Title. Genome wide association studies define minimal set of loci for prediction of penicillin and tetracycline susceptibility in *Neisseria gonorrhoeae*

Authors. Tatum D. Mortimer, PhD[1]*, Jessica J. Zhang, AB[1]*, Kevin C. Ma, AB[1], Yonatan H. Grad, MD[1,2]

[1] Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[2] Division of Infectious Diseases, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

*These authors contributed equally

Corresponding author:
Yonatan H. Grad
Department of Immunology and Infectious Diseases
Harvard T.H. Chan School of Public Health
665 Huntington Ave
Building 1, Room 715
Boston, Massachusetts 02115
ygrad@hsph.harvard.edu

ABSTRACT

Background. While *Neisseria gonorrhoeae* poses an urgent public health threat because of increasing antimicrobial resistance, much of the circulating population remains susceptible to historical treatment regimens. Point-of-care diagnostics that report susceptibility could allow for reintroduction of these regimens, but development of such diagnostics has been limited to ciprofloxacin, for which susceptibility can be predicted from a single locus.

Methods. We assembled a dataset of 12,045 *N. gonorrhoeae* genomes with phenotypic resistance data for tetracycline (n = 3,611) and penicillin (n = 6,935). Using conditional genome wide association studies (GWAS), we sought to define genetic variants associated with susceptibility to penicillin and tetracycline. We evaluated the sensitivity and specificity of these variants for predicting susceptibility and non-resistance in our collection of gonococcal genomes.

Findings. In our conditional penicillin GWAS, the presence of a genetic variant defined by a non-mosaic *penA* allele without an insertion at codon 345 was significantly associated with penicillin susceptibility and had the highest negative effect size of significant variants ($p = 5.0 \times 10^{-14}$, $\beta = -2.5$). In combination with the absence of bla_{TEM}, this variant predicted penicillin susceptibility with high specificity (99.8%) and moderate sensitivity (36.7%). For tetracycline, the wild type allele at *rpsJ* codon 57, encoding valine, was significantly associated with tetracycline susceptibility ($p = 5.6 \times 10^{-16}$, $\beta = -1.61$) after conditioning on the presence of *tetM*. The combination of *rpsJ* codon 57 allele and *tetM absence* predicted tetracycline susceptibility with high specificity (97.2%) and sensitivity (88.7%).

Interpretation. As few as two genetic loci can predict susceptibility to penicillin and tetracycline in *N. gonorrhoeae* with high specificity. Molecular point-of-care diagnostics targeting these loci have the potential to increase available treatments for gonorrhea.

Funding. National Institute of Allergy and Infectious Diseases, the National Science Foundation, and the Smith Family Foundation

Introduction.

Gonorrhea, caused by infection with *Neisseria gonorrhoeae*, is the second most reported notifiable infection in the United States at a rate of 188.4 cases per 100,000 people in 2019[1] and increasing antibiotic resistance has made it an urgent public health threat.[2] Treatment is empiric, and resistance has limited the recommended treatment in the US to ceftriaxone, an extended spectrum cephalosporin (ESC).[3]

Despite the emergence of multidrug resistant strains,[4,5] a large fraction of clinical isolates remain susceptible to multiple antibiotics.[1] Data from the Gonococcal Isolate Surveillance Project (GISP), the US Centers for Disease Control and Prevention's sentinel surveillance system for antibiotic resistance in *N. gonorrhoeae*, reported that in 2019 44.5% of clinical isolates were not resistant to any tested antibiotics, 64.6% were non-resistant to ciprofloxacin (MIC < 1 µg/mL), 72.2% were non-resistant to tetracycline (MIC < 2 µg/mL), and 87.2% were non-resistant to penicillin (MIC < 2 µg/mL).[1]

Point-of-care diagnostics that inform on antibiotic susceptibility may help forestall the emergence and spread of resistance by enabling a shift from empiric to tailored treatment and expanding the number of antibiotics used to treat *N. gonorrhoeae* infections.[6–8] The observation that ciprofloxacin susceptibility can be predicted with high specificity and sensitivity based on *gyrA* codon 91 has led to the development of molecular tests that query this locus; the SpeeDx ResistancePlus GC, for example, was recently approved for clinical use in Europe and granted 'breakthrough designation' by the FDA.[9–11] However, extension of this sequence-based approach to other antibiotics has been stymied, as they lack single locus determinants of susceptibility and resistance.

