sequencing effort; for example human DNA in studies analysing clinical samples for pathogen diagnostics[92].

Amplicon-based approaches enrich specific nucleic acid templates during library preparation to increase sensitivity for relevant but less abundant target sequences at reduced cost. A commonly used single target for profiling bacterial communities is the 16S rRNA gene which possesses both highly conserved regions facilitating the use of universal primers, and hypervariable regions which can be used to discriminate between bacterial taxa[98,99]. Increasingly, targeted ultra-highly multiplexed PCR panels have been developed to enable amplicon-based evaluation of the presence/absence and diversity of specific features of a metagenome associated with a key phenotype, for example AMR gene diversity[100-102], or of a subset of organisms associated with disease, for example respiratory viruses. One example is Illumina AmpliSeq™ which can be used with both Illumina- and community-curated panels (primer pools). These large panels enable capture of a broader range of targets than universal primers or multiplex PCR, whilst theoretically optimising sensitivity and sequencing resource, and are attractive to researchers as they come with defined laboratory protocols and user-friendly

bioinformatic pipelines. However, despite internal validation, Illumina-curated panels have demonstrated mispriming events leading to false positive calls of targets; the community-curated panels have largely not undergone any internal validation. Additionally, targeted methods may be prone to amplification biases, and by their nature only focus on specific features of the wider metagenome. Differences in reference database curation and mapping approaches between available bioinformatic pipelines may also influence results[103,104].

Combining targeted approaches such as 16S rRNA gene sequencing and ultra-highly multiplexed PCR/amplicon sequencing of AMR gene targets represents a potential alternative to deep shotgun metagenomics for profiling species and AMR gene diversity in polymicrobial samples. I therefore assessed the performance of shotgun metagenomics, versus 16S rRNA gene sequencing and the AmpliSeq for Illumina Antimicrobial Resistance Research Panel (henceforth "AmpliSeq") in reconstituting the true taxonomic composition and resistome of a well-defined DNA mock microbial community. I also compared several bioinformatics approaches to characterising species and AMR gene profiles, including an in-house approach to AmpliSeq data analysis (AmpliSeek). With performance on the mock community as a reference, I then applied the best approach to seven untreated wastewater samples to quantify performance in the context of "real-world" sample complexity. I aimed to highlight potential limitations of each method and recommend potential use cases.

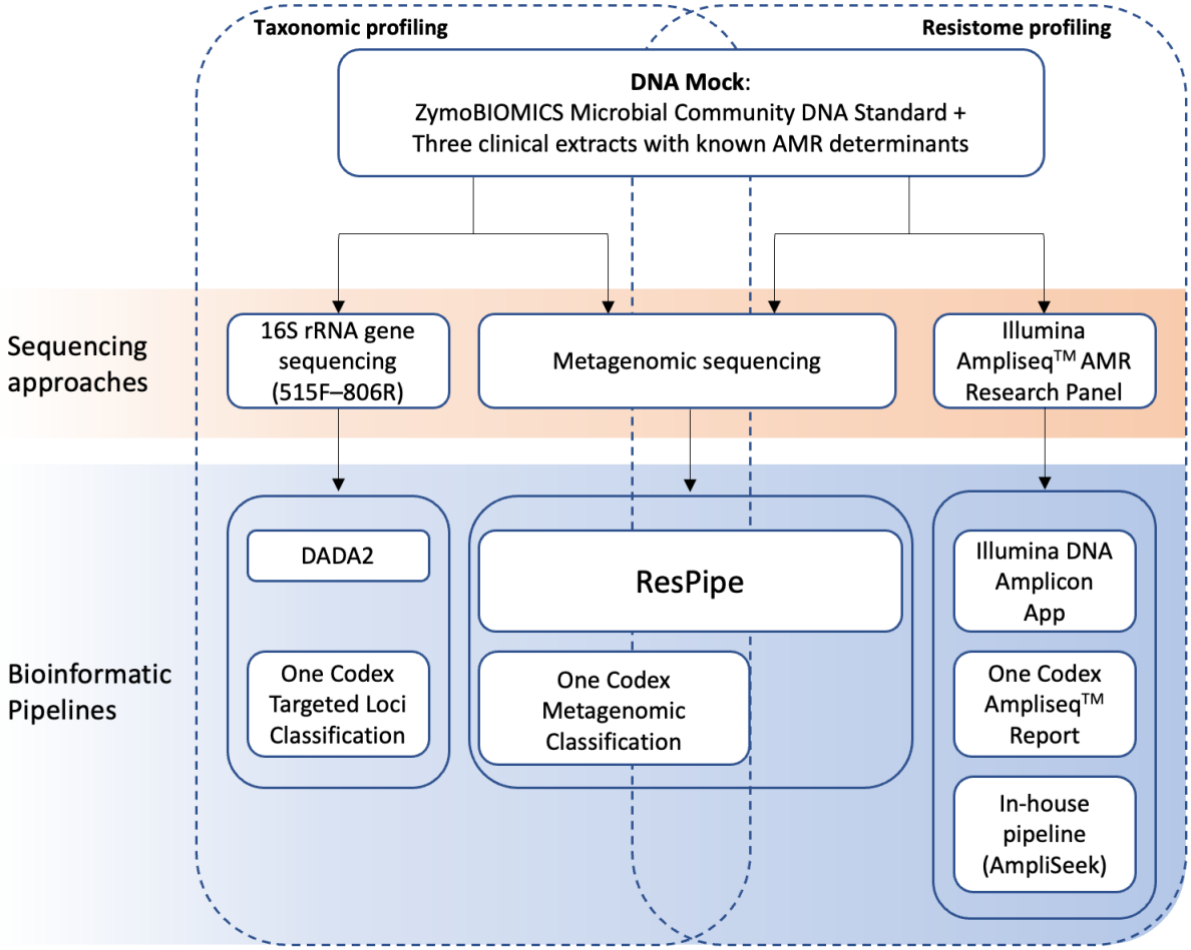
3.2 Methods

3.2.1 DNA mock community preparation and wastewater samples

Metagenomic, 16S rRNA gene and AmpliSeq sequencing were all conducted on aliquots from the same mock DNA sample (Figure 5). The mock was prepared by combining the ZymoBIOMICS Microbial Community DNA Standard (Zymo Research Corporation, Irvine, USA) with three bacterial isolate DNA extracts chosen to enrich the mock sample for clinically relevant AMR genes; these bacteria had been characterised with whole genome sequencing as part of a previous study evaluating human bloodstream infections[105].

Wastewater samples comprised seven metagenomic DNA extracts from wastewater influent collected as part of local surveillance in Oxfordshire, UK, 2019, and stored as pellets at -80°C. Metagenomic DNA was extracted using the PowerSoil kit (QIAGEN, Hilden, Germany) according to manufacturer protocols.

Figure 5 - Schematic overview of evaluated sequencing approaches and respective bioinformatic pipelines tested for taxonomic and resistome profiling of the DNA mock.



3.2.2 Sequencing

Shotgun metagenomic sequencing on mock DNA community was conducted by Novogene Co (Beijing, China) on the NovaSeq6000 (Illumina, San Diego, USA), generating 150 bp paired-end (PE) reads with a target sequencing depth of 20 million reads (6Gb). AmpliSeq was conducted using the community-designed Illumina for AmpliSeq AMR Research Panel and the Library PLUS kit with manufacturer guidelines; libraries were pooled and sequenced on the Illumina MiniSeq (150 bp PE reads) with replicate libraries

prepared for the mock sample. 16S rRNA gene sequencing utilised 515F–806R primers for library preparation and libraries were sequenced on the Illumina MiSeq (250 bp PE reads) as previously described[106].

Wastewater samples underwent sequencing as described above, however, shotgun sequencing was conducted by the Wellcome Trust Centre for Human Genetics (Oxford, UK) with a depth of ~75 million reads (~23Gb) per sample based on previous deep sequencing and rarefaction analyses demonstrating this was the minimum sequencing depth required to capture most AMR gene diversity in this sample type from the same sampling site[90].

Raw sequencing data are available under ENA project PRJEB52722 (https://www.ebi.ac.uk/ena/browser/view/PRJEB52722)[107]. R scripts and markdown renders are available at https://github.com/KaibondChau/Chau_etal_Ampliseek[108]. The AmpliSeek pipeline is available at https://github.com/KaibondChau/ampliseek[109].

3.2.3 Software

I tested seven bioinformatic pipelines (Figure 5, 6). Two pipelines were compared for metagenomic-based taxonomic profiling (ResPipe v1.4.0 and One Codex Metagenomic classification v8/13/2021), two for 16S rRNA-based taxonomic profiling (DADA2 v1.16 [assignTaxonomy] and One Codex Targeted Loci classification v4/15/2021), and three for AmpliSeq-based resistome profiling (Illumina BaseSpace DNA Amplicon App v0.7.12, One Codex AmpliSeq Report v1/17/2019 and an in-house BBTools wrapper pipeline [AmpliSeek] developed as part of this study available at https://github.com/KaibondChau/ampliseek). AmpliSeek performs trimming (BBDuk2) and merging (BBMerge) of reads produced by AmpliSeq sequencing before stringent mapping to AmpliSeq panel target sequences (BBMapSkimmer). Reference databases used by each pipeline are detailed in Figure 6.

Geneious Prime v2021.2.2 was used for detailed *in silico* characterisation of the mock AMR profile by visualising the relevant AMR reference sequences and reads mapping to these sequences for each methodological approach.

Figure 6 - Bioinformatic pipeline versions and associated reference databases

Sequencing approach	Profiling strategy	Bioinformatic pipeline	Reference database
Shotgun metagenomics	Taxonomy	ResPipe v1.4.0	RefSeq "Complete genome" assembly level sequences v91
		One Codex Metagenomic classification v8/13/2021	One Codex Genomes database
	Resistome	ResPipe v1.4.0	CARD v3.03
16S rRNA	Taxonomy	One Codex Targeted Loci classification v4/15/2021	One Codex Targeted Loci database
		DADA2 v1.16	SILVA SSU
AmpliSeq	Resistome	Illumina BaseSpace DNA Amplicon App v0.7.12	Illumina AmpliSeq panel reference genomes and manifest target sequences v1.0
		One Codex AmpliSeq Report v1/17/2019	One Codex-curated AmpliSeq marker sequences
		In-house pipeline (AmpliSeek)	Illumina AmpliSeq panel manifest target sequences v1.0

3.2.4 Taxonomic profile analysis

I used two previously described error metrics[110] to quantify differences between theoretical and pipeline-reported distribution of bacterial genera present in the mock sample, a modified mean absolute proportion error (MAPE) based on theoretical vs reported read differences, and Bray-Curtis dissimilarity[111] which further considers the total richness present. Both MAPE and Bray-Curtis scores range between 0 and 1, where 0 represents identical composition. For ease of interpretation, scores are presented as 1-MAPE and 1-Bray-Curtis, where higher values represent better accuracy.

Since no theoretical truth existed when characterising wastewater, diversity metrics were used to compare differences between sequencing-pipeline combinations. Chao1 was used as an estimate of taxonomic richness by representing total count of unique observed taxa. Pielou's evenness was used to assess how similarly represented observed taxa were in abundance estimates (constrained to [0,1] with low values representing unbalanced community estimates where few taxa constitute the majority of abundance). Shannon index was used to quantify both richness and evenness as a measure of overall community complexity where higher values represent increased entropy and therefore complexity. A two-tailed Welch paired *t* test (allowing for different variances in the two samples) was used to determine significance between differences in diversity metrics.

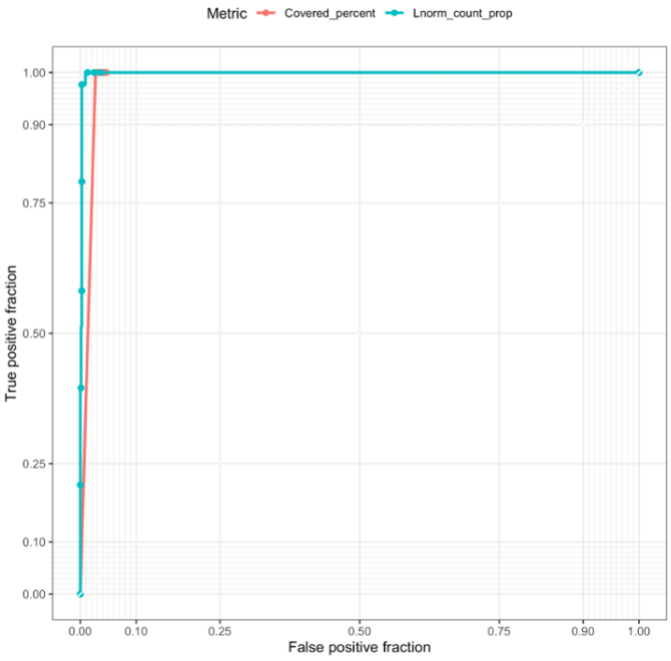
3.2.5 Sensitivity and specificity calculations for AMR gene detection

The mock DNA community was annotated with AMR target sequences present in the metagenomic and AmpliSeq reference databases using 100% identity threshold to

determine a "strict" *in silico* truth. These annotations were compared against pipeline calls using scoring thresholds defined by each pipeline (i.e. thresholds at which AMR targets are called present). For AmpliSeek, scoring thresholds were optimised to maximise sensitivity and specificity based on the "strict" *in silico* truth using receiver operating characteristic curves and Youden's index[112] (Figure 7). Established stringent scoring thresholds for AMR target presence were used for ResPipe (lateral coverage=1)[90] and the One Codex AmpliSeq Report (coverage ≥ 85% and identity ≥ 95%)[113].

Figure 7 - AmpliSeek scoring thresholds

Left: Receiver operator characteristic curve for mock AmpliSeek output metrics in relation to true positives and false positives based on the known "strict" *in silico* truth. Top right: Optimal thresholds as determined by Youden's index. Lnorm_count_prop denotes the proportion of length-normalised count for a specific AmpliSeek target. Bottom right: Combined thresholds to determine target status.



Metric	Value	Specificity	Sensitivity
Covered_percent	99.77	0.97	1
Lnorm_count_prop	0.000427	0.99	1

Target status	AmpliSeek thresholds
Present	Covered_percent => 99.77 AND Lnorm_count_prop =>0.000427
Absent	Covered_percent < 99.77 AND Lnorm_count_prop < 0.000427
Potentially present	Covered_percent => 99.77 OR Lnorm_count_prop =>0.000427

Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for AMR target detection were calculated by categorising AMR genes called by each method as true positive, true negative, false positive and false negative, when compared with the *in silico* truth (see above). By adapting this *in silico* truth with respect to the reference AMR gene database used for each method, I fairly compared overall methodological performance (i.e. the CARD v3.0.3 database of 2605 sequences for metagenomic AMR gene profiling approaches and the AmpliSeq reference panel of 815 [AmpliSeek] or 814 [One Codex] amplicon sequences which target 478 AMR genes). However, AmpliSeek scoring thresholds were optimised to the mock dataset and therefore AmpliSeek performance is not directly comparable to the DNA amplicon app and One Codex Ampliseq Report where scoring thresholds were not optimised to the mock dataset.

Mock DNA sequences were also annotated at 98% and 90% identity thresholds to determine “intermediate” and “relaxed” *in silico* truth respectively to investigate AmpliSeq false positives, i.e. to include more AMR gene matches in the truth set as present where these were similar to a gene in the reference database.

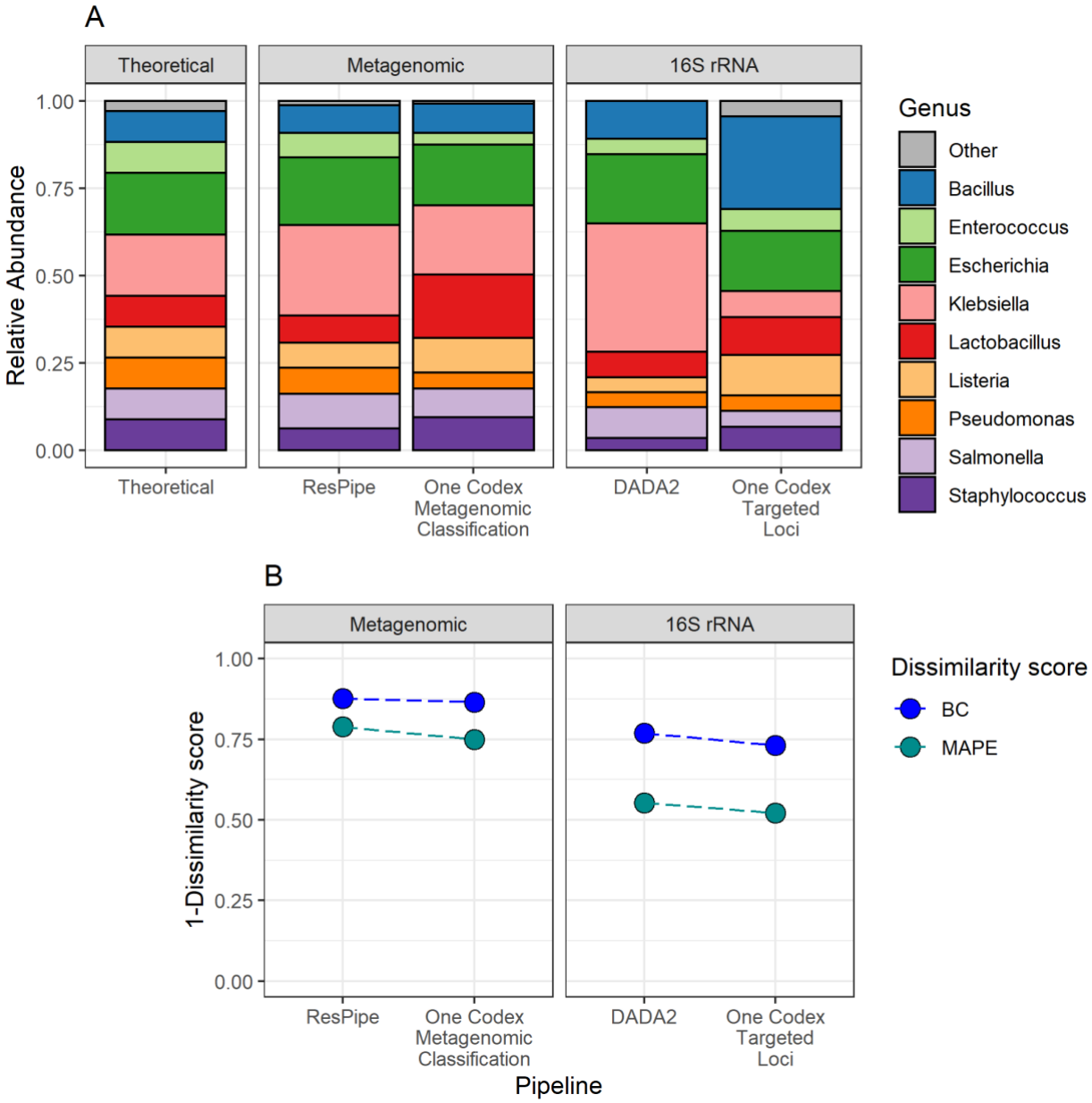
3.3 Results

3.3.1 Taxonomic profiling

All bacterial genera present in the mock DNA community were correctly detected by each sequencing method-pipeline combination, but the relative abundance of different genera was variably under- and overestimated for each combination (Figure 8A). This was most notable for 16S rRNA gene-based classification of *Klebsiella* which was underestimated by One Codex Targeted Loci classifier and overestimated by DADA2. *Bacillus* and *Lactobacillus* were also inaccurately estimated for One Codex Targeted Loci and Metagenomic estimates respectively. Regardless of pipeline, shotgun metagenomics more accurately reconstituted true taxonomic composition over 16S rRNA gene-sequencing (Figure 8B; mean 1-MAPE=0.77 vs 0.54, and mean 1-Bray-Curtis=0.87 vs 0.75). Differences in performance between pipelines within each dataset was small, with ResPipe (1-MAPE:0.79; 1-Bray-Curtis:0.88) marginally outperforming One Codex metagenomic classifier (0.75; 0.86) for shotgun sequencing data, and DADA2 (0.55; 0.77) outperforming the One Codex Targeted Loci classifier 0.52; 0.73) for 16S rRNA gene sequencing data.

Figure 8 - (A) Theoretical taxonomic composition of DNA mock as compared to actual generated pipeline outputs faceted by sequencing approach.

Genus "Other" represents genera present at <1% relative abundance in pipeline outputs. For the theoretical facet, "Other" represents DNA composition from yeasts included in the commercial DNA standard. (B) Taxonomic profiling error scores represented by 1-Bray-Curtis and 1-MAPE where 1.00 is perfect reconstitution of theoretical genera abundances.



3.3.2 AMR gene profiling

Annotation of the mock DNA sequences with approach-specific reference databases identified 77 CARD and 43 AmpliSeq target sequences present at 100% similarity. These were all detected with 100% sensitivity by all respective methods/pipelines (Figure 9), except the DNA Amplicon App, which performed poorly (22/43 known AmpliSeq targets detected; 51%). For precision, ResPipe (specificity=0.99; PPV=0.72) performed best, followed by AmpliSeek (0.98; 0.69), DNA Amplicon App (0.98; 0.65) and One Codex AmpliSeq Report (0.94; 0.49) (Figure 9B,C). However, notably, I found that most false positive hits obtained (67/75; 89%) arose from ≤2% nucleotide variation from reference AMR sequences (AmpliSeek [18/19], DNA Amplicon App [12/12] and One Codex AmpliSeq Report [37/44]) (Figure 10). AmpliSeq false positives were investigated by comparing pipeline calls to "intermediate" and "relaxed" mock *in silico* truth (see Methods section 3.2.5). I considered AmpliSeq reference sequences present in the mock with 98-100% identity as target variants (Figure 10) and assessed these separately to account for the limited sequence diversity amongst reference sequences for specific targets. An additional 37 reference sequences were identified in the mock at 98-100% identity. When these target variants were included in overall performance metrics, One Codex AmpliSeq Report retained 100% sensitivity with increased specificity and PPV (0.99; 0.92). However, AmpliSeek and DNA Amplicon App were less able to detect target variants, with overall sensitivities of 0.76 and 0.43 respectively. However, AmpliSeek sensitivity improved to 0.90 (Figure 10 - grey fill) when including less stringent scoring thresholds (i.e. "potentially detected" calls).

Figure 9 - (A) Detection sensitivity of known AMR genes present in the mock at 100% identity.

NB the reference AMR gene database used to annotate the mock was specific to, and hence different for, each approach. Error bars represent 95% confidence intervals (CIs), and facets divide sequencing approach and bioinformatic pipelines. (B) Specificity of AMR gene detection presented as for (A) and with magnified y-axis in (C).

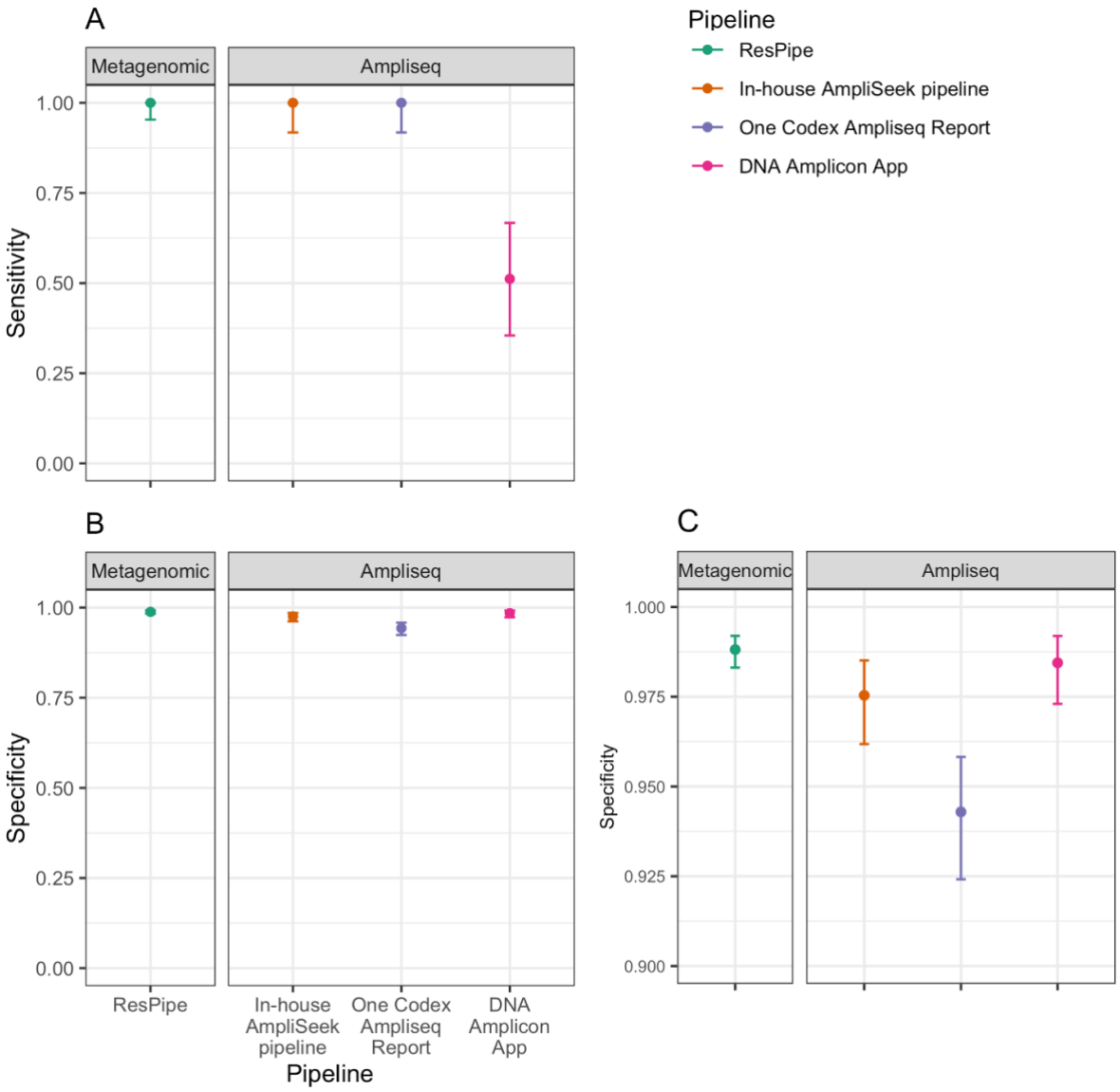
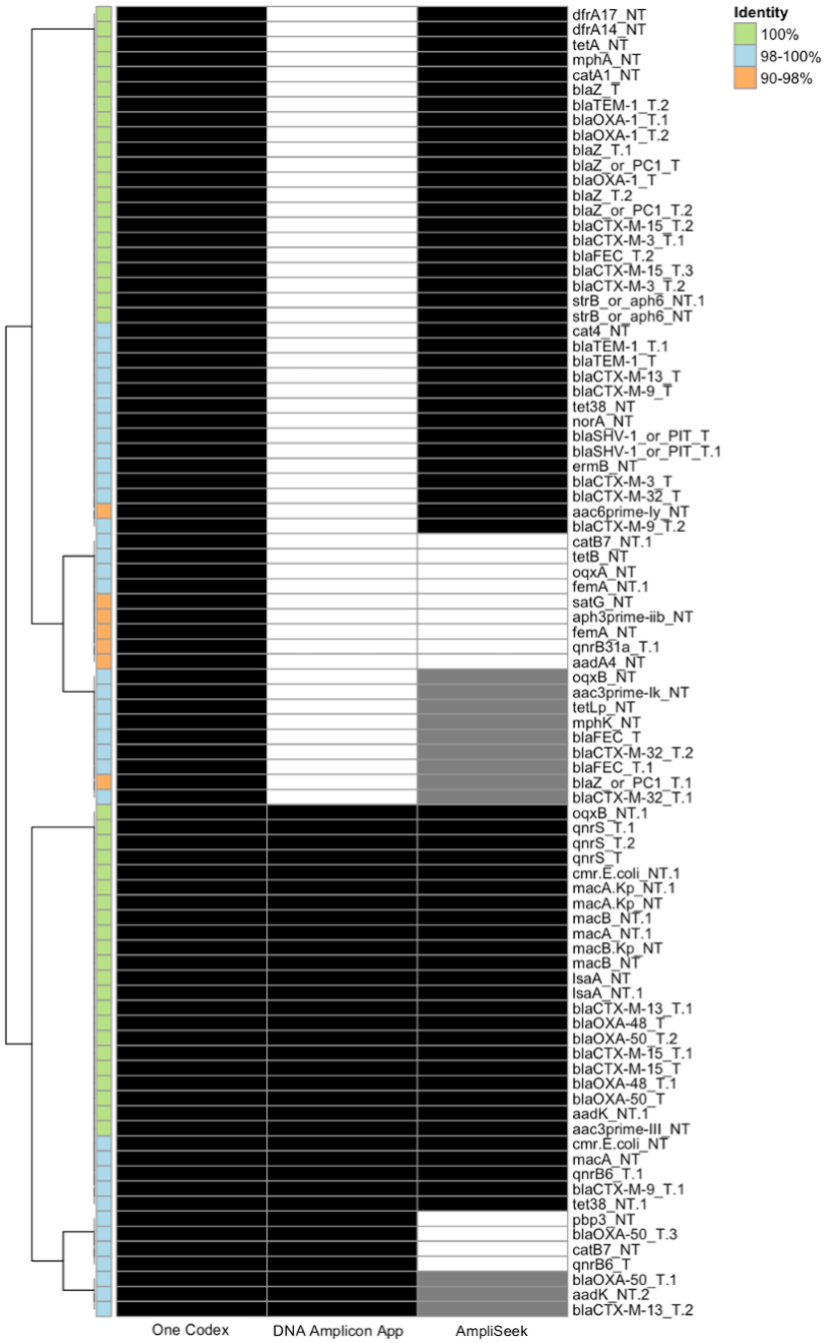


Figure 10 - Heatmap of AmpliSeq pipeline calls labelled by reference marker ID and sequence identity.

