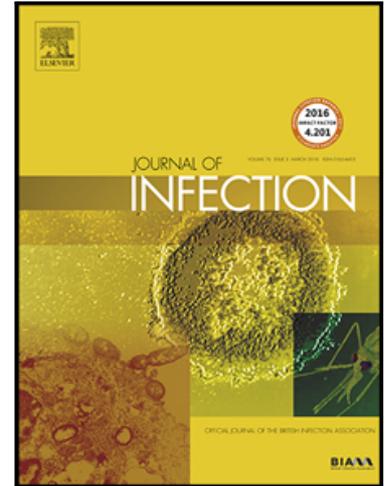


## Accepted Manuscript

*Neisseria gonorrhoeae* molecular typing for understanding sexual networks and antimicrobial resistance transmission: a systematic review

Katy Town , Hikaru Bolt , Sara Croxford , Michelle Cole ,  
Simon Harris , Nigel Field , Gwenda Hughes

PII: S0163-4453(18)30085-9  
DOI: [10.1016/j.jinf.2018.02.011](https://doi.org/10.1016/j.jinf.2018.02.011)  
Reference: YJINF 4064



To appear in: *Journal of Infection*

Received date: 3 November 2017  
Revised date: 11 February 2018  
Accepted date: 19 February 2018

Please cite this article as: Katy Town , Hikaru Bolt , Sara Croxford , Michelle Cole , Simon Harris , Nigel Field , Gwenda Hughes , *Neisseria gonorrhoeae* molecular typing for understanding sexual networks and antimicrobial resistance transmission: a systematic review, *Journal of Infection* (2018), doi: [10.1016/j.jinf.2018.02.011](https://doi.org/10.1016/j.jinf.2018.02.011)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Highlights**

- Combined molecular and epidemiological data can describe the spread of gonorrhoea
- Sexual networks can be inferred from molecular clusters of infection
- Gender and sexual orientation are commonly used to characterise these networks
- Application of these data within gonorrhoea control interventions is limited
- Future studies should focus on evaluating molecular typing data in practice

ACCEPTED MANUSCRIPT

## ***Neisseria gonorrhoeae* molecular typing for understanding sexual networks and antimicrobial resistance transmission: a systematic review**

Katy Town (MSc)<sup>a,b</sup>, Hikaru Bolt (MSc)<sup>c</sup>, Sara Croxford (MSc)<sup>c</sup>, Michelle Cole (DBMS)<sup>d</sup>, Simon Harris (PhD)<sup>e</sup>, Nigel Field (MBPhD)<sup>b</sup> and Gwenda Hughes (PhD)<sup>a,c</sup>

A. National Institute for Health Research Health Protection Research Unit in Blood Borne and Sexually Transmitted Infections at University College London in partnership with Public Health England and in collaboration with the London School of Hygiene and Tropical Medicine, Mortimer Market Centre, Capper Street, London, WC1E 6JB, UK

B. Centre for Molecular Epidemiology and Translational Research, Institute for Global Health, University College London, Mortimer Market Centre, Capper Street, London, WC1E 6JB, UK

C. HIV/STI Department, National Infection Service, Public Health England, 61 Colindale Avenue, London, NW9 5EQ, UK

D. Antimicrobial Resistance and Healthcare Associated Infections (AMRHAI) Reference Unit, National Infection Service, Public Health England, 61 Colindale Avenue, London, NW9 5EQ, UK

E. The Wellcome Trust Sanger Institute, Genome Campus, Cambridge, CB10 1SA, UK

**Corresponding author:** Katy Town, Third Floor, Mortimer Market Centre, Capper Street, London, WC1E 6JB, UK. [katy.town.15@ucl.ac.uk](mailto:katy.town.15@ucl.ac.uk) +44(0)2083277493

## **Summary**

### **Objectives**

*Neisseria gonorrhoeae* (NG) is a significant global public health concern due to rising diagnoses rates and antimicrobial resistance. Molecular combined with epidemiological data have been used to understand the distribution and spread of NG, as well as relationships between cases in sexual networks, but the public health value gained from these studies is unclear. We conducted a systematic review to examine how molecular epidemiological studies have informed understanding of sexual networks and NG transmission, and subsequent public health interventions.

### **Methods**

Five research databases were systematically searched up to 31<sup>st</sup> March 2017 for studies that used sequence-based DNA typing methods, including whole genome sequencing, and linked molecular data to patient-level epidemiological data. Data were extracted and summarised to identify common themes.

### **Results**

Of the 49 studies included, 82% used NG Multi-antigen Sequence Typing. Gender and sexual orientation were commonly used to characterise sexual networks that were inferred using molecular clusters; clusters predominantly of one patient group often contained a small number of isolates from other patient groups. Suggested public health applications included using these data to target interventions at specific populations, confirm outbreaks, and inform partner management, but these were mainly untested.

### **Conclusions**

Combining molecular and epidemiological data has provided insight into sexual mixing patterns, and dissemination of NG, but few studies have applied these findings to design or evaluate public health interventions. Future studies should focus on the application of molecular epidemiology in public health practice to provide evidence for how to prevent and control NG.

## Introduction

*Neisseria gonorrhoeae* is a sexually transmitted pathogen of significant public health concern due to rising diagnosis rates, particularly in men who have sex with men (MSM), and the emergence of resistance to all classes of antimicrobials used for treatment.<sup>1,2</sup> It is important to better understand how and why *N. gonorrhoeae* and resistant infections spread within sexual networks in order to design targeted and rational interventions to control transmission.

Traditional epidemiological methods using surveillance data and patient questionnaires of reported behaviour to understand and describe sexual networks have been used to improve our understanding of the distribution and transmission of *N. gonorrhoeae*, including the concentration of *N. gonorrhoeae* in specific groups at risk.<sup>3,4</sup> However, these methods are limited in their ability to determine whether infected individuals with the same epidemiological characteristics are part of the same transmission network. Establishing direct epidemiological links is challenging because gonorrhoea is frequently asymptomatic, especially in females, which may lead to missing (undiagnosed) cases in most datasets.

Molecular technologies can be used to group isolates according to similarities in their genetic data and infer relatedness between isolates.<sup>5</sup> By combining these molecular data with epidemiological data, hypotheses about the transmission of *N. gonorrhoeae* within and across different sexual networks can be tested. The use of molecular epidemiology in the analysis of *N. gonorrhoeae* populations is a rapidly evolving field but there has been no systematic assessment of its public health value.