Penicillin (PCN) and tetracycline (TET) were the recommended therapies for gonorrhea until the 1980s, when the prevalence of high level resistance increased enough to prompt a switch in the empiric treatment regimen.[12,13] Resistance to PCN and TET can be both chromosomal and plasmid mediated. Chromosomally-encoded resistance arises from mutations modifying the antibiotic targets—*rpsJ* for TET resistance and *penA* and *ponA* for PCN—and mutations in the porin *porB* and in the efflux pump *mtr* operon.[14] The plasmid-borne beta lactamase bla_{TEM} confers high level PCN resistance and the ribosome protection protein *tetM* confers TET resistance. Despite previously being first-line gonorrhea treatments for decades, molecular diagnostics for PCN and TET susceptibility have been less commonly studied. Proposed diagnostics or targets of molecular surveillance for PCN susceptibility have focused on bla_{TEM},[15,16] which performs poorly in the setting of chromosomally-encoded resistance, *porB,*[17] which neglects important target modifying mutations in *penA*, or resistance-associated *penA* alleles[18] rather than susceptibility-associated alleles. Similarly, assays targeting *tetM* have been developed, but they have not incorporated chromosomally-encoded tetracycline resistance.[16]

While there are multiple pathways to resistance for each drug,[14] the key goal for sequence-based diagnostics is to predict susceptibility—rather than resistance—with high specificity. Therefore, here we sought to identify a concise set of loci that are associated with PCN and TET susceptibility using genome wide association studies (GWAS) and evaluate their predictive performance in gonococcal clinical isolates.

Methods.

Global and validation dataset assembly. We collected publicly available whole genome sequencing (WGS) data (n = 12,045) and PCN (n = 6,935) and TET MICs (n = 5,727) from clinical *N. gonorrhoeae* isolates.[19–41] Genomes were assembled and resistance-associated alleles were called as previously described.[42] Specifically, alleles at known resistance loci were identified from variant calls after mapping to a reference genome for *rpsJ* codon 57, encoding the RpsJ V57M mutation associated with tetracycline resistance, and BLASTN[43] of *de novo* assemblies for accessory gene content using the sequences NG_068038.1 for bla_{TEM} and MG874353.1 for *tetM. penA* alleles were typed according to the NG-STAR database,[44] and the presence of insertions in *penA* were determined based on *de novo* assemblies*.* For 2116 isolates, TET MICs were reported as ≤ 4 µg/mL or ≤ 8 µg/mL. These were excluded them from phenotype-based analyses, since we could not classify them as susceptible or resistant. Additionally, Our final global dataset included 12,049 genotyped isolates, 3,611 with TET MICs and 6,935 with PCN MICs. To validate our results, we additionally assembled 1479 genomes from CDC's 2018 Gonococcal Isolate Surveillance Program (GISP) collection,[45] representing the first five viable isolates collected each month from urethral specimens at sentinel surveillance sites across the United States.

Genome wide association studies. To identify variants associated with PCN and TET susceptibility, we performed conditional GWAS[46] incorporating the presence of high effect size plasmid-mediated resistance. The GWAS employed a linear mixed model and were run using Pyseer v 1.2.0[47] with default allele frequency filters on 124,070 unitigs (unique sequences representing SNPs, insertions, deletions, and changes in gene content) for the PCN GWAS and 158,087 unitigs for the TET GWAS generated from GATB v 1.3.0. Most datasets reported PCN MICs within the range of 0.06 - 32 µg/mL. Isolates with PCN MICs reported as '>4' or '>2' were not included in the GWAS analysis since the precise MIC was unknown; the final PCN GWAS dataset size was 6220 isolates after excluding isolates with missing genotypic or phenotypic data. Similarly isolates with imprecise TET MICs were excluded (e.g. "≤ 4", "≤ 8"); the final dataset size for the TET GWAS was 3,453 isolates after excluding isolates with missing genotypic or phenotypic data. The GWAS incorporated isolate dataset of origin, country of origin, and presence of plasmid-encoded resistance determinants ($bla_{TEM,}$ *tetM*) as fixed effect covariates. We used a recombination-corrected phylogeny constructed with Gubbins v 2.3.4[48] to generate a similarity matrix, which was included as a random effect to correct for population structure. Unitigs from the GWAS were mapped and annotated using Pyseer annotate_hits_pyseer with the WHO_N genome[49] (GCA_900087725.2), which encodes the *β*-lactamase-encoding plasmid (African pblaTEM-1) and the *tetM*-encoding plasmid (Dutch), as the reference. The output was used to generate Manhattan plots in R v 4.0.3 with ggplot2.[50]

Sensitivity, specificity, and validation. We evaluated the sensitivity and specificity of resistance alleles to predict PCN and TET susceptibility using CLSI breakpoints for susceptibility (PCN MIC ≤ 0.06 µg/mL, TET MIC ≤ 0.25 µg/mL) and non-resistance (PCN MIC < 2 µg/mL, TET MIC < 2 µg/mL) in both the global and validation datasets. We additionally used isolate metadata from the 2018 GISP collection to estimate the prevalence of isolates with susceptibility associated genotypes across patient groups (sexual behavior, race/ethnicity). Statistical analyses were perfomed in R v 4.0.3 using infer v 0.5.4 (https://infer.tidymodels.org/).