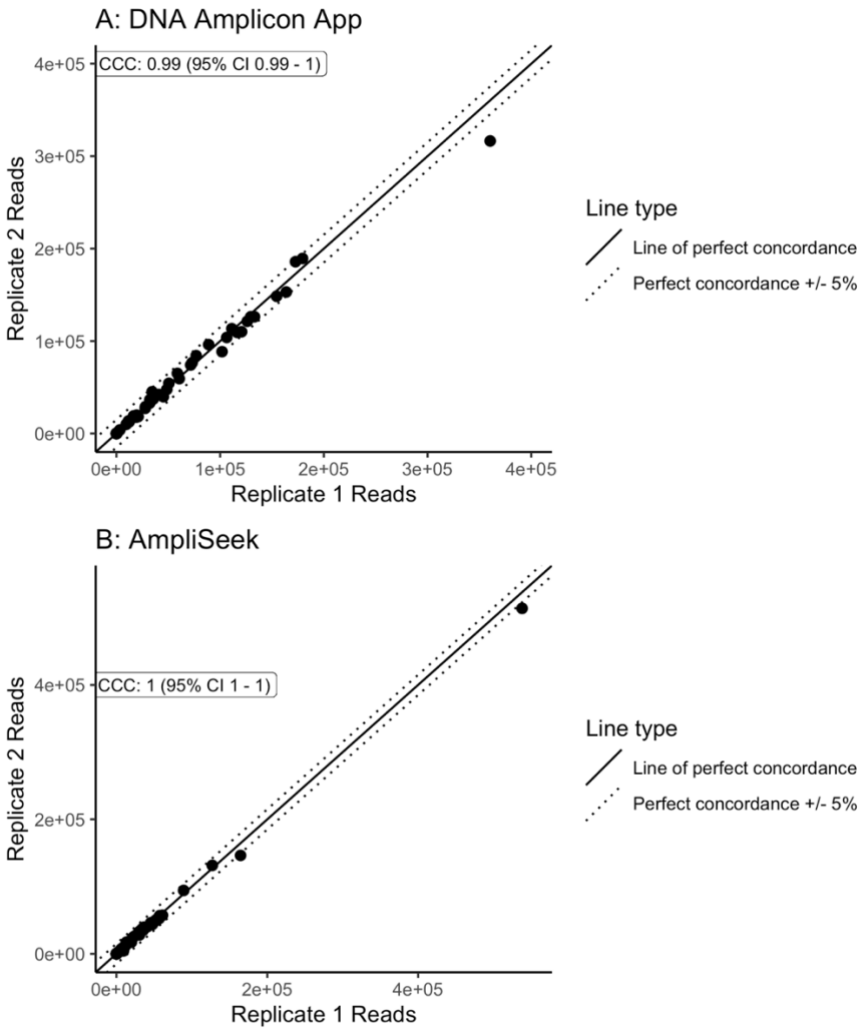
Each column represents a bioinformatic pipeline output where black and grey fill reflect detected and potentially detected calls respectively. Labels (green, blue, orange) denote the sequence identity of reference markers when annotated in the mock DNA community.



Replicate AmpliSeq libraries produced highly consistent results with no impact on performance metrics for all pipelines tested (Figure 11).

Figure 11 - Replicate results of AmpliSeq profiling.

Two aliquots of mock community DNA were independently prepared for AmpliSeq to assess replicability of ultra-high multiplexed PCR and library prep. Read counts for AMR targets between replicates were highly concordant for both the DNA Amplicon App (Lin's concordance correlation coefficient [CCC]=0.99) (panel A) and in-house AmpliSeek (CCC=1.00) (panel B). Interestingly, the read difference per AMR target between replicates was reduced in AmpliSeek (median=502; IQR: 14-2162) compared to the DNA Amplicon App (median=1746; IQR: 304-5466). Disagreements (i.e. reads mapping to a target in only one replicate) were rare for both DNA Amplicon App (n=2) and AmpliSeek (n=1), and associated to very low read counts (≤2). Variations in read count or target coverage did not impact overall performance metrics for presence/absence calls for either pipeline. Read-level data was not available from the One Codex AmpliSeq report, but overall performance metrics did not differ for One Codex replicates either.

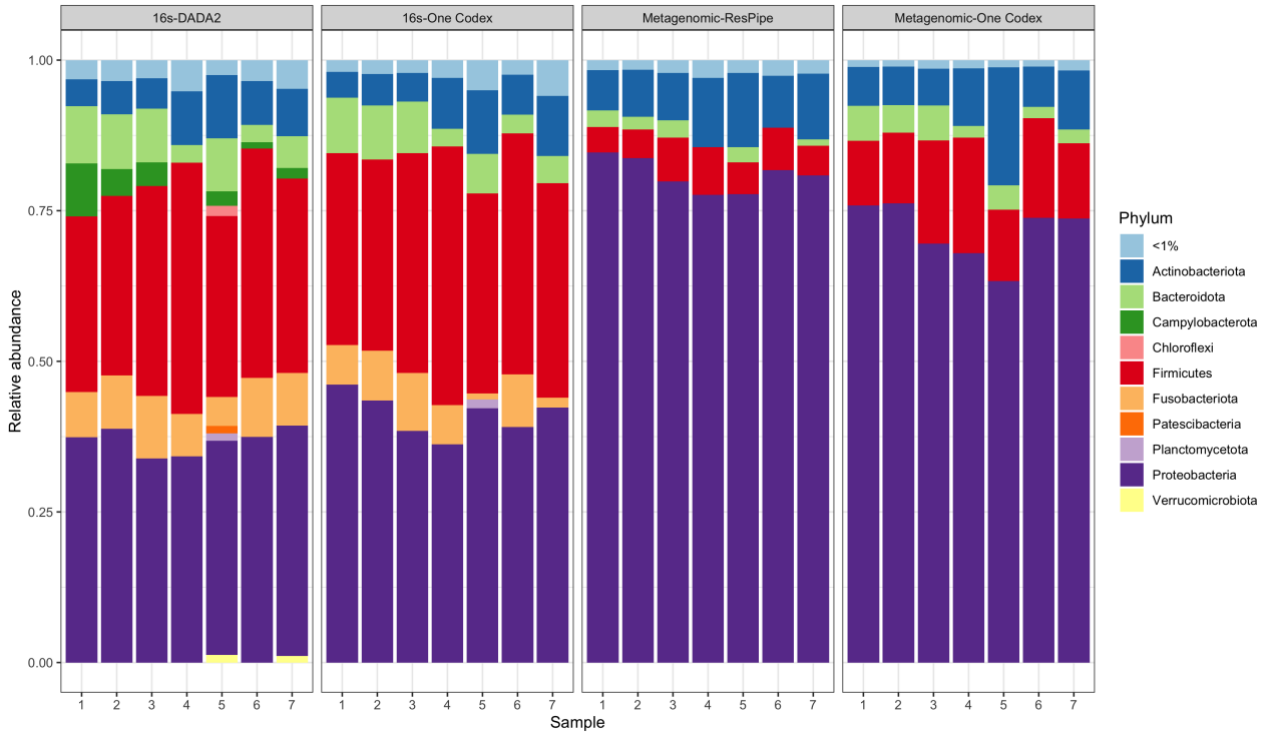


3.3.3 Real-world wastewater samples

3.3.3.1 Taxonomic profiles

Consistent with findings from my mock microbial community benchmarks, taxonomic profiles of “real-world” wastewater samples differed significantly between 16S rRNA and shotgun sequencing, with large differences in relative abundance and detection of core taxa (Figure 12). For metagenomics, Proteobacteria were estimated as most abundant across all samples, with relative abundance approximately two-fold higher than reported by 16S rRNA gene sequencing. Conversely, relative abundance of phyla detected by 16S rRNA gene sequencing across all samples were reduced in metagenomic estimates (Firmicutes, Bacteriodota) or entirely missed/found at <1% relative abundance (Fusobacteriota, Campylobacteriota). Interestingly, this was not the case for Actinobacteriota which was estimated with higher relative abundance than 16S rRNA gene-based estimates by both shotgun pipelines. 16S rRNA gene-based estimates of phylum abundance were more even, with variable detection of Chloroflexi, Patescibacteria, Planctomycetota and Verrucomicrobiota across samples. Interestingly, Campylobacteriota was not detected using the One Codex Targeted Loci pipeline but reported by DADA2.

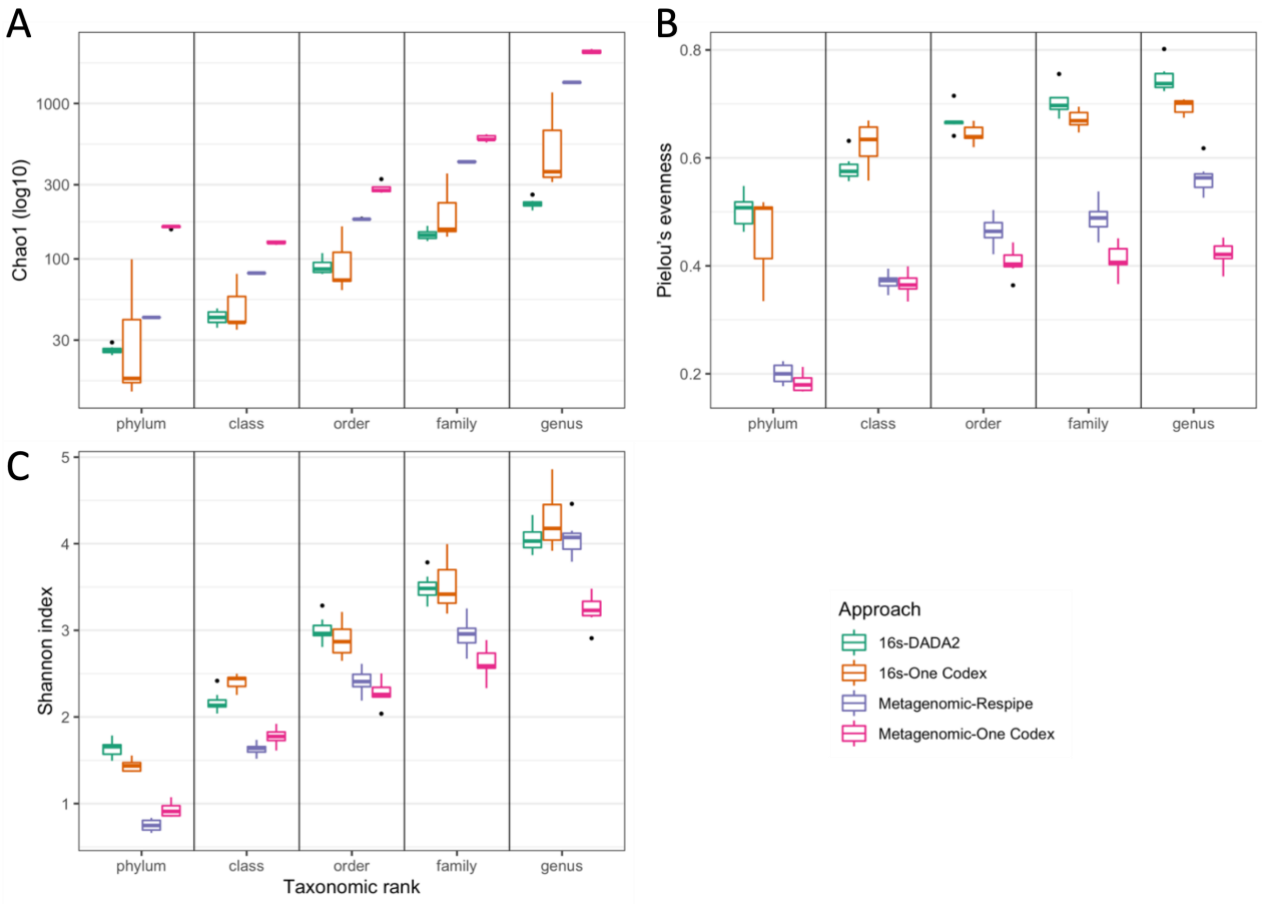
Figure 12 - Relative abundance of phyla for wastewater samples (1-7) faceted by sequencing approach and pipeline combination.



Metagenomic approaches consistently reported significantly higher richness (Chao1) than 16S rRNA gene sequencing approaches at all taxonomic ranks tested ($p<0.002$; mean increase of 1354 genera), with the One Codex Metagenomic Classifier identifying significantly more unique taxa than ResPipe ($p<0.001$; mean increase of 787 genera) (Figure 13A). Conversely, 16S rRNA gene sequencing approaches produced profiles with significantly more even representation of taxa (Pielou's evenness; $p<0.001$) than metagenomic approaches (Figure 13B). Despite higher richness in metagenomic results, overall community complexity (Shannon index) was significantly higher for profiles generated by 16S rRNA gene sequencing than metagenomic approaches ($p<0.007$); except for ResPipe at the genus level. (Figure 13C). Interestingly, One Codex Targeted Loci classifier (16S rRNA gene data) produced taxonomic profiles with higher variability across the wastewater samples than all other sequencing-pipeline combinations (Figure 13ABC).

Figure 13 - Distribution of diversity metrics across all wastewater samples coloured by sequencing approach and pipeline combination.

Panels: 13A reports richness via log10 Chao1, 13B reports abundance distribution via Pielou's evenness, 13C reports overall alpha diversity via Shannon index.



3.3.3.2 AMR profiles

Across all wastewater samples, AmpliSeq One Codex Report identified a total of 3220 AmpliSeq targets (2889 called as present, 331 called as probable) while AmpliSeek identified a total of 2157 (1970 called as present, 187 called as potentially present), and the shotgun metagenomics-ResPipe pipeline identified a total of 526 CARD sequences with lateral coverage of 1 (i.e. 100% nucleotide identity over the full length of the reference sequence). When deduplicated to account for targets shared across

wastewater samples and redundant AmpliSeq amplicon design, the AmpliSeq One Codex Report detected the most unique AMR genes (n=367), followed by AmpliSeek (n=300) and shotgun metagenomics-ResPipe (n=132), respectively.

The distributions of AmpliSeq targets identified by the One Codex Report and AmpliSeek showed the same ratio of increased present calls as the evaluation of the defined mock (i.e. 1.4-fold more AmpliSeq targets called as present by One Codex than AmpliSeek). Similarly consistent with the methodological comparisons on the mock sample, almost all AMR genes called as present by One Codex but absent by AmpliSeek were hits identified at <100% nucleotide identity (817/826; 99%; median nucleotide identity: 98% [IQR:97-99%]). The nine targets identified as present and at 100% identity by the One Codex pipeline but called absent by AmpliSeek possessed low numbers of mapped reads by One Codex (median of one read mapping only). The single AmpliSeq target not targeted by the One Codex AmpliSeq Report (aph2prime-Ia_NT) was called as present by AmpliSeek across all wastewater samples.

3.4 Discussion

In this study, I have compared the performance of metagenomic and targeted sequencing and different bioinformatics approaches in accurately reconstructing taxonomic and AMR gene composition of a defined mock microbial community, and a set of polymicrobial "real-world" wastewater samples containing complex microbial communities. I highlight that results are significantly influenced by both the sequencing and bioinformatics approach used, and this is an important consideration for studies using these methods.

3.4.1 Mock community evaluation

Metagenomics outperformed targeted 16S rRNA sequencing and more accurately estimated true relative abundance of mock genera regardless of the bioinformatic pipeline utilised. While 16S rRNA gene sequencing successfully captured all mock genera, relative abundances differed significantly from the known, true, composition. This was not unexpected as targeted approaches are subject to primer and amplification biases already known to potentially distort taxonomic profiles[114-116], however, the degree of bias has not been previously quantified. This was most apparent for One Codex Targeted Loci results where *Klebsiella* and *Bacillus* abundances were respectively under- and overestimated by up to three-fold. Conversely, shotgun metagenomics is less susceptible to these biases, and when combined with statistical methods for normalising abundances (as used by both ResPipe and One Codex Metagenomic Classifier), accurately estimated mock genus distributions. For profiling taxonomy, ResPipe and DADA2 slightly outperformed the One Codex Metagenomic and Targeted Loci classifiers. Since classification methods were similar within each sequencing approach (i.e. k-mer based for shotgun metagenomics and alignment-based for 16S rRNA gene profiling), differences in the reference taxonomic databases are most likely responsible for differences in performance[103], consistent with previous studies demonstrating classification is dependent on database choice[110,114].

Almost all metagenomics and AmpliSeq approaches detected all relevant unique AMR genes known to be present in the mock microbial community. However, the DNA Amplicon App performed poorly, reporting no reads mapping to 21 AmpliSeq target

sequences known to be present in the mock and reported as present by the One Codex AmpliSeq Report and AmpliSeek. This appears attributable to differences in the read mapping workflow rather than alignment algorithm *per se* since both One Codex AmpliSeq Report and the DNA Amplicon App use the BWA algorithm to align reads. However, instead of directly aligning reads to the reference AmpliSeq target sequences as performed by One Codex and AmpliSeek, the DNA Amplicon App first aligns reads to a reference database of whole genomes containing the reference AMR genes[117]. I noted that this initial genome-based screen appeared to erroneously filter out reads which would otherwise map directly to sequences present in the AmpliSeq AMR panel. However, full details of the DNA Amplicon App workflow are unclear and detailed evaluation of this pipeline was therefore beyond the scope of this study.

For AmpliSeq based methods, when focussing on AMR genes known to be present in the mock with 100% nucleotide sequence identity, AmpliSeek slightly outperformed One Codex AmpliSeq Report as the latter reported more false positives. False positives have been previously described with the AmpliSeq panels and attributed to mispriming events associated with primers erroneously binding to non-target sequences[118]. However, mispriming did not appear to be the main cause in this study since most false positives shared high sequence identity to the reference AMR genes (>98%), indicating that these are more likely related to pipeline-specific differences in calling genes present/absent. In fact, the highly related nature of most false positive calls made by the One Codex AmpliSeq Report would be consistent with the default 2% error BWA mapping threshold used in the pipeline to account for sequencing error. The difficulty in calling

presence/absence of related AMR genes where the reference sequences used as markers are very similar is highlighted in my analyses: For example, the amplicon reference sequence for *bla*$_{FEC}$ (blaFEC_T) shares high nucleotide identity with amplicon reference sequences denoting *bla*$_{CTX-M-3}$, *bla*$_{CTX-M-15}$ and *bla*$_{CTX-M-32}$ (blaCTX-M-3_T.2 [99%], blaCTX-M-15_T.3 [99%] and blaCTX-M-32_T.1 [98%]). Modifying the mapping threshold may avoid these issues since recent work has suggested the median error rate of Illumina MiniSeq sequencing reads may be as low as 0.6%[119]. For AmpliSeek I retained stringent mapping of reads as the default to enable differentiation between highly similar reference markers, guard against mispriming, and call specific alleles with high confidence, but this threshold can be relaxed depending on the users' requirements.

3.4.2 "Real-world" wastewater samples

On real-world samples, where the true diversity was much higher than in the mock community, differences in the relative abundance of taxa and the number of unique AMR genes observed between approaches when profiling the mock community were greatly amplified. In these samples, metagenomic approaches identified more unique taxa than 16S rRNA gene profiling at all taxonomic ranks tested. This represents a commonly cited key advantage of shotgun sequencing, where higher richness of bacterial taxa can be reported than for 16S rRNA-based methods[120-122], and is particularly important in the characterisation of complex microbial communities[121]. Taxonomic classification is also heavily dependent on the reference database used[104]. In this respect, the One Codex Metagenomic Classifier yielded significantly higher richness than ResPipe likely owing to respective reference species database sizes of ~115,000 and ~20,000 complete genomes.

Taxonomic richness reported by ResPipe was also seemingly capped across most ranks with minimal variation between wastewater samples which may be a result of saturating the reference database as previously suggested[90].

Metagenomic abundance estimates were largely skewed towards a small number of taxa; in contrast, 16S rRNA gene-based estimates were consistently more even across all ranks regardless of pipeline, and more sensitive to the detection of taxa missed by shotgun metagenomic outputs. This likely partly reflects the sensitivity trade-off of shotgun sequencing where sequencing effort can be swamped by high abundance taxa and/or taxa with large genomes; especially when sequencing depth is low and community complexity is high[92], which can result in missing the presence of low-abundance, but significant, taxa[123]. However, users of 16S rRNA gene sequencing for taxonomic profiling need to be aware of the likely greater divergence of profiles from the truth. In my hands, different 16S rRNA gene pipelines also notably impacted results; for example, Campylobacteriota were absent from One Codex outputs but classified as present using DADA2, further highlighting differences across reference databases.

Interestingly, despite being more sensitive in profiling AMR targets in the mock, metagenomics detected less than half the number of AMR genes reported by targeted AmpliSeq approaches. This is potentially driven by low abundance AMR genes which are missed by an agnostic approach, but directly targeted and amplified by AmpliSeq. However, I used the strictest definition of AMR gene presence requiring read matches covering all of the reference sequence (lateral coverage=1), so my metagenomic profiling

was a conservative estimate. As discussed, sequencing depth impacts the sensitivity of target detection in complex samples, with the relative simplicity of the mock composition (n=11 bacterial genomes, 77 CARD sequences, 43 AmpliSeq sequences) masking this limitation.

As seen in the evaluation on the mock, the less stringent mapping thresholds of the One Codex Report produced more present calls than AmpliSeek due to calling target variants as present. In fact, the proportion of increased present calls made by One Codex Report compared to AmpliSeek was consistent between mock results and wastewater results. These findings support similar performance of both AmpliSeq pipelines in wastewater as in the mock despite significantly increased complexity, where AmpliSeek results represent increased confidence in the presence of the exact marker sequence.

3.4.3 Limitations

I have not evaluated all methods for profiling taxonomic and resistome composition and therefore potentially missed approaches with increased performance. However, my pragmatic evaluation focussed on accessible approaches (i.e. kit-based and all-in-one pipelines) which are commonly used, attractive to researchers and have not undergone comparative evaluation. The relative simplicity of the mock DNA community may mask sensitivity issues owing to sequencing depth but use of *in silico* simulated datasets would otherwise omit assessment of library preparation amplification reactions. To mitigate this, I enriched the mock with additional AMR determinants and presented performance differences between the mock and “real-world” samples. My in-house pipeline AmpliSeek

was internally validated on the same mock dataset used for performance metrics which inherently optimises performance, however, the “real-world” samples represented an external validation confirming similar performance with and without scoring threshold optimisation. I also guarded against wastewater metagenomic sensitivity issues by using optimised depth as determined by previous ultradeep pilot sequencing performed on wastewater form the same sampling site[90]. My resistome characterisation focusses on presence/absence but a major advantage of shotgun metagenomics is quantitative measurement of genes which I have not explored here. Additional factors such as read length and approaches to normalization may also optimise each approach but exploring these were beyond the scope of this study.

In conclusion, a complementary sequencing approach encompassing 16S rRNA or shallow-metagenomic sequencing and the AmpliSeq AMR panel represents a potential lower-cost alternative to deep shotgun sequencing for taxonomic and resistome profiling of highly diverse, complex, polymicrobial samples. Accuracy of profiling was also to some extent dependent on the bioinformatics pipeline used; for AmpliSeq AMR gene profiling I have optimised an in-house pipeline (AmpliSeek) which outperformed the other methods tested on this dataset and is open source. However, further validation of this in-house pipeline is needed, since it was also developed on the same dataset.

I recommend careful consideration and further validation of sequencing approaches and bioinformatic pipelines in the context of study goals and sample type to be analysed. Validating sequencing and bioinformatic approaches remains challenging due to

difficulties in representing the true complexity of “real-world” samples; any conclusions drawn should always be considered in the context of the methods used.

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

Note: This chapter has previously been published as a preprint[124]. I authored the text and figures below (which are reproduced here largely unaltered from the published work) with input from supervisors commensurate with the amount of input/advice that would be considered appropriate for a DPhil thesis. All authors have agreed to the inclusion of this published work in my thesis.

4.1 Introduction

Antimicrobial resistance (AMR) represents a significant challenge to global health[6]. Surveillance efforts are essential to monitor trends, identify emergence and develop interventions[7]. 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[37]. Recent AMR-focussed WBE studies have surveyed global AMR gene distributions[43], identified associations between AMR genes in clinical and wastewater samples[49,50] and characterised the relationship between wastewater AMR abundance, antimicrobial usage and clinical isolate phenotype[125]. 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[42], by simultaneously capturing healthcare- and community-associated populations, and reflecting carried organisms thought to silently comprise most of the true AMR burden. Genotypic/sequencing-based

approaches to WBE may also provide useful information on specific AMR-associated genetic determinants and high-risk clones[44] whilst presenting a means to standardise AMR detection methods such as through metagenomic-based profiling of wastewater resistomes[39,47].

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[41]. Several studies have aimed to validate specific methodology for AMR-focussed WBE such as DNA extraction[61], or quantify the effect of freeze-thawing[60] or healthcare effluents[125], 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[41,52,53]. However, the composition of influent may vary significantly over short periods due to human behaviour[55] (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[51]. Despite this, grab sampling remains the most common wastewater sampling method used by AMR-focussed WBE studies[41].

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. I 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, I characterised changes in influent composition in relation to metadata such as time-of-day, flow rate and nutrient concentrations. I aimed to highlight limitations of each method and recommend potential use cases.

4.2 Methods

4.2.1 Wastewater sampling

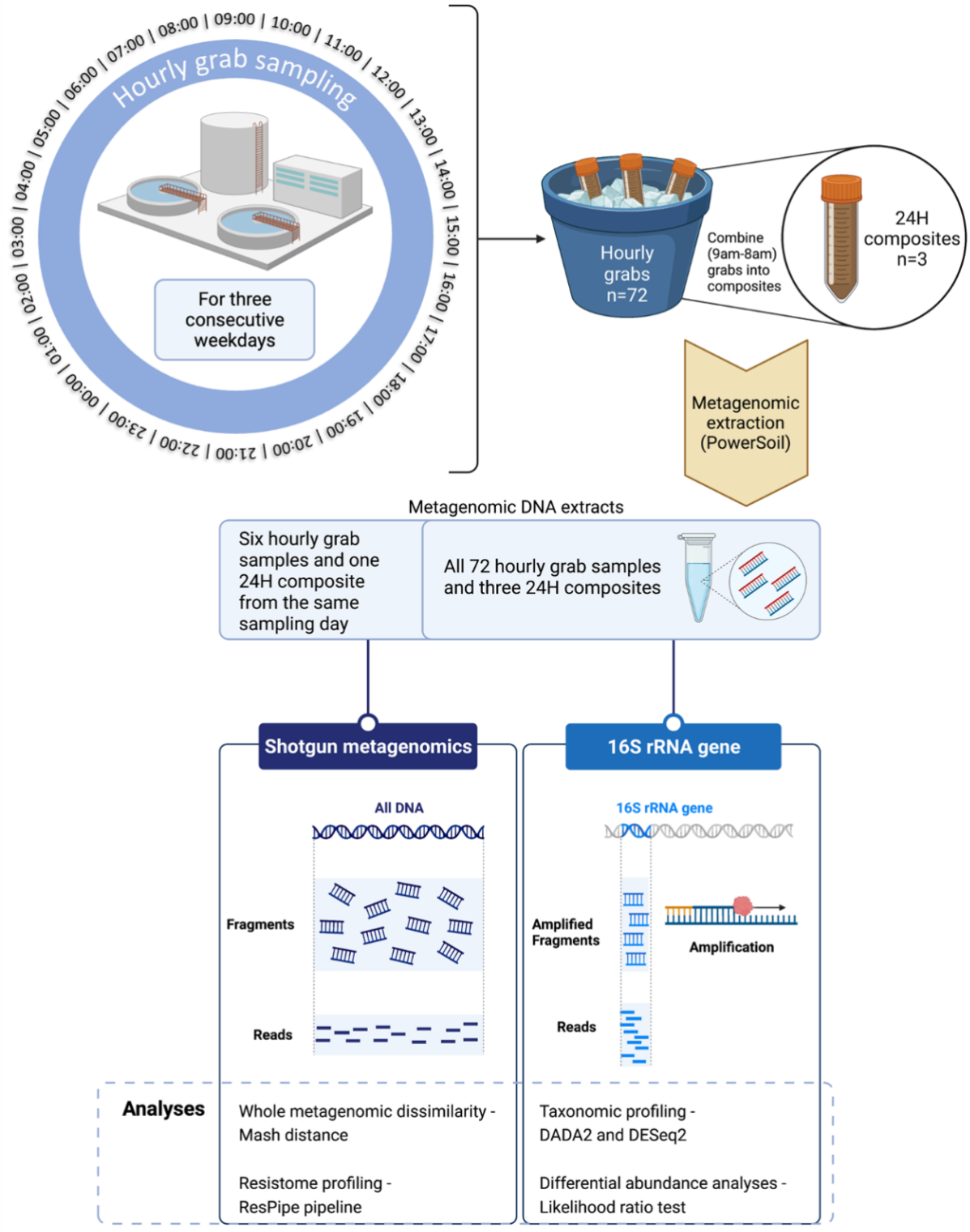
Wastewater influent was sampled from the crude inlet (post-mechanical screening out of large solids) of a large UK-based wastewater treatment work (WWTW) with population equivalent 223,435 and consented flow of 50,985 m^3/day. The WWTW received urban wastewater and healthcare effluent from multiple sites 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]) (Figure 14). 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[126]. 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.

4.2.2 Sequencing

All samples (n=75) underwent 16S rRNA gene sequencing using 515F-806R primers on the MiSeq (Illumina, San Diego, USA) as previously described[106] 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[90]. 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). R scripts and markdown renders are available at https://github.com/KaibondChau/Chau_etal_Hires.

Figure 14 - Schematic overview of study methods and analyses



4.2.3 Computational methods

Taxonomic processing of 16S rRNA gene sequence data was conducted using DADA2 v1.16 as previously described[106]. 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[127] and DESeq2 v1.34.0[128] 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[129]. Principal component analyses were conducted on variance stabilised counts using R packages stats v4.1.2, factoextra v1.0.7[130], FactoMineR v2.4[131] and vegan v2.6.2[132] 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 focussed on sampling time.