In this systematic review, we aimed to use published literature to investigate how linked *N. gonorrhoeae* epidemiological and molecular typing data can enhance our understanding of sexual networks and pathogen transmission and how this information has been used within public health interventions to control gonorrhoea infections.

## Methods

### Search strategy and selection criteria

In this systematic review (reported according to the international prospective register of systematic reviews: PROSPERO 2016 CRD 42016037238),<sup>6</sup> studies were included if sequence based DNA typing methods were used (Multi-locus Sequence Typing (MLST), *N. gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) and Whole Genome Sequencing (WGS)) (see Supplementary material for Glossary of Terms) and the typing data were linked to patient-level epidemiological data, including patient demographic data (e.g. gender, age, and ethnicity), sexual behaviour data (e.g. sexual orientation, sex work, condom use, number of partners, sex abroad) and/or clinical data (e.g. symptoms, site of infection, concurrent STIs, HIV status). We focused on the sequence-based DNA typing techniques described above, as these are currently the most commonly used and recommended<sup>7</sup> techniques for molecular epidemiological studies. Studies were excluded if the articles were not in English, if the typing data were only linked to the geographical location and/or date of the isolates and not to patient epidemiological data, if typing data were only used to

investigate the pathobiology of *N. gonorrhoeae*, or to develop methods for typing of *N. gonorrhoeae*.

We searched Web of Science, Scopus, MEDLINE, EMBASE, and the Cochrane library from inception to 31<sup>st</sup> March 2017. Search terms included the typing methods of interest (NG-MAST, MLST, and WGS) and terms related to *Neisseria gonorrhoeae*. The full search strategies are detailed in the Supplementary Material, including details of conference abstract books that were searched for relevant studies.

The title, abstract and then full text of studies were assessed against the eligibility criteria. Where no full text was found, study authors were contacted by email to request a copy of the paper. Authors of conference abstracts were also contacted for the poster or presentation slide set.

#### Data extraction and synthesis

Data were extracted using a standardised form and summarised in a descriptive table. Common themes were identified and are presented using a narrative approach. Risk of bias was assessed considering the potential for (1) selection bias affecting the interpretation and generalisability of results based on the isolates with typing data available and the wider population that the study aimed to represent, (2) missing data bias, and whether data were likely to be missing at random or not, and (3) reporting bias, particularly for reported behavioural variables.

#### Dual independent review

Each article was reviewed, data extracted and risk of bias assessed independently by two reviewers (KT and HB, or KT and SC); disagreements were resolved at a meeting between the reviewers.

## **Results**

#### Article screening

Our search strategy identified 4,759 studies, of which 2,101 were unique (Figure 1). Most studies were excluded following title and abstract review (94%; 1,982/2,101). Of the remaining 119 studies, 50 met the study eligibility criteria (42%; 50/119). One study was not reviewed as the full-text could not be accessed.<sup>8</sup> Studies were most commonly excluded at this stage because no patient data were reported (74%; 51/69). Nine conference abstracts were identified, four of which had been subsequently published. Authors for the remaining five were contacted for the presented poster or slide set but none replied (Supplementary material). Hence, 49 studies were included in this systematic review.

#### Description of included studies

Over one third of studies used isolates collected in the UK (37%; 18/49),<sup>9-26</sup> from consecutive patients during the study time period (35%; 17/49),<sup>11,14-17,21,22,24,27-35</sup> or limited inclusion of isolates to those known to be resistant to a specific antimicrobial (35%; 17/49).<sup>9,10,13,18-20,23,36-45</sup> The number of isolates sampled ranged from six to 3,326, with studies covering a time period of less than six months to nine years (Supplementary Material).

Most studies used NG-MAST (82%; 40/49),<sup>9-22,26-33,35-44,46-53</sup> two used MLST (4%; 2/49),<sup>34,54</sup> and seven used WGS (14%; 7/49),<sup>23-25,45,55-57</sup> (Figure 2). The WGS studies also determined the NG-MAST and/or MLST sequence type *in silico* and compared typing methods. Although the clustering identified through NG-MAST and MLST methods broadly correlated with the clustering determined by WGS-based phylogenetic analyses, the WGS analyses showed that the other methods sometimes misclassified clustering. For example, Didelot *et al.*<sup>25</sup> showed that isolates with an identical NG-MAST type that occurred within the same time period and location, and which would otherwise have been considered to be part of the same sexual network, were actually genetically distinct when WGS was used.

#### Risk of bias

Most studies applied specific selection criteria to the isolates chosen for typing, which mainly related to the setting and location isolates were selected from, the time period for data collection, or the selection of isolates with a specific antimicrobial susceptibility profile. This may limit the generalisability of the results. Most studies typed  $\geq 70\%$  of isolates eligible under the predetermined selection criteria of the study. The main reasons not all isolates were typed were that the isolate could not be retrieved or the typing method failed. In general missing epidemiological data were minimal. However, most studies did not define how patient-level variables were collected, and it was therefore often difficult to assess reporting bias. It is likely that variables related to sexual behaviour, such as sexual orientation, were self-reported, which increases the risk of reporting bias. Further details for each study are presented in the supplementary material.