Results.

Since plasmid-encoded resistance determinants are known to contribute to high level resistance for both PCN and TET, we used conditional GWAS to identify additional variants contributing to susceptibility. Here, we focused on significant variants associated with increased susceptibility (i.e., negative effect size, β). We found that a unitig (*penA*_01) corresponding to non-mosaic *penA* alleles without the insertion at codon 345 was significantly associated with PCN susceptibility (Supplementary Figure 1A, $p = 5.0 \times 10^{-14}$, $\beta = -2.5$). After conditioning on the presence of *tetM*, we found that a unitig corresponding to the wild-type allele at *rpsJ* codon 57, encoding valine, was significantly associated with TET susceptibility (Supplementary Figure 1B, $p = 5.6 \times 10^{-16}$, $\beta = -1.61$). Significant unitigs also mapped to *porB* and a loss of function variant in *mtrC* for both antibiotics; however, effect sizes were lower than unitigs mapping to antibiotic targets.

We used the presence of *penA*_01 combined with the absence of bla_{TEM} to predict PCN susceptibility in our global dataset (Figure 1A). We found that this susceptibility-associated genotype predicted PCN susceptibility and non-resistance with high specificity but low sensitivity (Table 1). For TET susceptibility prediction, we identified isolates with the wild-type allele at *rpsJ* codon 57 combined with the absence of *tetM* (Figure 1B). This combination predicted TET susceptibility and non-resistance with high specificity and moderate sensitivity (Table 1).

Since PCN and TET MICs were not reported for all isolates, we identified these mutations in our full genomic dataset: 2.1% had the PCN susceptibility-associated genotype, and 15.9% of isolates had the TET susceptibility-associated genotype. The prevalence of these genotypes varied across genomic epidemiology studies (Supplementary Table 1).

To validate our observations in a relatively unbiased dataset from the United States, we assembled a recently published collection of *N. gonorrhoeae* genomes from CDC's Gonococcal Isolate Surveillance Program[45]; in this collection, isolates were not selected for sequencing based on their susceptibility phenotypes. First, we verified that the *penA* sequence identified in the GWAS (*penA*_01) also identified isolates with non-mosaic *penA* alleles without the 345 insertion in the validation dataset. In this dataset, 100% of isolates with the *penA*_01 encoded non-mosaic *penA* alleles without the insertion when the full length *penA* allele was examined.

We also calculated sensitivity and specificity for prediction of PCN and TET susceptibility and non-resistance in the GISP collection (Figure 1, Table 1). Similar to results from the global collection, specificity was high for both antibiotics and CLSI cutoffs. Sensitivity increased for PCN prediction and decreased for TET prediction, reflecting different proportions of isolates falling into the susceptible and intermediate MIC categories in the global and validation datasets.

Since PCN and TET MICs were not reported for all isolates, we identified these mutations in our full genomic dataset: 2.1% had the PCN susceptibility-associated genotype, and 15.9% of isolates had the TET susceptibility-associated genotype. The prevalence of these genotypes varied across genomic epidemiology studies (Supplementary Table 1).

To validate our observations in a relatively unbiased dataset from the United States, we assembled a recently published collection of *N. gonorrhoeae* genomes from CDC's Gonococcal Isolate Surveillance Program[45]; in this collection, isolates were not selected for sequencing based on their susceptibility phenotypes. First, we verified that the *penA* sequence identified in

the GWAS (*penA*_01) also identified isolates with non-mosaic *penA* alleles without the 345 insertion in the validation dataset. In this dataset, 100% of isolates with the *penA*_01 encoded non-mosaic *penA* alleles without the insertion when the full length *penA* allele was examined.