Metagenomic sequencing data was processed through ResPipe[90] v1.4.0 for AMR genotyping, and Mash v2.3 for comparing overall similarity via metagenome distance estimation[133] [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 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.

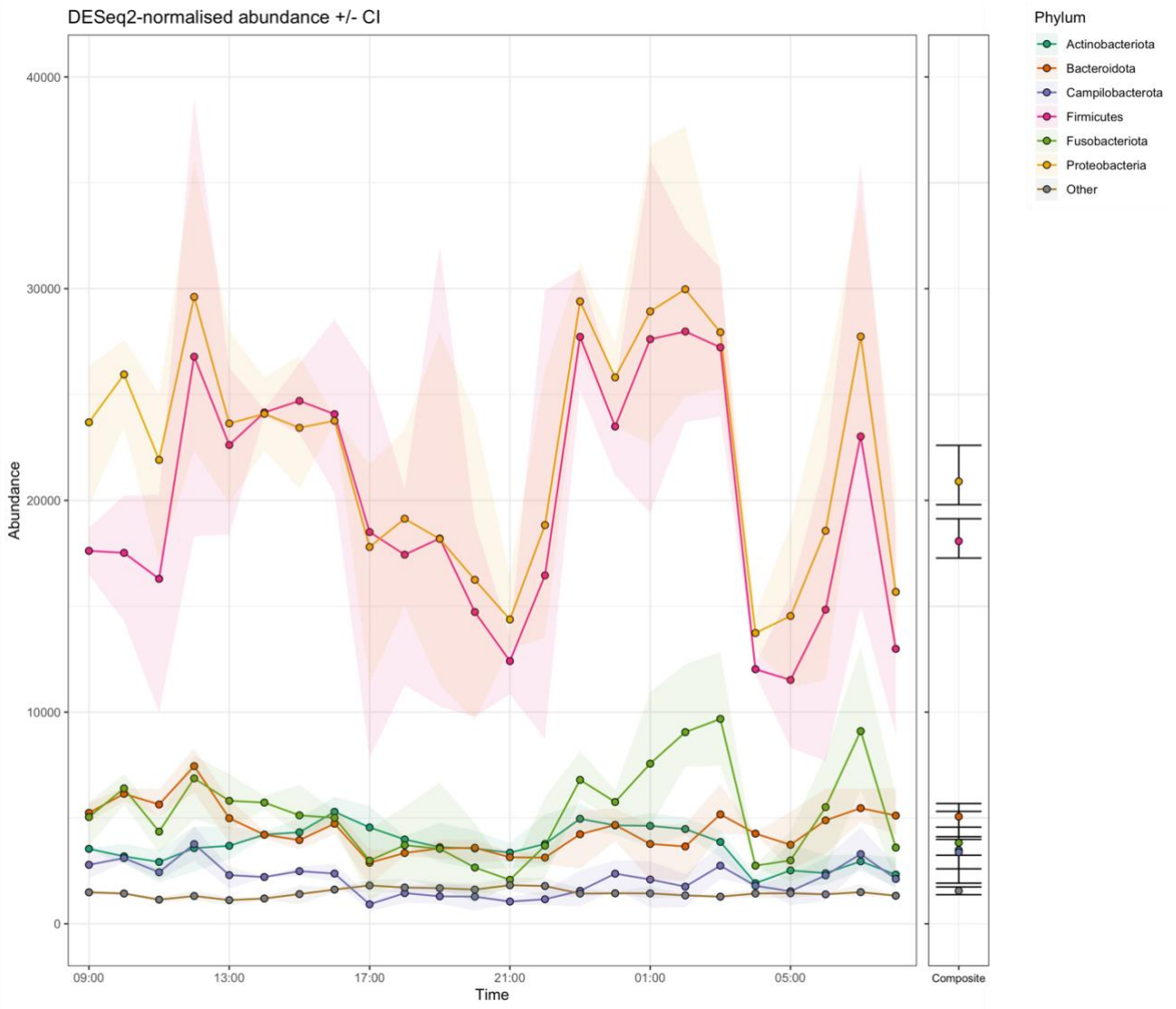
4.3 Results

4.3.1 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 (Figure 15), 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.

Figure 15 - Temporal taxonomic abundance fluctuations.

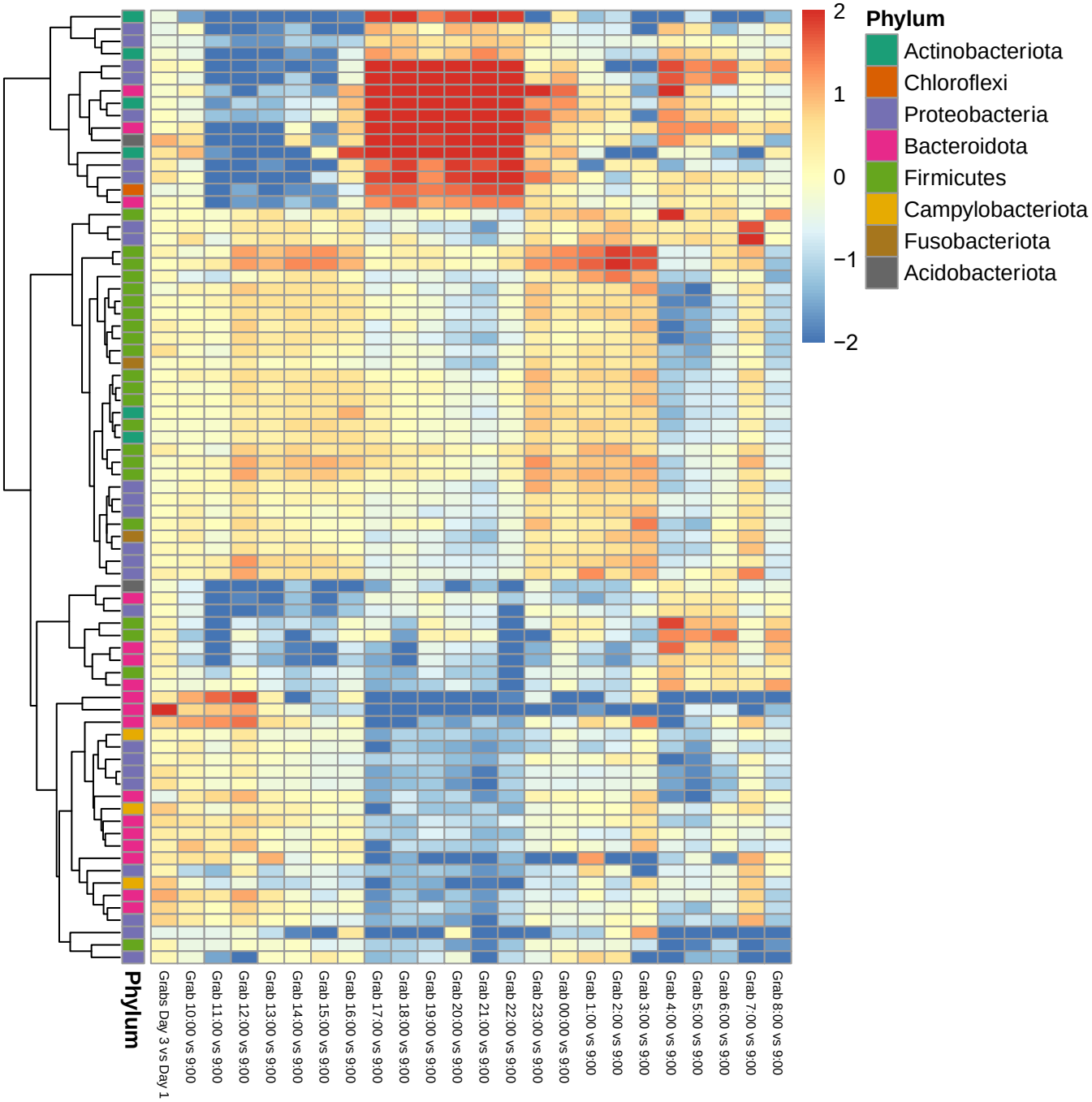
Each point represents the mean DESeq2-normalised abundance of Phyla (coloured) across the three sampling days for hourly grab samples (left) and 24H composites (right). Shaded areas (grab) or intervals (composite) convey 95% confidence intervals.



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) (Figure 16). 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).

Figure 16 - Log 2-fold changes in abundance for ASVs undergoing significant abundance changes across grab samples.

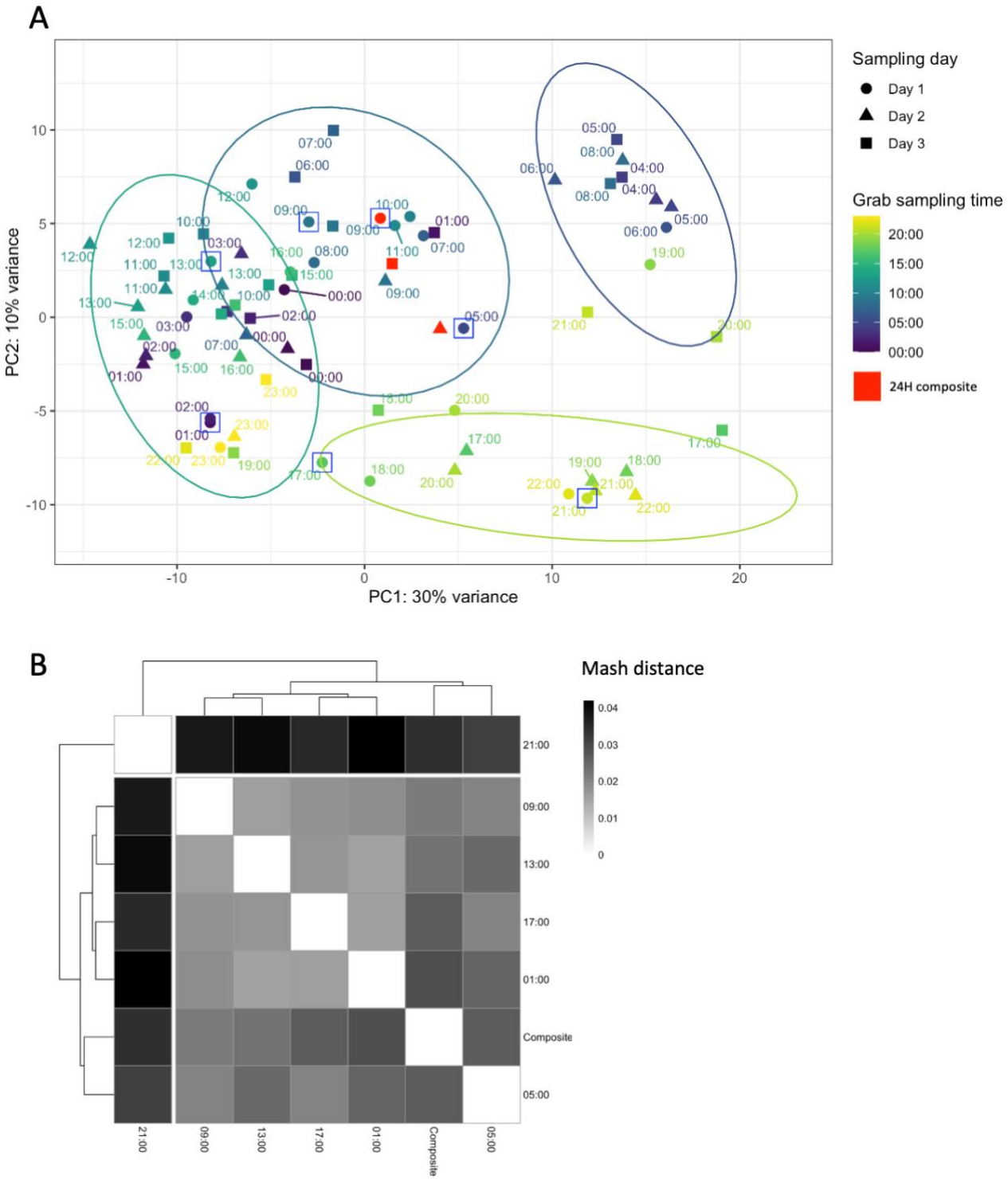
Each row represents an ASV identified as significant by likelihood ratio testing in at least one pairwise comparison between samples (columns). Maximum threshold of change was set at 2 and -2 to aid visualisation of significant abundance differences (i.e. values at 2 and -2 may represent >2 and <-2 respectively). Phylum-level classification of ASVs is annotated by colour (left).



Accordingly, Principal component analysis of 16S rRNA gene-based profiles demonstrated association between time of grab sampling and taxonomic variation (Figure 17). 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 (Figure 15), 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 17 - 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.



Mash distance of full metagenomic content for a subset of seven samples from sampling day one showed the same clustering pattern by Principal component analysis on 16S rRNA gene data alone (Figure 17B). 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 Principal component analysis, 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.

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).

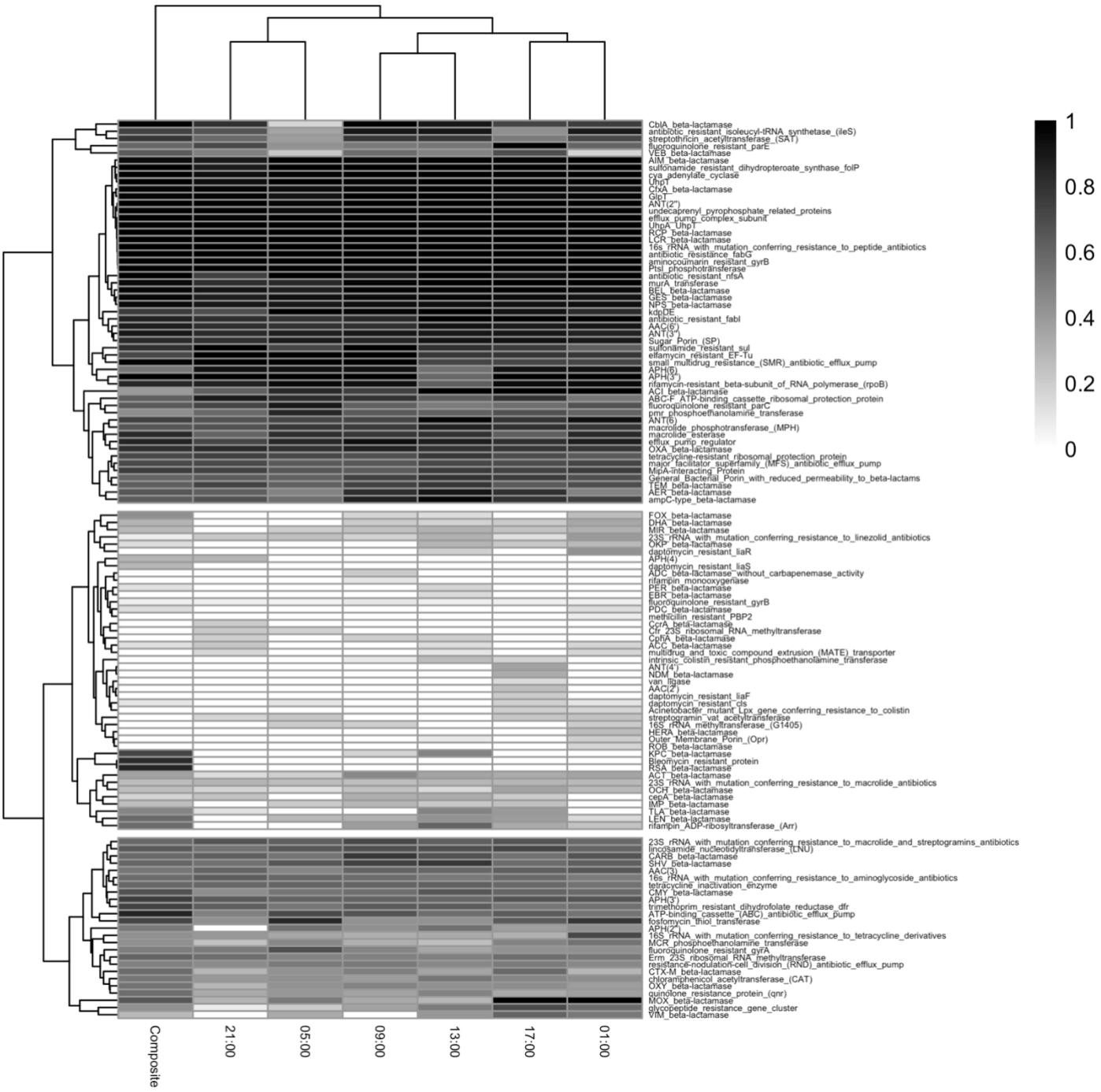
4.3.2 Resistome variation

From the seven samples undergoing metagenomic sequencing, in total 122 AMR gene families were identified at various levels of lateral coverage (i.e. mapping completeness) (Figure 18). Of the 122 AMR gene families, single grab samples identified between 90/122 (74%) and 101/122 (83%), while the 24H composite identified 101/122 (83%). When comparing all AMR gene families identified in all grab samples to all those identified in the 24H composite, grab samples identified a median six (IQR:5-8) gene families not seen

in the composite (total: 36 across grab samples of which 21 were uniquely identified), and the composite identified three AMR gene families 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 AMR gene families and identifies more false-positives than 24H composite sampling.

Figure 18 - Heatmap of AMR gene family lateral coverage across samples from the same sampling day.

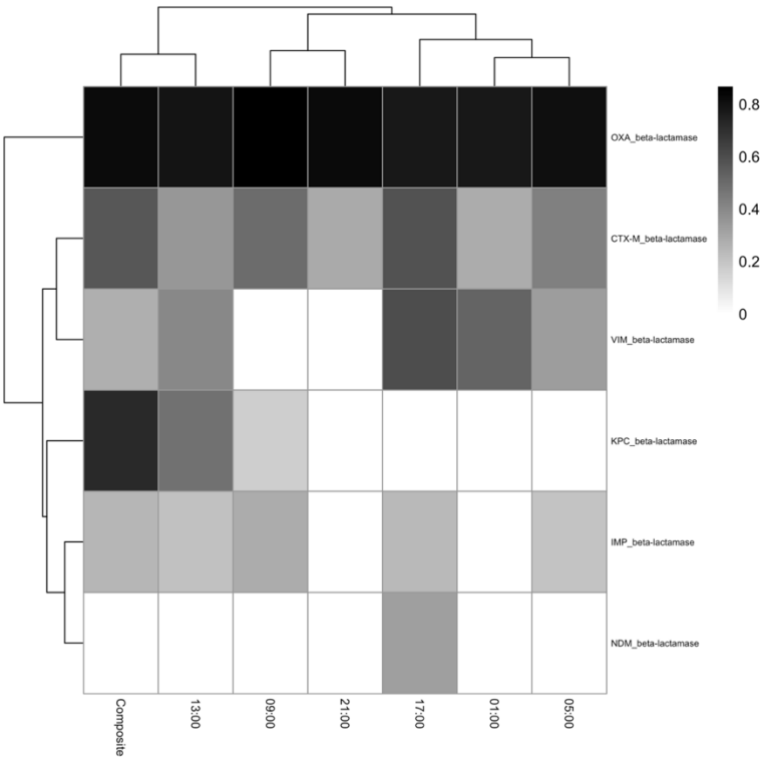
Row-wise hierarchical clustering divides gene families into consistently high lateral coverage (top), intermittently 0 lateral coverage (middle) and variable lateral coverage (bottom).



AMR gene families identified hierarchically clustered into three main groups based on lateral coverage across samples: predominantly high, variable and intermittently identified. 53/122 (43%) AMR gene families 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 gene families of significant human clinical importance had lateral coverage following these groupings (Figure 19), 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 19 - Heatmap of selected AMR gene family lateral coverage across samples from the same sampling day.

Subset of significant clinically relevant AMR gene families of particular interest to public health surveillance.



4.4 Discussion

In this study, I 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. I 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.

4.4.1 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[134]. 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[45], its composition is also influenced by the sewer environment during conveyance to WWTWs[38], and by contributions from other wider environmental and anthropogenic sources. Accordingly, my abundance estimates for wastewater phyla are not completely consistent with estimates of Firmicutes and Bacteroidota in direct human metagenomic studies (typically ~90% abundance[134]), but are similar to other wastewater estimates[135-137]. Campylobacteriota, Acidobacteriota and Chloroflexi also significantly

fluctuated, and have been previously reported in humans, soil and wastewater respectively[138]. Interestingly, Chloroflexi significantly increased in abundance during periods of lower flow rate, possibly indicating increased sewer community contributions[139].

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. I purposefully avoided sampling on weekends and wet weather days to prevent previously described factors such as weekend household behaviour[135] and surface water runoff/infiltration[58] from impacting influent composition. Without external factors such as these, I find that sampling one weekday in my 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[38]. My findings corroborate previously reported diurnal compositional fluctuations associated with flow rate related to human behaviour/household discharge patterns[55]. However, in addition to morning peaks reported by Guo et al., I 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 my 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 interpretating results, but are often overlooked[41].

4.4.2 Resistome profiling

The 122 AMR gene families 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[90]) in more than one sample, AMR gene families were generally also captured at 1 or high lateral coverage by most other samples (i.e. consistent detection), as seen for 53/122 (43%). 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 AMR gene families (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 gene families 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 AMR gene families (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[33,135]. However, the low lateral coverage seen for most intermittent gene families may alternatively represent mapping artifacts; in this case, the AMR genes may not be truly present.

While grab sampling identified several more AMR gene families in total across all six individual timepoints, the 24H composite more consistently identified gene families with high lateral coverage. Almost all AMR gene families 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 gene families 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

AMR gene family such as *bla*$_{KPC}$ is missed by most or all the six grab samples but captured with high lateral coverage in the composite, I can assume the gene 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, *bla*$_{VIM}$, *bla*$_{IMP}$ and *bla*$_{KPC}$ 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 AMR gene families (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[90] which may explain why certain single grab samples identified AMR gene families missed in composite samples. For example, *bla*$_{NDM}$ 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[140]) 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, AMR gene families is reduced in the composite.

4.4.3 Limitations

Resource limitations meant that I was 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[141,142]. 16S rRNA gene sequencing may also be prone to amplification biases and normalisation issues. To mitigate these issues, I undertook metagenomic sequencing of a subset of samples which supported my 16S rRNA gene-based conclusions and enabled us to additionally evaluate the impact of sampling approach on resistome profiles. I 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[90]; I sequenced at an optimised depth previously determined by ultradeep sequencing of wastewater from the same WWTW as my study. My 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.

In 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, I 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 AMR gene 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[35]. 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.

5. Fine-scale spatial and longitudinal survey of AMR prevalence, antibiotic prescribing and antibiotic residues across Oxfordshire WWTWs

5.1 Introduction

Surveillance efforts are crucial for combating the increasing threat of antimicrobial resistance (AMR)[6,7]. Understanding occurrence, transmission and dynamics of AMR precedes the development of effective policies and interventions. Current surveillance systems typically rely on data generated through clinical microbiology laboratories, limiting most data collected to phenotypic results for specific pathogens causing infections. This narrow clinical snapshot hampers the surveillance of high-risk AMR-associated clones and specific AMR-associated genetic determinants[44] which are frequently disseminated 'silently' through colonisation rather than infection, whilst inherently biasing surveillance efforts towards healthcare-associated populations[42].

Wastewater-based epidemiology (WBE) of AMR is an attractive alternative surveillance approach due to its potential for monitoring large populations at scale with relative ease[37]. Wastewater represents a pooled sample which reflects the source population[38], and can be analysed to generate rich population-representative datasets with reduced logistical or ethical barriers otherwise encountered during large-scale individual-level sampling[43]. Wastewater may also act as a convergence point of both healthcare- and community-associated populations[45], thus potentially avoiding selection bias and

capturing AMR prevalence of commensal organisms in healthy individuals, thought to constitute a major proportion of the true AMR burden[19].

Analysing wastewater samples with culture-independent profiling techniques such as metagenomic sequencing also affords additional advantages over phenotypic assays[41]. For example, wastewater resistomes may be agnostically probed for all known AMR genes using continuously updated databases, facilitating future reanalyses of sequencing datasets for targets yet to be discovered[39,47]. Metagenomes also provide information beyond AMR since datasets can be flexibly compared to reference databases such as RefSeq and MGnify for taxonomic and protein-homology profiling respectively. Culture-independent methods also increase profiling resolution since whole genome sequencing of single cultured isolates will generally only capture a tiny fraction of the true diversity present[90]. However, metagenomic approaches are more resource intensive and bioinformatic deconvolution of host species/strains and genetic context of relevant genes (e.g. location on mobile plasmid vs static chromosome) is challenging when using short-read data[90].

WBE has established prior utility in licit/illicit drug monitoring[40] and the monitoring of polio[32] but experienced a surge in interest during the SARS-CoV-2 pandemic[36] and most recently the human monkeypox (hMPXV) virus epidemic and potential polio resurgence[143]. However, WBE of AMR is relatively new with respect to method development and validation, with high heterogeneity in methods and study design likely contributing to systematic differences in interpretation (see Chapter 2)[41]. Several studies

have aimed to validate methodological aspects of WBE of AMR such as wastewater sampling (see Chapter 4)[124], DNA extraction[61], sequencing approaches (see Chapter 3)[91] and the effect of sample freeze-thawing[60], but much remains unexplored. One further such aspect is the spatial and temporal scale on which WBE for AMR surveillance is effective as this will dictate the overall utility of wastewater sentinel sites. Recent AMR-focussed WBE studies have suggested that single wastewater samples from large cities are effective for reflecting country-level AMR prevalence in transnational snapshot surveys[42,49,50]. However, few studies have conducted finer-scale longitudinal surveys to assess whether WBE is effective for capturing regional trends within a country.

I therefore conducted a fine-scale longitudinal survey of wastewater influent from five sites within the same region of the UK (Oxfordshire). I utilised deep metagenomic sequencing to characterise the taxonomy and resistome of wastewater samples collected at five timepoints over one year. Additionally, I explored sources of resistome variation by modelling contemporaneous local primary care antimicrobial prescribing as well as antibiotic residue concentrations in the wastewater. I aimed to identify whether significant compositional differences were associated with sampling site and or timepoint, and whether wastewater from a city in the region sufficiently represented AMR prevalence observed in other sites within the same region.

5.2 Methods

5.2.1 Site-level feature analyses and site selection

5.2.1.1 Data sources

Data regarding Oxfordshire wastewater treatment works (WWTWs) and farms was provided as a data extract and Geographic Information System map from the REHAB project[62] (the environmental REsistome: confluence of Human and Animal Biota in antibiotic resistance spread). Exact coordinates of farms were omitted due to confidentiality agreements with landowners. Postcodes for primary care general practices in Oxfordshire were retrieved from NHS Digital datasets[144] on November 22, 2018. Boundaries of practice catchment areas were also retrieved from the NHS Data Catalogue[145]. These represent individual practice submissions including those with incomplete geometry and are deemed as a "reasonably accurate reflection of the contract". Data on practice-level prescribing, number of registered patients and practice postcodes was retrieved from NHS Digital datasets for the Oxfordshire Clinical Commissioning Group[146]. These prescribing datasets list all medicines, dressings and appliances prescribed in a given month by Average Daily Quantities which reflect daily dose for an individual patient and can be used to compare prescribing activity across settings. Prescribing data was included for 01/01/2018-31/12/2018. Residential care home postcodes were retrieved from the Care Quality Commission care directory[147] and reflect all nursing and residential homes operating on May 1, 2019. Boundaries for Oxfordshire county and Oxfordshire Lower Layer Super Output Areas (LSOAs; divided localities of 1500 population) were retrieved from the Office for National Statistics Open Geography Portal[148]. Data on Indices of Multiple Deprivation (v2019) by LSOA were

retrieved from the Ministry of Housing, Communities & Local Government Open Data platform[149].

5.2.1.2 Data processing

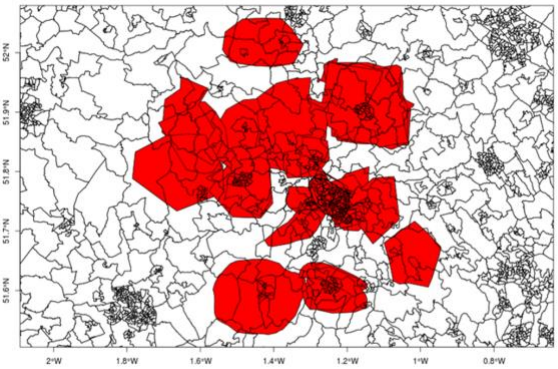
Residential care home and practice postcodes were geocoded into coordinates using the Google Maps Geocoding API via the R package ggmap[150]. The Google Maps Static Map API was then used to generate maps with specified coordinates, boundaries and features. Generated maps were overlaid onto the REHAB Geographic Information System map at the same scale to estimate proximity of farms to sites under consideration. Based on viewing wastewater pipeline maps of Oxfordshire (Thames Water, Reading, UK), a conservative 4.5km radius was used to approximate sewersheds (i.e. the area served by a WWTW) for the purpose of identifying features proximal to a given WWTW. Radii were accurately plotted by conversion into projection-capable raster formats to account for inequality between longitude and latitude scales. Prescribing data were cleaned to remove closed practices and practice codes referring to community doctors instead of primary care locations. Adjusted monthly Average Daily Quantities were calculated by antibiotic code (Anatomical Therapeutic Chemical Classification), practice code and month-specific patient denominators, and plotted with 95% confidence intervals for amoxicillin and amoxicillin combinations (amoxicillin-clavulanate, co-amoxiclav) prescriptions which have been previously associated with driving resistant infections in Oxfordshire[151]. Practice catchment data requiring cleaning was identified by individually sub-setting and mapping sites of interest using the Google Maps Static Map API. When required, catchment data was cleaned by dissolving erroneous internal polygons and

expanding remaining geometry to fit general practice (GP) boundaries. Practice catchments were then overlaid onto Oxfordshire LSOA boundaries to identify LSOAs of interest. LSOA IDs were extracted for overlapping areas per practice and respective deprivation ranks compared.