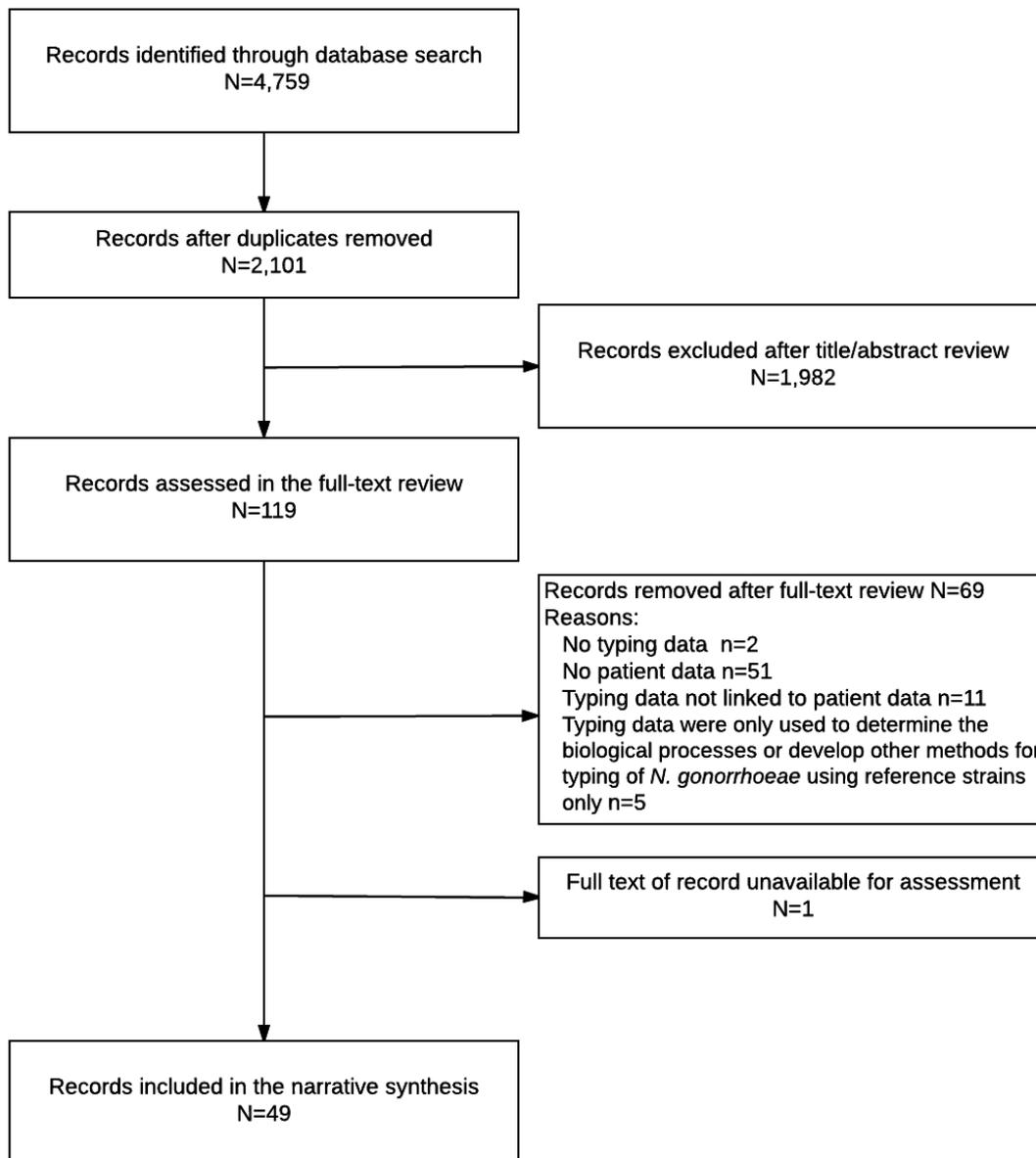
#### Patient-level epidemiological data

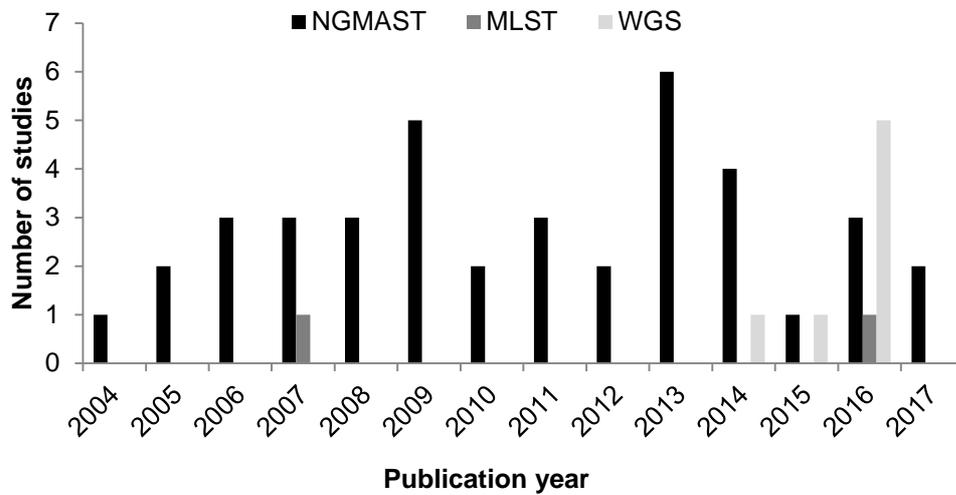
The most commonly reported and linked patient-level epidemiological data were gender (reported 100%; 49/49<sup>9-57</sup>, linked 90%; 44/49<sup>9,11-44,46-49,53-57</sup>) and sexual orientation (reported 82%; 40/49<sup>9-11,13-23,25-30,32,33,36-42,45-53,55,57</sup>, linked 71%; 35/49<sup>9,11,13-23,25-30,32,33,36-42,46-49,53,55,57</sup>).

Other patient-level epidemiological data (not always linked to molecular data) included age, sexual partner history (such as the number of sexual partners), location of patient residence or clinic, site of infection, travel associated sex, STI history, patient HIV status, patient ethnicity, infection symptoms and sexual behaviour (Figure 3).

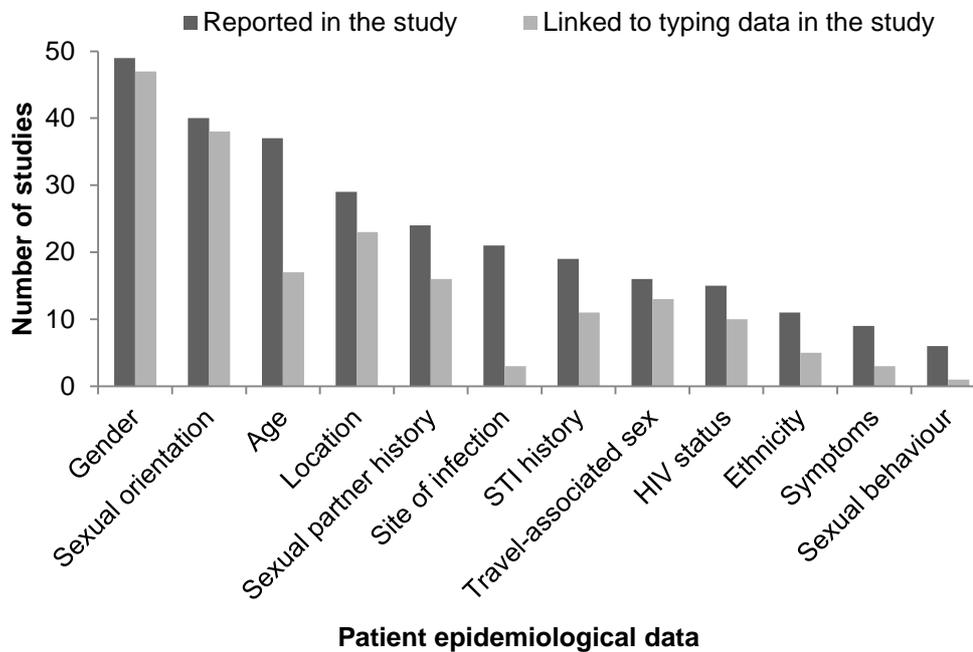
**Figure 1: PRISMA flowchart indicating the systematic selection of journal articles for inclusion in this review**

PRISMA flowchart for the conference abstract book is presented in the Supplementary Material.