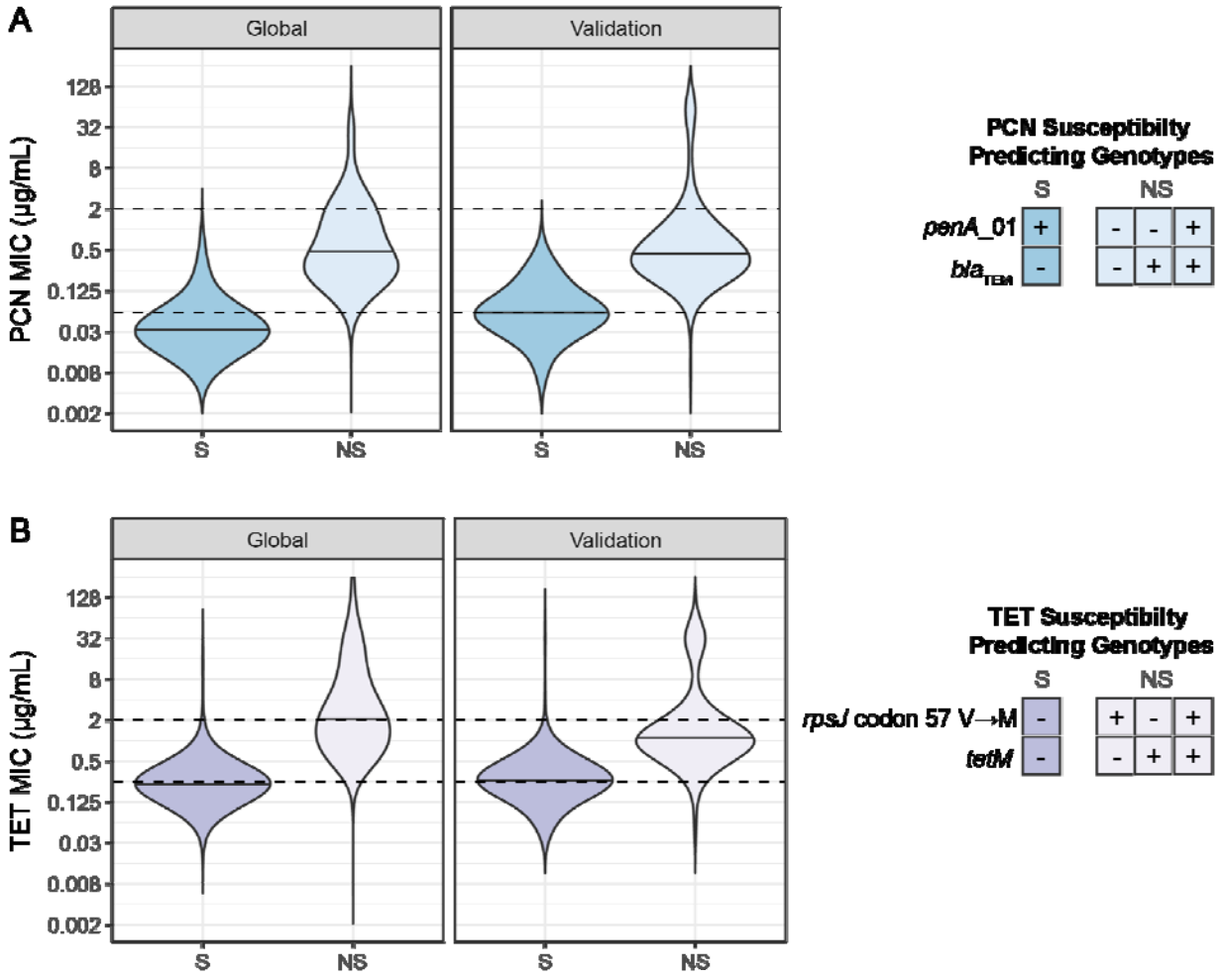

Figure 1. Penicillin and tetracycline MICs are lower in isolates with susceptibility associated genotypes in global and validation datasets. Dashed lines indicate CLSI breakpoints for susceptibility and resistance. A) Penicillin MICs of isolates with *penA*_01 and without *blaTEM* (S, dark blue) compared to isolates with one or more of these determinants (NS, light blue). B) Tetracycline MICs of isolates with wild-type *rpsJ* (57V) and without *tetM* (S, dark purple) compared to isolates with one or more of these determinants (NS, light purple).

We also calculated sensitivity and specificity for prediction of PCN and TET susceptibility and non-resistance in the GISP collection (Figure 1, Table 1). Similar to results from the global collection, specificity was high for both antibiotics and CLSI cutoffs. Sensitivity increased for PCN prediction and decreased for TET prediction, reflecting different proportions of isolates falling into the susceptible and intermediate MIC categories in the global and validation datasets.