Of the 92 WWTWs operating within Oxfordshire, 13 were shortlisted based on their accessibility and geographic location with respect to general practices and other WWTWs (Figure 20). Sites were shortlisted if GPs were present within 4.5km and minimally overlapped with other shortlisted sites, thereby optimising WWTW coverage of nearby GP populations and minimising overlap of covered populations between WWTWs.

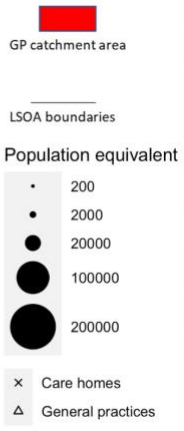
Figure 20 - Site selection summary

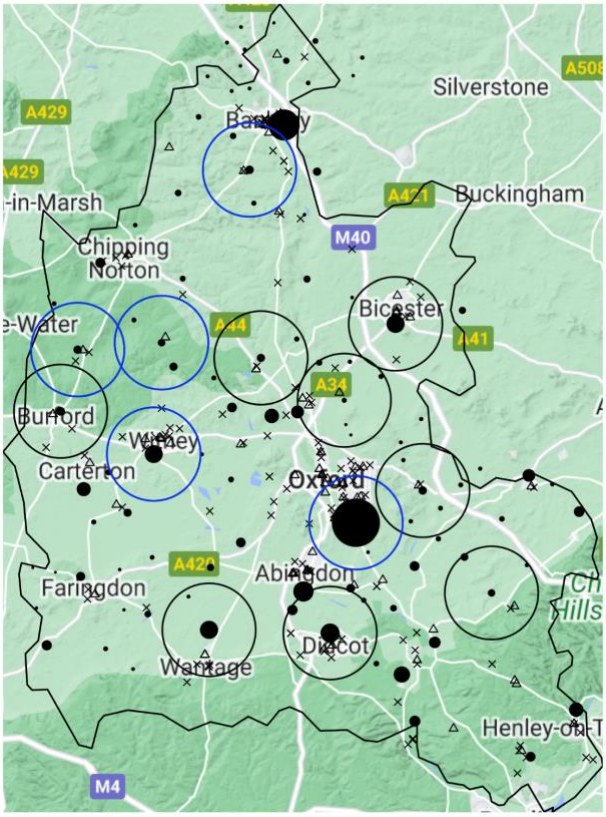
Right: Map of Oxfordshire with WWTWs represented by black points where size reflects population equivalent, GP practices by black crosses, and residential care homes by black triangles. Circles around WWTWs reflect a 4.5km radii around shortlisted sites where final sites purposefully selected are highlighted in blue. Top left: LSOAs of Oxfordshire with GP catchment areas overlayed (red) for shortlisted sites. Bottom left: summary of key characteristics of final site selections.



Final site selections

WWTW	Size (PE)	Features
Oxford	City 223,435	Largest site and population
Witney	Town 26000	High no. of care homes Abbatoir
Charlbury	Village 2863	Only site with no care homes closeby
Bloxham	Village 4100	High historic prescribing
Milton UW	Village 4045	Low historic prescribing



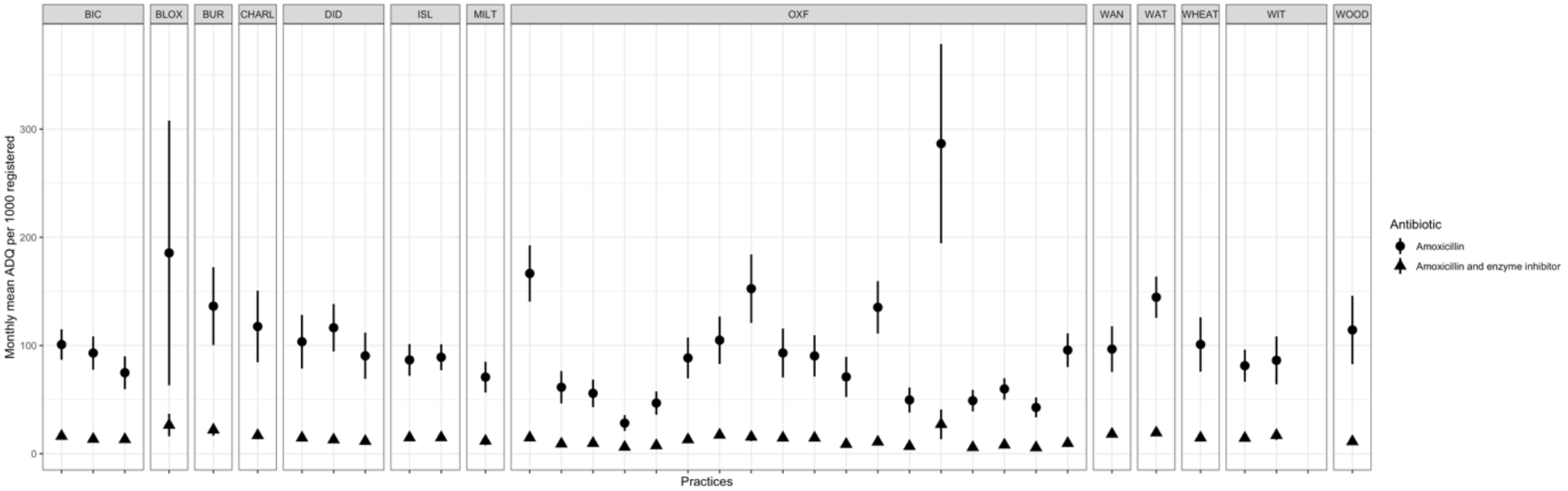


5.2.1.3 Practice-level antibiotic prescribing

Between January 2018 and December 2018, mean monthly average daily quantities across practices located close to the 13 shortlisted WWTWs (normalised to registered patients) varied 10-fold for amoxicillin (range 28.3-286.6) but less so for amoxicillin-clavulanate (co-amoxiclav) (5.6-27.2) (Figure 21). One Oxford practice had significantly higher mean amoxicillin prescribing rates than all other practices except Bloxham which was skewed by an extreme prescribing peak in Jan 2018 (646.8). Co-amoxiclav prescriptions were significantly lower in comparison to amoxicillin but broadly reflected the same patterns across practices.

Figure 21 - Historic practice-level prescribing of shortlisted sites

Mean GP practice-level prescribing of amoxicillin and co-amoxiclav in 2018 with 95% confidence intervals for practices within 4.5km of shortlisted sites (facets).



5.2.1.4 Presence of care homes, agriculture and abattoirs

Of the 49 care homes in the shortlisted sewersheds, most were located in Oxford (n=11), Witney (n=8) and Didcot (n=6) (Figure 20). Due to confidentiality, exact farm numbers/locations are not presented. However, all shortlisted sites had a similar number of poultry (n<5) and large ruminant farms (n<3). Bloxham had the largest number of small ruminant farms (n>10) compared to all others (n<6). Lastly, pig farms were only present for Watlington, Wantage and Didcot (n<3). A single abattoir was present across sites, and processes cattle, sheep and pigs with effluent flowing directly into Witney WWTW.

5.2.1.5 Relative Deprivation

Overall, most sites contain LSOAs within the upper ranks of the Index of Multiple Deprivation score, indicating the least deprivation (rank of most deprived area=1). As expected, larger areas denser in population have more LSOAs with a wider spread of ranks (Oxford, Bicester and Didcot) but no significant differences were noted.

5.2.1.6 Site selection summary

Five WWTWs were selected for this study based on purposive sampling to maximise diversity in key features associated with AMR (Figure 20). Sites were additionally excluded when uncertainty of catchment area was high, or sampling was logistically unfeasible. Briefly, the five final WWTW sites were purposively sampled to represent areas with high and low historic primary care prescribing, presence of healthcare facilities, farms and abattoirs, and a range of WWTW population equivalents (i.e. size of

site and volume of wastewater processed which reflects population size of the catchment area as well as industrial input).

5.2.2 Wastewater sampling and processing

Wastewater influent was sampled from the crude inlet of each selected WWTW (n=5), located after mechanical screening of large solids but before any treatment processes. 24H composites were collected via ice-filled autosamplers (Hach AS950, Colorado, USA) positioned to sample mid-stream flow to avoid stagnant material and recycled effluent. Samples were transported and processed on ice within three hours of collection. Each WWTW was sampled longitudinally between 19/11/2019 and 12/11/2020, with five timepoints (Nov-Dec 2019 [T1], Jan-Feb 2020 [T2], Jun-Jul 2020 [T3], Aug-Sep 2020 [T4], Oct-Nov 2020 [T5]). The capacity to perform sampling Feb-Jun was impacted by the national SARS-CoV-2 lockdowns to restrict movement that were implemented in March 2020.

Wastewater samples were centrifuged (5,300RPM, 10min, 4°C) and pellets stored at -80°C before metagenomic DNA extraction using the DNeasy PowerSoil kit following manufacturer specifications (QIAGEN, Hilden, Germany). In parallel, sample aliquots also underwent solid phase extraction as previously described[152]. Briefly, solid phase extraction cartridges were prepared in duplicate for each WWTW-timepoint sample by spiking 50 ml of wastewater with 50ng of internal standards to target 54 drugs and 23 metabolites (Table 2). Spiked samples were then filtered (Whatman GF/F) and loaded onto Oasis HLB cartridges under vacuum. Vacuum-dried cartridges were wrapped in foil

and stored at -20°C before shipment to the University of Bath Department of Chemistry for elution and UPLC-MS/MS as previously validated[152]. Measured residue concentrations (µg/L) were normalised using daily flow rates to generate daily load (g/day) and WWTW population equivalents for normalised daily load (mg/1000 inhabitants/day). Residue concentrations of antibiotic targets were compared to predicted no-effect concentrations (PNEC) calculated by Bengtsson-Palme et al.[153]. In brief, these represent the lowest concentration of an antibiotic before selection of resistance to said antibiotic is likely to occur. Prior validation of analyte peak identification facilitated a quality assessment for each measurement considering signal to noise[152] where measurements with lower noise are more confidently quantitative.

Table 2 - Table of targeted chemicals for residue measurement – modified from Holton et al.[152]

Grouping	Chemical	Class	Target type
Sulfonamide &	Sulfadiazine	Sulfonamide	Parent
Trimethoprim	Sulfapyridine	Sulfonamide	Parent
	Sulfamethoxazole	Sulfonamide	Parent
	Sulfasalazine	Sulfonamides	Parent
	Trimethoprim	Trimethoprim	Parent
	N-acetyl sulfadiazine	Sulfonamide	Metabolite
	N-acetyl sulfapyridine	Sulfonamide	Metabolite
	N-acetyl sulfamethoxazole	Sulfonamide	Metabolite
	4-hydroxy-trimethoprim	Trimethoprim	Metabolite
Macrolide	Azithromycin	Macrolide	Parent
& Lincomycin	Erythromycin	Macrolide	Parent
	Clarithromycin	Macrolide	Parent
	N-desmethyl azithromycin	Macrolide	Metabolite
	N-desmethyl erythromycin A	Macrolide	Metabolite
	N-desmethyl clarithromycin	Macrolide	Metabolite
Penicillin	Amoxicillin	Penicillin	Parent
	Ampicillin	Penicillin	Parent

Grouping	Chemical	Class	Target type
	Penicillin G	Penicillin	Parent
	Penicillin V	Penicillin	Parent
	Amoxicilloic acid	Penicillin	Metabolite
	Ampicilloic acid	Penicillin	Metabolite
	Penicilloic G acid	Penicillin	Metabolite
Cephalosporin	Cefalexin	Cephalosporin	Parent
	Cefixime	Cephalosporin	Parent
	Ceftiofur	Cephalosporin	Parent
	Ceftriaxone	Cephalosporin	Parent
Monobactam	Aztreonam	Monobactam	Parent
Carbapenem	Imipenem	Carbapenem	Parent
	Meropenem	Carbapenem	Parent
Quinolone	Besifloxacin	Quinolone	Parent
	Danofloxacin	Quinolone	Parent
	Enrofloxacin	Quinolone	Parent
	Gatifloxacin	Quinolone	Parent
	Lomefloxacin	Quinolone	Parent
	Moxifloxacin	Quinolone	Parent
	Nadifloxacin	Quinolone	Parent
	Nalidixic acid	Quinolone	Parent
	Norfloxacin	Quinolone	Parent
	Prulifloxacin	Quinolone	Parent (prodrug)
	Sarafloxacin	Quinolone	Parent
	Hydroxy-norfloxacin	Quinolone	Metabolite
	Ofloxacin N-oxide	Quinolone	Metabolite
	Desmethyl-ofloxacin	Quinolone	Metabolite
	Ulifloxacin	Quinolone	Metabolite
TB (1st line)	Isoniazid	Isoniazid	Parent
	Pyrazinamide	Pyrazinamide	Parent
	Ethambutol	Ethambutol	Parent
	Rifampicin	Rifamycin	Parent
	Rifabutin	Rifamycin	Parent
	Isonicotinic acid	Isoniazid	Metabolite
	Acetyl-isoniazid	Isoniazid	Metabolite
	5-Hydroxy-pyrazinoic acid	Pyrazinamide	Metabolite
	25-desacetyl rifampicin	Rifamycin	Metabolite
	25-O-desacetyl rifabutin	Rifamycin	Metabolite
TB (MDR)	Capreomycin IA ≠	Aminoglycoside	Parent

Grouping	Chemical	Class	Target type
	Capreomycin IB ≠	Aminoglycoside	Parent
	Gentamycin C1 ≠	Aminoglycoside	Parent
	Gentamycin C1a ≠	Aminoglycoside	Parent
	Gentamycin C2 C2a C2b ≠	Aminoglycoside	Parent
	Kanamycin A ≠	Aminoglycoside	Parent
	Streptomycin A ≠	Aminoglycoside	Parent
	D-cycloserine	Isoxazole	Parent/Metab.
TB (other)	Delamanid	Nitroimidazole	Parent
	Bedaquiline	Diarylquinoline	Parent
	Linezolid	Oxazolidinone	Parent
	Thalidomide	Thalidomide	Parent
Amphenicol	Chloramphenicol	Amphenicol	Parent
	Florfenicol	Amphenicol	Parent
	2-Amino-1-(4-nitrophenyl)-1,3-propanediol	Amphenicol	Metabolite
Cycline	Doxycycline	Cycline	Parent
	Oxytetracycline	Cycline	Parent
Nitrofuran	Nitrofurantoin	Nitrofuran	Parent
	1-(2-nitrobenzylidenamino)-2,4-imidazolidinedione	Nitrofuran	Metabolite
Azole	Metronidazole	Azole	Parent
	Hydroxy-metronidazole	Azole	Metabolite
Antiretroviral	Emtricitabine	ARV	Parent
	Lamivudine	ARV	Parent

5.2.3 Sequencing

All samples (n=25) underwent metagenomic sequencing on the NovaSeq6000 (Illumina, San Diego, USA) 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 prior ultradeep sequencing and rarefaction analyses[90] demonstrating this as the minimum depth required to capture most AMR gene diversity in samples from one of the WWTWs sampled in this study.

5.2.4 Contemporaneous primary care antibiotic prescribing data

Contemporaneous monthly primary care practice prescribing data based on the National Health Service Business Services Authority English Prescribing Dataset was extracted from OpenPrescribing[154]. Practices were limited to those in close proximity to 4/5 (Figure 20) selected WWTWs (n=6), excluding the largest WWTW (Oxford) whose catchment area and associated practices could not be reliably determined. Data were extracted for all antibacterial drugs using the British National Formulary section 5.1 as well as antibacterial classes corresponding to the chemical residue analysis targets (i.e. macrolides, tetracyclines, cephalosporins, nitroimidazoles, sulfonamides). Denominators of total registered patients were also extracted to calculate items prescribed per 1,000 patients. Two-monthly mean prescribing metrics (raw items and per 1,000 patients) were calculated for each WWTW and timepoint, generally representing one month either side of wastewater sampling date.

5.2.5 Metagenomic analysis

5.2.5.1 AMR profiling

To address the difference in sequencing effort across samples which varied between 92 to 161 million reads, all samples were randomly subsampled to 92,289,277 reads (the minimum observed). Reads were subsequently processed using ResPipe[90] (https://gitlab.com/hsgweon/ResPipe; v.1.6.0 with Comprehensive Antibiotic Resistance Database [CARD] v.3.2.3)[94] for AMR gene profiling. AMR gene profiles for all read sets were normalised using Fragments Per Kilobase Million (FPKM); briefly, the relative abundance of each AMR gene was multiplied by 1,000,000 to give a “Fragment/Reads Per

Million". Next, to account for the variation in reference AMR gene lengths, the FPMs of each AMR gene were divided by gene length in kb to result in FPKM[90].

The curation of the CARD database ensures that each AMR gene entry is supported by clear experimental evidence of elevated minimum inhibitory concentrations (MIC) against defined antibiotics[155]. This metadata was used to map and aggregate FPKMs into categories associated with resistance to specific antibiotic or antibiotic class.

5.2.5.2 Taxonomic profiling

Taxonomic profiling was conducted using an updated version of ResPipe (v1.4.1) which incorporates established approaches of Kraken2[110] and Bracken[156] which are not part of the ResPipe v.1.6.0 workflow. Bracken-adjusted counts for taxonomic ranks from phylum to species were assessed using PhyloSeq[127] v1.38.0 and DESeq2[128] v1.34.0 to address heteroscedasticity. Counts for each taxonomic rank underwent parametric, local and mean model fitting to identify best rank and approach combination. Dispersion estimates for the rank of genus with local model fitting produced the best fit and was further supported by reduced sparsity. The alternate estimation of size factor method was also employed to account for zero inflation of metagenomic count data following established recommendations[129].

Although rarefaction of taxonomic count data is controversial for differential abundance estimation[129], normalisation remains useful for visualisation of taxonomic distributions

between samples. For this, I used an improved normalisation method (scaling with ranked subsampling) which has demonstrated improvements in reproducibility, preservation of taxonomic frequencies and alpha diversity over random subsampling without replacement[157]. Briefly, this approach scales counts to the sample with the lowest sequencing effort (analogous to relative abundance transformation) before processing scaled counts into integers whilst preserving population structure and total count.

5.2.6 Statistical analysis

Principal Coordinate Analysis was conducted on Bray-Curtis dissimilarity matrices generated for AMR gene FKPM and variance stabilised counts of genera using vegan[132] v2.6.2. Hierarchical clustering was performed using factoextra[130] v1.0.7 and FactoMineR[131] v2.4 to identify clusters and illustrate ellipses. The association between resistome and taxonomic dissimilarities was assessed using Procrustes analysis.

5.2.6.1 Differential abundance testing

Differentially abundant taxa were identified using a consensus approach combining DESeq2 and ALDEx2[158] following recommendations that a single testing method alone may be insufficient for ensuring robust and reproducible interpretations[159]. My consensus approach combined the higher statistically powered DESeq2 with the conservative but more accurate ALDEx2[160,161], and avoids high false discovery rates seen in methods such as edgeR and LEfSe[159]. Features identified by both methods are therefore robust to variations in analysis and are more confidently attributable to true sample-wise

differences. The same dataset, study design formula focussed on WWTW, and comparator groupings (i.e. with reference to Timepoint T1 and the Oxford WWTW) were used for both methods. Timepoint T1 and Oxford WWTW were used as references to provide comparison to the start of the sampling period and the largest WWTW site respectively. Significance thresholds of adjusted p-values <0.05 indicating modest evidence and <0.01 indicating strong evidence were used for DESeq2 (Benjamini-Hochberg adjustment, FDR <0.1) while an effect size =>1 was used for ALDEx2 with 200 Dirichlet Monte-Carlo Instances as per developer recommendations[162].

5.2.6.2 Generalised linear modelling

Generalised linear models were used to assess the relationship between sample-level AMR gene abundance, primary care prescribing and antibiotic residue measurements. Four similar models were developed to assess both indirect and direct effects. Indirect effect models used unrestricted AMR abundance as the outcome (dependent variable), and unrestricted antibacterial drug prescribing and antibiotic residue concentrations as the exposures (independent variables), while outcomes and exposures for direct effect models was limited to six AMR classes targeted by chemical analysis (macrolides, tetracyclines, cephalosporins, sulfonamides, nitroimidazole) (separate models for each class). AMR abundance data comprised total FPKM (indirect models) or aggregated FPKM of genes conferring resistance to a specific antimicrobial class (direct models). Prescribing data was either all antibacterial drugs (indirect models) or restricted to each antibacterial class of interest (direct models), respectively (see Methods 5.2.4). Antibiotic residue data was considered as the mean of concentration (µg/L), WWTW daily load (g/day) or

WWTW normalised daily load (mg/1000 inhabitants/day) either for all targets (indirect models) or restricted to targets of specified antimicrobial classes (direct models). Additional categorical variables included WWTW site (five levels) and timepoint (five levels) to facilitate assessment of variation attributable to each site or timepoint. The underlying error structure of each dependent variable was iteratively tested with Gaussian (untransformed, log10 and Box-Cox transformation), Inverse-gaussian and Gamma distributions to identify optimal model fit based on the Akaike Information Criterion and the Bayesian Information Criterion. Log10 transformation with Gaussian distribution was then used for all models, with coefficients and 95% confidence intervals exponentiated for interpretation as fold-change per comparison of different levels of each categorical variable or per unit higher continuous measurement. Before model fitting, collinearity of fixed effects was assessed using Spearman correlations, and only one variation of a given fixed effect utilised when rho >0.9. “Univariable” models considered each antibiotic prescribing and residue factor adjusting for time and WWTW site alone. Backwards elimination was then conducted for full models containing all covariates of interest with a p>0.1 exit criterion to identify associations between antibiotic prescribing and residues with at least some evidence (0.1 exit criteria used to facilitate control of confounding and given the relatively small sample size). WWTW and timepoint were forced into all models as these were design parameters.

5.3 Results

5.3.1 Distribution of taxa

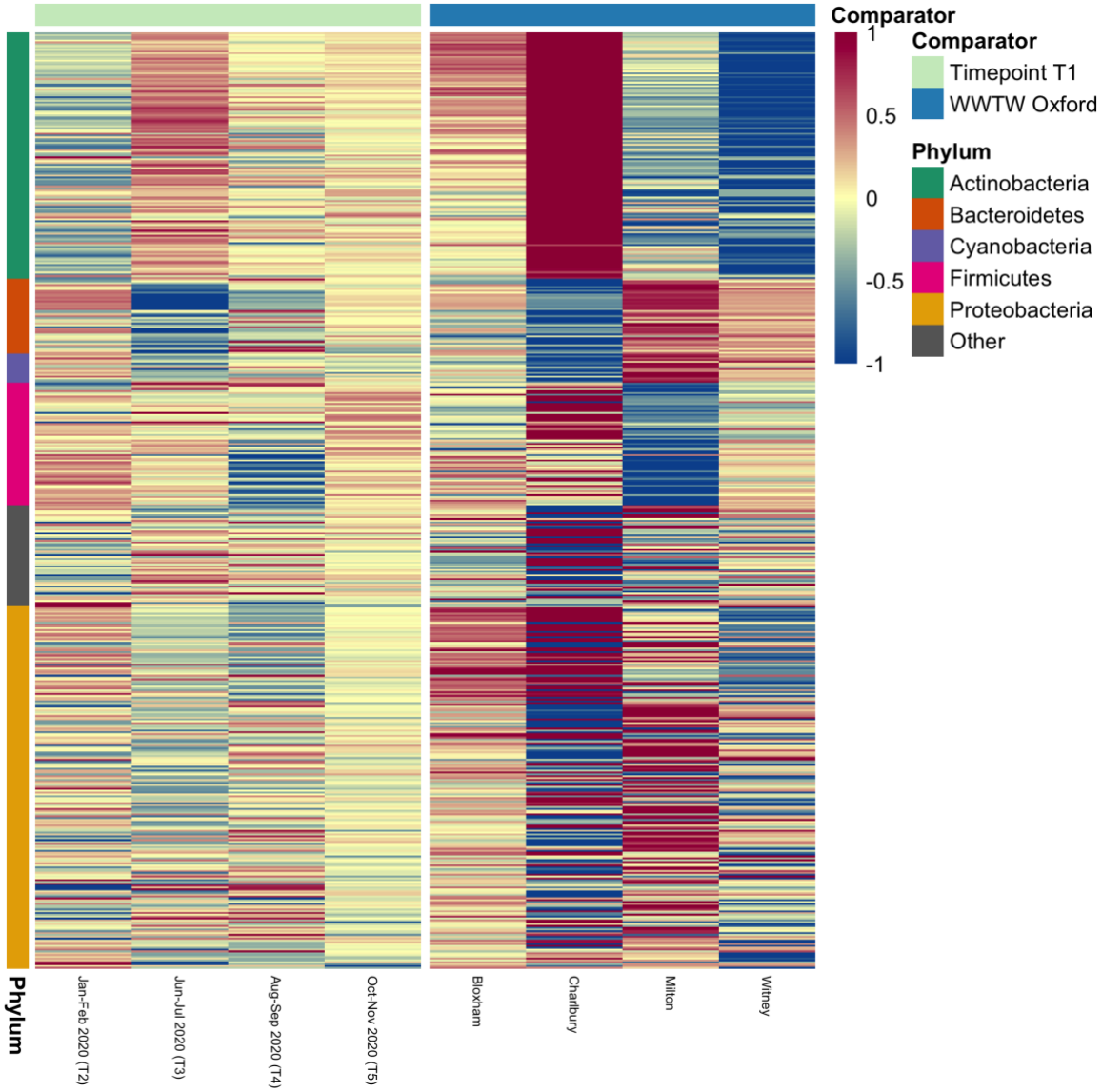
Wastewater influent was analysed from five sites over one year (five timepoints; 25 site-timepoints in total), covering one city, two towns and two villages in Oxfordshire. A total of 1501 unique genera were identified across all samples with similar richness between samples (range: 1300-1381). Proteobacteria dominated the taxonomic distributions of all samples where the genera *Acidovorax*, *Aeromonas*, *Pseudomonas*, *Acinetobacter* and *Faecalibacterium* were most abundant. Bray-Curtis dissimilarities did not differ significantly when stratifying by site or timepoint although dissimilarities were numerically lower between samples taken from the same site or timepoint (i.e. samples taken from the same site/timepoint were more similar as judged by the numeric similarity measures, although formally there was no statistical evidence of difference).

However, in contrast to comparisons using Bray-Curtis dissimilarities, a subset of genera showed significant differences in abundance between sites and/or timepoints as determined by consensus-based differential abundance testing. DESeq2 likelihood-ratio testing identified 896 affected genera (adjusted p-value <0.05; 763/896 with stronger evidence p<0.01) while the more conservative ALDEx2 identified 574. The two methods were concordant for 513 genera reflecting an 89% agreement between ALDEx2 and DESEq2 hits. Of these 513 genera, the most abundant were in the phyla Proteobacteria (199/513; 39%), Actinobacteria (135/513; 26%), Firmicutes (67/513; 13%) and Bacteroidetes (41/513; 8%) (Figure 22). Most differentially abundant genera were identified in site-level comparisons (median=110; IQR=85-170) rather than comparisons

between timepoints (19; 12-22) suggesting seasonal differences were relatively minor and experienced by all sites equally. With respect to the reference comparator WWTW Oxford, differentially abundant genera were most pronounced in samples from Charlbury (n=343/513 significantly different genera) followed by Milton-under-Wychwood (Milton n=112), Witney (n=108), and Bloxham (n=14). With respect to the reference timepoint, T1 (Nov-Dec), across all sites, the number of differentially abundant genera increased before falling over time (T2 [Jan-Feb]:21, T3 [Jun-Jul]:26, T4 [Aug-Sep]:16, T5 [Oct-Nov]:1). Interestingly only one genus reached the effect size threshold for T5 (Oct-Nov) which was otherwise closest to T1 (Nov-Dec) in terms of genus distributions and abundance, consistent with it being taken at a similar time of the year.

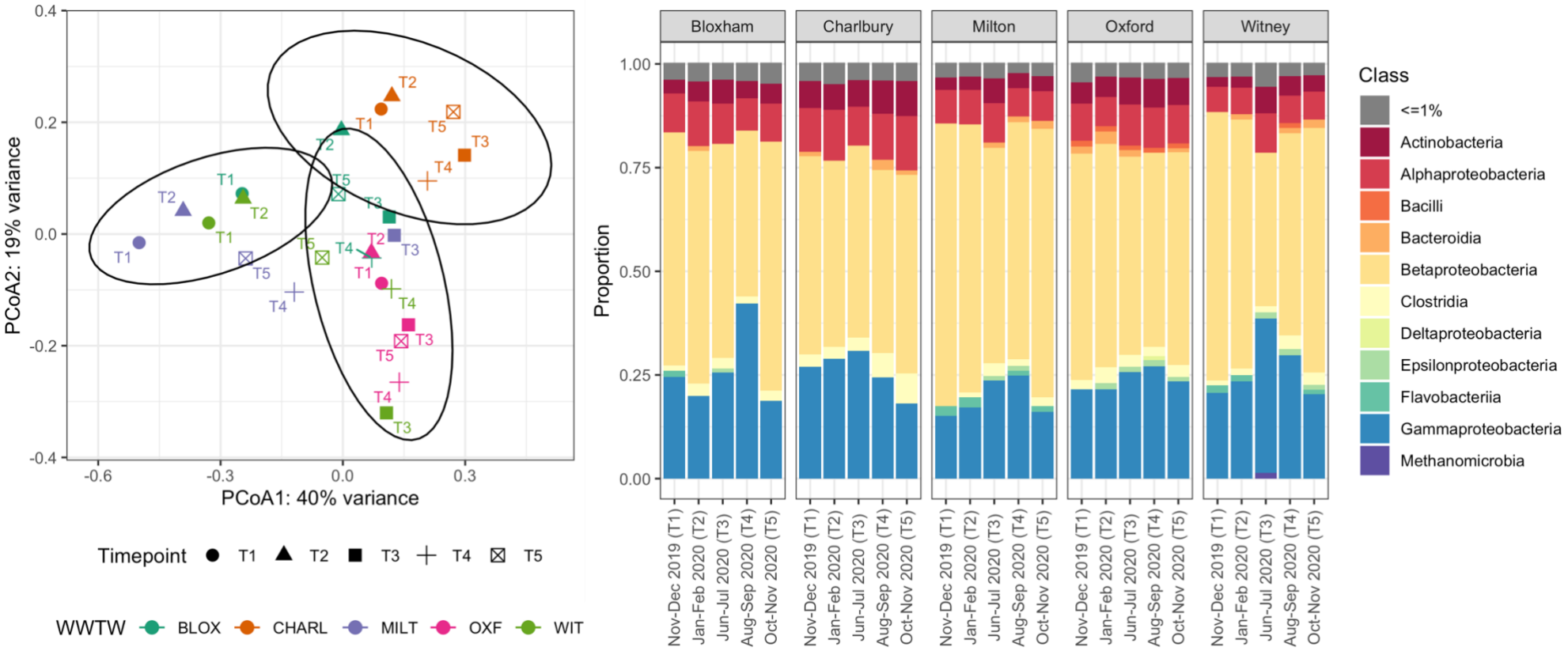
Figure 22 - Heatmap of significantly differentially abundant genera identified by consensus-based DESeq2 and ALDEx2 testing.