**Figure 2: Number of studies by typing method over time (N=49)**

NGMAST = *Neisseria gonorrhoeae* multi-antigen sequence typing  
 MLST = multi-locus sequence typing  
 WGS = whole genome sequencing

**Figure 3: Reporting and linkage of patient epidemiological data to typing data (N=49)**

## Summary of study findings

### *Identifying clusters of infection, sexual networks and transmission links*

The majority of studies that used NG-MAST or MLST defined a cluster by the presence of two or more isolates with identical sequence types.<sup>9,10,16,29,31,42</sup> Closely related sequence types (defined as varying genetically by  $\leq 1\%$ ) were sometimes grouped into larger categories called genogroups based on sequence similarity. Most studies identified a large number of different sequence types: many of which had not been identified before ('novel' sequence types) and/or were present only once in the isolate set ('unique' sequence types).<sup>29-31,53</sup> For example, Chisholm *et al* (2013)<sup>29</sup> found 406 NG-MAST sequence types amongst 1,066 isolates. Over half of these were novel (216/406) and/or unique (281/406). WGS studies used maximum likelihood or Bayesian phylogenetic statistical modelling techniques with information on the number of single nucleotide polymorphisms (SNPs) to identify similarities and allocate isolates into clusters.

The molecular data were also used to provide evidence about the likelihood that transmission of *N. gonorrhoeae* had occurred between two patients.<sup>9,11,12,14,16,17,23-25,38,53</sup> The degree of concordance between sexual contact pairs (confirmed through partner notification data) and genetic data, such as NG-MAST sequence type, was  $>85\%$  in these studies.

### *Describing sexual networks using typing and patient-level epidemiological data*

Patient-level epidemiological data were used to characterise and compare clusters of infection (Figure 3). For example, Horn *et al.*<sup>31</sup> found that isolates classified as genogroup 25 were predominantly from women, whereas those classified as genogroup 1407 were predominantly from men. In general, only the most common sequence types, which formed the largest clusters, were described in this way.<sup>10,11,15-19,21,27,31,32,40,47,48,50,53</sup> Where sequence types were statistically associated with one group of patients, such as MSM, authors inferred that this provided evidence of a discrete sexual network. Further examples are described below:

#### (i) Sexual orientation:

Overall, 38 studies investigated the relationship between sexual orientation and *N. gonorrhoeae* sequence type. Most (87%; 33/38) identified a difference in the sequence types circulating between MSM and heterosexual patients. For example, Cole *et al.*<sup>16</sup> found a statistically significant association between NG-MAST sequence types and sexual orientation: sequence types 147, 4, 1634 and 64 were more likely to be from MSM than heterosexual patients. However, often, clusters with a predominating patient characteristic also contained a small number of isolates from other patient groups. For example, Wong *et al.*<sup>32</sup> found that NG-MAST sequence type 547 was associated with MSM but four of the 16 isolates were from heterosexual men. There was evidence of the same sequence type being associated with different sexual orientation characteristics in other studies. For example, using isolates collected in 2003, Abu-Rajab *et al.*<sup>11</sup> found that NG-MAST sequence type 210 was associated with heterosexuals in Glasgow, Scotland. A year later, in 2004, Choudhury *et al.*<sup>15</sup> found the same NG-MAST sequence type was associated with MSM in London, England.

(ii) Sub-groups within MSM and heterosexual networks:

Half of the studies that included patient sexual orientation (17/38) investigated the existence of discrete transmission clusters within larger sexual networks by separately analysing the association between sequence types and patient variables within these sub-groups of the population. This enhanced analysis helped describe clusters of infection in more detail. For example, Choudhury *et al.*<sup>15</sup> found that isolates from heterosexual patients had NG-MAST sequence types that varied by ethnicity and age, indicating discrete sexual networks, but they did not identify similar patterns associated with specific characteristics in MSM. In contrast, Bernstein *et al.*<sup>27</sup> found that some NG-MAST sequence types were associated with specific characteristics in MSM. For example, isolates classified as NG-MAST sequence type 2992 were more likely to be from MSM who reported oral sex only, and other NG-MAST sequence types varied by reported sexual risk behaviour (three or more sexual partners compared to fewer sexual partners).

(iii) HIV status:

There were 10 studies that compared patient HIV status to the typing data to investigate whether the molecular data provided evidence of discrete sexual networks in people with and without HIV.<sup>15-17,27,32,42,46,47,50,53</sup> Not all studies found the same result, which is likely due to the different populations and sexual behaviours of these populations investigated in each study. For example Didelot *et al.*<sup>25</sup> found that *N. gonorrhoeae* was more likely to be transmitted between two HIV-positive patients than between a HIV-positive and HIV-negative patient, suggestive of serosorting sexual behaviour (the term for engaging in condomless anal intercourse with partners of the same known or presumed HIV status).<sup>58</sup> However, Bernstein *et al.*<sup>27</sup> and Cheng *et al.*<sup>47</sup> did not find an association between molecular clusters and HIV status.

(iv) Travel-associated sexual partnerships:

In 13 studies<sup>9-11,13,15,16,19,20,24,25,28,30,43</sup>, the molecular typing data was used to describe and investigate the role of travel-associated sexual partnerships on *N. gonorrhoeae* transmission and acquisition. Some studies found that unique sequence types were more likely to have been acquired outside of the local area, whereas clustered isolates were more likely to have been acquired within the local area<sup>10,15,19,20</sup>. Fernando *et al.*<sup>17</sup> found that patients attending a specialist STI clinic in Edinburgh, Scotland, infected with unique *N. gonorrhoeae* NG-MAST sequence types were more likely to have reported recent sexual contacts from outside the local area. Similarly, Martin *et al.* (2005)<sup>10</sup> found that the common strains types identified in London, England, were more likely to be from patients who did not report sex abroad.