Table 1. Sensitivity and specificity of genotypes for predicting PCN and TET susceptibility

Genotype	Phenotype	Global Dataset		Validation Dataset (GISP 2018[45])	
		Sensitivity	Specificity	Sensitivity	Specificity
*penA*_01 + absence of bla_{TEM}	PCN susceptible (MIC ≤ 0.06 µg/mL)	36.7%	99.8%	63.6%	98.9%
	PCN non-resistant (MIC < 2 µg/mL)	2.1%	100.0%	4.4%	100.0%
rpsJ WT + absence of *tetM*	TET susceptible (MIC ≤ 0.25 µg/mL)	88.7%	97.2%	78.2%	94.9%
	TET non-resistant (MIC < 2 µg/mL)	28.3%	99.7%	22.1%	99.5%

In addition to antimicrobial resistance phenotypes, GISP reports information on patient characteristics for each isolate collected. To analyze the utility of these genotypic markers in different patient populations, we calculated the prevalence of the susceptibility-associated genotypes across patient groups. Susceptible genotypes were more common among men who have sex with women (MSW) compared to men who have sex with men (MSM) and men who have sex with men and women (MSMW) for both PCN (χ^2 test, $p = 0.0035$) and TET (χ^2 test, $p < 0.0001$). The prevalence of the PCN susceptibility-associated genotype was 5.2%, 1.5%, and 2.2%, in MSW, MSM, and MSMW, respectively. For TET, the susceptibility-associated genotype was 20.6% in MSW, 9.6% in MSM, and 9.9% in MSMW. Additionally, the susceptibility-associated genotypes varied across race and ethnicity groups and were enriched in samples from Black men; however, prevalence of susceptibility-associated genotypes were not significantly different between race and ethnicity groups when MSM and MSW were considered separately (Supplementary Table 2).

Discussion. Here, we have used conditional GWAS incorporating known, high effect size variants[46] to identify targets for diagnostics addressing both plasmid and chromosomally mediated PCN and TET resistance. We found that the combination of *penA*_01, representing non-mosaic *penA* without an insertion at codon 345, and the absence of bla_{TEM} predicts PCN susceptibility and that the combination of *rpsJ* codon 57 and the absence of *tetM* predicts TET susceptibility. These loci defined the most susceptible isolates in our dataset and predicted susceptibility (PCN MIC ≤ 0.06 µg/mL, TET MIC ≤ 0.25 µg/mL) with high specificity and moderate sensitivity in both our global dataset and an unbiased collection from the United States. The addition of other loci (e.g., *mtr* and *porB*) may be needed to increase sensitivity for the higher cutoff (MIC < 2 µg/mL) but with as yet unclear impact on specificity.

These loci could additionally be used for culture-free molecular epidemiology and surveillance, as WGS directly from patient samples is not currently routine. Typing schemes such as NG-STAR[44] targeting resistance determinants have been developed; however, these schemes have not focused on loci specific to penicillin and tetracycline resistance.

Utility of a diagnostic or molecular surveillance targeting these loci may vary in different patient populations. For example, the prevalence of susceptibility associated genotypes varied across genomic epidemiology studies included in our global dataset, reflecting both enrichment of antibiotic resistant isolates in some studies as well as variable selection pressure from antibiotic use in different regions. WGS data from *N. gonorrhoeae* isolated in the United States, Europe,

and Australia make up the majority of available genomic data, and the composition of the *N. gonorrhoeae* population in other regions is unknown. Similar to other studies of the association between *N. gonorrhoeae* antibiotic resistance and patient demographics, prevalence of these susceptibility-associated genotypes vary across patient groups defined by sexual behavior and race/ethnicity in isolates collected by GISP.[34,37,39,51] In the United States, a diagnostic for PCN and TET susceptibility may be most useful in populations with higher prevalence of infection with susceptible isolates, such as MSW and women.

The alleles identified here from genomic analyses are promising targets for the development of POC molecular diagnostics for *N. gonorrhoeae* susceptibility to penicillin and tetracycline. Diagnostics that evaluate as few as two loci per drug could allow for the reintroduction into clinical use of these gonococcal treatment regimens.

Author Contributions. TDM and JJZ performed the GWAS and statistical analyses. TDM and KCM assembled genomic dataset. YHG supervised and managed the study. All authors contributed to data interpretation. TDM, JJZ, and YHG wrote the manuscript. All authors reviewed and edited the final manuscript. All authors had full access to the data in the study and were responsible for the decision to submit for publication.