Columns reflect timepoint or WWTW group-level comparisons to T1 [Nov-Dec] (left) and WWTW Oxford (right) respectively. Phyla of genera are denoted by colour along the y-axis. Dark blue to red (-1 to 1) continuous scale reflects ALDEx2 standardized distributional effect size metric (i.e. number of standard deviations separating the mean of groups). Effect sizes have been limited to -1/1 to denote qualitative significance for ease of interpretation.



Differential abundance results were also reflected in the principal coordinate analysis of Bray-Curtis dissimilarities, with Charlbury samples exhibiting the most separation into a distinct cluster from Oxford samples (Figure 23). Further consistent was the grouping of Oxford and Bloxham samples into the same cluster, and several Milton-under-Wychwood and Witney samples grouping into a separate cluster.

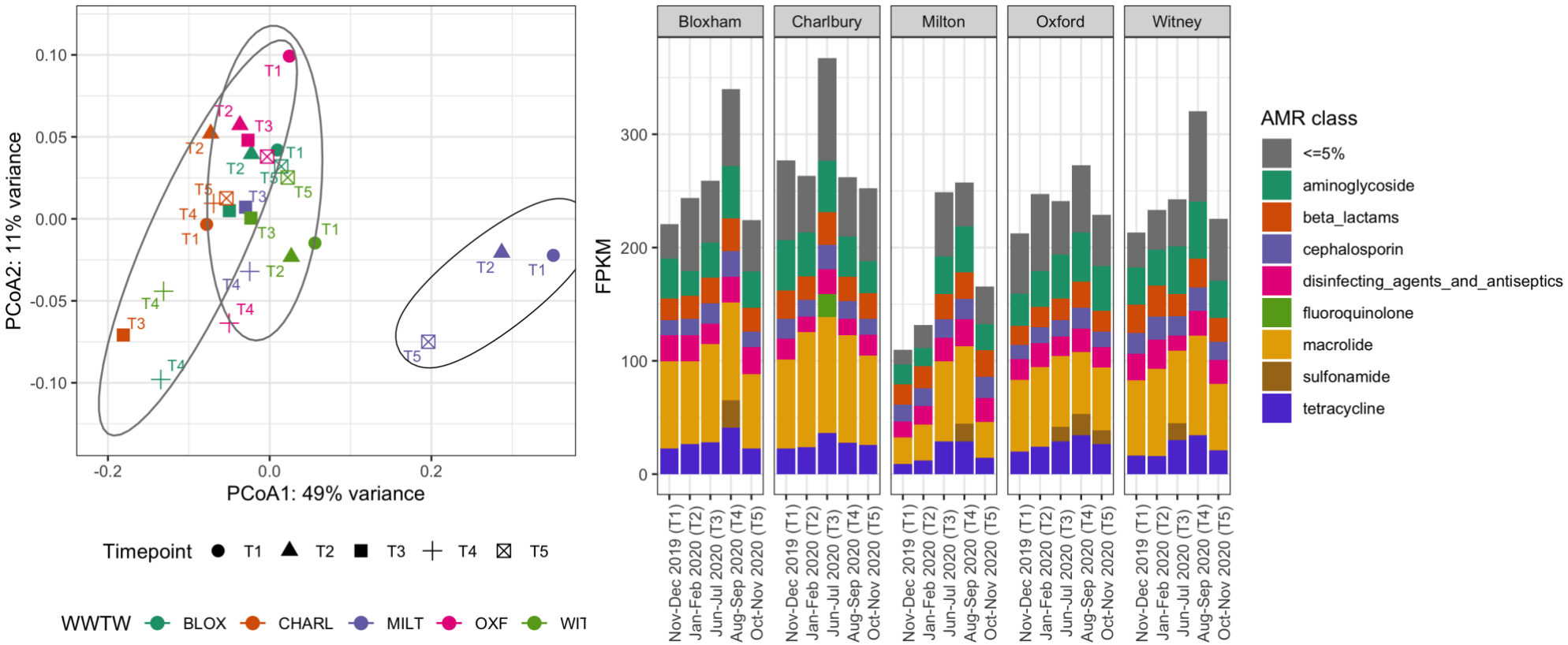
Figure 23 - Principal coordinate analysis (left) with hierarchical clusters (ellipses) and SRS-normalised relative abundance (right) of taxonomic composition



5.3.2 Distribution of AMR genes

Total AMR gene abundance and relative composition varied across timepoints and sites (Figure 24). Total abundance of AMR genes (fragments per kilo base per million fragments [FPKM]) across all sites was median 172 (IQR 164-189) at T1; rising to its highest point at T4 (median: 206; IQR:203-242), with statistically significant increases vs T1 at T4 as well as T3 (190; 189-200) and T2 (182; 174-188), before returning to a similar level as T1 at T5 (174; 166-175) (significance determined by two-way ANOVA [$p<0.01$]). Charlbury WWTW had the highest abundance (207; 206-211), significantly higher than Oxford WWTW (182; 175-186) and Milton-under-Wychwood (124; 102-189) which was also significantly lower than Bloxham (189; 174-200), Witney (188; 172-191) and Charlbury (207; 206-211). A total of 781/4870 unique AMR genes were identified across all sites and timepoints, representing 154/374 different AMR gene families represented in the reference database (CARD). Richness of AMR genes was significantly higher for samples collected from Oxford (413; 407-419) and during timepoint T4 (423; 413-433). Since different AMR genes may confer resistance to the same antimicrobial, FPKM abundances were consolidated into resistance scores corresponding to antimicrobial classes. Despite significant differences in total AMR gene abundance and richness between sites and timepoints, relative class-level resistance profiles were similar for most samples (Figure 24).

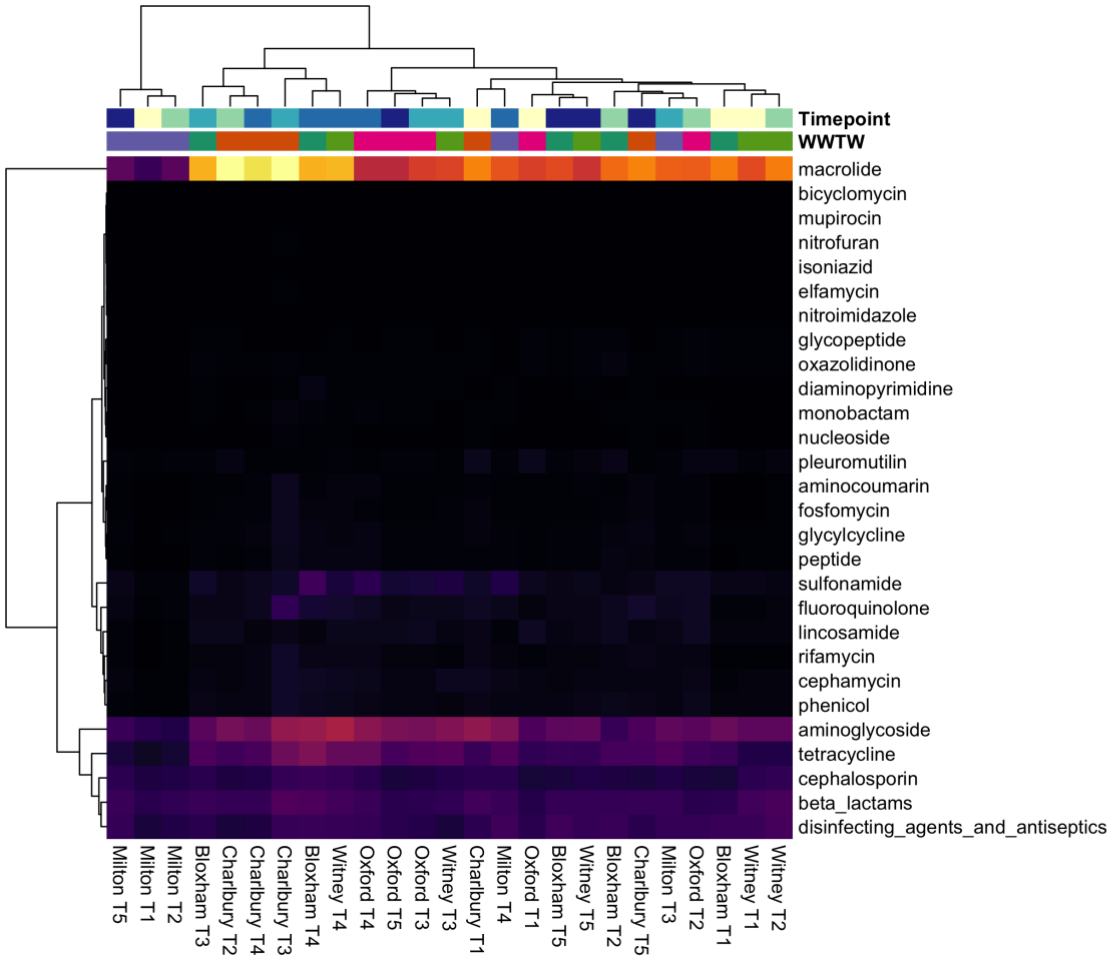
Figure 24 - Principal coordinate analysis of resistome composition (left) with hierarchical clusters (ellipses) and aggregated FPKM of AMR genes conferring resistance to antimicrobial classes (right).



AMR genes conferring resistance to macrolides, aminoglycosides, beta-lactams, tetracyclines and cephalosporins were most prevalent across samples. Of these, macrolide resistance consistently comprised the largest proportion of AMR genes across all samples, however, this was notably reduced in three samples from Milton-under-Wychwood. Interestingly, sulfonamide resistance genes intermittently comprised >5% of AMR gene totals in six samples, of which three were collected from Oxford. Hierarchical clustering of class-level resistance scores across samples reflected the reduced AMR abundance in Milton-under-Wychwood samples at T1, T2 and T5. Otherwise, sample resistomes did not clearly cluster by timepoint or WWTW (Figure 24 and Figure 25). However, sample resistomes were significantly and strongly associated with taxonomic composition (Procrustes: $p<0.001$; Correlation: 0.85).

Figure 25 - Heatmap of AMR gene FPKM aggregated by antimicrobial class

Each column reflects a separate wastewater sample where WWTW and timepoint are labelled on the x-axis (bottom) and annotated by colour (top).



5.3.3 Generalised linear modelling

A series of generalised linear models were fitted to investigate the associations between AMR gene abundance in samples, contemporaneous primary care prescribing of antibacterial drugs and levels of antibiotic residues in samples (see Methods 5.2.6). Both indirect and direct selection of resistance were assessed by including fixed effects of total or class-specific antibacterial drug prescribing and antibiotic residue measurements in

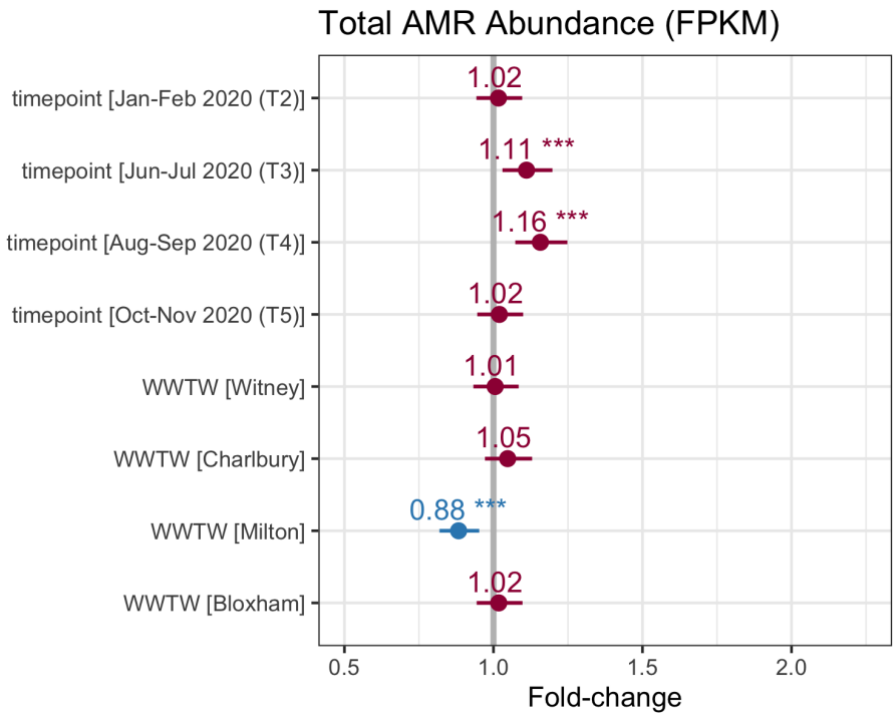
models of total AMR gene abundance or abundance of AMR genes conferring resistance to the relevant class respectively.

5.3.3.1 Association between wastewater AMR abundance, local primary care prescribing and wastewater antibiotic residues

No evidence of association was found between total AMR abundance and total antibacterial prescribing (1.01-fold higher per 1 item higher prescribing per 1000 patients (95% CI 0.99-1.02)) or residues (1.29-fold higher per 1 µg/L higher residue concentrations (95% CI 0.73-2.30)), suggesting that total prescribing (as a proxy for usage), wastewater antibiotic residues, and indirect selection (e.g. cross-resistance) were relatively minor drivers of total AMR abundance at this geographical scale. However, there was significant variation in total AMR abundance across timepoints and WWTWs in a multivariable model adjusting for both simultaneously, with significantly higher total AMR abundance at T3 (fold-change: 1.11; 95% CI: 1.03-1.20; p=0.01) and T4 (1.16; 1.07-1.25; p=0.002) compared with T1, and significantly lower levels in samples from Milton-under-Wychwood (0.88; 0.82-0.95; p=0.005) (Figure 26), consistent with the influence of other unmeasured environmental variables (e.g. climate).

Figure 26 - Generalised linear model assessing AMR against timepoint and WWTW.

Effect estimates (and 95% confidence intervals) from a generalised linear model assessing the relationship between the outcome of total AMR gene abundance, timepoint of sampling as compared to T1 (design parameter) and the WWTW sampled as compared to Oxford (design parameter). Factors associated with higher or lower fold-change are coloured in red and blue respectively. Asterisks denote level of significance where 0.01 (***) reflects strong evidence, 0.05 (**) moderate evidence and 0.1 (*) weak evidence.

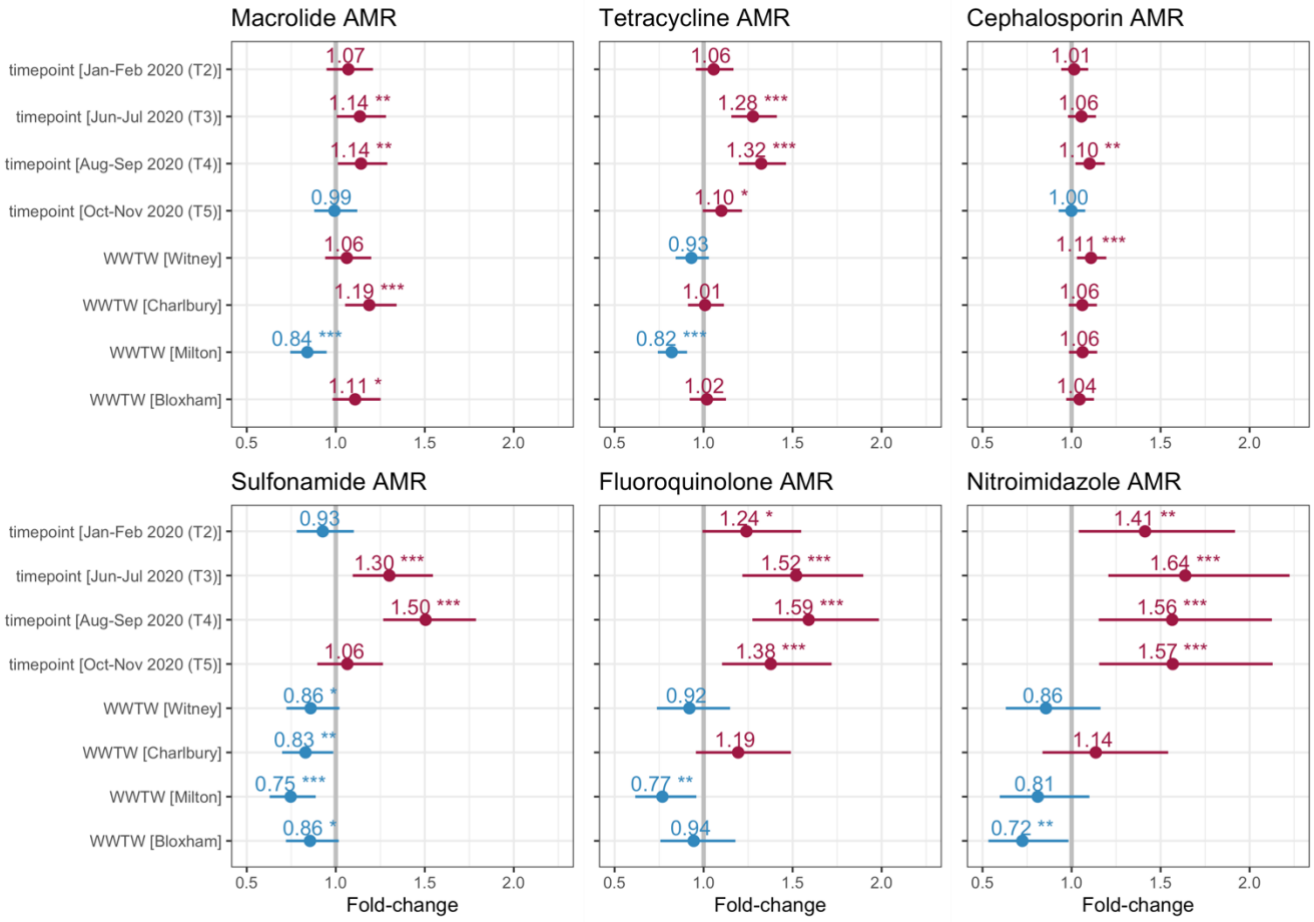


Similarly, in models evaluating antimicrobial class-specific impacts (i.e. direct selection), only the effects of timepoint and WWTW remained significant for 3/6 antimicrobial classes assessed (cephalosporins, fluoroquinolones, sulfonamides), with variable timepoint or WWTW effect sizes (i.e. estimated fold-change) depending on AMR class (Figure 27). Timepoints T3 (Jun-Jul 2020) and T4 (Aug-Sep 2020) were most commonly significantly enriched vs T1 (Nov-Dec 2019) for all six AMR classes with the largest effect for sulfonamide (fold-change: 1.50; 95% CI: 1.27-1.79; p<0.01, fluroquinolone (1,59; 1.27-1.98; p<0.01) and nitroimidazole (1.56; 1.15-2.13; p=0.01) resistance at T4. For WWTWs, Milton-under-Wychwood had significantly lower AMR abundance vs Oxford against macrolides (0.84; 0.74-0.95; p=0.01), tetracyclines (0.82; 0.74-0.90; p<0.01), sulfonamides

(0.75; 0.63-0.89; p<0.01) and fluoroquinolones (0.77; 0.62-0.96; p=0.03). Across timepoints, the largest fold-changes were observed for nitroimidazole resistance; then fluoroquinolone, sulfonamide, tetracycline, macrolide and cephalosporin. Across WWTWs, fold-changes were greatest for resistance to sulfonamides; then nitroimidazoles, fluoroquinolones, tetracyclines, macrolides and cephalosporins.

Figure 27 - Generalised linear models assessing class-level AMR against timepoint and WWTW

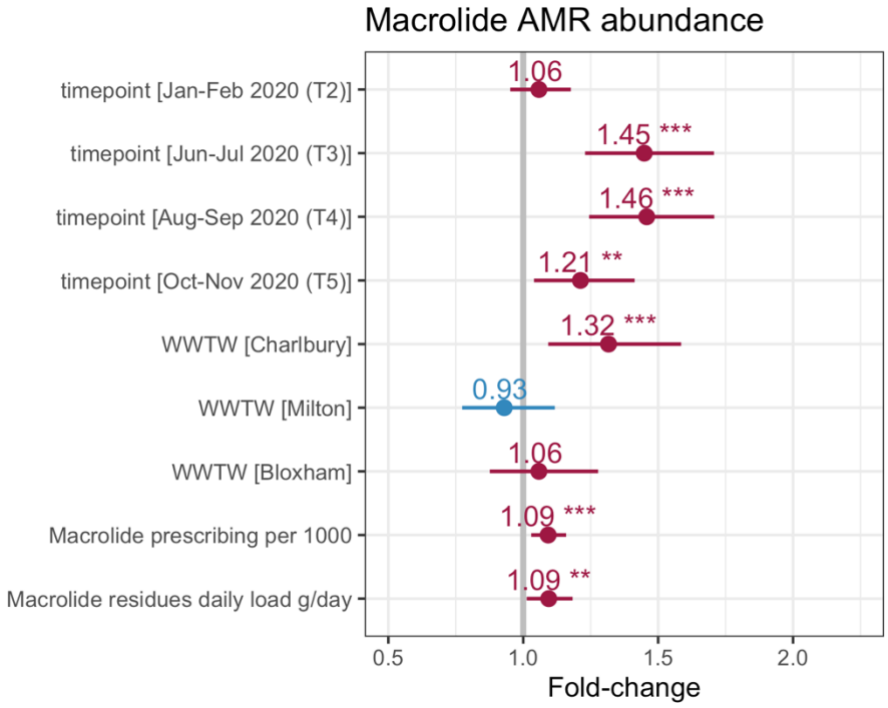
Effect estimates (and 95% confidence intervals) from generalised linear models assessing the relationship between the outcomes of AMR gene abundance for six antimicrobial classes, timepoint of sampling as compared to T1 (design parameter) and the WWTW sampled as compared to Oxford (design parameter). Factors associated with higher or lower fold-change are coloured in red and blue respectively. Asterisks denote level of significance where 0.01 (***) reflects strong evidence, 0.05 (**) moderate evidence and 0.1 (*) weak evidence.



For three antibiotic classes (macrolides, nitroimidazoles, tetracyclines) there was statistical evidence that either antibiotic residues or antibiotic prescribing had independent effects on AMR gene abundance, after adjusting for the effects of timepoint and WWTW shown above and after backwards elimination for all class-specific antibiotic prescribing and residue factors (Table 3). For macrolides, significant independent effects on wastewater FPKM of macrolide resistance genes were observed for both macrolide prescribing (1.09-fold higher per item higher prescribing per 1000 patients; 95% CI 1.01-1.18; p=0.01 [strong evidence]) and daily load of wastewater macrolide residues (1.09 per g/day higher; 1.01-1.18; p=0.04 [moderate evidence]) (Figure 28). Of note, after adjusting for these factors the lower macrolide AMR abundance associated with Milton-under-Wychwood attenuated substantially, suggesting that the overall lower macrolide AMR abundance in this WWTW could be explained by lower macrolide prescribing and/or residues. In contrast, the effect of timepoint increased, suggesting that – given changes in prescribing of macrolides over time – the effect of season on macrolide AMR abundance was as large as for other classes.

Figure 28 - Generalised linear model assessing macrolide AMR against timepoint, WWTW, macrolide prescribing and residue levels.

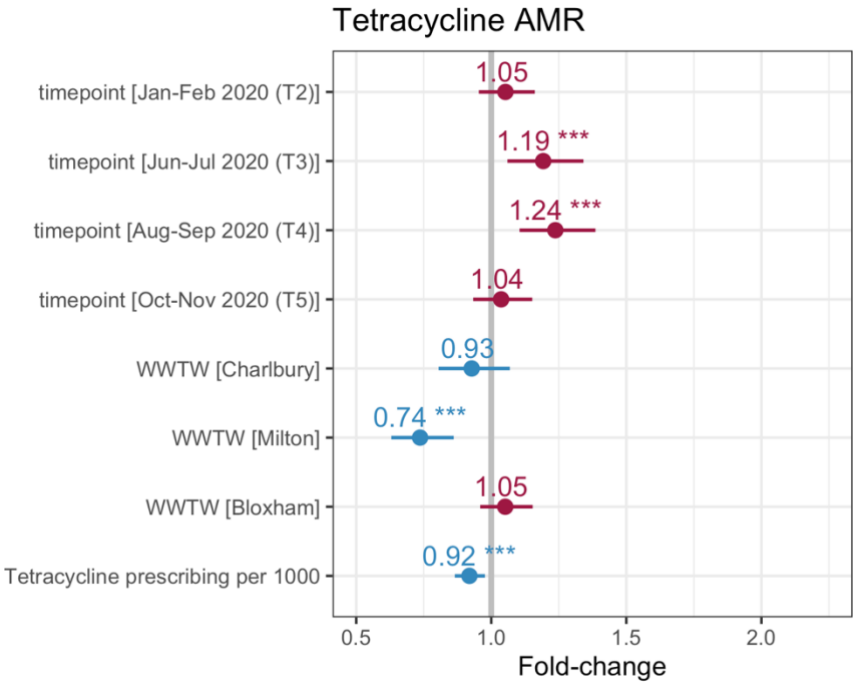
Effect estimates (and 95% confidence intervals) from generalised linear model assessing the relationship between the outcome of AMR gene abundance for genes conferring macrolide resistance, timepoint of sampling as compared to T1 (design parameter), WWTW sampled as compared to Witney (design parameter), contemporaneous primary care prescribing of macrolide items per 1000 patients (exposure) and measurements of macrolide residue daily load (exposure). Witney was used as the WWTW comparator since no prescribing data from Oxford was available and Witney represents the second largest WWTW sampled. Factors associated with higher or lower fold-change are coloured in red and blue respectively. Asterisks denote level of significance where 0.01 (***) reflects strong evidence, 0.05 (**) moderate evidence and 0.1 (*) weak evidence.



Interestingly, tetracycline resistance gene abundance was significantly negatively associated with tetracycline prescribing (Figure 29) although with a relatively small effect size (0.92 fold lower per one item higher tetracycline prescribing per 1000 patients; 95% CI 0.87-0.98; p=0.019).

Figure 29 - Generalised linear model assessing tetracycline AMR against timepoint, WWTW and tetracycline prescribing.

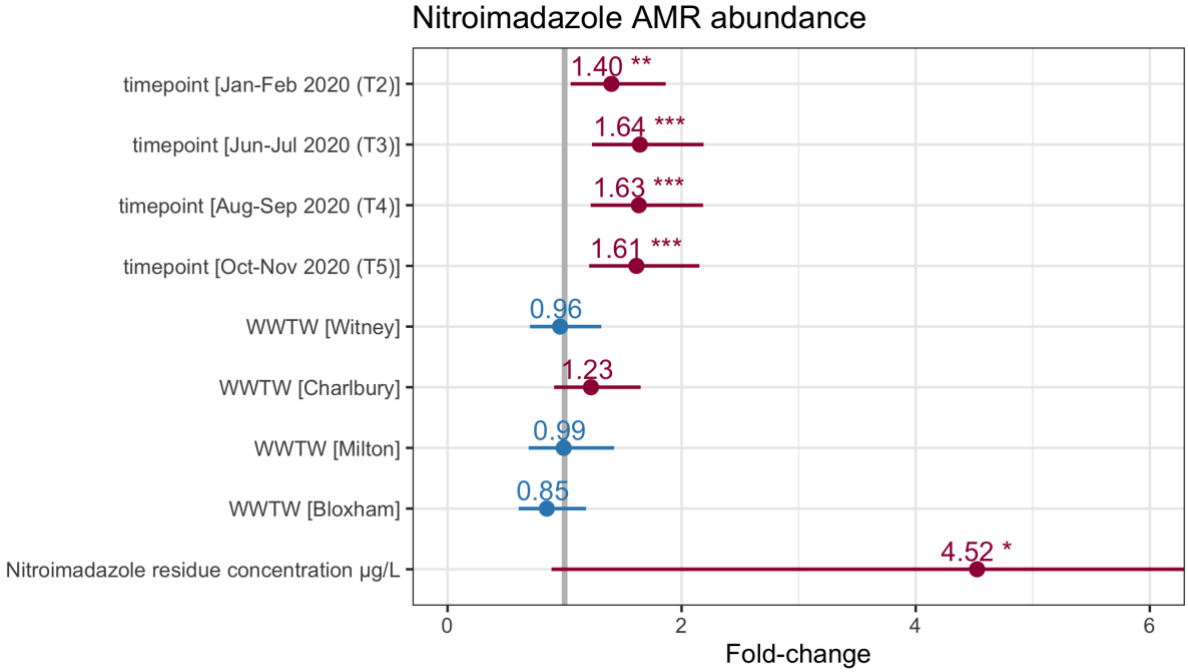
Effect estimates (and 95% confidence intervals) from generalised linear model assessing the relationship between the outcome of AMR gene abundance for genes conferring Tetracycline resistance, timepoint of sampling as compared to T1 (design parameter), WWTW sampled as compared to Witney (design parameter), prescribing of tetracycline items per 1000 patients (exposure). Witney was used as the WWTW comparator since no prescribing data from Oxford was available and Witney represents the second largest WWTW sampled. Factors associated with higher or lower fold-change are coloured in red and blue respectively with 95% confidence intervals. Asterisks denote level of significance where 0.01 (***) reflects strong evidence, 0.05 (**) moderate evidence and 0.1 (*) weak evidence.