*Describing antimicrobial resistance (AMR) within sequence types*

To describe and identify sequence types of *N. gonorrhoeae* associated with AMR, phenotypic antimicrobial susceptibility data were combined with molecular typing data in 83% of studies (41/49). These studies found that susceptible *N. gonorrhoeae* were genetically more diverse than resistant *N. gonorrhoeae*, as evidenced by the higher number of different NG-MAST sequence types in the susceptible clones.<sup>36,49,54</sup> Non-susceptible *N. gonorrhoeae* tended to be more clonal, for example, *N. gonorrhoeae* with decreased susceptibility to cefixime was usually detected in isolates identified as the NG-MAST

sequence type 1407.<sup>13,18,30,31,36,37,48,53</sup> Similar clonality of the gonococcal population was found for other antimicrobial resistant phenotypes, such as ciprofloxacin, penicillin and tetracycline. Azithromycin resistance was found to occur sporadically in the gonococcal population rather than be associated with a particular clone,<sup>42,55</sup> with the exception of high-level azithromycin resistant isolates (minimum inhibitory concentration >256mg/L) that were found in clusters of closely related gonococci.<sup>14,23,35,44</sup>

When the phenotypic data, molecular data and patient-level epidemiological data were all combined, this provided more insight into the distribution and spread of resistant *N. gonorrhoeae* infection in sexual networks or across populations.<sup>10,14,21,27,29,47,50,56,57</sup> For example, using WGS, Grad *et al.* (2014)<sup>57</sup> reported that *N. gonorrhoeae* with decreased susceptibility to cefixime first appeared on the west coast of the United States and then spread eastwards primarily within MSM networks, but with a small number of diagnoses in heterosexual patients. Authors often compared their findings to other studies to assess whether there were sexual networks operating between cities or countries and how this contributed to the spread of resistant infection. For example, Chisholm *et al.* (2009)<sup>14</sup> speculated that high-level azithromycin-resistant *N. gonorrhoeae* (MIC  $\geq$ 256 mg/L) identified in Liverpool, England, was imported from Scotland because the NG-MAST sequence types (ST649), AMR phenotype and sexual orientation of the infected patients matched those of cases previously reported by Palmer *et al.* (2008).<sup>21</sup>

#### *Public health application of findings from included studies*

Many studies stated that combining molecular typing data and patient-level epidemiological data could be used to support decisions about *N. gonorrhoeae* prevention and control activities (Table 1). Over half of the studies (54%; 26/49)<sup>9-11,15-22,24,25,27,32-34,39-41,44,50,51,56,57</sup> suggested that these data could be used to identify which patient groups to target for public health interventions in order to maximise the effectiveness of resources. Other public health uses of the typing data included confirming outbreaks of new strains, including AMR infection,<sup>9,20,23,26,52</sup> or using the data to evaluate existing public health interventions,<sup>9,10,22,30,49</sup> such as confirming the success of partner notification in a population. For example, Monfort *et al.* concluded that the high number of single STs in their sample was probably due to the lack of effective contact tracing. Similarly, many studies (22%; 11/49)<sup>11,12,15,16,19,22,24-26,51,57</sup> suggested that the molecular data could support contact tracing by providing complementary or confirmatory information.

Over one third of studies (39%; 19/49)<sup>9,11,16,17,19,23,26-29,32,34,36,37,42,45-47,55</sup> suggested the information could be used to tailor the clinical management that individual patients receive. For example, if a patient is infected with a strain associated with AMR, an ongoing outbreak or coinfection with another STI, they might be provided with enhanced clinical care, either informing antimicrobial choice or prioritising the patient for test-of-cure and partner notification. Several studies<sup>23,33</sup> highlighted the importance of undertaking localised and prospective molecular epidemiological studies in order to identify changing trends in sequence types that may be indicative of new gonorrhoea outbreaks.

**Table 1: Potential uses of molecular epidemiology to support or evaluate public health interventions suggested by reviewed studies**

| Public health intervention                          | Potential use of molecular epidemiology                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|-----------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Outbreak investigation                              | <ul style="list-style-type: none"> <li>- the similarity between NG DNA has been used to confirm whether cases are related and potentially part of an outbreak<sup>23,26,52</sup></li> <li>- modelling using WGS has been used to estimate the number of undetected cases in an outbreak<sup>25</sup></li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Sexual partner tracing                              | <ul style="list-style-type: none"> <li>- as with outbreak investigation, genetic similarities or differences between NG isolates might be used to inform sexual partner tracing and complement partner notification data<sup>11,12,15,16,19,22,24,25,51,57</sup></li> <li>- identifying likely transmission between isolates using WGS data has been used to determine the density of sexual networks and speed of transmission<sup>24,25</sup></li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                 |
| Antibiotic selection for patient management         | <ul style="list-style-type: none"> <li>- associations found between specific sequence types and antimicrobial resistance may be used to guide patient treatment where phenotypic resistance data are unavailable or delayed<sup>11,16,34,42,45,55</sup></li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Targeted health promotion/behavioural interventions | <ul style="list-style-type: none"> <li>- associations found between specific sequence types and sub-groups of the population might help to identify the sexual network and determine groups for targeting with specific public health interventions<sup>9,11,16,17,19,23,29,32,34,36,37,42,45-47,55</sup></li> <li>- the associations found between specific sequence types and sub-groups of the population might be useful for identifying and possibly quantifying the mixing between different groups, such as MSM and heterosexuals, which may help estimate the effect of tailored public health messages to one group vs the whole population<sup>15,16</sup></li> <li>- STI testing recommendations may be tailored based on whether the NG strain is associated with coinfection with another STI, including HIV<sup>11,17,24,27</sup></li> </ul> |

NG = *Neisseria gonorrhoeae*

MSM = men who have sex with men

## Discussion

This systematic review identified 49 studies between 2004 and 2017 that used molecular typing data linked to patient-level epidemiological data to describe sexual networks of *N. gonorrhoeae* and AMR infection. NG-MAST was the most frequently used typing method, but WGS has become increasingly popular in more recent years. These molecular data have informed our understanding about the mutability of the *N. gonorrhoeae* genome, the relationship between genotype and phenotype, and the pathogen population structure, which provide vital insights into the epidemiology of gonorrhoea and AMR spread. However, linked epidemiological data were often limited; with gender and sexual orientation the most commonly reported variables. Although details of sexual networks were inferred and the public health application of the study findings hypothesised, their value within public health interventions has yet to be systematically evaluated. Consequently, it remains unclear whether molecular epidemiological studies of *N. gonorrhoeae* are superior to traditional epidemiological studies without molecular data, such as routinely collected surveillance data and research studies using patient questionnaires of reported behaviour, when it comes to improving infection control or reducing the spread of *N. gonorrhoeae* and associated AMR.