Funding. TDM is supported by NIH NIAID (F32AI145157). YHG is supported by R01 AI132606 and R01 AI153521 and by the Smith Family Foundation. KCM is supported by NSF GRFP grant DGE1745303.

Conflicts of Interest. YHG is on the scientific advisory board of Day Zero Diagnostics, has consulted for Quidel and GSK, and has received grant funding from Merck, Pfizer, and GSK.

Data Sharing Statement. The analysis pipeline and data are available at github.com/gradlab/pcn_tet_susceptibility_gwas

References.

1 Centers for Disease Control and Prevention. Sexually Transmitted Disease Surveillance 2019. 2021.

2 Centers for Disease Control and Prevention (U.S.), National Center for Emerging Zoonotic and Infectious Diseases (U.S.). Division of Healthcare Quality Promotion. Antibiotic Resistance Coordination and Strategy Unit ., editors. Antibiotic resistance threats in the United States, 2019. 2019. https://stacks.cdc.gov/view/cdc/82532.

3 Cyr SS. Update to CDC's Treatment Guidelines for Gonococcal Infection, 2020. *MMWR Morb Mortal Wkly Rep* 2020; **69**. DOI:10.15585/mmwr.mm6950a6.

4 Eyre DW, Sanderson ND, Lord E, *et al.* Gonorrhoea treatment failure caused by a Neisseria gonorrhoeae strain with combined ceftriaxone and high-level azithromycin resistance, England, February 2018. *Eurosurveillance* 2018; **23**: 1800323.

5 Fifer H, Natarajan U, Jones L, *et al.* Failure of Dual Antimicrobial Therapy in Treatment of Gonorrhea. *New England Journal of Medicine* 2016; **374**: 2504–6.

6 Fingerhuth SM, Low N, Bonhoeffer S, Althaus CL. Detection of antibiotic resistance is essential for gonorrhoea point-of-care testing: a mathematical modelling study. *BMC Medicine* 2017; **15**: 142.

7 Tuite AR, Gift TL, Chesson HW, Hsu K, Salomon JA, Grad YH. Impact of Rapid Susceptibility Testing and Antibiotic Selection Strategy on the Emergence and Spread of Antibiotic Resistance in Gonorrhea. *The Journal of Infectious Diseases* 2017; **216**: 1141–9.

8 Turner KM, Christensen H, Adams EJ, *et al.* Analysis of the potential for point-of-care test to enable individualised treatment of infections caused by antimicrobial-resistant and susceptible strains of Neisseria gonorrhoeae: a modelling study. *BMJ Open* 2017; **7**: e015447.

9 Ebeyan S, Windsor M, Bordin A, *et al.* Evaluation of the ResistancePlus GC (beta) assay: a commercial diagnostic test for the direct detection of ciprofloxacin susceptibility or resistance in Neisseria gonorrhoeae. *Journal of Antimicrobial Chemotherapy* 2019; **74**: 1820–4.

10 Ltd SP. SpeeDx Receives CE-IVD Mark for ResistancePlus® GC Gonorrhea Test. GlobeNewswire News Room. 2018; published online Dec 6. https://www.globenewswire.com/news-release/2018/12/06/1662818/0/en/SpeeDx-Receives-CE-IVD-Mark-for-ResistancePlus-GC-Gonorrhea-Test.html (accessed April 28, 2021).

11 SpeeDx Receives FDA Breakthrough Designation for ResistancePlus® GC. 2019; published online June 11. https://www.businesswire.com/news/home/20190611005313/en/SpeeDx-Receives-FDA-Breakthrough-Designation-for-ResistancePlus%C2%AE-GC (accessed April 28, 2021).

12 1985 STD Treatment Guidelines. *Morbidity and Mortality Weekly Report: Supplement* 1985; **34**: iS–108S.

13 1989 Sexually Transmitted Diseases Treatment Guidelines. *Morbidity and Mortality Weekly Report: Recommendations and Reports* 1989; **38**: i–43.

14 Mortimer TD, Grad YH. Applications of genomics to slow the spread of multidrug-resistant Neisseria gonorrhoeae. *Ann N Y Acad Sci* 2018; published online June 6. DOI:10.1111/nyas.13871.

15 Buckley C, Trembizki E, Baird RW, *et al.* Multitarget PCR Assay for Direct Detection of Penicillinase-Producing Neisseria gonorrhoeae for Enhanced Surveillance of Gonococcal Antimicrobial Resistance. *J Clin Microbiol* 2015; **53**: 2706–8.