For nitroimidazoles, there was very weak evidence suggesting a very large effect of nitroimidazole residue concentrations on nitroimidazole resistance gene abundance (4.52-fold higher per 1 µg/L increase; 0.89-23.05; p=0.09 (Figure 30).

Figure 30 - Generalised linear model assessing nitroimidazole AMR against timepoint, WWTW and nitroimidazole residues.

Effect estimates (and 95% confidence intervals) from generalised linear model assessing the relationship between the outcome of AMR gene abundance for genes conferring nitroimidazole resistance, timepoint of sampling as compared to T1 (design parameter), WWTW sampled as compared to Oxford (design parameter), and measurements of nitroimidazole residue concentration (exposure). Factors associated with higher or lower fold-change are coloured in red and blue respectively with 95% confidence intervals. Asterisks denote level of significance where 0.01 (***) reflects strong evidence, 0.05 (**) moderate evidence and 0.1 (*) weak evidence. x-axis is truncated at to 6 (absolute value given in text)



There was no statistical evidence supporting direct selection of class-level AMR abundance for cephalosporins, fluoroquinolones, sulfonamides even when prescribing and residue data were fitted adjusting only for timepoint and WWTW (Table 3). As expected, these minimally adjusted respective exposures were significantly associated with macrolide, nitroimidazole and tetracycline AMR abundance outcomes as per respective multivariable models. However, a nonsignificant large effect was observed for macrolide residue concentrations (fold-change in class-level AMR abundance per 1 µg/L higher respective residue concentration: 1.52 [95% CI: 0.72-3.21]). Interestingly, this was also observed for tetracycline (1.43 [0.44-4.67]) and sulfonamide (1.22 [0.51-2.92]) AMR abundance. This effect was smaller for cephalosporins (1.06 [0.92-1.22]) but could not be

assessed for fluoroquinolones due to lack of a respective residue target. However, fluroquinolone prescribing per 1000 patients showed a similar effect size (1.24 [0.76-2.05]). However, backwards elimination showed that these associations were better explained by residue concentration rather than residue load.

Table 3 - Table of coefficients for univariable models assessing associations between prescribing and residues with class-level AMR abundance

Each exposure was fitted adjusting for design variables (WWTW and timepoint) – only coefficients for the exposure are presented here. Asterisks denote level of significance where 0.01 (***) reflects strong evidence, 0.05 (**) moderate evidence and 0.1 (*) weak evidence. Gray shading shows factors that were independently associated after backwards elimination considering all prescribing and residue factors.

	Coefficient	p-value	Fold-change (CI)
Macrolide			
Prescribing items	0.01	0.03**	1.01 (1.00-1.02)
Prescribing items per 1000	0.07	0.06*	1.07 (1.00-1.15)
Residues conc. (µg/L)	0.42	0.29	1.52 (0.72-3.21)
Residues load (g/day)	0.03	0.03**	1.03 (1.00-1.05)
Residues norm. load (mg/1000 inhabitants/day)	0.001983	0.099*	1.00 (1.00-1.00)
Nitroimidazole			
Prescribing items	-0.03	0.64	0.97 (0.88-1.08)
Prescribing items per 1000	-0.05	0.91	0.95 (0.41-2.20)
Residues conc. (µg/L)	1.51	0.09*	4.52 (0.89-23.05)
Residues load (g/day)	0.07	0.16	1.08 (0.98-1.19)
Residues norm. load (mg/1000 inhabitants/day)	0.005	0.06*	1.01 (1.00-1.01)
Tetracycline			
Prescribing items	-0.01	0.04**	0.99 (0.98-1.00)
Prescribing items per 1000	-0.08	0.02**	0.92 (0.87-0.98)
Residues conc. (µg/L)	0.36	0.56	1.43 (0.44-4.67)
Residues load (g/day)	0.02	0.35	1.02 (0.98-1.05)
Residues norm. load (mg/1000 inhabitants/day)	0.0009316	0.58	1.00 (1.00-1.00)
Cephalosporin			

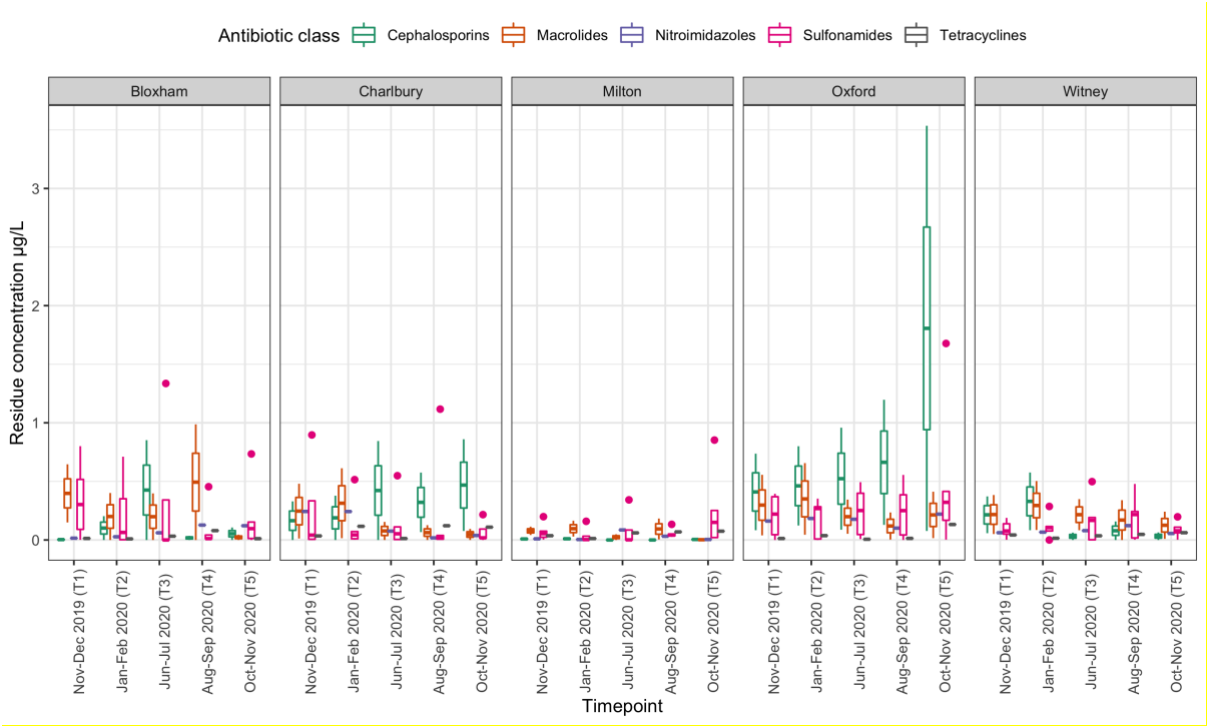
Prescribing items	-0.01	0.55	0.99 (0.97-1.02)
Prescribing items per 1000	-0.00352	0.97	1 (0.81-1.22)
Residues conc. (µg/L)	0.06	0.43	1.06 (0.92-1.22)
Residues load (g/day)	0.0007404	0.65	1.00 (1.00-1.00)
Residues norm. load (mg/1000 inhabitants/day)	0.0001411	0.58	1.00 (1.00-1.00)
Sulfonamide			
Prescribing items	-0.02	0.13	0.98 (0.96-1.00)
Prescribing items per 1000	-0.13	0.12	0.88 (0.76-1.02)
Residues conc. (µg/L)	0.20	0.66	1.22 (0.51-2.92)
Residues load (g/day)	0.01	0.40	1.01 (0.99-1.03)
Residues norm. load (mg/1000 inhabitants/day)	-0.00007626	0.96	1.00 (1.00-1.00)
Fluoroquinolone			
Prescribing items	0.04	0.37	1.04 (0.96-1.12)
Prescribing items per 1000	0.22	0.41	1.24 (0.76-2.05)

5.3.3.2 Antibiotic residue concentrations significantly vary across WWTWs

To assess the potential for AMR selection in wastewater across Oxfordshire, I focussed on concentration measurements of antibiotic targets with reference to established predicted no-effect concentrations (PNECs). Concentrations of antibiotic targets across the five antimicrobial classes detected fluctuated by site and timepoint (Figure 31). Across all timepoints and residues, concentrations (µg/L) were highest in wastewater from Oxford (median: 0.18; IQR: 0.04-0.38) and lowest from Milton-under-Wychwood (0.02; 0.003-0.07). However, concentrations varied greatly by antibiotic class, with for example, cephalosporin concentrations being highest from Oxford (0.43; 0.09-0.91), then Charlbury (0.20; 0.02-0.53), Witney (0.06; 0.01-0.14), Bloxham (0.003; 0-0.09) and Milton-under-Wychwood (0; 0-0.007).

Figure 31 - Antibiotic residue concentrations across sites and timepoint.

Mean antibiotic residue concentrations for each antibiotic class are represented by coloured boxplots indicating median, upper and lower quartiles. Facets denote WWTW site where the x-axis indicates timepoint of sampling.

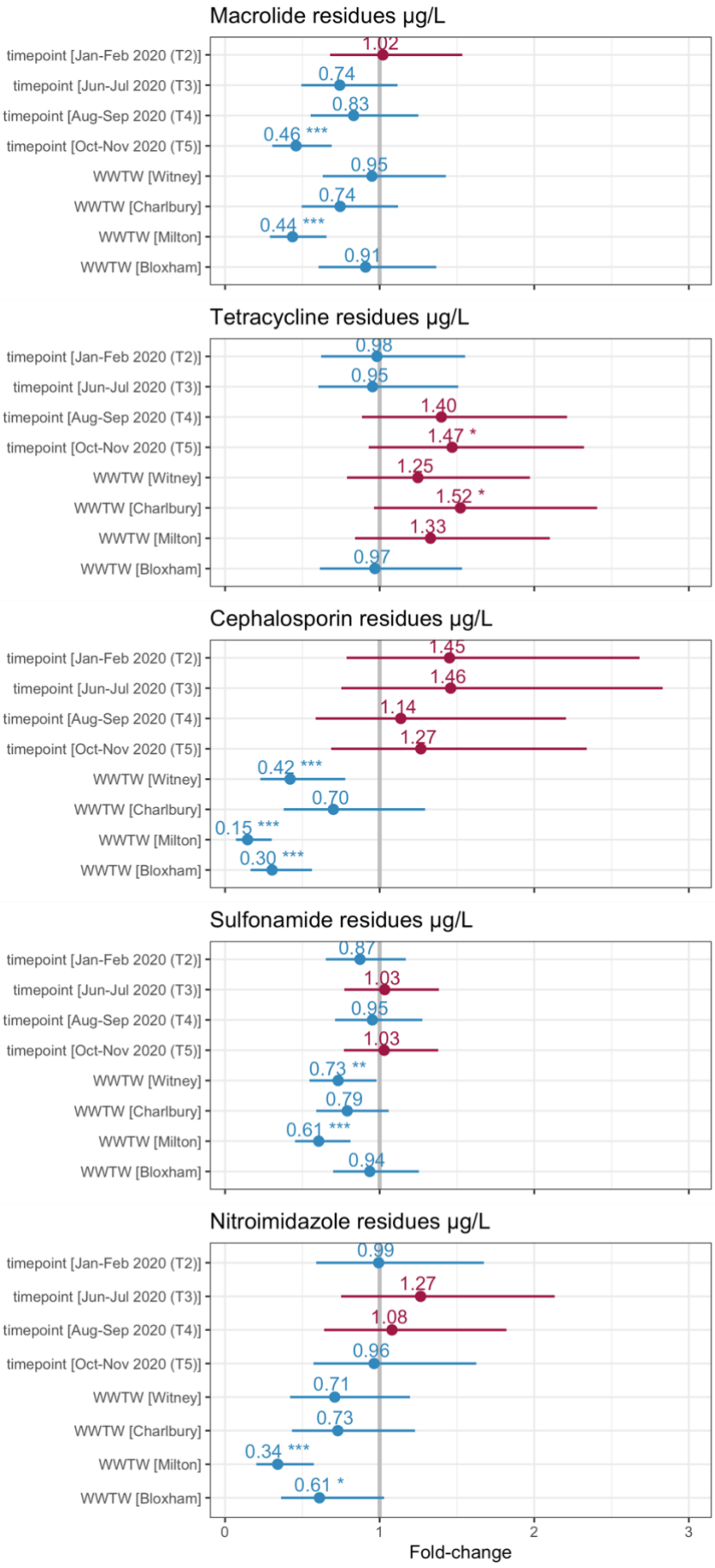


Multivariable modelling of class-level antibiotic residues as the outcome with WWTW, timepoint and class-level primary care prescribing as exposures identified significant effects for WWTW and timepoint only – that is, I found no evidence that class-level primary care prescribing was associated with class-level antibiotic residues. Timepoint was only significant for macrolide and tetracycline concentrations where there was strong evidence for lower macrolide residue levels (0.46 fold-change; 95% CI 0.31-0.69; p<0.01) at T5 (Oct-Nov 2020) vs T1 (Nov-Dec 2019), and very weak evidence for higher tetracycline residues at T5 vs T1 (1.47; 0.93-2.32; p=0.10). Concentrations of cephalosporins, nitroimidazoles and sulfonamides were only significantly associated with WWTWs. Compared to Oxford, Milton-under-Wychwood, Witney and Bloxham had strong evidence for large fold-decreases for at least one antibiotic class. Consistent with

significantly lower total AMR abundance (Figure 32), Milton-under-Wychwood was strongly associated ($p<0.01$) with lower residue concentrations for all classes except tetracyclines.

Figure 32 - Generalised linear model assessing class-level antibiotic concentrations against timepoint and WWTW

Effect estimates (and 95% confidence intervals) from generalised linear model assessing the relationship between the outcome of residue concentrations, timepoint of sampling (design parameter) and WWTW sampled (design parameter). Factors associated with higher or lower fold-change are coloured in red and blue respectively with 95% confidence intervals. Asterisks denote level of significance where 0.01 (***) reflects strong evidence, 0.05 (**) moderate evidence and 0.1 (*) weak evidence.



As expected from the multivariable models, there was no evidence of association between class-level antibiotic prescribing and antibiotic residues adjusting only for timepoint and WWTW. However, large non-significant effects were noted for cephalosporin (0.39 [95%CI 0.08-1.94]) and nitroimidazole (0.66 [95% CI 0.15-2.92]) prescribing per 1000 patients (Table 4).

Table 4 - Table of coefficients for univariable models assessing associations between class-level prescribing and residue concentrations.

Each exposure was fitted adjusting for design variables (WWTW and timepoint) – only coefficients for the prescribing exposures are presented here.

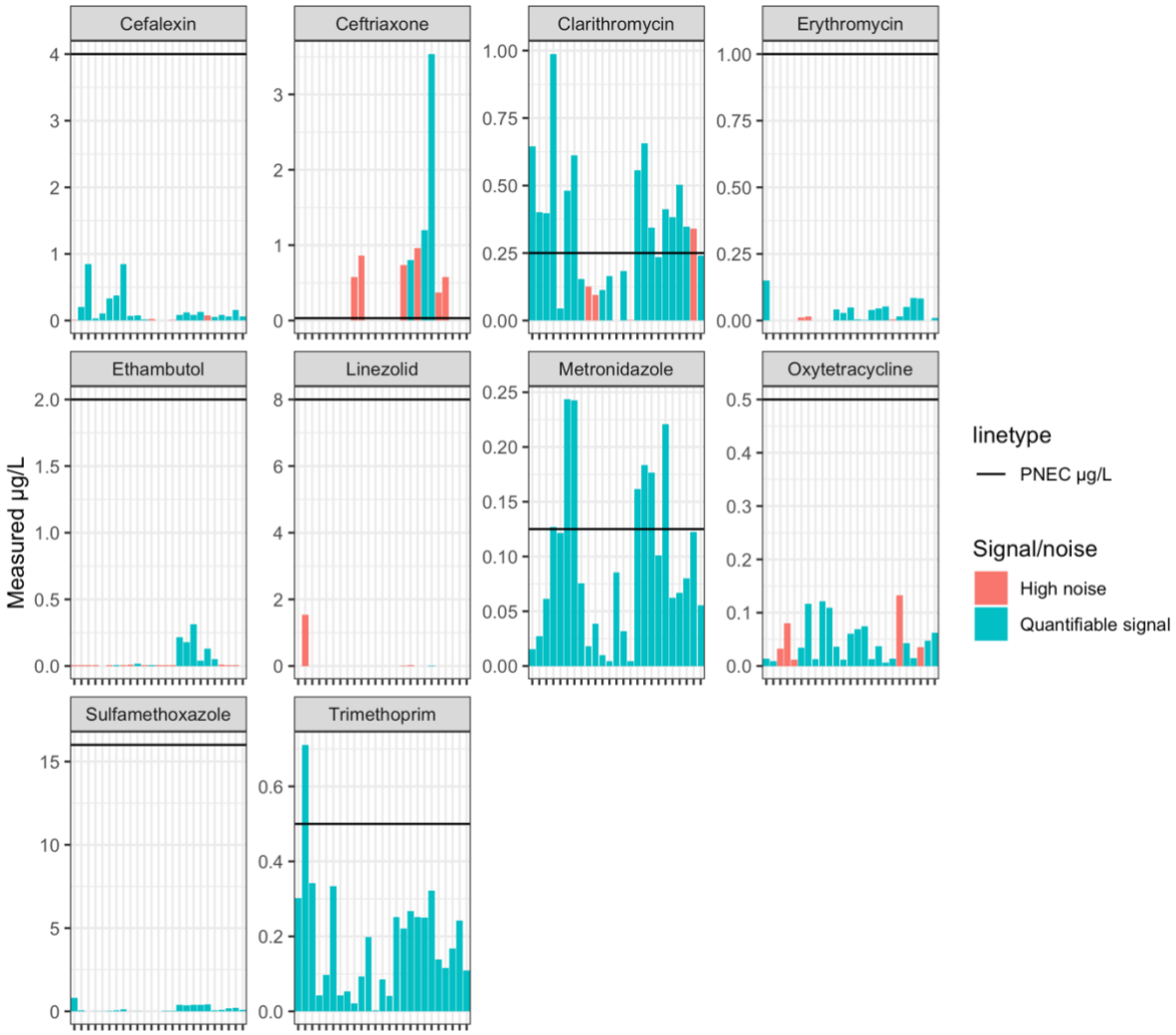
	Coefficient	p-value	Fold-change (CI)
Macrolide			
Prescribing items	-0.01	0.56	0.99 (0.96-1.02)
Prescribing items per 1000	-0.02	0.87	0.98 (0.76-1.26)
Tetracycline			
Prescribing items	0.02	0.35	1.02 (0.98-1.05)
Prescribing items per 1000	0.17	0.25	1.19 (0.90-1.56)
Cephalosporin			
Prescribing items	-0.12	0.22	0.89 (0.74-1.06)
Prescribing items per 1000	-0.94	0.28	0.39 (0.08-1.94)
Sulfonamide			
Prescribing items	-0.002157	0.93	1.00 (0.95-1.04)
Prescribing items per 1000	-0.05	0.71	0.95 (0.72-1.25)
Nitroimidazole			
Prescribing items	-0.05	0.62	0.95 (0.79-1.15)
Prescribing items per 1000	-0.41	0.6	0.66 (0.15-2.92)

5.3.4 Antibiotic residue concentrations reaching selective levels

Of the ten antibiotics detected directly, four exceeded previously reported predicted no-effect concentrations (PNECs)[153] in one or more wastewater sample (Figure 33). However, 7/31 measurements over PNEC were low-quality and may not be fully representative.

Figure 33 - Antibiotic residue concentration measurements for 10 antibiotics detected directly as their original chemical form.

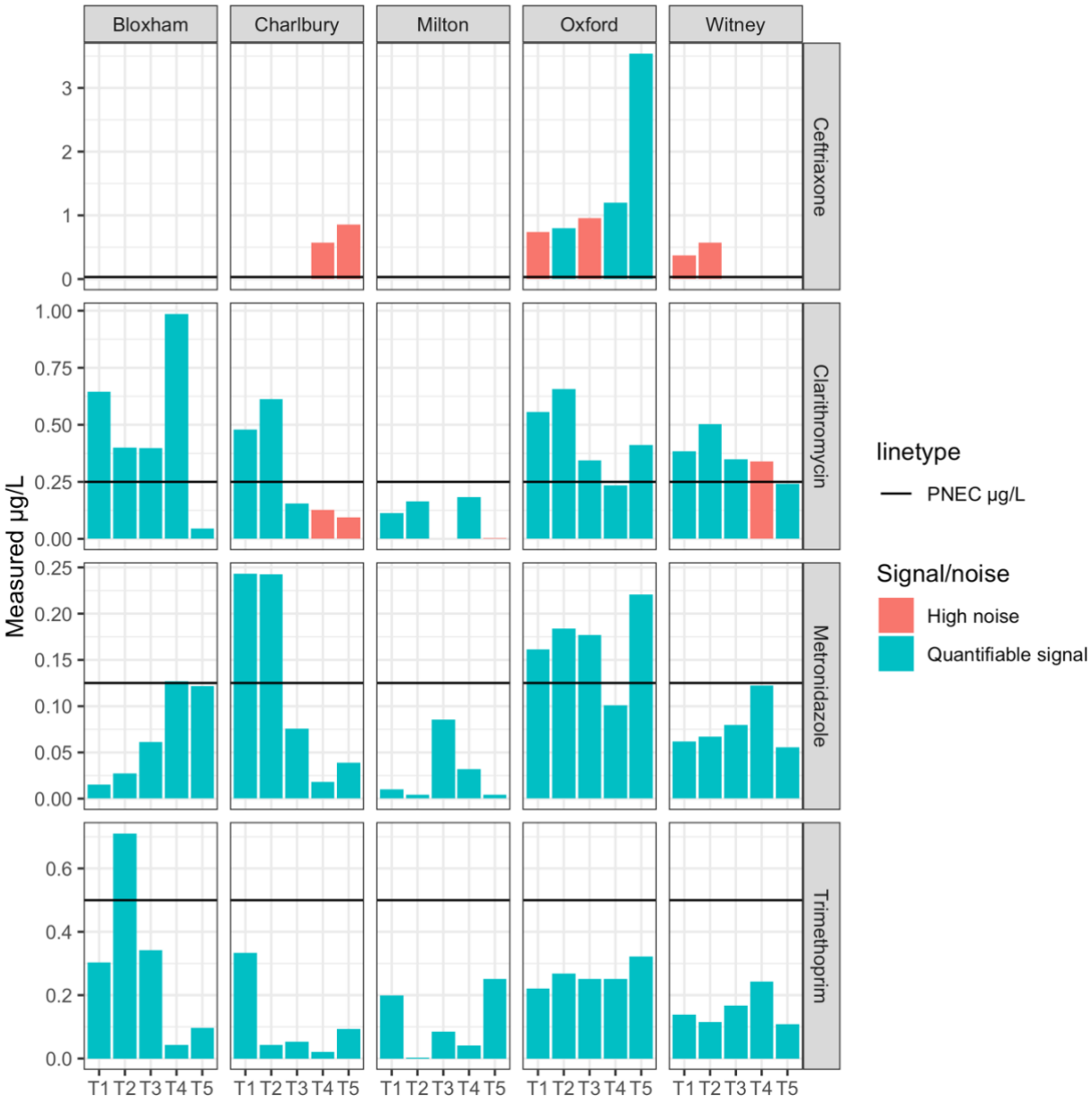
Facets are for each antibiotic with a horizontal line to reflect predicted no-effect concentration for that antibiotic. Within a facet, each bar denotes a single wastewater sample and is coloured by measurement quality where red indicates possible unreliable quantification (high noise), and blue indicates reliable quantification (low noise).



Clarithromycin residues most commonly exceeded PNEC (n=14; 1/14 high-noise), followed by ceftriaxone (n=9; 6/9 high-noise), metronidazole (n=7) and trimethoprim (n=1) (fig.). PNECs were reached in wastewater samples originating mostly from Oxford (n=13; 2/13 high-noise) and Bloxham (n=6), with fewer examples for Charlbury (n=6; 2/6 high-noise), Witney (n=6; 3/6 high-noise) and Milton-under-Wychwood where no samples reached PNECs for any residue. Consistent with modelling results (Figure 34), timepoint effects appeared minimal.

Figure 34 - Antibiotic residue concentration measurements for four antibiotics exceeding respective PNECs

Facets are for each antibiotic with a horizontal line to reflect predicted no-effect concentration for that antibiotic. Within a facet, each bar denotes a single wastewater sample and is coloured by measurement quality where red indicates possible unreliable quantification (high noise), and blue indicates reliable quantification (low noise).



5.4 Discussion

In this longitudinal survey, I characterised differences in wastewater taxonomy and resistome across five WWTW sites, representing discrete sewersheds and GP catchments, located in Oxfordshire over one year. Additionally, I compared resistome composition to site-specific primary care antibiotic prescribing and wastewater antibiotic residue measurements. Taxonomic and resistome differences were primarily WWTW site-driven but I also observed strong evidence of seasonality, with perhaps counterintuitively, highest resistance levels between June to September. Resistomes varied significantly in both abundance and composition across some sites and timepoints, with significant associations with primary care prescribing for macrolides, tetracyclines and macrolide- and nitroimidazole-associated residue levels only. These findings represent important considerations for future development and utility of WBE-based AMR surveillance.

5.4.1 Significant differences in taxonomy and resistomes across sites and timepoints

Taxonomic profiles were dominated by Proteobacteria across all samples where the most abundant genera included both human- and environmentally-associated bacteria (*Acidovorax*, *Aeromonas*, *Pseudomonas*, *Acinetobacter*, *Faecalibacterium*), in agreement with previous findings that wastewater composition is not an exact reflection of human faecal content but a confluence of wider environmental and non-anthropogenic sources too[45,134,136]. While taxonomic compositional dissimilarities (Bray-Curtis) were not significantly associated with timepoint or site overall, comparison of samples from the same site showed more similarity than between sites, suggestive of site-level differences which are retained across time and potentially attributable to site size[163]. Similarly, more

similarity in taxonomic profiles was seen amongst samples collected at the same timepoint which reflects a potential site-wide seasonality effect whereby wastewater temperatures predictably structure sewer communities[164]. These high-level observations were confirmed by consensus-based differential abundance testing which identified 513 genera undergoing significant abundance changes. With samples from the largest WWTW (Oxford) as the reference, I have shown that wastewater from a single city in a region does not necessarily represent wastewater from other closely co-located communities in the same region, refuting previous suggestions[43]. For example, WWTW in Charlbury (approx. 12 miles from Oxford) and Witney (10 miles) were respectively enriched and depleted for almost all Actinobacteria genera identified as differentially abundant when compared to Oxford. Interestingly, samples from Milton-under-Wychwood (17 miles) also diverged from Oxford to a similar degree as Witney whereas Bloxham (19 miles) showed minimal divergence despite a ~45-fold population difference and the most geographically distinct location. Differential abundance testing results further supported site-wide seasonality as divergence in taxonomic profiles from the reference first winter timepoint (T1) increased until T3 (summer) before decreasing to being almost entirely similar at the final winter timepoint (T5) where only a single genus was identified as differentially abundant.

Both AMR gene total abundance and richness in particular were associated most clearly with timepoint, showing significant stepwise increases from T1-T4 before a significant drop at T5, approximately the same time of year as T1. As with taxonomy[164], seasonality is known to influence wastewater AMR potentially through temperature fluctuations[48]

which is congruent with the highest abundance/richness being during summer (T3-T4) in my samples. This is also consistent with seasonal summertime peaks of *Escherichia coli* bloodstream infections in the UK[165]. However, other underlying factors include dilution from increased winter rain events[166] and seasonal variation of antibiotic prescribing, which is highest in the winter, requiring a potential lag effect from prescribing to downstream impacts on wastewater[167,168]. Strong site-level effects were also observed with highest total AMR gene abundance and richness in Charlbury and Oxford, respectively. This may relate to localised high carriage and shedding of AMR genes in Charlbury whereas for Oxford, a higher diversity of AMR genes may be expected due to the vastly increased population served. While not always significant, samples from Milton-under-Wychwood generally showed reduced AMR gene abundance and richness which may also reflect low primary care prescribing of antibiotics historically and during the sampling period; however Bloxham was not AMR-enriched despite extreme historic prescribing peaks. The most prevalent resistance classes observed in my samples align with previous global wastewater resistome surveys[43] as well as with specific findings of enriched macrolide resistance wastewater samples in Europe[49]. These findings further support hypotheses that dominant constituents of the wastewater resistome are broadly conserved across different WWTWs due to shared factors of the wastewater environment[48,169], but are further modulated by local trends such as macrolide resistance prevalence in Europe. No strong geographic or timepoint clustering of sample resistomes was evident other than for Milton-under-Wychwood which appeared to be driven by lower levels of resistance to macrolides. However, sample resistomes were significantly correlated to taxonomic composition using Procruste analysis in agreement with previous studies[48,125,169], and further highlighting the potential resistome-structuring role of

specific taxa through their association with AMR genes and AMR-associated mobile genetic elements such as plasmids[62,96,170].