Our review used systematic methods to identify and assess the relevant literature. The use of precise search terms applied to large medical literature databases reduced the risk of missing studies. Use of two independent reviewers to select studies and extract data minimised the risk of observer bias. A limitation of this review is that only English language studies were included.

The included studies highlight that molecular typing data are useful for understanding sexual networks, but also that there is often more than one explanation for the findings, which might make it difficult to determine appropriate public health actions. For example, a particular strain of *N. gonorrhoeae* identified in both MSM and heterosexuals might be explained by bridging between these sexual networks, or *de novo* independent generation of the sequence type. Similarly, where no sexual networks of sub-populations were identified, such as within the MSM community, this might be interpreted as being driven by disassortative sexual mixing and indicative of high rates of partner change, or it might be that there was insufficient epidemiological information collected to distinguish separate sexual networks.

We found it important to understand the sampling strategy used by molecular epidemiological studies to interpret the results and application within public health interventions. There is a risk of low internal and external validity if authors use restrictive sampling criteria, such as a specific AMR phenotype, or use a convenience sample without consideration of the isolates which are excluded. Inappropriate extrapolation is compounded if studies compare findings or use data from other studies without consideration of the different sampling strategies used. In order to improve the validity and use of molecular epidemiological studies, the sampling strategy should aim to reduce the chance of selection bias, for example by including consecutive isolates across a representative sample of the population. However, we note that the aim and objectives of the studies identified in this review varied and were not always intended to identify sexual networks.

We found conflicting results with respect to associations identified between particular patient variables and sequence types, such as whether different strains of *N. gonorrhoeae* are

circulating amongst people living with HIV compared to those not living with HIV. These differences are likely due to the different populations and sexual behaviours of people in the areas under investigation. Given that populations and their behaviours may differ markedly, care should be taken when extrapolating findings from one area to another, and local analyses may be required. Consequently, the public health action from these findings may also be specific to local areas.

Many of the studies have hypothesised how molecular data might support the development and evaluation of public health interventions. Some studies demonstrated how molecular data were used to confirm clustering of infections in an outbreak. However, many of the other suggested public health applications, such as tailoring and targeting health promotion messages to specific groups or by changing clinical management protocols (for example by recalling patients back for test-of-cure depending on the infecting strain of the molecular data) are yet to be applied or properly evaluated.

In conclusion, linking DNA-based sequence typing data to patient-level epidemiological data can improve scientific knowledge on the structure of sexual networks and dissemination of *N. gonorrhoeae* and associated AMR. Future molecular epidemiological studies should, however, place greater emphasis on determining specific applications of molecular typing within public health interventions for controlling gonorrhoea. Greater collaboration between those involved in molecular epidemiological research and public health practitioners might facilitate evaluation of the public health potential of these increasingly prevalent technologies.

#### **Author contribution**

KT, HB, MC, SC, NF and GH developed the protocol for the study. KT, HB and SC conducted the literature review and data extraction. All authors contributed to the writing of the manuscript.

#### **Acknowledgements**

**NIHR HPRU in Blood Borne and Sexually Transmitted Infections Steering Committee:** Caroline Sabin (Director), Anthony Nardone (PHE Lead), Catherine Mercer, Gwenda Hughes, Jackie Cassell, Greta Rait, Samreen Ijaz, Tim Rhodes, Kholoud Porter, Sema Mandal and William Rosenberg.

#### **Declaration of interests**

Dr. Cole reports grants from The European Centre for Disease Prevention and Control which included funding for molecular typing, outside the submitted work.

Dr. Harris reports grants from Wellcome Trust, during the conduct of the study.

#### **Role of the funding source**

The research was funded by the National Institute for Health Research Health Protection Research Unit in Blood Borne and Sexually Transmitted Infections at University College London in partnership with Public Health England and in collaboration with the London School of Hygiene and Tropical Medicine. The funder of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report or decision to

submit the manuscript for publication. The corresponding author (KT) had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## References