16 Fayemiwo SA, Müller EE, Gumede L, Lewis DA. Plasmid-Mediated Penicillin and Tetracycline Resistance Among Neisseria gonorrhoeae Isolates in South Africa: Prevalence, Detection and Typing Using a Novel Molecular Assay. *Sexually Transmitted Diseases* 2011; **38**: 329–33.

17 Buckley C, Trembizki E, Donovan B, *et al.* Real-time PCR detection of Neisseria gonorrhoeae susceptibility to penicillin. *J Antimicrob Chemother* 2016; **71**: 3090–5.

18 Vernel-Pauillac F, Merien F. A Novel Real-Time Duplex PCR Assay for Detecting penA and ponA Genotypes in Neisseria gonorrhoeae: Comparison with Phenotypes Determined by the E-Test. *Clinical Chemistry* 2006; **52**: 2294–6.

19 Alfsnes K, Eldholm V, Olsen AO, *et al.* Genomic epidemiology and population structure of Neisseria gonorrhoeae in Norway, 2016–2017. *Microbial Genomics* 2020. DOI:10.1099/mgen.0.000359.

20 Buckley C, Forde BM, Trembizki E, Lahra MM, Beatson SA, Whiley DM. Use of whole genome sequencing to investigate an increase in Neisseria gonorrhoeae infection among women in urban areas of Australia. *Scientific Reports* 2018; **8**: 1503.

21 Cehovin A, Harrison OB, Lewis SB, *et al.* Identification of Novel Neisseria gonorrhoeae Lineages Harboring Resistance Plasmids in Coastal Kenya. *J Infect Dis* 2018; **218**: 801–8.

22 De Silva D, Peters J, Cole K, *et al.* Whole-genome sequencing to determine transmission of Neisseria gonorrhoeae: an observational study. *The Lancet Infectious Diseases* 2016; **16**: 1295–303.

23 Demczuk W, Lynch T, Martin I, *et al.* Whole-Genome Phylogenomic Heterogeneity of Neisseria gonorrhoeae Isolates with Decreased Cephalosporin Susceptibility Collected in Canada between 1989 and 2013. *J Clin Microbiol* 2015; **53**: 191–200.

24 Demczuk W, Martin I, Peterson S, *et al.* Genomic Epidemiology and Molecular Resistance Mechanisms of Azithromycin-Resistant Neisseria gonorrhoeae in Canada from 1997 to 2014. *J Clin Microbiol* 2016; **54**: 1304–13.

25 Eyre DW, De Silva D, Cole K, *et al.* WGS to predict antibiotic MICs for Neisseria gonorrhoeae. *J Antimicrob Chemother* 2017; **72**: 1937–47.

26 Ezewudo MN, Joseph SJ, Castillo-Ramirez S, *et al.* Population structure of Neisseria gonorrhoeae based on whole genome data and its relationship with antibiotic resistance. *PeerJ* 2015; **3**: e806.

27 Fifer H, Cole M, Hughes G, *et al.* Sustained transmission of high-level azithromycin-resistant Neisseria gonorrhoeae in England: an observational study. *Lancet Infect Dis* 2018; **18**: 573–81.

28 Grad YH, Kirkcaldy RD, Trees D, *et al.* Genomic epidemiology of Neisseria gonorrhoeae with reduced susceptibility to cefixime in the USA: a retrospective observational study. *The Lancet Infectious Diseases* 2014; **14**: 220–6.

29 Grad YH, Harris SR, Kirkcaldy RD, *et al.* Genomic Epidemiology of Gonococcal Resistance to Extended-Spectrum Cephalosporins, Macrolides, and Fluoroquinolones in the United States, 2000–2013. *J Infect Dis* 2016; **214**: 1579–87.

30 Harris SR, Cole MJ, Spiteri G, *et al.* Public health surveillance of multidrug-resistant clones of Neisseria gonorrhoeae in Europe: a genomic survey. *The Lancet Infectious Diseases* 2018; **18**: 758–68.

31 Kwong JC, Chow EPF, Stevens K, *et al.* Whole-genome sequencing reveals transmission of gonococcal antibiotic resistance among men who have sex with men: an observational study. *Sex Transm Infect* 2017; : sextrans-2017-053287.

32 Lan PT, Golparian D, Ringlander J, Van Hung L, Van Thuong N, Unemo M. Genomic analysis and antimicrobial resistance of Neisseria gonorrhoeae isolates from Vietnam in 2011 and 2015–16. *J Antimicrob Chemother* DOI:10.1093/jac/dkaa040.

33 Lee RS, Seemann T, Heffernan H, *et al.* Genomic epidemiology and antimicrobial resistance of Neisseria gonorrhoeae in New Zealand. *J Antimicrob Chemother* 2018; **73**: 353–64.