5.4.2 Generalised linear modelling

5.4.2.1 Do city-level WWTW wastewater resistomes reflect regional WWTWs?

Timepoint and WWTW were associated with total AMR gene abundance, with significant increases at T3 and T4 as well as significantly lower levels from Milton-under-Wychwood vs Oxford; congruent with seasonality and low Milton-under-Wychwood historic primary care prescribing. Models stratifying by most prevalent AMR classes uncovered significant class-specific associations of varying magnitudes across the WWTW sampled. For example, in comparison to Oxford, Witney had higher class-level abundance of cephalosporin resistance genes (fold-change: 1.11; 95% CI: 1.03-1.20); Charlbury had lower abundance of class-level sulfonamide (0.83; 0.70-0.99) but higher abundance of class-level macrolide resistance genes (1.19; 1.05-1.34); Milton had lower abundance of class-level fluoroquinolone (0.77; 0.62-0.96), sulfonamide (0.75; 0.63-0.89) and macrolide resistance genes (0.84; 0.74-0.95). These findings suggest that while wastewater from a single large WWTW can reflect total wastewater AMR prevalence from multiple regional WWTWs (all except Milton-under-Wychwood in my samples), class-level AMR prevalence can differ significantly at specific regional WWTWs and potentially represents distinct localised population-level AMR prevalence, or site-specific effects that are acting at the treatment works level.

5.4.2.2 Does higher primary care prescribing associate to higher wastewater AMR prevalence?

I found no evidence that total AMR gene abundance across samples was associated with primary care prescribing of all antibacterial drugs or residue concentrations of all drug targets measured. This suggests that indirect selection (i.e. co- and cross-selection) is not a primary driver of regional wastewater resistance, and supports similar findings at the country[43] and hospital[125] level. For models assessing direct selection (e.g. association between prescribing of a specific antibiotic class and abundance of AMR genes conferring resistance to that class in wastewater), only macrolide resistance was identified as significantly increasing with macrolide prescribing and daily macrolide residue loads. This suggests that primary care prescribing of macrolides and ensuing macrolides residues are potential key drivers of wastewater macrolide resistance in this study region whereas primary care prescribing of tetracyclines, cephalosporins, sulfonamides, fluoroquinolones and nitroimidazoles plays a smaller role in modulating respective AMR gene abundances in wastewater. The association with macrolide resistance is intriguing; notably, single macrolide dosing for trachoma control has been associated with long-term increases in gut microbiome resistance in multiple studies[171], confirmed in a very large recent multi-country randomised controlled trial (MORDOR), suggesting that resistance may be particularly persistent following administration of drugs from this class[172]. Rather than being associated with primary care prescribing, abundance of these five AMR classes could instead be associated with secondary prescribing in hospitals or outpatient facilities[125], but analyses of these prescribing data was beyond the scope of this study. Furthermore, overall abundance of AMR genes conferring resistance to these classes was lower than for macrolides and may have reduced discriminatory power in the already

limited sample size for this set of models following omission of Oxford samples which lacked clear primary care prescribing data.

5.4.2.3 Wastewater antibiotic residues and selection of AMR

Mean antibiotic residue concentrations of macrolides, tetracyclines, sulfonamides, nitroimidazoles and cephalosporins (outcomes) were significantly associated with the WWTW sampled (exposure), with an additional significant timepoint (exposure) effect for macrolide and tetracycline residues only. These findings suggest that the potential for AMR selection in wastewater varies significantly across WWTWs on this geographic scale, but that seasonality played a lesser role. This is interesting since seasonality of prescribing is well-documented and seasonality of AMR abundance was observed in this study. This may suggest that antibiotic concentrations in wastewater do not significantly drive the selection of AMR and instead, selection occurs in the human gut with a lag period from prescribing. Consistent with this, no significant associations were found between class-level antibiotic residue concentration and contemporaneous primary care prescribing of that antibiotic even when fit adjusting only for timepoint and WWTW. A potential explanation may be that during conveyance to the wastewater sampling point, specific residues may experience varied processes of biodegradation and sorption such as accumulation[173] which would be expected to attenuate temporal flux of antibiotics to WWTWs, thus partially buffering concentrations against changes in prescribing. Dilution associated with wet weather of specific timepoints likely also plays a factor which would be interesting to explore with meteorological data. I have also only included residue concentrations for antibiotics in their original forms in the models investigating the

relationship between residues and prescribing which may not be fully representative of prescribing items due to metabolism and extraction at varied rates[174]. Further investigation of antibiotic metabolite residue measurements may identify prescribing-residue associations, but these were omitted for direct comparability to predicted no-effect concentrations for assessment of AMR selection potential.

Concentrations of clarithromycin residues reached or exceeded predicted no-effect concentrations (PNECs)[153] in more than half of all samples, and were primarily observed in samples from Bloxham, Oxford and Witney. This is congruent with the high proportion of AMR genes conferring resistance to macrolides in these sites and suggests that class-level AMR selection may be a driving factor in addition to population-level prevalence. This is also supported by the fact that no samples from Milton-under-Wychwood exceeded clarithromycin or any other tested PNEC, and were notably reduced in total AMR abundance and relative proportion of macrolide ARGs.

5.4.3 Limitations

While metagenomic sequencing represents a key strength of my study by facilitating the quantitative identification of all-known taxonomic and AMR targets in an agnostic manner, sensitivity is a concern when considering highly diverse samples such as wastewater[90]. To guard against sensitivity issues, I sequenced to an optimised depth as previously determined by ultradeep sequencing of wastewater from sites in my study area[90]. However, normalisation and differential abundance testing of metagenomic data represents a controversial field in of itself due to previous reports of heteroscedasticity,

loss of statistical power and variations in the performance of different testing approaches. To mitigate these issues, I utilised optimised normalisation and scaling methods as well as consensus-based differential abundance testing to identify conservative but true sample-wise differences. However, this conservative approach may mean missing additional significant differences between samples, but I contend this is an acceptable trade-off given the risk of false positives otherwise.

My statistical models assessing associations with primary care prescribing had to omit samples from Oxford due to uncertainty in WWTW and practice catchments, reducing statistical power to detect significant associations. However, Oxford samples were assessed in all other models as the comparator, and Witney was used in lieu as the second largest site for these models. Nevertheless, despite my study being large compared to the literature, with five sites sampled at five timepoints, power to detect associations is relatively limited. In particular, models treating site as a random effect failed to converge in the vast majority of cases, meaning that I had to use fixed effects for site and timepoint to adjust for the repeated measurements within my dataset.

5.4.4 Conclusion

Wastewater influent from WWTWs in the same region varies significantly in both taxonomic and resistome composition which may affect wastewater-based epidemiology (WBE) of region-level AMR prevalence based on the target WWTW. I found that wastewater from a single large city site was only partially representative of wastewater from towns and villages in the same region, and that regional AMR abundance was in

part driven by site-specific primary care prescribing, at least for some antibiotic classes. Therefore, reliance on a single large WWTW for regional WBE may lack resolution to monitor localised trends which are otherwise captured by local WWTW sampling. This also highlights the potential for fine-scale applications of WBE. An interesting and highly relevant perspective for future work is to consider whether the wastewater samples are more clearly associated with total and class-specific AMR prevalence in clinical isolates obtained from populations in the respective sewersheds for each WWTW rather than prescribing or wastewater antibiotic residues; this has not been explored as part of this thesis. Additionally, I found significant seasonality of both total and specific AMR gene abundance which should be considered when comparing wastewater from different timepoints, climates or hemispheres to avoid erroneously attributing differences solely to population-level differences. Further studies investigating the scale at which wastewater-based epidemiology of AMR remains effective are warranted to determine the utility of this promising surveillance approach.

6. Conclusions and future work

In this thesis I have used metagenomic sequencing methods to investigate several fundamental questions of methodology and study design regarding the application of wastewater-based epidemiology (WBE) to population-level AMR surveillance. My systematic review (Chapter 2) represents the first meta-analysis of evidence supporting WBE for AMR surveillance as well as a synthesis of shortcomings and knowledge gaps in the existing literature requiring research for the development of WBE into a robust AMR surveillance tool. By conducting this thorough appraisal, I was able to tailor my remaining experimental chapters to specifically address gaps in knowledge whilst employing best practices gleaned from the review.

Unlike wastewater-based monitoring of targets such as specific chemicals, poliovirus and SARS-CoV-2, the wastewater environment may represent a significant reservoir and source of AMR itself. This complex ecological dynamic is a potential reason why there is some scepticism of the wider applicability of WBE to AMR surveillance, and concern that the wastewater AMR signal may not be as readily associated with human AMR prevalence[38]. For instance, detection of SARS-CoV-2 in wastewater is more-or-less unambiguously indicative of excretion by the wastewater-generating population since the virus is not known to replicate in the wastewater environment and is relatively rapidly degraded[175]. Whilst analysis of archival wastewater samples collected before the SARS-CoV-2 pandemic detected the virus weeks before reported clinical cases (Illustrative of the potential WBE has for early warning), no signal was detectable in earlier samples

taken several months prior to the observed clinical emergence of SARS-CoV-2 in humans[176]. AMR, however, is a naturally ubiquitous phenomenon and is detectable regardless of human AMR infections – well illustrated by the isolation of resistant bacteria from pristine environments (e.g. Antarctic glacial ice cores) devoid of human presence or activity[177]. Therefore, the favourable microbial growth conditions of sewerage systems (nutrient-rich, moist, warm) and local selection pressures likely promote a baseline level of AMR selection and emergence regardless of human input as supported by studies identifying natively present AMR bacteria from storm sewers which do not cater to human sanitary waste[178]. As such, the sewer environment harbours a microbial community and resistome which is potentially distinct from human excreta (and/or may "adopt" and propagate resistance genes from human isolates[179]) and this may confound the ability to detect the AMR signal generated by the target population. This is a key concern for WBE-based AMR surveillance as a general approach since interpretations made on the basis of a wastewater sample may not decisively reflect the population but instead the sewer community. However, my results demonstrating high concordance between wastewater and human AMR prevalence across highly heterogenous study settings and methods acts as a strong proof of principle for WBE-based AMR surveillance. Whilst previous studies had identified this link or relied upon it implicitly, I was able to validate and quantify the wastewater-human AMR prevalence relationship and identify potential optimising study features. Notably, my findings in Chapter 2 suggesting that WBE performance is partly species- and AMR mechanism-dependent were validated by experimental work showing hospital wastewater AMR gene abundances only reflected clinical resistance in Enterococcaceae or Staphylococcaceae, but not Enterobacteriaceae isolates[125].

In addition to specific species or AMR determinants, performance is likely also affected by specific sewershed features such as presence of unique or enriched AMR point-sources as discussed in Chapter 2. The reporting of such features was generally poor across studies which precluded a synthesis in my review, but this underpinned my decision to purposively sample WWTW sites which would potentially produce such distinct AMR signals (e.g. abattoirs, nursing homes) in my experimental study. Whilst I identified significant differences in taxonomic and resistome composition between WWTWs which may be associated with differences in sewershed features, additional work to specifically investigate the potential impact of features such as care homes, agriculture and an abattoir has not yet been completed and is discussed in the following section covering future work on the Chapter 5 dataset. The observations of significantly different levels and types of antibiotic residues across sites was particularly interesting given the lack of association with current primary care prescribing. The reasons for these differences are far from clear but may reflect unassessed variables such as secondary care prescribing and movement of populations which experienced a sudden change partway through sampling due to the national SARS-CoV-2-associated lockdown. It may also be the case that further consideration of antibiotic metabolites and alternate aggregation of prescribing data with normalisation strategies better suited to modelling this relationship are needed. Additionally, it may be that the relationship is not contemporaneous and instead dependent on a lag period which I have not accounted for. However, a comparison of results from Chapter 5 with global[43] and European wastewater metagenome datasets[50] supports the theory that a core wastewater-associated resistome may be a function of local effects on the wastewater microbiome sustained by

features of the sewerage network which are then modulated by AMR gene contributions (in the form of waste) from the local population within the sewershed. To investigate this further, future studies exploring the use of sanitary sewers instead of combined sewer systems would be of great interest, and I would hypothesise that samples from sanitary sewers would be more representative of population AMR and possibly more sensitive since noise from surface/storm-water run-off (i.e. environmental contributions) would be reduced. I was not able to explore this due to the Victorian era combined sewerage system in place across most of the UK (and Oxfordshire), however, modern separated systems have been installed since the 1960s and represent an interesting study target to investigate this hypothesis.

Metagenomic sequencing provides a potential solution to the limitations of culture-based methods regarding representation of extremely diverse environmental samples such as wastewater, and has been used to good effect in several AMR-focussed WBE studies to date[43,49,57,88]. However, challenges remain with this technology, including high resource cost, sensitivity concerns and the complexities of bioinformatic deconvolution. My assessment of metagenomic methods for WBE-based AMR surveillance (Chapter 3) highlights the variability that can be introduced by processing methods, and makes suggestions for optimised strategies involving shotgun sequencing and an alternative complementary targeted profiling approach to detect AMR genes, evaluating these on both a well-defined mock and real-world wastewater samples. I demonstrated that an approach combining amplicon methods targeting 16S rRNA genes and an AMR gene panel is a feasible alternative to deep shotgun metagenomics as well as potentially being

more sensitive and cost-effective (dependent on sequencing depth). This is important as these methods represent a more readily accessible option since library preparation and sequencing can be conducted in-house with relatively inexpensive sequencers (e.g. MiniSeq) as opposed to deep shotgun sequencing (e.g. NovaSeq). To this end I also developed and released an open-source pipeline alternative for processing AmpliSeq sequence data designed to be more accessible and as accurate as currently available free or paid options respectively. However, additional work is required to validate this pipeline fully as thresholds for mapping were optimised to the same test dataset.

As identified in Chapter 2, most previous wastewater AMR studies have relied on grab sampling for the collection of wastewater - likely due to the convenience and significantly lower workload/cost compared to more comprehensive methods. However, the taxonomic composition of wastewater had been shown to significantly fluctuate on a short ~4H timescale[55]. Naturally, this can be expected to have implications for WBE analyses based on single timepoint grab samples but prior to the work carried out in this thesis, this was incompletely demonstrated. These findings drove my assessment of short-term compositional changes in wastewater (Chapter 4) providing insights into fluctuations at much higher temporal resolution than previously, using metagenomic profiling of total genomic content and resistome composition. By demonstrating that daily taxonomic composition undergoes a repeating diurnal pattern across the three days evaluated, I confirm and expand previous suggestions that samples from a single day are representative of a single adjacent day[39]. However, the temporal extent to which a sample remains representative requires additional investigation since it appears highly

dependent on the specific site and the AMR target of interest based on two-monthly observations in Chapter 5.

I found an unexpected result in Chapter 4 where recurring taxonomic abundance peaks at ~1am were observed across all three days in stark contrast to human waking periods thought to primarily drive flow. Samples from this period counterintuitively clustered with samples taken during morning peak use periods ~8am as opposed to forming distinct time period clusters as was seen for the evening and early morning samples. This interesting result is heavily illustrative of the often-overlooked role that WWTW and sewerage infrastructure plays in driving wastewater composition. Since I conducted the fieldwork and WWTW sampling myself, I was able to query site managers and process engineers directly to learn that during periods of high flow (e.g. 8am) exceeding the capacity of the WWTW, excess wastewater is mechanically delayed through various systems for later processing during periods of low flow (e.g. 1am). In the case of this WWTW, six pumping stations and nine rising mains divert excess flow to a balancing tank instead of the WWTW when capacity is reached. This stored wastewater then only enters the WWTW at designated low flow periods – explaining both the human behaviour-unrelated abundance peak and clustering of samples since the wastewater is essentially from the same time period. This complex interaction with the sewerage system demonstrates how interdisciplinary collaboration in this field is absolutely critical. Another example is during the wastewater sampling process itself where the exact location wastewater is collected from can potentially affect results depending on the WWTW site layout. Here, I am specifically referring not to pre- or post-treatment but

rather to the multiple potential sampling locations along the crude inlet system (i.e. sampling raw influent). By consulting with process technicians specific to each WWTW, I was able to position the autosampler to pump wastewater from areas of the crude inlet devoid of recirculating effluent (used to clean screening machinery) and input from storage tanks which are utilised during maintenance and store septic wastewater unlikely to represent the "fresh" flow from the community. Without this useful input, collected wastewater and analyses may have been significantly confounded by unknown variables. Having learnt about these systems and their apparent importance, it was interesting to note that very few studies have explored these factors and their impact regarding wastewater sampling for AMR or disease profiling. However, gaining access to WWTWs can be difficult and I was fortunate to meet staff enthusiastic about the surveillance approach who were willing to dedicate their time to consult and offer their expertise on methods. Future studies should carefully consider the expertise of collaborating WWTW staff where possible so as to optimise the sampling design and benefit the interpretation of results.

I directly applied findings from Chapters 2-4 to my longitudinal survey of Oxfordshire WWTWs (Chapter 5) by using validated 24H composite sampling and metagenomic sequencing/bioinformatic analyses, purposively sampling WWTWs with potential unique signals such as high/low historic prescribing or the presence of abattoirs, and addressing the lack of fine-scale longitudinal studies with extensive contemporaneous clinical and antibiotic residue data. This dataset extends beyond work conducted during this thesis but in my initial analyses, I was able to refute findings from previous studies suggesting

that sampling a single large WWTW is representative of smaller sites in the same region. However, these other studies focused on a much larger, global scale, investigating differences in AMR prevalence between countries which may be significantly more pronounced (e.g. especially when considering low- and middle-income countries (LMICs)), but this does pose the interesting question of scale with respect to WBE AMR surveillance. Most previous studies have focussed on the global or transnational scale, but application on a finer-scale may prove useful for local intervention as it has for WBE-based chemical analyses, SARS-CoV-2 and poliovirus. To this end, I identified significant differences between WWTWs within the same region which may reflect localised town- or village-level AMR prevalence, with further evidence suggesting local primary care prescribing of macrolides may act as a direct driver of macrolide resistance. However, this requires confirmation by comparison of wastewater resistomes to clinical AMR prevalence data I collected as part of the longitudinal dataset. I had originally planned to conduct these analyses as part of this thesis but the Covid-19 pandemic and resulting UK lockdown caused significant delays especially regarding sampling of WWTWs and metagenomic sequencing. These analyses will therefore now be conducted as future work.

6.1 Future work

In addition to the wastewater metagenomes, antibiotic residue concentrations and primary care prescribing data discussed in this thesis, I have also collected and sequenced contemporaneous primary care urinary tract infection (UTI) and bacteraemia isolates originating from each site which will reveal a snapshot of circulating sequence types and AMR determinants in the co-located human populations. Additionally, the Infections in

Oxfordshire Research Database (https://oxfordbrc.nihr.ac.uk/research-themes-overview/antimicrobial-resistance- and-modernising-microbiology/infections-in-oxfordshire-research-database-iord/) provides a rich spatiotemporally matched clinical dataset including all hospital admissions, microbiology test results and secondary care prescribing within Oxfordshire. With this extended dataset, I plan to first assess if spatiotemporally matched metagenomes can be used to reconstitute genomes from the contemporaneous primary care infections by mapping metagenomic reads to assembled infection isolate genomes as previously validated by Gweon et al[90]. Despite a collection of 567 genomes (470 UTI, 97 bacteraemia), fully reconstituting an isolate genome is unlikely due to the sheer diversity present in wastewater and in that instance, I plan to utilise the approach of k-mer containment via Mash Screen which enables the identification of partial reconstitution in the form of shared metagenome-isolate k-mer hashes. The clinical data extract will be used as the outcome in a Bayesian generalized linear model with a taxonomy-adjusted AMR metric derived from the wastewater metagenomes acting as the exposure. This approach is rooted in a previous study validating and optimising such a model for predicting AMR in infections caused by Enterobacterales using pooled population-level faecal metagenomic profiling[155]. By quantifying the accuracy of each site in reflecting local AMR prevalence, I plan to demonstrate if site-level features contribute "AMR noise" as well as then specifically investigate potential feature-associated AMR determinants (e.g. animal AMR from sites with abattoir or agriculture). Whilst the SARS-CoV-2 pandemic presented a significant challenge to completing this thesis, it has also imparted unique qualities to the samples collected and datasets generated throughout. In particular, the national lockdown took place in the midst of longitudinal sampling for Chapter 5 which presents an unexpected opportunity to

investigate the impact of significant human behavioural changes on wastewater AMR composition. The effects of the sudden change in population movement, hospital admissions, and public gatherings will be challenging but interesting to unpick and investigate. For example, I anticipate the strong seasonality effect observed in Chapter 5 is not purely climate-driven but also a function of the lockdown measures since it does not fully align with previous work, and the observed highest wastewater AMR abundance coincided with easing of lockdown measures.

6.2 Concluding remarks

In conclusion, in this thesis I have identified and contributed to addressing several fundamental questions regarding the use of wastewater-based epidemiology for population-level AMR surveillance, methods for sampling and processing data, as well as generating a large comprehensive dataset poised to further evaluate the promise of this surveillance approach. A common theme arising from this work is the importance of validating individual methods to systematically identify sources of variation when developing workflows such as WBE. Whilst the approach has been thrust into the spotlight and many studies are now being published, the overall lack of detailed reporting regarding study methods and site features hinders meaningful synthesis of results across studies. Iterative development of optimal practices is contingent on clearer reporting of high-quality metadata. Additionally, the need for interdisciplinary collaboration cannot be overstated, especially for this field where much can be gained from colleagues in areas of chemical and viral surveillance as well as the often-overlooked expertise of individuals working outside of research but within the wastewater industry such as WWTW

managers, process technicians and engineers. This thesis and body of work has therefore incrementally contributed to the knowledge base and development of this surveillance approach. The complexity involved in the analysis and deconvolution of wastewater for AMR surveillance is significant, but the potential benefits to optimising this strategy to contribute to the mitigation of antimicrobial resistance are also substantial.

Bibliography

1 Fleming, A. *Nobel Lecture*, <https://www.nobelprize.org/prizes/medicine/1945/fleming/lecture/> (1945).

2 Rammelkamp, C. H. & Maxon, T. Resistance of Staphylococcus aureus to the Action of Penicillin. *Proceedings of the Society for Experimental Biology and Medicine* **51**, 386-389, doi:10.3181/00379727-51-13986 (1942).

3 WHO. *Antimicrobial resistance: global report on surveillance.*, <https://apps.who.int/iris/bitstream/handle/10665/112642/?sequence=1> (2014).

4 Holmes, A. H. *et al.* Understanding the mechanisms and drivers of antimicrobial resistance. *Lancet* **387**, 176-187, doi:10.1016/s0140-6736(15)00473-0 (2016).

5 Murray, C. J. L. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **399**, 629-655, doi:10.1016/S0140-6736(21)02724-0 (2022).

6 O'Neill, J. *Tackling Drug-Resistant Infections Globally: Final Report and Recommendations*, <https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf> (2016).

7 WHO. *Global action plan on AMR: Objective 2*, <https://apps.who.int/iris/bitstream/handle/10665/193736/9789241509763_eng.pdf?sequence=1> (2019).

8 Coggon, D., Rose, G. & Barker, D. *Epidemiology for the uninitiated*. (2009).

9 Graunt, J. *Natural and Political Observations Made Upon the Bills of Mortality*. (1676).

10 Snow, J. & Richardson, B. *Snow on cholera*. (New York : The Commonwealth Fund, 1936).

11 Thacker, S. B. & Berkelman, R. L. Public health surveillance in the United States. *Epidemiol Rev* **10**, 164-190, doi:10.1093/oxfordjournals.epirev.a036021 (1988).

12 UKHSA. *Notifiable diseases and causative organisms: how to report*, <https://www.gov.uk/guidance/notifiable-diseases-and-causative-organisms-how-to-report> (2022).

13 Mandl, K. D. *et al.* Implementing syndromic surveillance: a practical guide informed by the early experience. *J Am Med Inform Assoc* **11**, 141-150, doi:10.1197/jamia.M1356 (2004).

14 UKHSA. Syndromic surveillance: systems and analyses. (2022).

15 Ginsberg, J. *et al.* Detecting influenza epidemics using search engine query data. *Nature* **457**, 1012-1014, doi:10.1038/nature07634 (2009).

16 Lazer, D., Kennedy, R., King, G. & Vespignani, A. The Parable of Google Flu: Traps in Big Data Analysis. *Science* **343**, 1203-1205, doi:doi:10.1126/science.1248506 (2014).

17 Lampos, V. *et al.* Tracking COVID-19 using online search. *NPJ Digit Med* **4**, 17, doi:10.1038/s41746-021-00384-w (2021).

18 Aiello, A. E., Renson, A. & Zivich, P. N. Social Media– and Internet-Based Disease Surveillance for Public Health. *Annual Review of Public Health* **41**, 101-118, doi:10.1146/annurev-publhealth-040119-094402 (2020).

19 Hay, S. I. *et al.* Measuring and mapping the global burden of antimicrobial resistance. *BMC Medicine* **16**, 78, doi:10.1186/s12916-018-1073-z (2018).

20 Marineli, F., Tsoucalas, G., Karamanou, M. & Androutsos, G. Mary Mallon (1869-1938) and the history of typhoid fever. *Ann Gastroenterol* **26**, 132-134 (2013).

21 Chisholm, R. H. *et al.* Implications of asymptomatic carriers for infectious disease transmission and control. *R Soc Open Sci* **5**, 172341, doi:10.1098/rsos.172341 (2018).

22 Wertheim, H. F. *et al.* The role of nasal carriage in Staphylococcus aureus infections. *Lancet Infect Dis* **5**, 751-762, doi:10.1016/s1473-3099(05)70295-4 (2005).

23 Riggs, M. M. *et al.* Asymptomatic Carriers Are a Potential Source for Transmission of Epidemic and Nonepidemic Clostridium difficile Strains among Long-Term Care Facility Residents. *Clinical Infectious Diseases* **45**, 992-998, doi:10.1086/521854 (2007).
24 Furuya-Kanamori, L. *et al.* Heterogeneous and Dynamic Prevalence of Asymptomatic Influenza Virus Infections. *Emerg Infect Dis* **22**, 1052-1056, doi:10.3201/eid2206.151080 (2016).
25 Brinkac, L., Voorhies, A., Gomez, A. & Nelson, K. E. The Threat of Antimicrobial Resistance on the Human Microbiome. *Microb Ecol* **74**, 1001-1008, doi:10.1007/s00248-017-0985-z (2017).
26 Sheppard, A. E. *et al.* Nested Russian Doll-Like Genetic Mobility Drives Rapid Dissemination of the Carbapenem Resistance Gene blaKPC. *Antimicrob Agents Chemother* **60**, 3767-3778, doi:10.1128/aac.00464-16 (2016).
27 Hess, A. S. *et al.* Prior colonization is associated with increased risk of antibiotic-resistant Gram-negative bacteremia in cancer patients. *Diagn Microbiol Infect Dis* **79**, 73-76, doi:10.1016/j.diagmicrobio.2014.01.022 (2014).
28 ARHAI. *Implementation of modified admission MRSA screening guidance for NHS*, <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/345144/Implementation_of_modified_admission_MRSA_screening_guidance_for_NHS.pdf> (2014).
29 PHE. NHS population screening explained. (2021).
30 Bhangar, S. *et al.* Chamber bioaerosol study: human emissions of size-resolved fluorescent biological aerosol particles. *Indoor Air* **26**, 193-206, doi:10.1111/ina.12195 (2016).
31 Shahi, F., Redeker, K. & Chong, J. Rethinking antimicrobial stewardship paradigms in the context of the gut microbiome. *JAC Antimicrob Resist* **1**, dlz015, doi:10.1093/jacamr/dlz015 (2019).
32 Asghar, H. *et al.* Environmental surveillance for polioviruses in the global polio eradication initiative. *Journal of Infectious Diseases* **210**, S294-S303, doi:10.1093/infdis/jiu384 (2014).
33 Michael-Kordatou, I., Karaolia, P. & Fatta-Kassinos, D. Sewage analysis as a tool for the COVID-19 pandemic response and management: the urgent need for optimised protocols for SARS-CoV-2 detection and quantification. *J Environ Chem Eng* **8**, 104306-104306, doi:10.1016/j.jece.2020.104306 (2020).
34 Larson, R. C., Berman, O. & Nourinejad, M. Sampling manholes to home in on SARS-CoV-2 infections. *PLoS One* **15**, e0240007, doi:10.1371/journal.pone.0240007 (2020).
35 Schang, C. *et al.* Passive Sampling of SARS-CoV-2 for Wastewater Surveillance. *Environmental Science & Technology* **55**, 10432-10441, doi:10.1021/acs.est.1c01530 (2021).
36 Ahmed, W. *et al.* First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: A proof of concept for the wastewater surveillance of COVID-19 in the community. *Science of the Total Environment*, doi:10.1016/j.scitotenv.2020.138764 (2020).
37 Choi, P. M. *et al.* Wastewater-based epidemiology biomarkers: Past, present and future. *TrAC Trends in Analytical Chemistry* **105**, 453-469, doi:10.1016/j.trac.2018.06.004 (2018).
38 Fahrenfeld, N. & Bisceglia, K. J. Emerging investigators series: sewer surveillance for monitoring antibiotic use and prevalence of antibiotic resistance: urban sewer epidemiology. *Environmental Science: Water Research & Technology* **2**, 788-799, doi:10.1039/c6ew00158k (2016).
39 Aarestrup, F. M. & Woolhouse, M. E. J. Using sewage for surveillance of antimicrobial resistance. *Science* **367**, 630-632, doi:10.1126/science.aba3432 (2020).