1. Newman L, Rowley J, Vander Hoorn S, et al. Global Estimates of the Prevalence and Incidence of Four Curable Sexually Transmitted Infections in 2012 Based on Systematic Review and Global Reporting. *PLoS One* 2015; **10**(12): e0143304.
2. Wi T, Lahra MM, Ndowa F, et al. Antimicrobial resistance in *Neisseria gonorrhoeae*: Global surveillance and a call for international collaborative action. *PLoS Med* 2017; **14**(7): e1002344.
3. Savage EJ, Marsh K, Duffell S, Ison CA, Zaman A, Hughes G. Rapid increase in gonorrhoea and syphilis diagnoses in England in 2011. *Euro Surveill* 2012; **17**(29).
4. Sonnenberg P, Clifton S, Beddows S, et al. Prevalence, risk factors, and uptake of interventions for sexually transmitted infections in Britain: findings from the National Surveys of Sexual Attitudes and Lifestyles (Natsal). *Lancet* 2013; **382**(9907): 1795-806.
5. Field N, Cohen T, Struelens MJ, et al. Strengthening the Reporting of Molecular Epidemiology for Infectious Diseases (STROME-ID): an extension of the STROBE statement. *Lancet Infect Dis* 2014; **14**(4): 341-52.
6. Town K, Bolt H, Cole M, Harris S, Field N, Hughes G. A systematic review of *Neisseria gonorrhoeae* (NG) DNA typing techniques to inform public health understanding of NG transmission in sexual networks and the risk factors for NG acquisition. 03/04/2016 ed: PROSPERO; 2016.
7. Unemo M, Dillon JA. Review and international recommendation of methods for typing *neisseria gonorrhoeae* isolates and their implications for improved knowledge of gonococcal epidemiology, treatment, and biology. *Clin Microbiol Rev* 2011; **24**(3): 447-58.
8. Todd K, Durrheim D, Pickles R, et al. Using epidemiological and molecular methods to investigate an outbreak of gonorrhoea associated with heterosexual contact in Newcastle, NSW, Australia. *Sex Health* 2007; **4**(4): 233-6.
9. Martin IMC, Ison CA, Aanensen DM, Fenton KA, Spratt BG. Rapid sequence-based identification of gonococcal transmission clusters in a large metropolitan area. *Journal of Infectious Diseases* 2004; **189**(8): 1497-505.
10. Martin IMC, Ison CA, Aanensen DM, Fenton KA, Spratt BG. Changing epidemiologic profile of quinolone-resistant *Neisseria gonorrhoeae* in London. *Journal of Infectious Diseases* 2005; **192**(7): 1191-5.
11. Abu-Rajab K, Palmer HM, Scoular A, et al. To what extent does *Neisseria gonorrhoeae* multiantigen sequence typing of gonococcal isolates support information derived from patient interviews? *International Journal of STD and AIDS* 2009; **20**(6): 414-7.
12. Bilek N, Martin IM, Bell G, Kinghorn GR, Ison CA, Spratt BG. Concordance between *Neisseria gonorrhoeae* genotypes recovered from known sexual contacts. *Journal of Clinical Microbiology* 2007; **45**(11): 3564-7.
13. Chisholm SA, Alexander S, Desouza-Thomas L, et al. Emergence of a *Neisseria gonorrhoeae* clone showing decreased susceptibility to cefixime in England and Wales. *Journal of Antimicrobial Chemotherapy* 2011; **66**(11): 2509-12.
14. Chisholm SA, Neal TJ, Alawattagama AB, Birley HD, Howe RA, Ison CA. Emergence of high-level azithromycin resistance in *Neisseria gonorrhoeae* in England and Wales. *Journal of Antimicrobial Chemotherapy* 2009; **64**(2): 353-8.
15. Choudhury B, Risley CL, Ghani AC, et al. Identification of individuals with gonorrhoea within sexual networks: a population-based study. *Lancet* 2006; **368**(9530): 139-46.
16. Cole MJ, Thomas DR, Chisholm SA, et al. Molecular epidemiology of gonorrhoea in Wales (UK). *Sexually Transmitted Infections* 2013; **89**(3): 267-72.
17. Fernando I, Palmer HM, Young H. Characteristics of patients infected with common *Neisseria gonorrhoeae* NG-MAST sequence type strains presenting at the Edinburgh genitourinary medicine clinic. *Sexually Transmitted Infections* 2009; **85**(6): 443-6.

18. Ison CA, Town K, Obi C, et al. Decreased susceptibility to cephalosporins among gonococci: Data from the Gonococcal Resistance to Antimicrobials Surveillance Programme (GRASP) in England and Wales, 2007-2011. *The Lancet Infectious Diseases* 2013; **13**(9): 762-8.
19. Palmer HM, Young H. Dramatic increase in a single genotype of TRNG ciprofloxacin-resistant *Neisseria gonorrhoeae* isolates in men who have sex with men. *International Journal of STD and AIDS* 2006; **17**(4): 254-6.
20. Palmer HM, Young H, Martin IMC, Ison CA, Spratt BG. The epidemiology of ciprofloxacin resistant isolates of *Neisseria gonorrhoeae* in Scotland 2002: A comparison of phenotypic and genotypic analysis. *Sexually Transmitted Infections* 2005; **81**(5): 403-7.
21. Palmer HM, Young H, Winter A, Dave J. Emergence and spread of azithromycin-resistant *Neisseria gonorrhoeae* in Scotland. *Journal of Antimicrobial Chemotherapy* 2008; **62**(3): 490-4.
22. Risley CL, Ward H, Choudhury B, et al. Geographical and demographic clustering of gonorrhoea in London. *Sexually Transmitted Infections* 2007; **83**(6): 481-7.
23. Chisholm SA, Wilson J, Alexander S, et al. An outbreak of high-level azithromycin resistant *Neisseria gonorrhoeae* in England. *Sexually Transmitted Infections* 2016; **92**(5): 365-7.
24. 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**(11): 1295-303.
25. Didelot X, Dordel J, Whittles LK, et al. Genomic analysis and comparison of two gonorrhoea outbreaks. *mBio* 2016; **7**(3).
26. Foster K, Cole M, Hotonu O, et al. How to do it: lessons identified from investigating and trying to control an outbreak of gonorrhoea in young heterosexual adults. *Sex Transm Infect* 2016; **92**(5): 396-401.
27. Bernstein KT, Marcus JL, Barry PM, et al. Characteristics of males infected with common *neisseria gonorrhoeae* sequence types in the gonococcal isolate surveillance project, san francisco, California, 2009. *American Journal of Epidemiology* 2013; **178**(8): 1289-95.
28. Chen SC, Yin YP, Dai XQ, Unemo M, Chen XS. First nationwide study regarding ceftriaxone resistance and molecular epidemiology of *Neisseria gonorrhoeae* in China. *Journal of Antimicrobial Chemotherapy* 2016; **71**(1): 92-9.
29. Chisholm SA, Unemo M, Quaye N, et al. Molecular epidemiological typing within the European Gonococcal Antimicrobial Resistance Surveillance Programme reveals predominance of a multidrugresistant clone. *Eurosurveillance* 2013; **18**(3).
30. Hjelmevoll SO, Golparian D, Dedi L, et al. Phenotypic and genotypic properties of *Neisseria gonorrhoeae* isolates in Norway in 2009: Antimicrobial resistance warrants an immediate change in national management guidelines. *European Journal of Clinical Microbiology and Infectious Diseases* 2012; **31**(6): 1181-6.
31. Horn NN, Kresken M, Korber-Irrgang B, et al. Antimicrobial susceptibility and molecular epidemiology of *Neisseria gonorrhoeae* in Germany. *International Journal of Medical Microbiology* 2014; **304**(5-6): 586-91.
32. Wong WW, Huang CT, Li LH, Chiang CC, Chen BD, Li SY. Molecular epidemiological identification of *Neisseria gonorrhoeae* clonal clusters with distinct susceptibility profiles associated with specific groups at high risk of contracting human immunodeficiency virus and syphilis. *Journal of Clinical Microbiology* 2008; **46**(12): 3931-4.
33. Serra-Pladevall J, Barberá MJ, Callarisa AE, Bartolomé-Comas R, Andreu A. Differences in *Neisseria gonorrhoeae* population structure and antimicrobial resistance pattern between men who have sex with men and heterosexuals. *Epidemiology and Infection* 2017; **145**(2): 379-85.
34. Trembizki E, Wand H, Donovan B, et al. The Molecular Epidemiology and Antimicrobial Resistance of *Neisseria gonorrhoeae* in Australia: A Nationwide Cross-Sectional Study, 2012. *Clinical Infectious Diseases* 2016; **63**(12): 1591-8.