34 Mortimer TD, Pathela P, Crawley A, *et al.* The distribution and spread of susceptible and resistant Neisseria gonorrhoeae across demographic groups in a major metropolitan center. *Clin Infect Dis* DOI:10.1093/cid/ciaa1229.

35 Peng J-P, Yin Y-P, Chen S-C, *et al.* A Whole-genome Sequencing Analysis of Neisseria gonorrhoeae Isolates in China: An Observational Study. *EClinicalMedicine* 2019; **7**: 47–54.

36 Ryan L, Golparian D, Fennelly N, *et al.* Antimicrobial resistance and molecular epidemiology using whole-genome sequencing of Neisseria gonorrhoeae in Ireland, 2014–2016: focus on extended-spectrum cephalosporins and azithromycin. *Eur J Clin Microbiol Infect Dis* 2018; **37**: 1661–72.

37 Sánchez-Busó L, Golparian D, Corander J, *et al.* The impact of antimicrobials on gonococcal evolution. *Nature Microbiology* 2019; **4**: 1941–50.

38 Thomas JC, Seby S, Abrams AJ, *et al.* Evidence of Recent Genomic Evolution in Gonococcal Strains with Decreased Susceptibility to Cephalosporins or Azithromycin in the United States, 2014-2016. *J Infect Dis* DOI:10.1093/infdis/jiz079.

39 Town K, Harris S, Sánchez-Busó L, *et al.* Genomic and Phenotypic Variability in Neisseria gonorrhoeae Antimicrobial Susceptibility, England - Volume 26, Number 3—March 2020 - Emerging Infectious Diseases journal - CDC. DOI:10.3201/eid2603.190732.

40 Williamson DA, Chow EPF, Gorrie CL, *et al.* Bridging of Neisseria gonorrhoeae lineages across sexual networks in the HIV pre-exposure prophylaxis era. *Nat Commun* 2019; **10**: 1–10.

41 Yahara K, Nakayama S-I, Shimuta K, *et al.* Genomic surveillance of Neisseria gonorrhoeae to investigate the distribution and evolution of antimicrobial-resistance determinants and lineages. *Microb Genom* 2018; **4**. DOI:10.1099/mgen.0.000205.

42 Ma KC, Mortimer TD, Hicks AL, *et al.* Adaptation to the cervical environment is associated with increased antibiotic susceptibility in Neisseria gonorrhoeae. *Nature Communications* 2020; **11**: 4126.

43 Camacho C, Coulouris G, Avagyan V, *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* 2009; **10**: 421.

44 Demczuk W, Sidhu S, Unemo M, *et al.* Neisseria gonorrhoeae Sequence Typing for Antimicrobial Resistance, a Novel Antimicrobial Resistance Multilocus Typing Scheme for

Tracking Global Dissemination of N. gonorrhoeae Strains. *J Clin Microbiol* 2017; **55**: 1454–68.

45 Reimche JL, Chivukula VL, Schmerer MW, *et al.* Genomic analysis of the predominant strains and antimicrobial resistance determinants within 1479 Neisseria gonorrhoeae isolates from the U.S. Gonococcal Isolate Surveillance Project in 2018. *Sex Transm Dis* 2021; published online May 14. DOI:10.1097/OLQ.0000000000001471.

46 Ma KC, Mortimer TD, Duckett MA, *et al.* Increased power from conditional bacterial genome-wide association identifies macrolide resistance mutations in Neisseria gonorrhoeae. *Nature Communications* 2020; **11**: 5374.

47 Lees JA, Galardini M, Bentley SD, Weiser JN, Corander J. pyseer: a comprehensive tool for microbial pangenome-wide association studies. *Bioinformatics* 2018; **34**: 4310–2.

48 Croucher NJ, Page AJ, Connor TR, *et al.* Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucl Acids Res* 2015; **43**: e15–e15.

49 Unemo M, Golparian D, Sánchez-Busó L, *et al.* The novel 2016 WHO Neisseria gonorrhoeae reference strains for global quality assurance of laboratory investigations: phenotypic, genetic and reference genome characterization. *J Antimicrob Chemother* 2016; **71**: 3096–108.

50 Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016 https://ggplot2.tidyverse.org.

51 Bauer HM, Mark KE, Samuel M, *et al.* Prevalence of and Associated Risk Factors for Fluoroquinolone-Resistant Neisseria gonorrhoeae in California, 2000–2003. *Clin Infect Dis* 2005; **41**: 795–803.