40 González-Mariño, I. *et al.* Spatio-temporal assessment of illicit drug use at large scale: evidence from 7 years of international wastewater monitoring. *Addiction* **115**, 109-120, doi:10.1111/add.14767 (2020).
41 Chau, K. K. *et al.* Systematic review of wastewater surveillance of antimicrobial resistance in human populations. *Environ Int* **162**, 107171, doi:10.1016/j.envint.2022.107171 (2022).
42 WHO. *Global Antimicrobial Resistance Surveillance System (GLASS) Report Early implementation 2017-18*, <https://apps.who.int/iris/bitstream/handle/10665/279656/9789241515061-eng.pdf?ua=1> (2018).
43 Hendriksen, R. S. *et al.* Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nature communications* **10**, 1124-1124, doi:10.1038/s41467-019-08853-3 (2019).
44 Tacconelli, E. *et al.* Surveillance for control of antimicrobial resistance. *Lancet Infect Dis* **18**, e99-e106, doi:10.1016/s1473-3099(17)30485-1 (2018).
45 Newton, R. J. *et al.* Sewage reflects the microbiomes of human populations. *mBio* **6**, 1-9, doi:10.1128/mBio.02574-14 (2015).
46 Fernández, M. D. B. *et al.* Environmental surveillance of norovirus in Argentina revealed distinct viral diversity patterns, seasonality and spatio-temporal diffusion processes. *Science of the Total Environment* **437**, 262-269, doi:10.1016/j.scitotenv.2012.08.033 (2012).
47 Wright, G. D. The antibiotic resistome: The nexus of chemical and genetic diversity. *Nature Reviews Microbiology*, doi:10.1038/nrmicro1614 (2007).
48 Su, J. Q. *et al.* Metagenomics of urban sewage identifies an extensively shared antibiotic resistome in China. *Microbiome* **5**, 84, doi:10.1186/s40168-017-0298-y (2017).
49 Karkman, A., Berglund, F., Flach, C.-F., Kristiansson, E. & Larsson, D. G. J. Predicting clinical resistance prevalence using sewage metagenomic data. *Communications biology* **3**, 711-711, doi:10.1038/s42003-020-01439-6 (2020).
50 Pärnänen, K. M. M. *et al.* Antibiotic resistance in European wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Science Advances* **5**, doi:10.1126/sciadv.aau9124 (2019).
51 Reinthaler, F. F. *et al.* Resistance patterns of Escherichia coli isolated from sewage sludge in comparison with those isolated from human patients in 2000 and 2009. *Journal of Water and Health* **11**, 13-20, doi:10.2166/wh.2012.207 (2013).
52 Tong, J. *et al.* Microbial community evolution and fate of antibiotic resistance genes along six different full-scale municipal wastewater treatment processes. *Bioresource Technology*, doi:10.1016/j.biortech.2018.10.079 (2019).
53 Zhang, L., Cheng, Y., Qian, C. & Lu, W. Bacterial community evolution along full-scale municipal wastewater treatment processes. *Journal of Water and Health* **18**, 665-680, doi:10.2166/wh.2020.092 (2020).
54 Jakobsen, L. *et al.* Characterisation, dissemination and persistence of gentamicin resistant Escherichia coli from a Danish university hospital to the waste water environment. *Environment International* **34**, 108-115, doi:10.1016/j.envint.2007.07.011 (2008).
55 Guo, B., Liu, C., Gibson, C. & Frigon, D. Wastewater microbial community structure and functional traits change over short timescales. *Science of the Total Environment*, doi:10.1016/j.scitotenv.2019.01.207 (2019).
56 Shchegolkova, N. M. *et al.* Microbial Community Structure of Activated Sludge in Treatment Plants with Different Wastewater Compositions. *Front Microbiol* **7**, 90, doi:10.3389/fmicb.2016.00090 (2016).
57 Pehrsson, E. C. *et al.* Interconnected microbiomes and resistomes in low-income human habitats. *Nature* **533**, 212-216, doi:10.1038/nature17672 (2016).

58 Shanks Orin, C. *et al.* Comparison of the Microbial Community Structures of Untreated Wastewaters from Different Geographic Locales. *Applied and Environmental Microbiology* **79**, 2906-2913, doi:10.1128/AEM.03448-12 (2013).
59 Honda, R. *et al.* Estimated discharge of antibiotic-resistant bacteria from combined sewer overflows of urban sewage system. *npj Clean Water* **3**, 15, doi:10.1038/s41545-020-0059-5 (2020).
60 Poulsen, C. S., Kaas, R. S., Aarestrup, F. M. & Pamp, S. J. Standard Sample Storage Conditions Have an Impact on Inferred Microbiome Composition and Antimicrobial Resistance Patterns. *Microbiol Spectr* **9**, e0138721, doi:10.1128/Spectrum.01387-21 (2021).
61 Knudsen, B. E. *et al.* Impact of Sample Type and DNA Isolation Procedure on Genomic Inference of Microbiome Composition. *mSystems* **1**, doi:10.1128/mSystems.00095-16 (2016).
62 Shaw, L. P. *et al.* Niche and local geography shape the pangenome of wastewater- and livestock-associated Enterobacteriaceae. *Sci Adv* **7**, doi:10.1126/sciadv.abe3868 (2021).
63 Davies, T. J. *Whole Genome Sequencing and Resistance in Escherichia Coli*, University of Oxford, (2019).
64 Davies, T. J. *et al.* Reconciling the Potentially Irreconcilable? Genotypic and Phenotypic Amoxicillin-Clavulanate Resistance in Escherichia coli. *Antimicrob Agents Chemother* **64**, doi:10.1128/aac.02026-19 (2020).
65 Lal Gupta, C., Kumar Tiwari, R. & Cytryn, E. Platforms for elucidating antibiotic resistance in single genomes and complex metagenomes. *Environment International* **138**, 105667, doi:10.1016/j.envint.2020.105667 (2020).
66 AGISAR. Critically Important Antimicrobials for Human Medicine 6th Re.vision 2018. Ranking of medically important antimicrobials for risk management of antimicrobial resistance due to non-human use. Report No. 9789241515528, (2018).
67 Higgins, J. P. T. *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ (Online)*, doi:10.1136/bmj.d5928 (2011).
68 Woodruff, T. J. & Sutton, P. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health Aff (Millwood)* **30**, 931-937, doi:10.1377/hlthaff.2010.1219 (2011).
69 Morgan, R. L. *et al.* GRADE: Assessing the quality of evidence in environmental and occupational health. *Environment International* **92-93**, 611-616, doi:10.1016/j.envint.2016.01.004 (2016).
70 Haghi, F., Shirmohammadlou, N., Bagheri, R., Jamali, S. & Zeighami, H. High frequency of vancomycin-resistant enterococci in sewage and fecal samples of healthy carriers. *Open Biotechnology Journal* **13**, 1-5, doi:10.2174/1874070701913010001 (2019).
71 Moradigaravand, D. *et al.* Genomic survey of Clostridium difficile reservoirs in the East of England implicates environmental contamination of wastewater treatment plants by clinical lineages. *Microbial genomics* **4**, doi:10.1099/mgen.0.000162 (2018).
72 Pignato, S. *et al.* Molecular epidemiology of ampicillin resistance in Salmonella spp. and Escherichia coli from wastewater and clinical specimens. *Foodborne Pathogens and Disease* **7**, 945-951, doi:10.1089/fpd.2009.0504 (2010).
73 YoungKeun, K. *et al.* Surveillance of antimicrobial resistance among Escherichia coli in wastewater in Stockholm during 1 year: does it reflect the resistance trends in the society? *International Journal of Antimicrobial Agents* **45**, 25-32, doi:10.1016/j.ijantimicag.2014.09.016 (2015).
74 Jørgensen, S. B. *et al.* A comparison of extended spectrum beta-lactamase producing Escherichia coli from clinical, recreational water and wastewater samples associated in time and location. *PLoS ONE* **12**, e0186576-e0186576, doi:10.1371/journal.pone.0186576 (2017).

75 Urase, T., Okazaki, M. & Tsutsui, H. Prevalence of ESBL-producing Escherichia coli and carbapenem-resistant Enterobacteriaceae in treated wastewater: a comparison with nosocomial infection surveillance. *Journal of Water and Health* **18**, 899-910, doi:10.2166/wh.2020.014 (2020).
76 Colomer-Lluch, M. *et al.* Antibiotic resistance genes in bacterial and bacteriophage fractions of Tunisian and Spanish wastewaters as markers to compare the antibiotic resistance patterns in each population. *Environment International* **73**, 167-175, doi:10.1016/j.envint.2014.07.003 (2014).
77 Meir-Gruber, L. *et al.* Population Screening Using Sewage Reveals Pan-Resistant Bacteria in Hospital and Community Samples. *PloS one* **11**, e0164873-e0164873, doi:10.1371/journal.pone.0164873 (2016).
78 Rahimi, F. & Bouzari, M. Biochemical fingerprinting of methicillin-resistant staphylococcus aureus isolated from sewage and hospital in Iran. *Jundishapur Journal of Microbiology* **8**, 1-5, doi:10.5812/jjm.19760v2 (2015).
79 Mourkas, E. *et al.* Gene pool transmission of multidrug resistance among Campylobacter from livestock, sewage and human disease. *Environmental microbiology* **21**, 4597-4613, doi:10.1111/1462-2920.14760 (2019).
80 Ojer-Usoz, E., González, D. & Vitas, A. I. Clonal diversity of ESBL-producing Escherichia coli isolated from environmental, human and food samples. *International Journal of Environmental Research and Public Health* **14**, 676-676, doi:10.3390/ijerph14070676 (2017).
81 Saifi, M., Pourshafie, M. R., Dallal, M. M. S. & Katouli, M. Clonal groups of high-level gentamicin-resistant Enterococcus faecium isolated from municipal wastewater and clinical samples in Tehran, Iran. *Letters in applied microbiology* **49**, 160-165, doi:10.1111/j.1472-765X.2009.02559.x (2009).
82 Hutinel, M. *et al.* Population-level surveillance of antibiotic resistance in Escherichia coli through sewage analysis. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* **24**, 1800497-1800497, doi:10.2807/1560-7917.ES.2019.24.37.1800497 (2019).
83 Huijbers, P. M. C., Larsson, D. G. J. & Flach, C. F. Surveillance of antibiotic resistant Escherichia coli in human populations through urban wastewater in ten European countries. *Environmental Pollution* **261**, 114200-114200, doi:10.1016/j.envpol.2020.114200 (2020).
84 Zaheer, R. *et al.* Surveillance of Enterococcus spp. reveals distinct species and antimicrobial resistance diversity across a One-Health continuum. *Scientific Reports* **10**, 3937-3937, doi:10.1038/s41598-020-61002-5 (2020).
85 Zarfel, G. *et al.* Comparison of extended-spectrum-β-lactamase (ESBL) carrying Escherichia coli from sewage sludge and human urinary tract infection. *Environ Pollut* **173**, 192-199, doi:10.1016/j.envpol.2012.09.019 (2013).
86 Raven, K. E. *et al.* Genomic surveillance of Escherichia coli in municipal wastewater treatment plants as an indicator of clinically relevant pathogens and their resistance genes. *Microbial genomics* **5**, doi:10.1099/mgen.0.000267 (2019).
87 Adator, E. H. *et al.* Whole Genome Sequencing Differentiates Presumptive Extended Spectrum Beta-Lactamase Producing Escherichia coli along Segments of the One Health Continuum. *Microorganisms* **8**, doi:10.3390/microorganisms8030448 (2020).
88 Hendriksen, R. S. *et al.* Pathogen surveillance in the informal settlement, Kibera, Kenya, using a metagenomics approach. *PLoS ONE* **14**, e0222531-e0222531, doi:10.1371/journal.pone.0222531 (2019).
89 Gouliouris, T. *et al.* Detection of vancomycin-resistant Enterococcus faecium hospital-adapted lineages in municipal wastewater treatment plants indicates widespread

distribution and release into the environment. *Genome research* **29**, 626-634, doi:10.1101/gr.232629.117 (2019).
90 Gweon, H. S. *et al.* The impact of sequencing depth on the inferred taxonomic composition and AMR gene content of metagenomic samples. *Environmental Microbiomes*, doi:10.1186/s40793-019-0347-1 (2019).
91 Chau, K. *et al.* Evaluation of metagenomic, 16S rRNA gene and ultra-plexed PCR-based sequencing approaches for profiling antimicrobial resistance gene and bacterial taxonomic composition of polymicrobial samples. *bioRxiv*, 2022.2005.2012.491637, doi:10.1101/2022.05.12.491637 (2022).
92 Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J. & Segata, N. Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol* **35**, 833-844, doi:10.1038/nbt.3935 (2017).
93 O'Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **44**, D733-745, doi:10.1093/nar/gkv1189 (2016).
94 Alcock, B. P. *et al.* CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res* **48**, D517-d525, doi:10.1093/nar/gkz935 (2020).
95 Fan, Y. & Pedersen, O. Gut microbiota in human metabolic health and disease. *Nat Rev Microbiol* **19**, 55-71, doi:10.1038/s41579-020-0433-9 (2021).
96 Matlock, W. *et al.* Genomic network analysis of environmental and livestock F-type plasmid populations. *ISME J* **15**, 2322-2335, doi:10.1038/s41396-021-00926-w (2021).
97 AbuOun, M. *et al.* A genomic epidemiological study shows that prevalence of antimicrobial resistance in Enterobacterales is associated with the livestock host, as well as antimicrobial usage. *Microb Genom* **7**, doi:10.1099/mgen.0.000630 (2021).
98 Gilbert, J. A., Jansson, J. K. & Knight, R. The Earth Microbiome project: successes and aspirations. *BMC Biol* **12**, 69, doi:10.1186/s12915-014-0069-1 (2014).
99 Emery, D. C. *et al.* 16S rRNA Next Generation Sequencing Analysis Shows Bacteria in Alzheimer's Post-Mortem Brain. *Front Aging Neurosci* **9**, 195, doi:10.3389/fnagi.2017.00195 (2017).
100 Guernier-Cambert, V., Chamings, A., Collier, F. & Alexandersen, S. A preliminary investigation of major bacterial antibiotic resistance genes present in Australian fecal samples using shotgun metagenomics and targeted sequencing. *bioRxiv*, 2020.2003.2019.999250, doi:10.1101/2020.03.19.999250 (2020).
101 Agrawal, S., Orschler, L., Sinn, J. & Lackner, S. High-throughput profiling of antibiotic resistance genes in wastewater: comparison between a pond system in Namibia and an activated sludge treatment in Germany. *J Water Health* **18**, 867-878, doi:10.2166/wh.2020.018 (2020).
102 Urbaniak, C. *et al.* Detection of antimicrobial resistance genes associated with the International Space Station environmental surfaces. *Sci Rep* **8**, 814, doi:10.1038/s41598-017-18506-4 (2018).
103 Ye, S. H., Siddle, K. J., Park, D. J. & Sabeti, P. C. Benchmarking Metagenomics Tools for Taxonomic Classification. *Cell* **178**, 779-794, doi:10.1016/j.cell.2019.07.010 (2019).
104 de Abreu, V. A. C., Perdigão, J. & Almeida, S. Metagenomic Approaches to Analyze Antimicrobial Resistance: An Overview. *Front Genet* **11**, 575592-575592, doi:10.3389/fgene.2020.575592 (2021).
105 Lipworth, S. *et al.* Ten-year longitudinal molecular epidemiology study of Escherichia coli and Klebsiella species bloodstream infections in Oxfordshire, UK. *Genome Medicine* **13**, 144, doi:10.1186/s13073-021-00947-2 (2021).

106 Seaton, F. M. *et al.* Long-Term Drought and Warming Alter Soil Bacterial and Fungal Communities in an Upland Heathland. *Ecosystems*, doi:10.1007/s10021-021-00715-8 (2021).
107 ENA. *Evaluation of sequencing approaches to profile taxonomy and resistome of polymicrobial samples*, <https://www.ebi.ac.uk/ena/browser/view/PRJEB52722> (2022).
108 GitHub. *Chau_etal_Ampliseek*, <https://github.com/KaibondChau/Chau_etal_Ampliseek> (2022).
109 GitHub. *AmpliSeek*, <https://github.com/KaibondChau/ampliseek> (2022).
110 Lu, J. & Salzberg, S. L. Ultrafast and accurate 16S rRNA microbial community analysis using Kraken 2. *Microbiome* **8**, 124, doi:10.1186/s40168-020-00900-2 (2020).
111 Bray, J. R. & Curtis, J. T. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* **27**, 325-349, doi:10.2307/1942268 (1957).
112 Youden, W. J. Index for rating diagnostic tests. *Cancer* **3**, 32-35, doi:10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3 (1950).
113 Smith, C. *Bioinformatics Details of the AmpliSeq™ Report*, <https://docs.onecodex.com/en/articles/3764465-bioinformatics-details-of-the-ampliseq-report> (2022).
114 Balvočiūtė, M. & Huson, D. H. SILVA, RDP, Greengenes, NCBI and OTT — how do these taxonomies compare? *BMC Genomics* **18**, 114, doi:10.1186/s12864-017-3501-4 (2017).
115 Fouhy, F., Clooney, A. G., Stanton, C., Claesson, M. J. & Cotter, P. D. 16S rRNA gene sequencing of mock microbial populations- impact of DNA extraction method, primer choice and sequencing platform. *BMC Microbiol* **16**, 123, doi:10.1186/s12866-016-0738-z (2016).
116 Brooks, J. P. *et al.* The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol* **15**, 66, doi:10.1186/s12866-015-0351-6 (2015).
117 Illumina. *BaseSpace DNA Amplicon v2.0 Online Help*, <https://support.illumina.com/help/BaseSpace_App_DNA_Amplicon_v2_OLH_1000000041403/Content/Source/Informatics/Apps/AlignerWholeGenomeBWA_swBS_appDNAA.htm> (2018).
118 McCall, C. M. *et al.* False positives in multiplex PCR-based next-generation sequencing have unique signatures. *J Mol Diagn* **16**, 541-549, doi:10.1016/j.jmoldx.2014.06.001 (2014).
119 Stoler, N. & Nekrutenko, A. Sequencing error profiles of Illumina sequencing instruments. *NAR Genomics and Bioinformatics* **3**, doi:10.1093/nargab/lqab019 (2021).
120 Durazzi, F. *et al.* Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. *Scientific Reports* **11**, 3030, doi:10.1038/s41598-021-82726-y (2021).
121 Brumfield, K. D., Huq, A., Colwell, R. R., Olds, J. L. & Leddy, M. B. Microbial resolution of whole genome shotgun and 16S amplicon metagenomic sequencing using publicly available NEON data. *PLoS One* **15**, e0228899, doi:10.1371/journal.pone.0228899 (2020).
122 Eloe-Fadrosh, E. A., Ivanova, N. N., Woyke, T. & Kyrpides, N. C. Metagenomics uncovers gaps in amplicon-based detection of microbial diversity. *Nat Microbiol* **1**, 15032, doi:10.1038/nmicrobiol.2015.32 (2016).
123 Tessler, M. *et al.* Large-scale differences in microbial biodiversity discovery between 16S amplicon and shotgun sequencing. *Sci Rep* **7**, 6589, doi:10.1038/s41598-017-06665-3 (2017).
124 Chau, K. *et al.* High-resolution characterisation of short-term temporal variability in the taxonomic and resistome composition of wastewater influent. *bioRxiv*, 2022.2007.2008.499324, doi:10.1101/2022.07.08.499324 (2022).

125 Perry, M. R. *et al.* Secrets of the Hospital Underbelly: Patterns of Abundance of Antimicrobial Resistance Genes in Hospital Wastewater Vary by Specific Antimicrobial and Bacterial Family. *Front Microbiol* **12**, 703560, doi:10.3389/fmicb.2021.703560 (2021).
126 Bowes, M. J. *et al.* Weekly water quality monitoring data for the River Thames (UK) and its major tributaries (2009–2013): the Thames Initiative research platform. *Earth Syst. Sci. Data* **10**, 1637-1653, doi:10.5194/essd-10-1637-2018 (2018).
127 McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**, e61217, doi:10.1371/journal.pone.0061217 (2013).
128 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
129 McMurdie, P. J. & Holmes, S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLOS Computational Biology* **10**, e1003531, doi:10.1371/journal.pcbi.1003531 (2014).
130 Kassambara, A., and Mundt, F. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. *R package version 1.0.7.*, https://CRAN.R-project.org/package=factoextra (2020).
131 Lê, S., Josse, J. & Husson, F. FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software* **25**, 1 - 18, doi:10.18637/jss.v025.i01 (2008).
132 Oksanen, J., Simpson, G.L., Blanchet, FG., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., Caceres, M., Durand, S., Evangelista, H.B.A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M.O., Lahti, L., McGlinn, D., Ouellette, M., Cunha, E.R., Smith, T., Stier, A., Ter Braak, C.J.F and Weedon, J. vegan: Community Ecology Package. *R package version 2.6-2*, https://CRAN.R-project.org/package=vegan (2022).
133 Ondov, B. D. *et al.* Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biology* **17**, 132, doi:10.1186/s13059-016-0997-x (2016).
134 Rinninella, E. *et al.* What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms* **7**, 14, doi:10.3390/microorganisms7010014 (2019).
135 Polo, D. *et al.* Making waves: Wastewater-based epidemiology for COVID-19 – approaches and challenges for surveillance and prediction. *Water Research* **186**, 116404, doi:10.1016/j.watres.2020.116404 (2020).
136 Numberger, D. *et al.* Characterization of bacterial communities in wastewater with enhanced taxonomic resolution by full-length 16S rRNA sequencing. *Scientific Reports* **9**, 9673, doi:10.1038/s41598-019-46015-z (2019).
137 Azli, B. *et al.* Metagenomics Insights Into the Microbial Diversity and Microbiome Network Analysis on the Heterogeneity of Influent to Effluent Water. *Frontiers in Microbiology* **13**, doi:10.3389/fmicb.2022.779196 (2022).
138 Speirs, L. B. M., Rice, D. T. F., Petrovski, S. & Seviour, R. J. The Phylogeny, Biodiversity, and Ecology of the Chloroflexi in Activated Sludge. *Frontiers in Microbiology* **10**, doi:10.3389/fmicb.2019.02015 (2019).
139 Santo Domingo, J. W. *et al.* Molecular survey of concrete sewer biofilm microbial communities. *Biofouling* **27**, 993-1001, doi:10.1080/08927014.2011.618637 (2011).
140 UKHSA. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR)*, <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1069632/espaur-report-2020-to-2021-16-Nov-FINAL-v2.pdf> (2021).

141 Langille, M. G. I. *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology* **31**, 814-821, doi:10.1038/nbt.2676 (2013).
142 Manor, O. & Borenstein, E. Systematic Characterization and Analysis of the Taxonomic Drivers of Functional Shifts in the Human Microbiome. *Cell Host Microbe* **21**, 254-267, doi:10.1016/j.chom.2016.12.014 (2017).
143 Nelson, B. What poo tells us: wastewater surveillance comes of age amid covid, monkeypox, and polio. *BMJ* **378**, o1869, doi:10.1136/bmj.o1869 (2022).
144 NHS Digital. *GP and GP practice related data*, <https://digital.nhs.uk/services/organisation-data-service/file-downloads/gp-and-gp-practice-related-data> (2022).
145 NHS Digital. General practice data collections. (2022).
146 NHS Digital. Practice Level Prescribing Data. (2022).
147 Care Quality Commission. Locations regulated by CQC. (2022).
148 Office for National Statistics. Open Geography Portal. (2022).
149 Ministry of Housing Communities & Local Government. English Indices of Deprivation 2019 - LSOA Level. (2022).
150 Kahle, D. & Wickham. ggmap: Spatial Visualization with ggplot2. *The R Journal* **5**, 144-161 (2013).
151 Vihta, K. D. *et al.* Trends over time in Escherichia coli bloodstream infections, urinary tract infections, and antibiotic susceptibilities in Oxfordshire, UK, 1998-2016: a study of electronic health records. *Lancet Infect Dis* **18**, 1138-1149, doi:10.1016/s1473-3099(18)30353-0 (2018).
152 Holton E. & Kasprzyk-Hordern B. Multiresidue antibiotic-metabolite quantification method using ultra-performance liquid chromatography coupled with tandem mass spectrometry for environmental and public exposure estimation. *Anal Bioanal Chem* **413**, 5901-5920, doi:10.1007/s00216-021-03573-4 (2021).
153 Bengtsson-Palme, J. & Larsson, D. G. J. Concentrations of antibiotics predicted to select for resistant bacteria: Proposed limits for environmental regulation. *Environment International* **86**, 140-149, doi:https://doi.org/10.1016/j.envint.2015.10.015 (2016).
154 Bennett Institute for Applied Data Science. *OpenPrescribing.net*, <https://openprescribing.net> (2022).
155 Auguet, O. T. *et al.* Population-level faecal metagenomic profiling as a tool to predict antimicrobial resistance in Enterobacterales isolates causing invasive infections: An exploratory study across Cambodia, Kenya, and the UK. *EClinicalMedicine* **36**, 100910, doi:10.1016/j.eclinm.2021.100910 (2021).
156 Lu J, Breitwieser FP, Thielen P & SL, S. Bracken: estimating species abundance in metagenomics data. *PeerJ Computer Science*, e104, doi:10.7717/peerj-cs.104 (2017).
157 Beule, L. & Karlovsky, P. Improved normalization of species count data in ecology by scaling with ranked subsampling (SRS): application to microbial communities. *PeerJ* **8**, e9593, doi:10.7717/peerj.9593 (2020).
158 Fernandes, A. D., Macklaim, J. M., Linn, T. G., Reid, G. & Gloor, G. B. ANOVA-Like Differential Expression (ALDEx) Analysis for Mixed Population RNA-Seq. *PLOS ONE* **8**, e67019, doi:10.1371/journal.pone.0067019 (2013).
159 Nearing, J. T. *et al.* Microbiome differential abundance methods produce different results across 38 datasets. *Nature Communications* **13**, 342, doi:10.1038/s41467-022-28034-z (2022).
160 Hawinkel, S., Mattiello, F., Bijnens, L. & Thas, O. A broken promise: microbiome differential abundance methods do not control the false discovery rate. *Brief Bioinform* **20**, 210-221, doi:10.1093/bib/bbx104 (2019).

161 Calgaro, M., Romualdi, C., Waldron, L., Risso, D. & Vitulo, N. Assessment of statistical methods from single cell, bulk RNA-seq, and metagenomics applied to microbiome data. *Genome Biology* **21**, 191, doi:10.1186/s13059-020-02104-1 (2020).
162 Gloor, G. *et al. ANOVA-Like Differential Expression tool for high throughput sequencing data*, <https://rdrr.io/bioc/ALDEx2/f/vignettes/ALDEx2_vignette.Rmd> (2022).
163 Kim, Y. K. *et al.* The capacity of wastewater treatment plants drives bacterial community structure and its assembly. *Scientific Reports* **9**, 14809, doi:10.1038/s41598-019-50952-0 (2019).
164 LaMartina, E. L., Mohaimani, A. A. & Newton, R. J. Urban wastewater bacterial communities assemble into seasonal steady states. *Microbiome* **9**, 116, doi:10.1186/s40168-021-01038-5 (2021).
165 Deeny, S. R., van Kleef, E., Bou-Antoun, S., Hope, R. J. & Robotham, J. V. Seasonal changes in the incidence of Escherichia coli bloodstream infection: variation with region and place of onset. *Clinical Microbiology and Infection* **21**, 924-929, doi:https://doi.org/10.1016/j.cmi.2015.06.023 (2015).
166 Joseph, S. M., Battaglia, T., Maritz, J. M., Carlton, J. M. & Blaser, M. J. Longitudinal Comparison of Bacterial Diversity and Antibiotic Resistance Genes in New York City Sewage. *mSystems* **4**, e00327-00319, doi:doi:10.1128/mSystems.00327-19 (2019).
167 Sun, L., Klein, E. Y. & Laxminarayan, R. Seasonality and temporal correlation between community antibiotic use and resistance in the United States. *Clin Infect Dis* **55**, 687-694, doi:10.1093/cid/cis509 (2012).
168 Caucci, S. *et al.* Seasonality of antibiotic prescriptions for outpatients and resistance genes in sewers and wastewater treatment plant outflow. *FEMS Microbiol Ecol* **92**, fiw060, doi:10.1093/femsec/fiw060 (2016).
169 Munck, C. *et al.* Limited dissemination of the wastewater treatment plant core resistome. *Nat Commun* **6**, 8452, doi:10.1038/ncomms9452 (2015).
170 Lira, F., Vaz-Moreira, I., Tamames, J., Manaia, C. M. & Martínez, J. L. Metagenomic analysis of an urban resistome before and after wastewater treatment. *Sci Rep* **10**, 8174, doi:10.1038/s41598-020-65031-y (2020).
171 O'Brien, K. S. *et al.* Antimicrobial resistance following mass azithromycin distribution for trachoma: a systematic review. *Lancet Infect Dis* **19**, e14-e25, doi:10.1016/s1473-3099(18)30444-4 (2019).
172 Doan, T. *et al.* Macrolide and Nonmacrolide Resistance with Mass Azithromycin Distribution. *New England Journal of Medicine* **383**, 1941-1950, doi:10.1056/NEJMoa2002606 (2020).
173 Jarnheimer, P. A. *et al.* Fluoroquinolone antibiotics in a hospital sewage line; occurrence, distribution and impact on bacterial resistance. *Scand J Infect Dis* **36**, 752-755, doi:10.1080/00365540410021027 (2004).
174 Mouton, J. W. & van den Anker, J. N. Meropenem clinical pharmacokinetics. *Clin Pharmacokinet* **28**, 275-286, doi:10.2165/00003088-199528040-00002 (1995).
175 WHO. *Water, sanitation, hygiene, and waste management for SARS-CoV-2, the virus that causes COVID-19*, <https://www.who.int/publications/i/item/WHO-2019-nCoV-IPC-WASH-2020.4> (2020).
176 La Rosa, G. *et al.* SARS-CoV-2 has been circulating in northern Italy since December 2019: Evidence from environmental monitoring. *Sci Total Environ* **750**, 141711, doi:10.1016/j.scitotenv.2020.141711 (2021).
177 Hwengwere, K. *et al.* Antimicrobial resistance in Antarctica: is it still a pristine environment? *Microbiome* **10**, 71, doi:10.1186/s40168-022-01250-x (2022).
178 Hamilton, K. A. *et al.* Antimicrobial-resistant microorganisms and their genetic determinants in stormwater: A systematic review. *Current Opinion in Environmental Science & Health* **16**, 101-112, doi:https://doi.org/10.1016/j.coesh.2020.02.012 (2020).

179 Flach, C.-F., Hutinel, M., Razavi, M., Åhrén, C. & Larsson, D. G. J. Monitoring of hospital sewage shows both promise and limitations as an early-warning system for carbapenemase-producing Enterobacterales in a low-prevalence setting. *Water Research* **200**, 117261, doi:https://doi.org/10.1016/j.watres.2021.117261 (2021).