35. Ni C, Xue J, Zhang C, Zhou H, van der Veen S. High prevalence of *Neisseria gonorrhoeae* with high-level resistance to azithromycin in Hangzhou, China. *Journal of Antimicrobial Chemotherapy* 2016; **71**(8): 2355-7.
36. Carannante A, Latini A, Cusini M, et al. Update on antimicrobial susceptibility and genotype of *Neisseria gonorrhoeae* isolated in Italy. *Diagnostic Microbiology and Infectious Disease* 2012; **72**(3): 288-90.
37. Carannante A, Renna G, Conte ID, et al. Changing antimicrobial resistance profiles among *Neisseria gonorrhoeae* isolates in Italy, 2003 to 2012. *Antimicrobial Agents and Chemotherapy* 2014; **58**(10): 5871-6.
38. Lundback D, Fredlund H, Berglund T, Wretling B, Unemo M. Molecular epidemiology of *Neisseria gonorrhoeae*- identification of the first presumed Swedish transmission chain of an azithromycin-resistant strain. *Apmis* 2006; **114**(1): 67-71.
39. Ota KV, Ng LK, Melano RG, et al. Identification of sexual networks through molecular typing of quinolone-resistant *Neisseria gonorrhoeae* in Ontario, Canada. *Sexually Transmitted Diseases* 2011; **38**(9): 811-4.
40. Singh AE, Gratrix J, Read R, et al. *Neisseria gonorrhoeae* multiantigen sequence typing is beneficial in further characterizing gonococcal populations in Alberta, Canada. *Sexually Transmitted Diseases* 2013; **40**(9): 744-50.
41. Starnino S, Dal Conte I, Matteelli A, et al. Trend of ciprofloxacin resistance in *Neisseria gonorrhoeae* strains isolated in Italy and analysis of the molecular determinants. *Diagnostic microbiology and infectious disease* 2010; **67**(4): 350-4.
42. Starnino S, Stefanelli P. Azithromycin-resistant *Neisseria gonorrhoeae* strains recently isolated in Italy. *Journal of Antimicrobial Chemotherapy* 2009; **63**(6): 1200-4.
43. Stevens K, Zaia A, Tawil S, et al. *Neisseria gonorrhoeae* isolates with high-level resistance to azithromycin in Australia. *Journal of Antimicrobial Chemotherapy* 2015; **70**(4): 1267-8.
44. Yuan LF, Yin YP, Dai XQ, et al. Resistance to azithromycin of *Neisseria gonorrhoeae* isolates from 2 cities in China. *Sexually Transmitted Diseases* 2011; **38**(8): 764-8.
45. Jacobsson S, Golparian D, Cole M, et al. WGS analysis and molecular resistance mechanisms of azithromycin-resistant (MIC  $\geq$  2 mg/L) *Neisseria gonorrhoeae* isolates in Europe from 2009 to 2014. *Journal of Antimicrobial Chemotherapy* 2016; **71**(11): 3109-16.
46. Chen CC, Yen MY, Wong WW, et al. Tracing subsequent dissemination of a cluster of gonococcal infections caused by an ST1407-related clone harbouring mosaic penA alleles in Taiwan. *Journal of Antimicrobial Chemotherapy* 2013; **68**(7): 1567-71.
47. Cheng CW, Li LH, Su CY, Li SY, Yen MY. Changes in the six most common sequence types of *Neisseria gonorrhoeae*, including ST4378, identified by surveillance of antimicrobial resistance in northern Taiwan from 2006 to 2013. *Journal of Microbiology, Immunology and Infection* 2015.
48. Jeverica S, Golparian D, Maticic M, Potocnik M, Mlakar B, Unemo M. Phenotypic and molecular characterization of *Neisseria gonorrhoeae* isolates from Slovenia, 2006-12: Rise and fall of the multidrug-resistant NG-MAST genogroup 1407 clone? *Journal of Antimicrobial Chemotherapy* 2014; **69**(6): 1517-25.
49. Monfort L, Caro V, Devaux Z, Delannoy AS, Brisse S, Sednaoui P. First *Neisseria gonorrhoeae* genotyping analysis in France: Identification of a strain cluster with reduced susceptibility to ceftriaxone. *Journal of Clinical Microbiology* 2009; **47**(11): 3540-5.
50. Starnino S, Suligoj B, Regine V, et al. Phenotypic and genotypic characterization of *Neisseria gonorrhoeae* in parts of Italy: detection of a multiresistant cluster circulating in a heterosexual network. *Clinical Microbiology & Infection* 2008; **14**(10): 949-54.
51. Unemo M, Sjostrand A, Akhras M, et al. Molecular characterization of *Neisseria gonorrhoeae* identifies transmission and resistance of one ciprofloxacin-resistant strain. *Apmis* 2007; **115**(3): 231-40.
52. Lahra MM, Ward A, Trembizki E, et al. Treatment guidelines after an outbreak of azithromycin-resistant *Neisseria gonorrhoeae* in South Australia. *The Lancet Infectious Diseases* 2017; **17**(2): 133-4.

53. Florindo C, Pereira R, Boura M, et al. Genotypes and antimicrobial-resistant phenotypes of *Neisseria gonorrhoeae* in Portugal (2004-2009). *Sexually Transmitted Infections* 2010; **86**(6): 449-53.
54. Perez-Losada M, Crandall KA, Bash MC, Dan M, Zenilman J, Viscidi RP. Distinguishing importation from diversification of quinolone-resistant *Neisseria gonorrhoeae* by molecular evolutionary analysis. *BMC Evolutionary Biology* 2007; **7**(84).
55. 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. *Journal of Infectious Diseases* 2016; **214**(10): 1579-87.
56. 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. *Journal of Clinical Microbiology* 2015; **53**(1): 191-200.
57. 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**(3): 220-6.
58. Kennedy CE, Bernard LJ, Muessig KE, et al. Serosorting and HIV/STI Infection among HIV-Negative MSM and Transgender People: A Systematic Review and Meta-Analysis to Inform WHO Guidelines. *J Sex Transm Dis* 2013; **2013**: 583627.

ACCEPTED MANUSCRIPT